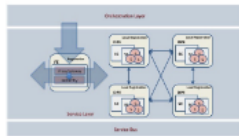


SCIA Data Fusion

- Driven by ~100 min SCIA orbit files
- Re-sample "footprint" records
- MODIS & ECMWF fusion
- Generate statistics

CLARA Service Oriented Architecture



NASA Information And Data System

Sciamachy Data Fusion

NASA
Jefferson Lab
CLARA Framework

ESTF2016
Earth Science Technology Forum 2016

NIAADS Team

- PI: Christopher Columbus Johnson (JPL)
- Co-PI: Jeffrey M. Brantner (JPL)
- Co-PI: Jeffrey J. Carlisle (JPL)
- Co-PI: Jeffrey J. Carlisle (JPL)
- Co-PI: Jeffrey J. Carlisle (JPL)
- Co-PI: Jeffrey J. Carlisle (JPL)

Orchestrator Framework

- Multi-core & multi-node service compositions
- Service setup & teardown
- Error handling
- YAML configuration
- No domain-specific knowledge

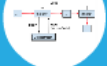
SCIA Prototype (AWS)

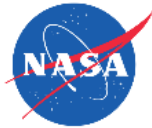


Processing Phases



Multi-core & Multi-node





NAIADS

Jefferson Lab CLARA Framework



NASA Information And Data System

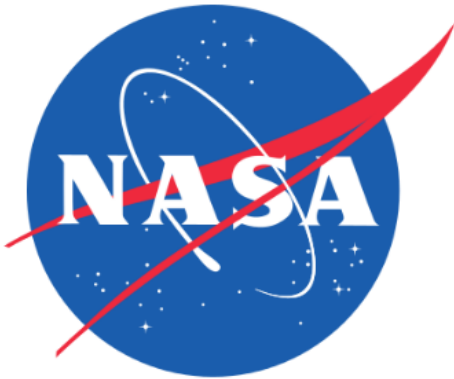
Sciamachy Data Fusion

Format



NAIADS Team

- PI: Constantine Lukashin (NASA)
- CO-I: Vardan Gyurjyan (JLab)

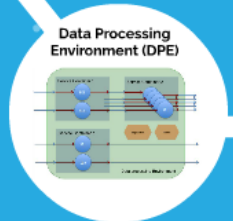
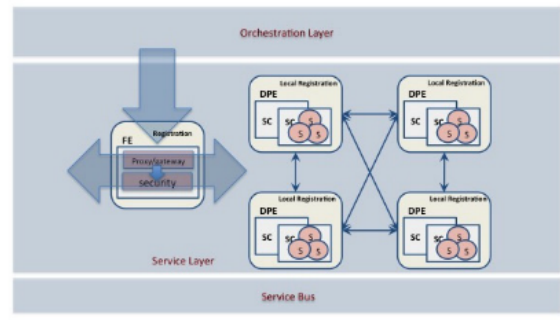
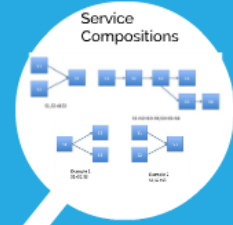


