CEOS
Committee on Earth Observation Sa
Working Group on Information Systems and S

Interoperable Catalogue System (ICS) System Design Document (SDD)

CEOS
Working Group on Information Systems and Services
Protocol Task Team

Doc. Ref.: CEOS/WGISS/PTT/SDD
Date: June 1998
Issue: Version 1.4

AUTHORITY

Issue:	Version 1.4
Date:	June 1998

This document has been approved for publication by the Protocol Task Team of the Committee on Earth Observation Satellites (CEOS) and reflects the consensus of the Protocol Task Team technical panel experts from the CEOS member agencies.

This document is published and maintained by:
Goddard Space Flight Center
National Aeronautics and Space Administration
http://edhs1.gsfc.nasa.gov

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DOCUMENT STATUS SHEET

Version	Date	Comments
0.1	May 1996	Interoperable Catalogue Systems and Services System Architecture Document (ECS White Paper 170-WP-007-001). Annotated Outline of SAD prepared for PTT Meeting, May 1996, Tokyo.
1.0	September 1996	Document renamed ICS SDD. RIDs from Tokyo PTT meeting incorporated. Document prepared for PTT Meeting, September 1996, Annapolis. Final scope and purpose. Initial technical completeness review.
1.1	November 1996	RIDs from Annapolis PTT meeting incorporated. Document prepared for PTT Meeting, December 1996, Ispra. Document to be reviewed by PTT using formal RID process.
1.2	March 1997	RIDs from Ispra PTT-4 meeting incorporated. Document prepared for ratification by PTT.
1.3	February 1998	Document prepared for RID review during PTT.
1.4	June 1998	RIDs from Salzburg PTT-8 meeting incorporated. Document prepared for ratification by PTT.

DOCUMENT CHANGE RECORD

Change No.	Date	Changes in Version 1.4 from Version 1.3
1	March 1998	Incorporated RIDs from PTT-8 meeting in Salzburg.

1. INTRODUCTION

1.1 SDD Purpose and Scope

ICS Compatibility: Explanatory

Major Space Agencies through the Committee for Earth Observation Satellites (CEOS) have recognized the importance of providing a seamless access to their catalogues via a new Interoperable protocol. This protocol is the Catalogue Interoperability Protocol (CIP) and is defined in a sister document: the CIP Specification - Release B. The CIP Specification defines how compliant catalogues need to interface through the low level Z39.50 protocol. This System Design Document (SDD) focuses on the design of a global system and the infrastructure that will link together agencies using the CIP protocol. This macro system is called the Interoperable Catalogue System or ICS.

The CEOS Protocol Task Team (PTT) is developing a suite of documents as defined in the PTT Development Plan [R1]*. The ICS User Requirements Document (URD) [R2] is where the CEOS agencies bring in their needs for functionality, constraints, etc. The PTT accepts input as requirements added to the URD. It is the PTT activity to interpret those requirements into the CIP Specification [R3], the ICS Guide Design and Protocol Specification [24], and the ICS SDD. As a result, if groups comply with the CIP Specification, the ICS Guide Design and Protocol Specification, and the SDD, then they will be compliant with the URD. This section defines the purpose of the SDD. The purpose of ICS is provided in Section 2.

The ICS SDD defines the elements and interfaces which comprise the CEOS ICS. Implmenentors of ICS elements should derive element requirements from the mandatory parts of the ICS System Design. This document defines relationships to the CEOS Network which is under development by the CEOS Network Subgroup. Assumptions are made in this document about existing systems which interface to ICS, e.g., existing agency systems will persistently store user orders. These assumptions cannot be required of the agency systems but represent the SDD's approach to the interface with ICS. The last section of the SDD defines a mimimum site configuration relative to the SDD content in order to meet CEOS policy. The SDD is written for a federation of CEOS agency systems, however other data provider federations could use the SDD as a template design for their system. The ultimate purpose of the SDD is to define a system which can be implemented and operated by the CEOS federation to provide data and services to users.

* Citations for references are provided in Section 1.5.

Each of the lowest level sections in the SDD are labeled for ICS compatibility. The applicability labels tell the reader if an ICS site must implement the content of the section in order to be compatible with other ICS sites. The determination of applicability of a section is based on adherence to CEOS policy, effective ICS operation, and assurance of interoperability. The following terms are used in the applicability labels.

- Mandatory - A site must comply with contents of the section to be ICS compatible.
- Mandatory As Applicable (MAA) - If a site is implementing the functionality in the section, then the site must comply with the section to be ICS compatible. A minimum site need not comply with any MAA section and could still be ICS compatible.
- Explanatory - The contents of the section describe a concept or background topic which may be useful to the reader to understand ICS. ICS compatibility is not determined by the contents of an explanatory section.

1.2 Organization of the SDD

ICS Compatibility: Explanatory

The SDD is organized into various views of the ICS. This approach is based on guidelines for SDDs, e.g., [R6] and existing SDDs, e.g., [R7]. After a discussion of the PTT approach to catalogue interoperability (Section 2), the SDD provides the following views of the ICS:

- Functional Framework (Section 3) - overall application view of the computing elements which comprise the functional architecture of ICS. The ICS Framework provides several configurations for ICS implementation at sites.
- Data View (Section 4) - provides an overall data framework indicating what data is held by the various ICS elements. Several data components are defined in the Retrieval Manager in order that ICS operations are performed correctly across ICS sites.
- Communications View (Section 5) - describes two models of the ICS design related to communications and networks. First a communication protocol stack is defined placing CIP on a TCP/IP stack. Next, ICS reliance on CEOSnet is discussed including identifying sites at which ICS Retrieval Managers are initially intended to be located and the traffic between the sites.
- Security View (Section 6) - This section provides a discussion of the security design for ICS. First, a security assessment of ICS is provided in terms of security assets, vulnerabilities and threats. Based on the assessment, the ICS security control methods are described. The methods are grouped into three categories: administration, physical and computing controls.
- Systems Management View (Section 7) - This section provides a list of topics which can serve as a basis for beginning the System Management design in the next ICS release.
- Architecture Verification (Section 8) - This section provides several methods for evaluating the system design. Through a series of scenarios, a definitive identification of inter-element interfaces, and distributed query performance estimates, this section demonstrates the completeness of the system design presented in earlier sections.
- ICS Minimum Site Configuration (Section 9) - This section defines the required minimum configuration for a site which wishes to be considered ICS compatible. The minimum site is based on the CEOS policy.

1.3 ICS/PTT Development Process

ICS Compatibility: Explanatory

This document was developed by the CEOS PTT. The PTT is part of the WGISS Access Subgroup within CEOS. The lead agency for compilation of PTT inputs and for final preparation of this version of the document was the National Aeronautics and Space Administration (NASA). A complete list of organizations participating in the PTT is provided in the PTT Terms of Reference (http://ceos.ccrs.nrcan.gc.ca/taskteam/cip.html)

1.4 A Guide To PTT Documents

ICS Compatibility: Explanatory

The CEOS Protocol Task Team (PTT) collects requirements for an interoperable infrastructure connecting Earth-observation data catalogues and access systems. It establishes agreement on these requirements and produces detailed specifications on which system implementations can be based. The PTT has named this interoperable infrastructure the CEOS Interoperable Catalogue System (ICS). The PTT output is captured and maintained in the documents listed below. A discussion of which document a user might read first is provided after the list of PTT documents.

- **ICS User Requirements Document (URD)**

  The ICS URD [R2] specifies user requirements for an interoperable infrastructure linking catalogue systems of different agencies. For this purpose it defines requirements for all interoperable components and the protocols needed for exchanging messages between them. (Currently, the URD does not reflect the http approach to guide and will be updated.)

- **ICS System Design Document (SDD)**

  The ICS SDD defines the elements and interfaces which comprise the CEOS Interoperable Catalogue System (ICS). The SDD provides diagrams showing the interrelations between ICS elements, scenarios to explain the dynamic interaction, a data model showing the data relations, the communications services utilized in ICS, and the system management approach for ICS. The SDD provides both design and tutorial information.

- **Catalogue Interoperability Protocol (CIP) Specification - Release B**

  The CIP Specification [R3] defines the interoperable protocol for exchanging messages related to data search and data ordering. CIP is defined as a profile of the ISO standard Z39.50 with extensions for distributed searching using the collections model. The specification defines all CIP messages, as well as the attributes used for searching and the elements needed for retrieval. CIP may be used outside of the CEOS ICS. The CIP Specification is the definitive source for determining CIP compliance. The CIP Specification provides the framework for exchanging data orders. Details on how to specify options on orders are defined in the Order Options Amendment to the CIP Specification [R28].

- **ICS Guide Design and Protocol Specification**

  The ICS Guide Design and Protocol Specification [R24] describes the ICS elements that support guide and the protocol of messages used for guide. The approach is based on the http protocol using virtual documents. The interaction of the Guide elements with CIP elements is discussed. Example scenarios describe the dynamic behavior. The ICS Guide Design and Protocol Specification is the definitive source for determining ICS Guide Protocol (IGP) compliance.

- **ICS Collection Manual**

  The ICS Collection Manual [R5] provides procedures and guidelines for the creation and maintenance of Collection Information contained in an ICS Retrieval Manager. The document provides sufficient detail to allow the ICS Site Administrators to manage the ICS Collection repositories according to the rules specified in the Collection Manual. The manual further provides guidelines for developing the ICS Collection Structure. Collectively, the procedures and guidelines; which can be applied to any/all implementation strategies, if followed, will ensure data interoperability.

- **ICS Valids Document**

  The ICS Valids Document [R4] defines the list of valid keywords for the enumerated search attributes used by CIP. The valids document provides the procedures for controlling the list of valids either based on coordination with other standardization groups or through rules and procedures for ICS only valids

A starting point for most users of PTT documents will be the SDD which provides tutorial information about how CIP, IGP and Collections are used in ICS. An implementer who wants the details of CIP messages may want to go directly to the CIP Specification. This is also true for an IGP implementer who may want to go directly to the Guide Design and Protocol Specification. Someone who is responsible for organizing the data for an agency may want to browse the SDD to understand the ICS data model and then proceed to the details in the Collections Manual and the Valids Document. If technical input to the PTT direction is desired, reviewing the URD and proposing new User Requirements is the right approach.

To avoid conflicting information, the overlap of material between PTT documents has been minimized. Specific guidelines for splitting information between the SDD and the CIP Specification were used and are presented in Table 1-1.

Table 1-1. Split of Information between CIP Specification and ICS SDD

Dimension	CIP Specification	ICS SDD
Scope	Application Protocol	Full Protocol Stack System Management
Data	Internal content of CIP messages	CIP Messages by Name only
Element Interactions	Session between a Target and Origin	Use of multiple CIP sessions across multiple ICS Elements

Additional information about PTT activities and documents can be found at:
http://ceos.ccrs.nrcan.gc.ca/taskteam/cip.html

1.5 Glossary

1.5.1 Acronyms

ICS Compatibility: Explanatory

The following acronyms are used in this document:

ADD	Architecture Design Document
ANSI	American National Standards Institute
APDU	Application Protocol Data Unit
AS	Access Subgroup (Part of CEOS-WGISS)
ASN.1	Abstract Syntax Notation.1
BER	Basic Encoding Rules
BNSC	British National Space Centre
CA	Certification Authority
CCRS	Canada Centre for Remote Sensing
CCSDS	Consultative Committee for Space Data Systems
CDB	Collection Data Base
CEO	Centre for Earth Observation (European Commission)
CEO-ES	CEO- Enabling Services
CEOS	Committee on Earth Observation Satellites
CINTEX	Catalogue INTeroperability Experiment
CINTOPS	Catalogue INTeroperability OPerationS
CIP	Catalogue Interoperability Protocol
CM	Configuration Management
CMT	Collection Management Tool
CNES	Centre National d'Etudes Spatiales (France)

CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
DB	Data Base
DBMS	Data Base Management System
DEM	Digital Elevation Map
DIF	Directory Interchange Format
DLR	Deutsche Forschungsanstalt fur Luft-und Raumfahrt
DNS	Domain Name Service
DOGS	Directory of GEO Servers
ECS	EOSDIS Core System
EO	Earth Observation
EOC	Earth Observation Center (NASDA)
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System (NASA)
ESA	European Space Agency
ESRIN	European Space Research Institute (ESA)
FGDC	Federal Geographic Data Committee (USA)
GCMD	Global Change Master Directory
GI	Guide Indexer
GS	Guide Serv er
GSFC	Goddard Space Flight Center (NASA)
GT	Guide Translator
HITS	Hughes Information Technology Systems, Inc.
HTML	Hyper Text Mark-up Language
HTTP	HyperText Transfer Protocol
ICS	Interoperable Catalogue System
IGP	ICS Guide Protocol
ID	Identifier
IP	Internet Protocol
IRE-RAS	Institute of Radio Engineering. - Russian Academy of Science
ISA	ICS Site Administrator
ISO	International Standards Organization
LaRC	Langley Research Center (NASA)
LDAP	Lightweight Directory Access Protocol
MAC	Message Authentication Code
MACAO	Member Agency Control Authority Office
MCT	Monitoring and Control Tool
NASA	National Aeronautics and Space Administration (US)
NASDA	National Space Development Agency (Japan)
NOAA	National Oceanic and Atmospheric Administration (US)
NRSC	National Remote Sensing Center
NSRS	Natural Environment Research Council (BNSC)
OHS	Order Handling System
OSI	Open Systems Interconnection

PTT	Protocol Task Team (Part of CEOS- WGISS-AS)
PKI	Public Key Infrastructure
QA	Quality Assurance
RFC	Request For Comment
RM	Retrieval Manager
RPN	Reverse Polish Notation
SATAN	Security Administrator Tool for Analyzing Networks
SDD	System Design Document
SFDU	Standard Formatted Data Unit
SSM	Site System Management
SNMP	Simple Network Management Protocol
SST	Sea Surface Temperature
TBD	To Be Determined
TBR	To Be Resolved
TBS	To Be Supplied
TCP	Transmission Control Protocol
TN	Technical Note
UDB	User Data Base
UPS	User Profile System
UR	User Requirement
URD	User Requirements Document
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
US	United States
USGS	U.S. Geological Survey
WGISS	Working Group on Information Systems and Services (Part of CEOS)
WWW	World Wide Web

1.5.2 Definitions

ICS Compatability: Explanatory

This section provides definitions of the terms related to ICS:

Archive	An archive of EO data can hold various types of data ranging from satellite images and climatological products processed from the images, to observation data and climatological statistics. An archive may also contain information describing the EO data and also supplementary data such as design documentation, algorithm object and source code, technical reports, user manuals, etc. There is likely to be a database management system for maintenance and low level access to the data. The archive will, in general, be accessed by a front end archive server that then presents the data as requested by the Retrieval Manager.
Catalogue Interoperability	The ability to provide a Data User with the appearance of a single, unified catalogue for all participating data providers. In order to provide catalogue interoperability all paricipating data providers must support at least one common method (i.e., API) for accessing functions such as authentication, directory, inventory, guide and order. Each supplier may support additional consumer functional interfaces to support their private data users
Catalogue System	A catalogue system provides services such as inventory, browse, directory, order and guide, which may be supplemented by further services, but should contain at a minimum, inventory. The CIP is the protocol that shall enable the many services of many catalogue systems to inter-operate. Usually a catalogue system resides at a particular agency or data provider facility but may be distributed across catalogue sites.
Catalogue Translator	One of three types of ICS *Translators*. *Catalogue Translator* converts CIP messages into a data providers protocol for the services of Inventory, Directory, and Browse.
CEOS Agency Systems	The data provider systems of CEOS member agencies and their affiliates.
CIP Client	A software element composed of a *Presentation Layer*, a *Local User Management Layer*, and an *Application Layer*. Only the *CIP Client Application Layer* is part of ICS.
CIP Client Application Layer	Part of the *CIP Client* which deals directly with CIP including creating CIP messages and includes off-the-shelf Z39.50 communication software.
CIP Client Local User Management Layer	Part of the *CIP Client* which provides functionally for local data management, e.g. saving a result set, or converting result sets into orders.
CIP Client Presentation Layer	Part of the *CIP Client* dealing with how information is presented to the human user, including all issues related to HMI, as well as dealing with certain format specific issues, e.g., displaying browse imagery.
CIP Operation	Based on Z39.50 definition, an initiating CIP request message and the corresponding terminating response, along with intervening related messages. Multiple operations may occur within a *CIP Session*.
CIP Session	A set of CIP messages exchanged between an *Origin* and *Target* beginning with an initialization message and ending with a close message between which the *Origin* and *Target* maintain state information concerning the interaction. (A *CIP Session* is a Z39.50 Z-Association.)
CIP Message	A unit of information transferred between an origin and a target whose format is specified as a Z39.50 Application Protocol Data Unit (APDU) possibly containing CIP specific APDUs in the external portion of a Z39.50 APDU.

Collection	A grouping of item descriptors that have commonality. A collection consists of a number of attributes that describe the collective contents of the collection, the values of these attributes can then be searched on to select items of interest to the user. Collections also have members; these are the unique identifiers of the items that are grouped by the collection rather than their collective descriptions. As collection members can be identifiers of other collections, a hierarchy of collections and product/guide members can be established, therefore permitting a flexible and powerful organization of data.
Collection Management Tool	Used by the RMA for tasks involved with populating and maintaining the data in the Retrieval Manager. These tasks involve translating collection or directory information into CIP collection format and checking for valid entries.
Existing Agency Client	Software elements which interact with a data providers *Catalogue* using the data providers protocol.
Guide data	Data that is available to the user to enhance understanding of the EO data, spacecraft, instrument, etc., and hence make a detailed analysis of whether the product data will be of value for a particular application. Guide data may also contain information necessary for processing the product data further.
ICS Client	Component of ICS that provides user access to CIP services via a CIP Client, and access to ICS Guide System services via an HTTP Client.
ICS Gateway	A software element which provides non-ICS clients access to ICS catalogues, e.g., searches by *Existing Agency Clients* to a *Catalogue* are sent to a *Retrieval Manager* using an *ICS Gateway*.
Item Descriptor	Used by CIP to represent items. The descriptor is a set of attributes. Item descriptors are of the following types: product, collection, guide, user.
Monitoring and Control Tools	Provides the machine-to-machine interface for integrating the operations of the *Retrieval Manager* with the operations of SSM.
OHS Translator	One of three types of ICS *Translators*. *OHS Translator* converts CIP messages into a data providers protocol for the purposes of order specification, quotation, and request.
Order Handling System	That part of a data provider which provides services associated with ordering products and guide.
Origin	Based on the Z39.50 definition for origin, where an *Origin* is that portion of a client or middleware which may initiate a *CIP session* with a *Target*. ICS elements capable of acting as *Origins* are *CIP Clients, Retrieval Managers, ICS Gateway* and *HTTP/CIP Gateways*.
Other CIP Based Federations	Groups of organizations other than CEOS using CIP to provide catalogue interoperability to data providers. These providers are interoperable with CEOS using CIP. These other federations may or may not use the ICS system design as a basis of the federation design.
Other Z39.50 Based Services.	Catalogue interoperability services provided using Z39.50 Version 3 but not necessarily CIP.
Primary Order	A CIP order between a *CIP Client* and a *Retrieval Manager*. (See *Secondary Order*)
Product data	A unique aggregation of data generated from information held in, or to be held in an archive (for predicted products). It can be located and retrieved by a user via CIP, possibly following further processing, such as map projection, sub-setting, band selection, etc., after or during extraction of the raw data as stored in the archive.

Registered Collections	Are owned by the EO data provider and described by a collection descriptor. These collections support the full range of CIP access services including discovery, navigation, location and searching. There are two types of registered collections which are distinguished by the purpose for which they were created, Archive Collections and Theme Collections.
Retrieval Manager	A *Retrieval Manager* services (and may be installed at) each catalogue site, it is used to integrate together the local catalogue systems and provide communication between users and other catalogue site *Retrieval Managers*. It is anticipated that each catalogue site is at least connected to one *Retrieval Manager* and that *Retrieval Manager* would 'know about' or 'own' a number of collections. The data within these collections would be the responsibility of that *Retrieval Manager*, with external collections referenced only and managed by their respective *Retrieval Managers*. The *Retrieval Managers* at each catalogue site would also communicate with each other using the CIP. The *Retrieval Manager* would then also communicate with local catalogue servers, such as archives and inventories, within its own site to services requests received from users. Another key function of the *Retrieval Manager* is to route search queries to other relevant *Retrieval Managers* and consolidate the search results before returning them to the user.
Retrieval Manager Administrator	The human operator that performs all tasks needed to establish and maintain a *Retrieval Manager*. In practice this is more than one person as the tasks are various types: scientist for collection definition, data base expert for maintaining CDB, system operator for diagnosing and correcting operational activities, etc. For convenience purposes all of these tasks are performed by the *RMA*.
Secondary Order	A CIP order is created by a *Retrieval Manager* in response to a *primary order* and may be either to another *Retrieval Manager* or to an *OHS Translator*
Site System Management	That part of a data provider which provides coordinated, on-line management of the distributed processing environment.
Target	Based on the Z39.50 definition for target, where a *Target* is that portion of middleware and servers which accept requests for *CIP sessions* from an *Origin*. ICS elements capable of acting as *Targets* are *Retrieval Managers* and *Translators*.
Task Package	The set of attributes that describe an activity which is started by an Extended Services Request. Based on Z39.50 definition for a Task Package.
Theme Collections	This type of collection may be set up by data providers or users who want to organize some of their data into groupings which differ from their provider archive collections (i.e. from the baseline inventory), for the convenience of their users, for example, based on the geographical area covered, the scientific discipline supported by the data, the instrument type. Etc.
Translators	Software element which converts CIP into the protocols used by a data provider. Three *Translators* are identified in ICS: *Catalogue Translator, OHS Translator, UPS Translator*.
UPS Translator	One of three types of ICS *Translators*. *UPS Translator* converts CIP messages into a data providers protocol for the user information, e.g., address, and for authentication.
Unregistered Collections	Are likely to be created and owned by an end user of EO products that has created and populated the collection to obtain a single source of thematic information. This will then enable further analysis or easy access by themselves or other users. These are collections of potentially quite disparate item descriptors of interest to a relatively small user community researching a particular theme, i.e. in the example, the mid-west flood of 1993. Unregistered collections need not be described by a collection descriptor and cannot be discovered in a search of ICS collections.

User Profile System	That part of a data provider which provides services associated with user information and authentication.
User Session	A user session represents the interaction between a human user and the CIP (i.e. a Retrieval Manager), which has in general been established by the user via a terminal which may be running a WWW client or a specific CIP client Man-Machine Interface. This, within CIP-B, will be an authenticated session, e.g. with the transmission and acceptance of a user name and password. The term 'log on' is used within this document to mean the establishing of user interaction session with a catalogue system supported within the CIP domain. In this context, the word 'session' should not be interpreted in terms of communications sessions or states. A user interaction session is independent of the underlying communications layer. Note further, that as physical communications layers can be broken, it is likely that a user identifier will need to be retained within the CIP domain and therefore, user identifiers will be exchanged.
Users	The user represents the combination of a real human user and the client software that the human user is using to interface with the Retrieval Manager. The CIP is not concerned directly with the client software, although the CIP has to be able to support the tasks that the user wishes to achieve, and it is anticipated that in general, the actual CIP should be transparent to the user. This is analogous to a human not directly interacting with TCP/IP, but being aware that it satisfies the task of data transfer. The only exceptions to this may be when the user is controlling a query or when error or state information is generated by the CIP under anomalous circumstances. The user has a number of important properties, such as a unique identifier, option, etc. (of course this does not exclude an anonymous user having a set of default properties). Generally, the user can perform three types of tasks, either place a query or place an order or control a session. Note that there is also a special type of user in the CIP domain and that is a non-human user such as another Retrieval Manager or a scheduler. These could for example place orders as required, without human interaction, apart from the initial set-up of the schedule.

1.6 References

ICS Compatability: Explanatory

[R1] *Protocol Task Team (PTT) Development Plan*, CEOS/WGISS/PTT/Plan, Issue 1.0, 19 July 1996, Committee on Earth Observation Satellites /Protocol Task Team

[R2] *Interoperable Catalogue System (ICS) User Requirements Document (URD)*, CEOS/WGISS/PTT/ICS-URD, Issue 2.2, March 1997, Committee on Earth Observation Satellites /Protocol Task Team

[R3] *Catalogue Interoperability Protocol (CIP) Specification - Release B*, CEOS/WGISS/PTT/CIP, Issue 2.3, to be ratified February 1998, Committee on Earth Observation Satellites /Protocol Task Team

[R4] *ICS Valids Document*, April 1998, Committee on Earth Observation Satellites /Protocol Task Team

[R5] *ICS Collection Manual*, CEOS/WGISS/PTT/CM, Issue 1.0, April 1997, Committee on Earth Observation Satellites /Protocol Task Team

[R6] *Technical Architecture Framework for Information Management (TAFIM), Volume 3: Architecture Concepts and Design Guidance*, Version 2.0, 30 June 1994, US Defense Information Systems Agency Center for Architecture.

[R7] *Summary of the ECS System Design Specification*, ECS Technical Paper Number 19400436, June 1994, Hughes Information Technology Systems (HITS) ECS Project Office. (document is available at http://edhs1.gsfc.nasa.gov/)

[R8] *CEOSnet Resource Planning and Coordination Document*, CEOS Network Subgroup, Draft Version 0.4, April 1997

[R9] *Information Retrieval (Z39.50): Application Service Definition and Protocol Specification*, ANSI/NISO Z39.50-1995, Official Text, July 1995, Z39.50 Maintenance Agency.

[R10] *DPRS Technical Note/002*, LUK.502.EC21317/TN002, Issue 0.1, 20 September 1995, Logica UK Ltd.

[R11] *ICS Collection Technical Note*, CEOS/WGISS/PTT/CTN, Version 1.0, 8 July 1996, Committee on Earth Observation Satellites, Protocol Task Team

[R12] *Internetworking with TCP/IP: Volume 1, Principles, Protocols and Architecture,* Douglas E. Comer, Prentice-Hall, 1995.

[R13] *Z39.50 Information Retrieval Protocol in the Internet Environment,* Clifford A. Lynch . (Document is available at http://lcweb.loc.gov/z3950/agency/)

[R14] *ISO 8824 - Information Processing Systems* - Open Systems Interconnection - Specifications for Abstract Syntax Notation One (ASN.1), 1987

[R15] *ISO 8825 Information Processing Systems* - Open Systems Interconnection- Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1) 1987

[R16] *RFC 1340 Assigned Numbers*. Reynolds, J.K.; Postel, J.B. July 1992

[R17] *CEO Enabling Services Usage Scenarios Document*, CEO/ES/WP 2200/SD-1/184, Document Issue 1, July 1996, Centre for Earth Observation (CEO).

[R18] *Hypertext Transfer Protocol -- HTTP/1.0 Specification*, RFC Number 1945, May 1996, T. Berners-Lee, R. Fielding, and H. Frysty.

[R19] *ICS Query Performance Estimation, NASA ECS White Paper*, Document 170-WP-014-001 Hughes Information Technology Systems, August 1997.

[R20] *Internet Official Protocol Standards,* Internet Architecture Board Standards Track, Document RFC #2200, Standard #1, J. Postel, Editor, June 1997

[R21] *Queuing Systems*, Volume 1: Theory, Leonard Kleinrock, John Wiley, 1975.

[R22] *ECS User Scenario Functional Analysis*, NASA ECS White Paper, Document 194-00548TPW, Hughes Information Technology Systems, October 1994

[R23] *The Directory - ITU-T Authentication Framework*, ITU-T (formerly CCITT), Recommendation X.509, 1994 (Version 3 extensions)

[R24] *Interoperable Catalogue System (ICS) Guide Design and Protocol Specification*, CEOS/WGISS/PTT/GDPS, Issue 1.0, January 20, 1998, Committee on Earth Observation Satellites /Protocol Task Team

[R25] *Z39.50 Application Profile for the Content Specification for Digital Geospatial Metadata or "GEO"*, Verssion 2.1, April 1998, Federal Geographic Data Committee

[R26] *Transmission Control Protocol*, IETF Document RFC #793, Standard #7, DARPA Internet Program, September 1981

[R27] *Network Security CEOS Context*, CEOS Network Subgroup, (Annex 1 by Grant Miller) September 1997.

[R28] *Order Options Amendment to the CIP Specification*, CEOS/WGISS/PTT/CIP-ORD, Planned to be released October 1998, Committee in Earth Observation Satellites

This page intentionally left blank.

2. CATALOGUE INTEROPERABILITY

This section provides an overview of key issues in catalogue interoperability. The following topics are addressed in this section:

- Purpose and scope of catalogue interoperability
- ICS Concepts
  - Collections data model
  - CIP as a Z39.50 profile
  - Browse data in CIP
  - Product ordering and security
  - Guide Documents in ICS
- Levels of Compliance to CIP, ICS and IGP

2.1 Purpose and Scope of Catalogue Interoperability

ICS Compatibility: Explanatory

The Committee on Earth Observation Satellites (CEOS) is comprised of international space agencies. CEOS promotes the interoperability of space agency catalogues through the definition and development of interoperability concepts. By enhancing the standardization of EO data and information management services, CEOS enables the catalogue services to be more accessible and usable to data providers and data users world wide. EO catalogues services, as defined by CEOS, are as follows:

- search and retrieval of information about EO data products
- order of EO data products
- searches and retrieval of Guide documents that complement the EO data products

Catalogue Interoperability in this context is defined as: the ability to provide a Data User with the appearance of a single, unified catalogue for all participating data providers. In order to provide catalogue interoperability all participating data suppliers must support at least one common method(i.e., API) for accessing functions such as authentication, directory, inventory, guide and order. Each supplier may support additional consumer functional interfaces to support their private data users

Catalogue interoperability may extend beyond just the members of CEOS in promoting data access within a wider community of EO data providers and eventually to non EO data providers.

2.2 ICS Concepts

This section introduces key concepts for the understanding of the CIP, IGP and ICS.

2.2.1 Design Approach: CIP Space, IGP Space and ICS

ICS Compatibility: Explanatory

The PTT design approach considers catalogue interoperability as the loose coupling of a federation of existing catalogue systems using a set of common protocols. The approach provides users the services available at all sites regardless of which site the user established a connection with.

- The Catalogue Interoperability Protocol (CIP) standardizes the services needed for interaction between users and catalogues of EO data products
- The ICS Guide Protocol (IGP) standardizes the services needed for a user to discover EO Related documents (i.e., guide documents).
- The Interoperable Catalogue System (ICS) is a design that uses CIP and IGP as the common protocols between data providers and users of the data.

The objective of implementing CIP and ICS is to provide more users with access to more data more easily. The ICS domain can be seen in Figure 2-1 as divided into two virtual domains;

- ‘CIP domain’ within which CIP messages, consisting of requests and responses, are exchanged between architectural elements.
- ‘IGP domain’ within which IGP messages, consisting of requests and responses, are exchanged between architectural elements.

These two domains are separable and can exist independently from each other. To allow access to both domains, an ICS client was designed with a CIP Client component and an IGP Client Component. To enable consistency in the ICS domain a Collections Management Tool (CMT) was defined to update the contents of both the CIP and IGP domains simultaneously.

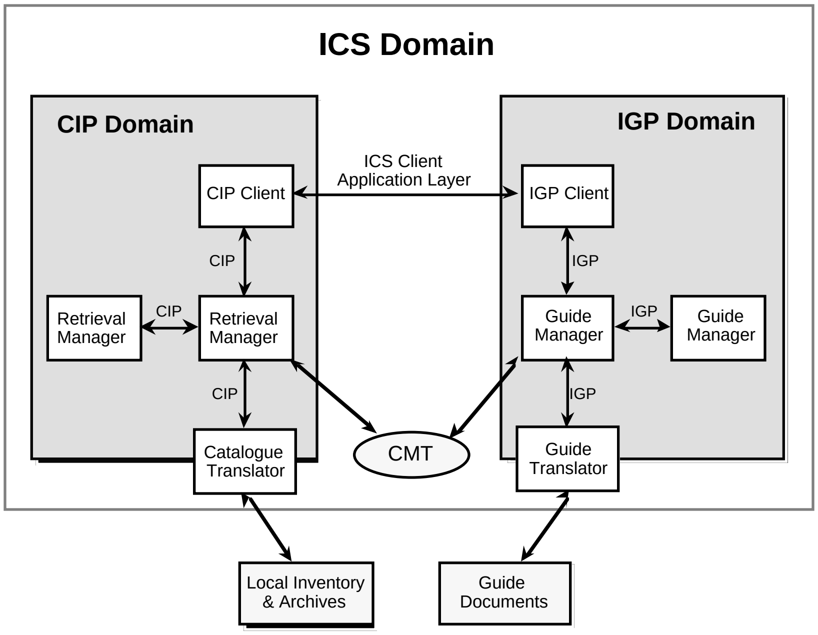


Figure 2-1. ICS Domain

To support transparent access to multiple catalogues, a three tier structure was used to design the ICS space. Clients exchange messages with a middleware layer which in turn interacts with multiple catalogue servers. The middleware provides the routing and translation services to allow client requests to be presented at the multiple heterogeneous catalogues. The middleware is of two types of elements: Managers and Translators. Managers provide an access point for clients and route the requests to the various servers. Translators, bound with the clients and servers, translate CIP or IGP to and from the native protocol of the client or server. Future client and server developments may use CIP or IGP directly and hence not require translators

This approach supports a diversity of clients, and servers. Clients may be used directly by a human user or may be an agency system acting on behalf of a user. Depending on the design of an existing catalogue system, services may be provided by different servers and translators. Because the routing service provided by the Middleware is independent of the type of service, separate translators may be provided for inventory, browse, ordering, and user profiles. This architecture is also applicable for small data providers, such as university research groups, who are unable to provide adequate middleware at their site but still wish to join the ICS domain. Their local catalogue, inventory and guide documents can be made available to the ICS community by the inclusion of appropriate descriptions within another agency's middleware.

CIP Space is a protocol centric view of catalogue interoperability and provides for the loosest coupling needed to achieve catalogue interoperability among a wide community of EO data providers and users of EO data. A range of design solutions is permitted by the CIP and IGP spaces. To provide for a higher degree of uniform services at the cost of additional agreements between agencies, additional design criteria for interoperability are defined in the ICS design document. The additional design definitions pertain to the allocation of functionality and data amongst components, agreement on an underlying communication protocol, and agreement on how to conduct distributed system management of ICS. The difference between CIP Space and the ICS is depicted in Figure 2-2. CIP Space is defined by those CEOS agencies and other federations and organizations which provide catalogue services using CIP and/or guide service using IGP. Those CEOS agencies which provide services, communications and systems management compatible with the ICS design make up the ICS. Its should be noted that while all ICS members must implement CIP, guide handling is considered an optional element of ICS and an ICS members may choose not to implement IGP. Note that other federations may choose to use the ICS design as the basis for their federation.

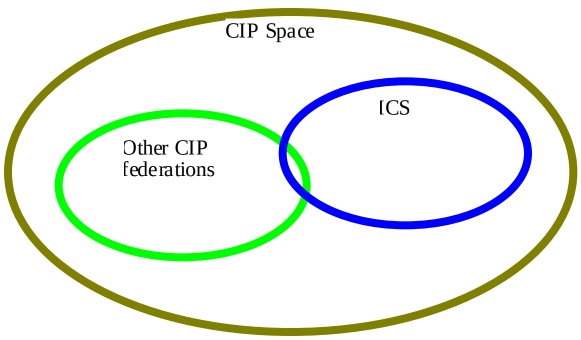


Figure 2-2. VENN diagram of CIP Space and ICS

Assuming query and result routing between geographically dispersed sites (see Figure 2-1), an agreed middleware layer and its interfaces to users and providers needs to be in place. To define such a system, the PTT have established the following CIP, IGP and ICS standards:

- Search Standardization - the functions and procedures how search can be invoked and executed both by end user clients and middleware; the search language and it's syntactical rules;
- Retrieval Standardization - the procedure for retrieval of query results and other objects which need to be presented to the client; an exact description of what is the retrieval content, its format and meaning.
- Attribute Standardization - the list of fields or attributes which can be searched and the definition of their semantic meaning. In addition CIP provides a local attribute mechanism which allows the data producer to extend the set of attributes to include more information to assist the user in the selection of EO products of interest.

- Dynamic Client Configuration Standardization - a mechanism which is used for dynamic client configuration, i.e. to make a client understand which functionality can be invoked at the server side, which attributes are understood, and what are their meanings.
- Order Procedures Standardization - a common way of defining product lists and associated order, packaging and online/offline delivery options; communication of price and accounting information; order and delivery addresses.
- Security Policies Standardization - a reliable and secure mechanism for user authentication and authorization.
- Guide Document Discovery and Retrieval Standardization - a mechanism which allows full text or fielded searches on documents relating to EO data holdings. This mechanism uses HTTP protocols and technology to enable the use of standard web browsers and discovery services such as Alta Vista to discover and retrieve EO related documents of interest. This service provides the general WWW community access to documents describing EO data. This service also provides pointers to EO data related to each document.

2.2.2 Collections Data Model

ICS Compatibility: Explanatory

In an interoperable catalogue environment it is essential to organize metadata by distinguishing user and provider views as well as archive-oriented and theme-oriented structures. It is important to define a mapping between the different views and structures which often will lead to a hierarchical relationship of collections. The ICS data model is based on the notion of collections. A collection may contain descriptors for data products or descriptors for other collections. In addition to the value of collections for presentation of data organization to users, collections provide the mechanism for routing distributed searches. When a collection contains both local and remote members, the Retrieval Manager may search the local site as well as sending the search on to the remote site.

The collection concept is visualized in Figure 2-3 below. The collections in the diagram are numbered so that their relationship can be easily seen; they do not represent the naming of collections in an actual implementation. The terminal collections (labeled ‘1.x’) group the product descriptors as is appropriate. As can be seen the collections can overlap each other and product descriptors can appear in more than one collection. Above the terminal level collections, there are non-terminal collections that group together any number of other collections. The grouped collections do not all have to be at the same hierarchical level and this grouping of collections can continue to any hierarchical level, with existing collections being included at any other arbitrary level. A non-terminal collection could group together terminal collections and other non-terminal collections (as the link between collections 3.1 and 1.5 shows). Also, a terminal collection could exist without a relationship to a higher collection (i.e. collection 1.9), or a non-terminal collection could exist with no relationship to lower collections, in other words a collection without members (i.e. collection 2.5). Collection 1.9 can not be reached by a hierarchical search, but could be located if its URL was made public (an example of such a collection may be a persistent result set or a collection under construction).

Collections can be used to group data together which have a similar semantic theme. All collections support the search mechanisms defined in the CIP. CIP defines two types of searches which a CIP user may request:

- Collection Search: finds collections of interest without searching collections containing products
- Product Search: finds individual product descriptors which may eventually lead to the order of an actual product.

Additionally, the user may request that the search be contained locally to the target Retrieval Manager (i.e., a local search), or request that the search be propagated to other Retrieval Managers based on the collections (i.e., a distributed search).

Two types of collections are defined: registered collections and unregistered collections.

- *Registered collections* are owned by the EO data provider and described by a collection descriptor. These collections support the full range of CIP access services including discovery, navigation, location and searching. There are two types of registered collections which are distinguished by the purpose for which the were created: Archive collections and Theme collections.
  - Archive collections: : This type of collection is likely to be created by data providers to organize their archives and facilitate access to the product descriptors (i.e. analogous to an inventory containing inventory entries
  - Theme collections: This type of collection may be set up by data providers or users who want to organize some of their data into groupings which differ from their provider archive collections (i.e. from the baseline inventory), for the convenience of their users, for example, based on the geographical area covered, the scientific discipline supported by the data, the instrument type, etc.

- *Unregistered collections* are likely to be created and owned by an end user of EO products that has created and populated the collection to obtain a single source of thematic information. This will then enable further analysis or easy access by themselves or other users. These are collections of potentially quite disparate item descriptors of interest to a relatively small user community researching a particular theme, i.e. in the example, the mid-west flood of 1993. Unregistered collections need not be described by a collection descriptor and their collection attributes cannot be discovered in a search of ICS collections.

An unregistered collection may be registered by an ICS data provider. The process of registration may vary significantly between ICS data providers, but minimally requires the creation of a collection descriptor and the insertion of that collection descriptor into the ICS data providers collection structure. Generally the registration of a collection will also involve a scientific review and the transfer of ownership of the collection from the end user to the ICS data provider to ensure accuracy and long-term availability of the registered collection. Further description of the registration process can be found in the ICS Collections Manual [R5].

Note that these category definitions are not mandatory for the CIP to operate, but help to distinguish collection categories for discussion purposes. The CIP does not distinguish between the categories (the Retrieval Manager does, however) and the same CIP search and retrieval services are applicable to all collections. The Retrieval Manager does make use of collections for routing of distributed queries. Standardization of collection definitions is provided as part of the ICS design.

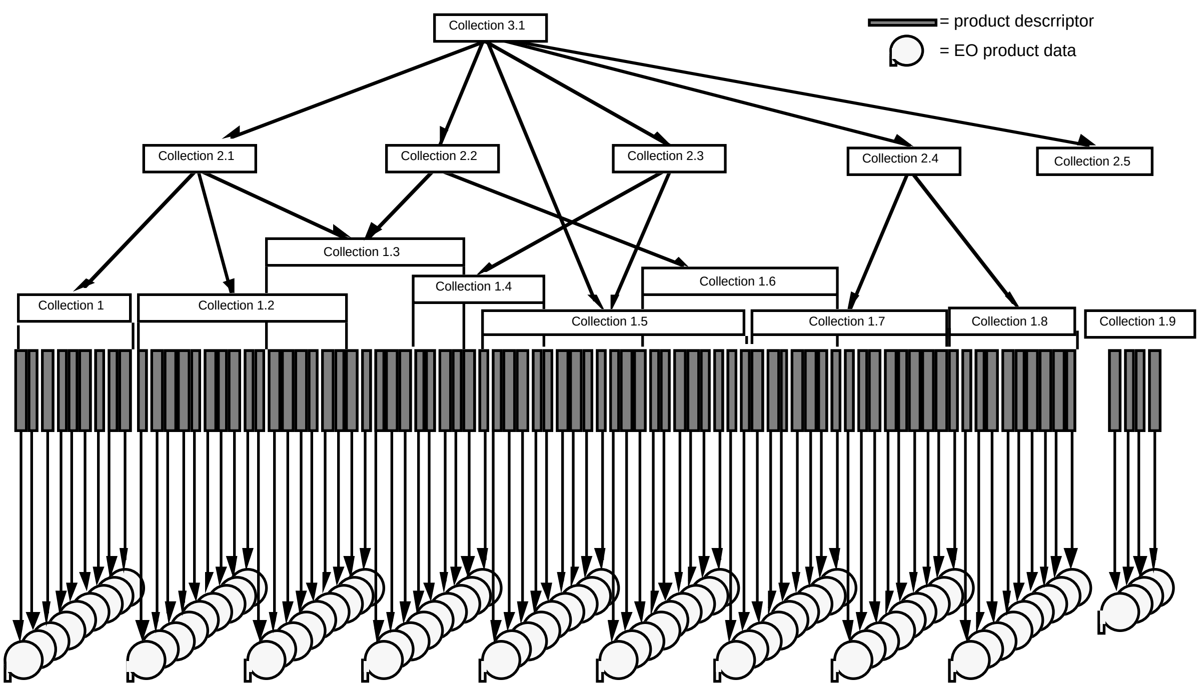


Figure 2-3. The Concept Of a Collections

2.2.3 CIP as a Z39.50 Profile

ICS Compatibility: Explanatory

Based on a set of user requirements and an analysis of existing communication standards, Z39.50 was selected as the base protocol for CIP. The Z39.50 protocol [R9] is designed for information search and retrieval within a generic domain which, together with the powerful services and data structures it supports, makes it an ideal basis of an EO domain search and retrieval protocol.

CIP has exploited and extended the services of Z39.50 to provide distributed searching, extensions to attribute set definitions, and the definition of a secure ordering service. CIP is a profile of Z39.50, i.e. it defines the use of the Z39.50 facilities within the CIP domain and defines the attributes that are used to search and present EO information. Other Z39.50 profiles include GILS, GEO and the Digital Collections Profile. CIP extends Z39.50 for distributed searching by supporting the collection data model discussed in Section 2.2.2 which allows hierarchies of related collections to be constructed and searched.

The GEO Profile supports Geographic Information Systems (GIS) applications and thus is of special interest to users of EO data. For this reason an alignment of the CIP and GEO profiles was made. The objective of this alignment was to allow both GEO and CIP clients to search and retrieve records from databases defined by either profile, and thereby maximize interoperability. The alignment was helped by the similarity of the spatial and temporal attributes of the metadata, but needed to take into account the different data models in CIP and GEO. It should be emphasized that the CIP/GEO interoperability is for search on the intersection of CIP and GEO attributes and the retrieval of item descriptors. There is no interoperability on the more advanced functions of CIP such as ordering and security. Additional support for compatibility is provided by the requirement that Retrieval Managers must support access by any Z39.50 Version 2/3 compatible client.

2.2.4 Browse Data in CIP

ICS Compatibility: Explanatory

Browse data helps users to evaluate EO products. Browse data are typically reduced resolution or summary data versions derived from the EO product data itself. Browse data are delivered to the user via two different mechanisms, dynamically over the network during a user query session, and as an EO product order. The second case allows users to order the Browse data from an archive system to be delivered separately from their query session. This means that the user can then store and access the data locally rather than dynamically over a network. It is important to note that although most catalogue systems will provide some form of reduced data retrieval, it is not a mandatory CIP service. The form and content of browse data is dependent on the nature of the associated EO data and the data selection criteria necessary for a science discipline to evaluate the EO data. Browse data in the CIP is seen as one of the following forms:

- Browse attribute - simple attribute containing the actual Browse data.
- Browse compound - compound attribute containing attributes describing the Browse data, including the simple Browse data object.

2.2.5 Product Ordering and Security

ICS Compatibility: Explanatory

CIP supports a wide range of ordering services, e.g. the specification of order options and provisions for authentication and non-repudiation of orders. A user can retrieve the order options associated with a product, where order options may be processing as well as packaging options. CIP allows a local order handling system to define or refer to its own local order options. But the CIP also contains a standard mechanism to define order options..

