Institute of Remote Sensing and Digital Earth
Chinese Academy of Sciences

A Cloud-based Remote Sensing Data Integration and Production System

Lizhe Wang

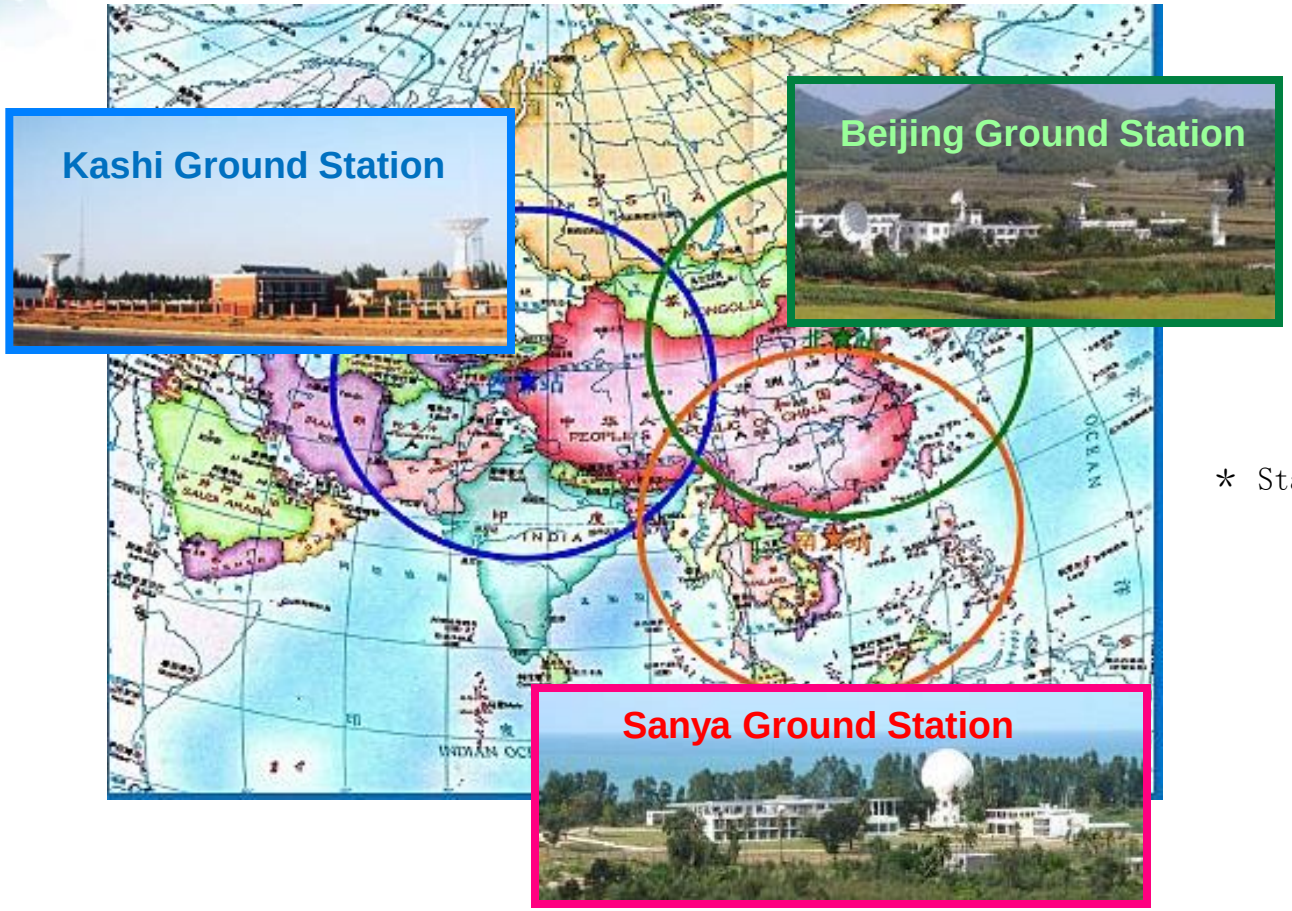
Institute of Remote Sensing and Digital Earth (RADI)
Chinese Academy of Sciences (CAS)
wanglz@radi.ac.cn

Outline

1. Earth Observation in China
2. MCCPS： RS Data Management，Processing
   - System Overview
   - Data Integration across Multiple data Centers
   - Regional to Global Quantitative Inverse Products
   - Dynamic RS Workflow Processing Across Data Centers
   - Global RS Data Products Show
3. Lesson Learned (challenges)
4. PIPSCloud: Towards Cloud processing
   - System Overview：System Infrastructure
   - DS-RS：RS Data as a Service
   - PS-RS：Distributed Processing as a Service
   - VS-RS: RS Virtual Environment as a Service

Earth Observation in China

- Before 2007: *1* Single Remote Sensing Satellite Ground Stations
- Year 2010 : *3* Remote Sensing Satellite Ground Stations



* Stations are Owned by *RADI*

Earth Observation in China

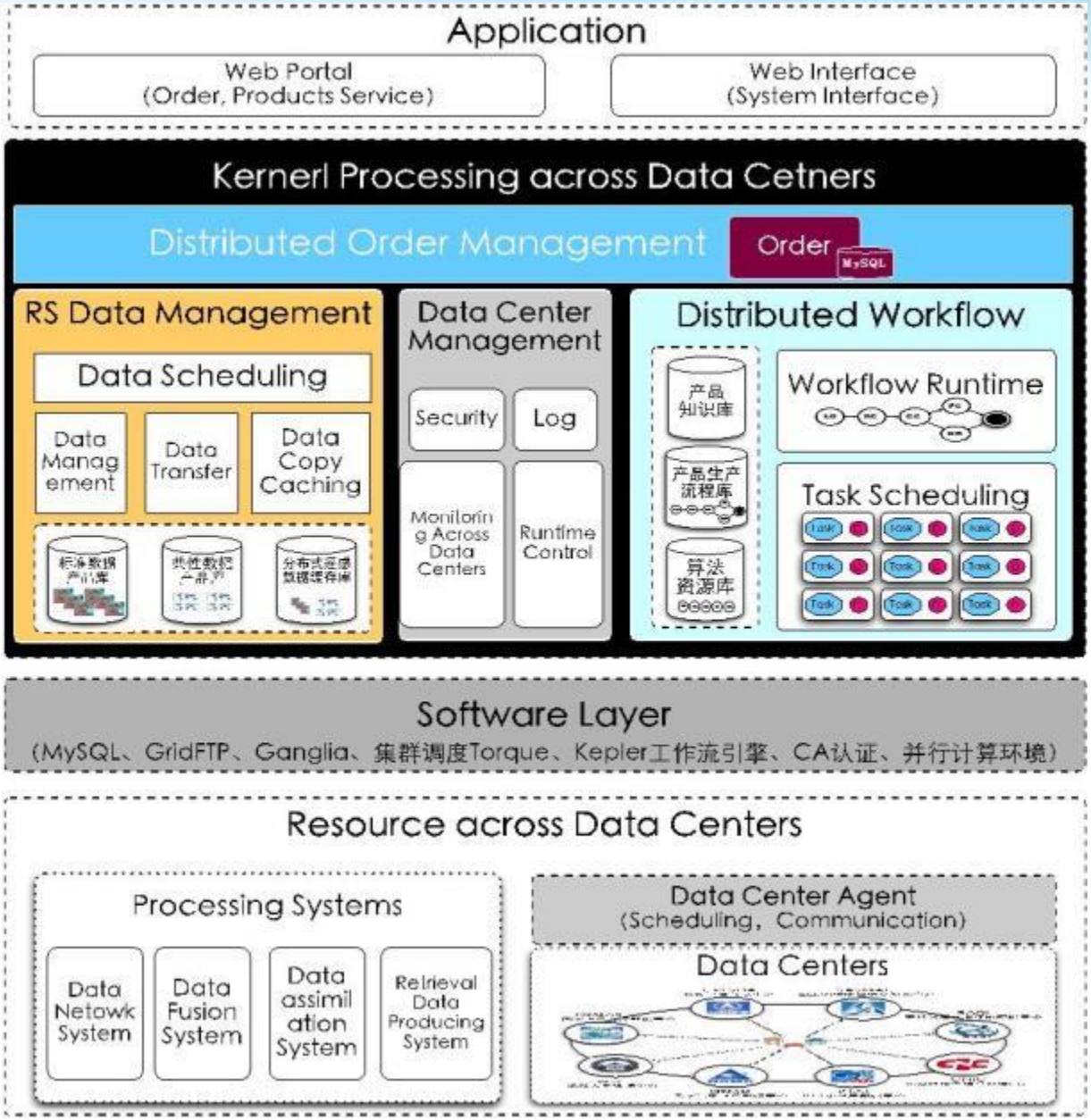
- Before 2007: *1* Single Remote Sensing Satellite Ground Stations
- Year 2010 : *3* Remote Sensing Satellite Ground Stations
- Currently: *6* Remote Sensing Satellite Ground Stations

