Computer Technology Makes the Service of GeoScience Processing System More Convenient

Shengyang Li, Wanfeng Zhang, Zhiwen Liu

Technology and Engineering Center for Space Utilization, CAS

March 1, 2016

Contents

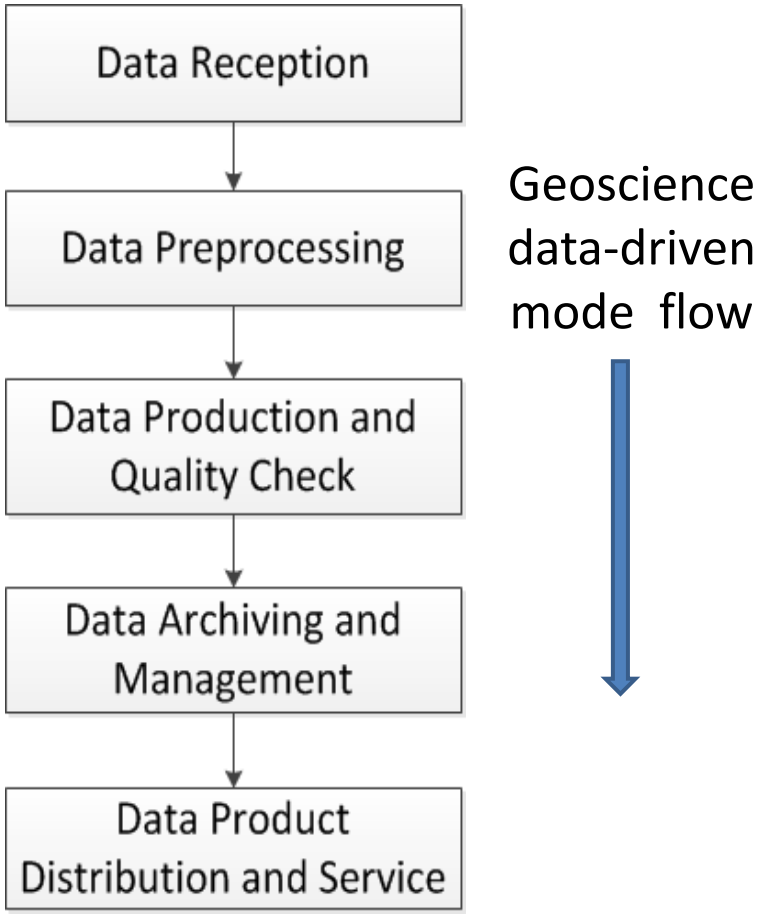
Background

- With the developments of various sensors in **earth observation (EO)** and increasing abilities of acquiring geoscience data, **data centers (DCs)** which possessed different types of data have been formed.

- **Challenge such as data and algorithms automatic selection need to be overcome** when research demands from numerous remote sensing application fields become more and more diversified.

Background

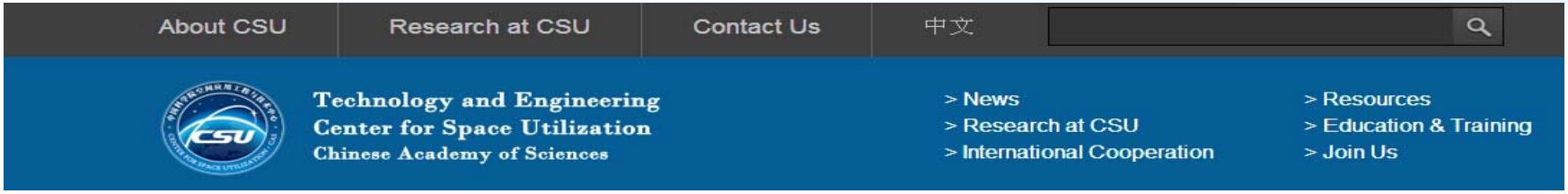
- The main **processing mode of Remote Sensing Products (RSP) in Geoscience is data-driven mode**.
- For specific remote sensing satellite mission, the data-driven mode is generally used in a uniform way of data processing.
- Processing and providing service of Multi-source data product would require on-demand mode.