A user can request a quote for a specific order and submit the order. The order process is monitored by the Retrieval Manager and can be queried later by the user to determine the status of the order. To support ordering of data for which a user must have privileges or for orders which the user will be charged, a authentication scheme has been defined. The authentication supports digital signatures using either a shared (symmetric) key approach or an public (asymmetric) key approach. Authentication allows the Retrieval Manager to identify the user with an appropriate level of confidence and enables the Retrieval Manager to log the authenticated user requests to provide non-repudiation. The CIP security approach avoids the need to transfer password information over the network. Future enhancements to CIP anticipate the ability to support the transfer of financial information to support billing.

If a provider decides to create a collection containing browse products, a user would be able to order browse products via the CIP. In this case the browse product would be considered as any other data product and not any longer as “browse” as defined in the previous section.

2.2.6 Guide Documents in ICS

ICS Compatibility: Explanatory

Much of the metadata for EO data collections is not easily stored in a structured form. This information is stored in documents called Guides. Since Guide documents provide information that is required for the understanding of some EO data collections, they must easily be accessed via ICS mechanisms that provide search and retrieval of catalogs. Guide documents also provide a human readable descriptions of EO data collections and are often used by new EO data users as a discovery mechanism to identify collections of interest. It is the goal of the ICS to make this discovery mechanism as simple and widely available as possible to extend the uses of EO data to communities which have not traditionally used EO data. This goal has resulted in the definition of a guide system in ICS which uses ICS Guide Protocols (IGP) based on HTTP and enables general purpose Internet Search/Discovery Engines such as Alta Vista to locate EO Guide documents either by free text or attribute value searches.

This system is not based on Z39.50 and is not a mandatory capability of an ICS node. However there is a strong linkage between the CIP client/retrieval manager and the HTTP based client and indexing method for Guide Documents. To allow coordinated access to catalogues and documents an ICS client was designed with a CIP Client component and an IGP Client Component. The ingest of documents and collections into the ICS is coordinated by the Collection Management Tool (CMT) to assist in maintaining the consistency of the collection descriptor and the HTTP index that enable search and access of Guide Documents. Further details of this ingest process are discussed in the ICS Collections Manual [R5]. The specific design of the guide system can be found in the ICS Guide Design and Protocol Specification [R24].

2.3 Levels of Compliance to CIP, IGP and ICS

ICS Compatibility: Explanatory

The ICS SDD and CIP Specification are detailed documents with many services and mechanisms specified. As discussed in Section 2.1, agencies may choose to implement a wide range of these services in their CIP Clients and Retrieval Managers. It is critical for the designers and implementors of these software components to understand what capabilities are critical to the minimal operations of the ICS and must be implemented in all components versus those capabilities which are optional. In addition it is assumed that various CIP based components will be available either as shareware or commercial software. The developers of ICS or other CIP federations will need a method to categorize and select among these available components. For this reason, compliance levels have been defined within both the ICS SDD and the CIP Specification.

- *ICS Compliance,* which is defined in Section 9 of the ICS SDD, discusses the required minimum configuration for a site which wishes to be considered ICS compatible. The minimum site is based on the CEOS policy.
- *CIP Compliance,* which is defined in the CIP Specification, discusses specific CIP messages and parameters must be supported by a CIP Client or an RM.

These compliance concepts are interdependent since in order to support a specific ICS service, the RM and CIP Client must support the CIP messages which enable that service.

IGP has only one compliance level which is full compliance so no system that does not implement the full IGP Specification can be considered IGP compliant.

3. FUNCTIONAL VIEW

The ICS Functional View describes the partitioning of ICS software elements, application level protocols used between the elements, and a discussion of how the elements might be arranged on physical hosts. Included is a Functional Framework which shows the arrangement of ICS elements and identifies the services provided by each element.

3.1 Architecture Foundations

This section discusses several topics which lay the ground work for the Functional Framework which is presented in the next section. This section discusses the conceptual design which was used during the requirements development; discusses the ICS as a generic three-tier architecture; discusses how the choice of Z39.50 as the underlying protocol for CIP affects ICS; and discusses how guide document management is performed using http as a base protocol.

3.1.1 ICS Domain Models

ICS Compatability: Explanatory

The ICS domain consists of two protocol domains, the CIP Domain and the ICS Guide Protocol (IGP) Domain as shown in Figure 3-1. The ICS URD [R2] defines ICS around a common protocol, CIP, and the elements which speak CIP. The URD discusses the CIP domain as a virtual 'CIP space' within which CIP messages, consisting of requests and responses, are exchanged between architectural elements: clients, servers and middleware. The CIP space is bounded by interfaces to translators, that is, software that converts CIP messages into external formats. (See also Section 2 for a discussion of CIP space.) As described in the ICS Guide Design and Protocol Specification [R24], the IGP provides document management using IGP messages that are exchanged between clients, servers, and middleware using an architectural approach that parallels CIP. Thus, the ICS Client is represented by the CIP Client and the HTTP Client; ICS middleware is represented by the CIP Retrieval Manager and Catalogue Translator along with the IGP Guide Manager and Guide Translator; and the ICS-related servers are represented by the respective local inventory, archive, and guide documents. The partitioning of functionality between the architectural components is discussed further in the next section

The URD [R2] also defines the ICS from a systems perspective in terms of providing a seamless interoperable environment whereby users can conveniently access all available data, whoever it belongs to, requiring both technical and administrative problems to be addressed by ICS. The ICS also must provide management of data access; user authentication, user session management, problems of routing (sending client requests to the correct server) and collation (accumulation of different sets of results). This leads to the concept of middleware, which sits between the client(s) and server(s).

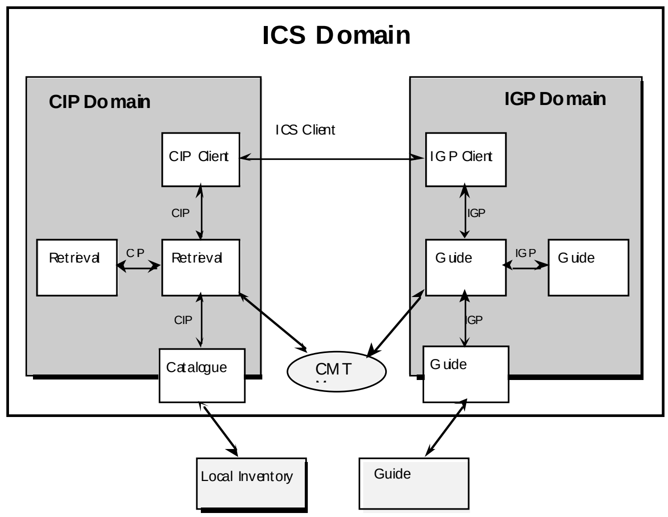


Figure 3-1. ICS Domain

The URD also comments that dividing the functionality between client, middleware, and CIP specific components of the server, is not to be resolved at the level of the URD. A range of architectural solutions can be envisaged, with different distributions of functionality and emphasis, and these should be traded off and evaluated at the level of system design analysis. These issues are addressed within this SDD.

3.1.2 Three Tier Architecture

ICS Compatability: Explanatory

The ICS approach is a prime example of a generalized concept for distributed information systems which is referred to as the Three-Tier interoperability architecture (see Figure 3-2). The three-tiered application model splits an application into its three logical component types, clients, middleware, and servers. The three tiers refer to the three logical component parts of an application, not to the number of machines used by the application. There may be any number of each of the component types within an application. Application components can be shared by any number of application systems. The application components communicate with each other using a distributed computing infrastructure. The three-tier architecture supports provider sites which manage their own resources and processes.

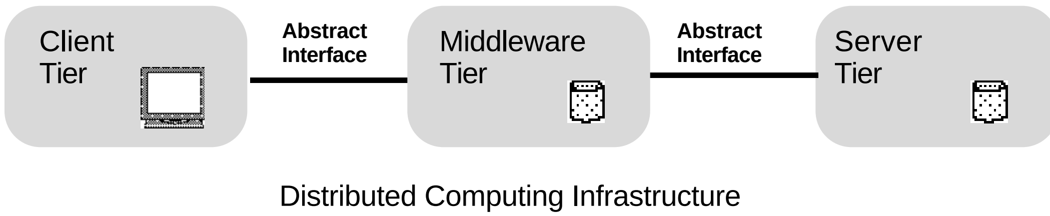


Figure 3-2. Three Tier Architecture

The Client tier consists of the user environment. A client component contains the logic which presents information to an external source and obtains input from that source. In most cases the external source is a human end user working at their own computer, although the external source might also be process-oriented. The client logic generally provides menus of options to allow the user to navigate through the different parts of the application, and it manipulates the input and output fields on the display device. Frequently, the presentation component also performs a limited amount of input data validation.

Middleware applications are entities that are used to model “business” processes. For ICS the “business” is catalogue interoperability. ICS middleware models the data and processes necessary for providing the client transparent access to heterogeneous catalogues at physically distributed locations. This is accomplished through interaction with the communication structure and an abstract interface, i.e., CIP.

The Server tier contains the logic which interfaces either with a data storage system, or with some other type of external data source such as a data feed or an external application system. Server functions to a data provider are generally invoked by the middleware, although in some applications they may be invoked directly by a presentation component.

A common mistake made in planning for three-tier architectures is assuming that, because a client/server application has three logical tiers, it must be implemented in three physical tiers. This misconception is unfortunate because it limits opportunities to capture application-specific requirements in the areas of user platform, network support, platform price/performance, development tools, and management capabilities. The three logical components can be distributed in many different ways to provide optimum configurations for application maintenance and support. The key is that once the three-tier architecture has been embraced in the design of an application, the partitioning of that application into its physical instantiation should be done in a way that optimizes performance, security, integrity, maintainability, and management.

Distributed computing infrastructures for three tier architectures is a current topic of much research and development that promises to provide transparent, dynamic access to services. For this purpose ICS relies instead on a proven and deployed infrastructure of Internet services, including DNS directory service. These are described in Section 5.

3.1.3 Z39.50 as a Base Protocol

ICS Compatability: Explanatory

Communications using the CIP are stateful, client/server sessions based on the Z39.50 Protocol [R9]. Z39.50 is an ANSI standard for information Retrieval. The Z39.50 protocol was chosen based on analysis of the URD requirements versus the services already defined in the Z39.50-1995 specification [R10]. CIP uses a subset of the Z39.50 messages to define an EO profile. CIP provides a set of EO specific attributes, as well as an extension of Z39.50 messages, through the provision of search control commands particular to the *collections* which are defined for ICS. CIP is the interface protocol for communications for collection and product searching.

3.1.4 HTTP as a Base Protocol

Communications in the IGP domain are accomplished using a set of messages that are based on the Hypertext Transfer Protocol (HTTP) protocol. The HTTP is an application-level protocol for distributed, collaborative, hypermedia information systems. It is a generic, stateless, object-oriented protocol which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods. A feature of HTTP is the typing and negotiation of data representation, allowing systems to be built independently of the data being transferred. HTTP allows an open-ended set of methods that indicate the purpose of a request. It builds on the discipline of reference provided by the Uniform Resource Identifier (URI), as a location (URL) or name (URN) , for indicating the resource to which a method is to be applied. Messages are passed in a format similar to that used by Internet mail as defined by the Multipurpose Internet Mail Extensions (MIME). Within the scope of the IGP Domain, IGP messages based on the HTTP protocol are passed between the HTTP Client, the Guide Manager, and the Guide Translator in order to provide guide document searching and retrieval.

3.2 ICS Functional Framework

The ICS Functional Framework provides the top-level design of the application processes which make up ICS. In this section, the context of ICS is described first. Next, the ICS Elements are introduced according to their role in the three-tier architecture introduced in the previous section. Lastly the ICS Functional Framework which shows several sites is presented.

3.2.1 Context of ICS

ICS Compatability: Explanatory

The context of ICS is defined by the interfaces to systems external to ICS. Graphically this is displayed in Figure 3-3. (Note that in the convention of context diagrams, interfaces between external items are not shown, e.g., no interface is shown between Users and CEOS Agency Systems, as these are out of the scope of ICS.) There are five types of external systems with which ICS has interfaces: *Users, CEOS Agency Systems, Web Indexers, Other CIP Based Federations, GEO Based Systems, and Other Z39.50 Based Services.*

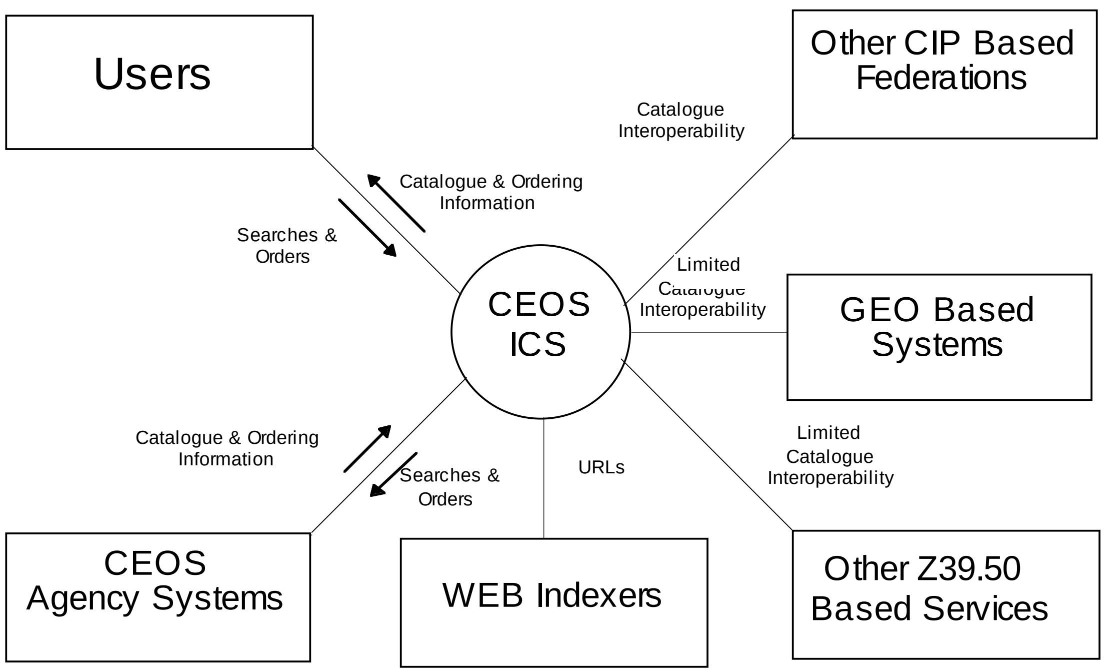


Figure 3-3. ICS Context Diagram

Users in Figure 3-3 are the human users which interface with ICS using a variety of clients. *Users* submit searches and orders to ICS and ICS responds with catalogue and ordering information respectively.

CEOS Agency Systems in Figure 3-3, are those data providers which provide catalogue services to ICS. When searches and orders are appropriately routed to the *CEOS Agency Systems*, they will respond with catalogue and ordering information respectively. *CEOS Agency Systems* includes affiliate CEOS agencies as well. Services assumed to be part of the *CEOS Agency Systems* are a catalogue service, order handling, user profiling, archiving, and site system management services.

Web Indexers in Figure 3-3, such as Alta Vista, can access ICS document collections via URLs.
Other CIP Based Federations in Figure 3-3, are groups of other data providers which provide catalogue interoperability using CIP. These other federations, may or may not have a design similar to ICS.

GEO Based Systems in Figure 3-3, are groups of data providers which use GEO-based Z39.50 profiles. Significant effort has been undertaken to bring CIP and GEO into a level of alignment that maximizes interoperability.

Other Z39.50 Based Services in Figure 3-3, are other groups of data providers which provide interoperability using Z39.50 but not necessarily using CIP. Catalogue interoperability with these providers will be limited to those supported by the specific Z39.50 profile which the providers are using.

3.2.2 Introduction of ICS Elements

Before describing how the ICS elements are connected to make a system, the individual elements are introduced. The elements are organized following the three tier architecture introduced earlier: Clients, Middleware and Servers.

In addition to the elements described in this section, the ICS is hosted on a variety of hardware platforms and operates over a variety of networks. These issues are considered in other sections of the SDD, e.g. Sections 5 and 8.3.

3.2.2.1 ICS Client

ICS Compatability: Explanatory

The main element of the client tier is the *ICS Client* (Figure 3-4). The *ICS Client* is used by a human user to compose searches and orders which are directed to the middleware elements. The *ICS Client* consists of a presentation layer, a local management layer, an application layer, and two communication protocol clients for CIP and IGP.

The CIP client and IGP client layers accept internal messages from other parts of the *ICS Client* and create messages formatted in the CIP and IGP protocols. Also they receive CIP and IGP messages and create internal messages. The *ICS Client* may also provide access to other protocols. For example, the client may incorporate FTP as a mechanism for accessing products made available through the ordering process.

The *application layer* manages the other parts of the *ICS client* including state management most of the user perceived functionality. The *application layer* accepts inputs from the user via the *presentation layer.* The *application layer* does all local processing on the data, e.g., query formulation, and passes a message to the CIP or IGP client for transmission to the server. The *application layer* configures the *ICS Client* using data from the *Retrieval Manager* Explain database. The *application layer* is the bridge between the CIP domain and the IGP domain. For example, a URL for a guide document (that it related to a collection) received via CIP is accessed using IGP.

The *presentation layer* deals with how the information is presented to the user, including all issues related to HMI, as well as dealing with certain format specific issues, e.g., displaying browse imagery.

The *local management layer* provides functionally for local data management, e.g. saving a result set, or converting result sets into orders.

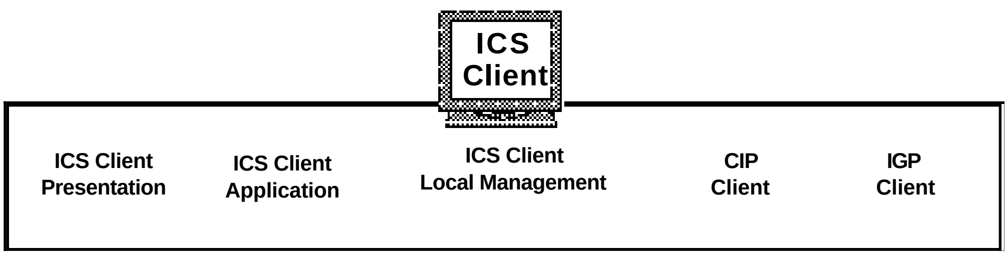


Figure 3-4. ICS Client

3.2.2.2 Retrieval Manager

ICS Compatability: Explanatory

The middleware tier provides the ICS mechanisms which enable access to the back end data providers that supply product and/or related document information. The ICS elements which comprise the middleware tier are the *Retrieval Manager* and a series of *Translators* for product and document access and the *Guide Manager* for document access.

The *Retrieval Manager* is the main element of ICS (Figure 3-5). The *Retrieval Manager* holds the *collections* which are key to the distributed searching. The *Retrieval Manager* is key as it is the common element which routes the CIP messages. Furthermore the *Retrieval Manager* enables the following key ICS services:

- dynamic configuration of clients with semantic attributes
- *collection* searching

- routing to distributed sites
- access to diverse and heterogeneous catalogues
- flexible extensions and incorporation of additional services
- generation of 'secondary' searches from original searches for theme collections
- monitoring of 'secondary' searches
- retrieval and combination of search results from 'secondary' searches

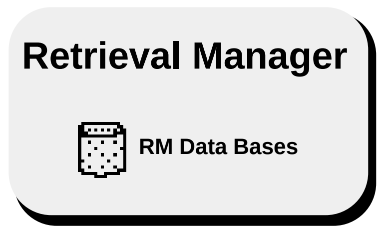


Figure 3-5. Retrieval Manager

3.2.2.3 CIP Translators

ICS Compatability: Explanatory

The second middleware element type are the *CIP translators* (See Figure 3-6). *CIP translators* provide a *Retrieval Manager* with access points to data providers. *CIP Translators* convert CIP messages into the local messages of the data provider. The portion of a *CIP translator* which handles CIP is common to other *CIP translators* across ICS. The back half of a *CIP translator* is unique to the data provider and the internal protocols that are supported by data provider.

CIP translator configurations will be defined based on the local data provider needs. Some data providers will use a single translator to convert all CIP messages to the data provider’s local protocol messages. Other data providers will have multiple *CIP translators* to support different CIP messages. For the multiple translator approach the following types have been defined:

- *Catalogue Translator*
- *Order Handling System (OHS) Translator*
- *User Profile System (UPS) Translator*

Given the variety of *CIP Translator* configurations, the *CIP Translator* elements are shown generically in Figure 3-6.

If a data provider uses CIP as the internal protocol, a *Translator* is not required.

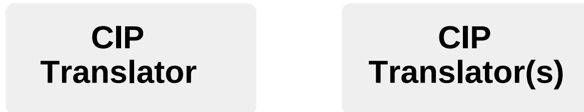


Figure 3-6. CIP Translators

3.2.2.4 Guide Manager Elements

ICS Compatability: Explanatory

The third type of middleware is the ICS Guide Manager as depicted in Figure 3-7. This consists of two components, the Guide Server and the Guide Indexer. The Guide Server interacts with the IGP Client and the Guide Translator to perform searches and retrievals for documents. It also maintains the Guide Metadata Database and the Attributes Defaults File for guide management. When a document is inserted into ICS, the Guide Indexer controls the indexing of the guide document so that free-text and fielded searches can be performed against the document via the ICS Client.

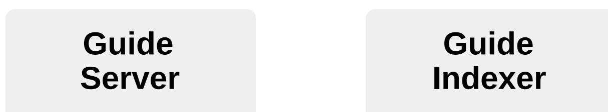


Figure 3-7. ICS Guide Manager

3.2.2.5 Guide Translator

ICS Compatability: Explanatory

Similar to the *CIP translators*, a *Guide translator* element is defined in ICS (See Figure 3-8). The Guide Translator converts document search and retrieval services for the local Guide Document Archive.

Figure 3-8. Guide Translator

3.2.2.6 Ser ver Elements

ICS Compatability: Explanatory

The last tier of the three tier architecture includes the server elements (Figure 3-9). Server elements are the elements which manage access and storage of the EO data of interest to the users. Strictly speaking the server elements are outside of the scope of ICS because they belong to data providers and are contained in the CEOS Agency Systems block on the ICS Context Diagram. Four server elements were introduced previously: Catalogue, OHS, UPS, and Guide Document Archive. Additional elements are the Archive and the Site System Management (SSM). The Translators may not directly interact with the archive but for complete depiction of the ICS design, an archive holding the actual product data is needed. The SSM is present for those data providers which provide coordinated, on-line management of the distributed processing environment.

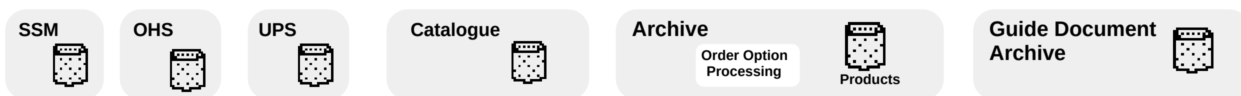


Figure 3-9. Server Elements

3.2.2.7 Other ICS Elements

ICS Compatability: Explanatory

There are other important elements which do not neatly fit into the three-tier architecture but are required for the ICS (Figure 3-10). The *ICS Site Administrator (ISA)* is the human operator that performs all tasks needed to establish and maintain a *Retrieval Manager* and, potentially, a *Guide Manager*. In practice, this is more than one person as the tasks are of various types: a scientist for *collection* definition, a database expert for maintaining the *Retrieval Manager* or *Guide Manager* databases, a system operator for diagnosis and correcting operational activities, etc. For convenience purposes all of these tasks are performed by the *ISA*.

The *Collection Management Tool (CMT)* (Figure 3-10), perform tasks involved with populating and maintaining the data in the Retrieval Manager. These tasks involve translating *collection* or directory information into CIP *collection* format and checking for valid entries. If the ICS site is maintaining guide documents, then another task includes entering guide URLs in the collection descriptors to relate guides to collections. The *CMT* is used in conjunction with the data base administration tools which the *Retrieval Manager* provides.

The *Monitoring and Control Tools (MCT)* (Figure 3-10), provide the machine-to-machine interface for integrating the operations of the *Retrieval Manager* with the operations of a site. Through the *MCT*, the monitor and control operations of the *Retrieval Manager* become part of the *SSM* operations for the data provider. (The *MCT* has been identified as an element in the ICS but specific functionality will be detailed in future releases of the SDD.)

Figure 3-10. Other ICS Elements

3.3 Catalogue Interoperability Protocol (CIP) Domain Design

The Catalogue Interoperability Protocol (CIP) is the main method by which the objective of catalogue interoperability is achieved in ICS. CIP is the main interface method between ICS elements. CIP is elaborated in the CIP Specification [R3]. The CIP Specification defines the services and attributes which are specific to EO data catalogues and common across *Retrieval Managers* and *Translators* which speak CIP.

3.3.1 CIP System Design

The following sections discuss configurations of CIP elements from the perspectives of conformance to the protocol, interoperability with existing systems, and discovery of data within ICS.

3.3.1.1 Maximum ICS CIP Site

ICS Compatability: Explanatory

This section defines how the ICS elements are configured into a data provider site that needs search capability for collections and products. The elements were defined in Section 3.2. The site described here is a maximum site in that it contains all ICS elements (Figure 3-11) from the CIP domain. Particular sites have subsets of the elements. Variety amongst sites is addressed in the next section. The top half of the figure shows those elements which are strictly part of ICS. The bottom half of the figure shows those elements which exist within the *CEOS Agency Systems* and for the purposes of this discussion are labeled as ICS Related Elements.

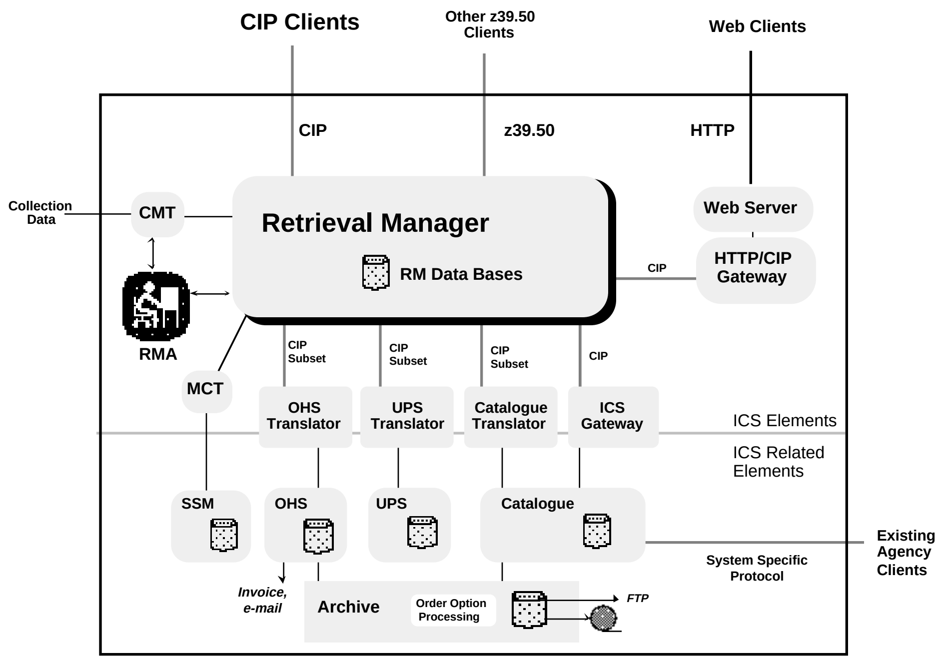


Figure 3-11. Maximum ICS CIP Site

The heart of an ICS site is the *Retrieval Manager*. Virtually all CIP communication through the site passes through the *Retrieval Manager.* A variety of Clients are shown interacting with the *Retrieval Manager.* A connection to an *ICS Client* is shown. WWW access is provided through a Web Server and *HTTP/CIP Gateway* interacting with the *Retrieval Manager*. The *Other Z39.50 Clients* are those clients which communicate using a profile of Z39.50, Version 3, but not CIP. These clients will be able to access some but generally not all of the *Retrieval Manager* services. An additional client is also shown in Figure 3-11, *Existing Agency Clients* will continue to access the site's catalogue. The site *catalogue* provides access to ICS through an *ICS Gateway*. This allows users of existing agency clients to perform searches and ordering of ICS items.

The *Retrieval Manager* interacts with a series of *Translators* to access the data providers services. Figure 3-11 shows one appraoch to translator configuration at a site (See Section 3.2.2.3 for other approachs). Between the *Retrieval Manager* and the *Translators*, a subset of CIP messages particular to the *translator* are used for communications. Only a subset is supported because the *Retrieval Manager* routes queries to the *Catalogue* and orders to the *OHS*, therefore an *OHS Translator* is not designed to respond to queries. Between the *Translators* and the provider site services, the protocols for communication within the site are used. The data provider elements perform the necessary catalogue services or interact with other site resources, e.g., archive, to respond to the requests coming from the *Translators*.

Several other elements are shown in Figure 3-11 which pertain to keeping the ICS site operational. The *ISA* is shown interacting with the *Retrieval Manager* to establish and maintain the *Retrieval Manager*. This is done through a Graphical User Interface (GUI) provided by the *Retrieval Manager*. The *ISA* uses the *CMT* to populate and maintain the data in the *Retrieval Manager*. To support integration of *Retrieval Manager* operations with the *SSM*, an *MCT* is provided to monitor an control *Retrieval Manager* operations as directed by the *SSM*.

3.3.1.2 ICS as a System

ICS Compatability: Explanatory

Figure 3-12 shows the ICS Functional Framework. The framework shows a variety of arrangement of elements at ICS sites. Specific arrangement of ICS elements provided at ICS sites will vary over time. That is Clients, *Retrieval Managers* and even Agency holdings will be added, or taken away, dynamically, during the ICS operations. The ICS must enable this dynamic configuration. Therefore the ICS Framework shown in Figure 3-3 is only representative of how the ICS will be implemented at any given time.

The ICS Framework shows four sites which emphasize different roles at the sites. The four roles are described in this section: *Retrieval Manager* with an Included Catalogue, *Retrieval Manager* as a gateway to Multiple Catalogues, *Retrieval Manager* with only Theme Collections, and *Retrieval Manager* as a Gateway to an Agency Data System.

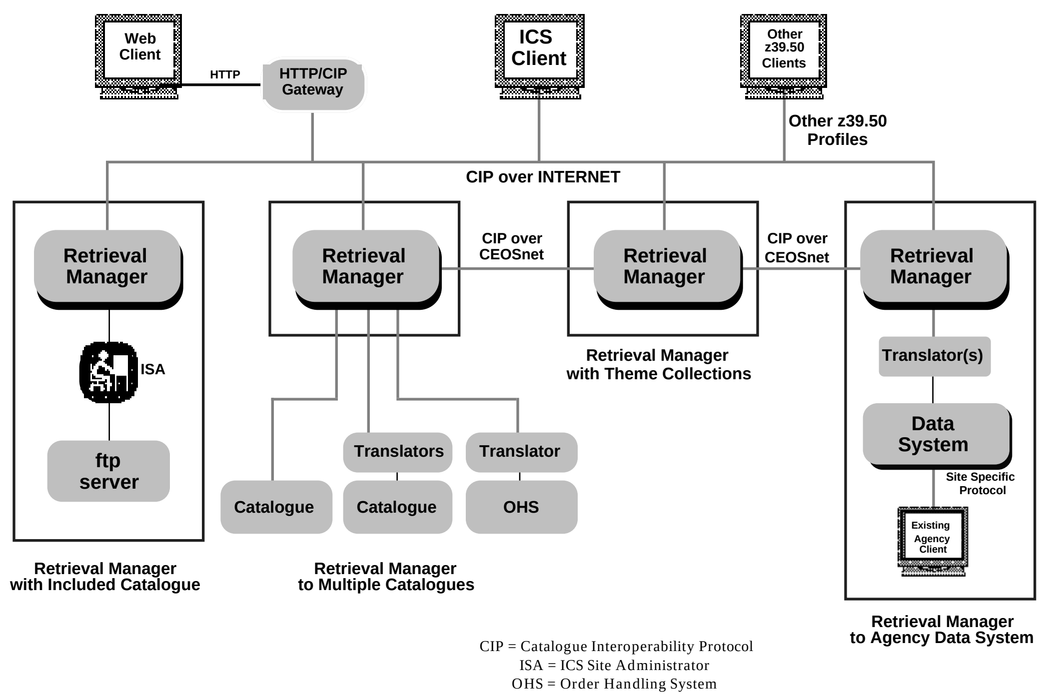


Figure 3-12. ICS CIP Functional Framework

Retrieval Manager with an Included Catalogue. The middle configuration in Figure 3-12 shows how a *Retrieval Manager* is used to provide access to a more limited archive with minimal resources. In this case, the *Retrieval Manager* is a standalone system which provides the *collections* and *product* searches. No *Translator*s are used. The *Retrieval Manager* contains the metadata for *collections* and *products*. The user then retrieves the actual *products* via ftp, for example, from the archive. Ordering and authentication need not be implemented at this site. Setting up this form of a *Retrieval Manager* may be simplified by implementing the *Retrieval Manager* databases as a set of ASCII indexed files. Administration of the indexed files is accomplished through the *Retrieval Manager.*

Retrieval Manager as a gateway to Multiple Catalogues. A *Retrieval Manager* may be used as a gateway to multiple CIP accessible catalogues. The *Retrieval Manager* serves as the access point for a cluster of catalogues. The cluster of catalogs for the *Retrieval Manager* are considered local to that *Retrieval Manager* even though they may be geographically distant. The *Retrieval Manager* holds the collections which include products from the local catalogues. The *Retrieval Manager* may also reference collections in other *Retrieval Managers*, i.e., remote collections. Collection searches are performed in the *Retrieval Managers*. Product Searches are forwarded to the catalogues in the cluster.

Retrieval Manager with only Theme Collections. An additional variant of the *Retrieval Manager* implementation is an ICS router node in the ICS Framework. A *Retrieval Manager* as a router functions as a "well known" ICS resource enabling users to find other *Retrieval Manager*s for specific data providers. An ICS Router is a middleware node containing *collections* but no *product* metadata. For Release B, an ICS router will be used for the ICS Global Node providing a "well known" resource to access all ICS *Retrieval Managers* (see Section 3.3.1.4 for a discussion of the ICS Global Node).

Retrieval Manager as a Gateway to an Agency Data System. For Agencies with existing catalogue and archive systems, the *Retrieval Manager* holds the *collections* which are particular to the ICS domain. The *product* metadata which is in the agency catalogue database and is accessed through a *Catalogue Translator* dynamically to satisfy queries. Likewise, the *Retrieval Manager* in this case holds user management data for only the ICS domain users (which will include the other *Retrieval Managers*). The *Retrieval Manager* interfaces to existing *User Profile Systems* and *Order Handling Systems* through *Translators*.

Multiple configurations for *Retrieval Manager* -to-*Translator* interfaces are supported. As shown in Figure 3-12, a *Retrieval Manager* can direct CIP messages to more than one*Translator*. The converse is not true, namely a *translator* can not be served by more than one *Retrieval Manager*. Limiting the *Translator* in this fashion reduces the burden on the *translator* to manage CIP messages from more than one source, making the *translator* a simpler design. The purpose of the multiple transaltors for a single *Retrieval Manager* allows clusters for*catalogues* to be served by a single *Retrieval Manager*. The multiple catalogues may be at different sites but under the same administration. Or a site may host a *translator* and *catalogue* and rely on a different site and organization to manage the *Retrieval Manager.* If a *catalogue* needs to provide access to multiple *Retrieval Managers*, multiple *translators* are provided as the existing *catalogue* is designed to handle multiple clients.

3.3.1.3 CIP/GEO Interoperability

This section describes the approach to supporting interoperation between the CIP domain and GEO - another Z39.50 based protocol.

3.3.1.3.1 Overview

ICS Compatability: Explanatory

One reason for basing CIP on Z39.50 is the compatibility with other Z39.50 based systems. The GEO Profile [R25] is the basis for several systems which provide access to geospatial metadata conforming to the US Federal Geographic Data Committee (FGDC) standard. CIP - Release B and the GEO profile have been aligned to support interoperability on a limited basis between GEO clients and CIP servers and between CIP clients and GEO servers. This alignment increases the data accessible to both GEO and CIP users.

There are differences between the GEO and CIP models. GEO is a 2-Tier, client-server profile. CIP extends the Z39.50 protocol to a 3-Tier, distributed model (see Figure 3-11). The GEO data model consists of datasets held at GEO servers. GEO datasets correspond to CIP Products. CIP Products are aggregated into Collections. CIP allows aggregation of Collections to support Theme Collections and to support distributed searching. Due to these differences in architectures and data models, there must be constraints on which collections/servers are targeted when crossing the CIP/GEO domains. These constraints are discussed in the following sections.

References to CIP Collections in GEO systems is done using a CIP collection URL as defined in [R3]. References to GEO servers in CIP Retrieval Managers is done using a URL as defined in [R25]. A GEO data administrator is responsible for maintaining the directory information of CIP collections in their system. The ISA is responsible for maintaining the collection database entries for GEO targets listed in the ISA's Retrieval Manager.

3.3.1.3.2 GEO to CIP Interoperability

ICS Compatability: Explanatory

To make CIP data available to GEO users, ICS must provide GEO based systems access to Retrieval Managers. To support access to distributed GEO servers, several GEO based systems maintain some form of directory, which can be referred to as a Directory Of GEO Server

(DOGS). DOGS serve as an access point for a client to send a single search which will be directed to one or more GEO servers (See Figure 3-13). A DOGS fetches the results from the multiple GEO servers and return a single response to the user's client. For CIP Interoperability, Provider Archive Collections are listed in the DOGS. (Note that Theme Collections cannot be listed in the DOGS as the GEO data model does not allow collections of collections.) The DOGS may allow for discovery of the collections within the DOGS and may also allow for searches to be propagated from the DOGS to the GEO servers.

When a DOGS determines that a GEO search should be targeted at a CIP Provider Archive Collection, a Z39.50 Version 2 session is established with the Retrieval Manager which provides access to the collection and a search request using the GEO attribute set is sent. Because the GEO and CIP attribute sets have been partially aligned the Retrieval Manager is able to pass the search request to the appropriate catalogue as a CIP local product search request. (For details on the behavior of non-aligned attributes, see the CIP Specification [R3].)

The arrangement of elements in the GEO world allows distributed searching using a 2-Tier protocol. Distributed searching is accomplished using two, 2-Tier client-server interactions with the DOGS in the middle. The DOGS accept requests from a user's client via an application protocol, e.g., http. Using the directory of servers held in the DOGS, the DOGS forms Z39.50 sessions with one or more servers using the GEO profile. Each session is a client-server session with a GEO server. The DOGS are the key item in these approaches to distributed searching.

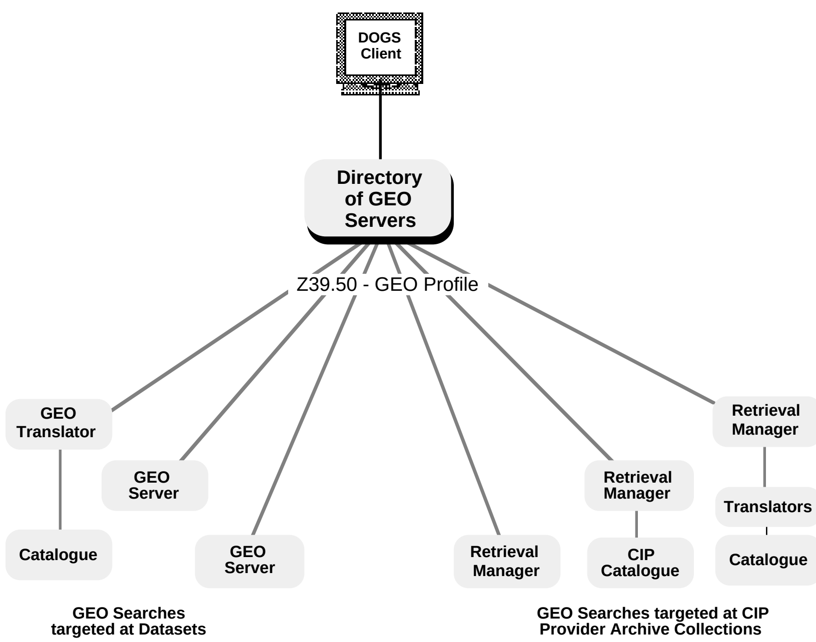


Figure 3-13. CIP Interoperability with GEO

3.3.1.3.3 CIP to GEO Interoperability

ICS Compatability: Explanatory

To make GEO data available to CIP users, GEO based systems must provide access to GEO servers for ICS *Retrieval Managers*. GEO servers provide search information retrieval services on GEO datasets. The GEO datasets are listed in a *Retrieval Manager* collection database as a product in a collection. These “GEO collections” can then be targeted for searching either directly or through inclusion into Theme collections. When a user targets a product search at a collection which contains a GEO dataset, the *Retrieval Manager* will establish a Z39.50 Version 2 session with the GEO target. The product search will be sent using the CIP attribute set.

The interoperability approach allows mirroring a DOGS registry as a collection of product archives within a *Retrieval Manager* as shown in Figure 3-14. There is the potential here to build a richer collection structure of GEO servers than is currently available through a DOGS registry. GEO servers could be grouped into thematic sub-collections. This structure would only be visible to true *ICS Clients* via collection hierarchies. A *ICS Client* would submit a product search to a selected set of GEO servers via the collection hierarchy.

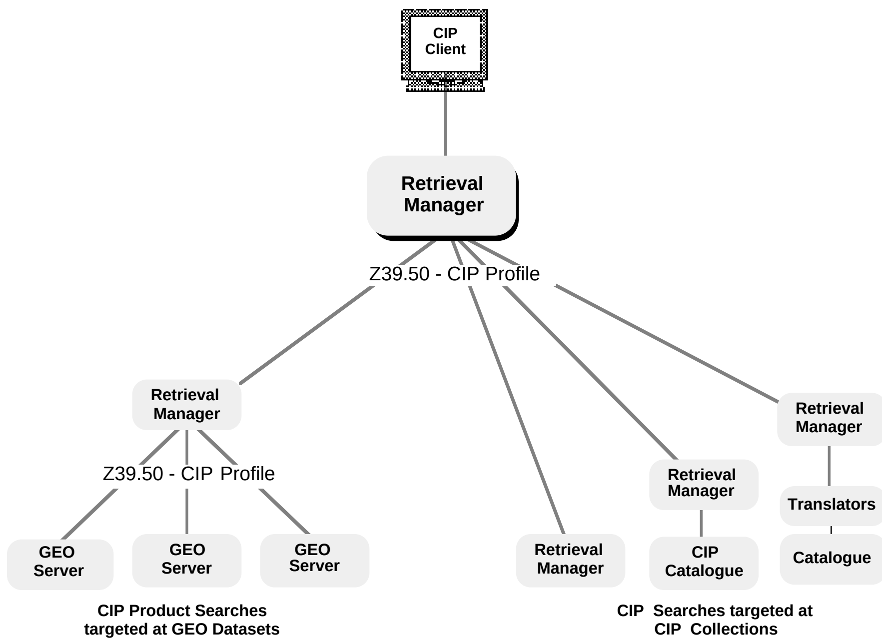


Figure 3-14. GEO Interoperability with CIP

3.3.1.4 Global Node Site

ICS Compatability: Explanatory

This section defines the concept of an ICS Global Node site. Because ICS provides the framework to link many different data providers, it becomes increasingly important to provide a convenient way for users to *discover* the kinds of data (i.e., products and documents) that is available to them within the system. An ICS Global Node provides the ability to the user to discover collections across the system.

An ICS Global Node maintains information about each root node that it can access. The root node is defined by the root collection descriptor maintained in the Collection Database managed by a given RM. A Global Node Administrator maintains a small text description of each accessible root node along with the URL of the root node CDB so that the Global Node knows how to find it.

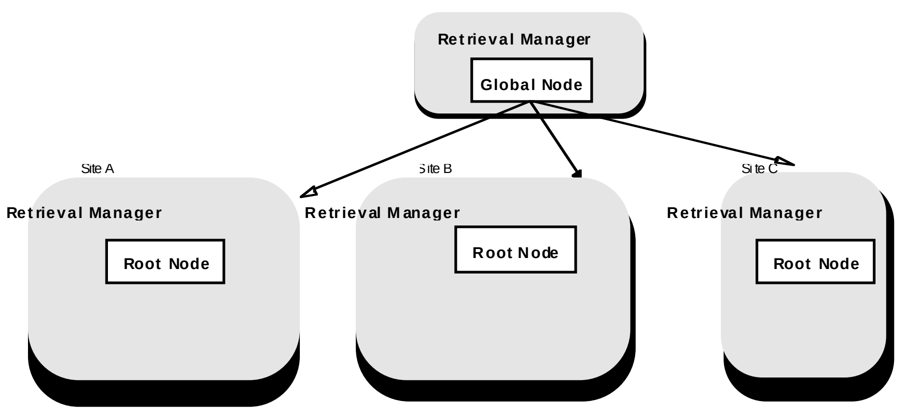


Figure 3-15. Global Node Site

Via an ICS Client, the user can perform a search against the Global Node to see a list of the available root nodes along with the text description. To view the collection descriptors for each root node, the Global Node issues a search request to each of the RMs to which it can connect. Should the user desire to see descriptors for collections below the root collection descriptors, a more refined collection search can be posed through the ICS Client. As shown in Figure 3-15, the Global Node again issues the search request to each of the RMs to which it can connect. In this case, each RM performs a “local” search on its collections. A collection descriptor is returned by the RM for each local collection that satisfies the given search criteria.

It is important to understand the local collection search constraint that must be implemented by the RM. For example, if the search criteria selects a theme collection which is comprised of local collections and one or more remote collections, the given RM returns only the collection descriptors for the local collections, assuming that they meet the search criteria. The reason for this is that the Global Node has issued the search request to all RMs. If the "remote" collections that comprise part of the theme collection satisfy the search criteria, then their collection descriptors will be returned by the RM where they are stored. This constraint prevents the problem of redundant searches and thus, redundant collection descriptors in the Global Node foundset.

From an architecture perspective, Global Nodes can be replicated within ICS. This permits distant sites to more easily perform discovery searches. Global Nodes themselves are not complete instances of RMs; they have a much pared down architecture. And, it is assumed that once a user decides to issue a product search based on the collection descriptor information retrieved by the Global Node, that the ICS Client then switches from a Global Node connection to a Retrieval Manager connection. In other words, the Global Node does not act as a "pass-thru" mechanism for product searches.