* Stations are Owned by *RADI*

Earth Observation in China

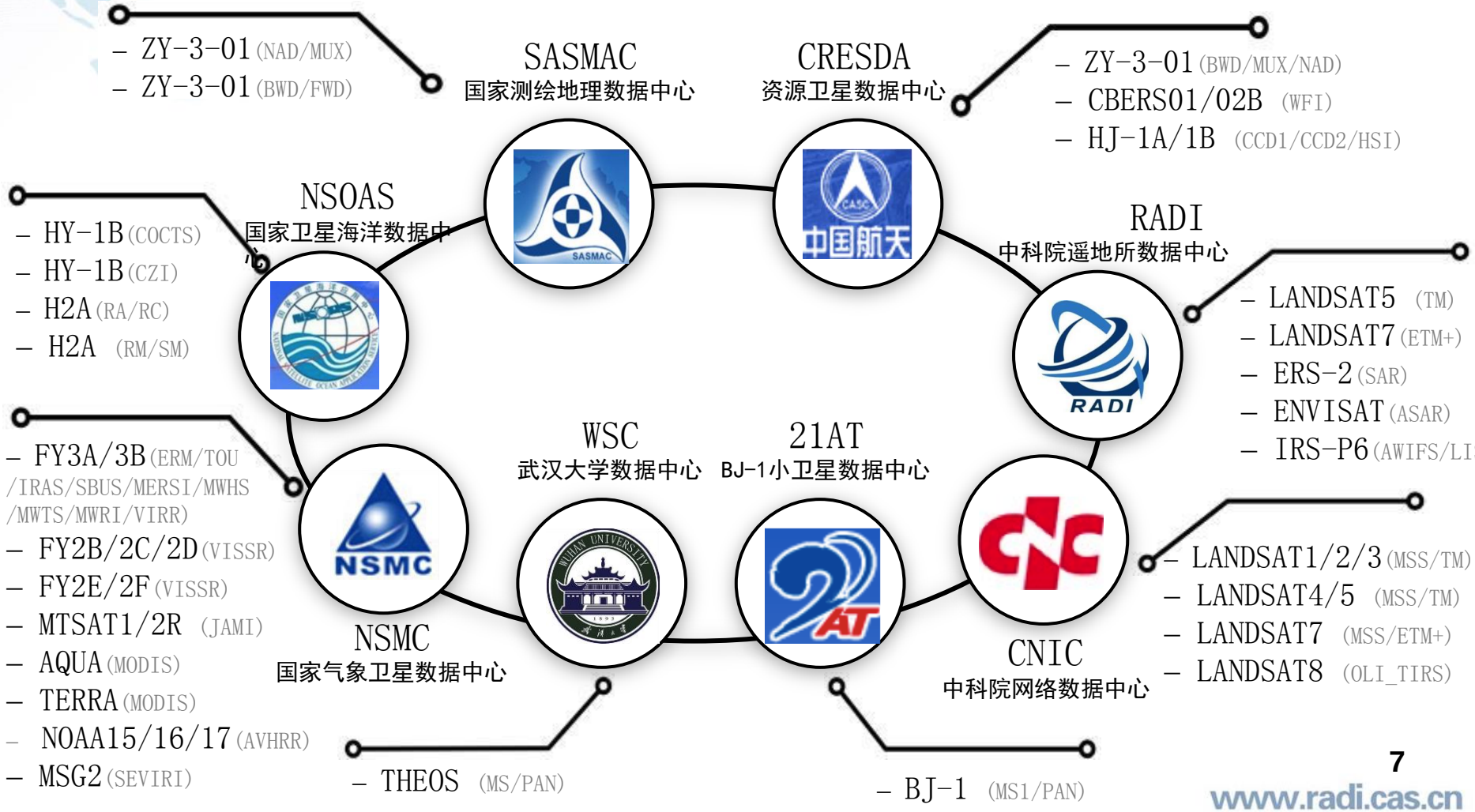
- Member of Landsat Ground Station Operations Working Group
- Acquiring, Processing and Distributing over *3.3 million* scenes of Remote Sensing Satellite datasets, since 1986
- *3* Stations covering the *whole* territory of China and *70%* of Asia
- Acquiring up to *> 500* Terabytes/year, data is soar up
- Downlink data rate up to *> 1 Gbps*
- Acquires data from *12* satellites currently, *50* satellites in 2025
- One of the *Largest* Remote Sensing Ground Station for civil use in the world