NAIADS

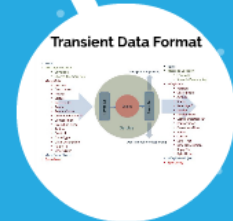
 **Jefferson Lab**

CLARA Framework

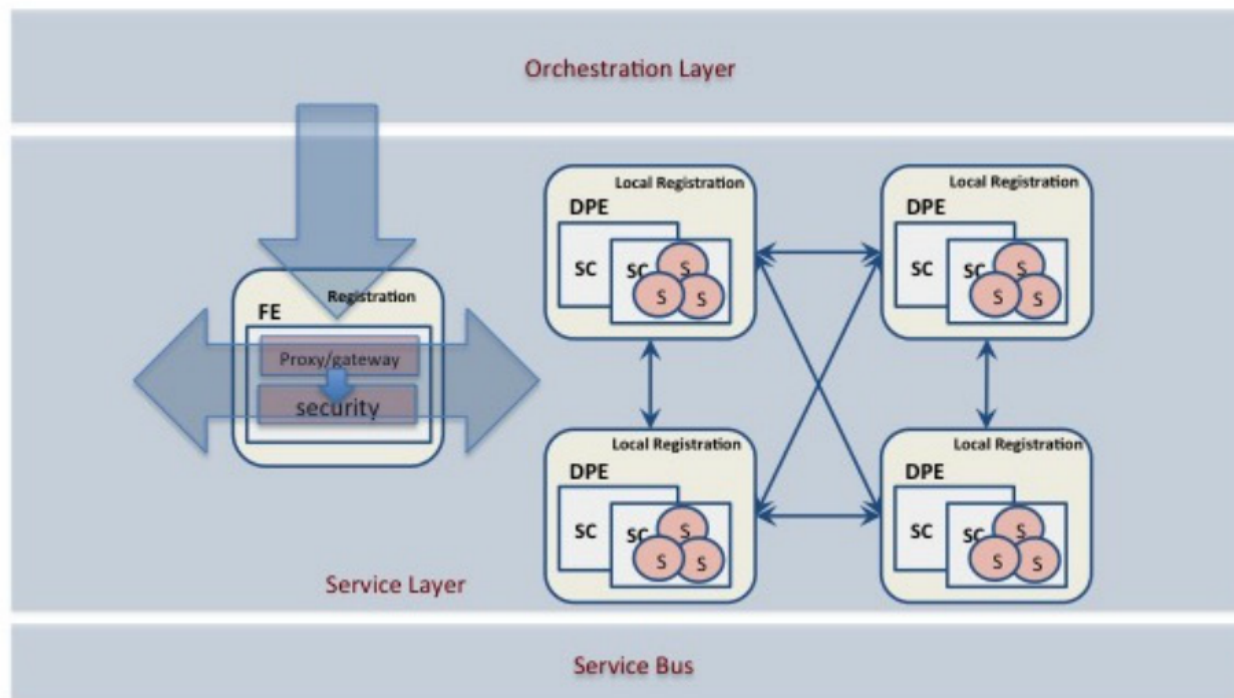
CLARA Service Oriented Architecture