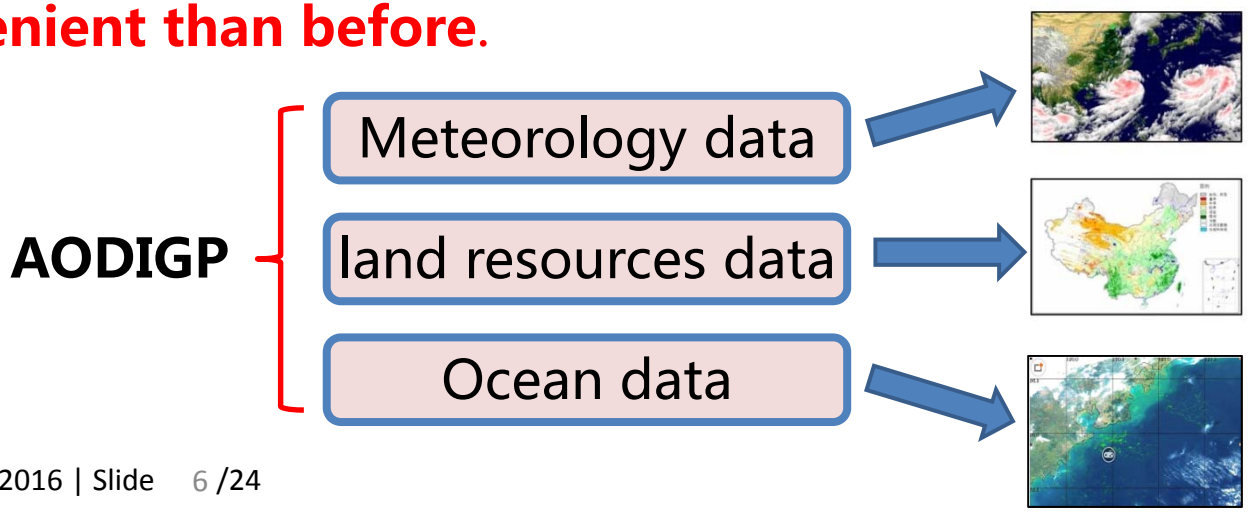
Background

- The European Space Agency (ESA) Grid Processing on Demand (G-POD) is a generic GRID-based operational environment.
  - Easy operation
  - Automatic processing
  - On-demand service based on diversified product types
    - MERIS Mosaic PR/COM
    - ASAR GM Antarctica Mosaics
    - GlobTemperature Daytime/Nighttime
- **On-demands mode** with convenient operation interface and multiple remote sensing data source become more and more indispensable.

Background

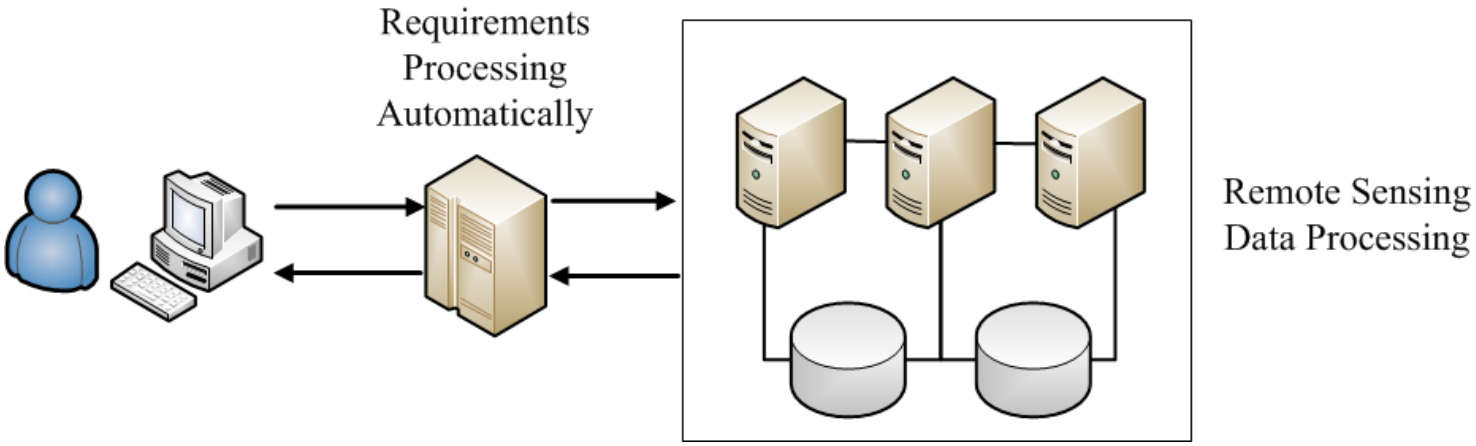


- We presented an Architecture On-demand for Data Intensive Geoscience Product (**AODIGP**) using computer technology.
- AODIGP' s purpose is analyzing user' s demands, matching computing and data source, and scheduling workflow **automatically**.
- Diverse geoscience products processing would be **more convenient than before**.