MCCPS-System Overview

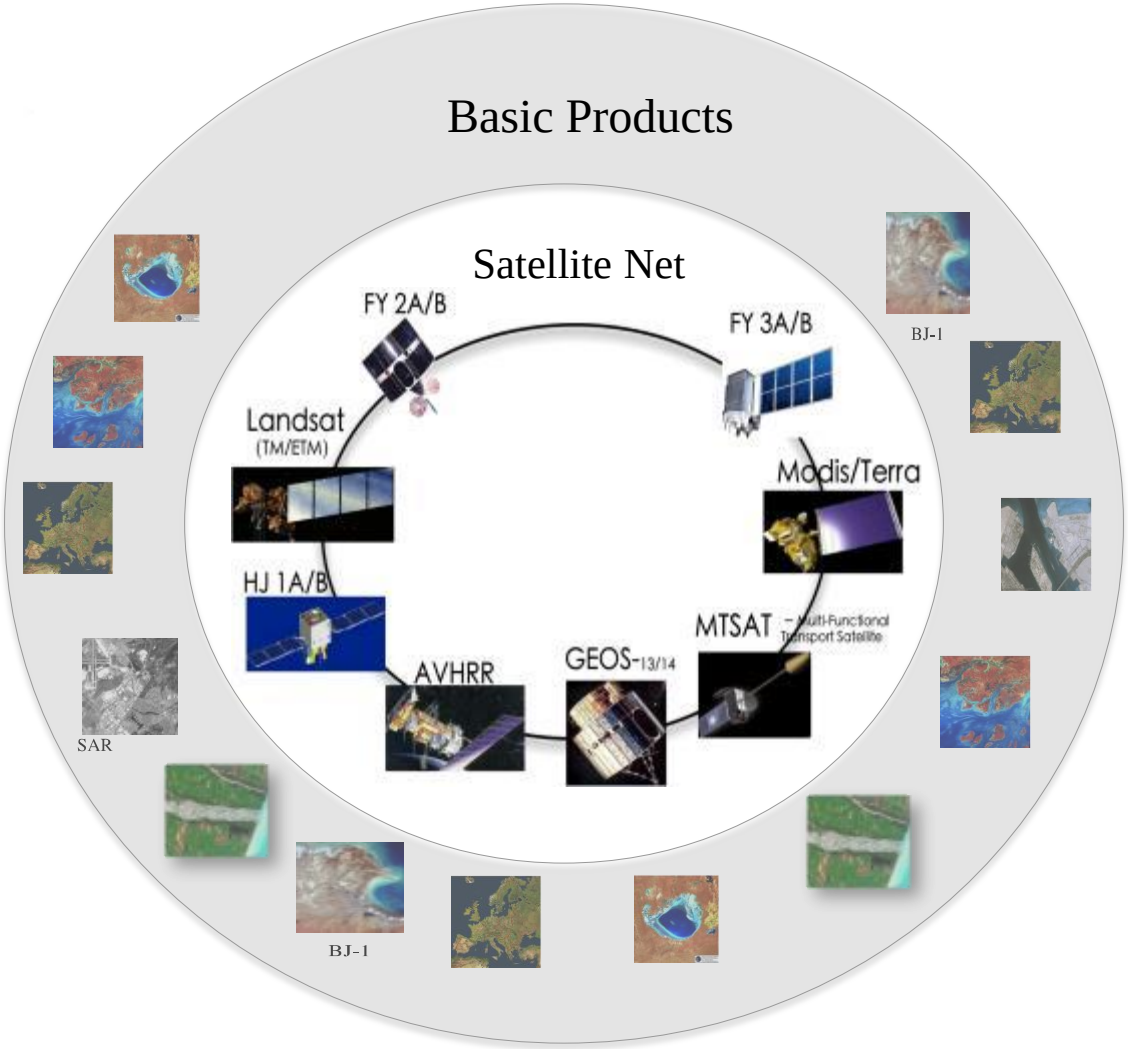


Data Integration across Multiple data Centers

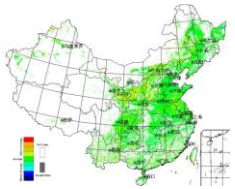
- Integrated *8* Data Centers

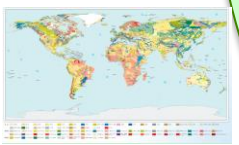


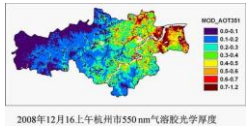
Regional to Global Quantitative Inverse Products

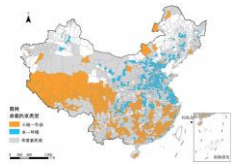


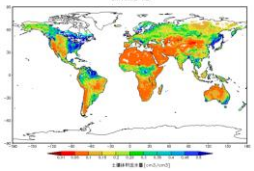
Regional to Global Quantitative Inverse Products