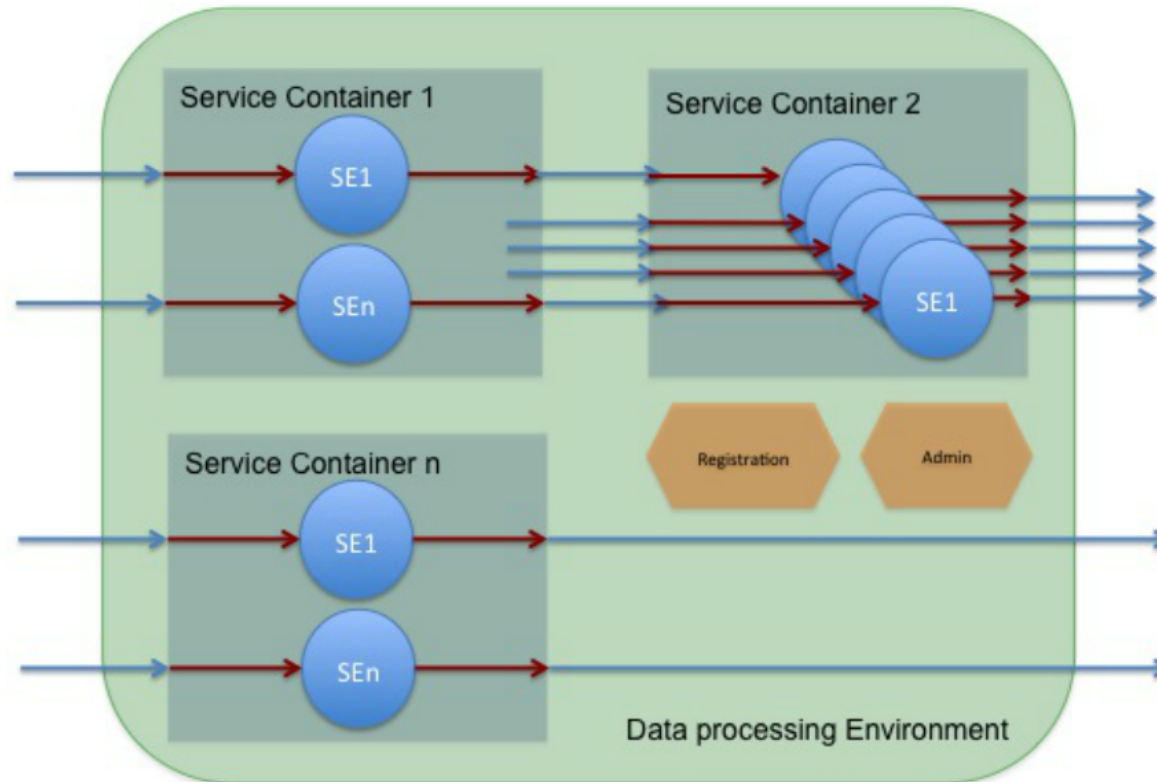
NA
Ar
Scian



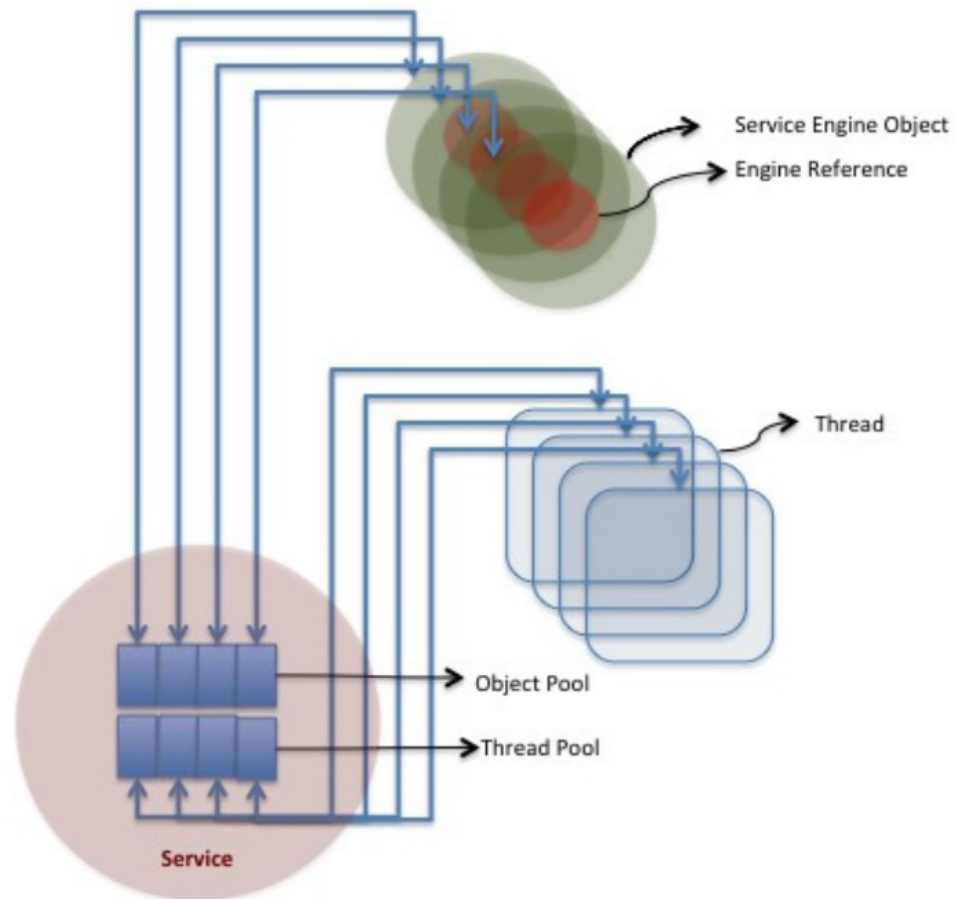
CLARA Service Oriented Architecture



Data Processing Environment (DPE)