3.3.2 CIP Protocol Overview

ICS Compatability: Explanatory (See R3 for CIP Compliance)

CIP messages are exchanged between an *origin* and a *target* as part of a *CIP Session*. An *origin* initiates a session with a *target*. Several ICS elements act as *origins*: *ICS Client, Retrieval Manager, ICS Gateway, and HTTP/CIP Gateway*. Several ICS elements act as *targets*: *Retrieval Managers* and *Translators.*

The messages which CIP uses are listed in Table 3-1. The messages are used in the SDD to demonstrated the use of CIP in ICS. The CIP Specification [R3] provides a controlled definition of the contents and meaning of each of the messages in Table 3-1.

Table 3-1. CIP Messages

CIP Message Name	Short Description of CIP Message
InitializeRequest	Used by an *origin* to establish a *session* with a *target*.
InitializeResponse	Generated by the *target* after reception of the *InitializeRequest*
SearchRequest	Defines the elements transferred for any search operation including the query, the target database, and a result set name
SearchResponse	Contains the response to a *SearchRequest*, in particular it contains the total number of records in the result set
PresentRequest	Specifies the characteristics of the records in a result set which are to be returned to the *origin.* Included in the *PresentRequest* is the range of records to be returned.
PresentResponse	Return of the records requested during a *PresentRequest*.
SegmentRequest	When the records requested by a *PresentRequest* will not fit in a single segment , the *target* returns multiple *SegmentRequest*, each containing a portion of the requested records.
DeleteResultSetRequest	Allows a *target* to request an *origin* to delete specified result sets created during a *CIP session*
DeleteResultSetResponse	Generated by the *target* in response to the *DeleteResultSetResponse*
AccessControlRequest	Allows a *target* to challenge the identity of an *origin* .
AccessControlResponse	Sent by the *origin* to the *target* in response to a “challenge” to the identity of the *origin* through an AccessControlRequest
ResourceControlRequest	Allows the *target* to provide information concerning the status of a query to the *origin*.
ResourceControlResponse	Allows the *origin* to indicate whether to continue a query or not.
TriggerResourceControlRequest	Used by the *origin* to request a status report or change the status of an existing query. May be used to cancel a query started by the *origin* .
ResourceReportRequest	Allows an *origin* to request a *target* to produce a resource report.
ResourceReportResponse	Generated by the *target* in response to the *ResourceReportRequest* .
ExtendedServicesRequest	Allows an *origin* to perform tasks which are executed outside the scope of Z39.50. CIP supports the following extended services: Persistent Result Sets, Persistent Query, Periodic Query Schedule , Database Update, CIP Ordering The CIP Ordering Extended Service can contain one of the following actions: *orderQuoteAndValidate, orderSubmit, orderMonitor,* and *orderCancel.*
ExtendedServicesResponse	Allows a *target* to send back the response related to the execution of tasks executed outside the scope of Z39.50 and initiated by an *ExtendedServicesResponse.*
Close	Allows an *origin* or *target* to abruptly terminate all active operations and to initiate the termination of the *CIP Session*

Session Initialization and Closure. CIP is a session based protocol. A *CIP session* begins with an initialization *operation* and ends with a closure *operation*. The initialization allows negotiation procedures to control the use of all services to be used during the session. The closure *operation* signals the end of the session and is done unilaterally by either the *origin* or *target*.

Search. The Search service is the means by which the *origin* sends a query to be evaluated against the EO data and the status of the query is returned. A CIP search consists of a number of fundamental parts:

- Search control parameters
- The search query using a Z39.50 language and containing a valid combination of CIP attributes
- The target database for the search
- The result set to be created at the *target* to hold the results of the search
- Information about the structure of a distributed result set

Search Control. Search control parameters allow the user to control the type of search requested. The user can select either a *collection* search or a *product* search. A *collection* search is a CIP search that is used to identify and retrieve *collection* definitions. A *product* descriptor search is a CIP search that is used to identify and retrieve *product* descriptors. Also, the user can select any search to be either local to the *Retrieval Manager* which holds the targeted *collection*, or allow the search to be distributed to other *Retrieval Manager*s as identified in the searched *collection.*

Query Language. CIP defines queries using a general purpose query language defined using Reverse-Polish-Notation (RPN). Note, however, that this does not preclude *Retrieval Managers* from supporting other additional Z39.50 query languages.

Search Attributes. CIP splits attributes into different types depending upon their function: Use, Relation, Position, Structure, Truncation and Completeness attributes. In a search query, the Use attribute identifies the access point against which the search term is to be matched. A set of Use attributes based on the EO domain are defined and controlled in the CIP Specification. The other types of attributes are used to provide additional match criteria in a query.

In addition to the CIP attributes defined in the CIP specification, each data provider may define its own local attributes. These local attributes are used in exactly the same way as the CIP attributes. The only restriction is that their understanding is limited to the domain of the data provider who defines them (instead of the whole CIP domain as for the CIP attributes). For instance, if a search query contains a local attribute, this local attribute will be applicable (i.e. will be recognized) only by the *collections* owned by the appropriate *Retrieval Manager*.

Result Set Information. As a distributed CIP search may result in results at multiple targets, the *searchResponse* contains a description of the distributed result set, e.g., the number of records in each of the lowest level result sets. With this information, the *origin*, can request records from specific result sets using a subsequent *presentRequest.*

Result Set Retrieval. Using a *presentRequest*, CIP returns records that have been located through execution of the search service. Once a CIP *searchRequest* has been submitted to the *Retrieval Manager*, performed at the *target*(s), and has responded successfully, a result set is made available for subsequent *presentRequests* by the *origin*. Optionally, an *origin* may request piggybacking in a CIP search request, in which case a small number of retrieval records are returned in the search response itself. The retrieval record formats for piggybacked records are the same as those described for present responses. By selecting from pre-defined element sets, the user can select the desired fields to be returned from the result set.

Extended Services. *Extended Services* provide a mechanism to define and monitor tasks that are executed outside Z39.50. An *Extended Service* (ES) defines a particular task which is related to information retrieval but is not defined within Z39.50. It allows an *origin* to create, modify or delete *task packages*, which are maintained by the *target* in a database - the *Extended Services database*. The task defined in a *task package* depends on the particular *Extended Service* used. For CIP the following *Extended Services* are provided: Persistent Result Sets, Persistent Query, Periodic Query Schedule, Database Update, and CIP Ordering. Note that an *ES response* does not necessarily signal the completion of a task, which may have a lifetime which exceeds the *CIP Session* during which it is initiated.

Ordering. CIP defines an *Extended Service* for ordering of EO *products*. CIP supports submitting an order specification for validation and to obtain a quote. A specific message is provided for affirmative submission of an order. Additionally, a user can monitor the status of a order and request to cancel the order. Options on an order are specified in the order specification of CIP order message. Details on order options can be found in the Order Option Ammendment to the CIP Specification [R28].

Authentication: CIP provides the capability to transfer authentication information between *target* and *origin* pairs of a *CIP Session*. This transfer allows *Retrieval Managers* to authenticate *ICS Clients* and other *Retrieval Managers*. The authentication may be done at the beginning of a *CIP Session* or may be requested by a *Retrieval Manager* for a specific *operation*. Additionally, a *Retrieval Manager* can request non-repudiation be enforced for a specific request.

Explain. The Explain service enables the capabilities of a target server to be ascertained by clients. Explain queries are targeted at the Explain database of a *Retrieval Manager*. The Explain database contains the attributes which can be used at the *Retrieval Manager* as well as key information about the EO content of the *Retrieval Manager*. The Explain database also provides information about the format of records returned to the *origin* by the *Retrieval Manager*.

Identifiers. The CIP Specification also defines syntax for identifiers of several key items in the CIP space: messages and *collections* (including *task packages*). For example, each operation request from a client (or *Retrieval Manager*) must have a unique reference identifier. This can be used in the tracking of an operation and is particularly important when a search is invoked which is passed onto remote *Retrieval Manager*s.

Unsupported Operations. Not all ICS elements will support all CIP operations. Negotiation of supported operations occurs during *session* initialization. ICS elements as part of initialization will identify the operations supported by an element. If a message for an unsupported operations is sent during the session, e.g., a CIP ordering message is sent to a Catalogue Translator, the ICS element shall respond to the unsupported message with a diagnostic message.

3.3.3 CIP Operations

Messages are passed between ICS elements primarily using CIP. CIP is a session based protocol for conducting the operations, e.g., search and order. This section shows how various messages are passed between the distributed ICS elements to accomplish a request by a user for a query or an order. Each section discusses the simpler “local” case, and then discusses the case when multiple *Retrieval Managers* are involved. The operations shown here are examples of the two main interactions in ICS - search and order - and do not demonstrate the full extent of how ICS is used.

In each scenario it is assumed that a session is begun before the scenario is described. A scenario is established by the Client sending an *InitializeRequest* to the *Retrieval Manager* and the *Retrieval Manager* replying with an *InitializeResponse*. Likewise when a *Retrieval Manager* must send a secondary message to a another *Retrieval Manager* or to a *Translator*, a session is formed between the two. Nominally, the Client will send a *close* message to end the session. There is no response from the *Retrieval Manager* to the Client for a *close*. When the *Retrieval Manager* receives a *close* it will send a *close* to any other ICS elements to which a session had been established for this user.

3.3.3.1 Queries

When the user wants to find records pertaining to a particular interest, there will be multiple CIP messages behind the scenes which the ICS Client may or may not make visible. The CIP messages provide the flexibility to retrieve only the information that the user wishes to view. For instance a query consists of two separate operations: a search operation and a present operation. The search operation returns the number of hits. The present returns the records in the particular format requested by the user.

The following two sections show how the combination of search and present operations are used to search ICS data. First a local query is described followed by a distributed query. The particular examples shown in this section are typical but not the only combination of operations which a user may cause to happen in ICS. For example, a user query session may include multiple searches and only one present.

3.3.3.1.1 Local Query

ICS Compatability: Mandatory

Figure 3-16 shows the passing of messages between ICS elements to accomplish a local query. The query is searching the *collections* held by a local catalogue. The ICS elements involved are an *ICS Client Application*, a *Retrieval Manager*, and a *Catalogue Translator*. The client targets a *searchRequest* at a *collection* held by the *Retrieval Manager*. The contents of the *collection* in this case are held in the local catalogue and so the *Retrieval Manager* passes a *searchRequest* to the *Catalogue Translator*. The *searchResponse* which is returned from the Catalogue *Translator* and is passed to the client contains the number of "hits" produced by the search.

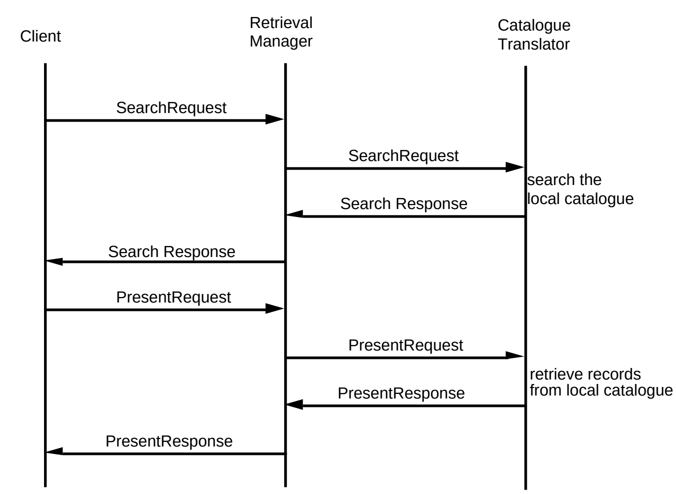


Figure 3-16. Local Query Operation Messages

Assuming that the user is satisfied by the number of hits indicated, the *ICS Client* is used to request the records in a specific format. This is accomplished by the client issuing a *presentRequest* against the result set which is held by the *Catalogue Translator*. The records are retrieved from the local catalogues and passed to the client via the *Retrieval Manager.*

The main function of the *Retrieval Manager* in this case has been to route messages to appropriate ICS elements based on the contents of the *collections*. Messages are routed based on what type of search was requested and where specific data is held. A user can request that a CIP search be either local to the *Retrieval Manager* or be propagated to other *Retrieval Managers* if the collection structure would warrant this for a specific query. The user may also choose between searching just the collections or searching the included products as well. If an ICS site holds the product descriptors in the local catalogue and not in the *Retrieval Manager,* the search will need to be forwarded to the catalogue if a Product Search was requested. How data should be distributed between the *Retrieval Manager* and the catalogue is dependent upon characteristics of an agency's data, e.g. number of products and volatility of *collections*. Each agency will determine the distribution of data while considering the aspects of maintaining replicated data versus the response time for processing a user request. These issues can be assessed at the time of configuring an ICS site. The CEOS PTT plans to provide guidance on these issues in the ICS Collections Manual [R5].

The scenarios for the *Navigate* and *Locate* services are done similar to the Search scenario above. When a user knows a URL for an item, e.g., a collection, the user can locate that item by sending a search to the *Retrieval Manager* indicated in the URL where a search is performed, e.g. against the collection database. Once a successful search has been completed, a present request can be made to retrieve the item descriptor. Similarly for the navigation service, it is assumed that user has an item descriptor at the ICS Client and wishes to change context to a related collection. Where the related collection is listed as part of the collection descriptor. To retrieve the related collection, a search for the collection's URL must first be performed successfully, followed by a present request and response.

Any CIP search operation that is performed to achieve a locate or navigate scenario is done as part of a CIP session. In order to minimize the overhead of session handling, the CIP Client checks if a session is already established with the Retrieval Manager to which the seachRequest will be sent. If a session already exists with the Retrieval Manager, the Client sends the searchRequest as part of that session otherwise a new session is established.

3.3.3.1.2 Distributed Query

ICS Compatability: Mandatory

The distributed query capabilities of the ICS are depicted by the messages in Figure 3-17. Similar to the previous section, the following operations are handled in a *CIP session* which is assumed to be established prior to the *operations* shown here.

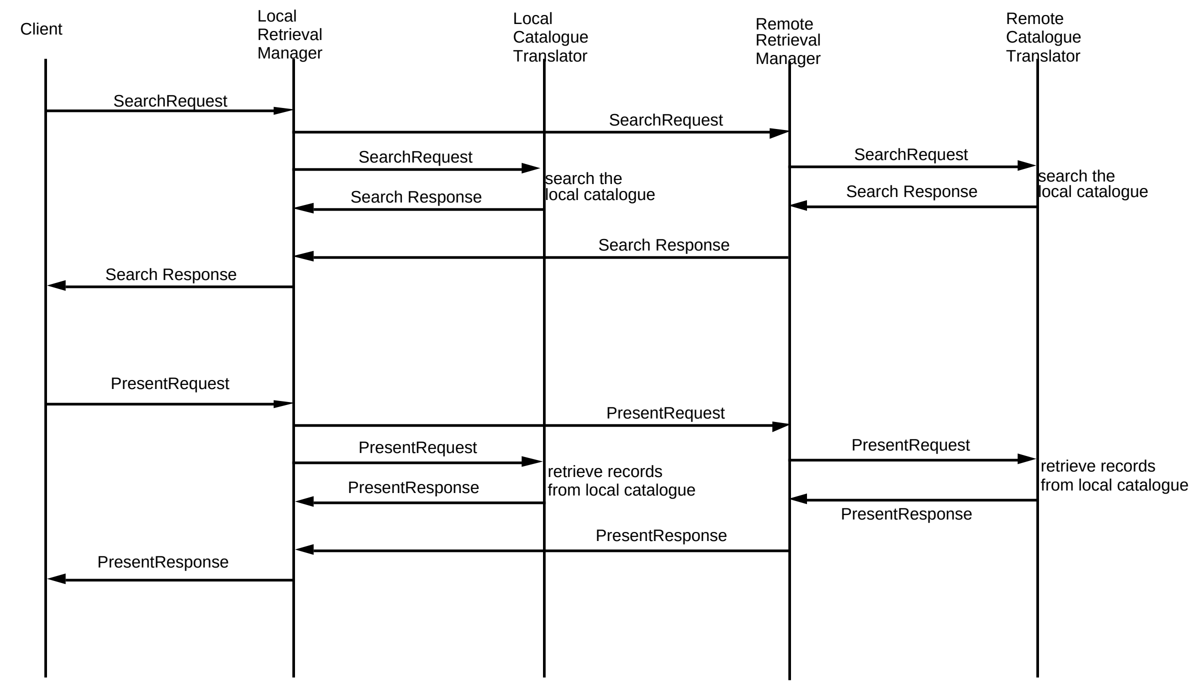


Figure 3-17. Distributed Query Operation Messages

When an *ICS Client* targets a *searchRequest* at a *collection* which includes both local *collections* and remote *collections*, the *Retrieval Manager* creates sub-searches which are targeted at the included *collections*. In this particular example, two separate *searchRequests* are created: one is sent to the local *Catalogue Translator*, the second is sent to a remote *Retrieval Manager*. The remote *Retrieval Manager* in turn determines, based on the *collection* contents, that a *searchRequest* needs to be sent to a *Catalogue Translator* which is local to the remote *Retrieval Manager*. *SearchResponses* are routed back to the initial *Retrieval Manager* which combines them, and a single *searchResponse* is sent to the client. The *searchResponse* contains information about the structure of distributed result sets, e.g., the number of records in each leaf level result set.

Similar to the splitting of a *searchRequest*, a *presentRequest* is split into sub-requests based on the contents of the result set. The *presentRequests* are routed to the distributed result sets and *presentResponses* are returned to the initial *Retrieval Manager*. A *presentRequest* can be directed to return all records from all of the distributed result sets. Because the sub-requests could each bring back large result sets, CIP provides the ability for an *ICS Client* to request records from specific distributed leaf result sets. An *ICS Client* composes such a detailed *presentRequest* using the description of the distributed result set found in the *searchResponse*.

An *ICS Client* can keep the user informed of the status of a search using the CIP Resource Control Facility. When in effect, Resource Control provides the structure for the *Retrieval Manager* and *ICS Client* to exchange messages about the status of an active operation, e.g., a distributed product search. Resource Control can be in effect for an entire *CIP session* or an *ICS Client* can trigger the Retrieval Manager to send a report for a specific operation only. If during initialization, Resource Control was negotiated to be in effect for the session, the *Retrieval Manager* will send *resourceControlRequests* to the *ICS Client* while an operation is active. The *resourceControlRequest* will provide a resource report including the status of the operation and may allow the client to stop the operation using a *ResourceControlResponse.* If resource control is not in effect for the session, the *ICS Client* can send a *TriggerResourceControlRequest* to which the *Retrieval Manager* will respond with a *ResourceControlRequest* including the resource report with status. For a distributed operation, the resource report contains information about sub-requests. This allows a user to determine, for example, that several of the distributed product sub-searches have been completed but the entire operation has not completed because of a single translator being down.

Users also want to know if a portion of a distributed seach has failed. When a sub-search returns with a failed status, the *Retrieval Manager* allows the other sub-searches to go to completion. If Resource Control is in effect for the session, the failure of a sub-search will be included in the *ResourceControlRequest* which is sent to the client. When all of the sub-searches are completed, a *searchResponse* is returned to the *ICS Client.*

After doing a distributed collection search, a user may target a product search at a specific remote collection. The user's client may need to access information about the collection in order to do an accurate query. The client does a *searchRequest* of the Explain database. The client makes this request against the local *Retrieval Manager*, which forwards the searchRequest on to the remote *Retrieval Manager* which holds the collection. The remote *Retrieval Manager* replies back to the local *Retrieval Manager* with the *searchResponse*. The local *Retrieval Manager* sends the *searchResponse* to the *Client*.

3.3.3.1.3 Global Node Queries

ICS Compatability: Mandatory

The global node query capability of the ICS is depicted by the messages in Figure 3-18. Like the previous sections, the following operations are handled in a *CIP session* which is assumed to be established prior to the *operations* shown here.

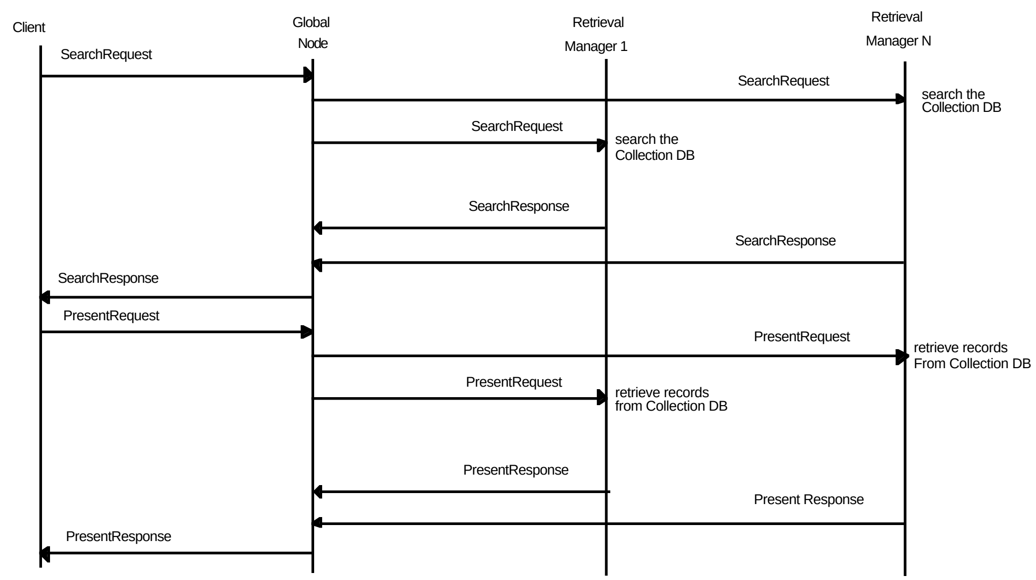


Figure 3-18. Global Node Query Operation Messages

When the *ICS Client* targets a *searchRequest* at a *Global Node*, the *Global Node* issues a *searchRequest* to each of the *Retrieval Managers* to which it can connect in order to search the Collection Database. The *searchRequest* parameter *additionalSearchInfo* indicates that the search is to be performed only locally. That is, if a theme collection is found that points to a remote collection, only the matching local component collections are returned as part of the *searchResponse.* Each *Retrieval Manager* then returns the *searchResponse* to the *Global Node.* The *Global Node* aggregates the *searchResponses* and returns the composite *searchResponse* to the *ICS Client.*

The handling of the *presentRequest* and subsequent *presentResponses* occurs in the same way as the *searchRequest* and *searchResponses* for a Global Node Query. The composite *presentResponse* contains the collection descriptors for all of the ICS collections that match the query criteria. This enables the user to "discover" collections across the entire ICS.

3.3.3.2 Ordering of EO Products

This section describes ordering of EO Products using CIP. The ordering process is shown here as fundamentally a two step process. First, a step of submitting an order specification resulting in validation and a quote. Second, the order is submitted for processing. There are multiple variations on this two step process. There may be multiple quotes requested before the user is happy with the order specification and then the order is submitted. On the other extreme the user may know exactly what they want and will directly submit the order without first asking for a quote. The above two step process assumes that the order items, i.e. the products which must be ordered, have already been identified. This is performed by querying collections and performing refinements on the queries until the desired order items and their related order options have been identified, by using the *Search* and *Present* facilities described in the previous section. A user can also monitor and cancel an order, but these operations are not shown in the following scenarios.

This section, like the last previous section, first shows a single *Retrieval Manager* case followed by the case in which two *Retrieval Manager*s are involved. There is a key difference between queries and orders. A query may be distributed to multiple *Retrieval Manager*s and *Catalogues*. Orders must contain *products* from a single *OHS*. When the order is composed by the client it will contain requests for items from only one site. The order may be routed from the local *Retrieval Manager* to a remote *Retrieval Manager,* but the order will only be presented to one *OHS*. If products need to be ordered from multiple OHS', several order requests must be submitted. For this reason we distinguish between "distributed" queries vs. "indirect" ordering.

Critical to ordering for some data providers will be the security mechanisms provided by CIP. Data providers will restrict some *products* to particular individuals and a monetary charge will be required for some *products*. CIP provides an authentication mechanism allowing a data provider to verify the identity of the person requesting a product. CIP also provides non-repudiation, which when invoked requires the user to submit an order which can not be disavowed at a later time. A detailed scenario using CIP authentication is provided in Section 6.

3.3.3.2.1 Direct Ordering

ICS Compatability: MAA

The direct ordering scenario by a user involves a *primary order*, from the *ICS Client* to the *Retrieval Manager*, and a *secondary order* to the local OHS *Translator* (See Figure 3-19). A *primary order* is defined to be between an *ICS Client* and a *Retrieval Manager*. A *secondary order* is created by an *Retrieval Manager* in response to a *primary order* and may be either to another *Retrieval Manager* or to an *OHS Translator*. Like in the previous scenarios, the messages shown occur as part of a *CIP Session* for which the initialization *operation* is not shown in the scenario.

After the user has specified the contents of the order, the client sends a CIP Ordering *extendedServicesRequest,* with the action of *orderQuoteAndValidate,* to the *Retrieval Manager*. To support the ordering *Extended Service*, the *Retrieval Manager* creates a *Task Package* which contains the order as well as additional information related to the *operation*. The *Task Package* will be updated based on the additional *Extended Services* messages as the scenario proceeds. The *Retrieval Manager* determines which OHS can fulfill the specified order. In this case, the *extendedServicesRequest* is routed by the *Retrieval Manager* to a local *OHS Translator*. The *OHS* provides the *Retrieval Manager* with a quote in the *extendedServicesResponse* which is returned to the client.

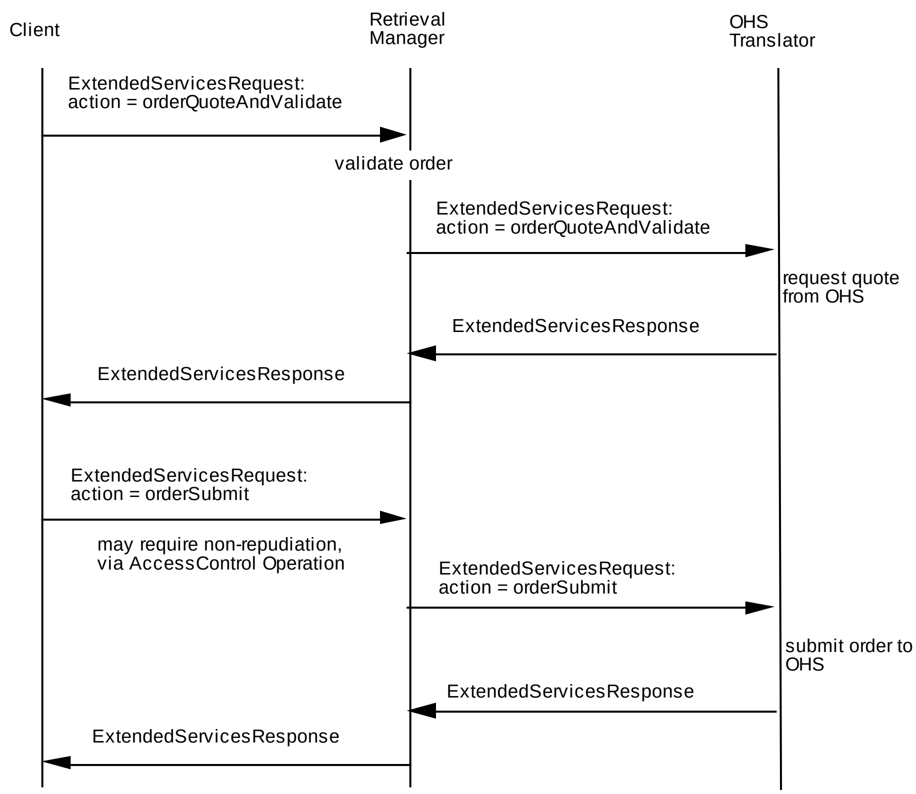


Figure 3-19. Direct Ordering

The scenario shown in Figure 3-19 continues with an order submittal. A user sends a CIP Ordering *extendedServicesRequest,* with the action of *orderSubmit,* when they are committing to the order. In some cases a *Retrieval Manager* will request that the order be submitted in a message which cannot later be denied, i.e. non-repudiation is required on the *orderSubmit*. Non - repudiation requires an *accessControlRequest* message from the *Retrieval Manager* to the client and an *accessControlResponse* message back from the client (not shown in Figure 3-19). The *Retrieval Manager* then submits the order using the *extendedServicesRequest* to the *OHS Translator*. An *extendedServicesResponse* is returned to the *Retrieval Manager* which in turn sends an *extendedServicesResponse* to the client.

The client may specify that the *extendedServicesRequest* operation be executed asynchronously with the *dontWait* option. This option is useful for cases where the client may not wish to wait for the *Task Package* to complete execution. In this case, the *extendedServicesResponse* indicates either acceptance of the requested *task* or rejection of the *task*. Execution of the *extendedServicesRequest* may take place outside the current CIP Session and the client then needs to query the *Retrieval Manager* with an Order Monitor request to retrieve status information.

Order status in ICS is a pull process, i.e., the current status of an order is pulled from the OHS whenever the user performs an Order Monitor request. Before this request, the order status contained in the Task Package in the *Retrieval Manager's* Extended Services Database may not necessarily be the most current status and hence Search and Present requests against this database are not supported. The *Retrieval Manager* updates the Task Package only when the client explicitly requests an update with an Order Monitor extended service request.

3.3.3.2.2 Indirect Ordering

ICS Compatability: MAA

An indirect ordering scenario by a user involves a *primary order*, a *secondary order* between a local *Retrieval Manager* and a remote *Retrieval Manager*, and another secondary order between the remote *Retrieval Manager* and the OHS *Translator*.

An indirect ordering scenario may have the local *Retrieval Manager* acting as either a proxy or a passthrough. In the proxy case, access rights of the Local *Retrieval Manager* on the Remote *Retrieval Manager* are used to order from the Remote *Retrieval Manager*. This allows a user to order *products* when the user has no privileges with the Remote *Retrieval Manager*. In the pass-through case, the Local *Retrieval Manager* routes the information to the remote *Retrieval Manager*, but the order is requested using the privileges of the user.

The messages passed for an Indirect Order (Figure 3-20) are very similar to the direct ordering case. The major difference is that when the client submits an *extendedServicesResponse* with action of either *orderQuoteAndValidate* or *orderSubmit* to the *Retrieval Manager*, the *Retrieval Manager* identifies that the order is to be filled by an *OHS* to which it does not have direct access. The *Retrieval Manager* then becomes an intermediary and passes the requests off to a Remote *Retrieval Manager*. Both the local and the remote *Retrieval Manager* will create a task package associated with the order request that it receives from the client or local *Retrieval Manager*, respectively.

Asynchronous operations take place as described in the Direct Ordering case.

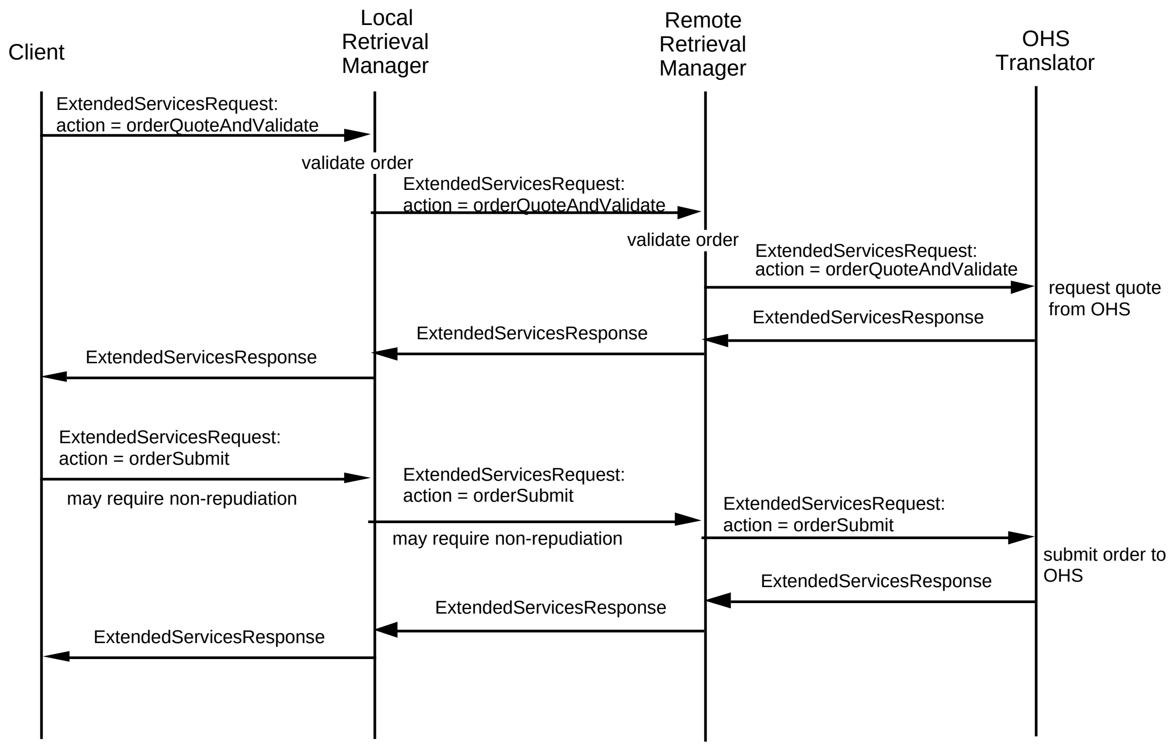


Figure 3-20. Indirect Ordering

3.4 ICS Guide Protocol (IGP) Domain Design

This section provides an overview of the design approach for incorporating the ICS Guide Protocol into ICS when a site needs to manage and access documents. The protocol integrates CIP and HTML document spaces, thereby allowing direct access to documents via distributed HTML gateways and minimizing the direct impact on document authors. It should be noted that the while the ICS framework supports the integration of this protocol, it only needs to be implemented when the site has documents, such as guide data, that need to be associated and viewed as part of a collection. The ICS Guide Design and Protocol Specification [R24] describes the terminology, goals, and detailed design for inserting documents into the system, managing the related metadata, searching for documents, and retrieving the documents.

Document services include both search and present services of the metadata and document content. A document service can be characterized by the following:

- document data can be searched through pre-defined metadata using Attribute/Keyword pairs;
- document data can also be searched by free text searches;
- documents can be retrieved through a known URL; and
- document data can be presented as metadata or as a textual document.

3.4.1 IGP System Design

The ICS Guide Protocol (IGP) incorporates use of a virtual document that can be used both inside and outside of the ICS document system. This approach has been selected because it permits both ICS internal indexing to provide for fielded and free-text searching within ICS and it allows general-purpose Web search engines to index and access ICS virtual documents. For a detailed discussion on the methodology for using general-purpose Web search engines to access ICS virtual documents, please refer to the ICS Guide Design and Protocol Specification [R24]. The following paragraphs describe Figure 3-21 which shows the ICS site architecture that permits insertion and indexing of documents as well as searching for and retrieval of the virtual documents by the ICS Client.

A "virtual document" is an HTML document generated by software for ICS. It is not stored anywhere, but is only served up at the time the software runs. The virtual document consists of the complete, original, document with necessary information added to it by the software, making it indexable and linking it to CIP collection information. Virtual documents are accessed via a CGI script URL which will point to the software that creates them on-the-fly. Indexing information will be embedded invisibly using HTML META tags.

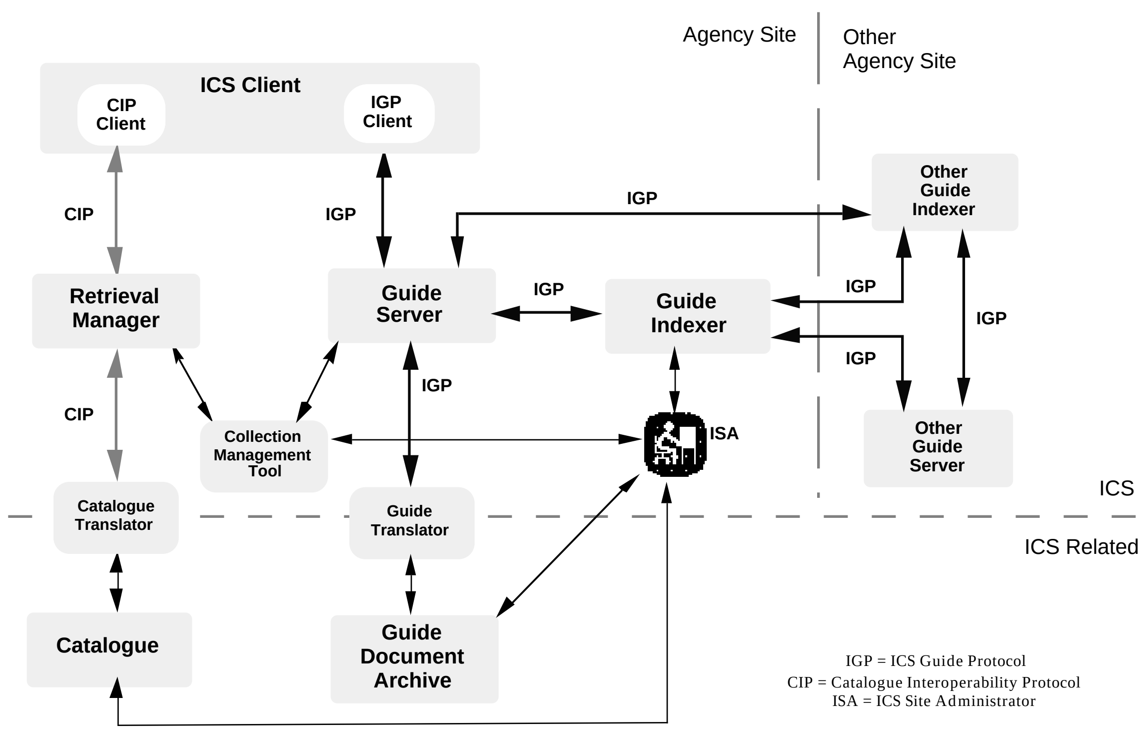


Figure 3-21. ICS Document Management Framework

3.4.1.1 Guide Insertion and Indexing

ICS Compatability: Explanatory

Two custom indexes, a general free-text index and a fielded index, are managed by the ICS Guide Indexer to permit ICS Clients to perform free-text searches and fielded document searches using CIP Attribute/Keyword pairs against ICS-only documents. These indices reference all documents known to each agency site that is part of ICS. Documents can be added to the system either in a "batch" mode, where a whole set of documents is used to establish a new set of indices, or in an "add" mode, where the existing indices are updated with one or more new documents. Batch mode is usually used when a new site is being added to ICS. The batch mode is executed to create the indices for all of the other (existing) agency sites. Add mode is then used for the documents that "belong" to the new site. Add mode is also used when new documents are added to an existing site.

When documents are indexed in batch mode, the ICS Guide Indexer (GI), a specialized Web-crawling index engine, creates the general free-text index and the fielded index for all physical documents known to each agency represented in the ICS. For example, suppose we had an ICS that had two site agencies, the "PDQ" agency and the "XYZ" agency. A new site, agency "ABC" is to be added. First, the batch mode is run to create the indices for agencies PDQ and XYZ at the ABC agency site. To perform this task, the GI starts the indexing by pointing at a top-level Web page that contains hyperlinks to all known agency site's *icsdoc* scripts. This Web page is mirrored at each ICS site. In order to run the *icsdoc* script, the GI requests that the Guide Server (GS) invoke the Guide Translator. The Guide Translator (GT) runs the *icsdoc* script to perform two separate tasks that are discussed in the following paragraphs.

When each *icsdoc* script is invoked, it returns a "directory" that is an HTML document consisting of hyperlinks to all documents reachable through the agency site, one link per line. The hyperlink to each document is NOT the URL of the physical document, but instead is a link back to the *icsdoc* script with path information to the document. Because the Guide Document Archive is empty for the ABC agency, there are no paths returned back from the GT through the GS to the GI. When the GI looks at the next *icsdoc* script URL on the Web page (icsdoc.html), it issues a batch mode request to GS at the XYZ agency site to create the portion of the directory that represents all of the paths to the physical files at that site. As discussed above, the GS at the XYZ agency site requests that its GT create the directory from the Guide Document Archive at the XYZ agency.

Then the ICS Guide Indexer follows these links. Via the Guide Server, the GI requests the Guide Translator to run the *icsdoc* script for each document, with the document path information sent into the script as a parameter; this information identifies a particular physical document that is located in the Guide Document Archive. In this case, the script returns a "virtual document", generated from the physical document referred to in the path information. For example, contacting the URL http://xyz.org/icsdoc/docs/erbe.html causes *icsdoc* to read and modify the erbe.html document and then return it as a virtual document. This virtual document contains all the HTML of the physical document, plus invisible CIP Attribute/Keyword pairs and a visible link to the appropriate data collection(s) that is created using information from a Guide to Collections mapping file that is accessible to the GT.

Next, the GI scans the virtual document for words and adds the document link to its database, with the words as keys. The virtual document is not stored at the index, just the link. All of the documents in the xyz and pdq "directory" are indexed in the same way to complete indexing for batch mode.

In addition, the GI manages updating of document metadata. When the GT creates the virtual document, it creates as much of the document metadata as it can from the document and performs the mapping between the local schema attributes and CIP attributes. This metadata is returned back to the GI, who then requests the DS to "fill in" any missing attribute information. The GS maintains a global attributes default file that provides information about mandatory CIP attributes and their default values. For cases where the metadata that is created from the virtual document is missing mandatory attributes, the GS supplies the missing attribute names and their default values. As a final "third" pass, the GI requests the GS to display a screen via the ICS Client containing the metadata derived from the document along with the default metadata that has been supplied by the GS. This gives the Site Administrator the opportunity to change any metadata that is deemed appropriate. Lastly, the GI requests that the GS store the document metadata.

After the free-text and fielded indices are created for the XYZ and PDQ agencies, the documents belonging to the new agency site ABC can be added. First, the ABC site manager adds the new documents to the Guide Document Archive shown in Figure 3-21. Or, the document can be made visible to the Guide Document Archive if it is not located at the site. Then, the ABC site manager uses a browser to tell the ICS Guide Indexer the URL for the new document. For example, the URL may be http://abc.org/icsdoc/newdoc.html. The GI then runs its Web-crawling index engine on the URL. Through the Guide Server, the GI requests that the Guide Translator retrieve the document from the Guide Document Archive and run the *icsdoc* script on it, because "icsdoc" is part of the URL. Just as in the batch mode example, the *icsdoc* script creates a virtual document and updates the fielded and free-text with searchable keywords from the document and the virtual document URL. In addition, the document metadata is updated as discussed previously.

Next, as shown in Figure 3-21, the ICS Guide Indexer notifies Guide Indexers at the other ICS agency sites (XYZ and PDQ in our example) that are represented in the icsdoc.html file (i.e., http://abc.org/icsdoc.html) to run a "add" mode request for agency site ABC with the URL that was specified to the browser. When the other Guide Indexers receive the add request, they proceed in exactly the manner specified for the ABC agency site, but do not notify other Guide Indexers. In this manner, all agency sites maintain up-to-date indices.

3.4.1.2 Guide Searching

ICS Compatability: Explanatory

Free-text and fielded searches are performed via requests issued from the ICS HTTP Client to the ICS Guide Server (GS). Search requests issued to the GS return a list of URL links to the ICS virtual documents that match the query criteria specified by the ICS Client.

For example, suppose a user issues a free-text search request through the ICS Client that requests all documents containing the term “OZONE GSFC”. The URL for the ICS free-text search request would look like the following for the ABC agency site:

http://abc.org/icssearch/free_text="OZONE GSFC"

The GS issues a search request against the free-text index to find the documents that have the term “OZONE GSFC” in them. The URLs for any documents that have that term in them are then returned to the GS. In turn, the GS returns the list of the URLs to the ICS Client for display to the user.

When the user desires to perform a fielded search, ICS Client requests a list of the valid Use attributes from the GS. The GS returns the names of the valid CIP attributes from the global attributes defaults file that can be used for searching. The ICS Client generates a form with empty text boxes for each CIP attribute. Suppose the user enters the following values in these fields:

Author: "John Doe"
PublishingDate: "June 1997"

The search URL would look like:

http://abc.org/icssearch?Author=JohnDoe&
PublishingDate=June1997

When the user submits the form from the ICS Client, the Guide Server will check the fielded index for document URLs that have the specified CIP Attributes. Finally, note that the system can provide both an ICS free-text and fielded search simultaneously. A text field on the fielded search form could be used to enter free text.

As in the free-text search, when the user selects one of the document links, they retrieve the virtual document served by the *icsdoc* script at the appropriate agency site. Then, *icssearch* would check for matching document URLs in the fielded index and the free text index.

3.4.1.3 Guide Document Retrieval

ICS Compatability: Explanatory

When the document search request returns a list of matching document URLs to the ICS Client, the user can select to retrieve a given document and its associated metadata. Upon selection, the ICS Client issues a document retrieval request to the Guide Server. The GS retrieves the metadata for the document from the Guide Metadata Database and requests the Guide Translator to create the virtual document using the *icsdoc* script and the URL selected by the user. The GT creates the document and it is made available to the user via the HTTP Client that is part of the ICS Client.

3.4.2 IGP Protocol Overview

ICS Compatability: Explanatory (See R24 for IGP Compliance)

IGP messages are exchanged between the ICS Client, Guide Server, Guide Indexer, and Guide Translator in order to manage document storage and retrieval within ICS. The messages which the IGP uses are listed Table 3-2. The messages are used in the SDD to demonstrated the use of IGP in ICS. The ICS Guide Design and Protocol Specification [R24] provides a controlled definition of the contents and meaning of each of the messages in Table 3-2.

Table 3-2. IGP Messages (1 of 2)

Message Name	How It Is Used
FieldSearchForm	ICS Client sends this message to ICS Guide Server, which runs *icssearch* script. This message causes script to read Attributes Default File, generate a fielded search form, and returns form to client to display.
GetCollection	Link from document to collection for an ICS Client. CIP Client returns collection search results to CIP Client.
GetCollectionGuest	Link from document to collection for a general Web browser. Links to CIP Client. Starts a session, shows collection search results.
GetDirectory	Indexing engines send this message to ICS Guide Server, which has Guide Translator run the *icsdoc* script without path information. Returns a list of URLs of all virtual documents reachable through the site.
GetVirtualDocument	Indexing engines, browsers, and ICS Clients send this message to ICS Guide Server, which has Guide Translator run the *icsdoc* script with the path information to a document. Returns a virtual document. ICS Clients append the "&ICS_CLIENT" to notify the Guide Translator which type of collection link to insert.
IndexAdd	ICS Guide Indexer sends this message to the ICS Guide Indexer at other sites, which instructs them to update their indexes with the new guide document.