Background

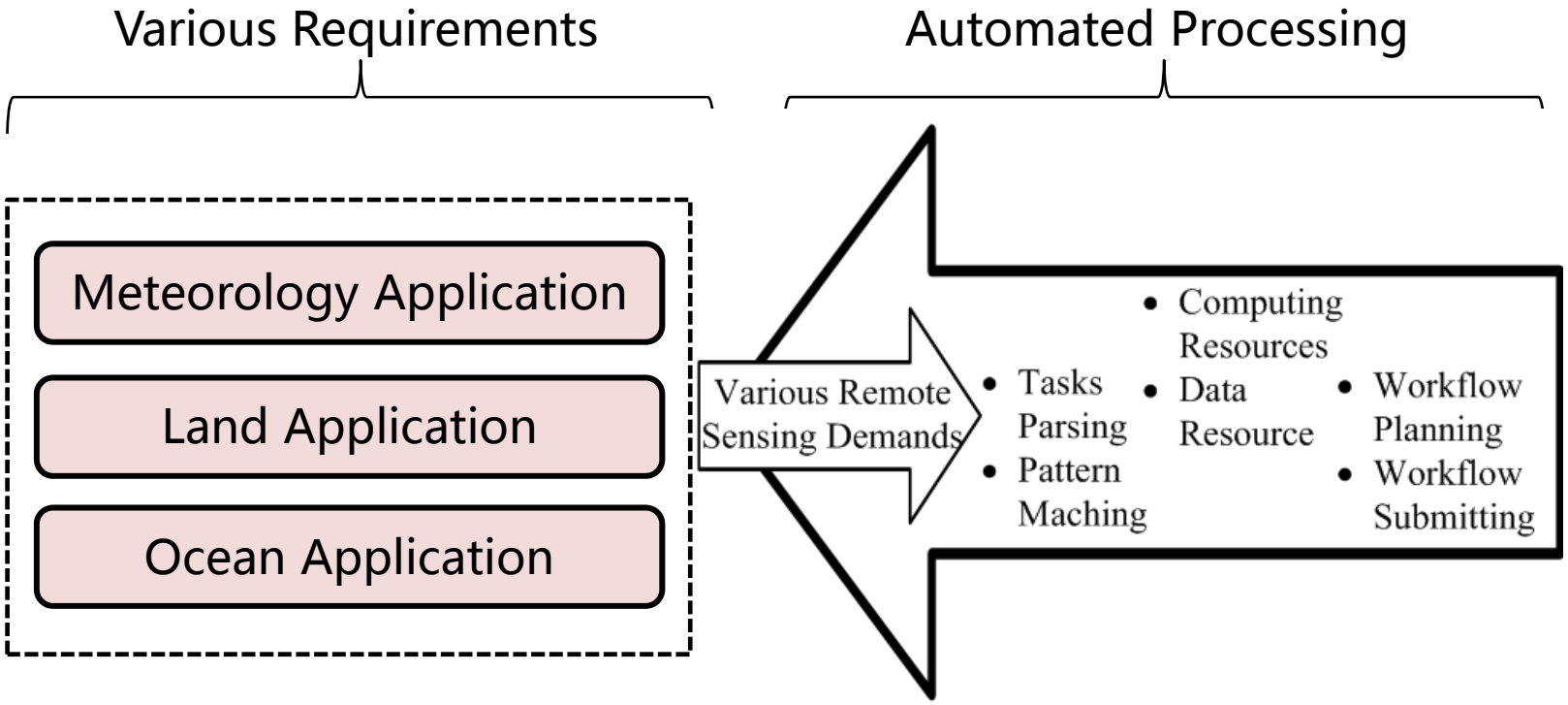
- AODIGP simplified variant geoscience product tasks into a **normalized task set**.
- **Empirical knowledge** about Remote Sensing Product in Common Properties (RSPCP) would be collected in knowledge base.



Contents

On-demand Architecture

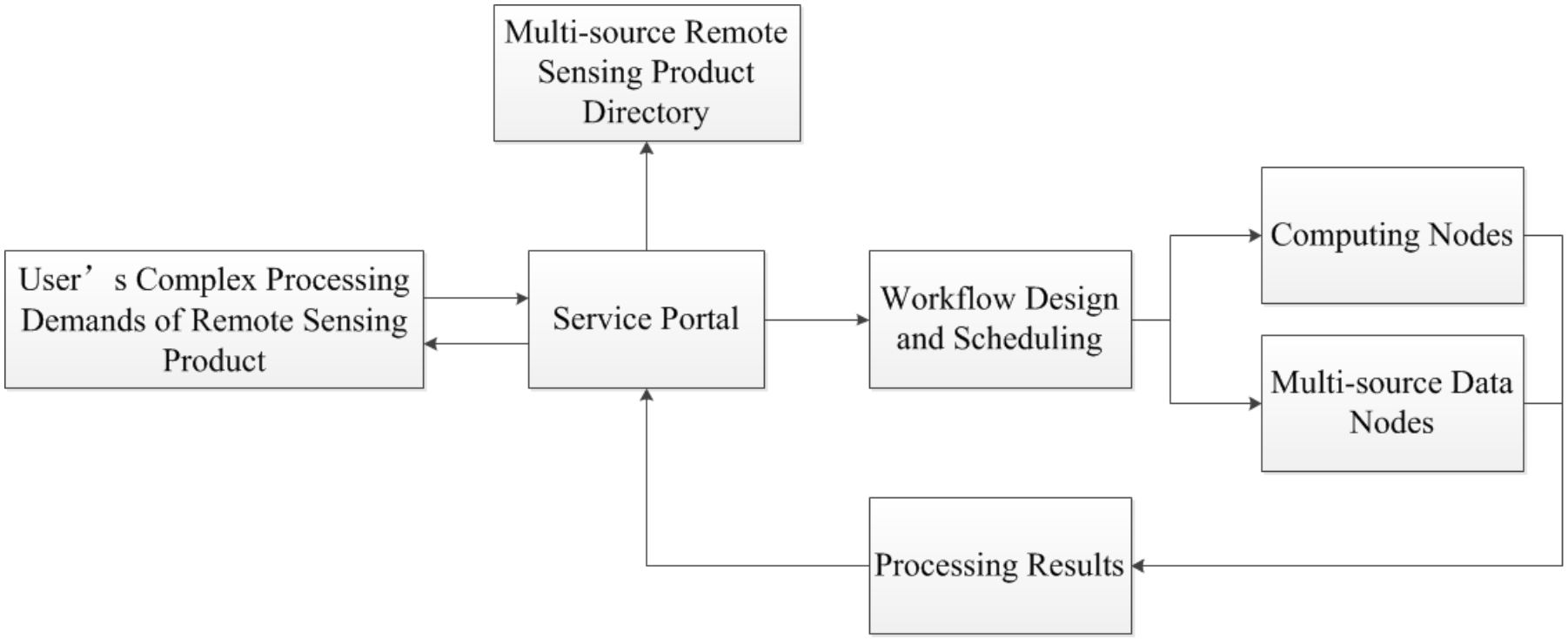
- Automated geoscience product processing in AODIGP is realized through requirements parsing, workflow matching, and data resource searching.



