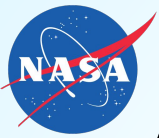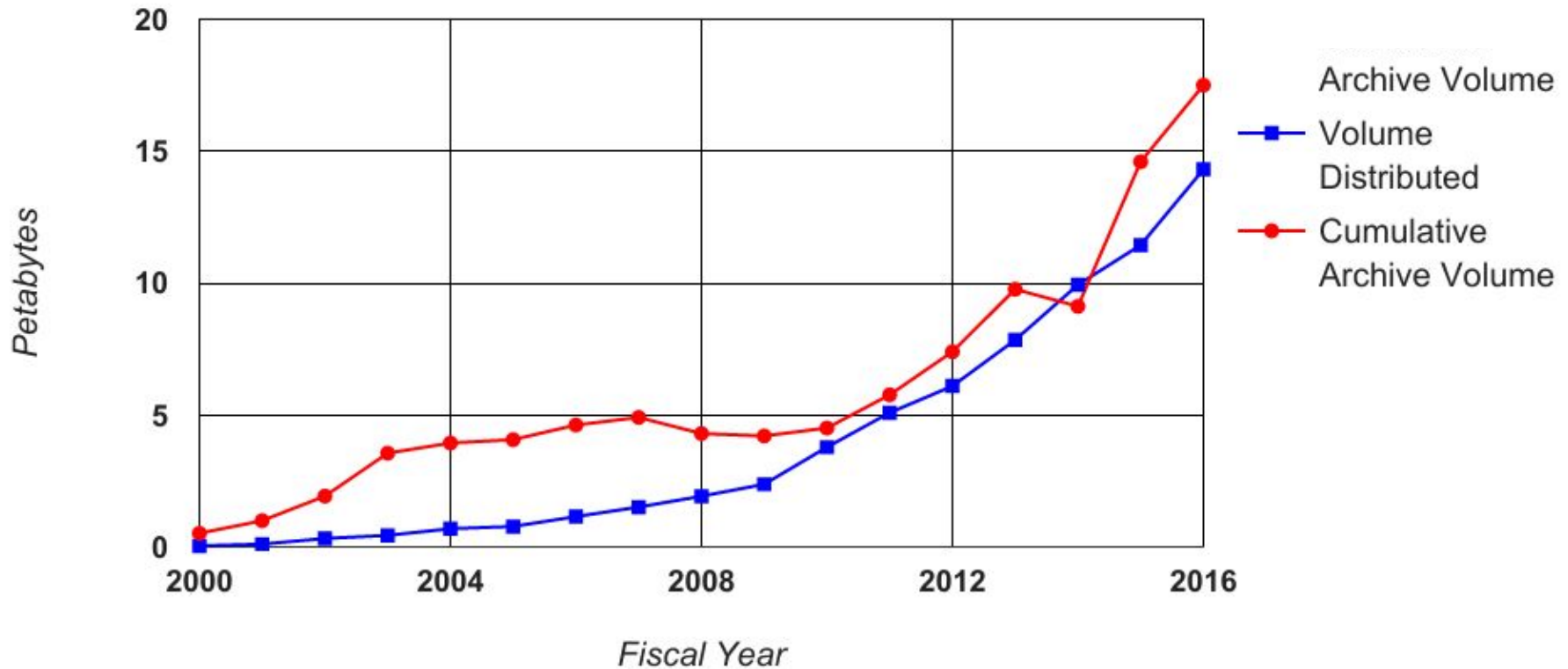# NASA EOSDIS
Cloud Prototype Systems

