Hyperspectral Imaging in Cloud Computing Environment

Sergey Popov
Image Processing System Institute RAS, Samara, Russia

Outline

• Issues of Traditional Approach
• Core principles of big data technology
• Cloud Architecture for Hyperspectral Imaging
• Distributed Hyperspectral Image (HSI)
• Processing Organization
Traditional Approach
Hyperspectral Imaging by Hadoop

- Row data move to HDFS
- Partitioning without responsiveness to inner structure of HSI
- Poor accordance MapReduce paradigm to basic HSI processing procedures
Core principles of big data technology

• Divide and Conquer
• Move Code to Data
• No UPDATE (Data Are Forever!)

• In-Memory Processing
• Lazy Evaluation
Architecture for Hyperspectral Imaging in the Cloud

- Data as a Service
- In-Memory Store
- Distributed Hyperspectral Image (DHSI)
- Federation of Frame-Services
- Coordination via Hierarchical Namespace
- Move Code to Data
- No UPDATE when Processing
Distributed In-Memory Store

- Hypercube in Frame Service construct at Startup
- Plug-In Modules to load data in different formats

- connectForData
- connectForTask

- Processes Executor

- HDFS
- HDF5
- GeoTIFF
Data Partitioning Strategies

Spatial dimension X

Spatial dimension Y

Spectral dimension

Spectral domain

Spatial domain
Data Partitioning Strategies

(a) 3x3 kernel computation split among 2 PE
(b) intro overlap to avoid inter-process communication in 3x3 kernel computation
Data Partitioning Strategies

- 2D Spatial Decomposition with Overlap
- Overlap Size
- Fault Tolerance
Data Partitioning Strategies
Coordination via Hierarchical Namespace

- Distributed coordination service ZooKeeper
Data Processing Strategies

• Task – One Operation
• Job – Graph Operations
Data Processing Strategies

- **Task** – One Operation
- **Job** – Graph Operations

![Diagram](image-url)

- Op1
- Op2
- Op3
- Op4
- Op5
- Op6
Data Processing Strategies

- Task – One Operation
- Job – Graph Operations
Prototype Experiments
Thank you for the attention!