The flowchart of on-demand processing towards multiple geoscience applications

On-demand Architecture

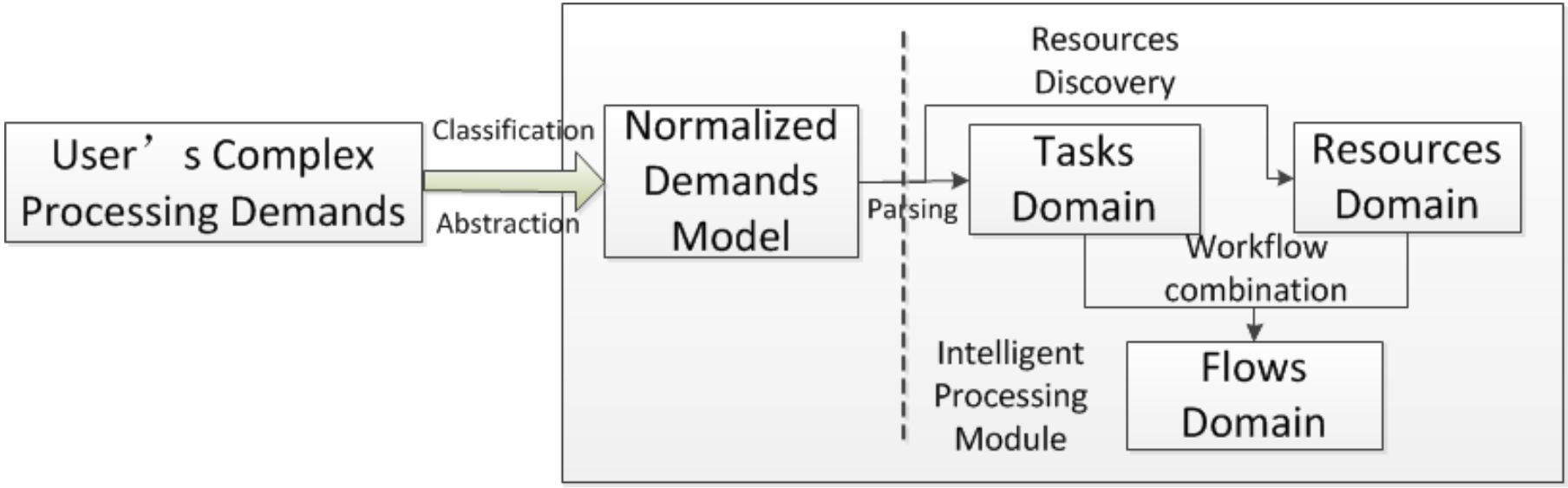
- General processing and service flow



The flowchart of general processing and service in on-demand mode

On-demand Architecture

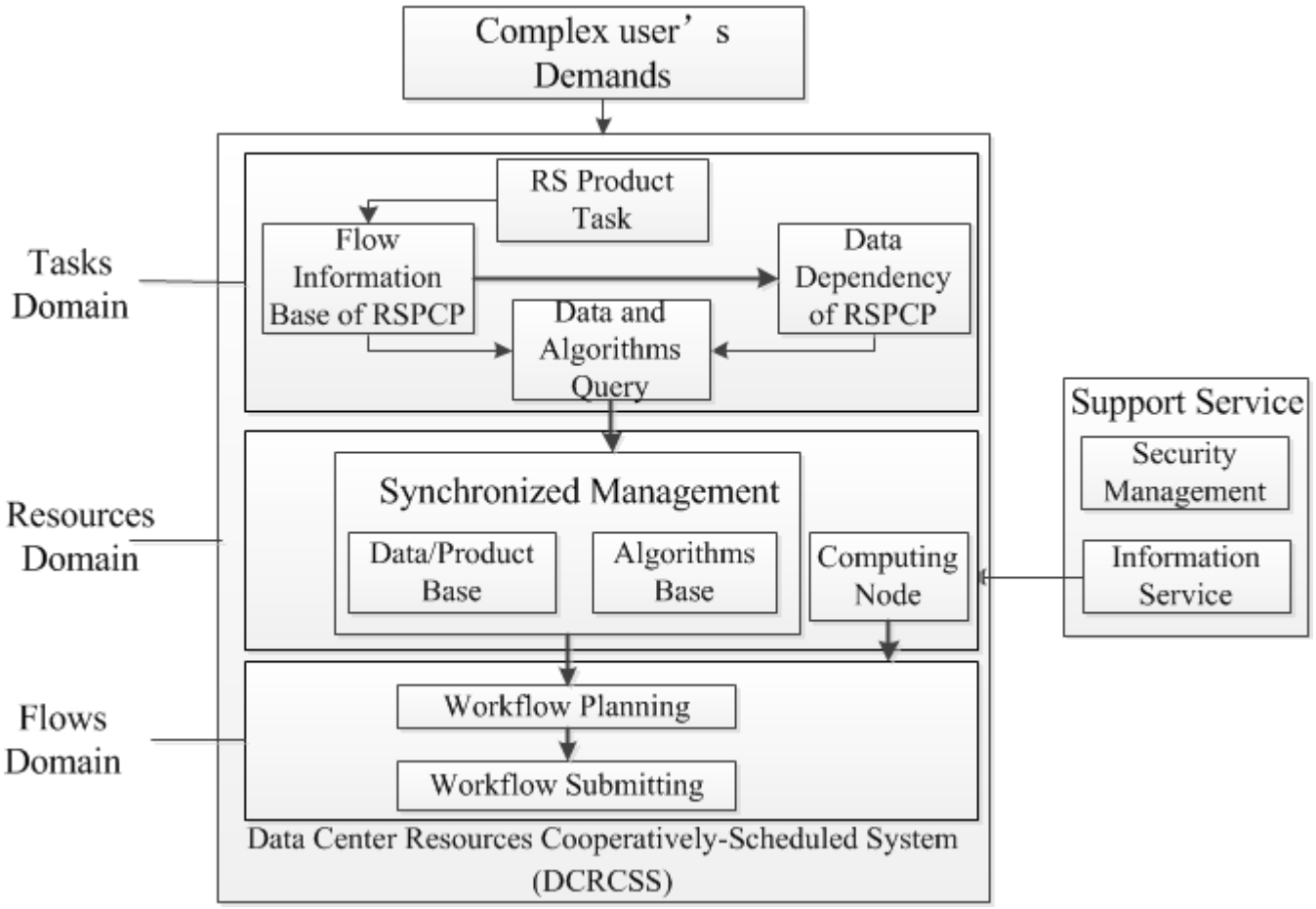
- In order to illustrate the on-demands architecture, fundamental components including **tasks domain, resources domain, and flow domain are obligated** .



The flowchart of complex processing demands parsing automatically

On-demand Architecture

- These three types of domains could be implemented through Data Center Resources Cooperatively-Scheduled System (**DCRCSS**) across multiple DCs.