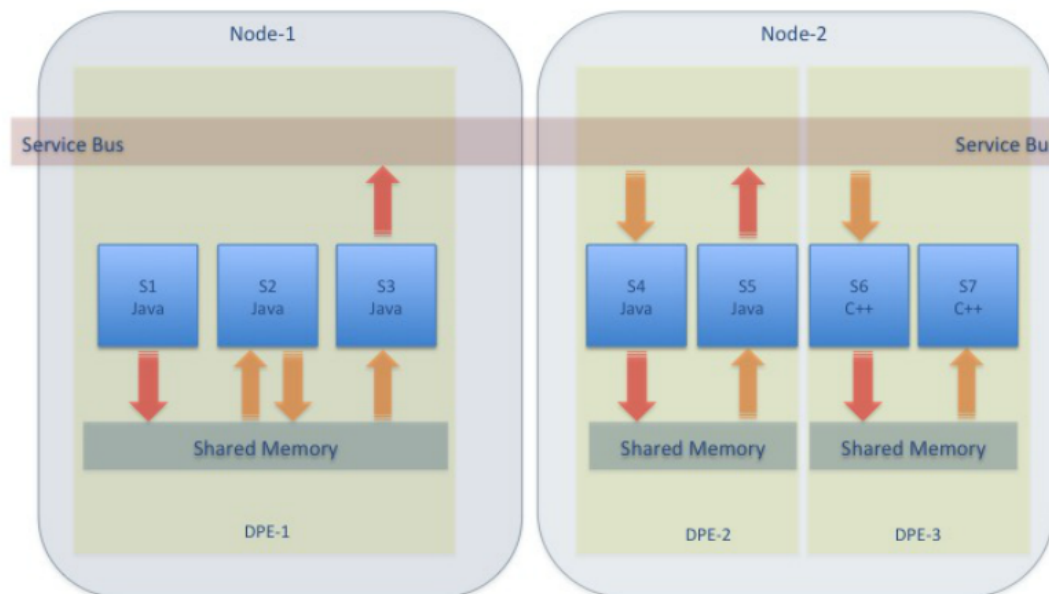


Service Threads @DPE

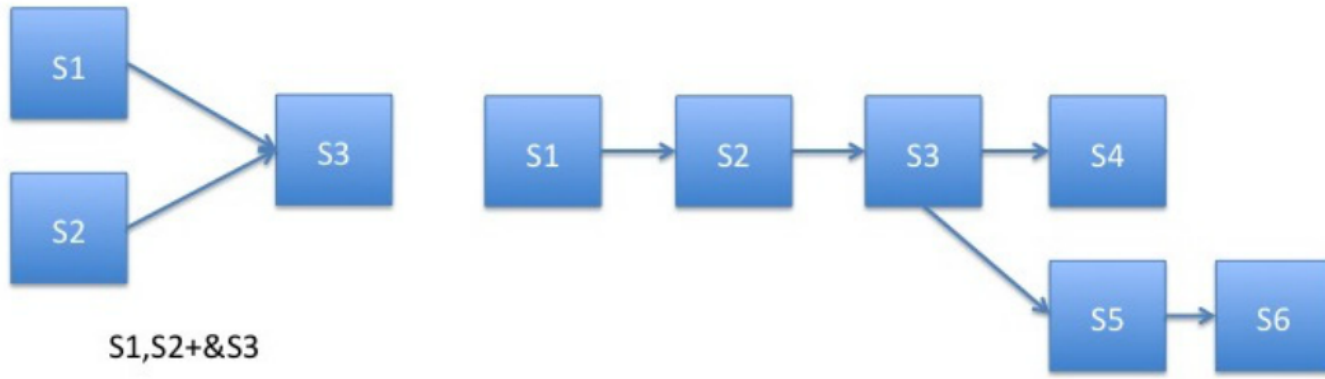


Service Deployment: Java, Python & C++

Data processing application composition (task service)