Table 3-2. IGP Messages (2 of 2)

Message Name	How It Is Used
IndexNew	An ICS Site Administrator sends this message manually (or browser client software sends it) to ICS Guide Indexer at his site, notifying it of a new guide to be indexed at the site.
IndexValues	ICS Site Administrator browser sends this message to ICS Guide Indexer to update the Guide Metadata Database with the information provided. This is done during guide ingest.
SearchFields	ICS Client sends this message to ICS Guide Server, which runs *icssearch* script. This queries the fielded index with the supplied values for specified attributes. If more than one attribute/value pair is specified, an "&" appears before each additional pair. Returns a list of URLs of virtual documents that have at least one pair in its entry in the Guide Metadata Database. (Note this implies logical OR between pairs). A search of both fielded and free-text indexes can be done by using "free_text" as one of the attributes and the text as the value.
SearchFreeText	ICS Client sends this message to ICS Guide Server, which runs *icssearch* script. This queries the free-text index with the supplied text. Returns a list of URLs of virtual documents that contain the text.
UpdateCollections	Collection Management Tool sends this message to ICS Guide Server when the ISA adds a guide_url to a collection in the Collections Data Base. The server has the Guide Translator run *ics_collection_update* script, which updates Collections Mapping File.
UpdateMetadata	ICS Guide Index sends this message to ICS Guide Server when it has Guide Attribute/Value pairs for a new document added to the system. The server has the Guide Translator run *ics_metadata_update* script, which updates the Guide Metadata Database.

3.4.3 IGP Operations

Messages are passed between ICS Guide elements primarily using IGP. IGP is based on the HTTP and is used to pass primarily indexing, search and guide retrieval requests. This section shows how various messages are passed between the distributed ICS elements to accomplish a request by a user for a search or a guide retrieval.

3.4.3.1 ICS Fielded Guide Search

ICS Compatability: MAA

Figure 3-22 shows the two sets of messages that are used to accomplish an ICS fielded guide search. First, the *ICS Client* issues the *FieldSearchFormRequest* to the *Guide Server* to create a fielded search form from the Attributes Default File. The form is passed back to the *ICS Client* via the *FieldSearchFormResponse* so that it can be displayed to the user. The user then specifies values for the desired search fields and the *ICS Client* issues a *SearchFieldsRequest* to the *Guide Server*. The *Guide Server* searches the fielded index to determine if there are any attribute/value matches. If so, the *Guide Server* returns a list of URLs of virtual documents that contain the matching attribute/value pairs. Although not depicted in Figure 3-22, it should be noted that both a fielded search and free-text search can be performed via the *SearchFieldsRequest*. In order to perform this, the user specifies "free_text" as one of the attributes and the search text as the value. The *Guide Server* then searches the free-text index in addition to the fielded index.

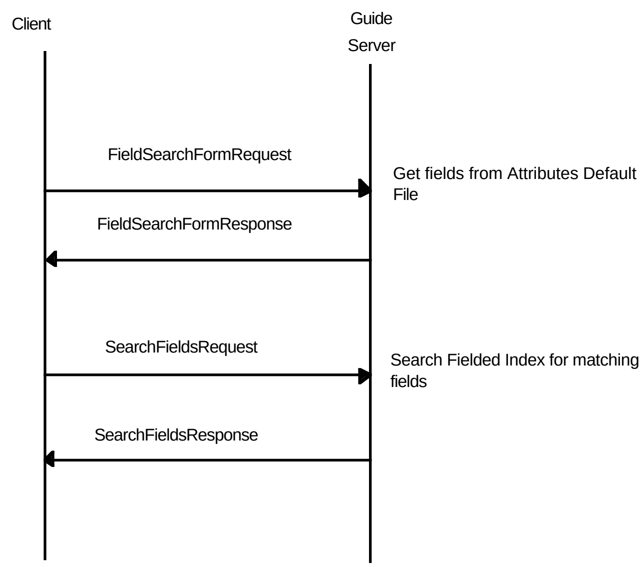


Figure 3-22. Fielded Guide Search

3.4.3.2 ICS Client Free-Text Guide Search

ICS Compatability: MAA

Free-text searches for guide documents are performed in ICS using the messages depicted in Figure 3-23. The *ICS Client* presents a simple form with a single text box to the user. The user enters text that must be in the list of guides that will be returned. When the user submits the form, the *ICS Client* issues a *SearchFreeTextRequest* to the *Guide Server*. The *Guide Server* then searches the free-text index with the supplied text to determine if there is a match. The list of URLs of virtual documents that contain the text is then returned to the *ICS Client* via the *SearchFreeTextReponse* message.

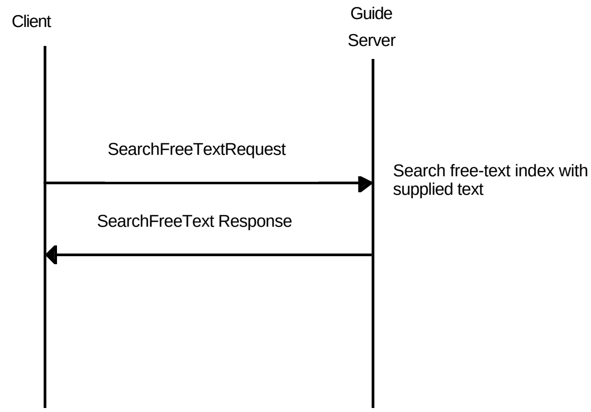


Figure 3-23. ICS Client Free-Text Guide Search

3.4.3.3 ICS Guide Retrieval from ICS Client

ICS Compatability: MAA

Retrieval of virtual guide documents is performed in ICS using the messages shown in Figure 3-24. Based on a list of URLs presented from either the fielded or free-text searches, the user can specify the URL of a virtual guide document for retrieval. The ICS Client passes the URL via the GetVirtualDocumentRequest to the GuideServer. The Guide Server routes the message to the Guide Translator to create the virtual document. The Guide Translator retrieves the Guide Document from the local store and retrieves the metadata from the Guide Metadata Database so that it can construct Meta tags and links for the related collections in the virtual document. It should be noted that each site contains free-text and fielded indices for all sites within ICS. If the user at site A requests retrieval of a document from site B, then it is the *Guide Server* and *Guide Translator* from site B that actually performs the retrieval. Once the virtual document is created by the *Guide Translator*, it is returned to the *ICS Client* via the *Guide Server* in the *GetVirtualDocumentResponse.*

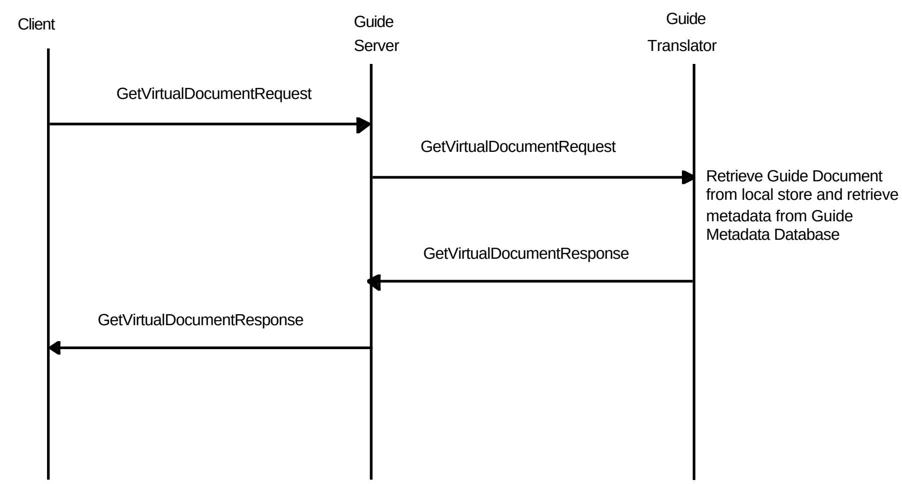


Figure 3-24. ICS Guide Retrieval from ICS Client

3.5 Identification of ICS Element Services and Interfaces

This section provides an identification of the services provided by each element shown in the ICS Framework. The remainder of this section provides the identification of services for the following elements:

- *Retrieval Manager*
- *ICS Client*
- *Catalogue Translator*
- *OHS Translator*
- *UPS Translator*
- *HTTP/CIP Gateway*
- *ICS Gateway*
- *ICS Site Administrator (ISA)*
- *Collection Management Tool (CMT)*
- *Monitoring and Control Tool (MCT)*
- *Guide Server*
- *Guide Indexer*
- *Guide Translator*

Section 3-6 describes services assumed to be provided by the non-ICS elements.

The services supplied by a given site, using ICS elements, will vary by site. The services in this section are defined assuming a maximum site. The minimum services needed to be provided by a site to be ICS Compatible are described in Section 9.

3.5.1 Retrieval Manager Services

ICS Compatability: MAA

The following services are provided by the *Retrieval Manager*:

- CIP Compatible Messaging based on Z39.50 messaging and session management: *origin* and *target*
- Session Management: Manages user requests including; queries and associated sub-queries, search routing, recursive search trapping, result set collation and presenting, order management.
- Collection Management: persistent storage of local *collection*; services to create, ingest and maintain local *collection*; attribute management; promotion of hot *collections* to theme *collections*.
- Explain Database: Persistent Store of Explain data
- Native Extended Services: Creation, modification and persistent storage of *Task Package* in response to *Extended Service* requests and responses.
- CIP Ordering Extended Services: A CIP customized Extended Service for ordering EO Products. Specific CIP Order Extended Service request types are Order Validation and Quotation, Order Submission, Order Monitoring and Order Cancellation.
- User Management: persistent storage of user information where user may be a human using a client or may be another *Retrieval Manager*.
- Operator Interface for *ICS Site Administrator* administration of the *Retrieval Manager.*
- Monitoring and Error Management: User session errors, CIP diagnostic messages, network diagnostics messages
- E-mail Client: E-mail to *ISA* for error messages.

The *Retrieval Manager* design is required to be a modular design allowing a site to disable specific services which are not be hosted at the site. The variation of *Retrieval Manager* services across sites is described in Section 3.3.1.2.

The *Retrieval Manager* has the functional interfaces indicated in Table 3-3.

Table 3-3. Retrieval Manager Interfaces (1 of 2)

Other ICS or Related Element	**Interface Description**
ICS Client	A *Retrieval Manager* is capable of forming *CIP sessions* with multiple *ICS Clients* where the *ICS Client* is the *origin* and the *Retrieval Manager* is the *target*. All CIP messages are supported by this interface.
Other z39.50 Clients	*Retrieval Manager* is capable of forming a z39.50 Z-association with other z39.50 clients, where the *Other Z39.50 Client* is the *origin* and *Retrieval Manager* is the *target*. All Z39.50, Version 2 and 3 sessions are supported by this interface with the exception of the following z39.50-1995 services: *scanRequest*, *scanResponse*, *sortRequest, sortResponse, extendedServicesRequest, extendedServicesResponse.*
Other *Retrieval Manager*s	A *Retrieval Manager* is capable of forming CIP sessions with other *Retrieval Manager*s where either *Retrieval Manager* may be the *origin* or the *target*. All CIP messages are supported by this interface.
ICS Gateway	A *Retrieval Manager* is capable of forming *CIP sessions* with multiple *ICS Gateways* where the *ICS Gateway* is the *origin* and the *Retrieval Manager* is the *target*. All CIP messages are supported by this interface.
Catalogue Translator	A *Retrieval Manager* is capable of forming CIP sessions with multiple Catalogue *Translator*s where the *Retrieval Manager* is the *origin* and the Catalogue *Translator* is the *target*. All CIP messages are supported by this interface with the exception that CIP Ordering Extended Service is not supported by this interface. (Note that other interface approaches are under investigation, e.g., direct function call, for those sites with tight coupling between the *Retrieval Manager* and *Translator*.)
OHS Translator	A *Retrieval Manager* is capable of forming CIP sessions with multiple OHS *Translator*s where the *Retrieval Manager* is the *origin* and the OHS *Translator* is the *target*. The following CIP messages are supported by this interface: *InitializeRequest, InitializeResponse, AccessControlRequest, AccessControlResponse, ExtendedServicesRequest, ExtendedServicesResponse, Close.* The only valid extended services type for this interface is CIP ordering

Table 3-3. Retrieval Manager Interfaces (2 of 2)

Other ICS or Related Element	Interface Description
UPS Translator	A *Retrieval Manager* may have an interface to a UPS *Translator* to request User Management Data including User Authentication Information. This interface may be accomplished using CIP. A Retrieval Manager may have an interface to a Certification Authority to perform public key directory lookups. This interface should use the X.509 Lightweight Directory Access Protocol (LDAP) [R23].
ICS Site Administrator (ISA)	The *Retrieval Manager* provides an operator interface to the *ISA*. The operator interface provides a graphical user interface to allow the *ISA* to monitor, control, diagnose, and maintain the operations of the *Retrieval Manager*.
Collection Management Tool	The *Retrieval Manager* allows the *CMT* to modify the EO content of the *Retrieval Manager Collection Data Base*
Monitoring and Control Tool	The *Retrieval Manager* allows an *MCT* interface to gather operation and error handling data from the *Retrieval Manager* and to responds to *MCT* commands for controlling the configuration and state of *Retrieval Manager* processes. This interface should be based on a standard protocol, e.g., SNMP, to be defined in later releases.

3.5.2 ICS Client Services

ICS Compatability: MAA

The following services are provided by an *ICS Client*:

- CIP Compatible Messaging. The *ICS Client* is able to send and receive all CIP messages.
- Search Formulation. The *ICS Client* supports the user in formulating queries, e.g., queries of *collection*s and *products*, incremental queries, conversion of geographic inputs into CIP queries.
- Result Set handling. The *ICS Client* supports the user in viewing temporary result sets, and conversion of a temporary result set to a persistent result set.
- Order Formulation. The *ICS Client* supports the user in specifying and reviewing quotes on orders, including conversion of a multi-site result set into single site orders.
- Authentication Support. The *ICS Client* supports the user in providing the user's credentials to form an authenticated session or to respond to an authentication request from the *Retrieval Manager*.
- Dynamic Configuration. The *ICS Client* supports the retrieval and caching of explain information from *Retrieval Managers* including definitions for attributes and record structures. The *ICS Client* provides the attributes in support of a user creating searches.
- User Profile Management. The *ICS Client* may support the persistent storage of user profile data.
- Document Search Formulation. The *ICS Client* supports the user in formulating free-text and fielded searches for documents within ICS.

Note that the list above intentionally does not include presentation services to the user.

The *ICS Client* has the functional interfaces indicated in Table 3-4.

Table 3-4. ICS Client Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	*ICS Client* is capable of forming CIP sessions with at least one *Retrieval Manager* where the ICS Client is the *origin* and the *Retrieval Manager* is the *target*. All CIP messages are supported by this interface.
ICS Guide Server	*ICS Client* is capable of forming an HTTP connection using IGP messages with the *ICS Guide Server* where the ICS Client is the client and the *ICS Guide Server* is the server. All IGP messages are supported by this interface.
Interfaces outside of ICS	The *ICS Client* may have interfaces to the following ICS related elements: the *Archive.* The *ICS Client* may access the *Archive* using ftp to transfer files from the data store for user access through the ICS Client.

3.5.3 Catalogue Translator Services

ICS Compatability: MAA

The following services are provided by the *Catalogue Translator*. Other CIP Translator configurations are described in Section 3.2.2.3. For example, if a single CIP Translator is used to interface to a data provider (see the “Retrieval Manager as a Catalogue Gateway” example in Figure 3-12) the services of a *Catalogue Translator*, *OHS Translator* and *UPS Translator* are merged.

- mapping of CIP attributes to the attributes of the local Catalogue (i.e., for product and guide)
- mapping of CIP query elements to local query elements
- mapping of local result sets into CIP result sets
- use of the local catalogue communication protocol
- query optimization

The *Translator* contains elements which are common across ICS sites, e.g., those parts of the *Translator* which speak CIP, and contains parts which are unique to the site, e.g., those parts which speak the local protocol.

Multiple *Translator*s may be provided to allow multiple *Retrieval Managers* to interface with the *Catalogue.*

The *Catalogue Translator* has the functional interfaces indicated in Table 3-5.

Table 3-5. Catalogue Translator Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	A Catalogue *Translator* is capable of being a CIP *target* for *CIP sessions* from a single *origin*. Typically the *origin* will be a *Retrieval Manager*. All CIP messages are supported by this interface except the CIP Ordering Extended Service. A catalogue translator will typically not support *collection* seaches or Explain searches. (Note that other interface approaches are under investigation, e.g., direct function call, for those sites with tight coupling between the *Retrieval Manager* and *translator*.)
Interfaces outside of ICS	The *Catalogue Translator* interfaces to the following ICS related elements: *Catalogue*

3.5.4 OHS Translator Services

ICS Compatability: MAA

The following services are provided by the *OHS Translator*.

- Convert CIP order objects to local order handling system objects
- use of the local *catalogue* communication protocol

The *Translator* contains elements which are common across ICS sites, e.g., those parts of the *Translator* which speak CIP, and contains parts which are unique to the site, e.g., those parts which speak the local protocol.

Multiple *Translator*s may be provided to allow multiple *Retrieval Managers* to interface with the *OHS*.

The *OHS Translator* has the functional interfaces indicated in Table 3-6.

Table 3-6. OHS Translator Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	An *OHS Translator* is capable of being a CIP *target* for one CIP session. Typically the *origin* will be a Retrieval Manager. The following CIP messages are supported by this interface: *InitializeRequest, InitializeResponse, AccessControlRequest, AccessControlResponse, ExtendedServicesRequest, ExtendedServicesResponse, Close.* The only valid extended services type for this interface is CIP ordering
Interfaces outside of ICS	The *OHS Translator* interfaces to the following ICS related elements: *Order Handling System*

3.5.5 UPS Translator Services

ICS Compatability: MAA

The following services are provided by the *UPS Translator*.

- Local user profile to CIP user profile
- Secure handling of authentication information
- Use of the local *catalogue* communication protocol for user information

The *Translator* contains elements which are common across ICS sites, e.g., those parts which interface to the *Retrieval Manager*, and contains parts which are unique to the site, e.g., those parts which speak to the local *UPS*.

Multiple *Translator*s may be provided to allow multiple *Retrieval Managers* to interface with a *UPS*.

The *UPS Translator* has the functional interfaces indicated in Table 3-7.

Table 3-7. UPS Translator Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	An *UPS Translator* is capable of being a CIP *target* for one *CIP session*. Typically the *origin* will be a Retrieval Manager. This interface may be accomplished using CIP.
Interfaces outside of ICS	The *UPS Translator* interfaces to the following ICS related element: *User Profile System*

3.5.6 HTTP/CIP Gateway Services

ICS Compatability: MAA

The following services are provided by the *HTTP/CIP Gateway*:

- CIP Messaging. The *HTTP/CIP Gateway* converts the user's inputs received via HTTP to CIP messages for transmission to a *Retrieval Manager*.
- HTML Generation. The *HTTP/CIP Gateway* dynamically generates HTML output from a set of templates for presentation of CIP information to the user. This may be based on CGI scripts (programs) capable of receiving input from - and sending HTML output to - any standard WWW browser.
- Session management. The *HTTP/CIP Gateway* maintains CIP session context in response to http messages which is a state-less protocol, e.g., preservation of information from screen to screen for a particular user session (e.g., result set IDs, user preferences). The HTTP protocol does not maintain session information, it is a one time passing of information. Virtually every single mouse click to a browser results in the gateway generating a single page from scratch.
- Security. The *HTTP/CIP Gateway* supports user authentication in the ICS.

- Explain Data Management. The gateway provides an Explain cache as a local cache for the Explain information provided by the *Retrieval Manager*
- *ICS Client* server. Some *HTTP/CIP Gateway*s may act as servers of portable client code. One example is a Java applet which would allow a Java enabled web-browser to run some or most of the *ICS Client* layers at the user's computer.

The *HTTP/CIP Gateway* has the functional interfaces indicated in Table 3-8.

Table 3-8. HTTP/CIP Gateway Application Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	*HTTP/CIP Gateway* is capable of forming CIP sessions with at least one *Retrieval Manager* where the *HTTP/CIP Gateway* is the *origin* and the *Retrieval Manager* is the *target*. All CIP messages are supported by this interface.
Interfaces outside of ICS	The *HTTP/CIP Gateway* may have interfaces to the following ICS related elements: *a web server*

3.5.7 ICS Gateway

ICS Compatability: MAA

The following services are provided by the *ICS Gateway*.

- mapping of local attributes to CIP attributes (e.g. for *product*, *guide* and *collection*)
- mapping of local query elements to CIP query elements
- mapping of CIP result sets into local result sets
- use of the local *catalogue* communication protocol
- The *ICS Gateway* converts the user's inputs received via local protocol to CIP messages for the *Retrieval Manager*.

The *ICS Gateway* contains elements which are common across ICS sites, e.g., those parts of the *ICS Gateway* which speak CIP, and contains parts which are unique to the site, e.g., those parts which speak the local protocol.

The *ICS Gateway* has the functional interfaces indicated in Table 3-9.

Table 3-9. ICS Gateway Application Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	*ICS Gateway* is capable of forming *CIP sessions* with at least one *Retrieval Manager* where the *ICS Gateway* is the *origin* and the *Retrieval Manager* is the *target*. All CIP messages are supported by this interface.
Interfaces outside of ICS	The *ICS Gateway* may have interfaces to the following ICS related elements: *catalogue* and *Existing agency client.*

3.5.8 ICS Site Administrator (ISA) Operations

ICS Compatability: MAA

The following operations are performed by the *ISA*:

- Collection Definition. The *ISA* creates the site *collections* in accordance with the ICS Collection Manual, including site *provider archive collections*, *provider theme collections*, links to *collections* held at other sites.
- Collection Maintenance. The *ISA* maintains the site *collections* in accordance with the ICS Collection Manual, including periodically checking the *collections* held in their *Retrieval Manager*s for consistency.
- *Retrieval Manager* data. The ISA maintains all data in the *Retrieval Manager* databases.
- *Retrieval Manager* Operations. The *ISA* monitors and corrects any incorrect operations of the *Retrieval Manager* at their site.
- Guide Document Definition. The *ISA* adds Guide Documents to ICS and updates related metadata.

The *ISA* has the operational interfaces indicated in Table 3-10.

Table 3-10. ISA Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	The *ISA* accesses the *Retrieval Manager* via an operator's interface. The operator interface provides a graphical user interface to allow the *ISA* to monitor, control, diagnose, and maintain the operations of the *Retrieval Manager*.
Collection Management Tool	The *ISA* uses a *CMT* operator interface to control ingest of data into the *CMT*, modification of the data, and insertion of CIP *Collections* into the *Collection Database*.
Catalogue Translator	The *ISA* maintains the data and monitors the operations of the *Catalogue Translator* at the site.
OHS Translator	The *ISA* maintains the data and monitors the operations of the *OHS Translator* at the site.
UPS Translator	The *ISA* maintains the data and monitors the operations of the *UPS Translator* at the site.
Guide Indexer	The *ISA* accesses the *Guide Indexer* via an operator's interface. The operator interface provides a graphical user interface to allow the *ISA* to add guide documents, update related metadata, or request batch update of the document indices.
Interfaces outside of ICS	The *ISA* may have interfaces to the following ICS related elements: *OHS, UPS, Catalogue, Archive, Guide Document Archive*. These interfaces are not discussed here as they are outside of ICS.

3.5.9 Collection Management Tool (CMT) Services

ICS Compatability: MAA

The following services are provided by the *CMT*:

- Conversion of files into CIP compatible format *Collections*
- Support maintenance of the *Collection Data Base*
- Provides mechanism to define linkage between *Collections* and *Guide Documents*

The *CMT* has the functional interfaces indicated in Table 3-11.

Table 3-11. Collection Management Tool Interfaces

Other ICS or Related Element	Interface Description
*Retrieval Manager*s	The *CMT* interfaces to the *Collection Database and Explain Database* of the *Retrieval Manager* to establish and maintain the EO content of the *collections* held in the data bases.
ISA	The *CMT* provides an operator interface to the *ISA*. The operator interface allows the *ISA* to control ingest of data into the *CMT*, modification of the data, and insertion of CIP *Collections* into the *Retrieval Manager*.
Guide Server	The *CMT* notifies the Guide Server with an *UpdateCollections* message so that the Collections Mapping File can be updated.
Interfaces outside of ICS	The *CMT* will have an interface to accept data which is to be made into CIP *collections*, e.g. EOSDIS V0 DAAC data set descriptions, DIFs. This interfaces is not discussed here as it is outside of ICS.

3.5.10 Monitoring and Control Tools (MCT) Services

ICS Compatability: MAA

The following services are provided by the *MCT*:

- Monitoring of *Retrieval Manager* operations and error handling.
- Commanding the *Retrieval Manager* to change states based on SSM commands

The *MCT* has the functional interfaces indicated in Table 3-12.

Table 3-12. Monitoring and Control Tool Interfaces

Other ICS or Related Element	Interface Description
Retrieval Manager	The *MCT* interfaces to the Retrieval Manager to gather operation and error handling data from the Retrieval Manager and to control the configuration and state of Retrieval Manager processes. This interface should be based on a standard protocol, e.g., SNMP, to be defined in later releases.
Translators	The *MCT* may interface to the translators to gather status and command changes similar to the interface with the *Retrieval Manager*. This particular interface is TBS.
Interfaces outside of ICS	The *MCT* may have interfaces to the following ICS related element: SSM. These interfaces are not discussed here as they are outside of ICS. This interface should be based on a standard protocol, e.g., SNMP, to be defined in later releases.

3.5.11 Guide Server Services

ICS Compatability: MAA

The following services are provided by the *Guide Server*:

- Manages free-text and fielded searches for guide documents
- Includes an HTTP Daemon
- Manages retrieval service for virtual guide documents
- Holds the Guide Metadata attribute defaults
- Holds the Guide Metadata Database

The *ICS Guide Server* has the functional interfaces indicated in Table 3-13.

Table 3-13. ICS Guide Server Interfaces

Other ICS or Related Element	Interface Description
ICS Client	The *Guide Server* is capable of forming IGP connection with the *ICS Client* where the *Guide Server* is the *origin* and the *Guide Server* is the target. This interface is an IGP interface which is based on HTTP
Guide Translator	The *Guide Server* interacts with the *Guide Translator* to pass messages to request Directory searching and retrieval of virtual documents. This interface is an IGP interface which is based on HTTP
Guide Indexer	The *Guide Server* interfaces with the local *Guide Indexer* and *Other Guide Indexers* to update Guide Metadata databases. This interface is an IGP interface which is based on HTTP
CMT	The Guide Server receives an *UpdateCollections* message from the *CMT* so that the Collections Mapping File can be updated.
Interfaces outside of ICS	The *ICS Guide Server* may have interfaces to the following ICS related element: SSM. These interfaces are not discussed here as they are outside of ICS. This interface should be based on a standard protocol, e.g., SNMP, to be defined in later releases.

3.5.12 Guide Indexer Services

ICS Compatability: MAA

The following services are provided by the *ICS Guide Indexer*:

- Creates and updates the ICS Free-Text index of ICS Guide Documents
- Creates and updates the Fielded index of ICS Guide Documents
- Lists all sites of ICS Guide Servers
- Contains an HTTP Daemon

The *Guide Indexer* has the functional interfaces indicated in Table 3-14.

Table 3-14. Guide Indexer Interfaces

Other ICS or Related Element	Interface Description
Guide Server	The *Guide Indexer* interfaces with the *Guide Server* to update the Guide Metadata database.
Other *ICS Guide Indexers*	The *Guide Indexer* interfaces with other *Guide Indexers in ICS* to let them know that a new URL has been added and that their indices need to be updated.
Other *Guide Servers*	The *Guide Indexer* interfaces with other *Guide Servers* to update the Guide Metadata databases.
Interfaces outside of ICS	The *Guide Indexer* may have interfaces to the following ICS related element: SSM. These interfaces are not discussed here as they are outside of ICS. This interface should be based on a standard protocol, e.g., SNMP, to be defined in later releases.

3.5.13 ICS Guide Translator Services

ICS Compatability: MAA

The following services are provided by the *ICS Guide Translator*:

- Gets directory of all URLs to all guide documents that are reachable through the agency site
- Retrieves virtual documents
- Updates collection and Guide Document metadata

The *Guide Translator* has the functional interfaces indicated in Table 3-15.

Table 3-15. ICS Guide Translator Interfaces

Other ICS or Related Element	Interface Description
ICS Guide Server	The *ICS Guide Translator* returns directories for guide documents and virtual documents to the *Guide Server.*
Interfaces outside of ICS	The *Guide Translator* has interfaces to the following ICS related element: *Guide Document Archive.* These interfaces are not discussed here as they are outside of ICS.

3.6 Identification of ICS Related Element Services

This section describes the services assumed to be provided by Elements related to ICS but not part of ICS. For these elements which are shown on the ICS Framework (Figure 3-9) but are not part of the ICS per se, the following identification of services are assumptions which will be used as the ICS development proceeds.

- *Catalogue*
- *Order Handling System (OHS)*
- *User Profile System (UPS)*
- *Archive*
- *Site System Management (SSM)*
- *Guide Document Archive*

3.6.1 Existing Catalogue Services

ICS Compatability: Explanatory

The following services are assumed to be provided by the *Catalogue*:

- Directory service provides descriptions of data sets containing high level information suitable for making an initial determination of the potential usefulness of a data set for some applications. This information will be used to populate ICS *collections.*
- Inventory service provides the information about the data held in an archive. This information will be used as ICS *product* descriptor data.
- Browse service provides access to *browse* data.

Note that a particular existing catalogue may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *Catalogue*.

3.6.2 Order Handling System (OHS) Services

ICS Compatability: Explanatory

The following services are assumed to be provided by an existing *OHS*:

- Persistent Storage of Orders originating from a *Retrieval Manager*
- Formulation of quotes on orders
- Statusing of orders
- Billing and accounting of users orders (for orders filled by the *OHS* and by orders for which the local *Retrieval Manager* acted as a proxy.)
- Notification to users of order filled
- Invoices to users

Note that a particular *OHS* may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *OHS*.

3.6.3 User Profile System (UPS) Services

ICS Compatability: Explanatory

The following services are assumed to be provided by an existing *UPS*:

- Persistent Storage of user profile data

Note that a particular *UPS* may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *UPS*.

For a *UPS* that is based on a public-key infrastructure (PKI), the following services are assumed to be provided by a Certification Authority: (CA)

- Publish the criteria for granting user certificates
- Granting, revoking and general certificate management functions
- storing root keys

3.6.4 Archive Services

ICS Compatability: Explanatory

The following services are assumed to be provided by an *Archive*:

- Persistent Storage of EO data
- Filling of orders: delivery via ftp or media

Note that a particular *Archive* may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *Archive.*

3.6.5 Site System Management (SSM)

ICS Compatability: Explanatory

The following services are assumed to be provided by an existing *SSM*:

- Management oversight of functional elements at the data provider site
- Monitoring of site services
- Commanding of site service configuration and modes

Note that a particular *Site System Management* may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *Site System Management*.

3.6.6 Guide Document Archive

ICS Compatability: Explanatory

The following services are assumed to be provided by an existing *Guide Document Archive:*

- Retrieval of documents from within the collection

Note that a particular *Guide Document Archive* may not provide all of these services. But, it is the assumption of the ICS SDD that if these services are provided by a data provider that they are provided by the *Guide Document Archive*.

This page intentionally left blank.

4. DATA VIEW

The Data View characterizes the data required to support the ICS functions as described in the Functional View Section of this document. The main data object in ICS is the *Collection*. Therefore, the main focus of this section is a description of the conceptual and logical nature of the *Collection.* To begin, Section 4.1 provides a motivation from an EO user's point of view for *Collections*. This section is then followed by a description of the *Collections* Model and a brief description of using *collections* with earth observation data. Section 4.4 addresses ction Concept Details while Section 4.5 describes how the users will interact with the model. This would then conclude the conceptual discussion of the *collection.* Section 4.6 primarily addresses the logical nature of the *Collection* by providing an overall description of the data components of ICS. This description includes an illustration of the ICS data components coupled with a brief textual narration. Section 4.4 briefly describes the document data.

4.1 Motivation for Collections

ICS Compatability: Explanatory

Users of Earth Observation data will be interested in subsets of data from various datasets and various locations. A user is typically interested in an assembly of data which best fits the topics which the user needs to investigate. This assembly requires the data to be organized in a certain way. It is this notion - that the users need for data does not uniformly overlap the organization - that provides the motivation for the collection concept.

The following are some examples of how EO Data consumers will want to access EO data:

- It is assumed that in general the user community will often desire access to specific groupings of ICS holdings. For example the user may request information from a single instrument, or for a specific wind-speed and direction data from AMSR, or of trace gas distribution data from MIPAS;
- Because agencies have responsibility for holding the data from certain sources, the physical location of data will be assumed to be geographically distributed around the world. Collections can be formed which can be associated with products from multiple geographic locations;
- It is assumed that the community will also require access which spans instruments. For example, a request for all sea surface temperature data, across MODIS, AMSR, and so on;
- It is assumed that inter-disciplinary work will call for access across science disciplines and instruments, relative to some specific time, space, or event coincident coverage.

Therefore, to address these somewhat divergent but related needs the concept of a *collection* was developed. This concept embraces the following fundamental concepts.

- The *collection* is an item
- The *collection* has members
- The members can be other items

These concepts allow the data to be organized in a manner which will facilitate the access needs of the user.

The following sections explore these concepts in further detail by introducing a model of the collections concept, describing the collection membership, and providing examples of user interactions with the collections.

4.2 Collections Model Overview

ICS Compatability: Explanatory

The collections concept permits a wide degree of flexibility in organizing EO data. A *collection* is described in ICS as a collection descriptor, i.e., a set of elements which describe the collection. The *collection* may have members which are reflected as included item descriptor identifiers. These identifiers serve to identify the associated *products,collections*, and *task packages,* through the *Included Product Descriptors,* , *Included Collection Descriptors,* or *Included Task Packages* respectively.

A *collection* supports the notion of associated data or membership either due to their physical location (i.e. database storage), or along the lines of subject themes, that are deemed useful organizational principles. The science data provider coupled with a *Retrieval Manager Administrator* should decide the appropriate organizational strategy. Certain rules will be specified in the ICS Collection Manual to provide the necessary guidance to assist this team in making the appropriate decision.

Some examples for arranging data follows. Traditionally EO data was arranged according to EO product types (Raw, Level 0, etc.). Within ICS *Collections*, data can be arranged according to agency, discipline, spectral range, instrument, processing level, geographical area, etc. The user can then direct a search in a more focused manner, by including or excluding particular collections. As a *collection* can contain either included *product descriptor* ID's, *included collection descriptor* ID's or included *persistent result set* or included *persistent query* , it is possible to associate a number of collections under a single theme as the data provider or user finds applicable. It is also possible to associate an existing set of *product descriptor* ID's, which are already associated to an existing *collection*, to a new *collection* by just referring to the existing *collection descriptor* ID within the definition of the new *collection descriptor*. Because Release B of CIP does not support mixed collections each *collection* may only include descriptors of a specific type; i.e. *Product*, *Collection , persistent query* or *persistent result set..*

Figure 4-1 illustrates this concept followed by Figure 4-2 which further illustrates the notion of a *collection* and its' relationship (association) with other collections.

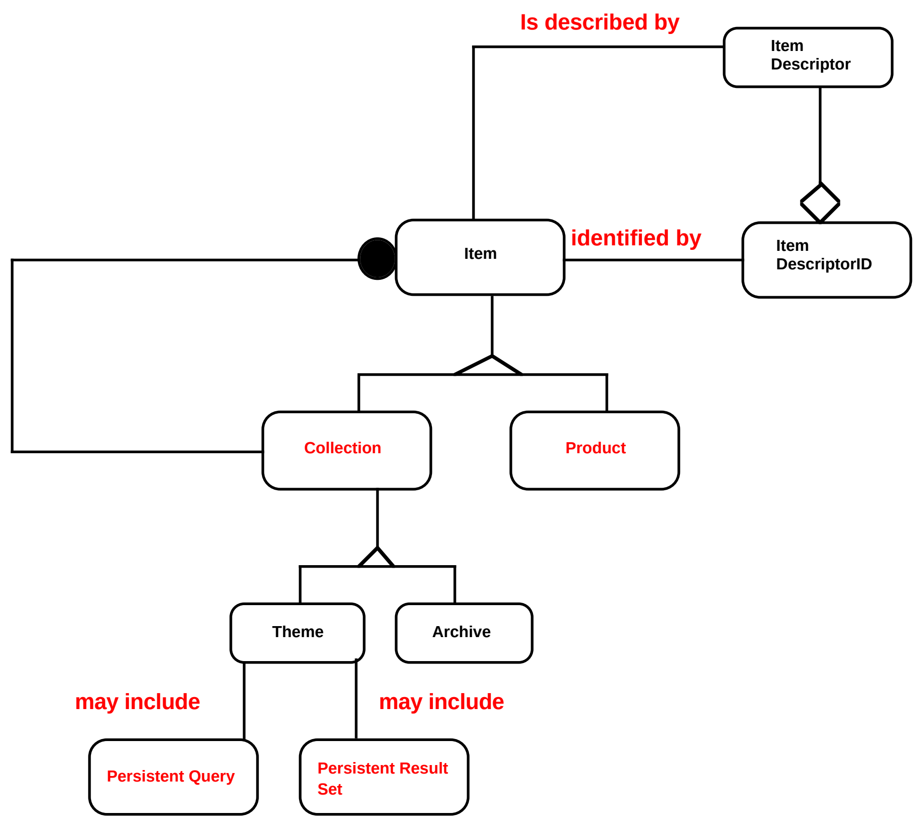


Figure 4-1. Single Instance of Collection in Figure 4-2.

The above model is an attempt to graphically represent the concept of a *collection* which is illustrated in Figure 4-2. First an item is equivalent to a collection in Figure 4-2 (i.e. collection 2.2). Each item is described by an item descriptor which logically consists of a group of elements which describe the overall nature of the item, such as, name, temporal and spatial extents. The items can be either a collection or product; the collection item can be either a theme or archive. Each item is identified by a unique item descriptor. Collections may contain zero or more items. These items can be other collections or products. The remaining sections will explore each of the above elements of the collection.

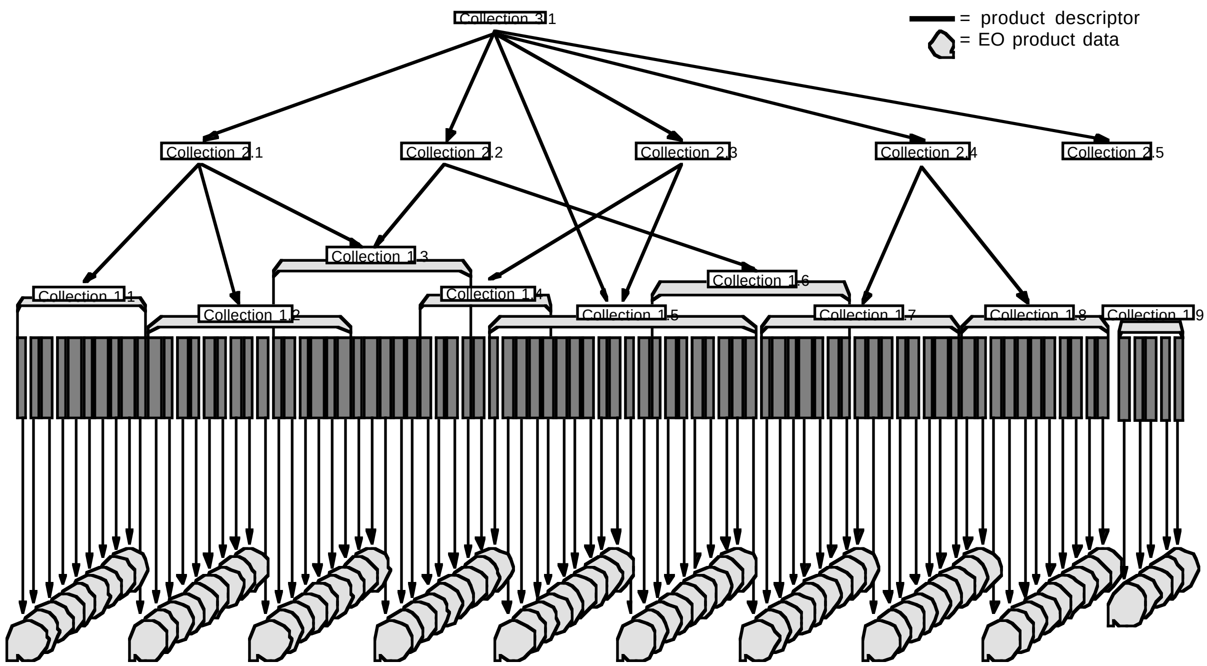


Figure 4-2. The "Collection Structure" Concept

Figure 4-2 expresses the logical relationship between multiple *collections*. As can be seen a web of interesting relationships can be formulated when using the collection concept. The collections in Figure 4-2 do not represent the naming of collections in an actual implementation. As can be seen the collections can overlap each other and can be associated with more than one collection. The central concept is that there are terminal collections, i.e., Collection 1.2 and non-terminal, i.e., Collection 2.2. The terminal collections in this illustration describe the information contained in the inventories and are defined as Archive Collections in ICS. The Archive Collection Included Item Descriptor Id's will reflect only the Included Product Descriptor ID's. Above the terminal level collections, there are non-terminal collections. These collections are typically constructed around a theme, or event, however not limited to these aggregations, which expresses the primary tenet of the associated collections. In ICS these are the Theme Collections. The included collections may be any existing collection.

There is no restriction on the number of included item descriptors within collection levels, therefore a non-terminal collection could group together terminal collections and other non-terminal collections. Also, a terminal collection could exist without a relationship to a higher collection (i.e. collection 1.9), or a non-terminal collection could exist with no relationship to lower collections, in other words a collection without members (i.e. collection 2.5). These conditions could occur in particular circumstances, for instance when a catalogue site has only one single collection, i.e. Collection 1.9, or when the collection is under construction, or is a Theme Collection, i.e., Collection 2.5. The specific details for constructing and maintaining the collections are described in the ICS Collections Manual.

An integral part of the collection structure is the Root Node. The overall purpose of the root node is to allow a *Search Request*, which was formulated at the origin by a user, to ultimately determine the EO *collection* that match the criteria specified in the query attributes of the *Search Request.* The root node serves as the logical mechanism which supports this service. The criteria specified in the query are evaluated against the contents of the Root Node and all of the Root Nodes' included collections to determine if the *collections* in the targeted *collection* tree satisfies the *Search Request*. The data objects model for the root node contains the *collection* descriptor object and attributes.

Specific rules and the required procedures for developing and maintaining a root node are identified and discussed in the ICS Collections Manual.

4.3 Example Of Using Collections With EO Data

ICS Compatability: Explanatory

Several examples of constructing collections based on the users needs are reflected in Figure 4-3. In this illustration a Collection has been established to include MODIS data from several missions. Also a Collection has been created based on the theme Sea Surface Temperature (SST), which associates existing SST collections from AMSR, AATSR, and MODIS. The intent of the subset caption is to show that only some of the item descriptors are included in the MODIS collection or AMSR on ADEOS II, or AATSR on ENVISAT, or SAR on Radarsat. Lastly, a Collection is defined based on an Andes Event. As illustrated, this collection includes collections which otherwise may not be related.

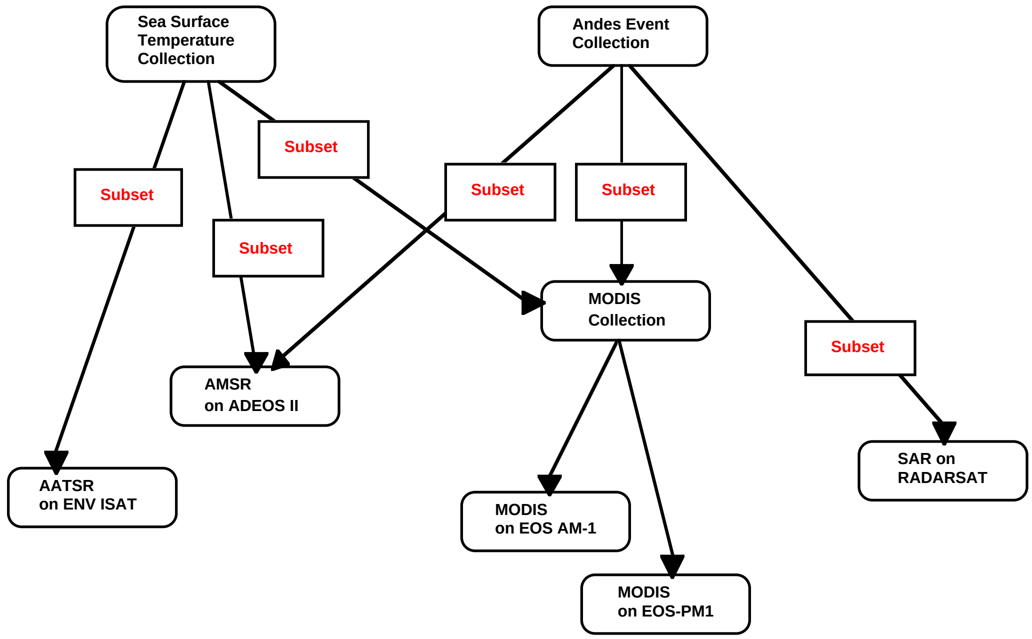


Figure 4-3. Creating Collections for Users

4.4 Collection Concept Details

In order to fully understand the power of the Collection Concept within ICS the following detailed information is provided in this section. The intent is to describe the collection descriptor classifications (categories), to provide some insight into collection membership, to discuss and define characteristics which are unique to this concept, and lastly, to describe how a user can interact with collections.