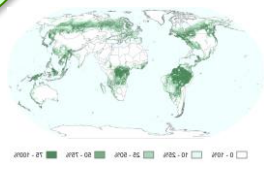










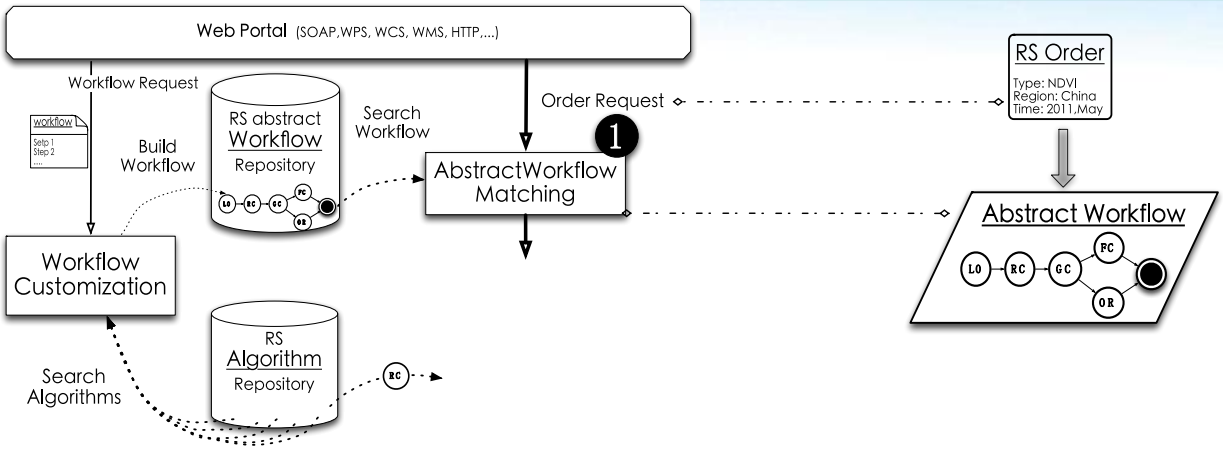


Dynamic RS Workflow Processing Across Data Centers

1

Workflow Match

- Workflow Interpretation
- Abstract→Concrete Workflow Mapping



Dynamic RS Workflow Processing Across Data Centers

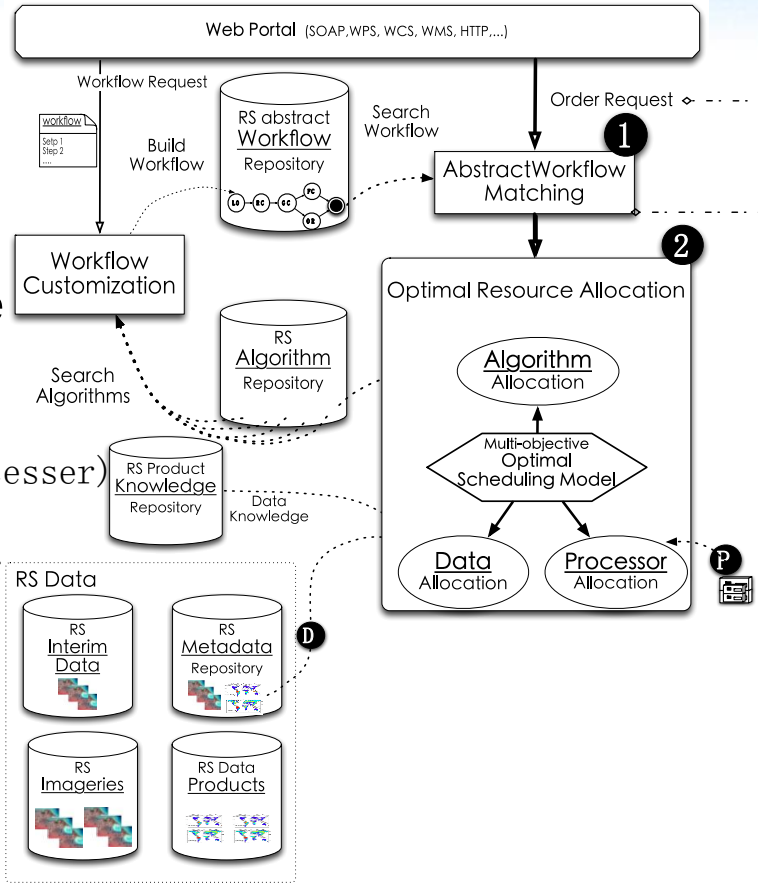
Workflow Match

- Workflow Interpretation
- Abstract Concrete Workflow Mapping

2 Resource Schedule

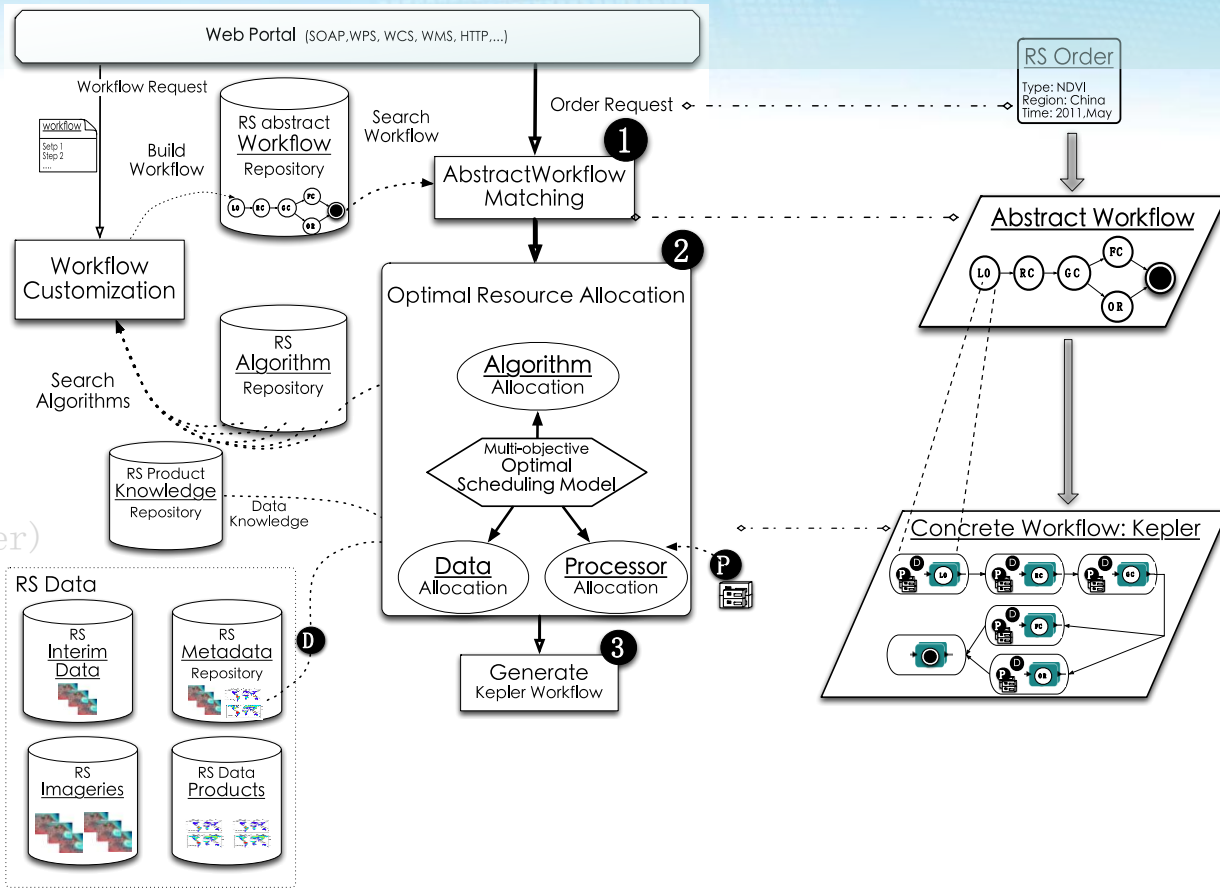
Dynamic Resource Assignment across Data Centers

- Data Center Selection (Processer)
- Algorithm Selection
- Data Discover and Preparing



Regional to Global Quantitative **Inverse** Products

1. Workflow Match
   - Workflow Interpretation
   - Abstract Concrete Workflow Mapping
2. Resource Schedule
   - Dynamic Resource Assignment across Data Centers
     - Data Center Selection (Processer)
     - Algorithm Selection
     - Data Discover and Preparing
3. Create Workflow



Dynamic RS Workflow Processing Across Data Centers

1 Workflow Match

- Workflow Interpretation
- Abstract Concrete Workflow Mapping

2 Resource Schedule

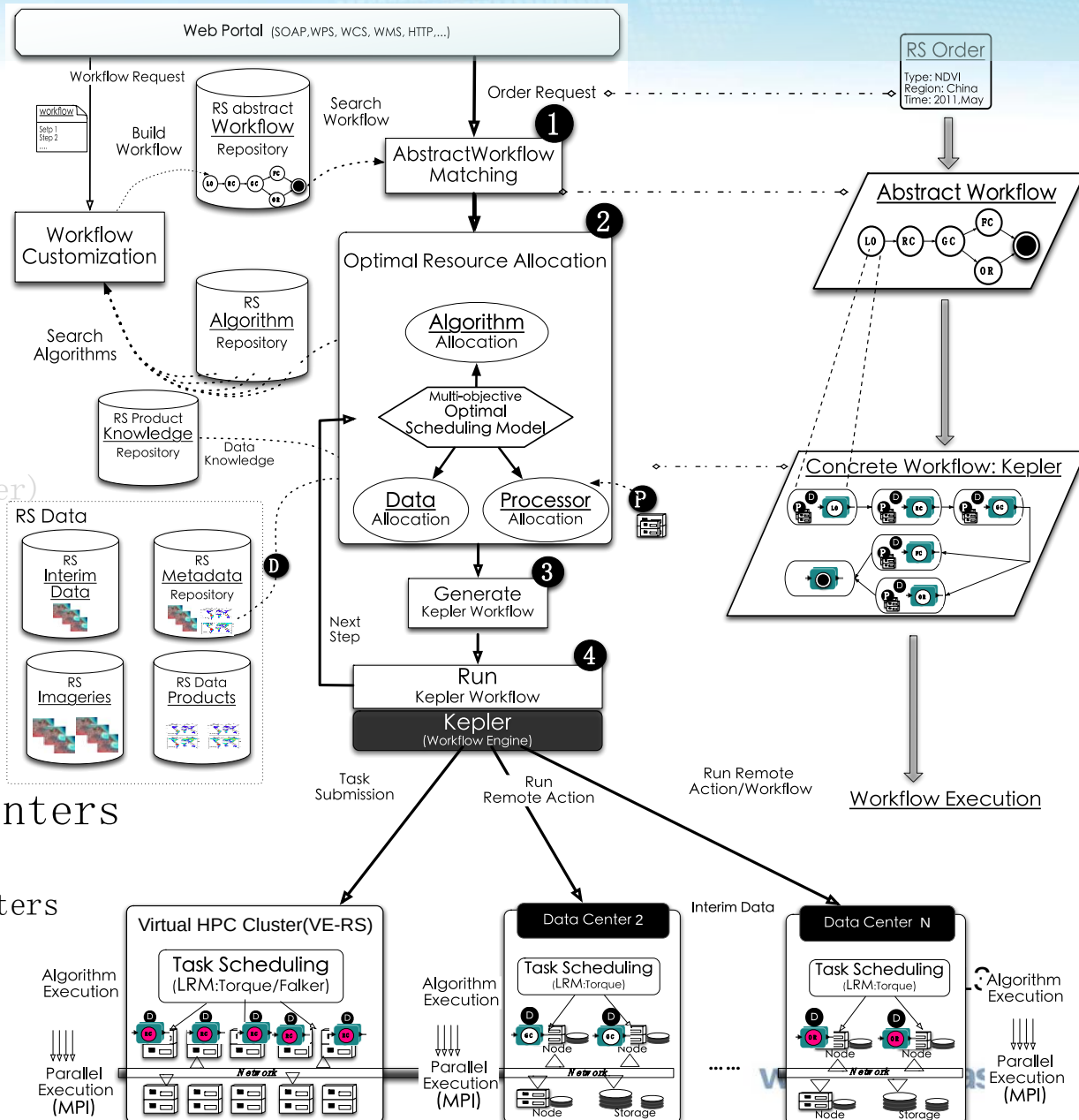
Dynamic Resource Assignment across Data Centers