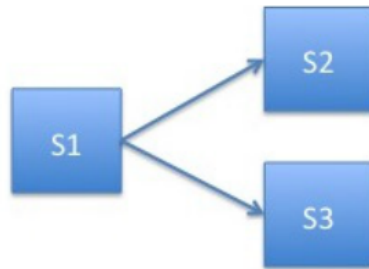


Service Compositions

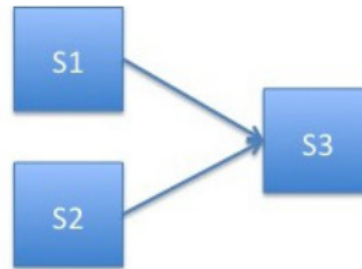


$S1, S2 \& S3$

$S1+S2+S3+S4; S3+S5+S6$



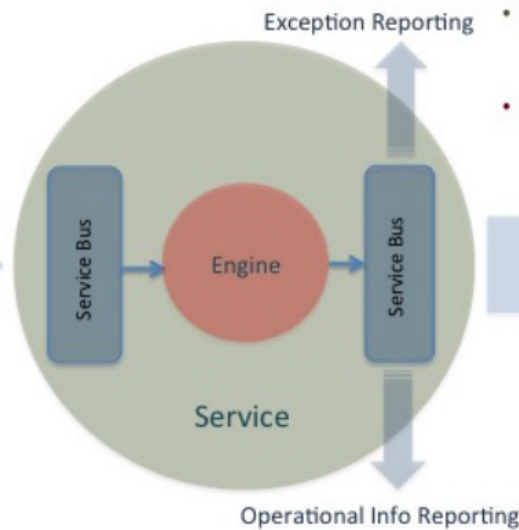
Example 1
 $S1+S2, S3$



Example 2
 $S1, S2+S3$

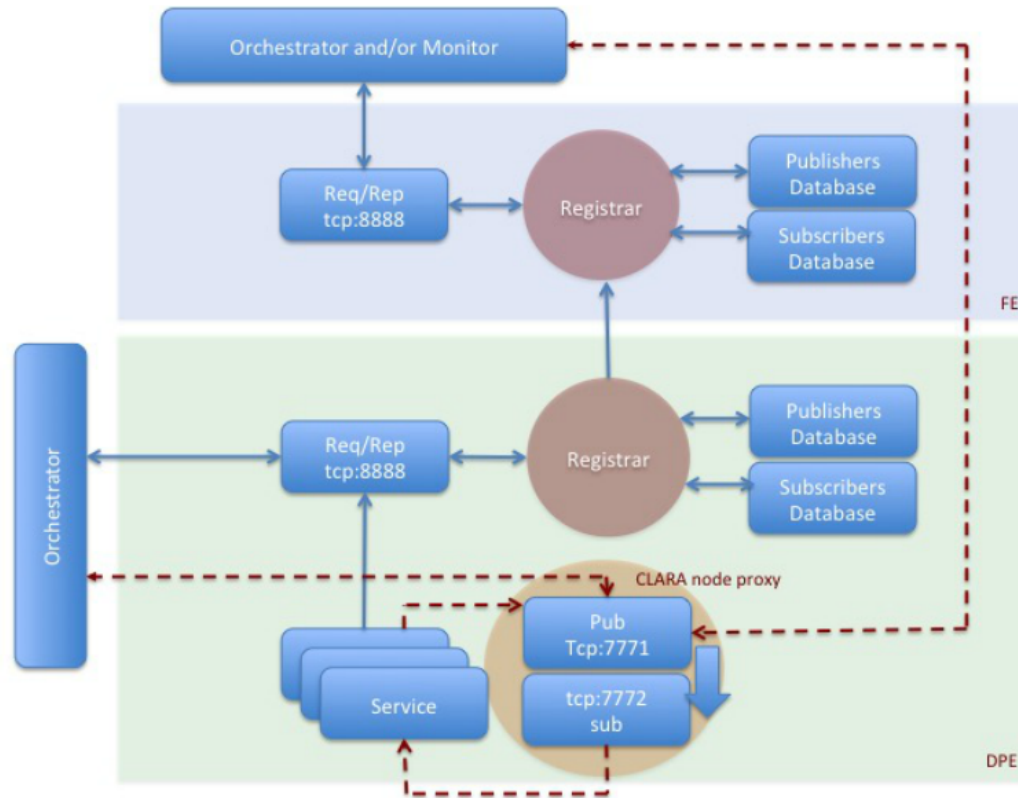
Transient Data Format

- Topic
- Message-Location
 - Envelope
 - Shared-Memory Key
- xMsgMeta
 - Version
 - Description
 - Author
 - Status
 - Severity-ID
 - Sender
 - Sender-State
 - Communication-ID
 - Composition
 - Execution-Time
 - Action
 - Control
 - Data-Type
 - Data-Description
 - Reply-To
 - Byte-Order
- xMsgData-Object
- **Byte-Array**