Chris Lynnes, Mark McInterney, Katie Baynes

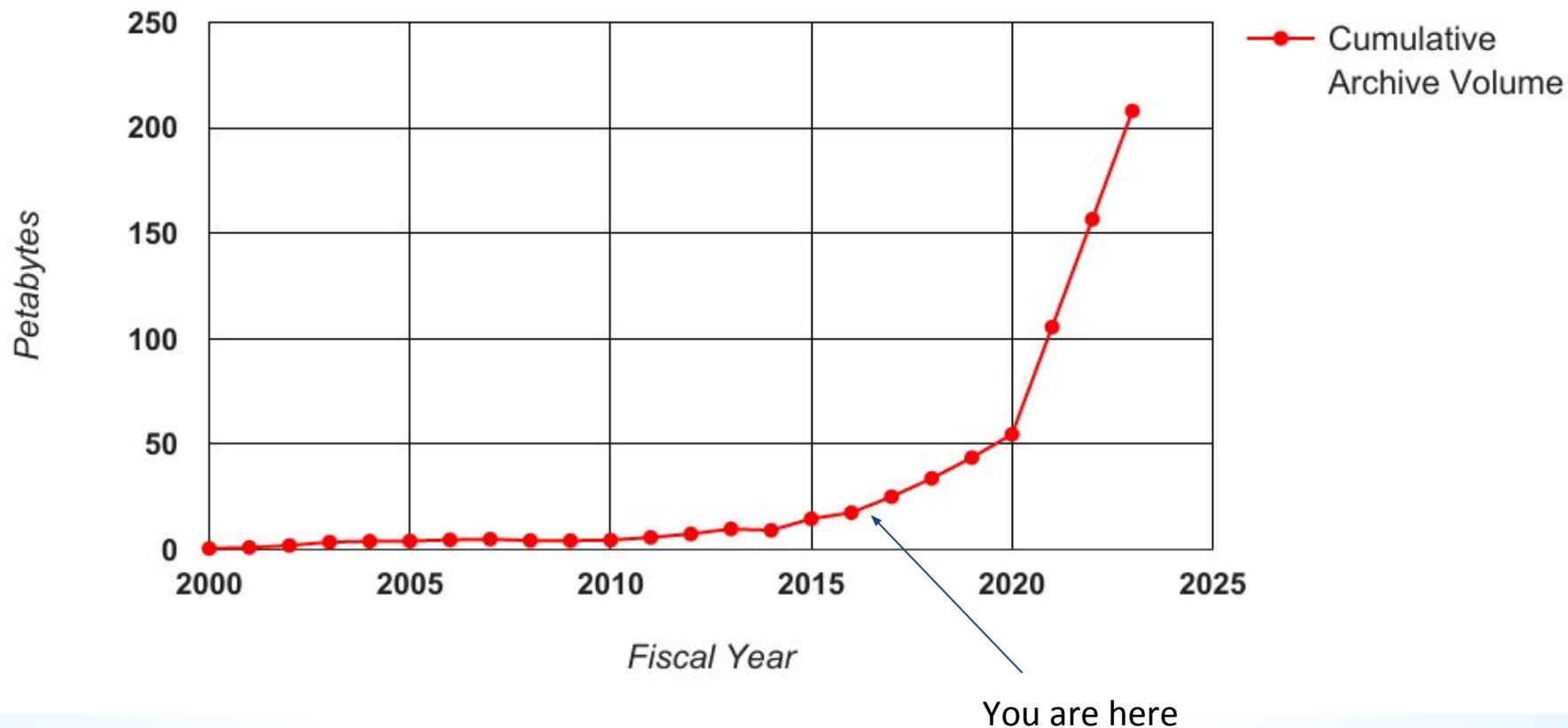NASA

Annual distribution is on the same order of magnitude as the total archive volume

And our archive is scheduled to grow dramatically

You are here

If these high volumes are a challenge for robust, experienced data centers to manage...

...what about the science end users?

Conceptual "Data Close to Compute"