Three domains construction in DCRCSS

Contents

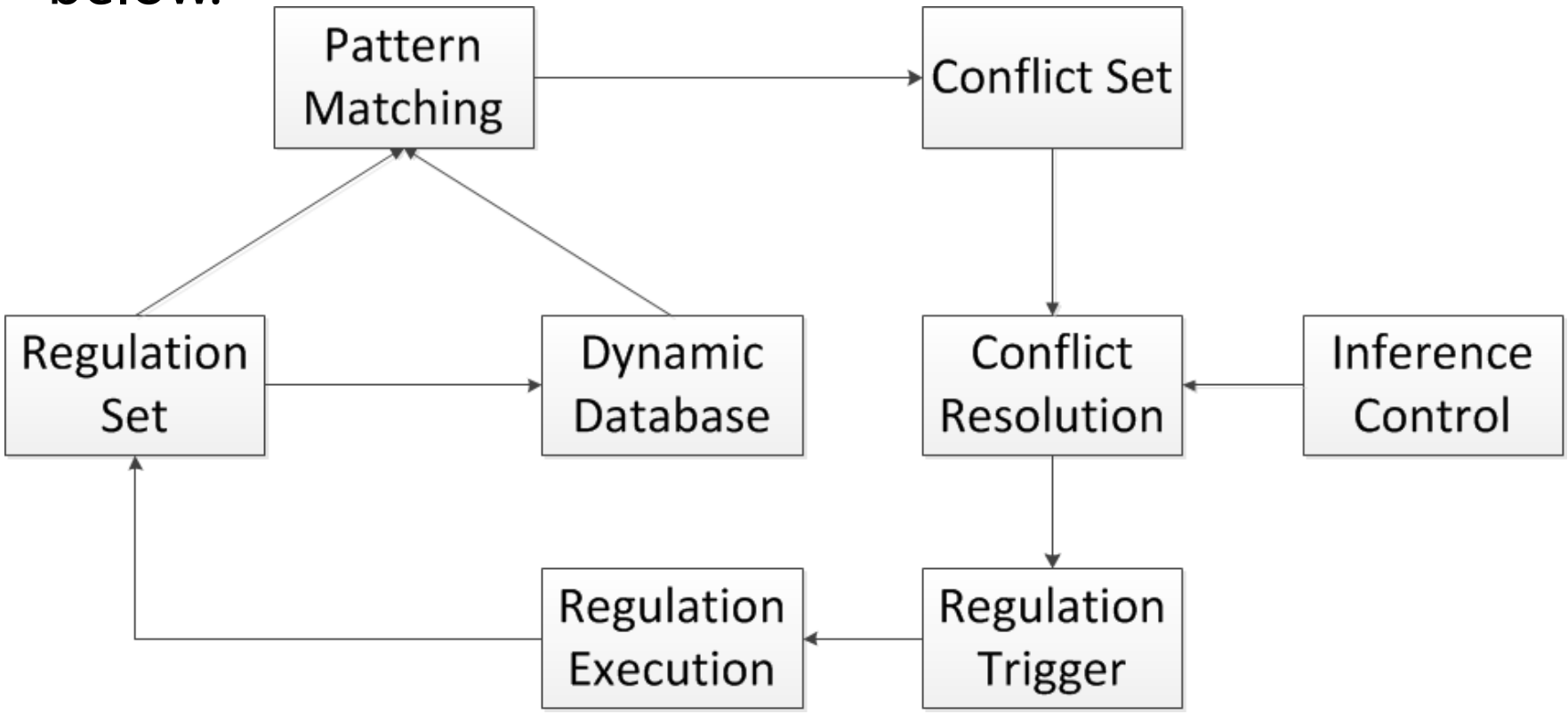
Experiments

1. Knowledge Base Implementation

- In order to implement the above three domains, knowledge base of RSPCP processing is the core part.
- To meet the requirements of multi-source remote sensing product processing, we take product classification rules into consideration when building **knowledge base**.