4.4.1 Collection Descriptor Classification

ICS Compatability: Explanatory

The flexibility of the collections concept requires an accurate classification scheme to allow the users of ICS to determine the characteristics of certain collections they are creating or viewing. In support of this requirement two major classification elements are used to classify *collections* in the ICS. The following describes each of these elements.

Type

- *Archive* - An *archive collection* reflects a single underlying physical archive that is "owned" by a single *Retrieval Manager*
- *Theme* - A *theme collection* reflects a common concept or theme and may include Provider Archive or Other Theme Collection members that are "owned" by several different *Retrieval Managers.*

Examples of collections which can be used to group data together which have a similar semantic theme can be seen in Figure 4-4 . The classification of those *collections* include:

1. *Collections* of SENSOR data by mission (e.g. MODIS from AM-1, SAR from RADARSAT, etc.) which reflects the organization of the underlying physical archives. *archive* would be the correct element classification.

2. A *collection* has been established to include MODIS data from several missions. A collection has been created based on the theme of Sea Surface Temperature (SST) including SST products from AMSR, AATSR and MODIS. This collection would be classified as a *Theme* collection.

The first type of *collection, Archive*, is likely to be created by data providers to organize their archives and facilitate access to the *product* descriptors (i.e. analogous to an inventory containing inventory entries). The second type of *collection, Theme*, may be established by data providers or Users who want to organize some of their data into groupings which differ from their *archive collections* (i.e. from the baseline inventory), for the convenience of their users. The following provides a more detailed explanation of the provider archive and theme collections.

Archive collections. An *archive collection* could be a dynamic collection where new item descriptors are automatically added to the collection by the catalogue system as new archive entries are created, i.e. for data from an existing, operational EO satellite. Alternatively, an *archive collection* could be a static fixed collection of historical data that is no longer being added to as the source satellite is no longer operational or the campaign is no longer active. An *archive collection* includes only item descriptors that are local.

Theme collections. This *collection* could be based on a covered geographical area, a scientific discipline supported by the data, or an instrument type, etc. The difference between *theme* and *archive collections* is that *archive collections* only contain homogeneous information, while *Theme Collections* may have item descriptors whose contents may vary depending on the nature of the collection. *Theme Collections* are non terminal. Theme Collections may also be created to include unregistered collections that have been created and preserved in the Extended Services Database. All of these options are dependent on how the Data Administrator decides to organize the data.

Note that these classification definitions are not mandatory for the CIP to operate, however to ensure data interoperability the *collections* described above must share identical semantical interpretation across ICS. They must also be under the control of the ISA which is one step towards ensuring interoperability. The *Retrieval Manager* makes use of *collections* for the routing of distributed queries. Although a *collection descriptor* data element is identified and defined to capture a collection classification, the primary intent of the categorization is to aid the ICS user in understanding the quality guidelines imposed on the *collection* data and metadata. Standardization of *collection* definitions is provided as part of the ICS design. Guidelines for the creation and maintenance of these collections in addition to further explanation of the characteristics of these collection classifications can be found in the ICS Collections Manual.

The two classifications of collections are illustrated in Figure 4-4.

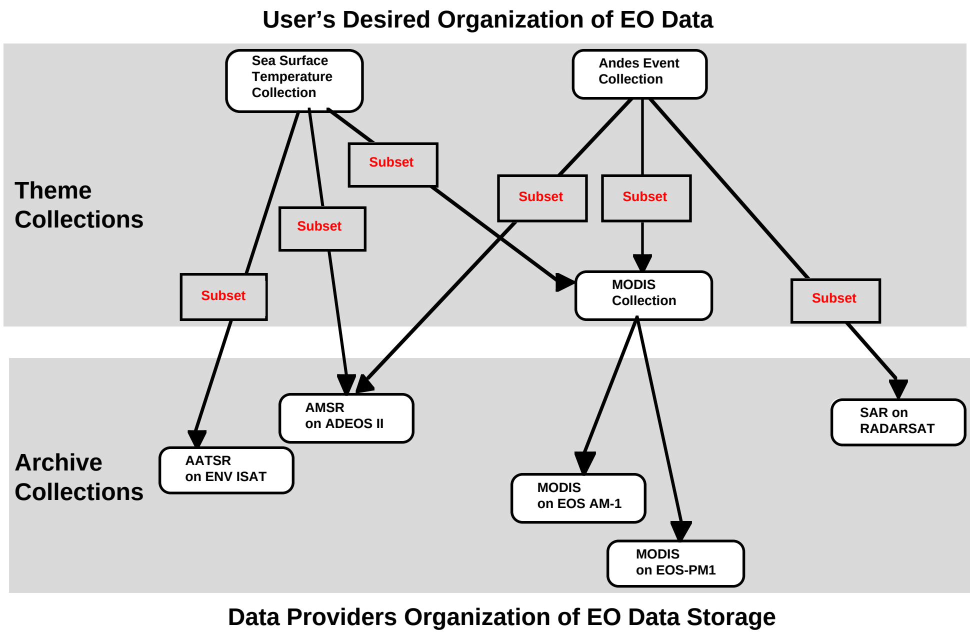


Figure 4-4. Theme and Archive Collections

4.4.2 Collection Members

ICS Compatability: Explanatory

Collection Membership in ICS is defined as including remote or local item descriptor ID's in the *collection descriptor* which defines the *collection*. A detailed process for identifying and including members will be described in the ICS Collections Manual. However, there are several fundamental rules which serve to guide this process.

1. A *product* descriptor identifier may be included in many Theme Collections however as a minimum it must be contained in one and only one *archive collection*.
2. *Archive collections* must include one or more *product* descriptors
3. *Theme collections* may include *collection* descriptors which may be an *archive* or *theme collection* descriptor.

4.4.3 Collection Characteristics

ICS Compatability: Explanatory

This section explores in more detail, the higher level concepts that were introduced in the previous sections. This is achieved by identifying and defining detailed *collection* characteristics to include the following:

- Commonality
- Evolution over time
- Terminal Collections
- Identifier
- Uniqueness
- Remote collections
- Related collections
- Related Guide Documents
- Local Attributes
- *Registered* Collections
- *Unregistered* Collections
- Included Task Packages

Commonality:

By definition, a *collection* is a grouping of items that have something in common. A *collection* may have members that have many or fully common attributes (*archive collection*), or a *collection* may have members that have a common semantic theme, though only a small subset of common attributes (*theme collection*).

Further CIP specifies a list of standard attributes that can be searched. Some of these standard attributes are mandatory for all *collection* members (different mandatory sets for different descriptor types), while some are optional (although commonly understood). Finally, some attributes can be locally defined.

Evolution over time:

Static members do not change over time. This can be a result of static underlying *collections* or the mechanism used to create the *theme collection* such as a Volcanic Eruption Theme Collection whose members will more than likely remain static over time.

Dynamic members may change over time based on changes in the included *collections*. It is envisioned that the majority of *collections* will not be static, but will evolve as ICS is used. This will occur in response to the way in which the user community wants to view relationships between the various data held by ICS, and the ways the CEOS Agencies wish to respond to those desires. Dynamic membership will require close supervision by the *ISC Site Administrator* to ensure that the *collection* descriptor information is current.

Terminal Collection

As *collections* can contain pointers to other collections, reflected in the included Item Descriptor ID's, there exists the concept of a 'collection structure' (see Figure 4-2), the leaves of the branches being product descriptors. The *collections* that include only *product* descriptors are termed 'terminal collections, however, Theme Collections which include Remote Collections may act as terminal collections at the RM that hosts the collection structure.

Identifier (In the following text, *item descriptor* includes *product*, or *collection descriptor):*

Each member of a *collection* (i.e. included item descriptor of any type) must have an identifier unique within all the *collections* in the *retrieval manager's* collection tree. This unique identifier can be seen as the name of the item descriptor. In particular, it will be a single unique identifier which includes the *Retrieval Manager* identifier, the *collection* identifier and the *collection* member identifier(included item descriptor ID's). For a *collection* descriptor, the *collection* identifier refers to a collection of any classification (*archive,or theme* ,), whereas for *product* descriptors, the identifier must refer to an *archive collection*. Specific format details can be found in the CIP Specification[R3].

Uniqueness:

A *collection* member (included item descriptor ID) may be a member of more than one *collection*, as illustrated in Figure 4-4, Provider Archive Collection AMSR on ADEOS II is a member of the Sea Surface Temperature Collection as well as the Andes Event Collection . However, duplicate members (included item descriptor ID's) must not be visible within a single *collection* . For example, Provider Archive Collection AMSR on ADEOSII could not appear twice as an included item descriptor ID in a collection that contained both the Sea Surface Temperature Collection and the Andes Event Collection . This property is known as elimination of duplicates to achieve uniqueness.

In the case of a *collection* which is a child of two or more included *collections*, any operation such as search,which traverses the *collection* tree from the top level collection will end up repeatedly visiting the child *collection*. The unique collection identifier provides a means of preventing repeated operations on the same *collection* . This is achieved by noting which tree nodes (unique identifiers) had been visited and then restricting access to those nodes (unique identifiers) for the same search. This functionality will be performed by the *Retrieval Manager.*

Remote Collections:

Remote collections are members of a collection hierarchy whose information is stored or maintained at a CIP site other than the one in which it is included as an included remote item.

Normally, a *collection* structure would be held in one place (say, as a database on a computer). A logical *collection* tree is where one or more of the members are held elsewhere - the complete *collection* tree thus spans multiple sites. The *collection* that is referencing(included item descriptor or included task package) a *collection* at the remote location is termed a 'remote member'.

Remote members do not have to maintain information about which members refer to them; remote members are indistinguishable from local *collection* members from the user's point of view. This concept is supported by the consistent use of URLs to identify *collections,* in the same manner as the complete WWW is seen by the user as a single database. A *Retrieval Manager* 'owns' those *collections* for which it stores and maintains the attributes, and only stores the pointers(Included Item Descriptor ID's/included task packages) to remote members, not their attributes and values.

No attribute or value of an attribute for a remote member, or the pointer (included item descriptor/task package name) to the remote member, can be guaranteed. The *Retrieval Manager* where the remote member is stored may not be available; the remote member may have changed its data structure (adding, changing or deleting attributes), or the remote member may have been deleted from the remote *Retrieval Manager*.

Related collections:

Collections may be related to one another without the need for a "parent-child" or the "include" construct. The relation may be through content or purpose, for example, and allows the spanning of one *collection* tree to another. A *collection* descriptor will contain a list (possibly empty) of related *collections* as part of its content.

Related Guides

It will be possible to relate Guide Documents to collections within the collection descriptor. This can be achieved by indicating the associated Guide URL in the Related Item ID data element of the collection descriptor.

Local Attributes

Local Attributes are collection specific characteristics about a collection. The existance of local attributes may be specified within a collection descriptor by setting the Local Use Attribute Flag Element. A flag of 0 indicates that the collection does not contain local attributes, 1 indicates that the local attributes are described within the collection descriptor whose values are captured in the Product Descriptor; 2 indicates that the local attributes have been described in the Explain and the values for the attributes are captured in the Product Descriptor. Local Attributes Using the Collections Database and Local Attributes Using Explain are addressed in Sections 4.6.1.2 and 4.6.2.2 respectively. Specific details regarding the creation/maintenance of local attributes are contained in the ICS Collection Manual.

Registered Collection -

Registered Collections contain item descriptors, implies persistence (i.e.) long lived and has a collection descriptor in the Collections Database. May be accessed via the full range of CIP access services including *discovery, navigation, location and searching.*

- *Archive collections*: : This type of collection is likely to be created by data providers to organize their archives and facilitate access to the product descriptors (i.e. analogous to an inventory containing inventory entries. May be static or dynamic
- *Theme collections*: This type of collection may be set up by data providers or users who want to organize some of their data into groupings which differ from their provider archive collections (i.e. from the baseline inventory), for the convenience of their users, for example, based on the geographical area covered, the scientific discipline supported by the data, the instrument type, etc. May be static or dynamic.

Unregistered Collections

Unregistered Collections (also known as result sets) - contains item descriptors, does not imply persistance and does not have a collection descriptor in the collections database. May only be accessed via the locate service or by the creator of the collection.

Persistent result sets - implies persistance, May be static or dynamic. May be accessed using *Locate* service.

Temporary result sets- implies no persistance (i.e. only available in the z-association in which it was created). Is static. Can only be accessed by creator in the same z-association in which it was created.

Included Task Packages

Task Packages may be included as members of a collection. The purpose for including Task packages would be to establish a relationship between the collection and the extended services result sets which may be a persistent query or an accumulation of result set item descriptors. Reference Section 4.6.4 for a description of how included task packages would be used by the Retrieval Manager.

4.5 User Interactions with Collections

ICS Compatability: Explanatory

This section describes multiple methods a user will have available using the CIP within the ICS domain to discover items in *collections*, e.g. *products, guide, browse,* etc. Four methods for using the *collection* structure are described:

- Collection Discovery
- Collection Navigation

- Collection Searching.
- Collection Locating.

Each of these methods differ based on what the user needs to know before using the method and what element of the *collection* descriptor is used by the method. An example is that for *collection* discovery the user needs to know very little of the ICS, compared with collection locating in which the user must have a specific *collection* ID to use the method.

These methods are depicted in Figure 4-5 which illustrates each of the *collection* usage methods identified above relative to a *collection* node. Each of the *collection* usage methods are described in the remainder of this section.

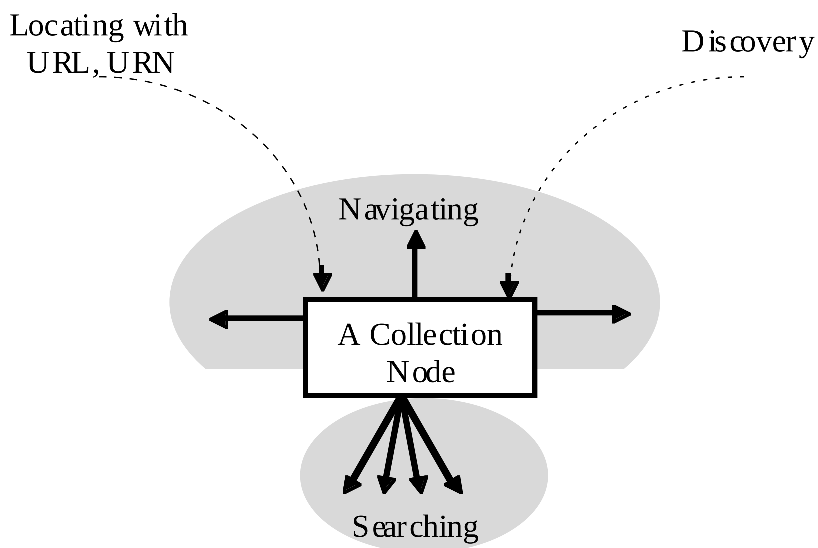


Figure 4-5. Collection Usage Methods

Collection Discovery:

Collection discovery allows a user of the ICS, having no prior knowledge of the *collection* structure, to be assured that all *collections* of interest to the user are found. It is assumed that the user will know of at least one *Retrieval Manager*. This method provides a user with an exhaustive method to determine all *collections* which may be of interest. The discovery of *collections* makes no assumptions about the user knowledge of any existing *collections* or other *Retrieval Managers*.

For CIP-B, *collection* discovery will be accomplished by the relatively simple method of searching a global *collection* which is discussed in Section 3. It is anticipated that information discovery methods currently being researched in WWW community may have value for accomplishing *collection* discovery in later releases.

Collection Navigation:

Collection navigation allows the user to examine *collections* related to a specified *collection* and switch context to the related *collection*. This is achieved by allowing the user to navigate through a collection structure resident on the user's client. When the user reaches the edge of the collection structure a search request will need to be made to the RM to retrieve the collection of interest. The related item descriptor is the data design construct which supports this functionality. The Collection Navigation service will apply only to the Archive collections and Theme Collections.

Collection Searching:

Collection searching is the core function of the CIP. With this method the user knows of a *collection* of interest as well as the *Retrieval Manager* where the *collection* descriptor for the *collection* resides. The user establishes a session with the *Retrieval Manager* and targets a search at the *collection* The search is then propagated down through the included item descriptors of the *collection* and any included *collections*.

Collection Locating:

Collection locating assumes that the user has a *collection* URL in their possession and wishes to find the *collection*, examine the *collection* elements, and perhaps target a search at the *collection*. Having a URL, the user can examine the URL and determine information about the *collections* location, the user can submit the URL to a *Retrieval Manager* which can locate the *collection* or determine that the *collection* no longer exists.

The following table summarizes the salient concepts which characterize the various collections that have been presented throughout Section 4.2.

The **Owner** is the individual(s) responsible for the management and ownership of the collection. **Ref in Coll Structure** implies that the collection can be included in an existing collection. **CIP Metadata** refers to the existence of data about the collection in the Collections Database. **Lifespan** is the anticipated longevity of the collection. Please note this is the anticipated longevity. Collections may be deleted under certain circumstances as identified in the CM. **Services** are the methods that may be used to access the collections and lastly **Attributes Homogeneity** refers to the likelihood that the characteristics which describe the collection will be similar.

Table 4-1. Collection Concepts Summary

	Owner	Ref in Collection Structure	CIP Metadata	Lifespan	Services Capable of Discovery	Attributes Homogeneity	Static or Dynamic
Registered Collections							
Archive Collection	Admin.	Yes	Yes	Long Term	Discovery Navigate Search Locate	High	Based on specific collection
Theme Collection	Admin.	Yes	Yes	Long Term	Discovery Navigate Search Locate	Medium	Based on specific collection
Unregistered Collections (i.e.result sets)							
Persistent Result Set	Admin. / User	Yes	No	Medium Term	Locate	Low	Based on specific implementation or collection
Temporary Result Set	User	No	No	single Z-association	None	Low	Static

4.6 ICS Data Framework

This section of the Data View will present the logical data components of the ICS. This will be achieved by describing and defining each data component within the overall Data Framework. This framework as illustrated in Figure 4-6 represents the logical data objects which will be managed and maintained in ICS. The emphasis, in the illustration, and this section, is on the data held in the *Retrieval Manager* which includes the:

- *Collections Database (CDB)* (Section 4.6.1),
- *Explain Database* (Section 4.6.2),
- *Extended Services Database* (Section 4.6.3),
- *Result Set Database (*Section 4.6.4).
- Session, Error and User Data (Section 4.6.5-4.6.7).

The data in the IGP Domain are defined in the Guide Design and Protocol Specification Document. However, Section 4.7 of this document provides an overview of the ICS Document Data.

The data required by elements other than the *Retrieval Manager*, for example external interfaces, are recognized in Section 4.7. Each data component will be described in the sections indicated above. These descriptions will include an overview of the contents of the database, an object model, which will illustrate the data objects, attributes and relationships for each of these data repositories, and lastly any unique design characteristics surrounding these databases. However, the CIP schema described in the CIP Specification [R3] is the definitive source for all schema definitions. Additionally, the valid domain of values associated with each attribute is identified and described in the ICS Valids Document [R4].

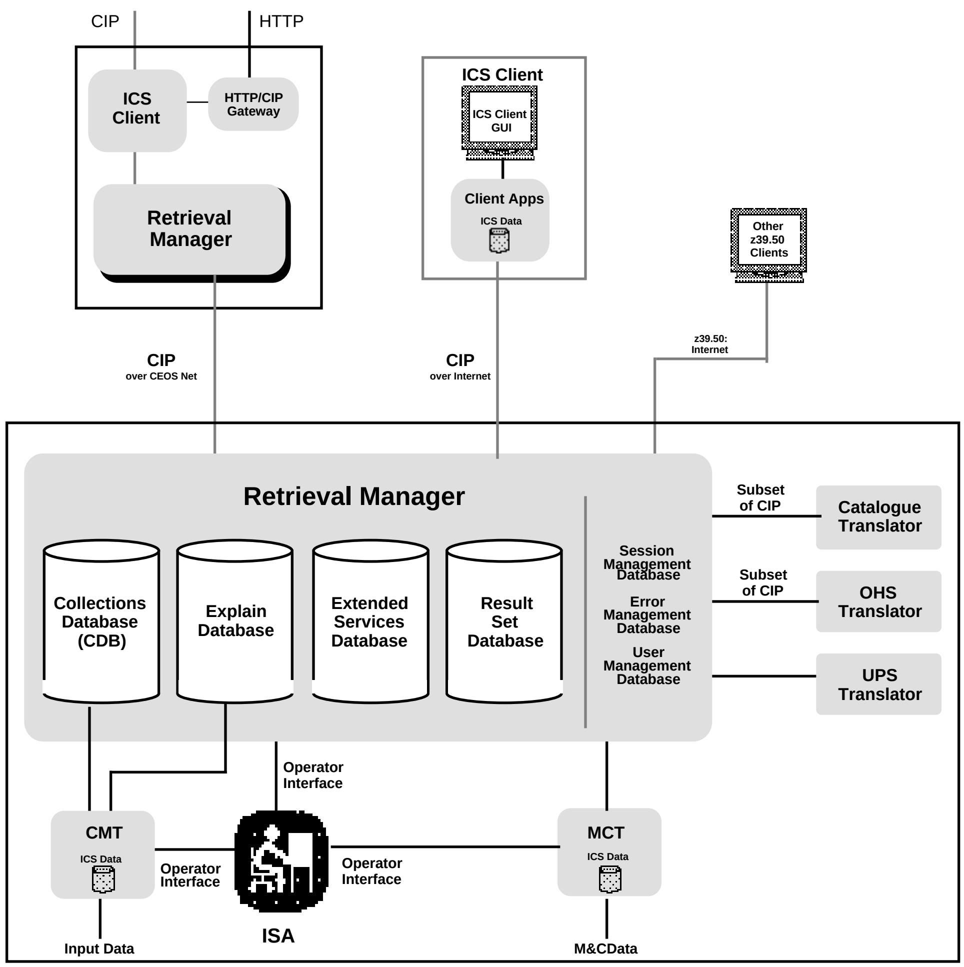


Figure 4-6. ICS CIP Data Framework

4.6.1 Collection Database(CDB)

The *collection database* contains the instances of the science metadata which are referenced from the model illustrated in Figure 4-1. The metadata coupled with the organization of the metadata within the *collections database* is used to assist the user in discovering, searching, and ordering Earth Observation information. The intent of the *collections database* is to satisfy these objectives within a framework which will allow maximum flexibility while minimizing change to the overall structure. This is achieved through the item descriptor which is a generalization of the specific classes of data which are supported under the *collection database* umbrella. The item descriptors supported within CIP's *collection database* are the *collection,and product,* Descriptors. The following definitions provide the context for each of these descriptors.

Collection Descriptors	Collection metadata serves to characterize the underlying product information/included collections. These characteristics consist of spatial, temporal, source, quality, document references, and keyword informational elements.
Product Descriptors	Product metadata characterizes the individual products contained in the archives. Each product can be described in terms of it's spatial, temporal, source and quality attributes.

4.6.1.1 Collection and Product Attributes

ICS Compatability: Explanatory

For any of the above descriptors, a standard set of attributes are required (see CIP Specification - Release B [R3]). Note that not all catalogue systems will support all attributes, but the minimum set of attributes as specified in the CIP Specification should be supported. In the case of *product* descriptors (the attributes of which are based upon currently existing inventory entry definitions), there may be attributes that are local to a particular catalogue system. The CIP will permit the transfer and searching of these attributes without needing to 'understand' their real world meanings.

An example of how the attributes can be used is presented in the tables below. (Note that the attributes represented in these examples are for illustrative purposes only. These examples do not include the full list of mandatory attributes. The complete result schema is defined in the CIP Specification Release B, Appendix C [R3], which indicates those attributes that are mandatory or optional.

The list of *product* descriptor attributes that might describe an AVHRR data product, and the corresponding values for a single instance of that product, might be:

Product Descriptor Attribute	Example Value
SensorName	AVHRR
EndDate	1994-07-10T17:20:43.123456Z
MissionId	NOAA-09
SpatialCoverage	-32.27, -79.39, 62.55, 65.32
ArchivingCentreId	OBERPFAFFENHOFEN
ItemDescriptorId	PID_103.47c6dec1

Finally, the EO collection descriptor for a *theme collection* may be part of a *collection* that contains all AVHRR data covering the European continent, a list of the *collection* descriptor attributes and their values that might describe such a collection could be:

Collection Descriptor Attribute	Example Value
ItemDescriptorId	cip://ciprm.esrin.esa.it/CID_121
ItemDescriptorName	AVHRR
CreationDate	1994-05-01T00:00.0
CollectionHierarchyPosition	NON-TERMINAL
CollectionCategory	Theme Collection
Purpose	This collection is a grouping together of all known AVHRR data that is available via the ESRIN and DLR EO data centres.
RevisionDate	1995-01-15T00:00.0
VersionId	2.3
SpatialCoverage	-10, -70, 35, 70
Progress	IN WORK
BrowseId	http://gds.esrin.esa.it/avhrr/collections/europe.jpg
IncludedCollectionDescriptors	cip://ciprm.esrin.esa.it/CID_226 cip://ciprm.dfd.dlr.de/CID_125

The list of attributes for each of the above item descriptor objects (*product* descriptor, and *collection* descriptor) will be different, as the information carried by these objects is inherently different. It is also unrealistic to assume that (apart from any mandatory subset) the list of attributes for any one object within one agency will be exactly the same within another agency; this is especially true for the *product* descriptor objects, where existing archives and inventory entries are very different.

The CIP Specification - Release B specifies a set of attributes for describing *collections*. Not all servers need to support all *collection* attributes, however a mandatory set must be specified (see Appendix C of [R3]). For *product* descriptors there are several mandatory attributes that shall ensure basic interoperability between *product* descriptors (see Appendix C of [R3]).

4.6.1.2 Local Attribute Definition Using the CDB

ICS Compatability: MAA

Each CIP Archive Collection, can be further characterized beyond the standard set of elements specified in the CIP Specification [R3]. This is achieved through the identification and description of local attributes. This capability will allow the developer of the collection to identify and record, either in the Explain Database and Collection Database, or only in the Collection Database, additional characteristics about a collection such as AverageBlack Body Temperatures, to give an example.

The process and overall procedure for capturing and recording this information is described in the Collections Manual. The data model for the Collection Database, Collections Module and further the local attributes are specified in Figure 4-7.1. From this figure one can determine the information that is necessary to describe the collection specific data that can not be described using the standard set of attributes. The key ISA decision surrounding the implementation of the local attribute concept is whether to record the local attributes characteristics in the Collection Database or in the Explain Database. These characteristics are the explanations of the local attributes such as the definition of the local attribute, the meaning, the name etc. The local use attribute flag element which is contained in the collection descriptor expresses the ISA decision and is used by the search and present services to determine the target (Collections Database or Explain) for the request.

4.6.1.3 CDB Support for CIP / GEO Interoperability

ICS Compatability: Explanatory

The collections database schema currently contains the necessary elements and use attributes to support data interoperability between the GEO and CIP. This was achieved by aligning the data between the two applications, paying careful attention to the semantical and structural differences and identifying a set of common characteristics which may be applied to both applications. Due to this alignment users of either CIP or GEO can formulate queries which may target each other and expect semantically similiar results. The specifics surrounding the data and functionality required to ensure this interoperability can be found in the CIP/GEO Specification [R].

4.6.1.4 ICS Valid Values

ICS Compatability: Mandatory(See R4)

The ICS has an associated list of valid values which are specified allowable terms for a select group of attributes. These valid values serve the following important purposes.

- Establishes a clear definition of the use attributes by providing a list of examples (valid values) that maybe used in populating the attributes.
- Maintains the integrity of the data for searchable attributes.
- Ensures data interoperability across sites where a variety of data value solutions could be applied to a given use attribute.

The ICS Valids Document [R4] contains the list of attributes and associated controlled list of valid values. This document also specifies the ICS procedures for maintaining the valids.

4.6.1.5 Collections Database(CDB) Data Model

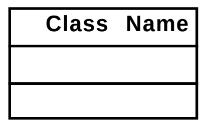
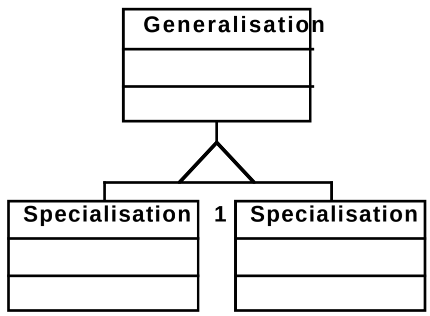
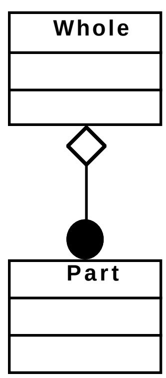
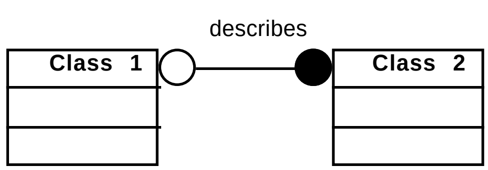
ICS Compatability: Mandatory(See R3)

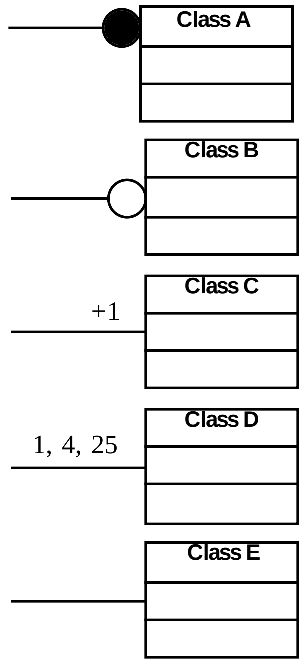
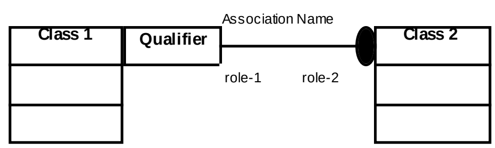
The Collection Database Data Object Model, which is illustrated in Figure 4-7.1 through 4-7.5, expresses the EO data objects, attributes, cardinality, and relationships between the objects. There are five modules which collectively represent the ICS CDB. The modules allow a partitioning of the entire CDB into manageable subsets. These models attempt to illustrate the relationship between the collections and other CIP objects such as the Explain Database and the Extended Services Database.

Description of the CIP objects and attributes are recorded in Appendix “B” and “C” of the CIP Specification - Release B [R3]. The ARS’s which is the basis for these models is also contained in the Rel B Specification, Section C.3.[R3] The Business rules associated with the creation and maintenance of these objects are defined in the ICS Collections Manual.

The following table identifies and describes the Object Modeling Technique(OMT) Rumbaugh notation that is used in the data models contained in Figures 4.7 - 4.15. Column one contains the OMT term, column two the graphical interpretation of the term and column three a brief description of the notation.

OMT Notation

Class-&-Object		Defines a class and its associated objects which provide an abstraction of something within the problem domain. The class contains attributes (data) and methods (functions) which are common between all objects in the class. Example: Collection is a class.
Generalization; **Specialization**		Defines generalization - specialization relations between classes. The specialization classes inherits the attributes and services of the generalization class. A generalization class can have one or many specializationÕs. A specialization can have one or more generalizations. Example: a Map Projection is a specialization of Planar System Class.
Whole-Part		Defines whole-part relations between classes. Example: Processing Options, Processing, and Scene Selection, are a part of Product Delivery Options.
Association **Name**		Identifies the relationships among classes... A relation has two labels, forward and inverse. For Example, Class 1 describes Class 2 is a forward direction,; the opposite would be Class 2 is described by Class 1 which is the inverse. These labels help to define the role of two or more classes with respect to the relation.

Multiplicity of Associations		Multiplicity specifies how many instances of one class may be related to a single instance of an associated class. For Example An instance of a collection may relate to one or more Spatial Coverage's. The examples on the left express the ways in which multiplicities may be specified: Class A=many(zero or more) Class B= optional(zero or one) Class C=one or more Class D=numerically specified Class E=exactly one
Qualified Association		Role names identify the participation of the objects in the overall association. For example The relationship between Collection and Contact states that each Collection must have at least one Contact who is an administrator. Qualifier is an attribute which serves to further qualify the association. For example the association between the Product class and Product Delivery Options class is qualified by the attribute Group ID

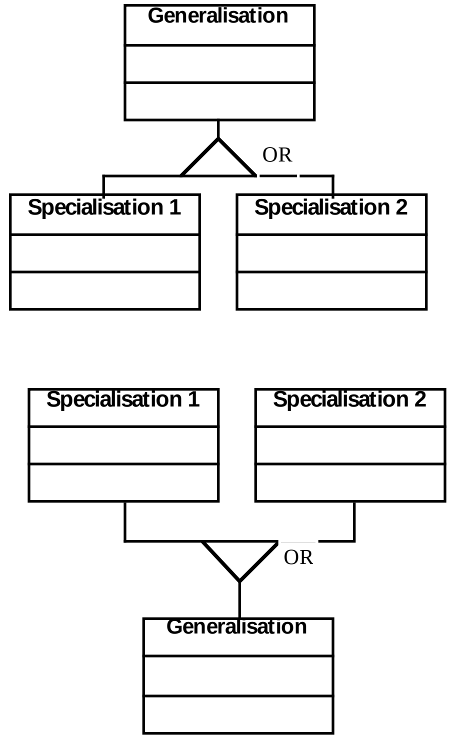
Exclusive Or		Exclusive "OR" indicates that only one of the specialization classes may exist for a given instance of the generalization class. This is graphically illustrated by indicating OR next to the specialization symbol.

Figure 4-7.1 reflects the collection module which illustrates the data objects, attributes, and associations related to the collection. This model reflects the breadth of classes and relationships that could exist for any instance of a collection. The off page classes, Spatial Reference, Spatial Coverage, Database Info(Explain) and Item Descriptors(Extended Services) are illustrated in other CDB figures, Explain Database or Extended Services Database figures. The Item Data and Browse Pointer are references to instances of browse information that are characterized outside of the CDB.

Figure 4-7.2 reflects the product module which illustrates the data objects, attributes, and associations related to the product. The same explanation provided above also applies to this module with the exception of the Database Info which does not relate to the product.

Figures 4-7.3 and 4-7.4 illustrates the spatial coverage and spatial reference modules respectively and their individual data objects, attribute, and associations. In the Spatial Coverage Module the Spatial Coverage class is a special purpose class which exists to add clarity to the module. It will not contain attributes and therefore will not have instances other than the indicated specialization's. The G Polygon is also reflected in this module to add clarity and will not contain instances other than through the specified relationships Outer G-Ring and Exclusion G-

Ring. The Spatial Reference Module contains several special purpose classes which are identical in purpose to the classes describe in the Spatial Coverage Module. The Horizontal Coordinate System, Vertical Coordinate System, Spatial Reference and Planar System are all special purpose classes which serve to add clarity/definition to the module.

Figure 4-7.5 reflects the Product Order information by illustrating the order objects and associated attributes. The relationship between this module and the Extended Services Database is also captured in the Order Module.

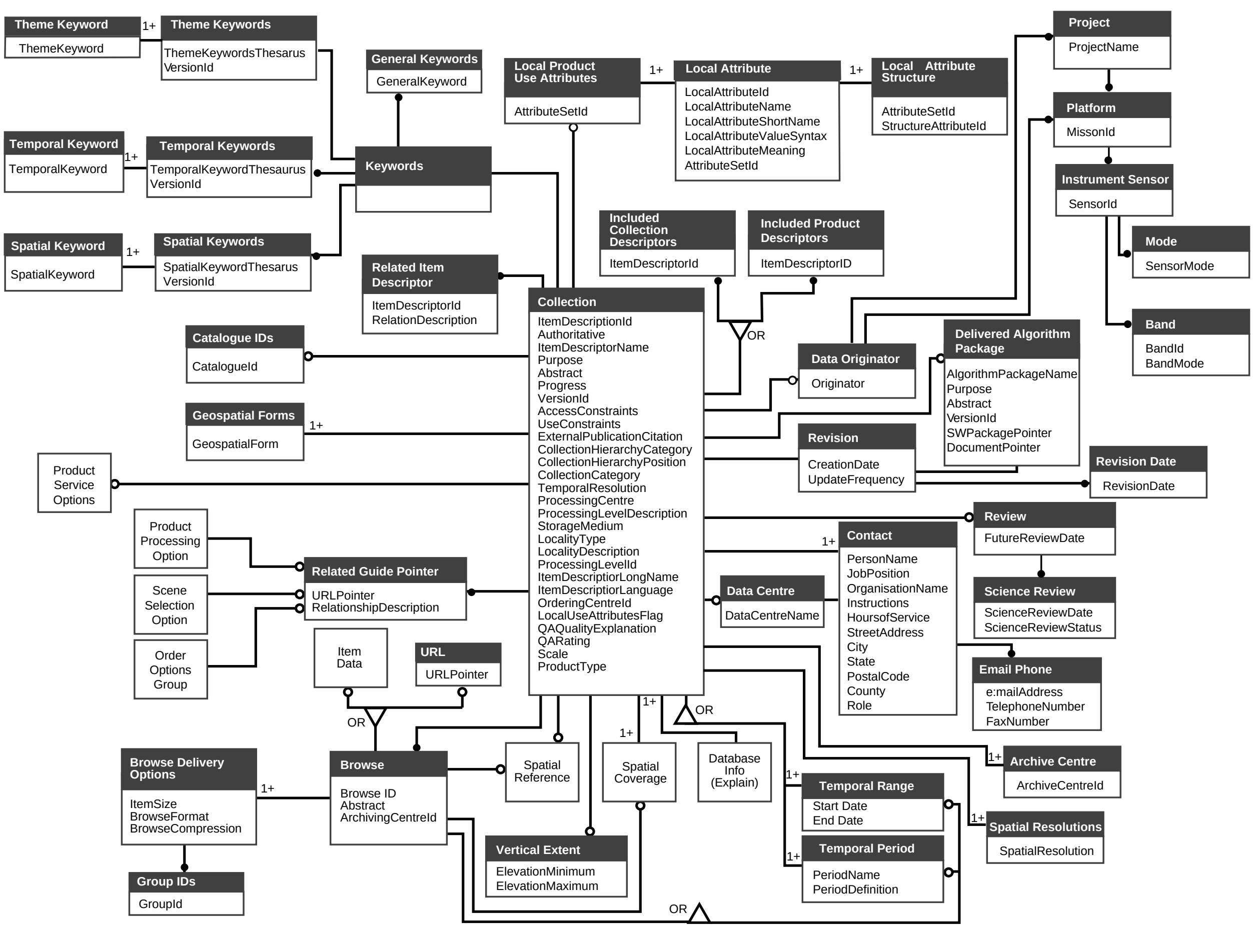


Figure 4-7.1. Collections Database Data Object Model - Collections Module

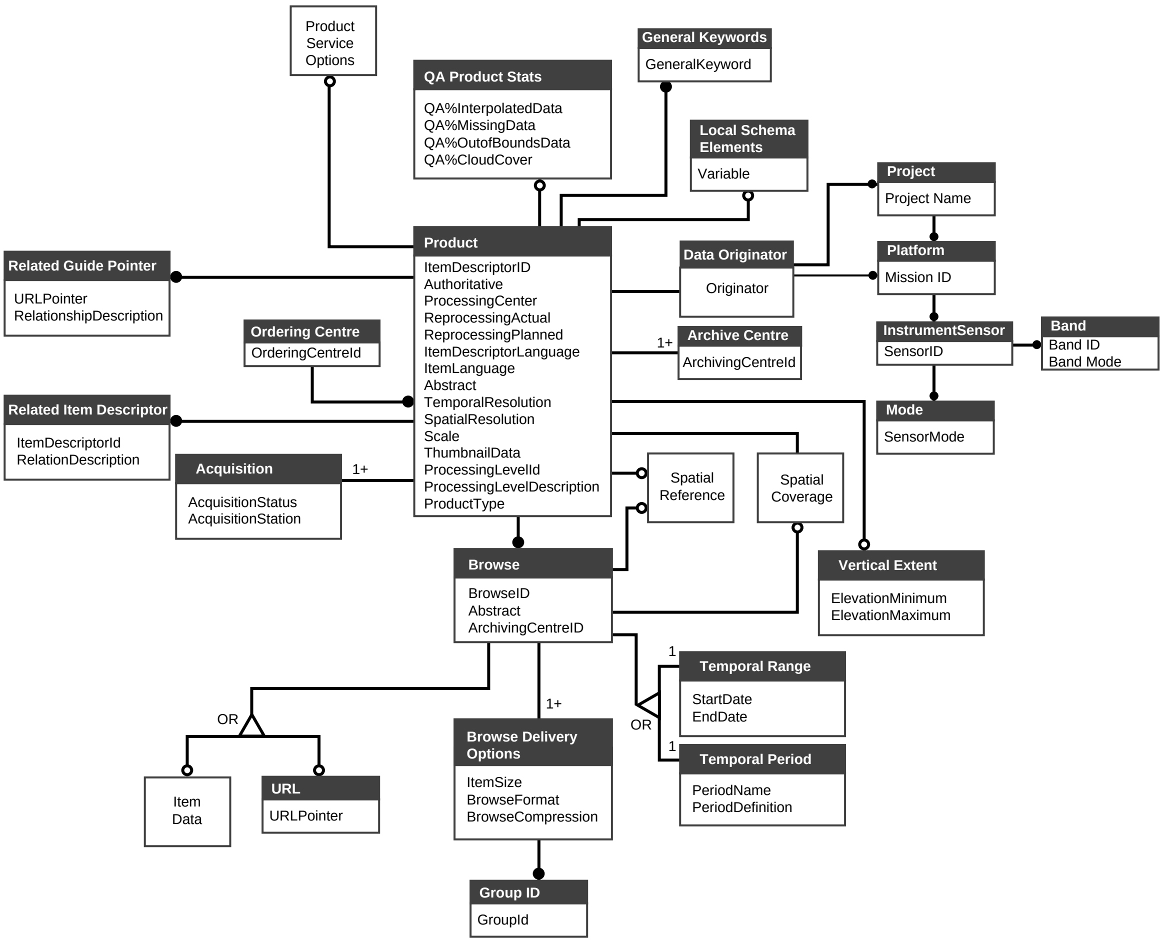


Figure 4-7.2. Collections Database Data Object Model - Product Module

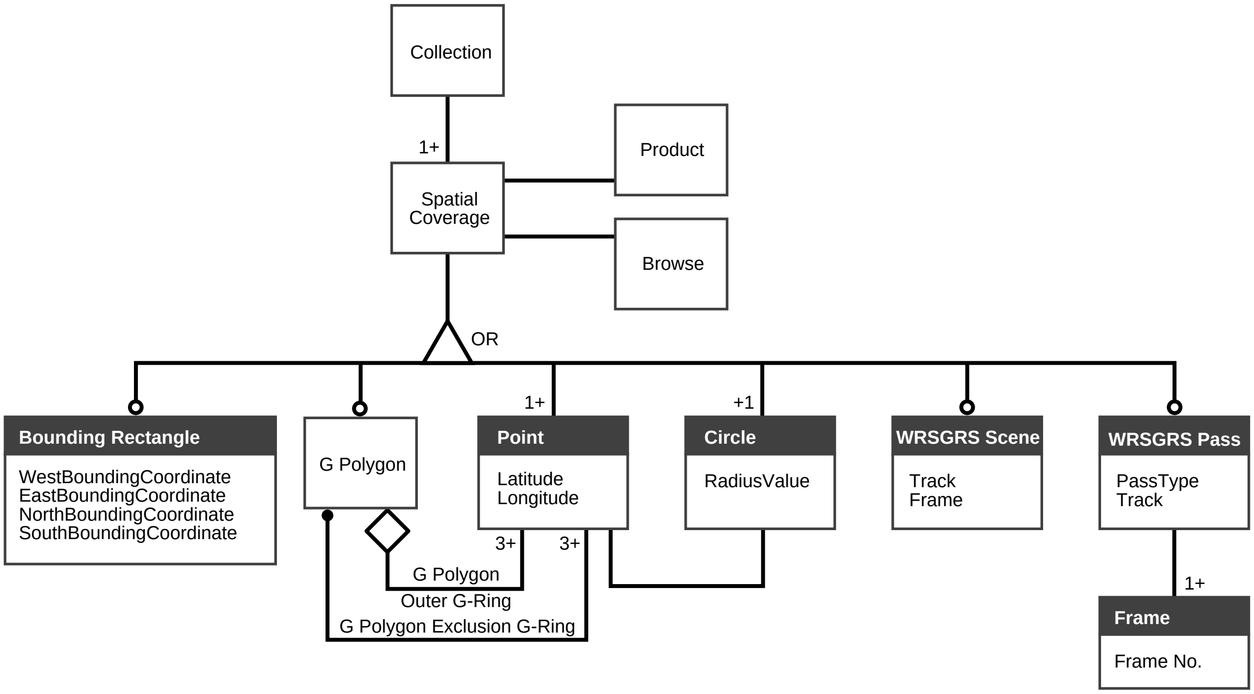


Figure 4-7.3. Collections Database Data Object Model - Spatial Coverage Module

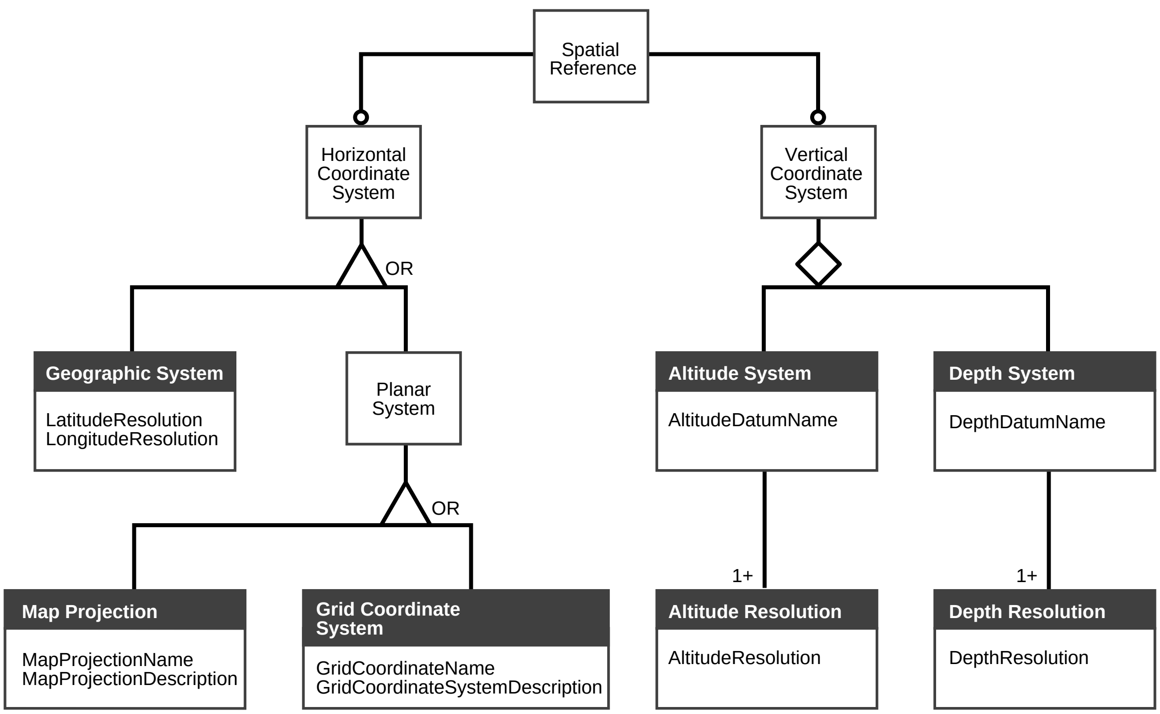


Figure 4-7.4. Collections Database Data Object Model - Spatial Reference Module

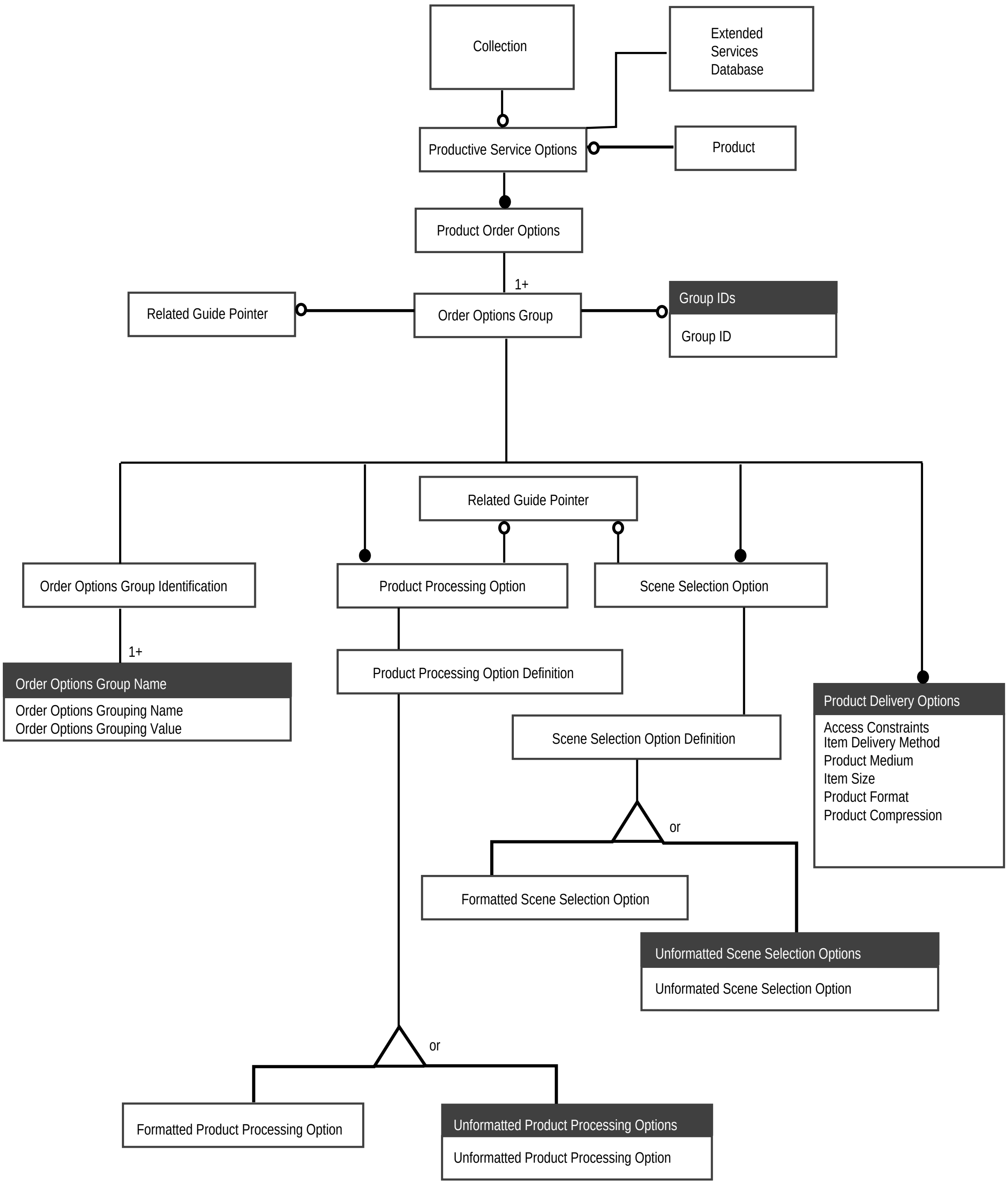


Figure 4-7.5. Collections Database Data Object Model - Order Module

4.6.1.6 Collection Census

ICS Compatability:Explanatory

This section provides statistics on the number of collections anticipated in ICS coupled with estimated metadata sizing per product.