Large Volume Data Storage
Datasets stored in cloud object storage

EOSDIS Applications
Applications using cloud compute, storage, and services

Scalable Compute
Provision, use, and terminate as needed

Community Applications
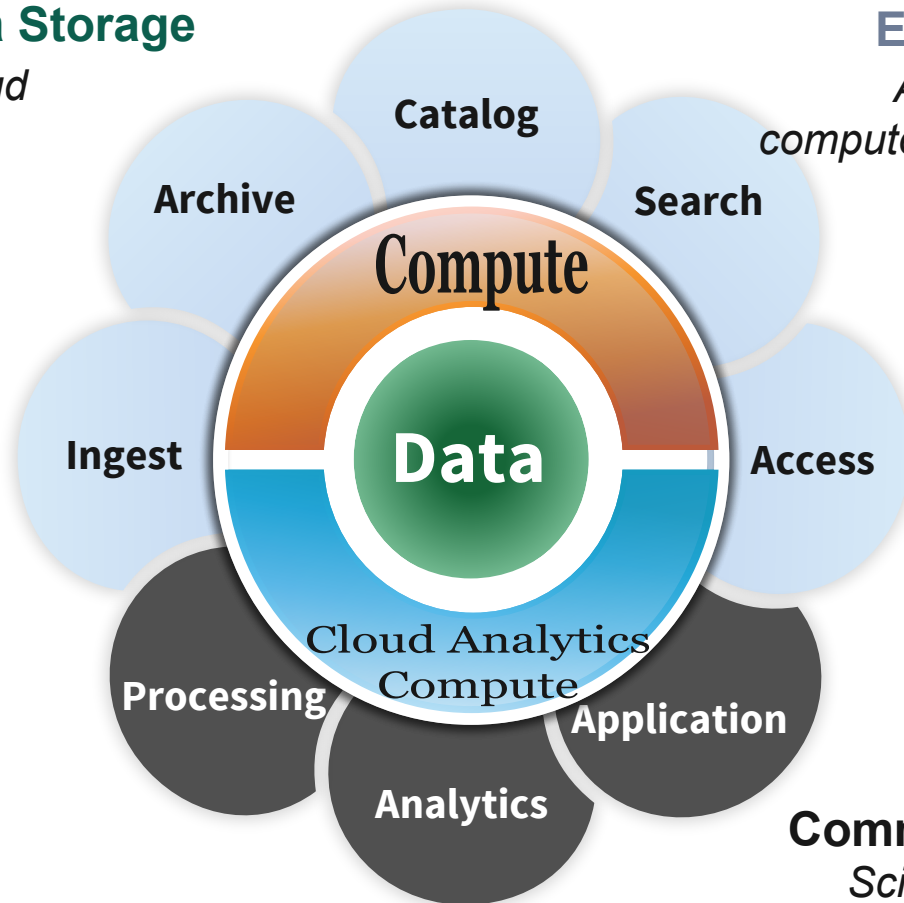*Science community brings algorithms to the data.*

Conceptual "Data Close to Compute"

Large Volume Data Storage
Datasets stored in cloud object storage

EOSDIS Applications
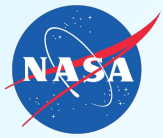*Applications using cloud compute, storage, and services*

Scalable Compute
Provision, use, and terminate as needed

Community Applications
Science community brings algorithms to the data.