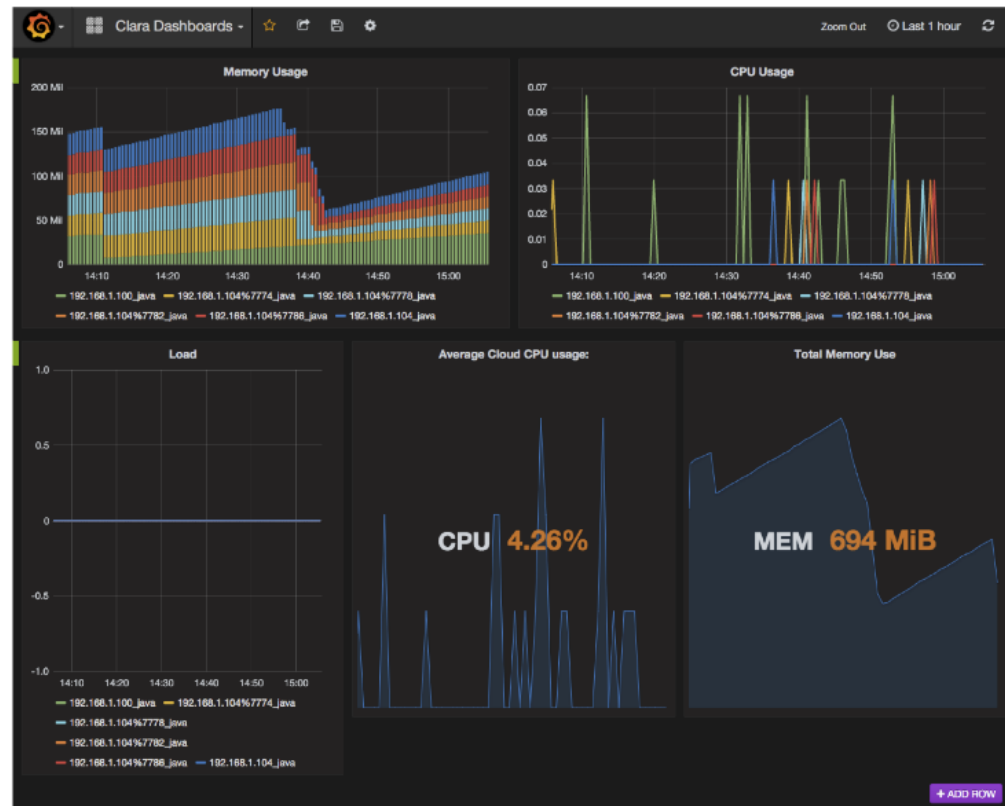


- Topic
- Message-Location
 - Envelope
 - Shared-Memory Key
- xMsgMeta
 - Version
 - Description
 - Author
 - Status
 - Severity-ID
 - Sender
 - Sender-State
 - Communication-ID
 - Composition
 - Execution-Time
 - Action
 - Control
 - Data-Type
 - Data-Description
 - Reply-To
 - Byte-Order
- xMsgData-Object
- **Byte-Array**

Registration & Monitoring



Web Dashboard



RESTful Interface



dpes

Show/Hide | List Operations | Expand Operations | Raw

GET

/dpes/

Find DPEs that match the optional query parameters

POST

/dpes/

Start new DPE(s)

...

containers

Show/Hide | List Operations | Expand Operations | Raw

GET

/dpes/{DPE_id}/containers/

Find all containers

POST

/dpes/{DPE_id}/containers/

Create a new Clara Container

...

services

Show/Hide | List Operations | Expand Operations | Raw

GET

/dpes/{DPE_id}/containers/{container_id}/services/

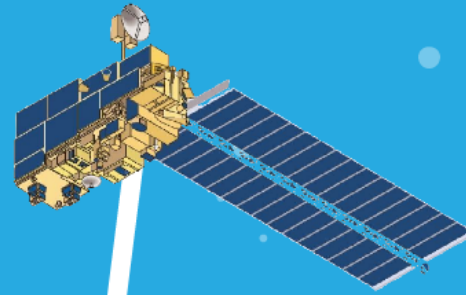
Find services that match the optional query parameters