It order to support the design analysis, a rough order of magnitude estimate of the number of collections in ICS is needed. A specific count of the collections is not necessary nor is a list of named collections. The estimates provided here are also not aimed at indicating an expectation that a particular agency will provide access to the number of collections indicated. The purpose of the estimate is to provide a high end estimate of what the CEOS collection space might grow to. This estimate is provided in Table 4-1.

The estimate reflected in Table 4-1 was discussed during the CEOS Catalogue Sub-Group meeting in May 1996. (The Catalogue Sub-group has been renamed the Access Sub-group.) It was the conclusion of the Access Sub-Group that the estimate in Table 4-1 significantly under estimated the number of collections which could be indexed using CIP. In particular, the GCMD experience indicates that there are millions of datasets, although a large portion of these datasets will never be accessible on-line. The number of datasets which may provide on-line directory metadata may be on the order of hundreds of thousands. And for datasets which would have inventory on-line, an estimate on the order of tens of thousands was felt to be accurate.

For now the best that can be said is that CIP should be sized for hundreds of thousands of collections with collection level metadata only. Further, CIP should be sized to hold tens of thousands of Collections containing product descriptors.

Table 4-1. ICS Collection Upper Bound for Design Sizing

Agency	Rough Estimate of the Number of Archive collections
European	
BNSC	**100 +/- 50**
CEO	**50 to 500**
CNES	
DLR	
ESA	**50**
Canadian	
CCRS	**500**
Japanese	
NASDA	**70**
US Agencies	
NASA (EOSDIS)	**Version 0:1000 Version 1: 2000**
NOAA	**1200**
USGS	
Total	**Roughly 5,420**

Also important to ICS element sizing is the size of product metadata, i.e. the size of a product descriptor. This number will certainly vary for the various products. For a first order estimate, an estimate from ECS is provided here. ECS is using an estimate of 2K/product. This estimate is being used for both ECS Release A (predominately Version 0 and TRMM data) and for the ECS DBMS prototype for ECS Release B (Release A plus EOS AM-1, Landsat-7, and others).

4.6.2 Explain Database

The Explain Database offers the ICS the structure and information necessary to respond to the Z39.50 Search, and Present Services. This component of the ICS data architecture “Explains” the data or information environment in which the ICS will operate. This environment consists of categories of information that the target supports. These categories would include databases, schema’s, record syntax, attribute sets, and extended services. For each of these categories, a detailed description, which is contained in the attributes, relationships, and associated data objects serve to provide the model under which the ICS will capture, store, and make available information about the ICS *Retrieval Manager*. The Explain database data model which is illustrated in Figures 4-11.1 to 4-11.2 graphically depicts these concepts in addition to the relationship between Explain and the Collections Database.. Additional details regarding each of the objects in the data model are described in the CIP Specification - Release B [R3]. The creation and maintenance of the Explain entries are described in the ICS Collections Manual.

4.6.2.1 Explain Database Data Model

ICS Compatability: Mandatory

The following briefly describes each of the data objects referenced in the Explain Data Model Figure 4.11.1. Collectively these objects serve to organize the Explain data into manageable entities which can be readily understood by data engineers.

Database Info

This object describes each collection in the logical CDB. The ISA must create an Explain Record in the Database Info for each CDB collections entry. This explain record should include the database name, which must be described according to the specification in the CIP Release B Specification [R3], Appendix D.2, Database Names. Access available, Record Count, Average Record Size, User Fee, and Max Record Size should also be specified in this record. Additionally, if the collection is a Root Collection then the content of the Keyword Attribute should reflect Root.

Schema Info

The Schema Info Data Object identifies the abstract record structures which are the elements that describe the collection referenced in the Database Info. There are currently two (2) record structures that are supported in the ICS:

- Collections
- Products

Each Database Info entry must relate to at least one of the above entries in the Schema Info.

Retrieval Record Details

This object contains the descriptive details about the elements of the retrieval record.

Record Syntax Info

The Record Syntax Info is used by the client to determine the format of the retrieval record. Four formats are currently available; Explain; Sutrs; GRS-1; and Extended Services. Each entry in the Database Info must relate to one or more Record Syntax's. The Record Syntax Info coupled with the Element Set Details (described below) allows the client to determine both the format and content of the retrieval record.

Element Set Details

This data object allows the client to determine the data that will be retrieved by specifying the Element Set Name. To achieve this the Database Info and Schema Info Records which were created for each collection must be related to an instance of the Element Set Details. Each Element Set Details is then related to an instance of an element(s) in the Per Element Set Details, Explain Records. For example the Element Set Detail Name = "FULL" has been related to the Per Element Set Details; Element Name = Item Descriptor ID for the Schema; Name ="Collection" for the Database Info;Name="XYZ"

There are six element set names:

- Full
- CIP Full
- Brief
- Browse
- Collection Member
- Options
- Local Attributes
- Summary

Per Element Set Details Details

This data object contains a description of each element and is related to the instance of the Database Info through the Schema Info and Element Set Details.

Attribute Set Info

The Attribute Set Info provides a description of each attribute set supported by the target. This set includes the OID, Name and description of the attribute set.

Attribute Type

This specifies the type of attribute i.e. Use Attribute, Truncation Attribute, Relation Attribute, etc. This object also includes the name and description of each attribute type.

AttributeDescription

Specifies the description of the attributes associated with the attribute type. For example an attribute type = “Relation” may have an attribute description = “Less Than” and Attribute Value = “1”. The attribute value is the ID# associated with the attribute description.

Query Types

This object specifies the query types supported by the RM. The RPN is the only Query Type currently available in the CIP.

Unit Info

The Unit Info specifies the unit system supported by the target system. The ISA establishes this information from the contents of the data that the RM is supporting.

4.6.2.2 Local Attribute Definition Using Explain

ICS Compatability: MAA

The Explain Database may be used to capture the definitions of the local attributes rather than the Collections Database (ref Section 4.6.1). The process for recording this information in Explain is described in detail in the Collections Manual [R5]. The important concept within the SDD is that the Explain Database will not require additional attributes, objects or relationships beyond what is specified in the Explain Data Model Figure 4.11.1, to support this requirement.

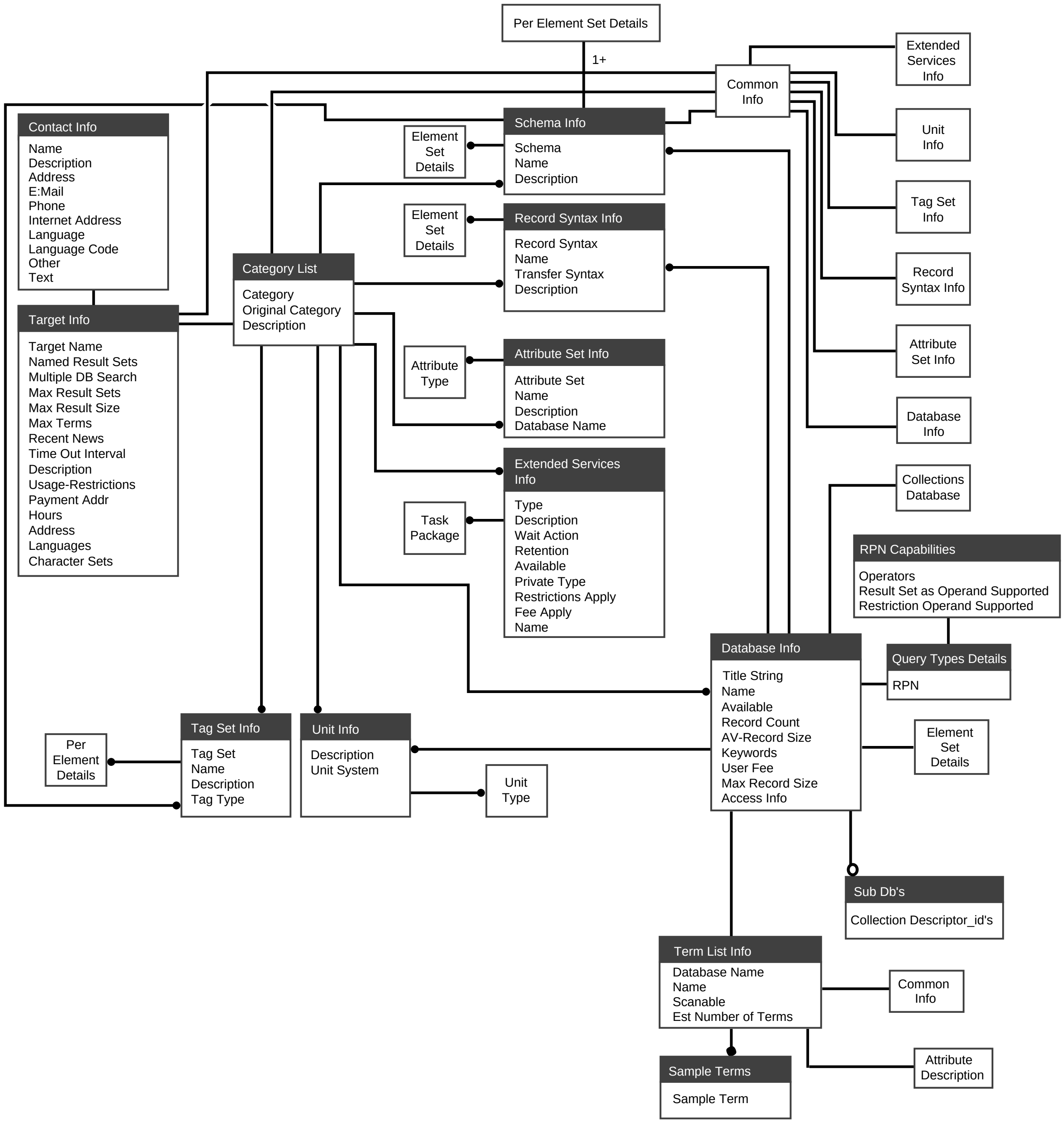


Figure 4-11.1. Explain Database

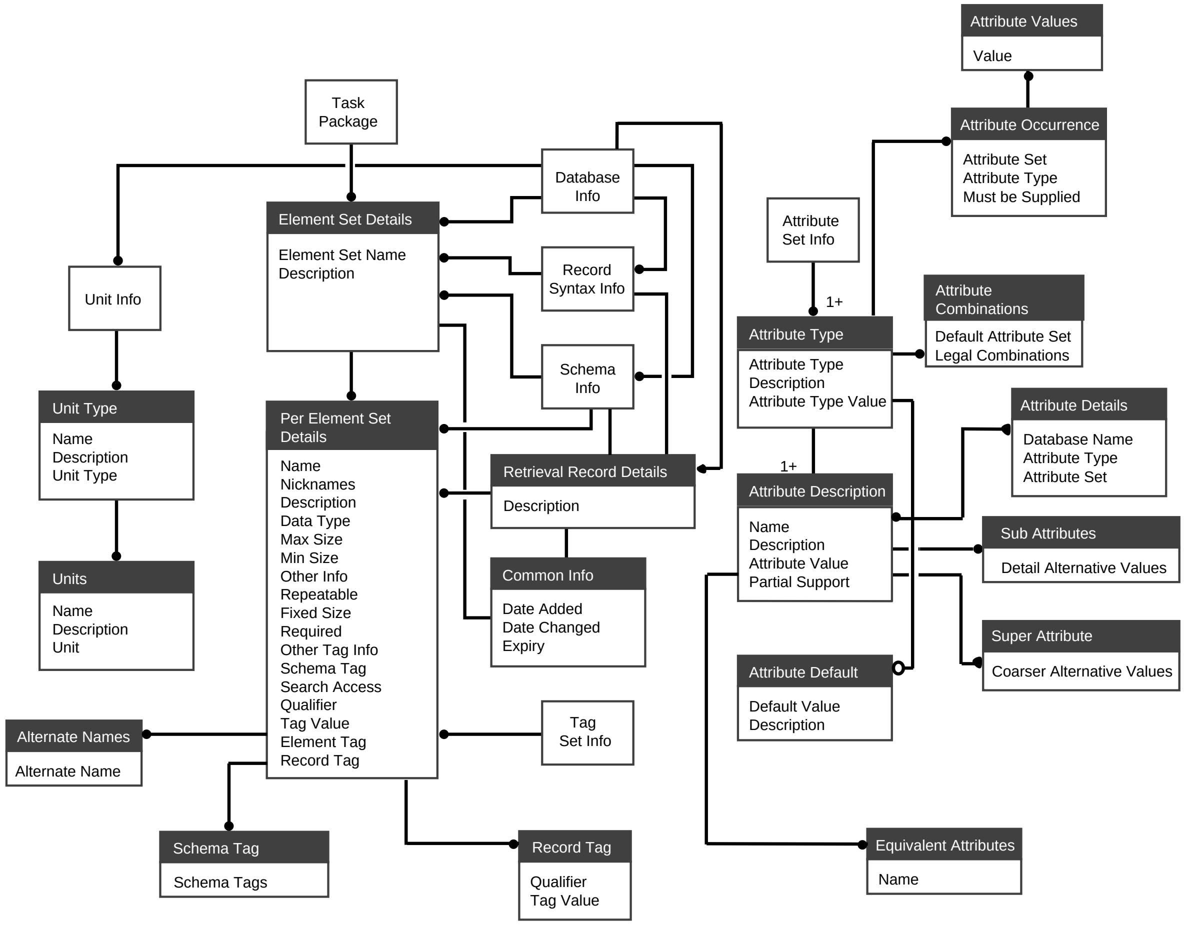


Figure 4-11.2. Explain Database

4.6.3 Extended Services Database

ICS Compatability: MAA

The Extended Services Database component of the ICS Data Architecture provides the information necessary to support the Z39.50 Extended Services Facility. This facility offers to ICS the extensions referred to as Task Packages, necessary to support the persistent storing and accessing of Order, Query, Database Update, and Result Set information. The Extended Services Database contains a description of each of these task packages. The persistent result set task package contains the administrative information about persistent result sets This would include the package ID which would be the ID of the Theme Collection referenced in the collections database.The Order Task Package contains the information necessary to respond to an order request from an origin. The Query Task Package provides the metadata associated with the persistent queries which are stored in a query file. Lastly, the Database Update Task Package provides a journal of the database activity associated with an ICS database. Each of these packages may have assigned a User or a Group of Users and their assigned privileges. Figure 4-12 graphically illustrates the Extended Services Database data object model. Additional details regarding each of the objects in the data model are described in the CIP Specification - Release B [R3]. Beyond these details the relationships between the extended services and the collections are illustrated in Figure 4-12.

The following briefly describes each of the major objects referenced in Figure 4-12.

TaskPackage
This specifies the administrative information (status, time of creation etc.) of the task package in the Extended Services Database.

Result Set
This object contains the characteristics of a persistent result set to include the Result Set Name and the Number of Records contained in the result set.

OrderSpecification
The order specifies the characteristics about a specific order such as the status, order date, Delivery Method, Package Medium, Pricing Information, and Billing Method.

Persistent Query
This object contains the characteristics associated with a persistent query such as the definition of the query, parameters associated with the execution of the query. This object is related to the database info object as the persistent query must target a collection.

Periodic Query Schedule
This object specifies the characteristics associated with a query schedule. This query schedule relates to both the Persistent query that it may execute and persistent result sets which the persistent query may generate.

DB Update
Specifies database update activity for a specified database.

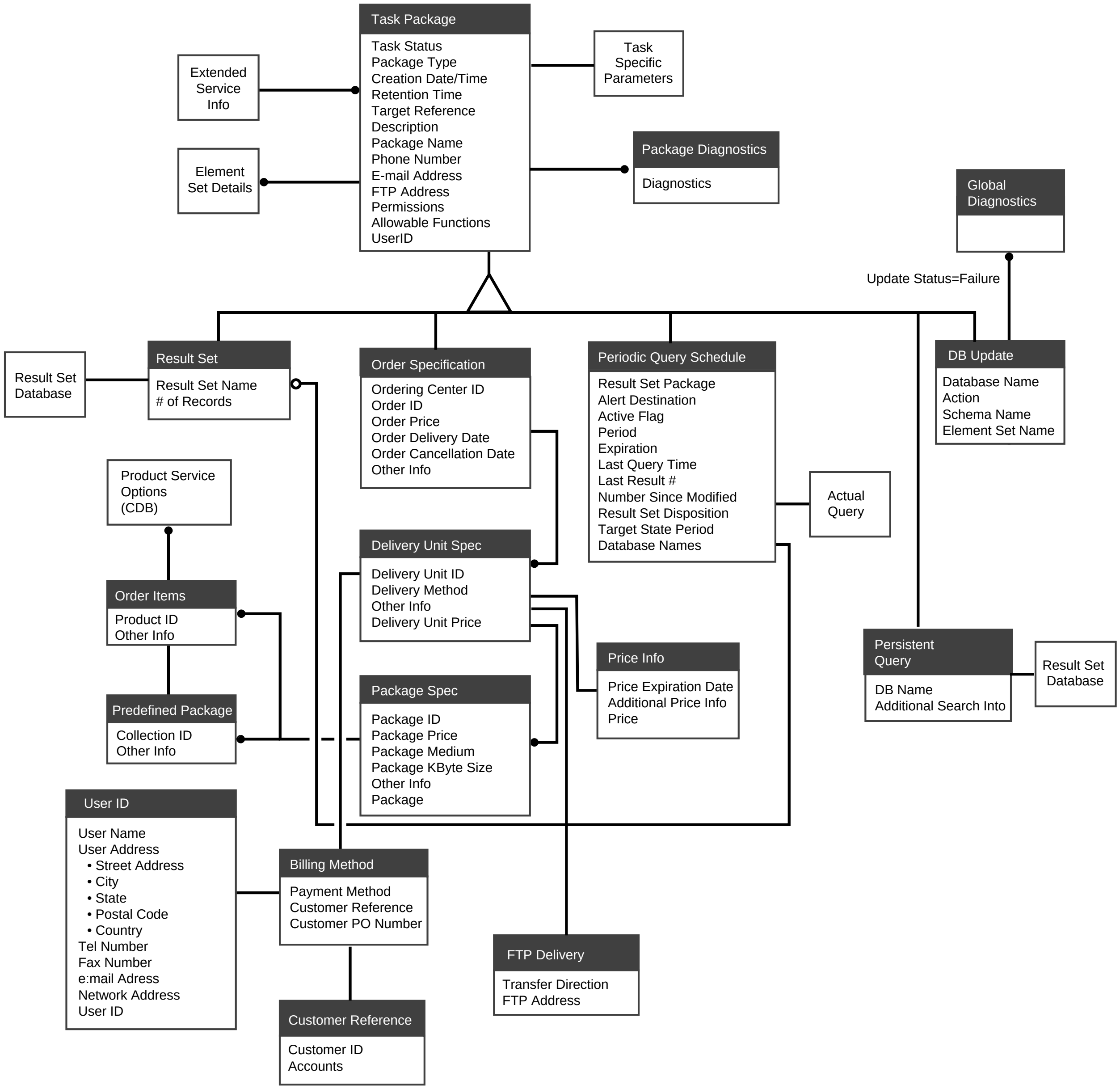


Figure 4-12. Extended Services Database

4.6.4 Persistent Result Set Database

ICS Compatability: MAA

The Persistent Result Set DB contains references to either collections that are in the result set or queries, that when executed, dynamically generates the list of collections. These references may be local or remote or a combination of both. When a collection is encountered which includes a task package reference by including an IncludedCollectionDescriptors, which references a persistent query, an extended services request is generated that targets the extended services database using the packageName referenced in the collection. An extended services response is generated in response to this request which contains a RPN Query. The RM appends the RPN Query(persistent query) to the original query and executes. The result set contains the intersection data, that is the data that satisfies both the persistent query and the original query.

The following Result Set Database Data Model captures the relationships between the Extended ServicesEnvironment, Result Set Database, Local Inventory and the Collections Database.

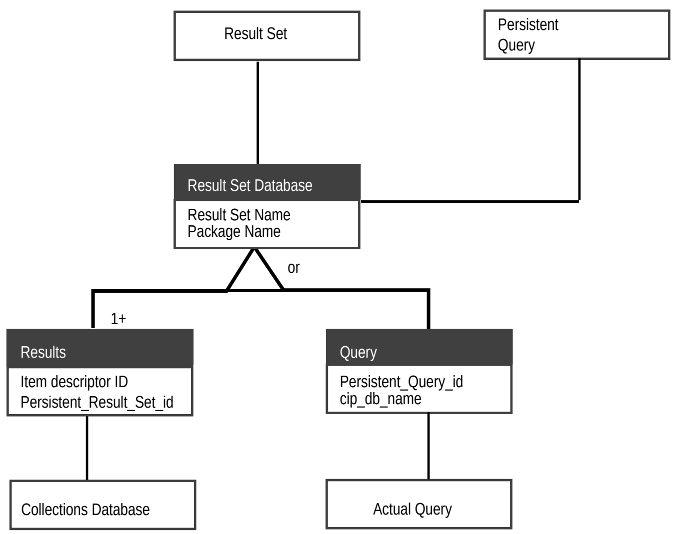


Figure 4-13. ResultSet Database Data Model

4.6.5 Session Management Database

ICS Compatability: Mandatory

The session management data will contain the information necessary to support Query, Results and General Session activity. The intent of this data are to provide the necessary information to support the interaction between the *Retrieval Manager* and the interfaces external to the *Retrieval Manager*.

- Query Data is characterized as information required to support the query processes to include both search and present. Additionally, Query Management information would also be included which would reflect Routing, and Status of Queries
- Results Data will include all of the information required to support results processing.
- General Session Information will include User Interaction Session logs which would contain the User ID, Location, Start and End Times, Profile of Tasks Performed, and Resources Used. Additionally, the CIP User Log which would capture Authentication Attempts, Successes, Failures, Service Options Accessed and Unsuccessful Log Ins would also be included, as well as command control information such as types of requests (order, batch, status query).

4.6.6 Error Management Database

ICS Compatability:Explanatory

The error management data are the data required to provide a description of error conditions which may occur during processing. This includes diagnostic messages, level of priority of error or failure and the location of the error within the system to include attribute error, software error, or hardware error. CIP Diagnostic Messages are listed in Appendix E of the CIP Specification - Release B [R3].

4.6.7 User Management Database

ICS Compatability:Explanatory

User Management data includes the User Profile information including the data required for authentication and user privileges. Additionally, *Retrieval Manager* profile data such as locations are also included in the user management database. The following model represents the user informatiion that will be captured in the User Management Database.

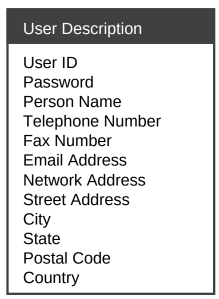


Figure 4-14. User Descriptor Model

4.7 IGP Document Data (Guide, Reference Papers etc.)

Although the Document protocol supports the insertion and retrieval of any type of document, the most common example from the Earth Observation (EO) perspective is the Guide Document. Guides are textual documents that are developed to assist the user in understanding the EO environment. Guide Documents provide additional detailed information about a specific mission, sensor, or organization. The most frequently covered guide topics include:

Data Centers - provides details about the center that holds the data.

Project/Campaign - describes projects and associated intensive field campaigns coordinated to collect data for a focused study.

Source/Platform - describes the source that held the instrument and sensors during data collection. Source is intended to communicate the data collection environment which includes satellite, aircraft, buoys, ships and ground station platforms. It also includes humans in the case of hand held sensors or human observations, paper or electronic surveys in the case of questionnaires, and computers in model analysis.

Sensor/Instrument - describes the instrument and its component sensor(s) that actually collected the data. This includes eye as a sensor, as well as, paper or electronic questionnaires used to collect data. It also includes computer model analysis that generate data.

Dataset - describes the dataset, collection, procedures, algorithms, and processing data.

Document Data will be captured in the local systems environment. The details for accessing and retrieving this information is specified in the ICS Guide Design and Protocol Specification [R24].

4.7.1 Guide Document Data Model

ICS Compatability: MAA

The following data model reflects the guide data for the Guide Indexer referenced in the ICS Guide Design and Protocol Specification [R24]. This model identifies each of the objects and further attributes that will be supported in addition to the relationships between the data objects.

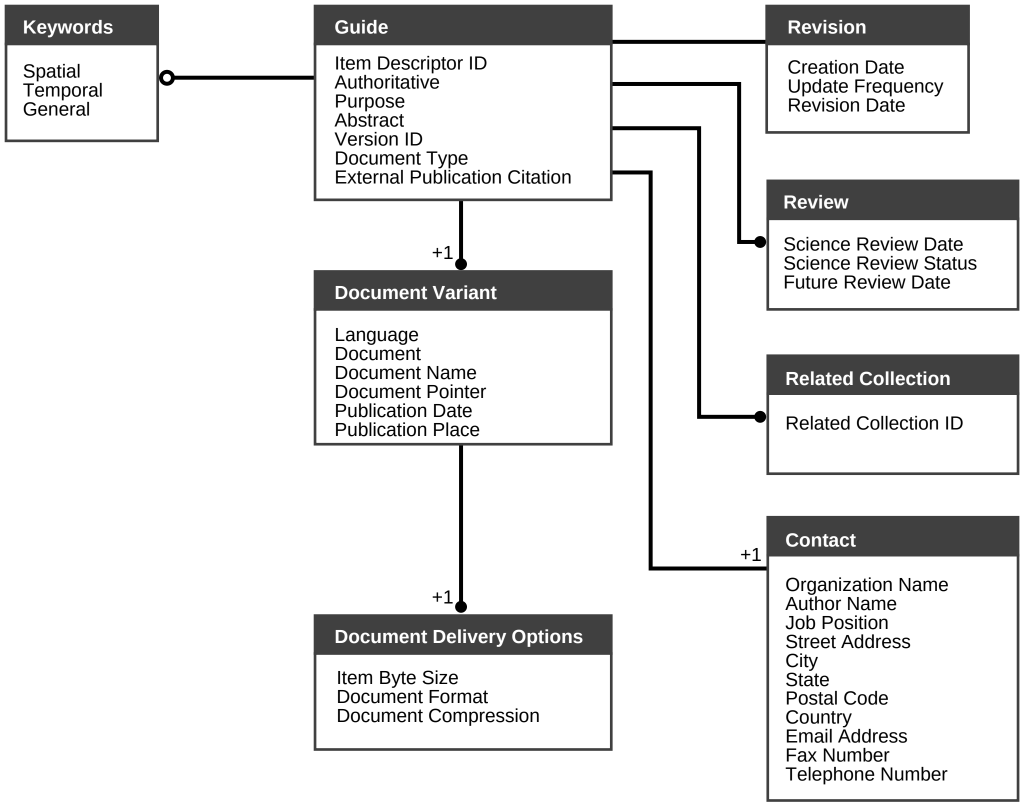


Figure 4-15. Guide Data Model

This page intentionally left blank.

5. COMMUNICATIONS VIEW

The Communication View provides two models of the ICS: 1) a communication framework showing the protocol stack used in ICS, and 2) a description of the CEOS Network and Internet usage between Retrieval Managers.

5.1 ICS Communication Framework

This section defines the communications stack to be used for implementing CIP over the networks between ICS *Retrieval Managers*.

5.1.1 ICS Utilization of TCP/IP

ICS Compatibility: Mandatory

ICS will be implemented using the TCP/IP set of protocols [R20 and R26]. Note that CIP is required to not be limited to TCP/IP (see requirement 336 in [R2]), allowing some CIP applications to be implemented over an OSI communication stack, for example. But, ICS shall exclusively be TCP/IP. In ICS, CIP and IGP will use TCP. ICS addressing will be according to Internet domain, host name and port id, format. The use of TCP services by CIP and IGP are defined in the following sections. An identical communication configuration is adopted by GEO Servers to allow interoperability.

5.1.2 TCP/IP Services

ICS Compatibility: Explanatory

This section provides an overview of the layering of CIP and HTTP over TCP/IP. This overview supports the discussion in the next sections on specific TCP calls by CIP and HTTP.

TCP/IP has 4 conceptual layers of software. Starting at the top is the application layer, supported by the Transport layer, supported by the Internet Layer, and the Network Interface Layer. A hardware layer supports this stack. (See [R12] for a more extensive review of TCP/IP.)

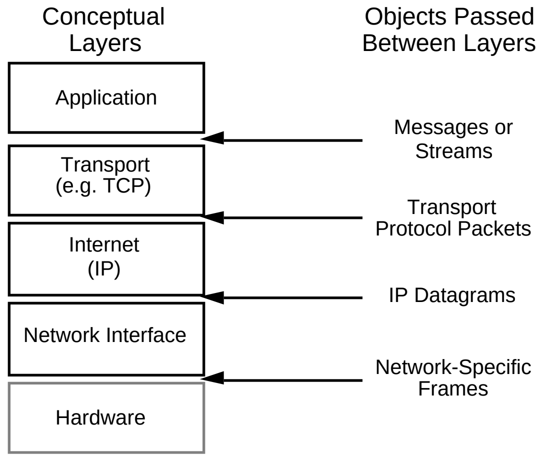


Figure 5-1. TCP/IP Internet Layering Model

The interface between the Application Layer and the Transport Layer for TCP is defined in [R26]. An Application must use the TCP Application Commands shown in Table 5-1 to send messages over the network. TCP accepts the commands listed in Table 5-1 from an application and must also return information to the application it serves. When an application wants to communicate with a remote application, TCP must establish a connection. A connection is a logical communication path identified by a pair of sockets, one at each machine. A socket is identified by an Internet address and a TCP port.

Table 5- 1. TCP Application Commands

Command	Description
Open	This command establishes a TCP connection for the application. A Passive open sets the connection to Listen for an incoming connection. An Active Open forms a connection with a specific remote socket.
Send	This command causes data to be sent on a connection.
Receive	This command allows data to be accepted from a connection.
Close	This command causes the connection specified to be closed.
Status	This command returns data to the application about the state of a connection.
Abort	This command causes all pending Sends and Receives to be aborted, the connection state is deleted, and message is sent to the TCP on the other side of the connection.

5.1.3 CIP Translators and TCP Communication Stack

ICS Compatibility: Explanatory

The implementation of the communication protocol stack for ICS is shown in Figure 5-2. CIP applications within the *Retrieval Manager* and *CIP Client* will create CIP messages which can be encoded using Z39.50. Z39.50 is an application level protocol and produces a byte stream which is passed to the Transport Layer using a TCP socket. TCP establishes a virtual circuit with the remote site and sends Transport Protocol Packets to the Internet layer. The Internet layer implemented with IP, passes IP datagrams to the Network Interface.

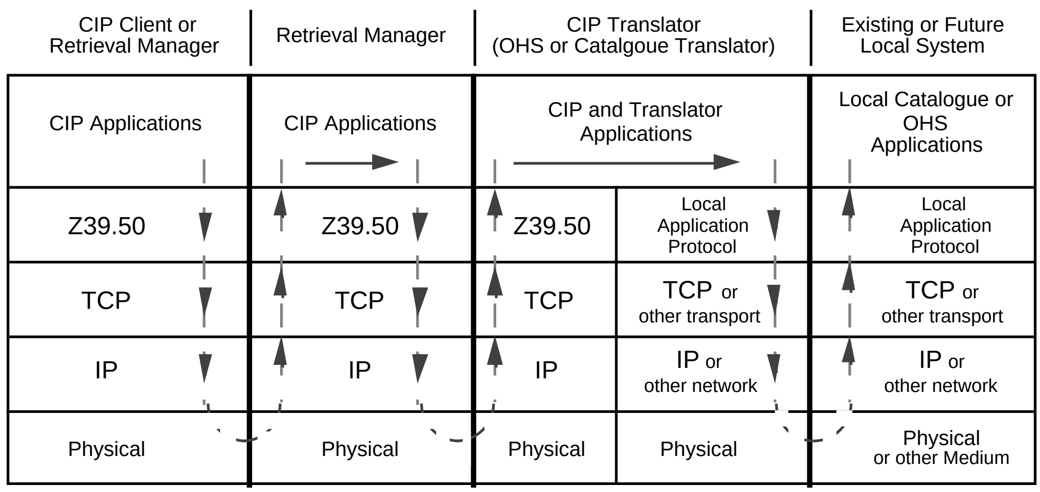


Figure 5-2. ICS Layered Communication Model

The specific example shown in Figure 5-2 shows a *CIP Client* (or a *Retrieval Manager*) establishing a *CIP Session* with a *Retrieval Manager*. Messages are passed down the protocol stack in the *CIP Client* with each layer encapsulating the message with its layer-specific information until they can be passed to the physical network to the other site. Note that there may be multiple devices in between the physical devices hosting the *Client* and the *Retrieval Manager*, e.g. routers. The messages are unpacked on the way up the protocol stack in the *Retrieval Manager*. In this particular example the message continues on from the *Retrieval Manager* to a CIP Protocol Gateway through a packing and unpacking cycle. In the application level of the CIP Protocol Gateway, the message may under go much translation to change the application semantics of the message from the CIP domain to the legacy system domain. The CIP Protocol Gateway then wraps the legacy application system message in the protocol specific to the legacy system communication stack. This allows the message to be passed from the Gateway to hosts within the local system, which in turn unwrap the message and answer the message at the application level.

5.1.3.1 Implementing CIP using TCP/IP

ICS Compatibility: Mandatory

This section defines the interface between CIP and TCP services. This section summarizes a memo by Clifford A. Lynch [R13] and extends the topics to CIP. (Lynch's memo was written for Z39.50 Version 2 and has not been updated for Version 3.) Table 5-2 provides the mapping of CIP messages to TCP Commands. When a CIP message is to be sent via TCP, the TCP commands listed in Table 5-2 are used.

Table 5-2. Mapping CIP Messages to TCP Commands

CIP Message	TCP Command
InitializeRequest	*Open* to establish a connection, then *Send* the *InitializeRequest* data using the open connection
InitializeResponse, SearchRequest, SearchResponse, PresentRequest, PresentResponse, SegmentRequest, DeleteResultSetRequest, DeleteResultSetResponse, AccessControlRequest, AccessControlResponse, ResourceControlRequest, ResourceControlResponse, TriggerResourceControlRequest, ResourceReportRequest, ResourceReportResponse, ExtendedServicesRequest, ExtendedServicesResponse	*Send* to an established connection
Close	*Send* the CIP *Close* data, then *Close* the TCP connection.

Connection. In the Internet environment, TCP Port 210 has been assigned to Z39.50 by the Internet Assigned Number Authority [R16]. To initiate a *CIP Session* with a *CIP Target* in the TCP/IP environment, a *CIP Origin opens* a TCP connection to port 210 on the *CIP Target* and then, as soon as the TCP connection is established, *sends* an *initializeRequest.* The TCP connection can be closed by either the *CIP Client* or the *Retrieval Manager* by sending a *close* message and then closing the TCP connection.

Encoding. The CIP specification and the Z39.50 standard specify application protocol data units (APDUs) in Abstract Syntax Notation One (ASN.1) [R14]. These APDUs include EXTERNAL references to other ASN.1 and non-ASN.1 objects such as those defining record transfer syntax to be used in a given application association. Standard Basic Encoding Rules (BER) [R15] are applied to the ASN.1 structures defined by the CIP profile and Z39.50 protocol to produce a byte stream that can be transmitted across a TCP/IP connection using the *Send* command.

As the approach above is based on a Z39.50 Implementor's Group agreement, the mapping of GEO messages to TCP commands is done the same as shown for the CIP messages in Table 5-2.

5.1.4 Distributed Session Management

ICS Compatibility: Mandatory

The need for Distributed Session Management when using CIP is illustrated in Figure 5-3. A user may request a search which results in sub-searches to other *Retrieval Managers*. To accomplish this, the *CIP Client* forms a *CIP Session* with a *Retrieval Manager*. In order to send the sub-query to the second *Retrieval Manager*, a second *CIP Session* between the two *Retrieval Managers* is established. In the first *Retrieval Manager*, the second *CIP Session* must be associated with the initial *CIP Session* between the *CIP Client* and the *Retrieval Manager*, so that results are returned appropriately to the client which initially requested the search. The set of *CIP Sessions* needed to achieve the users request is referred to as a *User Session*.

The GEO DOGS that is used to initiate a session with a *Retrieval Manager* maintains a connection for one user session and terminates the connection with a close. Hence, each time a GEO user requests *CIP domain* data, a new connection is established. Similarly, when a *Retrieval Manager* accesses the GEO Servers for data, a single Z39.50 (version 2) session is initiated for each user. At the end of the user session the connection is closed. All the GEO specific requests from the *CIP domain* are serviced via this interface.

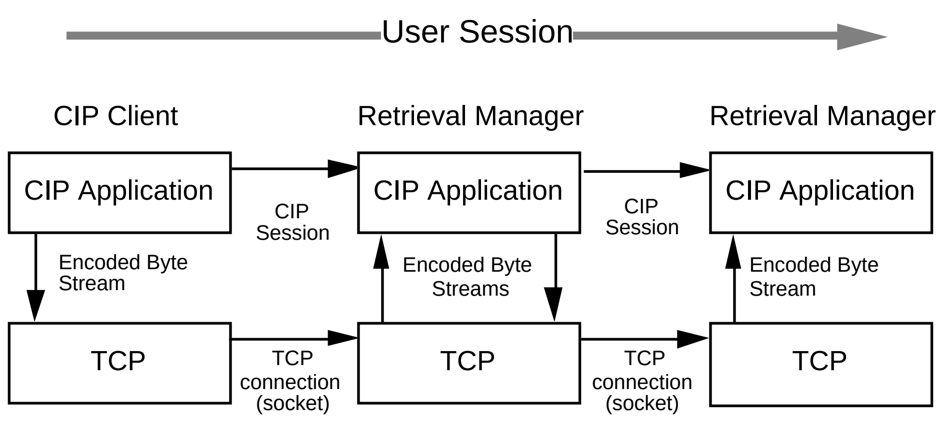


Figure 5-3. Distributed Sessions

The management of distributed sessions is accomplished in the following fashion.

- Individual *CIP Sessions* are based on Z39.50 associations. Z39.50 is a single client to single server protocol, i.e., it does not provide for distributed session management across several *Retrieval Managers*.
- CIP defines a unique, mandatory Reference ID for each message, which is used to manage message tracking.

- The *Retrieval Manager* is required to manage associations between incoming CIP messages and subsequent secondary messages to other ICS elements and to GEO Servers. This is accomplished using the session management logs defined in Section 4.

So, the management of distributed sessions is accomplished by the session logs which are established in the *Retrieval Managers*. A session log is established for a user during the initialization of a session. For each CIP request which the *Retrieval Manager* receives it maintains the association with any secondary request in the session log. When the secondary response is returned, the session log is consulted, the appropriate primary session is determined and the primary response is sent.

5.1.5 Implementing HTTP using TCP/IP

ICS Compatibility: Mandatory

IGP is implemented using the Hypertext Transfer Protocol (HTTP) Version 1.0 [R18]. HTTP is an application-level protocol for distributed, collaborative, hypermedia information systems. It is a generic, stateless, object-oriented protocol which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods. A feature of HTTP is the typing and negotiation of data representation, allowing systems to be built independently of the data being transferred

HTTP communication usually takes place over TCP/IP connections. The default port is TCP 80, but other ports can be used. This does not preclude HTTP from being implemented on top of any other protocol on the Internet, or on other networks. HTTP only presumes a reliable transport; any protocol that provides such guarantees can be used. Within ICS, IGP using HTTP is implemented over TCP/IP.

The HTTP/1.0 protocol is based on a request/response paradigm. A client establishes a connection with a server and sends a request to the server in the form of a request method, URI, and protocol version, followed by a MIME-like message containing request modifiers, client information, and possible body content. The server responds with a status line, including the message's protocol version and a success or error code, followed by a MIME-like message containing server information, entity meta-information, and possible body content. A separate TCP connection is established to fetch each URL. In this approach, the use of inline images and other associated data often requires a client to make multiple requests of the same server in a short amount of time.

5.1.6 Directory Services

ICS Compatibility: Mandatory

Key to a distributed three-tier architecture (Section 2) is the use of a directory service. As mentioned in Section 2, ICS uses a simple directory service for the conversion of high-level names to network address. Additional directory services are not currently used by ICS, e.g., to resource location service, Yellow Pages services, mail address lookup. Based on these requirements and based on the choice of TCP/IP, the Domain Name Service (DNS) is used in ICS.

To perform a search, a specific collection must be targeted in the search. The syntax which CIP requires for collection names follows a URL structure. The URL contains an Internet domain name for the Retrieval Manager on which the collection can be found. DNS contains information about the mapping of host and domain names, such as, "eos.nasa.gov", to IP addresses. DNS is maintained in a distributed fashion, with each DNS server providing name service for a limited number of domains. Also, secondary name servers can be identified for each domain, so that one unreachable network will not necessarily cut off name service.

An additional directory service is needed when a Retrieval Manager chooses to implement the Public Key Infrastructure for authentication. ICS uses the Lightweight Directory Access Protocol (LDAP) for interacting with a Certification Authority. LDAP is a TCP/IP version of the X.500 Directory Service which is defined by the ITU-T (formerly CCITT) [R23]. The ITU X.509 Directory - Authentication Framework defines authentication services are provided by the Directory to its users. The Directory can usefully be involved in meeting their needs for authentication and other security services because it is a natural place from which communicating parties can obtain authentication information of each other: knowledge which is the basis of authentication. The Directory is a natural place because it holds other information which is required for communication and obtained prior to communication taking place.