Common Metadata Repository

Cumulus

Earthdata Search

Compute

Data

Global Imagery Browse Service

Web Object Storage

Cloud Analytics Compute

Getting Ready for NISAR

NEXUS Giovanni CAaaS*

AppEEARS
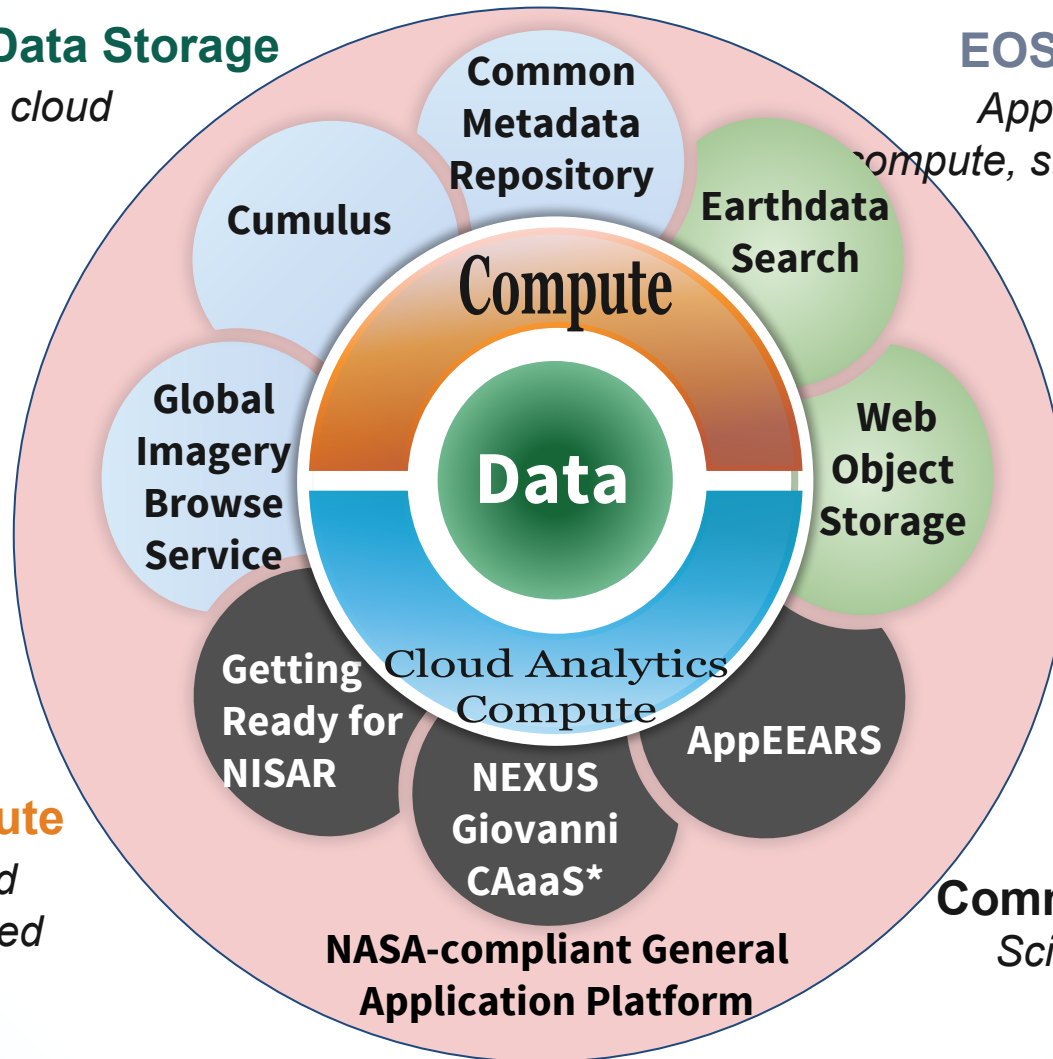
NASA-compliant General Application Platform

*CAaaS = Climate Analytics as a Service

Specific Benefits

Avoidance of expensive hardware purchase

First public EOSDIS application in the cloud

High-bandwidth edge server for radar data

Acceleration of OPeNDAP access to data

Approval to operate publicly

Acceleration of end-user analysis

On-demand regeneration of virtual products

Improvement of ingest performance

Serverless architecture demonstration

Common Metadata Repository

Cumulus

Earthdata Search

Global Imagery Browse Service

Compute

Data

Web Object Storage

Getting Ready for NISAR

NEXUS Giovanni CAaaS

Cloud Analytics Compute

AppEEARS

NASA-compliant General Application Platform

NASA-compliant General Application Platform: "Compliance-as-a-Service"

Software-as-a-Service

Compliance-as-a-Service
security controls, Authorization to Operate
governance
procurement
reliability and availability