- The product classification knowledge base is built according to **RSPCP categories**.
- Remote sensing data dependency, ability of computing nodes , and storage capacity also need to be saved in knowledge base.

Experiments

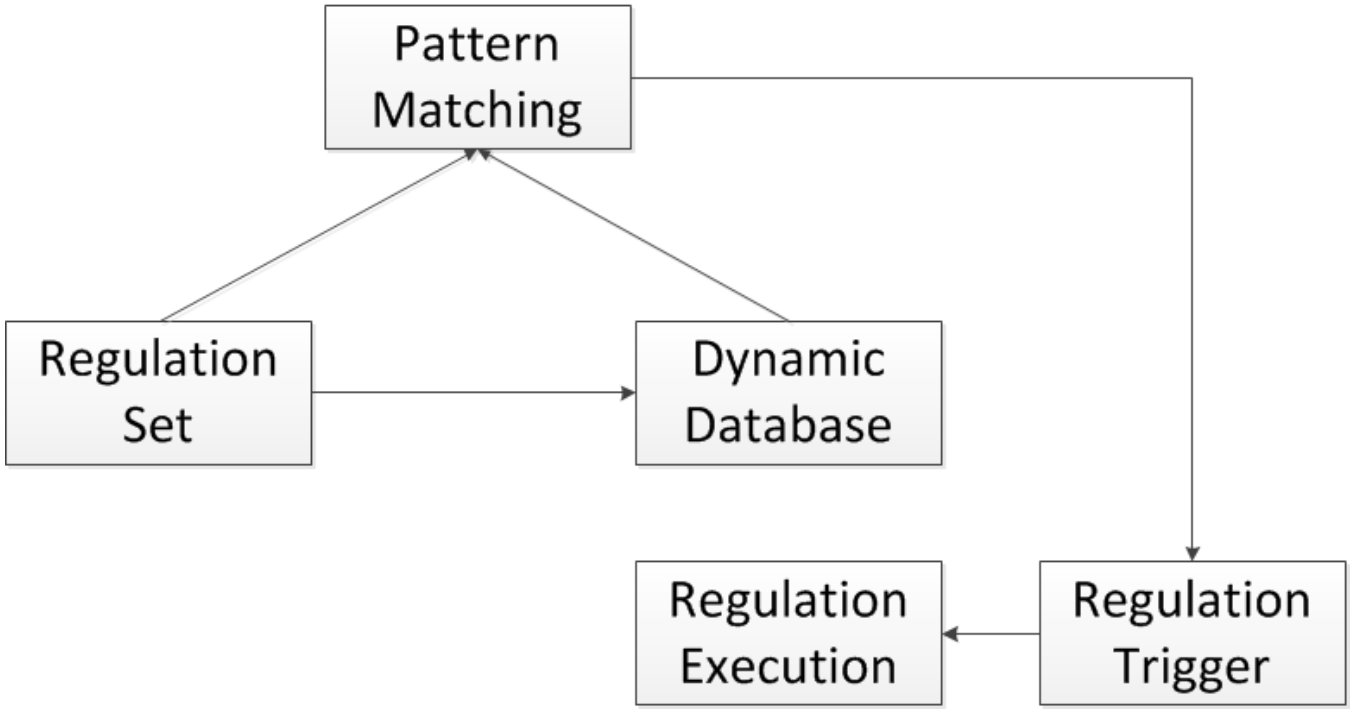
- Knowledge base can be created according to procedure below.