5.2 CEOS Network Connectivity

This section defines the connectivity between ICS *Retrieval Managers* provided by the CEOS Network. This section contains the following information:

- An overall CEOS Network Architecture
- Bandwidth for Browse Image Data

5.2.1 CEOS Network Architecture

ICS Compatibility: Explanatory

The CEOS Network Subgroup is developing a network for CEOS usage, called CEOSnet. An overview of how ICS uses CEOSnet is presented in this section. See [R8] for detailed information about the CEOSnet. The CEOS Network Subgroup maintains a WWW page at http://nic.nasa.gov/ceos-ns/index.html.

The CEOSnet architecture is shown in Figure 5-4. Within each participating agency or country, a *Retrieval Manager* is provided. It is the intent that the users in a country use that country's national Internet (or other national network resources) to access its local *Retrieval Manager*. To satisfy user requests, a *Retrieval Manager* may then need to access data from another participating *Retrieval Manager*. Two network alternatives are available to provide this access. One alternative is to use the worldwide Internet. While the worldwide Internet is quite ubiquitous, its performance is not dependably high.

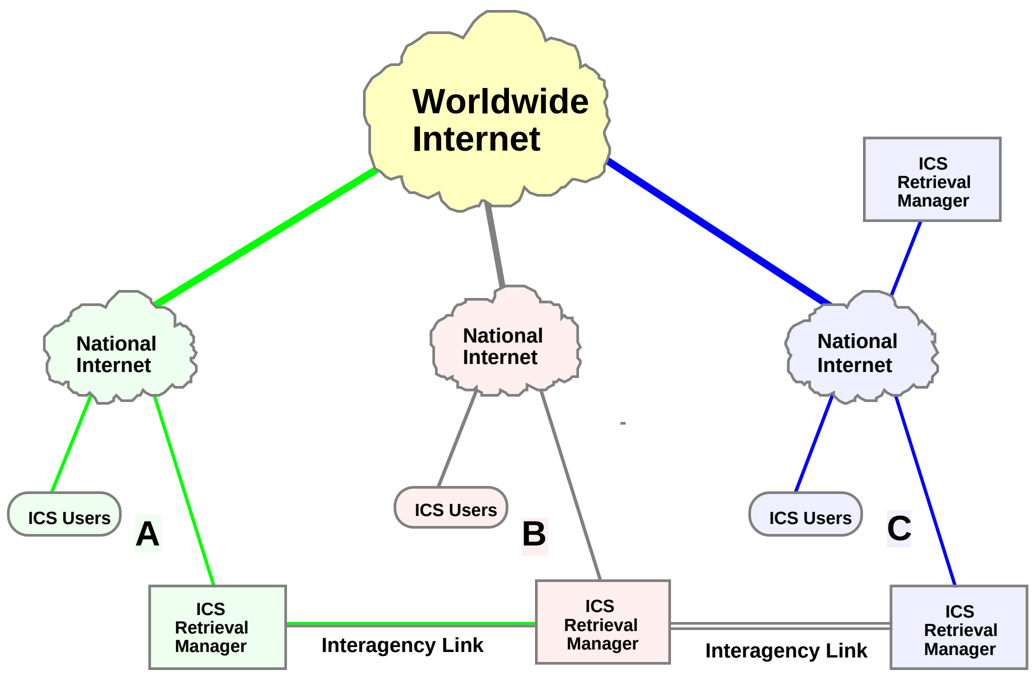


Figure 5-4. ICS Networks Model

On an individual, bi-lateral basis, the participating CEOS members may choose to implement private circuits between their facilities. These private circuits can be made available for access to *Retrieval Managers*, and may or may not also provide connectivity for other bilateral services. Collectively, these private circuits between participating CEOS organizations constitute the CEOS Network (CEOSnet).

CEOSnet is a limited access network coordinated and maintained by CEOS agencies, affiliates, and observers to support CEOS tasks and activities by providing access to and sharing of global Earth observation data and information. Access on or through the CEOSnet resources is only authorized when that access is in conjunction with CEOS approved activities and programs, the funding agencies for the CEOS Network links approve or agree to the use, and capacity on the CEOSnet links is adequate to support the use. Additional uses of CEOSnet will be allowed as approved by the CEOS agencies, affiliates and observers providing the CEOS Network resources. CEOSnet is not to be used for commercial gain or profit. (See [R27] for details on CEOSnet acceptable use policy.)

Note that the CEOSnet is thus not a separate network consisting of a distinct set of circuits and equipment, but instead is a logical network of components provided by participants and used for CEOS purposes.

The nominal data flows to satisfy a user query is:

1) User accesses local *Retrieval Manager* via national (or agency) network resources
2) *Retrieval Manager* contacts cooperating *Retrieval Managers* via CEOS net (preferred) and/or the worldwide Internet as appropriate.
3) Cooperating *Retrieval Managers* access local database resources to formulate response.
4) Responses are returned to originating *Retrieval Manager* via same network which carried the request.
5) Originating *Retrieval Manager* collects the responses, and delivers them to the user via the national or agency infrastructure.

Note that it is not essential for an organization to have dedicated circuits to have a participating *Retrieval Manager*. Subject to performance limitations, the worldwide Internet provides the connectivity required. In Figure 5-4, dedicated links are shown (as an example only) between agencies A and B, and between B and C. In this example there is no dedicated link between A and C.

Note that technically in Figure 5-4, it is possible for the *Retrieval Manager* at A to access C via the pair of dedicated links (A-B and B-C). However, policy issues must be resolved by each agency for this to be allowed. Otherwise, the worldwide Internet can provide connectivity between A and C.

The *ICS Client* has network options for accessing CIP items on remote *Retrieval Managers.* For example, assume that a user connected by a national Internet to *Retrieval Manager* B in Figure 5-4, performs a distributed collection search which matches a collection on *Retrieval Manager* C. If the user wishes to perform a product search on the *Retrieval Manager* C collection, the *ICS Client* could send the search either to *Retrieval Manager* B which would forward the request or the client could send the request directly to *Retrieval Manager* C. (Information in the collection identifier supports the search request in either path.) By sending the search to *Retrieval Manager* B, the interagency link between B and C is used which will probably outperform the path from B to C as it involves the Worldwide Internet. In general, *ICS Clients* will want to make use of the Interagency Links between *Retrieval Managers.*

The TCP/IP protocols will be used for communications between ICS *Retrieval Managers*. Interconnected routers in the CEOS Network must support network protocols which are robust and consistent with existent Internet protocols and standards (see [R20] for applicable standards). All interconnecting networks attaching to CEOS Network are at least IP-based relays to other networks or user sites. Additional network protocols or enhancement features are allowed but should not affect the network availability, performance, or interconnectivity.

5.2.2 Bandwidth Considerations

ICS Compatibility: Explanatory

A driver in the amount of CEOS Network bandwidth required to support ICS users is the retrieval of browse images. The CEOS Browse Task Team has begun to develop several parameters needed to support an estimate of the CEOS Network bandwidth necessary to support ICS user's request for browse images. Estimates for the browse image data are based upon the statistics from the existing guide data and the anticipated user behavior.

First, a bandwidth estimate for the browse data is presented. The first parameter in the Browse bandwidth estimation is the browse data size after compressed (to transfer via WAN). A maximum size is 1MB for EOS-HDF browse size with other agencies browse size is less than about 200 KB. So, an average browse size is estimated to be 500 KB.

The next parameter to be estimated is the user's expected time for transfer of a browse image. Current performance for DLR and the University of Rhode Island provides image retrieval from 45 to 60 seconds per browse data. The Browse Task Team suggests that a desirable requirement is 30 seconds.

The next parameter for estimating the bandwidth for a given ICS node is the frequency of browse retrieval requests. Given that CIP is a session oriented protocol, one approach to estimating the number of browse requests is to estimate the number of concurrent sessions. The number of sessions will have a diurnal variation, but only the peak number of sessions and the related peak number of requests are important for estimating the needed bandwidth. Also as the purpose of the worldwide ICS is realized, the diurnal variation of requests at a *Retrieval Manager* will be reduced as more users are retrieving data from around the globe. The ICS URD requires the *Retrieval Manager* be designed to be capable of supporting a peak load of a minimum of 100 concurrent user interaction sessions. It is expected that all ICS data providers should support a minimum of 30 simultaneous sessions for a *Retrieval Manager*, this includes remote sessions to remote *Retrieval Managers*. The CEOS Browse Task Team estimates an average of 10 sessions. This number is 1/3 minimum multiple sessions requirement for ICS URD as some sessions of CIP are used for catalogue retrieving and ordering.

If we assume that this 1/3 fraction of users is continually executing browse requests these numbers provide the following network bandwidth requirement:

500 KB * 10 sessions / 30 sec = 1333 Kbps

We use a similar method of bandwidth estimation for the guide data. The existing browse data has an average size of 40 KB. Using this basis and the method above we can see that the network bandwidth requirements for the guide data would compute to:

40 KB * 10 sessions / 30 sec = 106 Kbps

The combined bandwidth for the browse image and guide data will be:

106 Kbps + 1333 Kbps = 1440 Kbps ~ 1.5 Mbps (T1)

Of course, this estimate will change as the parameters are estimated by each agency. So, this bandwidth requirement is only a guideline.

This page intentionally left blank.

6. SECURITY VIEW

This section provides the security architecture for ICS. This section contains three parts. First, an assessment of the need for security in ICS is presented. This motivates the second part which presents the system design for a secure ICS.

6.1 ICS Security Assessment

The purpose of the ICS Security Assessment is to explain the need for appropriate security risk mitigation measures addressed in Section 6.2.

6.1.1 ICS Security Needs Overview

ICS Compatability: Explanatory

ICS is an international catalogue interoperablity program accessible to users on a global basis. Built on an open-computing network architecture to facilitate access by a wide variety of users, the driving requirements of its security architecture are integrity, availability, and confidentiality of its data assets. The project vision is open access to well-pedigreed data in which users may have confidence, while maintaining certain proprietary and administrative data in confidence.

Integrity. Because of the scientific nature of the project, integrity of project information is critical. The accuracy of the collection, product, guide and browse data must be stored and accessed in a fashion which maintains the integrity of the data. In addition, information to support ordering of ICS products and services must be protected against unauthorized access.

Users expect the system to guard against tampering with the source materials of the *Retrieval Managers* not only during storage, but during maintenance and distribution. Integrity threats are manifested through unauthorized access or use, leading to change, alteration or modification of information resources. The security architecture must provide adequate safeguards against these threats.

Availability. The ability to have assured use of project resources is vital to instill confidence in its users. Availability needs translate into two categories: fault tolerant features to preclude failure or operational disruption; and recovery actions, to enable timely resumption of operational activities and minimize the length of the disruption. Availability threats are manifested through denial of service events, either physical in nature (e.g., fire or loss of power) or logical (e.g., computer virus or other intrusive software).

Confidentiality. Confidentiality requirements exist because some of the information within ICS requires special protection. This includes specific user product requests and account information, and information of a private nature on individual users on the system. Information of this nature, if compromised, could result in damage or harm to ICS or to individuals. Confidentiality threats are manifested through access by unauthorized persons and authorized persons who have exceeded their privileges, e.g., unauthorized interrogation of user profile data.

6.1.2 Vulnerabilities

ICS Compatability: Explanatory

A vulnerability is a weakness in security procedures or controls that could be exploited by a threat. Vulnerabilities are often analyzed in terms of missing safeguards. Vulnerabilities contribute to risk because they may "allow" a threat to harm the system. Section 6-2, links the ICS security controls that counter the threats which may potentially exploit ICS vulnerabilities. There are seven categories of vulnerabilities which may impact ICS.

Software Vulnerabilities. Software vulnerabilities include: inadequate configuration management that permits program errors; unauthorized automated routines; and inadequacies in system and application software that may result in processing or calculation errors, or may allow unauthorized access to hardware, data, or programs. Given the collaborative nature of the software development of CEOS *Retrieval Managers*, these vulnerabilities may occur due to unintentional miscommunication.

Hardware Vulnerabilities. Hardware weaknesses include: improper operation of hardware; lack of proper hardware maintenance; inadequate physical security; and inadequate protection against natural disaster. Because all ICS hardware will be procured, installed and maintained by procedures outside of ICS, it is assumed that hardware security measures will be followed at each site. The availability of the *Retrieval Manager* at each site is contingent on this assumption. The ICS security design will need to protect against breaches of confidentially and integrity to other sites independent of hardware failures at a given site.

Data Vulnerabilities. Data vulnerabilities include inadequate access control that permits unauthorized access or authorized personnel to exceed privileges with the potential result of both accidental and malicious deletion, corruption, modification, or destruction of data, as well as theft. Of special concern to ICS are susceptibilities that impact the integrity of collections data, browse and guide data held by the *Retrieval Manager*, system configuration data, registration data, authentication data, and ordering data.

Administrative Vulnerabilities. Administrative vulnerabilities are associated with weaknesses in the effective administrative control of IT resources. They include inadequate or nonexistent administrative and security policies, guidelines, training, and controls; operating procedures (i.e., standard operating practices and procedures); management constraints, and accountability. As the administration of ICS is distributed throughout its agency members, it will be assumed that the availability of each site's *Retrieval Manager* will be dependent on the site's administrative practices. Some administrative practices will be defined by the PTT, e.g. Collection Manual [R5], but their application is dependent solely on the site's personnel. ICS must protect from losses of integrity or confidentiality from a lapse in a single site's administration.

Communications Vulnerabilities. Vulnerabilities associated with communications include: inadequate access control that allows unauthorized access to networks and communications circuits that could result in transmission interception and unauthorized access to network components; and inadequate measures to prevent circuit failure from both natural disaster and human activities, intentional and accidental, resulting in denial of service. ICS is dependent upon the CEOS Network as defined by the CEOS Network Sub-group. Individual site communication security is dependent upon routers for their sites.

Personnel Vulnerabilities. As used here, the term personnel means people who have an authorized association with ICS resources or facilities, such as: employees, certain guests and maintenance personnel, and authorized system users. This group of people is often referred to as "insiders". Insiders represent the greatest weakness in any system, including ICS, because they already have access, usually understand the system configuration and operation, and may be aware of existing vulnerabilities. Security weaknesses associated with insiders include inadequate physical and logical controls that allow an insider access to systems beyond which she or he has privileges; and inadequate administrative procedures or controls to minimize or detect accidents involving IT resources or IT resource theft, abuse, misuse, damage or destruction. Perhaps the greatest weakness involving insiders is that they are exposed to external influences and pressures that may provoke malicious acts against IT resources, such as destruction or theft.

Facility Vulnerabilities. Facility weaknesses include: inadequate physical security that permits accessibility by unauthorized persons which could lead to facility, and content, misuse, damage, or destruction, or theft of its contents; and inadequate protection against natural disaster that may result in the damage or destruction of the facility or its contents. Furthermore, poor facility maintenance and services, such as poor housekeeping, poor air quality, temperature extremes, and power fluctuations may result in damage to or destruction of IT resources.

As the ICS resources are distributed throughout its agency members, it will be assumed that the availability of each site's *Retrieval Manager* will be dependent on the site's facility practices. ICS must protect from losses of integrity or confidentiality from a lapse in a single site's facility.

6.1.3 Threats

ICS Compatability: Explanatory

ICS is concerned with threats that exploit the above vulnerabilities and have a detrimental impact on the integrity, availability, and confidentiality of its IT resources. Generally, threats to ICS resources come from two major sources: natural disaster and human activity.

- Natural disaster includes: airborne particles, cataclysm (earthquake, volcanic eruption, tidal wave, etc.) , fire, static electricity, and weather. Note that these threats will exploit vulnerabilities at specific sites affecting availability over which the ICS system design has no authority. But ICS must preclude lapses in confidentiality and integrity to other sites given a natural disaster, i.e. contain any threat to the single site.

- Human activity includes activity from both authorized persons and unauthorized persons.
  - Authorized persons are users, employees, and maintenance personnel who have some level of authorization to use or have access to ICS resources. Threats resulting from authorized activity may be accidental (an incident without malice) and intentional (a malicious act). This may include otherwise authorized persons who exceed their authority. This may also include errors or omissions in the software development or the intentional or accidental inclusion of malicious code, e.g. viruses.
  - Unauthorized persons are users or persons who do not have authorization to use or have access to ICS resources. Even though in theory, activity by unauthorized persons can be accidental, this section treats all such activity as intentional.

6.1.4 ICS Security Definitions

ICS Compatability: Explanatory

The following definitions are used in the ICS with respect to security concepts.

Authentication: Verification of the identity of a user or, validation of a communication (the second part provides for non-user based authentication, e.g. between *Retrieval Managers*).

Authorization: Permission, granted by a properly appointed person or persons, to perform some action.

Confidentiality: The protection of information from disclosure to those not intended to receive it.

Data Integrity: The assurance that data received is the same as data generated.

Domain: A system or portion of a system which has the same security policies and requirements. Individual agencies determine the boundaries of domains.

Non-repudiation: The ability of the receiver to prove that the sender of some data or of a request did in fact send the data even though the sender might later desire to deny ever having sent that data.

Proxy: A software agent that acts on behalf of a user.

Registration: The process whereby an individual submits required personal information to an agency and, in return, the agency provides the means (e.g. login name and password) necessary to perform authentication with the agency's system.

6.2 ICS Secure System Design

6.2.1 Overview of Secure System Design

ICS Compatability: Explanatory

The ICS security controls are divided into three groups: administrative, physical and computing.

- Administrative security controls are policies, guidelines, and practices and procedures designed to manage and implement security.
- Physical security controls are physical barriers or devices designed to prevent harm to or loss of IT resources and assets, such as access control card readers, intrusion detection systems, and fire suppression systems.
- Computing security controls (sometimes called technical security controls) are software mechanisms designed to prevent harm to or loss of data and information.

Table 6-1 maps the vulnerability categories, discussed above, against the security control categories. The remaining sections in this chapter describe the specific security controls in each control category of administrative, physical and computing.

Table 6-1. ICS Vulnerabilities versus Security Controls

Vulnerabilities	Security Control Category		
	Administrative	**Physical**	**Computing**
Software Vulnerabilities.	CEOS ICS Software CM ICS Event handling Security Testing	Site facility protections *	Standards on RM development Fault Handling
Hardware Vulnerabilities.	Site hardware administration *	Site facility protections *	RM Response to Unavailability of Remote RM
Data Vulnerabilities.	ISA back up procedures Authentication Information Management	Physical security of hardware *	Access control - users Access control - ISA RM DBMS data integrity functions Time Out Features Tamper Proofing Encryption
Administrative Vulnerabilities.	Collection Manual ICS Administration Manual ISA Training ICS Event handling System Rules for Users CEOS Authorization		RM Administration Independence Display System Rules for Users RM Activity Logs
Communication Vulnerabilities.		Site disaster prevention Physical security of hardware *	Network security
Personnel Vulnerabilities.	ISA Training Site personnel practices *		RM Administration Independence
Facility Vulnerabilities.		Site physical security * Site maintenance *	

* Site security controls are assumed to be in place for the sites where *Retrieval Managers* will be installed.

6.2.2 Administrative Security Controls

ICS Compatability: MAA

Administrative security controls include security policy, and other items and activities that are designed to manage and implement security policy. They should provide security guidance to users and ISAs who have some level of authorization to use or have access to ICS resources. This section defines the security controls which are listed in the Administrative Security Column of Table 6-1.

CEOS ICS Software CM. To support the establishment of *Retrieval Managers*, ICS software will be made available for reuse. This will be accomplished using a configuration controlled access point for the distribution of ICS software. Configuration Management (CM), from a security point of view, provides assurance that the software which is available is the correct version (configuration). Configuration management can be used to help ensure that changes take place in an identifiable and controlled environment and that they do not unintentionally harm any of the system's properties, including security. Changes to the software can have security implications because they may introduce or remove vulnerabilities. The ICS CM software distribution site, e.g., an ftp site, must not allow unauthorized modification of ICS code. Once an organization pulls code from an ICS reuse library, the ICS CM is longer in effect. It is possible for an organization to cause unintended erroneous action in ICS reuse code by modification outside of ICS. This will not threaten ICS due to the domain independence of the ICS elements, e.g., access to one *Retrieval Manager* does not automatically provide access to all *Retrieval Managers*.

ICS Event handling Because the ICS operations depends on the loosely associated ISAs, procedures for dealing with events in the ICS are defined in advance. ICS events include: adding a new *Retrieval Manager*, alerting ICS to detection of an intrusion at a *Retrieval Manager*, alert of a *Retrieval Manager* being off-line either due to a planned or unplanned cause. The main communication amongst ISAs will be via an ISA e-mail list. Although for security alerts, communications must be by a separate channel than e-mail, e.g., phone, as an e-mail alert may be intercepted. The procedures for ICS event handling will be detailed in the ICS Administration Manual.

Security Testing Security testing is conducted to ensure that the security features meet technical specifications and to locate vulnerabilities. Examples of security testing tools are: Security Administrator Tool for Analyzing Networks (SATAN) and Internet Scanner, a product of Internet Security Systems, Inc. (ISS). These tools are designed to discover weaknesses or holes in a UNIX based network and recommend fixes. Procedures for ICS security testing will be detailed in the ICS Administration Manual.

ISA Training In order to insure consistent ICS operations and adherence to procedures with security implications, ICS Training will be conducted for ISAs. This training will cover, at a minimum, the material in the ICS Administration Manual.

System Rules for Users ICS users cannot be expected to act responsibly with respect to ICS operations, unless they are aware of the system rules for users. These rules should clearly delineate responsibilities of and expectations for all individuals with access to the system. Often rules should reflect logical security controls in the system. For example, rules regarding authentication should be consistent with technical features in the system.

Management Authorization The authorization of a system to process information, granted by a management official, provides an important quality control. By authorizing processing in a system, a manager accepts the risk associated with it. Management authorization should be based on an assessment of management, operational, and technical controls. Since the SDD establishes the security controls, it should form the basis for management authorization, supplemented by more specific studies as needed. In addition, periodic review of controls should also contribute to future authorizations. Re-authorization should occur prior to a significant change in processing, but at least every three years. It is important to identify the appropriate management authorization for ICS. ICS is a CEOS activity and has the context defined by its interfaces (see ICS Context Diagram in Section 3). For ICS, the authorization must come from both the appropriate CEOS organization as well as individually by the agencies which host *Retrieval Managers*. The ICS is planned and authorized by the Access Sub-Group. Each agency operating a *Retrieval Manager* is represented in the Access Sub-Group.

6.2.3 Physical Security Control

ICS Compatability: Explanatory

Physical security controls are designed to guard against threats that result from both natural disaster (e.g., storms and resulting power outages) and human activity (e.g., fire, theft of hardware and software, physical access to a facility by unauthorized persons). All physical security controls listed in Table 6-1 are the responsibility of the sites which host a *Retrieval Manager* and therefore are not under the control of the ICS SDD.

ICS as a system is robust to a failure of physical security control at a single site, i.e., a security failure due to physical controls may cause a loss at the site but cannot result in a loss at another ICS site. Each *Retrieval Manager* is independent from a security authentication perspective such that the integrity of data is not threaten by a lapse in physical security at another site. From availability consideration, the loss of a particular *Retrieval Manager* could cause disruptions in the operations of ICS. The event handling measures described in Section 6.2.1 and the System Management in Section 7, provide for the response procedures which would be initiated in case of a *Retrieval Manager* failure due to a lapse in physical security control.

6.2.4 Computing Security Controls

A summary of Computing Security Controls is provided in the first part of this section. The remainder of the section describes the authentication in ICS and the Group Security Model.

6.2.4.1 Summary of Computing Security Controls

ICS Compatability: MAA

Computing security controls are software and firmware mechanisms used to limit access, detect intrusion, detect malicious logic and prevent its propagation, etc. This section defines the computing controls which are listed in the Computing Security Column of Table 6-1.

The following Computing Security Controls are required to be implemented in the *Retrieval Manager* based on requirements in the ICS URD [R2], Sections 3.1.2 and 3.2:

- Standards on *Retrieval Manager* software development
- *Retrieval Manager* Fault Handling
- Access Control (See also SDD, Section 6.2.4.2)
- *Retrieval Manager* DBMS data integrity functions
- Session Time Out control
- RM Administration Independence

The following Computing Security Controls are required to meet the ICS security design. Requirements will need to be added to the ICS URD [R2] to insure the ICS elements are compatible:

- Retrieval Manager Response to Unavailability of a Remote Retrieval Manager
- Display System Rules for Users

It is important to note that Network Security is covered as a CEOS Network Sub-Group topic. Network security covers the issues outside of the application level *Retrieval Manager* Security functions, e.g., IP address blocking in a communication router.

6.2.4.2 Authentication Mechanism

ICS Compatability: MAA

This section describes several topics which introduce the ICS design for authentication. First the two mechanisms provided by CIP are described, then two scenarios are described for a *target* authenticating an *origin*. This section address the application layer provision for authentication which CIP provides. It is also important to note that *Retrieval Managers* make use of IP address as a basic authentication of other *Retrieval Managers*. . In addition, the IP address of the clients can also be used for user authentication and validation. This concept is discussed later in this section.

To meet the ICS security requirements in the ICS URD [R2], a comprehensive approach to network security based on a well developed cryptographic mechanism is needed. Authentication occurs in the context of a *CIP Session.* A *CIP session* is composed of multiple CIP *operations.* A CIP *operation* consists of several messages. A CIP session is begun with an *initializeRequest* and ends with a *close*. The authentication protocol described below allows authenticated sessions as well as authentication for any specific operation.

Three mechanisms are provided in CIP for authentication: unencrypted symmetric key, secure symmetric key, and asymmetric key. IP address authentication can be used to augment and simplify the authentication mechanism. In the symmetric cases, both the *target* and the *origin* are holding the same key, e.g., a username/password. In the asymmetric case, the *origin* and *target* are holding different but related keys, e.g., the *origin* holds a private key and the *target* holds a public key. Both the secure symmetric key approach and the asymmetric key approach use a digital signature as the means to authenticate the *origin.* The digital signature contains information which could have only been constructed with the user's key. The IP address authentication is not based on any type of key based mechanism and is used only for clients that have a fixed address. This mechanism can simplify the authentication by using the encapsulated information of the TCP/IP wrapper to automatically authenticate the users.

The unencrypted symmetric key approach uses a username and password in the clear, that is, there is no confidentiality of the password as it is sent from the *Client* to the *Retrieval Manager*. This mechanism is the appropriate choice when minimizing the complexity of the *Client* is a driving requirement and the risk associated with unprotected transmission of the password is minimal. Note that this approach cannot be used between *ICS Retrieval Managers.*

In ICS, the secure symmetric key mechanism is the default approach for authentication. If a *Retrieval Manager* provides authentication, the secure symmetric approach must be provided at a minimum. To provide for an agency's specific needs, a *Retrieval Manager* may choose to provide asymmetric key or unencrypted symmetric key authentication. This approach is driven by the need to comply with laws regarding export of encryption algorithms, i.e. some countries restrict the public key approach for strong authentication.

The basic element of the symmetric key approach is a Message Authentication Code (MAC). A MAC is a key-dependent, one-way hash function. A secret, shared key (password) is required to form the hash and the hash cannot be decoded. Only someone with the identical key can verify the hash by performing the same hash operation and verifying that the result is identical. MACs are useful to provide authentication without privacy. The protocol does not use privacy as a basis, i.e., encryption is not used. The MAC for CIP is calculated using an MD5 hash. The MAC approach relies on a shared key between the *origin* and *target*. In this protocol the secret, shared key is a username/password which is particular to the user. How the *target* got the username/password is addressed in Section 7.

The asymmetric key approach is based on a set of related key pairs (public, private) used for the encryption and decryption of messages. A “digital signature” is obtained by encrypting a message hash combined with a timestamp with a private key. Such a message can be authenticated by a decryption based on the corresponding public key. In the CIP context, a client holds the private key and a Retrieval Manager holds (or has access to via a Certification Authority) the corresponding public key.

The IP address authentication is based on the nature of the Internet protocol. In most instances, the client that connects to the *Retrieval Manager* will have a fixed IP address and pass the address to the Retrieval Manager as part of their TCP/IP connection (the approach does not work for the clients systems that dynamically assign IP addresses). Once authenticated by a password mechanism, the IP address can be stored in the system area associated with the user. Thereafter, the IP address itself is used to automatically authenticate the user and set the user privileges without requesting the user for the password and going through a login procedure. This approach does assume that the user workstation is secure and only the authorized user uses the system to connect to the Retrieval Managers. If that is not the case, this approach is not recommended.

Use of CIP for authentication is shown below in two scenarios. The first scenario shows use of part of the protocol for authentication for a specific operation. The second scenario shows how an authenticated session is established during initialization of the session.

- **Authentication for an Operation**

This section addresses how a user would be challenged for authentication credentials based on a request for a CIP service, e.g., placing an order. It is assumed that the user’s session is not an authenticated session.

The messages for this authentication are shown in Figure 6-1. Specific contents of the messages are provided in the CIP Specification [R3].

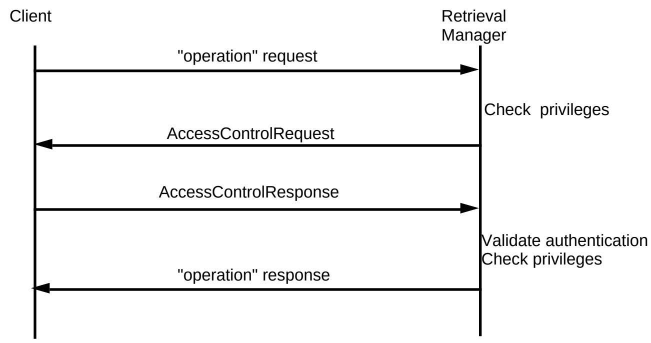


Figure 6-1. Authentication for an Operation

If Non-Repudiation is requested by the *Retrieval Manager*, another pair of messages would be needed in between *presentCredentials* and "operation" response. The first message would be from the *Retrieval Manager* to request a non-repudiated order from the client. The Client would reply with a non-repudiatable order message.

- **Authentication for a Session**

This section addresses how a user would begin a session with the intent to have an authenticated session. The authentication is a two step process. First there is a two step authentication followed by a negotiation of cryptographic options including use of a session key.

The messages for this authentication are shown in Figure 6-2. Specific contents of the messages are provided in the CIP Specification [R3].

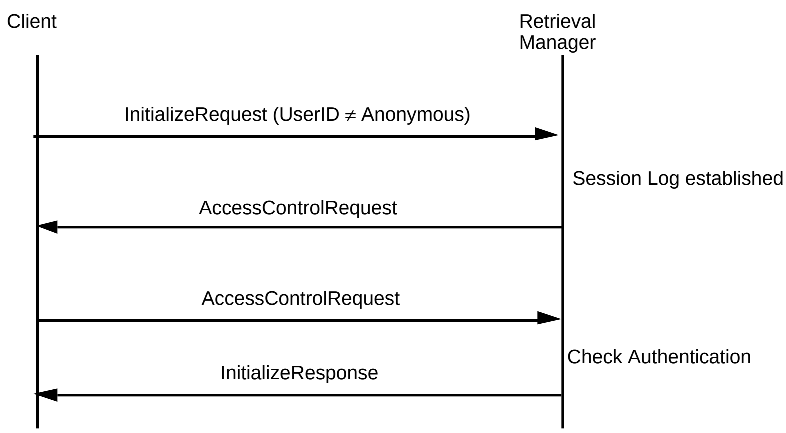


Figure 6-2. Authentication for a Session

6.2.4.3 Group Security Model

ICS Compatability: MAA

When security is considered in ICS - a distributed information system - the proliferation of credentials must be considered. The ICS has been designed to prevent a world-wide proliferation of usernames/passwords. The approach is to have the *Retrieval Managers* act as brokers for the users which they support. This avoids the user needing to be known at each *Retrieval Manager*. The *Retrieval Managers* serve as brokers based on two relationships: 1) a *Retrieval Manager* will support many users and 2) a *Retrieval Manager* is known to other *Retrieval Managers*.

The group security approach applies to any distributed session. The approach is most important when applied to ordering as ordering will require the highest security considerations. Figure 6-3 shows the focusing of orders by a home *Retrieval Manager*. “Home” in this case means that the user is registered at the *Retrieval Manager,* and for this example the user belongs to a group that has privileges to allow the order. Figure 6-3 shows multiple users sending order messages to the home *Retrieval Manager*. The home *Retrieval Manager* in turn, creates a secondary order message to the remote *Retrieval Manager* which holds the data. Each primary order from a user results in a secondary order between *Retrieval Managers*. The secondary orders are allowed based on the authentication between *Retrieval Managers* and the group membership of the local *Retrieval Manager*. The home *Retrieval Manager* maintains a cross reference of the primary order with the secondary order which was submitted to the remote *Retrieval Manager*.

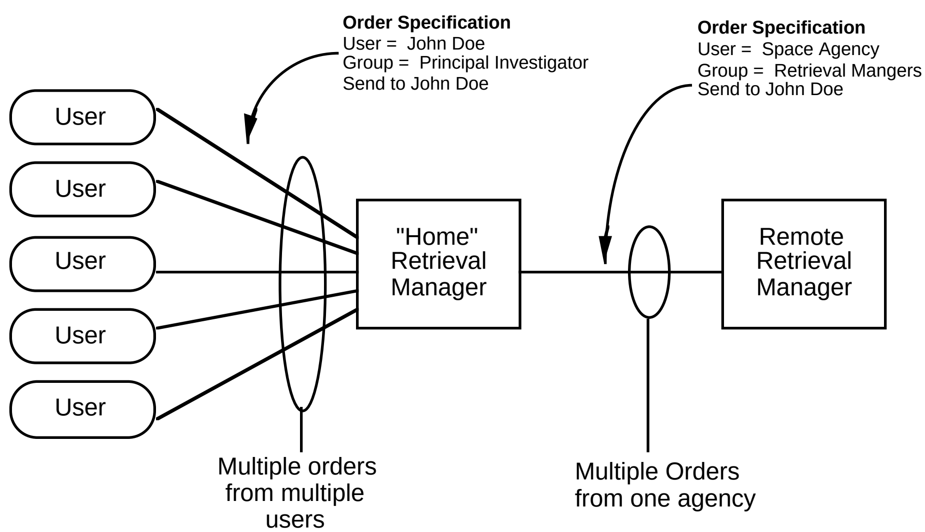


Figure 6-3. Group Ordering Model

The ordering described in Figure 6-3 and the associated text is the Order by Proxy approach. That is the *Retrieval Manager* is acting as the user's proxy to order the data on the users behalf. In the Order by Proxy case there is an agreement between agencies whereby one agency guarantees payment to another agency.

ICS also allows a second case labeled the Passthrough case. In the Passthrough Case there is a mechanism in the CIP to allow pass-through of information needed by an agency to perform its own authentication and authorization. There may be cases where agreements outlined in the Order by Proxy Case cannot be reached between agencies. Passthrough is a different method for ordering data that will provide convenience to the user. If a user who has a session established with the local *Retrieval Manager* (in Figure 6-3) wishes to order data from the Remote *Retrieval Manager* and is registered with a remote *Retrieval Manager*, the CIP can pass information (e.g. username/password) in a secure manner through the local *Retrieval Manager* to the remote *Retrieval Manager*. The remote *Retrieval Manager* then performs authentication and authorization for the user.

A scenario demonstrating the use of the Group Ordering design is shown in the next section.

6.2.4.4 Secure Indirect Ordering

ICS Compatability: MAA

This section demonstrates use of the CIP authentication mechanism to perform secure indirect ordering. Critical to secure Indirect Ordering is determining if the order should be a proxy or passthrough order. Indirect Ordering was introduced in Section 3.4.2 without discussing the security necessary if the ordered items require privileges based on authentication.

Indirect ordering involves an order submitted to a Local *Retrieval Manager* which must be forwarded to a Remote *Retrieval Manager* as the OHS which will process the order is accessible from the Remote *Retrieval Manager*. The CIP messages between the ICS elements for this scenario are shown in Figure 6-4.

Several assumptions are made concerning this scenario:

- The items specified in the order specification require privileges based on authentication, e.g., access to the data is restricted or access has a cost.
- The privileges of a user who is registered at the Remote *Retrieval Manager* will be better than those granted to the Local *Retrieval Manager* using a proxy order account, so passthrough is preferred if it is possible.
- Prior to the first step in the scenario, the User did a distributed search which necessitated initialization of a session between the Client and the Local *Retrieval Manager* and a session between the Local *Retrieval Manager* and the Remote *Retrieval Manager*. These initializations were done without authentication.

- The Local *Retrieval Manager* will support proxy ordering for this user at the Remote *Retrieval Manager* site. The scenario indicates where the Local *Retrieval Manager* can reject a proxy order if this assumption is not true.

The scenario demonstrates how the ICS RMs use CIP messages to determine if the order should be a passthrough or proxy order based on the privileges and successful authentications of the user and the Local *Retrieval Manager*. The *Retrieval Manager*s must interact on distributed information in order to resolve the decision about passthrough versus proxy. If all of the information was accessible to a single ICS element, that element could make the decision on order type based on Table 6-2. A design goal is to have the selection of proxy vs. passthrough as invisible to the User as possible. Table 6-2 assumes the Local *Retrieval Manager* can be authenticated by the Remote *Retrieval Manager*, otherwise secure indirect order is not possible.

Table 6-2. Passthrough vs. Proxy Order Decision Table

Authenticated User has privileges on Remote *Retrieval Manager***?**	**Authenticated Local** *Retrieval Manager* **has privileges on Remote** *Retrieval Manager***?**	**Authenticated User has privileges on Local** *Retrieval Manager***?**	**Type of Order**
Yes	Yes or No	Yes or No	Passthrough
No	Yes	Yes	Proxy
No	No	Yes or No	order not allowed
No	Yes or No	No	order not allowed

The steps of the scenario in Figure 6-4 are as follows:

1. A User sends an *ExtendedServicesRequest* for CIP ordering to the Local *Retrieval Manager* where the action is orderQuoteAndValidate. The order is of a direct type as it is only between the Client and the Local *Retrieval Manager* at this point.
2. The Local *Retrieval Manager* inspects the order specification and determines that the products are provided by an OHS associated with a Remote *Retrieval Manager*. The Local *Retrieval Manager* checks if the user name in the order specification is defined at the local site and, if so, prompts the user to provide a user name defined at the remote *Retrieval Manager*. The Local *Retrieval Manager* creates an *ExtendedServicesRequest* for a Passthrough order operation. Passthrough is used because, if valid, passthrough will allow the highest privileges. The passthrough order contains the user names for the Local *Retrieval Manager* and the User.
3. When the Remote *Retrieval Manager* receives the *ExtendedServicesRequest* it identifies that this is a passthrough order and not a direct order. The Remote *Retrieval Manager* checks if User is known to Remote *Retrieval Manager* UPS. If the User is known, the order will remain as a passthrough order. If the User is not known, the Remote *Retrieval Manager* checks if the Local *Retrieval Manager* is known to the Remote *Retrieval Manager* UPS. If the Local *Retrieval Manager* is know, the order will be changed to a proxy order. If neither

the User nor the Local *Retrieval Manager* are known, the order will be treated as a guest order or rejected with an access control request, asking that the User supply a user name that is defined at the Remote *Retrieval Manager* UPS.

4. The Remote *Retrieval Manager* submits the orderQuoteAndValidate *ExtendedServicesRequest* to the OHS Translator, with the possibly modified user information.
5. The OHS Translator interacts with the OHS to develop a quote for the order based on the user information provided in the order.
6. The OHS Translator sends an *ExtendedServicesResponse* to the Remote *Retrieval Manager* including the quote for the order.
7. The Remote *Retrieval Manager* sends an *ExtendedServicesResponse* to the Local *Retrieval Manager* including the quote for the order.
8. As part of the process of updating its task package database, the Local *Retrieval Manager* checks to see if the order was converted to a proxy order by the Remote *Retrieval Manager*. If the Local *Retrieval Manager* accepts serving as proxy for this User's order to the Remote *Retrieval Manager*, the Local *Retrieval Manager* will update the local task package and pass the order to the user. (If the Local *Retrieval Manager* decides that it will not allow a proxy order in this instance, the Local *Retrieval Manager* may either resubmit the order to the Remote *Retrieval Manager* as a guest order or the Local *Retrieval Manager* may reject the order and return an access control request, asking that the User supply a user name that is defined at the Remote *Retrieval Manager* UPS.)
9. The Local *Retrieval Manager* sends an *ExtendedServicesResponse* to the Client including the quote for the order.
10. The user now decides to submit the order that was quoted in the preceding steps. The Client sends an *ExtendedServicesRequest* for CIP ordering to the Local *Retrieval Manager*, where the action is orderSubmit.
11. The Local *Retrieval Manager* submits the *ExtendedServicesRequest* to the Remote *Retrieval Manager* where the action is orderSubmit
12. Because the order requires privileges, the Remote *Retrieval Manager* requires that the user be authenticated, so the Remote *Retrieval Manager* sends an *AccessControlRequest* to the Local *Retrieval Manager* requesting that the Local *Retrieval Manager* present its credentials
13. If the order is passthrough, the Local *Retrieval Manager* responds with an *AccessControlResponse* containing a MAC based on the Local *Retrieval Manager*'s user name and password. (If the order was proxy, the Local *Retrieval Manager* would first authenticate the User, using data in the Local *Retrieval Manager* UPS, before returning the *AccessControlResponse* to the Remote *Retrieval Manager*)

14. The Remote *Retrieval Manager* constructs a MAC using the user name and password for the Local *Retrieval Manager* stored in the Remote *Retrieval Manager* UPS to authenticate the Local *Retrieval Manager*. (If the authentication fails, neither proxy nor passthrough ordering is allowed. With a successful authentication of the Local *Retrieval Manager*, a proxy order would be submitted directly, i.e., jump to step 19)
15. For a passthrough order, the Remote *Retrieval Manager* sends an *AccessControlRequest* to the Local *Retrieval Manager* requesting authentication of the User.
16. The Local *Retrieval Manager* sends the *AccessControlRequest* to the Client requesting that the client present its credentials
17. The Client sends an *AccessControlResponse* to the Local *Retrieval Manager* containing a MAC constructed with the User's password.
18. The Local *Retrieval Manager* sends an *AccessControlResponse* to the Remote *Retrieval Manager* containing the MAC constructed by the Client with the User's password. The Remote *Retrieval Manager* authenticates the user by comparing the received MAC with a MAC constructed with the User's password held in the Remote *Retrieval Manager* UPS.
19. The Remote *Retrieval Manager* has now performed the authentications necessary to submit the privileged order. The Remote *Retrieval Manager* sends an *ExtendedServicesRequest* to the OHS Translator.
20. The OHS Translator interacts with the OHS to submit the order.
21. The OHS Translator sends an *ExtendedServicesResponse* to the Remote *Retrieval Manager* including the status of the order, e.g., "order being processed."
22. The Remote *Retrieval Manager* updates the order task package in the Remote *Retrieval Manager* ESDB and sends an *ExtendedServicesResponse* to the Local *Retrieval Manager*.
23. The Local *Retrieval Manager* updates the order task package in the Local *Retrieval Manager* ESDB and sends an *ExtendedServicesResponse* to the Client.
24. The Client displays the status to the User. The User smiles.

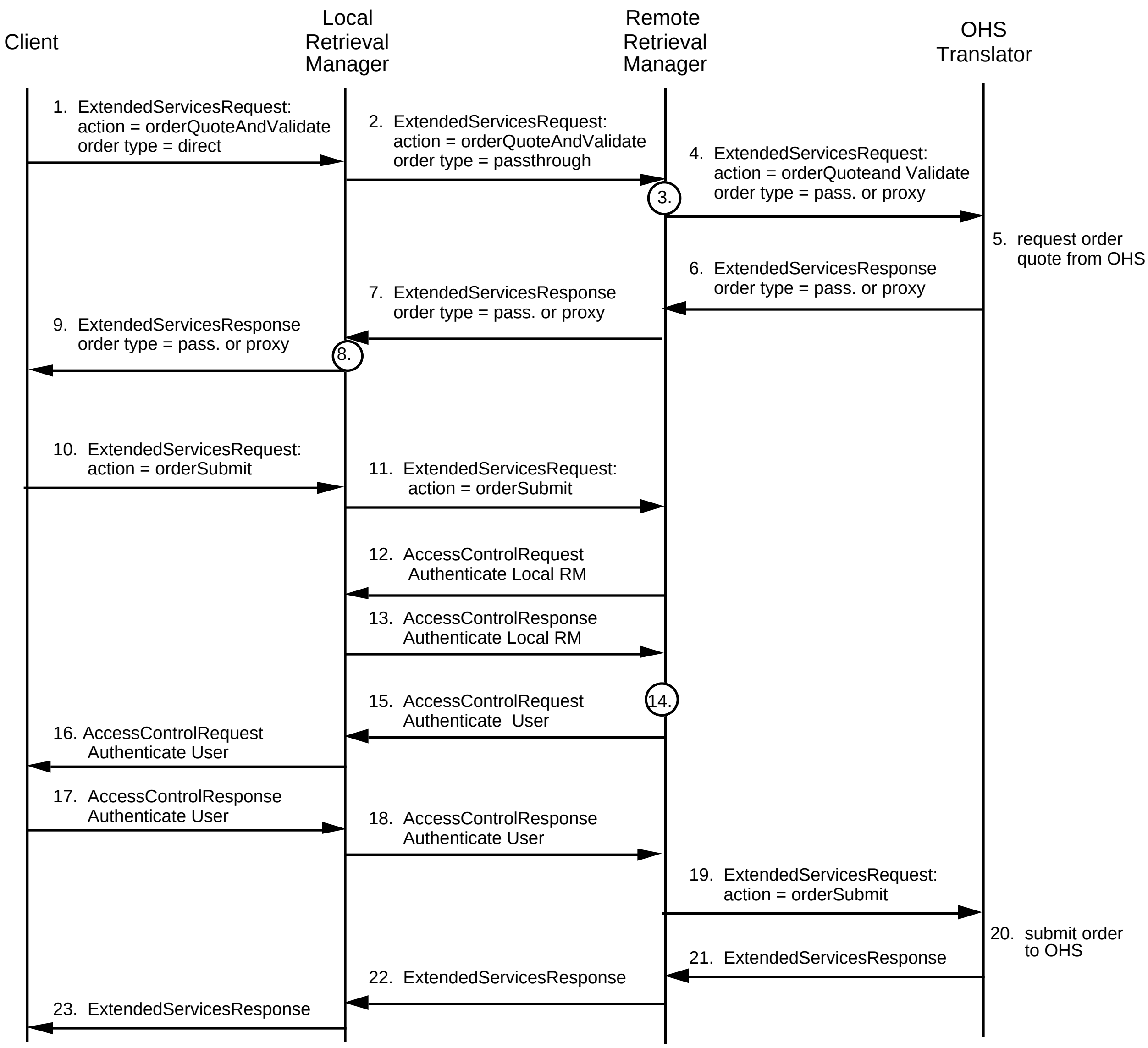


Figure 6-4. Secure Indirect Ordering

7. SYSTEMS MANAGEMENT VIEW (ICS)

The ICS Federation is a concept which describes a collection of ICS Sites. To preserve the integrity of the ICS Federation and provide some amount of assurance that each known ICS Site is compliant with the CIP and IGP Protocols, System Management functions have been identified and further described throughout the remaining subsections of Section 7. Individually these functions ensure a degree of interoperability among ICS Components; Collectively, they provide an overall picture of the compliance of the ICS System.