Platform-as-a-Service

Infrastructure-as-a-Service

EOSDIS
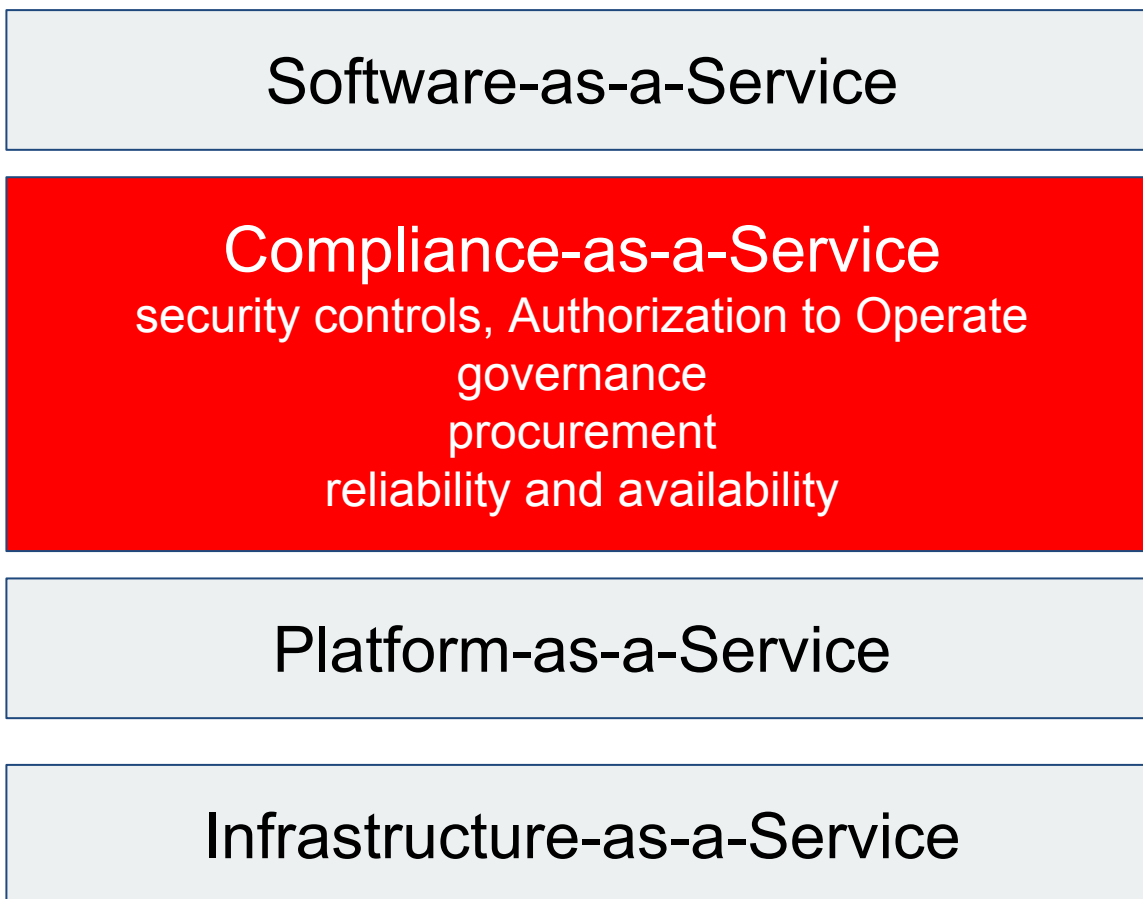
Key Cloud Benefits for Science Users

Immediate direct access to data
Immediate availability of compute

Lower Barrier to Entry

Dynamic scaling for bursting
Variety of CPU and GPU

Scalable Performance

Purchase compute as needed
Spot market for cost control
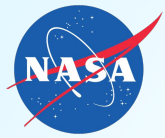No need for local copy of data