POST

/dpes/{DPE_id}/containers/{container_id}/services/

Create a new service in one container at one DPE

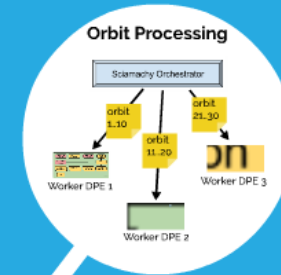
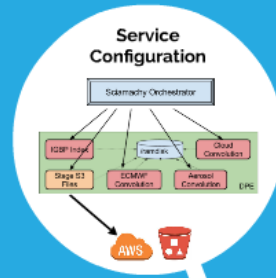
...



SCIA Data Fusion

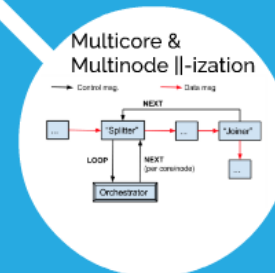
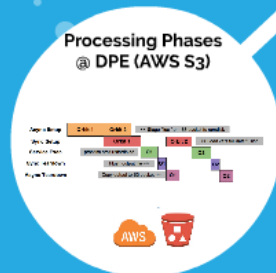
- Driven by ~100 min SCIA orbit files
- Re-sample "footprint" records
- MODIS & ECMWF fusion
- Generate statistics





Orchestrator Framework

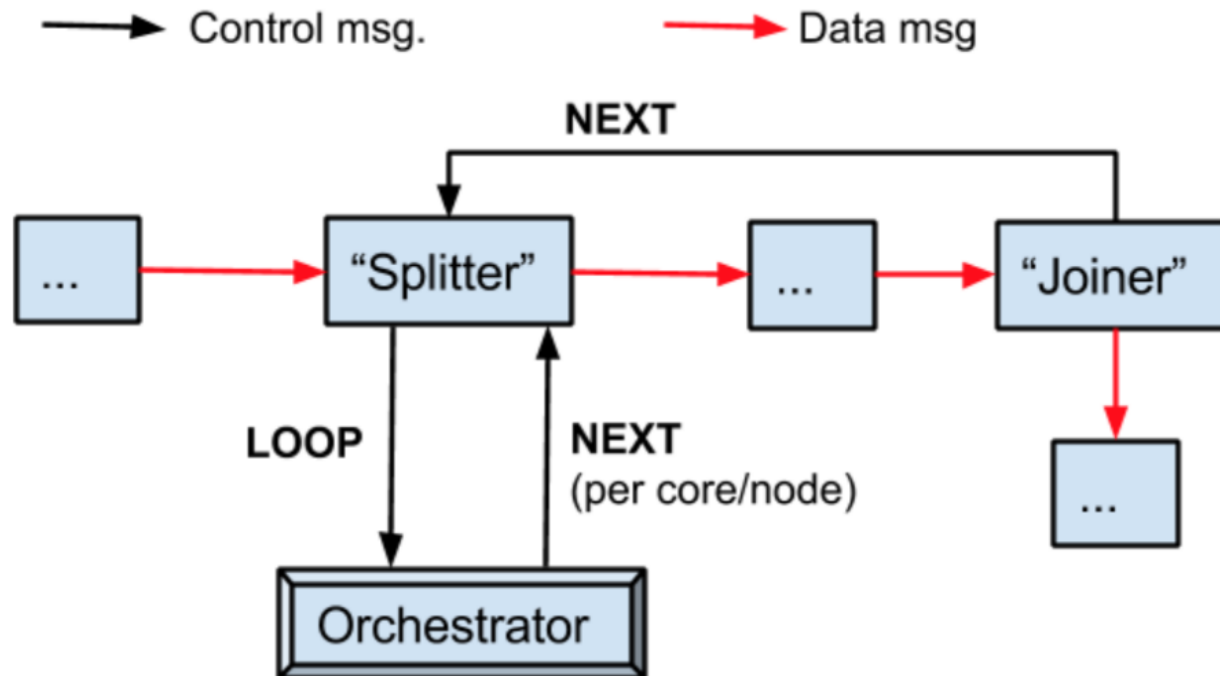
- Multi-core & multi-node service compositions
- Service setup & teardown
- Error handling
- YAML configuration
- No domain-specific knowledge