- Data Center Selection (Processer)
- Algorithm Selection

3 Create Workflow

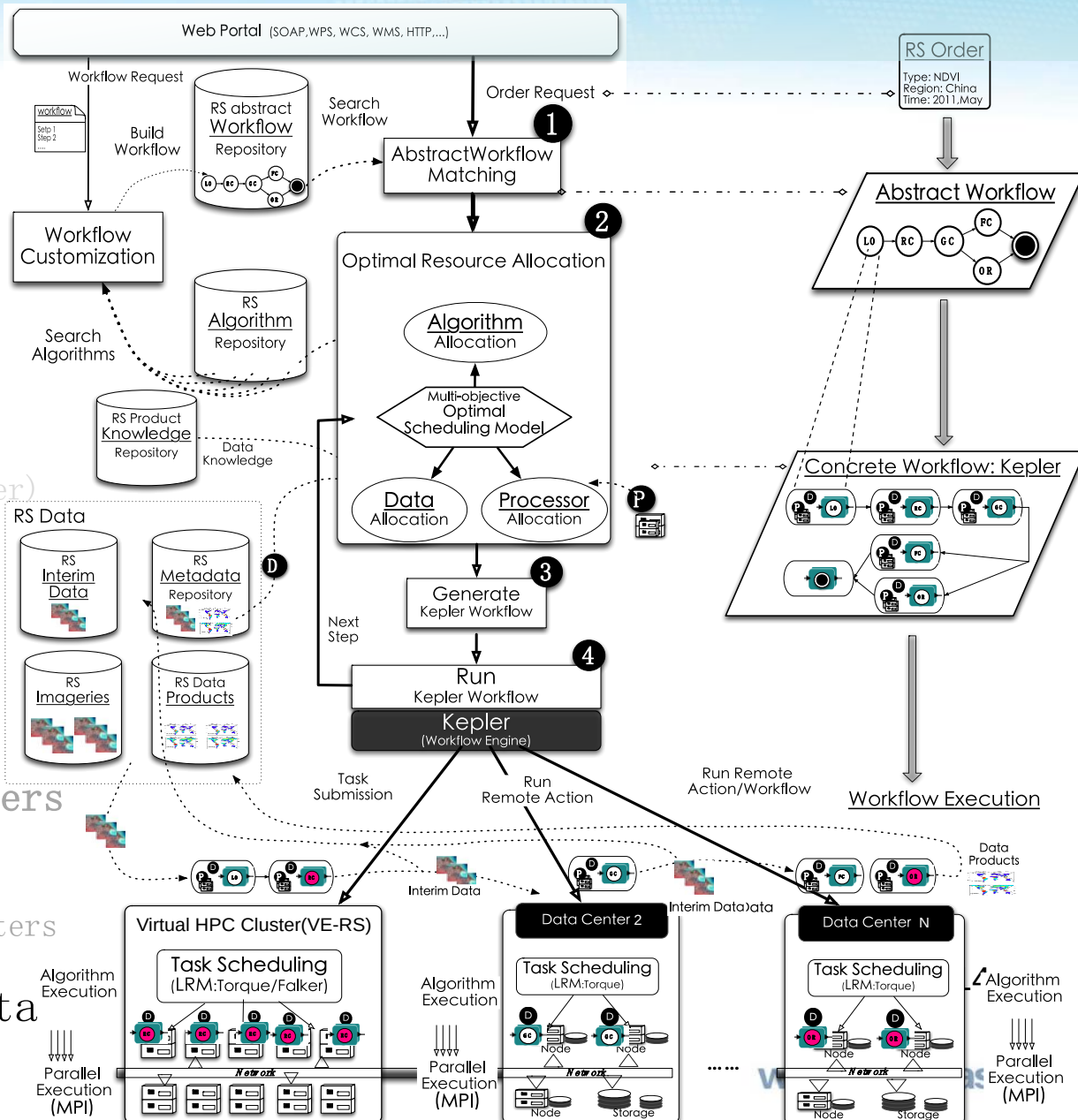
4 Processing Across Centers

- Kepler workflow engin
- Parallelism: among/inside Centers
- Caching: interim data

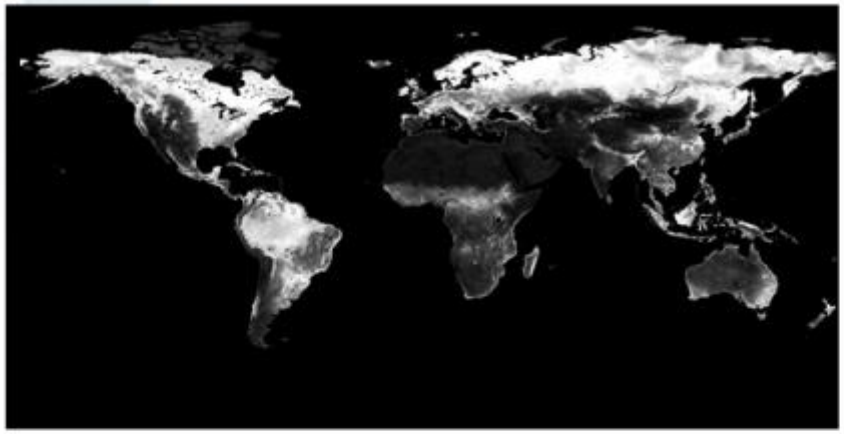


Dynamic RS Workflow Processing Across Data Centers

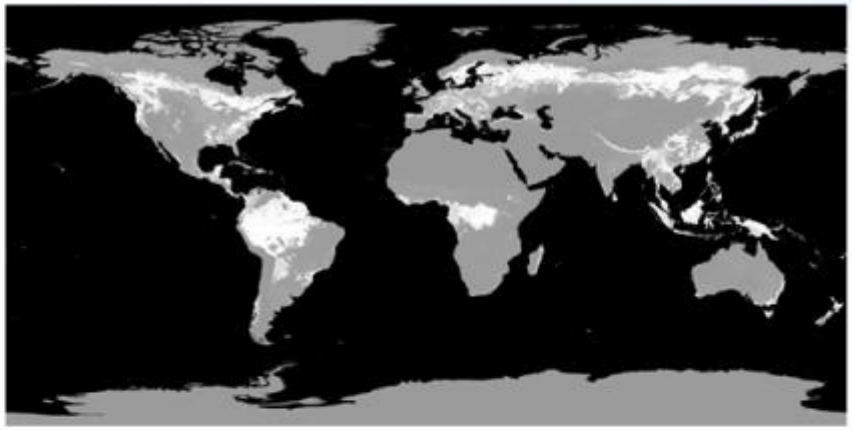
1. Workflow Match
   - Workflow Interpretation
   - Abstract → Concrete Workflow Mapping
2. Resource Schedule
   - Dynamic Resource Assignment across Data Centers
     - Data Center Selection (Processer)
     - Algorithm Selection
3. Create Workflow
4. Processing Across Centers
   - Kepler workflow engin
     - Parallelism: among/inside Centers
5. Caching: interim data