The flowchart of knowledge base creating process

Experiments

- The implement of Knowledge base inference is simplified.



The flowchart of knowledge base implement process

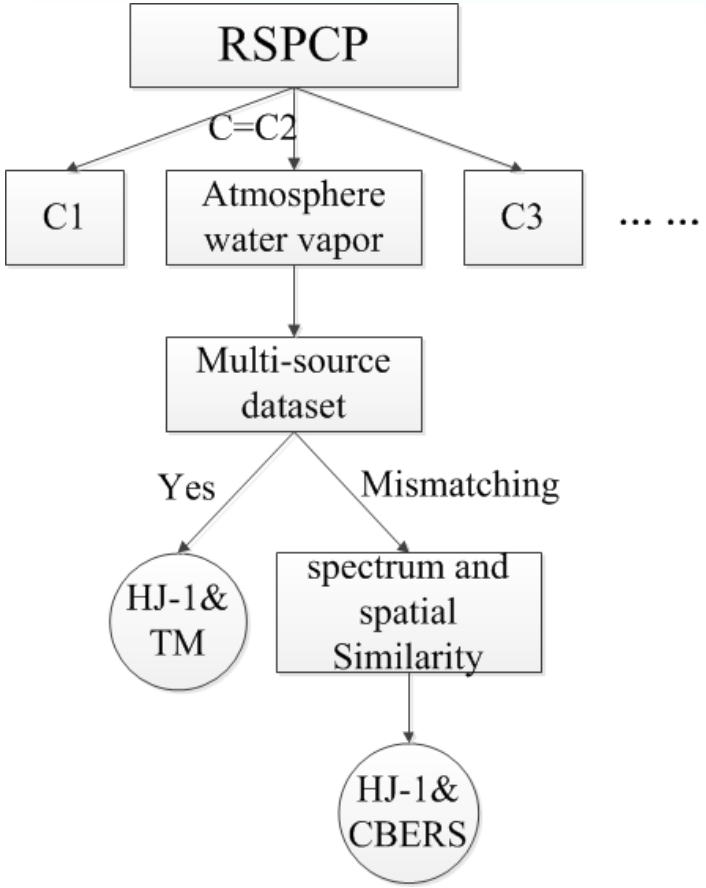
Experiments

- These four categories cover several different remote sensing products.

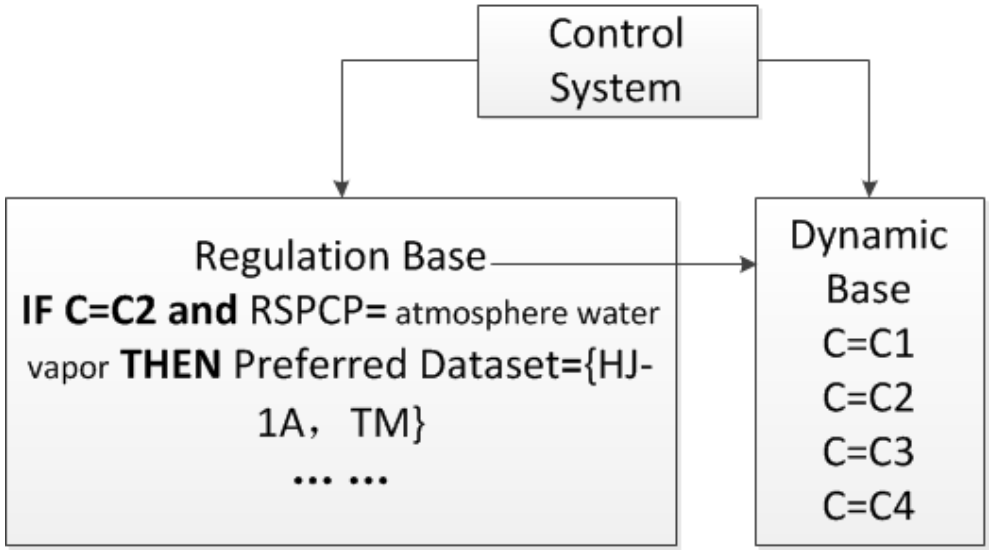
RSPCP

- **C1**:Mineral detecting key parameters
- **C2**:Radiation and hydrothermal parameters
  - Aerosol optics thickness
  - Atmosphere water vapor
  - BRDF
  - … …
- **C3**:Global snow changing parameters
- **C4**:Vegetation structure
  - Chlorophyll content
  - Leaf area index
  - Phenological phase
  - … …

Experiments



- **Once the data category is determined, the primary data source would be confirmed** based on production rules.