7.1 System Management Functions

ICS Compatibility: Explanatory

To ensure the stability and reliability of the ICS across sites while insuring site autonomy, a federation management approach to systems management has been selected. This federation management will

- Verify and Validate ICS systems that request to be included in the ICS Federation,
- Provide Daily Maintenance and Operation of the ICS Federation,
- Manage the evolution of the ICS Federation, and lastly,
- Assure Science Consistency of data held in the ICS Federation.

Collectively these four high level functions will provide the necessary structure to ensure that the tenets of the CIP and IGP Protocol and the Earth Observation Data are supported throughout the ICS Federation. Each are explored in more detail in the following sections.

7.2 Verification and Validation of ICS Systems

Before a new site can be added to the federation its compliance with the CIP and IGP has to be verified against the CIP and IGP Specifications. For this purpose an acceptance test will be performed which will evaluate the proposed ICS System against a set of ICS Functions. As a result of this acceptance test the ICS System will be marked as non-operational or as operational.

Each of the ICS Components; *Retrieval Manager, OHS Translator, and Catalog Translator*; will be tested through a series of functional tests which may span multiple components. The test scripts will be organized such that any component point of failure will be captured and recorded.

This acceptance testing that is performed will not replace the internal site testing that should be performed by the organization owning the ICS System. The acceptance testing will only begin when the agency requesting that the new system become a ICS Compliant System, certifies that the new system has passed all internal site testing.

The ICS Federation Tests results will be made available to the CEOS agency that requested the ICS verification and validation and to the ICS Federation Group responsible for Maintenance and Operations, and System Management.

The following sections provide an overview of the required testing. A detailed description of the acceptance tests will be provided in a separate document.

7.2.1 ICS Component Compliance Acceptance Test

ICS Compatibility: Mandatory

The acceptance test of an ICS Component will be based on a standardized set of test cases where every test case defines criteria which has to be instantiated for every particular test. The test cases are grouped with respect to the ICS Functions. i.e., Catalog, Browse etc.

Table 7-1 identifies each of the ICS Functions and related components that must be tested.

Table 7-1. Acceptance Testing Functions

ICS Function	ICS Components
LocalCollection Search	RM
RemoteCollectionSearch	RM
LocalCatalogSearch	RM, Catalog Translator
RemoteCatalogSearch	RM, Catalog Translator
DistributedCollectionSearch (5 or more RM's)	RM
DistributedCatalogSearch	RM, Catalog Translator
Evaluate results for Collection: Geographic Selection Temporal Selection Combinations of attributes	RM
Local Attribute Search	RM
Browse	RM
PassthroughOrder	RM,OrderTranslator
ProxyOrder	RM,OrderTranslator
Restricted Order	RM,OrderTranslator
Guest Access	RM
Status Messages for Errors	RM
Free Text Guide Search	RM,ICSGuideServer
Fielded Keyword Guide Search	RM,ICSGuideIndexer

Every new ICS Component is tested with all test cases which refer to a supported ICS Function. There are at least three possible test results for each test that is performed:

- The ICS System passes all tests. In this case the ICS System will be marked as operational for all specified services.

- The ICS System passes some tests for some components. The ICS Federation Testers will make the determination on the severity of the failed tests and decide whether to mark the protocol operational or developmental.

- Some tests fail for every function. In this case the system will be marked as non-operational.

The set of test cases may be extended to include other tests if other problems occur during the operation of the system.

The ICS Federation Testers will inform the organization that owns the system of the results of the test and schedule any future re-tests that may be necessary.
Experience with the current CINTEX federation has shown that several iterations of acceptance tests are sometimes necessary.

7.2.2 Overall Evaluation of ICS System Compliance

ICS Compatibility: Mandatory

The results of each of the above tests will be consolidated into an overall verification and validation report which will specify the details of each of the tests and the results. Further this report will specify the recommendations from the ICS Federation.

7.3 Maintenance & Operations (M&O)

Maintenance & Operations of an ICS Site will require extensive use of a Report Generation Service which will provide access to ICS Federation wide information. Types of reports that would be useful within ICS are as follows:

1) Performance/Compliance - Reports to describe short and long term trends in system operation in relation to established performance criteria.
2) Workload - Reports that provide statistics on the number of ICS invocations, or problem reports generated.
3) Accountability - Reports which provide audit and trace capability of significant events associated with users.
4) User Access Statistics - Reports which provide information for user modeling such as statistics on the types of searches users perform, number of users per month, number of product orders per user, the source and location of the orders, etc.

7.3.1 System Monitoring and Control (SM&C)

ICS Compatibility: Mandatory

The System Monitoring and control task performs the monitoring of all operational components (CIP Client, Retrieval Manager, Translators, etc.) within the ICS Federation. In addition, this task monitors the configuration of the components with respect to their location and the content. The SM&C task will establish procedures for dealing with unusual events such as component failure, new component addition, etc. in the ICS federation along with gathering statistics relating to the reliability/availability of the ICS agency member resources.

The SM&C function will monitor components on a regular interval by sending standard requests and receiving responses to these requests or by using customized scripts which invoke operating system commands to determine the status of a the ICS member agency components. The interval for conducting this activity will have been established by M&O.

The SM&C task will attempt to locate the problem if a component does not respond and inform the operator of the component of the problem. If the component will not return to operation for an extended time period, then the ICS Federation M&O will remove it from the configuration files of all gateways to make it invisible for the users until the problem has been corrected. The SM&C task will also monitor the gateways by accessing them or using automated tools to determine the status at an interval that is configurable by the M&O.

7.3.2 Valids Management

ICS Compatibility: Mandatory

The Valids Management task will provide a science and consistency check on system wide valids being defined on inventories accessible through the ICS federation. A consistency check is made to ensure that the appropriate and complete set of data requested is properly retrieved. The valids management task will review agency's proposed valids on the new inventory being introduced to the ICS federation and then approve the proposed valids or provide feedback to improve the science and consistency of the valids. The details of the process for valids management is captured in the ICS Valids Document [R4]

The Valids Management task may coordinate the local attributes being developed at various sites.

7.3.3 Federation Help Desk

ICS Compatibility: Mandatory

The federation level help desk function will perform the coordination of the various ICS Agency help desks. The federation level help desk will also document and maintain a recommended list of services to be provided by each agency's individual help desk organization. The federation help desk will also provide support to agencies establishing an agency help desk for the first time.

The individual agency help desks will provide support to users of the distributed catalog system. The federation help desk will provide input to the ICS group on how to get agency help desk contact information to the users and how the user queries can be routed to the responsible agency's help desk. If the responsible agency has no individual help desk, the federation help desk will decide how to provide coverage for these situations. The federation help desk will serve as the help desk for the agency help desks. Individual user questions should not reach the federation help desk directly. It should always be brought up by an agency help desk. One agency's help desk can be designated as the federation level help desk .

7.3.4 System Wide Interoperability Interface Management

ICS Compatibility: Mandatory

The System Wide Interoperability Interface management task will provide technical consulting to agencies on details of how to link their systems into the ICS, and for any software development or customization of components needed to link into the system wide interoperability interface layer. This task will manage the set of documentation to define the system wide interoperability layer.

Within System Wide Interoperability Interface Management , the component distribution function provides the capability for technical ICS system administrative staff (M&O) to distribute new or upgrades to ICS software, database schemas, documentation and related items. A centralized approach is the most efficient method for ensuring that each ICS site receives any and all updates to ICS components.

7.4 System Management (SM)

The System Management roles and responsibilities will focus on the evolution of the system. This includes the planning for the inclusion of new system or protocol components, the modification of existing components and removal of old components. It also includes the inclusion of new data and the update for the schemas to describe the data and the server. The SM arranges for any verification of new components. The SM also identifies all the required activities to ensure that system modifications are adopted in an orderly manner.

7.4.1 Decision Making Process

ICS Compatibility: Mandatory

The System Management is also responsible for collecting information on testing, system monitoring, and control and providing this information to the ICS federation. The system management role also includes

a) forwarding recommendations to the ICS federation
b) removal of components from the ICS federation operational baseline
c) addition of new components into the ICS federation operational baseline
d) coordinate meetings and teleconferences for ICS federation

The types of decisions and inputs that the ICS federation must provide are :

a) inclusion of new ICS site into the federation
b) recommended data access policies
c) schedule for system evolutionary components
d) removal of a ICS site from the federation
e) provide inputs to the agendas for teleconferences and meetings

8. ARCHITECTURE VERIFICATION

This chapter provides the results of several types of analysis which demonstrate that the system design described in the previous chapters will meet the ICS requirements and that the various architectural views are consistent. The analysis results in this chapter also provide design information for the developers of the particular elements of the ICS.

8.1 Query Performance Estimates

ICS Compatibility: Mandatory

Key to user satisfaction with ICS is the response times for searches. Based on the design presented in the previous sections and a set of design constraints listed in this section, estimates for the four types of CIP searches were made. Estimates for the query response times under nominal conditions are provided in Table 8-1. For detail on how the estimates were developed, see [R19]. The collection search estimates are comparable to WWW index searches, e.g., Infoseek and Alta Vista. The product search estimates are dominated by the *Existing Catalogue* search times, which are outside of ICS.

Table 8-1. Query Performance Estimates

	Collection Search	Product Search
Local Search	2 sec.	125 sec.
Distributed Search	13 sec.	136 sec.

To attain the query performance listed in Table 8-1, the design in the previous sections must be implemented as well as the following design constraints.

- Within a given *Retrieval Manager* the collection database has been well laid out and is fairly efficient with regard to the expected queries.
- The network time, including time to form Z–associations, is small compared to the time required to service a query for the network connection between a *Retrieval Manager*, the *Catalogue Translator* and the local catalogue system all at the same site.
- *Retrieval Managers* will send in parallel. For example, if a query being performed at RM_A requires subqueries to be performed by RM_B and RM_C, RM_A will send the subqueries to RM_B and RM_C simultaneously rather than sequentially. RM_B and RM_C will perform their subqueries in parallel and present their response to RM_A.
- All elements (*Retrieval Managers*, networks, translators, etc.) can be modeled as M/M/1 queues. If this is not the case, as would be the case if some element had multiple servers, then certain equations in the distributed query cases may have to be revisited. (See [R21] for a discussion of Queuing Theory.)
- *Retrieval Managers* are efficient in detecting and ignoring collection-to-collection overlap. A Retrieval Manager will not search a collection more than once during a given local collection search (regardless of whether the original query was local or distributed)
- *Retrieval Managers* will keep a CIP session open with their associated *Catalogue Translators*, such that a typical product search will not require an initialization delay between the *Retrieval Manager and Catalogue Translator.*

Table 8-2 defines the parameters used in estimating query performance within ICS. As the performance is dependent upon these parameters, the parameter values must be met to provide the query performance in Table 8-1.

Table 8-2. Performance Parameters and ICS Elements

Parameter	**Affected ICS Element**	**Units**	**Nominal**	**Range**	**Source**
Average collection query service time	*Retrieval Manager*	seconds	1		Logica
Retrieval Manager utilization	*Retrieval Manager*	n/a	0.5	0-0.99	Raytheon/2
Average response time required to search a local collection	*Retrieval Manager*	seconds	5		URD
Average time required to form a subquery from a query	*Retrieval Manager*	seconds	0.01		Logica
Average time required to establish a Z–association	*Retrieval Manager*	seconds	5	0-60	Logica & Raytheon/1
Average time to pass a message or response to the network	*Retrieval Manager*	seconds	0.01		Logica
Average catalogue translation service time	*Catalogue Translator*	seconds	1		Raytheon/1
Catalogue translator utilization	*Catalogue Translator*	n/a	0.5	0-0.99	Raytheon/2
Average number of collections per collection	*ISA*: collection structure	n/a	5	5-50	Delphi
Collection depth	*ISA*: collection structure	n/a	4	2-7	Delphi
Probability of overlapping collections	*ISA*: collection structure	n/a	0.7	0-0.95	Logica
Ratio of remote links to total number of links	*ISA*: collection structure	n/a	0.5	0-0.95	Logica
Average number of unique translators at a Retrieval Manager	*ISA*	n/a	1		Raytheon/2
Number of ICS sites (Number of Retrieval Managers)	*ICS System Management Group*	n/a	13	10-200	CINTEX Sites
Average response time of network transmission	*CEOSnet*	seconds	0.16		Germain
Average catalogue query service time	*Existing Catalogue* (Not part of ICS)	seconds	60	12-240	Delphi
Inventory catalogue system utilization	*Existing Catalogue* (Not part of ICS)	n/a	0.5	0-0.99	Raytheon/2

Legend for "Source" in Table 8-1:

Delphi	PTT Delphic study.
Germain	Estimate by CEOS Network Performance Test (Andy Germain)
Raytheon/1	Estimate based on translator prototype.
Raytheon/2	Estimate supplied by Raytheon Systems Corporation.
Logica	Estimate supplied by Logica.
URD	Requirement 374 in ICS URD [R2].

8.2 Scenarios

This section provides several scenarios showing the dynamic aspects of the ICS including how the scenarios are accomplished via interfaces between ICS elements and services of the ICS elements. Scenarios for both the user's and operator's activities are provided.

Scenarios in this section are organized into the following categories:

- User Scenarios
  - WWW Access to a *Retrieval Manager*
  - Existing Agency Client Access into ICS
  - Indirect Ordering Scenario
  - Guide Search - WWW Search Engine
  - Guide Search - ICS Client
  - Guide Retrieval
- Collection Population Scenarios
  - Making a CIP Compatible Catalogue Available
  - Collection Established for an Event
- System Management Scenarios

8.2.1 User Scenarios

8.2.1.1 WWW Access to a Retrieval Manager

ICS Compatibility: Explanatory

General

The scope of this scenario is to follow a user's confrontation with the system and to simulate whether a user can successfully retrieve helpful documents and EO-data from the ICS and related Catalogues.

This scenario is based on the CEO Enabling Services Scenario 1: Search for EO-data, documents and adverts [R17]

Assumptions

- The user is familiar with WWW browsers, has access to a web browser, is connected to the Internet and knows the HTTP address of a CIP/WWW Gateway.
- He has only a basic knowledge of remote sensing.
- The user works for a consultancy company that is specialized in identifying and evaluating potential water reservoirs in third world countries.

- The present project aims to build a water reservoir in Egypt. He has an explicit question he needs to be answered: Is a high resolution elevation model of this area available?
- He is interested in any documents related to this subject.
- The user has never entered the ICS before and is therefore not registered. The scenario does not include an authorization operation and the user is considered a member of the "guest" group.

Expected Outputs

The user wants to find out if a DEM of the target region is available. For the DEM data he hopes to find explicit meta-data that describes the data (i.e. browse) before he orders it. (He tries finding documents covering this subject.)

Step Sequence

1. The user enters the CIP/WWW Gateway home page for a *Retrieval Manager*.
2. He reads the short introduction about what the local *Retrieval Manager* can do for him.
3. He enters the "EO-data search" interface. The user has the choice between a local or remote collection search , a data search against the entire ICS holdings (using the global node) , a local search against the entire local site holdings (using the root node), or against popular topics (using the key access nodes).
4. The user chooses the local collection search of the root collection after reading the short definition supplied by the system. The form that is supplied contains fields in which the search attributes have to be entered.
5. The user selects DEM from the listed "product attributes".
6. He selects "display full valid definition" to make sure the acronym is he expected.
7. The system displays the full definition of a digital elevation model.
8. The definition is what the user expected and the user submits the local collection search.
9. After a syntax check of the search the system accepts the search.
10. The system returns a structured list of collections which point to data that incorporates DEMs. One of the listed collections is called: "Elevation Models of Africa; ESA; 1992".
11. The user chooses: continue with "EO- product search".
12. The system displays a form which contains attribute fields.
13. The user chooses the: "Elevation Model of Africa; ESA; 1992" and another elevation model terminal collections from the collection list. (The collection may be chosen in the collection search).
14. The user chooses the "select geographical region" option.
15. This gives him the following choices:
    - type global coordinates
    - draw area of interest by polygon
    - draw area of interest by point and radius
16. The user chooses "type global coordinate".
17. He enters the coordinates for Egypt.
18. He submits the EO-product search.
19. The system indicates that the search is valid after a successful syntax check has been performed.
20. (Note this step was deleted from the CEO scenario as CEOS policy is to not charge for searching.)

21. The user asks for a “status monitor” to display the status of the search.
22. A status monitor is displayed showing the two parallel searches (compare to step 13). The search “Elevation Models of Africa, ESA, 1992” is finished but the second search has not received any reply by the provider site yet.
23. The user selects “display product list”.
24. The results of the EO product searches are displayed in a list.
25. The information displayed contains information on the location, provider, title and the availability of on-line meta-data and browse data.
26. The user selects a product.
27. The user chooses the “transfer metadata” option.
28. The metadata is transferred and displayed.
29. The user chooses the “transfer browse product” option.
30. The browse product is transferred and displayed.
31. The user selects the “order product” option and orders the product delivered via ftp.
32. An FTP-session is started and the product is transferred to the machine of the user. *Steps 32 and 33 are outside the ICS, but they are mentioned here for completeness purposes*.
33. The user saves this file on his hard disk.
34. The user chooses the “save EO-data search” option.
35. The system saves the “data search ” (the system stores this internally while this session continues).
36. The user chooses the “save EO-data search results” option. (The result set is converted to a hot collection, held on the users client.)
37. The system saves the “EO-data search results”.

8.2.1.2 Existing Agency Client Access into ICS

ICS Compatibility: Explanatory

General

This scenario follows both the user steps and the system software steps for an agency's user to access and order products held in the ICS collections using his local client software and an ICS gateway. The scenario concentrates on the user visible processes and the interaction between the *ICS Gateway* and the ICS *Retrieval Managers*.

Assumptions

1. User is an experienced, registered User of his local Agency (Agency A) catalogue system and is using his local Client.
2. Agency A has 2 way interoperability with ICS that is transparent to the users of its systems.
3. The User has established a session with his local catalogue systems as a registered user
4. The user is interested in AVHRR data over Europe (resident in ESA ICS holdings)
5. The user has no ICS experience and is not a registered user
6. The local agency catalogue system has been configured to recognize ICS collections through a catalog/Gateway configuration
7. Agency A and ESA have a trust agreement for authentication.
8. ESA and Agency A have a bilateral agreement that allows Agency A to act as a proxy for its users to order data. ESA bills Agency A for data ordered by its users.
9. The scenario assumes the ICS Functional Framework shown in Figure 3-3.

Agency A architecture is similar to third site in Figure 3-9 i.e., Retrieval Manager as a catalogue gateway.

Expected Outputs

The user uses his local client (*Existing Agency Client)* and finds new data through the *Retrieval Manager* which is held at a site to which his site previously did not have access.

Scenario Sequence

Search Sub-Secnario

1. Human user initiates local catalogue system client and logs in as a registered user
2. The user develops a query to discover any products containing AVHRR data over Rome taken in the 1995-1997 time frame.
3. Based on the contents of the query and information from the local agency catalogue the query is sent to the ICS Gateway
4. *ICS Gateway* translates the query from the local query language (e.g., ESQL) to the CIP Query format (RPN) and translates local agency attribute names attributes into appropriate CIP attributes
5. *ICS Gateway* interacts with the local catalogue system to determine which ICS sites can satisfy the query
6. The *ICS Gateway* acts as a ICS *CIP Client* and establishes a session to the Agency A *Retrieval Manager* (RM) using CIP and sends the query to the Agency A RM
7. The RM then sends the query to the remote *Retrieval Managers* which pose queries to local collections and return results to the local agency ICS Gateway.
8. The *ICS Gateway* translates the returned result set into local format and returns the result set to the Client.
9. The human user evaluates the results and requests browse attributes for four data products of interest in ESA collections
10. Steps 3-8 repeated
11. The user decides to order the four products

Order Sub-Scenario

12. The local Client puts up an order specification form which the user fills out
13. The local Client sends an order quote request to the ICS Gateway
14. The *ICS Gateway* translates the order quote message into CIP format and the local product identifiers into appropriate CIP product identifiers
15. The *ICS Gateway* acts as a *CIP Client* and establishes a session to the Agency A *Retrieval Manager* (RM) using CIP and sends the quote request to the Agency A RM
16. The Agency A RM sets up an authenticated session with the ESA RM as Agency A via the CIP to obtain Agency A session options.
17. The Agency A RM sends the quote request to the ESA RM and requests a price estimate for the order.
18. The ESA RM forwards the quote request to the local Order Handling System (OHS) which returns the estimate to the ESA RM.
19. The ESA RM returns the quote to Agency A RM.
20. The Agency A RM forwards the estimate to the ICS Gateway
21. The *ICS Gateway* translates the returned message into local format and returns the message to the Client
22. The user reviews the estimate and requests the order be submitted

23. The *ICS Gateway* checks the user's credit via an interaction with the Agency A Billing and Credit Subsystem and on credit approval obtains an order number for local billing
24. The *ICS Gateway* translates the estimate approval format into CIP and forwards it to the Agency A RM which forwards it to the ESA RM which forwards it to the local OHS
25. The local OHS accepts the order and sends a status update of Order Acceptance to the ESA RM which forwards it to the Agency A RM which forwards it to the ICS Gateway
26. The *ICS Gateway* translates the message to the Agency A format and order number and informs the local Client and the local Billing and Credit subsystem of the order status
27. The human user logs off his local client and all sessions are terminated

8.2.1.3 Indirect Ordering Scenario

ICS Compatibility: Explanatory

General

The scope of this scenario is to show how a user places an order with a *Retrieval Manager* at which he is registered (*Retrieval Manager* - A) and, transparent to the user, the order is routed to another *Retrieval Manager* (*Retrieval Manager* - B) which has access to the products.

Assumptions

1. User has performed a search process and obtained the identifiers of a set of products which he wishes to order.
2. The user’s agency and the data holding agency have an agreement that allows the user’s agency to act as a proxy for its users to order data. The data holding agency will bill the user’s agency for the data. The user’s agency will bill the user for the data.
3. The order is placed with a *Retrieval Manager* (*Retrieval Manager* - A) with which the user is registered and has privileges sufficient to allow the order.
4. The Intermediary *Retrieval Manager* (*Retrieval Manager* - A) is authenticated with the *Retrieval Manager* which will fill the order (*Retrieval Manager* - B).
5. The order will be delivered in media, e.g. CD-ROM.
6. The user has established an authenticated session with *Retrieval Manager* - A.
7. *Retrieval Manager* - A has established and authenticated session with *Retrieval Manager*- B.

Expected Outputs

The user reviews a quote for the order, submits the order, and receives the products some time later.

Step Sequence

1. The *CIP Client* displays an order form which the user fills out.
2. User chooses to be charged by his own agency (not the agency that holds the data).
3. The *CIP Client* sends the order request to *Retrieval Manager* - A as part of an authenticated session.
4. *Retrieval Manager* - A determines the user’s privileges based upon the user’s group membership for that session. The group privileges allow this order request.
5. *Retrieval Manager* - A identifies the site from which data is being ordered (*Retrieval Manager* - B) based on information in the product identifiers in the order.

6. *Retrieval Manager* - A sends the order request and a group designation to *Retrieval Manager* - B as part of an authenticated session. (Note that the group designation in this session is for the *Retrieval Manager* - A session and may differ from the group of the client's session.)
7. *Retrieval Manager* - B determines the privileges based upon group membership. The group privileges allow this order.
8. *Retrieval Manager* - B forwards the order to the *OHS Translator* which converts the order into a local OHS order. The translator then passes the order to the OHS.
9. The OHS produces a quote and cost breakdown and assigns a quote number.
10. OHS sends quote, cost breakdown and quote number to *Retrieval Manager* - B (via the *OHS Translator*) which forwards it to *Retrieval Manager* - A which forwards it to the *CIP Client*.

[Steps 2-10 can be repeated by the user by changing and re-submitting the order specification to get data pricing information.]

11. The user reviews the quote and decides to submit the order. The *CIP Client* sends the order submittal to *Retrieval Manager* - A.
12. The group privileges for the user's session allow this order but determines that the order must be non-repudiatable. *Retrieval Manager* - A sends a non-repudiation request to the CIP Client
13. The *CIP Client* requests the user to confirm the order submittal, which the user does.
14. The *CIP Client* sends the order as a non-repudiatable message to *Retrieval Manager* - A.
15. *Retrieval Manager* - A sends order submittal to *Retrieval Manager* - B.
16. *Retrieval Manager* - B determines the group privileges associated with the session with *Retrieval Manager* - A and allows the order submittal.
17. *Retrieval Manager* - B passes the order submittal to the local *OHS* (via the *OHS Translator*).
18. OHS accepts the order, determines an expected order completion date, and returns an order submittal response to *Retrieval Manager* - B.
19. *Retrieval Manager* - B passes the order submittal response to *Retrieval Manager* - A.
20. *Retrieval Manager* - A notifies the local *OHS* (via the *OHS Translator*) that the order submittal has been accepted.
21. The OHS associated with *Retrieval Manager*-A does the accounting and billing associated with the order and the user's account.
22. *Retrieval Manager* - A passes the order submittal response to the CIP Client.
23. Possibly after the users session is closed, the OHS associated with *Retrieval Manager* - B fills the order and sends the bill to the OHS associated with *Retrieval Manager* - A.

8.2.1.4 Guide Search Scenario - WWW Search Engine

ICS Compatibility: Explanatory

General

A user will be able to discover ICS Guide documents through general Web search engines, such as Alta Vista, Yahoo, or Lycos,

Assumptions

In order for users of general Web search engines to find ICS guides the following must take place prior to the search:

- An ICS *Guide System* must be set up at least at one site.
- Guide documents must have been ingested into the system.
- The companies hosting Web search engines should be notified of the URL of one instance of the Sites Listing file icsdoc.html. Recall that this file will be mirrored at every ICS site.
- The Web search engine must have indexed the ICS guide documents, which it finds from icsdoc.html.

Expected Outputs

The user is presented with URL links to guide documents accessible in ICS which match the user's query.

Step Sequence

1. User submits a free-text search query to a general Web search site (such as Alta Vista). If they just enter strings that they expect will be in the ICS guides, they will find them. However, since these general Web search sites index all documents on the WWW, users will likely get matches of documents outside of ICS. They could narrow their search using known Guide Attribute/Keyword pairs. For instance: "AuthorName=John Doe".
2. The search engine looks up the URLs for the associated query strings in its index and displays the links for guide documents that contained the free-text. Note that while the user can use Guide Attribute/Keyword pairs because of the insertion of Meta tags, the search is still a free-text search. These search engines only use free-text.
3. The search engine responds back to the WWW Browser with a list of links.

8.2.1.5 Guide Search Scenario - ICS Client

ICS Compatibility: Explanatory

General

An ICS fielded guide search allows the user of an ICS Client to specify values for guide attributes, which will be in the list of guides returned. The IGP Client sends the FieldSearchForm IGP message to the ICS Guide Server.

Assumptions

- The user has a URL for an *ICS Guide Server.*

Expected Outputs

The user is presented with URL links to guide documents accessible in ICS which match the user's query.

Step Sequence

1. *ICS Client* connects to an ICS Guide Server using a *FieldSearchForm* message
2. The *ICS Guide Server* invokes the *icssearch* CGI script with the argument "searchform"
3. *icssearch* reads the Attributes Default File, generates a fielded search form HTML page, and sends it to the *ICS Client* to display. The form contains Guide Attribute names for labels for text boxes that the user can fill in. There is an additional text box with the label "Free Text" which the user can use to enter free-text to search. The form will have a submit button on it.
4. The user fills in the search form with desired entries
5. The ICS Client connects to the *Guide Server* and sends a *SearchFields* message.
6. The server responds to the SearchFields message by invoking *icssearch*, the ICS search engine, with the parameter value list.
7. The query return a list of guide URLs.
8. *icssearch* constructs an HTML page containing the list of guide URLs and passes it back to the *ICS Client*
9. The *ICS Client* displays the page of matching guide URLs to the user.

8.2.1.6 Guide Retrieval Scenario

ICS Compatibility: Explanatory

General

This scenario describes how a virtual guide document is created when requested by a client. The guide itself if pulled from the *Guide Document Archive* and additional information is added. The additional information includes metatags and URLs for related collections.

Assumptions

- The user has a URL for a guide document which is accessible through ICS. This URL may have been obtained from a previous search.

Expected Outputs

The guide document is displayed to the user in their stand-alone WWW browser or ICS Client.

Step Sequence

1. User initiates an HTTP connection using the URL for a browse document
2. Client connects to the ICS Guide Server at the site indicated in the URL, with the *GetVirtualDocument* IGP message.
3. Guide Server invokes *icsdoc* in the Guide Translator with path information of the guide document.
4. *icsdoc* looks up the guide in the Guide Metadata Database, retrieves the metadata for that guide, and constructs Meta tags
5. *icsdoc* reads the guide document from the *Guide Document Archive* using the path information supplied in the message.
6. *icsdoc* inserts the Meta tags into the header section of the guide

7. *icsdoc* looks up the guide in the Collections Mapping File, extracting the collection IDs for the guide. *icsdoc* constructs links to the data collections. The format of the collection link will depend on the type of client requesting the document, WWW Browser vs. ICS Client.
8. the Guide Server responds back to the Client with the guide document.

8.2.2 Collection Population Scenarios

This section contains scenarios concerning establishment of collections.

- Establishing collections in a Retrieval Manager for the first time
- Creating a new Provider Theme collection in response to a Earth Science Event

8.2.2.1 Making a CIP Compatible Catalogue Available

ICS Compatibility: Explanatory

General

A provider who works for a large company wants to make his scanned aerial pictures accessible through the ICS.

This scenario is roughly based on CEO-ES Scenario 4: Making a CIP compliant catalogue available.

Assumptions

The following assumptions apply to this scenario:

- The provider has an Internet connection and a UNIX workstation.
- The provider has an existing inventory system on a relational database system which catalogues all the data products archived at the company.
- The catalogue the provider wants to make available is an EO-data catalogue.

Expected Outputs

The catalogue of the provider becomes accessible for the ES users.

Step Sequence

1. The potential data provider attends a CEOS meeting and decides to host a *Retrieval Manager*
2. Reviews CEOS documentation describing ICS
3. Accesses the CEOS *Retrieval Manager* software distribution site
4. Retrieves RM source code
5. Configures the *Retrieval Manager* for his specific environment
6. Develops translator for local inventory searches reusing a skeleton *Catalogue Translator*.

7. Develops a gateway from the RM to his local Order Handling System (OHS) by reusing the local OHS provided with the RM
8. Establishes Provider Archive Collections for each of his existing datasets by developing Collection Descriptors for each of the datasets and an Explain database which lists the available services at his RM, and describes some product specific attributes which are used in the local catalogue systems to describe specific details of the individual photographs
9. Establishes a Provider Theme Collection which combines several existing datasets which contain aerial photographs and modifies the Explain Database to list the new Provider Theme Collection as a Key Access node
10. Establishes a root collection that references all the collections at the local site
11. Contacts the Global Collection Administrator via email and describes the datasets that have been established. The Global Collection Administrator advises the provider as to the keyword that should be used to characterize his collections and some remote collections that have similar themes
12. The provider modifies his root collection and his Provider collection to reflect the discussions with the Global Collection Administrator
13. Conducts local test of *Retrieval Manager* and collections
14. Send a message to the Global Collection Administrator announcing his desire to bring up a *Retrieval Manager* and join the ICS and puts the RM on-line to allow remote testing
15. The Global Collection Administrator performs integration testing on the new provider site, and when he is satisfied at the results, adds the root collection of the new provider site to the global collection and announces the availability of the new provider site and collections to the ICS community via email and the WWW.
16. The *ISA* at the new site continues to evolve his collections by adding new Provider Theme Archives and adding references to remote collections of interest to each of his collections

8.2.2.2 Collection Established for an Event

ICS Compatibility: Explanatory

General

This scenario demonstrates how a Provider Theme Collection is created. The scenario shows the interaction of the *ISA* and a scientist for the creation of the collection. The collection is created based on an event of particular weather patterns over the Andes.

This scenario is based on a scenario titled 'Climate, Erosion, and Tectonics in the Andes and other Mountain Systems,' which is ECS Scenario 22B, in [R22].

Assumptions

The following assumptions apply to this scenario:

- The user is a scientist at an agency which hosts a RM
- The user has discovered an interesting event analyzing data from an instrument.

Expected Outputs

The user wants to establish a collection for the event which his colleagues may access.

Step Sequence

1. A severe storm in the Andes is detected by user's review of MODIS Level 2 Imagery Products
2. User searches in ICS for other data products which overlap the event in temporal and spatial location.
3. The user establishes a hot collection based on the result set returned from his query. this results set contains data granules from MODIS, MISR, ASTER, GLAS, LANDSAT-7, ERS-2, ADEOS, and RADARSAT.
4. The user conducts incremental queries on the hot collection holdings to narrow the collection by determining the quality of each product and the degree of overlap with the storm track
5. The user establishes the final refined result set as a hot collection and deletes the earlier established hot collection
6. The user orders all the products remaining in his hot collection and applies various analysis techniques to detect the presence of the event in the other sensor data
7. The user uses hot collection editing tools to eliminate the products that do not show effects(e.g. landslides or floods) from the storm and contacts his local site *ISA* to request his hot collection be upgraded to a User Theme Collection.
8. The user writes a paper on the event and lists the URL for the collection requesting reviews by his colleagues to verify and augment his conclusions
9. The user receives several favorable reviews of his research and contacts his local *ISA* about the possibility of having the agency publicize and maintain the collection for long term preservation
10. The *ISA* and the agency science review board reviews the collection and the colleague comments and agrees to ingest the collection into the agency archive.
11. The *ISA* sends a form to the user requesting needed metadata about the collection. The user fills out the form and returns it to the *ISA*.
12. The *ISA* converts the user theme collection to a provider theme collection by upgrading the metadata, including the collection descriptor in the RM root collection.
13. The *ISA* advertises the collection via Email , bulletin boards and the CEO advertising service.
14. All ICS users are now able to access and search Andes Severe Weather Event Provider Theme Collection.

8.2.3 System Management Scenarios

ICS Compatibility: Explanatory

This section will contain scenarios concerning ICS System Management. As System Management is an ICS Release C focus, these section contains short descriptions which may be developed for Release C.

Retrieval Manager Registration. The scope of this scenario is to show how a *Retrieval Manager* becomes registered as an ICS *Retrieval Manager* and subsequently supports authenticated *CIP sessions* to allow ordering.

Planning a Retrieval Manager Outage. The scope of this scenario is to show how the *ISA*s work as a distributed management team. It is assumed that *ISA*s are trained in ICS Systems Administration procedures, a certain *Retrieval Manager* has been operational for some time and is remotely linked by many other *Retrieval Managers*, and the *Retrieval Manager* must be down from operations for several hours to change a piece of hardware. The expected output is that disruptions to ICS users are minimal and any users calls to *ISAs* are handled in an informed manner.

Response to Retrieval Manager Fault Condition. The scope of this scenario is to demonstrate the use of the *Retrieval Manager* monitoring functions by an operator in abnormal situations. It should be noted that the problems encountered in this scenarios are not representative of the ICS normal behavior and are presented here for illustration purposes. This scenario is based on CEO-ES Scenario 8: Middleware node operator - Use of the monitoring server.

8.3 Internal Interface Identification

Interfaces for the various ICS components have been stated in the multiple architectural views provided in the previous sections of the SDD. This section provides a summary of the interfaces between ICS components insuring consistency and completeness.

8.3.1 Retrieval Manager Interfaces

ICS Compatibility: MAA

The *Retrieval Manager* has the interfaces indicated in Table 8-3.

Table 8-3. Retrieval Manager Interfaces

Other ICS or Related Element	Interface	SDD Section
ICS Client	CIP sessions	3.5.1
ICS Client	Data Exchange	4.6
ICS Client	TCP/IP via National Internet	5.2.1
ICS Client	TCP/IP via World-wide Internet	5.2.1
Other z39.50 Clients	Z39.50, Version 2 or 3, sessions	3.5.1
Other z39.50 Clients	Data Exchange	4.6
Other *Retrieval Managers*	CIP sessions	3.5.1
Other *Retrieval Managers*	Data Exchange	4.6
Other *Retrieval Managers*	TCP/IP via National Internet	5.2.1
Other *Retrieval Managers*	TCP/IP via World-wide Internet	5.2.1
Other *Retrieval Managers*	TCP/IP via CEOSnet	5.2.1
Other *Retrieval Managers*	Secure Indirect Orders	6.2.4.4
Catalogue Translator	CIP sessions	3.5.1
Catalogue Translator	Data Exchange	4.6
OHS Translator	CIP sessions	3.5.1
OHS Translator	Data Exchange	4.6
UPS Translator	UPS Session	3.5.1
UPS Translator	Data Exchange	4.6
UPS Translator	Username/Passwords	6.2.4.3
ISA	Operator Interface	3.5.1
ISA	Operator Interface Data	4.6
ISA	Username/Passwords	6.2.4.3
Collection Management Tool	Collection Data Base Modification	3.5.1
Collection Management Tool	Data Exchange	4.6
Monitoring and Control Tool	*Retrieval Manager* Management	3.5.1
Monitoring and Control Tool	Data Exchange	4.6
ICS Gateway	CIP Sessions	3.5.1
All CIP Interfaces	Use of TCP/IP	5.1.1
All CIP Interfaces	CIP to TCP Mapping	5.1.3.1
Distributed CIP Interfaces	Session Management	5.1.4
DNS	Directory Service	5.1.6
Site Physical Facilities	Physical Security Control	6.2.3
Test Equipment	Functional Tests	7.2

8.3.2 ICS Client Interfaces

ICS Compatibility: MAA

The *ICS Client* has the functional interfaces indicated in Table 8-4.

Table 8-4. ICS Client Application Interfaces

Other ICS or Related Element	Interface	SDD Section
Retrieval Manager	CIP sessions	3.5.2
Retrieval Manager	TCP/IP via National Internet	5.2.1
Retrieval Manager	TCP/IP via World-wide Internet	5.2.1
Retrieval Manager	Data Exchange	4.6
HTTP/CIP Gateway	CIP sessions	3.5.2
Human user	Data to user	3.5.2
Guide Server	HTTP Connection	3.5.2
All CIP Interfaces	Use of TCP/IP	5.1.1
All CIP Interfaces	CIP to TCP Mapping	5.1.3.1
All IGP Interfaces	Use of TCP/IP	5.1.1
All IGP Interfaces	HTTP use of TCP	5.1.5
DNS	Directory Services	5.1.6

8.3.3 ICS Site Administrator (ISA) Interfaces

ICS Compatibility: MAA

The *ISA* has the operational interfaces indicated in Table 8-5.

Table 8-5. ICS Site Administrator Interfaces

Other ICS or Related Element	Interface	SDD Section
Retrieval Manager	Operator Interface	3.5.8
Retrieval Manager	Physical Security	6.2.3
Retrieval Manager	Operator Interface Data	4.6
Collection Management Tool	Operator Interface	3.5.8
Monitoring and Control Tool	Operator Interface	3.5.8
ICS User	Username/Password	6.2.4.2
Other *ISAs*	Inter-Agency Billing Agreements	6.2.4.3
Other *ISAs*	Inter-Agency Security Agreements	6.2.4.3
Other *ISAs*	Group Management	6.2.4.4
Other *ISAs*	Administrative Security Controls	6.2.2
Other *ISAs*	System Administration	7.
Interfaces outside of ICS	OHS, UPS, Catalogue, Archive, Guide Document Archive.	3.5.8
OHS Translator	Maintain and Operate	3.5.8
UPS Translator	Maintain and Operate	3.5.8
Catalogue Translator	Maintain and Operate	3.5.8
Guide Indexer	Maintain and Operate	3.5.8
ICS Federation Management	System Management Operations	7

8.3.4 Collection Management Tool (CMT) Interfaces

ICS Compatibility: MAA

The CMT has the interfaces indicated in Table 8-6.

Table 8-6. Collection Management Tool Interfaces

Other ICS or Related Element	Interface	SDD Section
Retrieval Manager	Collection Data Base Modification	3.5.9
Retrieval Manager	Data Exchange	4.6
ISA	Operator Interface	3.5.9
Guide Server	Update Guide Collection Mapping File	3.5.9
Interfaces outside of ICS	Data file ingest	3.5.9

8.3.5 Monitoring and Control Tool (MCT) Interfaces

ICS Compatibility: MAA

The MCT has the interfaces indicated in Table 8-7.

Table 8-7. Monitoring and Control Tool Interfaces

Other ICS or Related Element	Interface	SDD Section
Retrieval Manager	*Retrieval Manager* Management	3.5.10
Retrieval Manager	Data Exchange	4.6
ISA	Operator Interface	3.5.10
Guide Server	Monitor Status and send commands	3.5.10
Guide Indexer	Monitor Status and send commands	3.5.10
Guide Translator	Monitor Status and send commands	3.5.10
OHS Translator	Monitor Status and send commands	3.5.10
UPS Translator	Monitor Status and send commands	3.5.10
Catalogue Translator	Monitor Status and send commands	3.5.10
Interfaces outside of ICS	SSM	3.5.10

8.3.6 Guide Server Interfaces

ICS Compatibility: MAA

The Guide Server has the interfaces indicated in Table 8-8.

Table 8-8. Guide Server Interfaces

Other ICS or Related Element	Interface	SDD Section
ICS Client	IGP Connection	3.5.11
ICS Client	Data Exchange	4.7.1
Guide Indexer	IGP Connection	3.5.11
Guide Translator	IGP Connection	3.5.11
Other *Guide Indexers*	IGP Connection	3.5.11
CMT	Update Collection Mapping File	3.5.11
Interfaces outside of ICS	SSM	3.5.11
All IGP Interfaces	Use of TCP/IP	5.1.1
All IGP Interfaces	HTTP use of TCP	5.1.5
DNS	Directory Services	5.1.6
CEOSnet	Network Connectivity	5.2.1

8.3.7 Guide Indexer Interfaces

ICS Compatibility: MAA

The Guide Indexer has the interfaces indicated in Table 8-9.

Table 8-9. Guide Indexer Interfaces

Other ICS or Related Element	**Interface**	**SDD Section**
Guide Server	Update Guide Metadata Database	3.5.12
Guide Server	Data Exchange	4.7.1
Other *Guide Servers*	Update Guide Metadata Database	3.5.12
Other *Guide Indexers*	Notification of new URL	3.5.12
Interfaces outside of ICS	SSM	3.5.12
All IGP Interfaces	Use of TCP/IP	5.1.1
All IGP Interfaces	HTTP use of TCP	5.1.5
DNS	Directory Services	5.1.6
CEOSnet	Network Connectivity	5.2.1

9. ICS MINIMUM SITE CONFIGURATION

9.1 ICS Minimum Site Compatibility

ICS Compatibility: Mandatory

This section defines the required minimum configuration for a site which wishes to be considered ICS compatible. The minimum site is based on the CEOS policy which anticipates that members will provide the following: collection, product, and explain searches.

To be ICS compatible, a site must at least meet the following:

- Comply with the design in the SDD paragraphs with mandatory ICS Compatibility.
- Provide the elements listed in Section 9.2
- Support the CIP messages listed in Section 9.3
- Provide data listed in Section 9.5
- Support the operations listed in Section 9.6

Note that this section does not define CIP Compatibility, which is defined in the CIP Specification [R3]. The relationship between CIP and ICS compatibility is addressed in Sections 1.1 and 2.1.3 of the SDD.

9.2 ICS Elements for a Minimum Site

ICS Compatibility: Mandatory

For ICS compatibility, a site must provide the ICS elements listed in Table 9-1.

Table 9-1. ICS Elements for a Minimum Site

Element Name
Retrieval Manager
ISA

9.3 CIP Messages for a Minimum Site

ICS Compatibility: Mandatory

For ICS compatibility, a site's Retrieval Manager must support the CIP messages listed in Table 9-2.

Table 9-2. CIP Messages for a Minimum Site

CIP Message Name
initializeRequest
initializeResponse
searchRequest
searchResponse
presentRequest
presentResponse
segmentRequest
deleteResultSetRequest
deleteResultSetResponse
resourceControlRequest
resourceControlResponse
triggerResourceControlRequest
resourceReportRequest
resourceReportResponse
close

9.4 IGP Messages for a Minimum Site

ICS Compatibility: Explanatory

As support of Guide is not required at a minimum site, no IGP messages are required.

9.5 ICS Data for a Minimum Site

ICS Compatibility: Mandatory

For ICS compatibility, a site must provide the data listed in Table 9-3.

Table 9-3. ICS Data by for a Minimum Site

Element Name
CDB
Explain Database
Session Management Data
Use of ICS Valids

9.6 ICS Operations for a Minimum Site

ICS Compatibility: Mandatory

For ICS compatibility, a site must support the operations listed in Table 9-4.

Table 9-4. ICS Operations for a Minimum Site

Operational Activity
Retrieval Manager nominal operations: 24 hours per day, 7 days per week.
Trained *ISA* Staff: operations staff at the site trained as ICS Site Administrators.
ISA Response Time: maximum time of 4 hours for ISA at a site to begin responding to an ICS Event
Perform Collection Maintenance as defined in ICS Collection Manual