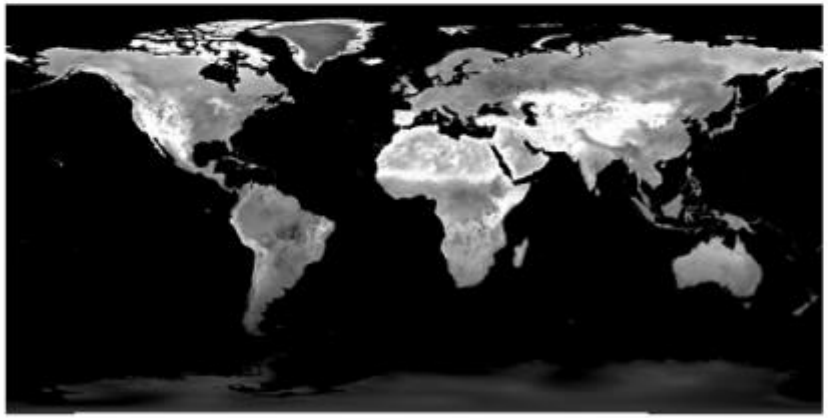
Global RS Data Products Show



1km Soil moisture index



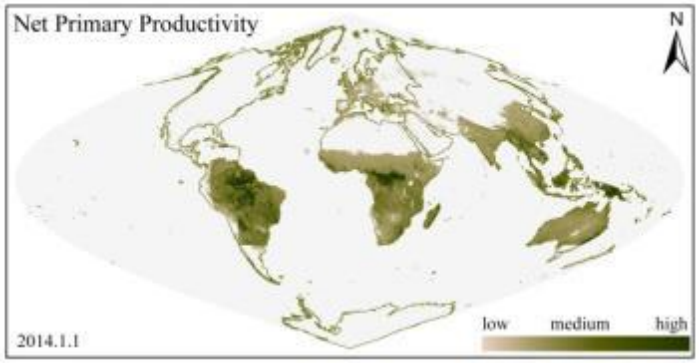
1km Aerodynamic roughness

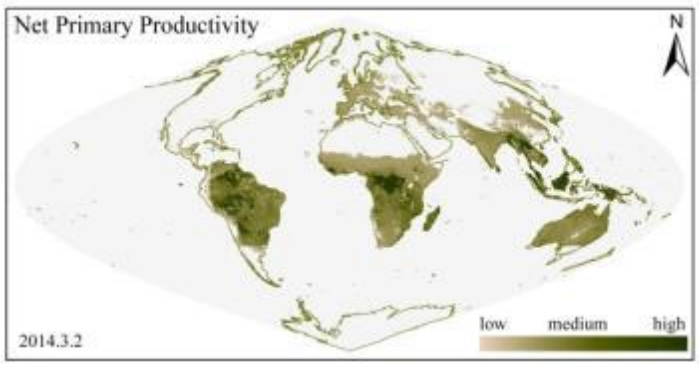


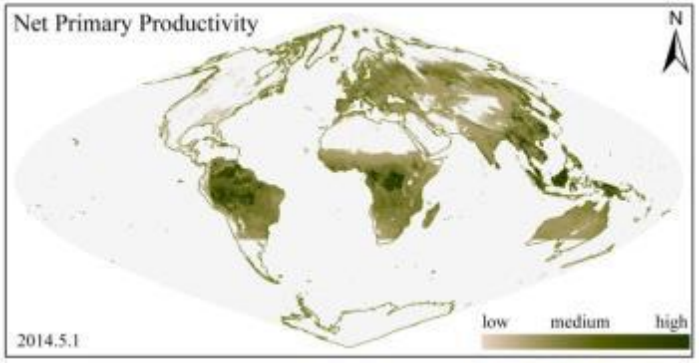
1km Sensible heat flux

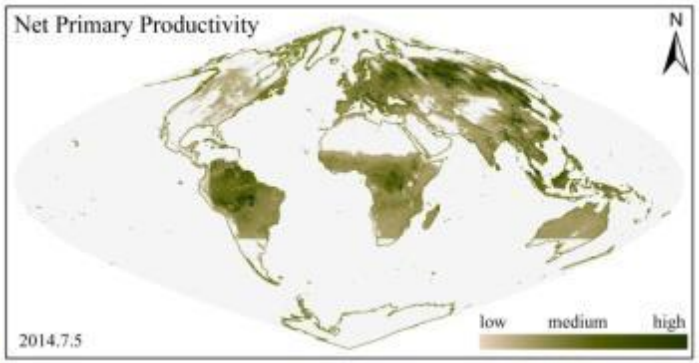
1km Latent heat flux

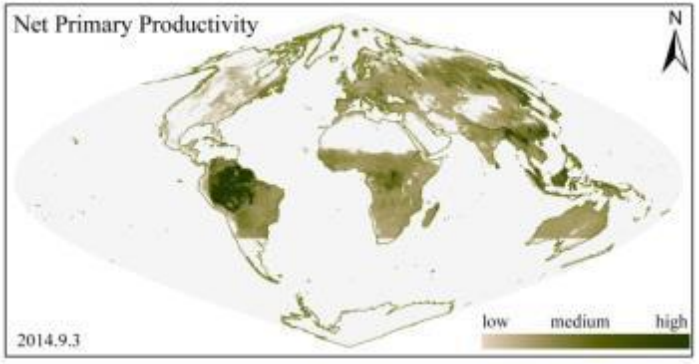
Global RS Data Products Show

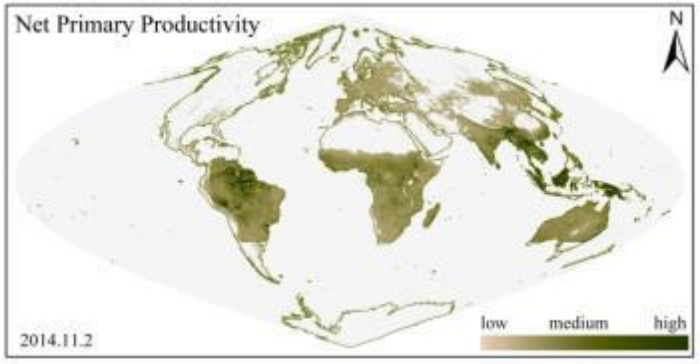












NPP is clearly tied to seasonal change, with productivity peaking in each hemispheres summer. The Boreal Forests of Canada and Russia, for example, experience high productivity in July and then a slow decline through fall and winter. Year-round, tropical forests in South America, Africa, Southeast Asia, and Indonesia have high productivity, not surprising with the abundant sunlight, warmth, and rainfall. This was well adaptive in natural animal growing.

Lesson Learned: Challenges

❶ RS Big Data Management

- *Massive Data* in Explosive Growth: TB to PB
- *Multi-sources*: Multi-sensor, Multi-resolution, Multi-Platform
- *Distribution* (Different *formats*): Across data centers or organizations
- *Diversities* of RS Data (Data *Relationships*):

❷ Large-scale RS Data Processing

- Large-region covered : Regional to Global
- Time Serial Analysis
- Collaborative across data centers

❸ Application-Specific HPC Platform

- *Limited Resources*: Computing nodes and Storages
- *Lack of RS-oriented Platform*: Lack RS software, not RS application-specific
- *Difficult and expensive to manage HPC Infrostructure*

pipsCloud — A Remote Sensing Cloud Platform

❶ *Cloud Services*

RS Data Processing Service

- On-line RS Processing
- Seasonal Auto-Producing
- Collaborative Processing across Data Centers

RS Data Service

- RS Big Data Service
- RS Metadata Service
- Thematic Data Subscription
- Data Harvest and Integration
- Data Tracing