The flowchart of Knowledge Base System Inference in Tasks Domain

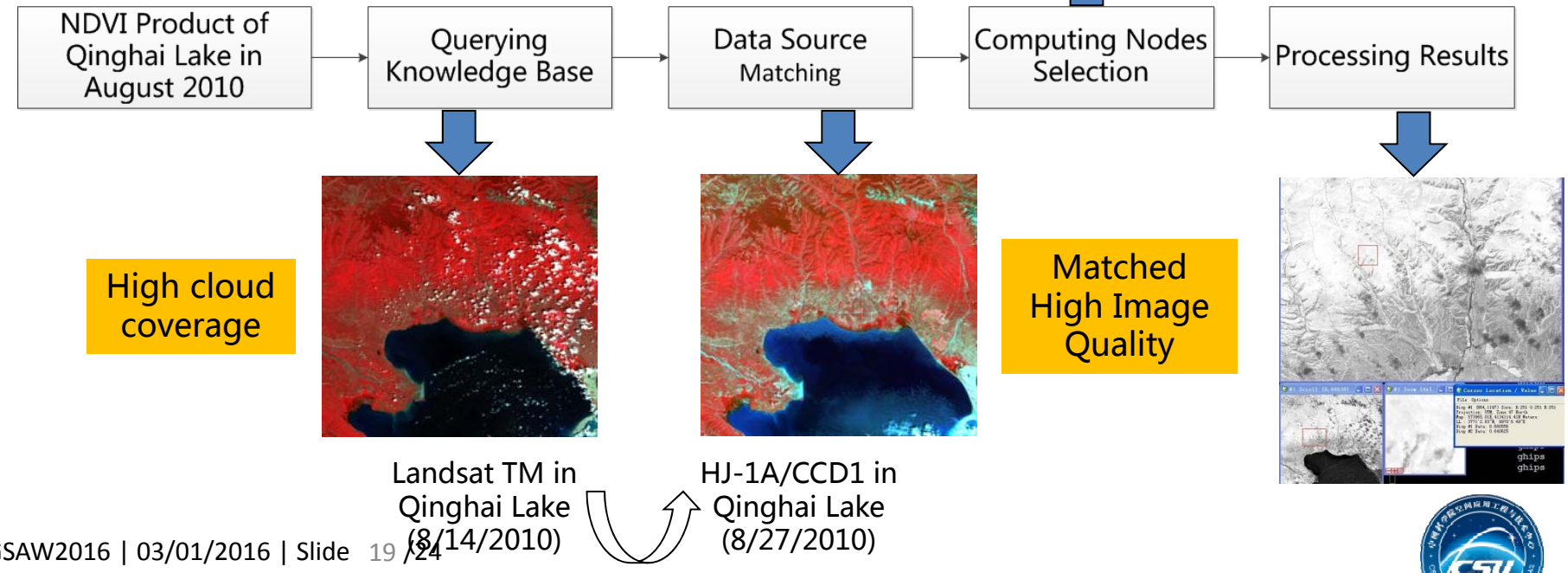
Experiments

2. Implementation of On-demands Architecture for Remote Sensing Products

- An example of NDVI production is illustrated as follows.

NDVI=(NIR-R)/(NIR+R)

Requirements



Optical Data similarity model

Experiments

- **Optical data similarity model** is illustrated as follows.

$$Sim(R_1,R_2)=\{\Delta_{\lambda(R_1,R_2)}, \Delta_{\frac{1}{2}\lambda(R_1,R_2)}, \Delta_{Reso(R_1,R_2)}, \Delta_{H(X)(R_1,R_2)}\}$$

- Four factors are considered
  - Central wavelength
  - Distance from central wavelength to boundary wavelength
  - Image entropy
  - Spatial resolution

$$\Delta_{\lambda(R_1,R_2)}=\left[\lambda_{start_{R_1}}+\frac{1}{2}(\lambda_{end}-\lambda_{start})_{R_1}\right]-\left[\lambda_{start_{R_2}}+\frac{1}{2}(\lambda_{end}-\lambda_{start})_{R_2}\right]$$

$$\Delta_{\frac{1}{2}\lambda(R_1,R_2)}=\left[\frac{1}{2}(\lambda_{end}-\lambda_{start})_{R_1}\right]-\left[\frac{1}{2}(\lambda_{end}-\lambda_{start})_{R_2}\right]$$

$$\Delta_{Reso(R_1,R_2)}=Reso_{R_1}-Reso_{R_2}$$

$$\Delta_{H(X)(R_1,R_2)}=H(X)_{R_1}-H(X)_{R_2}$$

Contents

Conclusion

Results

- User’s complex requirements are parsed and translated into a concrete workflow automatically ,
- Knowledge base is designed and implemented for multiple remote sensing product.
- **Processing of Remote sensing products becomes convenient with the help of three domains realized.**

Conclusion

■ Further Research

- RSPCP defined in this report is relatively limited, **more extensive remote sensing applications** would be imported into knowledge base.
- **Data similarities model should be optimized** by considering more comprehensive and dispensable factors.
- Our DCRCSS is implemented across virtual DCs computing environments, and will be adopted by **practical data centers** step by step.

Thanks !