Cost

Experiment replication enabled
Sharing of complex analysis workflows
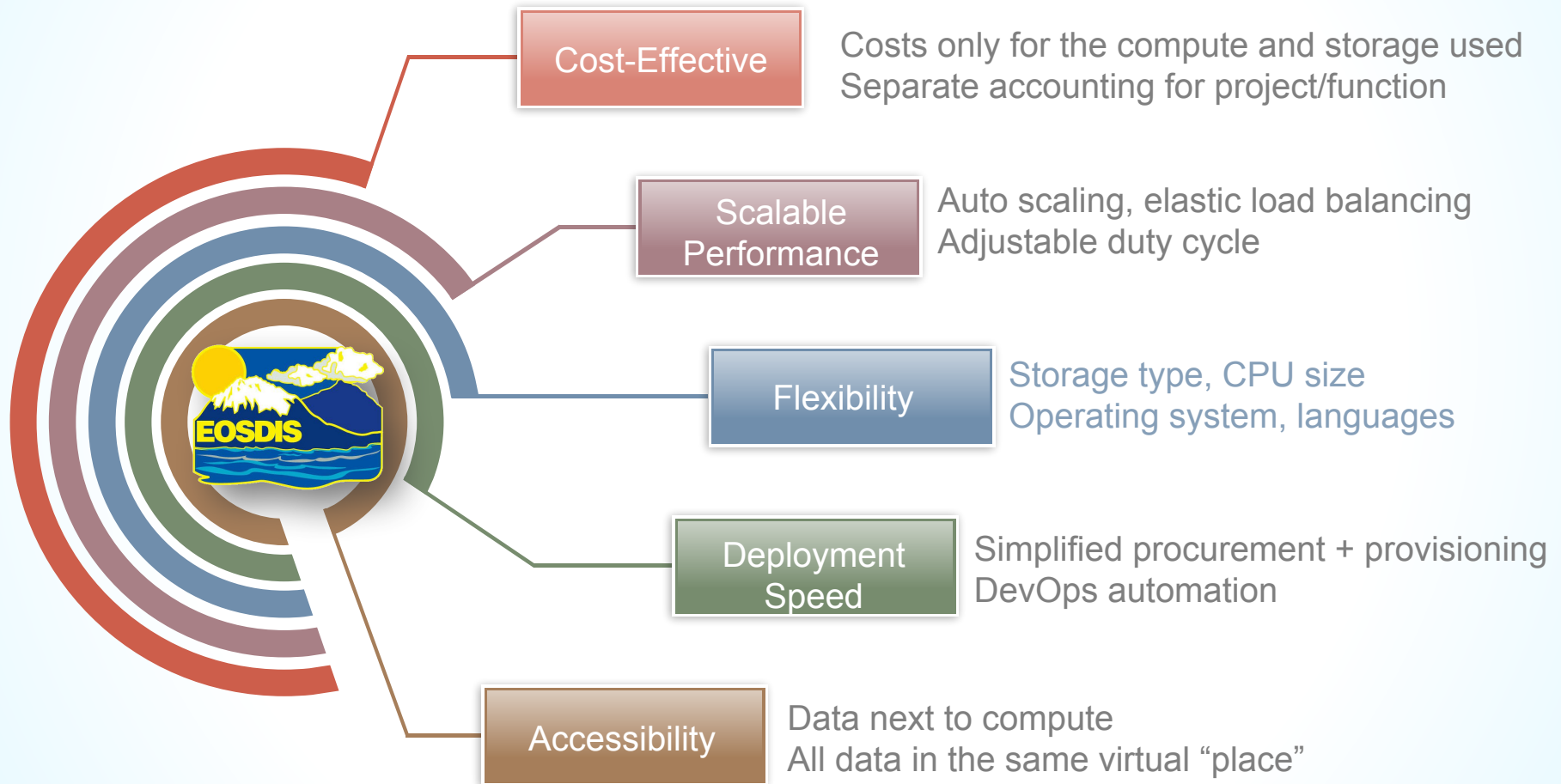
Repeatability & Sharing

Decrease in startup time for projects
Analysis at scale for anyone
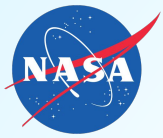Rapid iteration enabled

Promoting Science Discovery

Key Cloud Benefits for Data Systems (EOSDIS)

Cost-Effective
Costs only for the compute and storage used
Separate accounting for project/function

Scalable Performance
Auto scaling, elastic load balancing
Adjustable duty cycle

Flexibility
Storage type, CPU size
Operating system, languages

Deployment Speed
Simplified procurement + provisioning
DevOps automation

Accessibility
Data next to compute
All data in the same virtual "place"

Decision Considerations

High level decision consideration for individual project prototypes and capabilities to operationalize into AWS (commercial cloud)
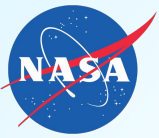
01 Cost — Is AWS (commercial cloud) affordable?.

02 IT Security — Is NASA IT Security compliance and tactical operations achievable in AWS (commercial cloud)?

03 Performance — Is performance equal to or better than current on-premises solutions?

04 Operational — Can we operate "Operationally" in AWS (commercial cloud), technical and business?

Key Lessons - So Far

- Cloud-native re-architecture is preferable to forklift/lift-and-shift

- Savings from Dev / Test environments

  - Full-workload testing: $N_{ops} = N_{perftest}$

  - On-Premise:
    $N_{ops} + N_{perftest} = 2N \times 24hr \times 7d = 336$ N-hr

  - Cloud:
    $(N_{ops} \times 24hr \times 7d) + (N_{perftest} \times 10hr \times 5d) = 218$ N-hr = 35% savings*

- Serverless architectures look promising

  - Reduce needed code down to business logic only