pipsCLOUD

RS-specific HPC Platform as a service

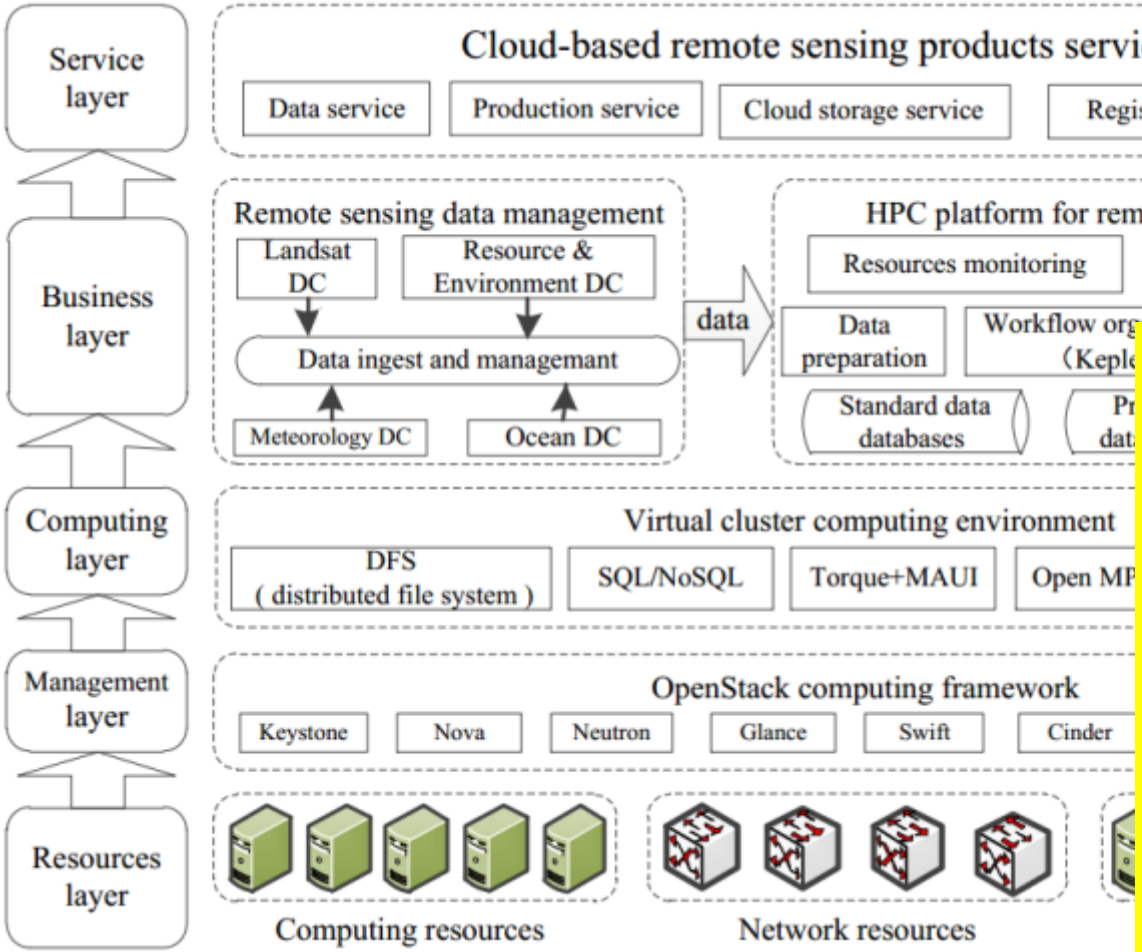
- On-demand HPC service
- *RS Soft-ware* Auto-deploy
- Mount Virtual RS Data Catalog and Storage
- VM and PM Provisioning

RS Processing System as a Service

- PIPS System Service
- On-demand System Deploy
- RS *IDE* Service integrate algorithm , workflow

pipsCloud — System Infrastructure

System Infrastructure of pipsCloud

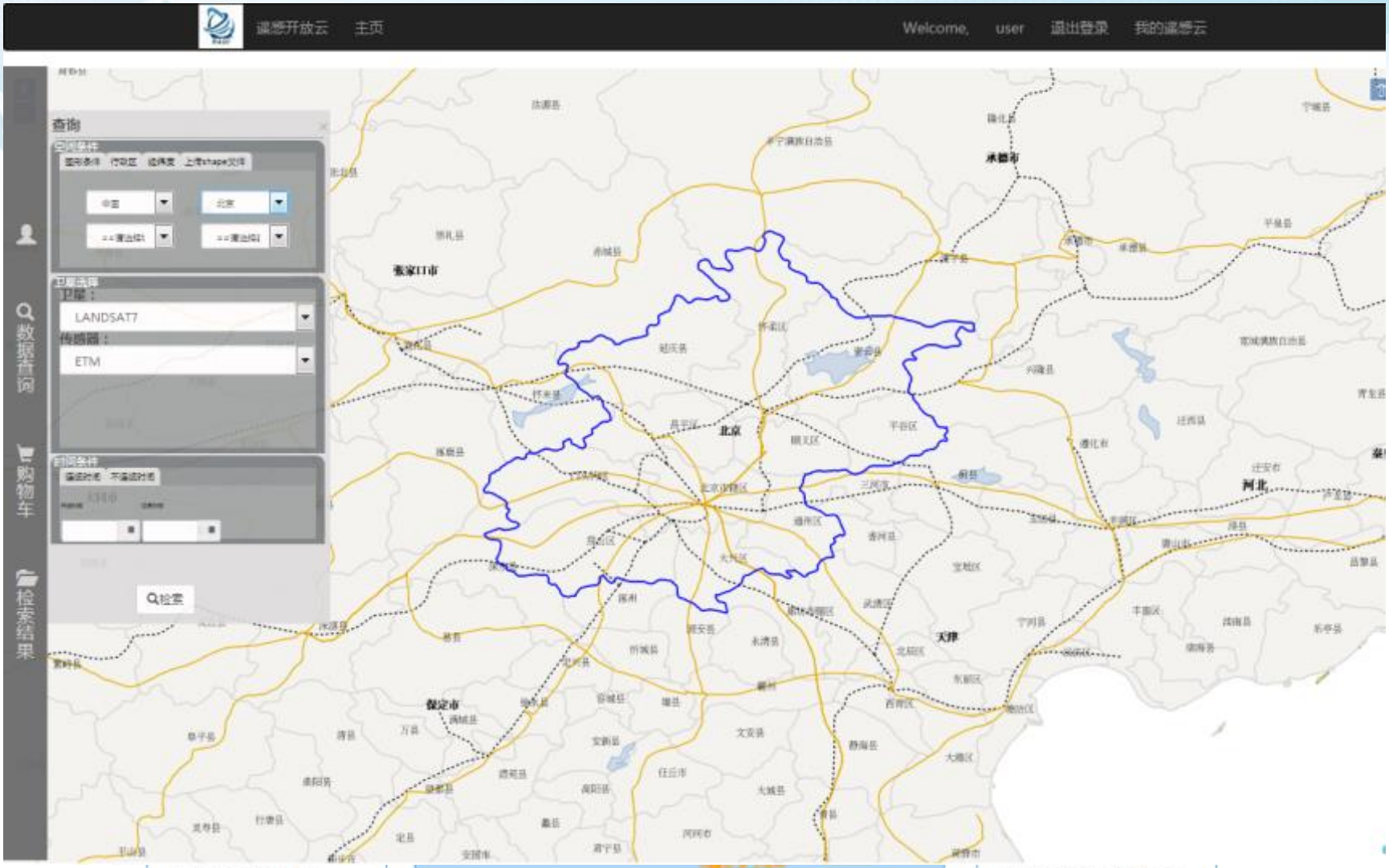


❶ DS-RS (Data Service)
- RS Data & Products Services
- Metadata Service
- RS Data Subscription

❷ PS-RS(Processing Service)
- RS Data Processing Services
- Subscribed auto-processing

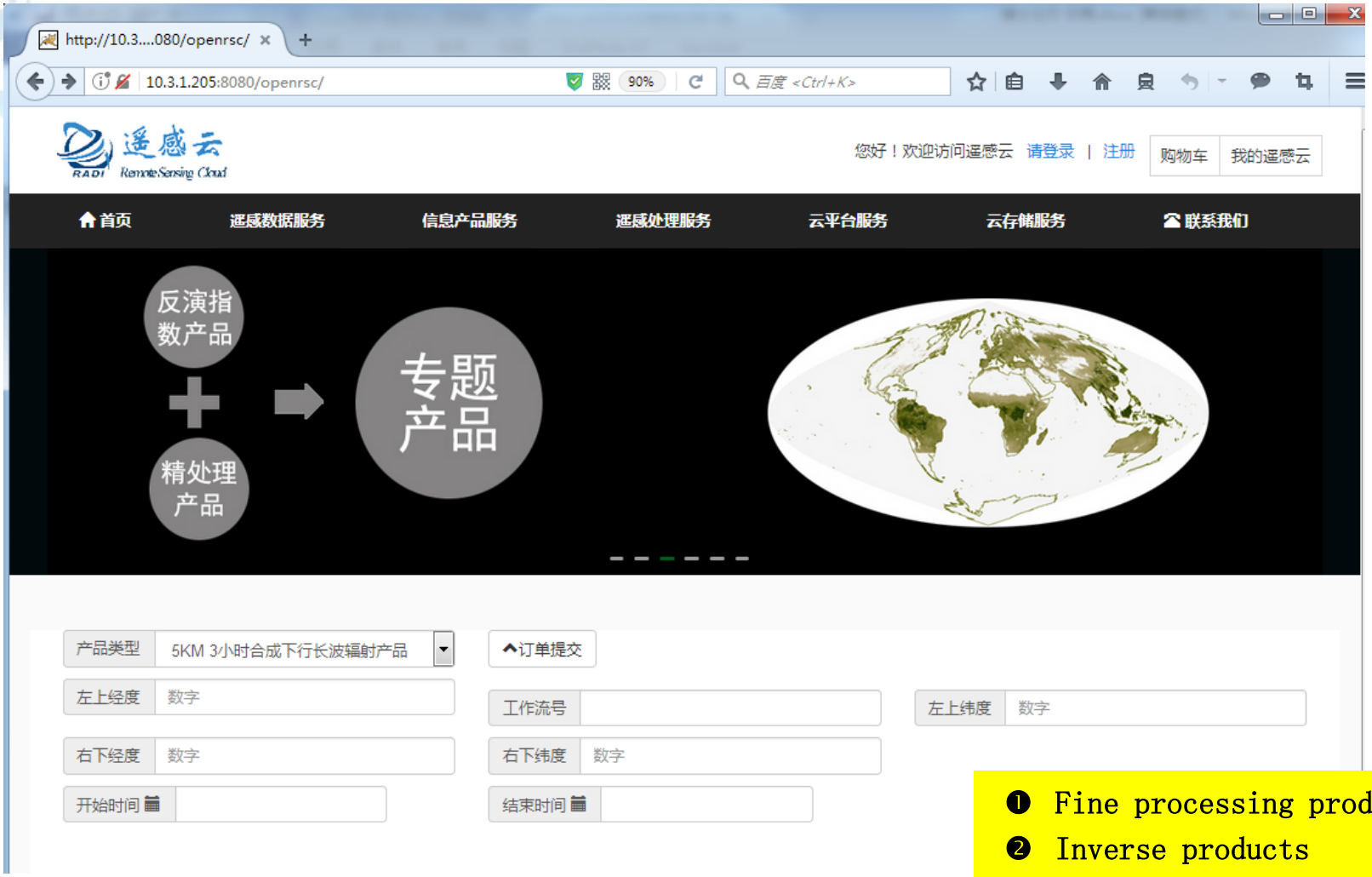
❸ VS-RS (Virtual pips System)
- PIPS Processing System
- Algorithm Integration
- Workflow customization and Integration

DS-RS— DAAS Portal



DS-RS— ProAAS Portal

- *Online-Processing and Production*



❶ Fine processing products
❷ Inverse products
❸ Thematic products

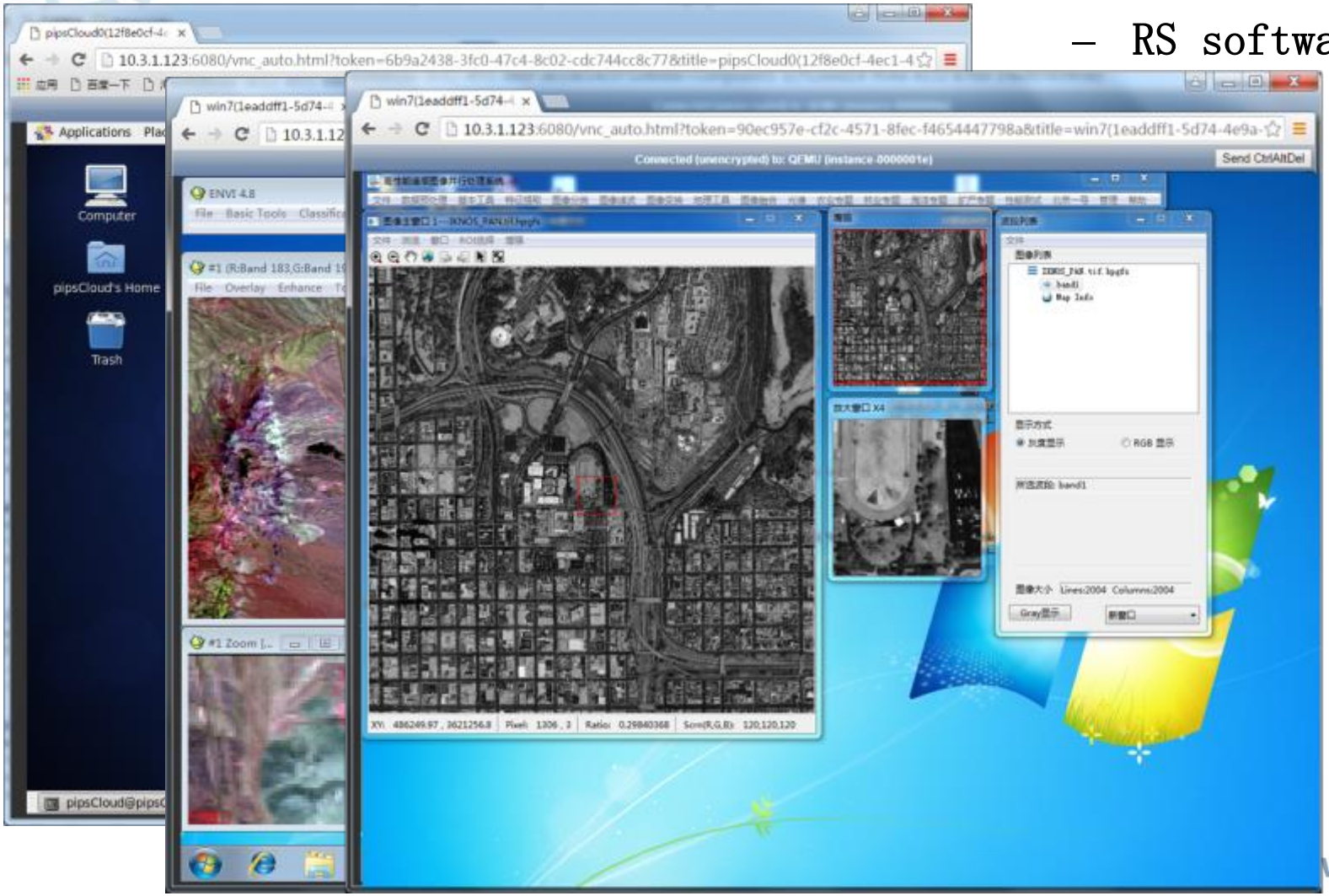
pipsCloud — Virtual RS HPC Environment

❹ *Virtual RS HPC Environment*

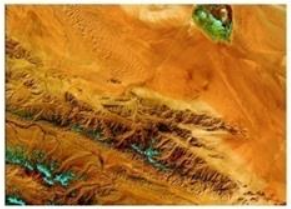
❶ Auto-deployment

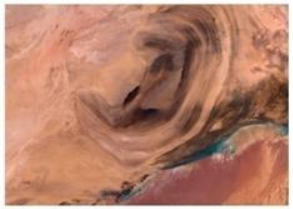
– HPC software,

– RS software



Thanks!






Institute of Remote Sensing and Digital Earth
Chinese Academy of Sciences

Add：No.9 Dengzhuang South Road,Haidian District,Beijing 100094,China
Tel：86-10-82178008 Fax：86-10-82178009
E-mail：office@radi.ac.cn
Web：www.radi.cas.cn