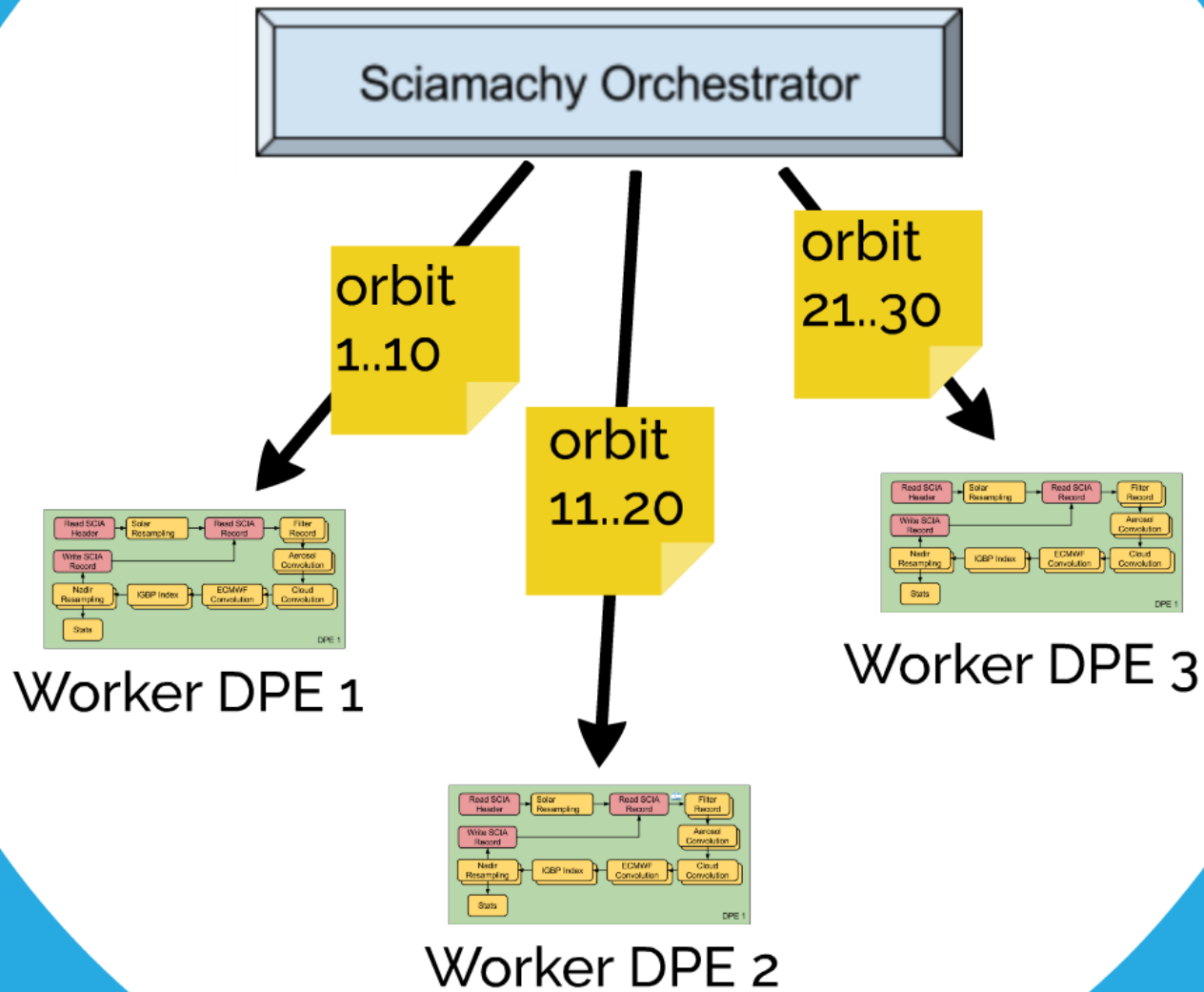
Orchestrator Framework

- Multi-core & multi-node service compositions
- Service setup & teardown
- Error handling
- YAML configuration
- No domain-specific knowledge

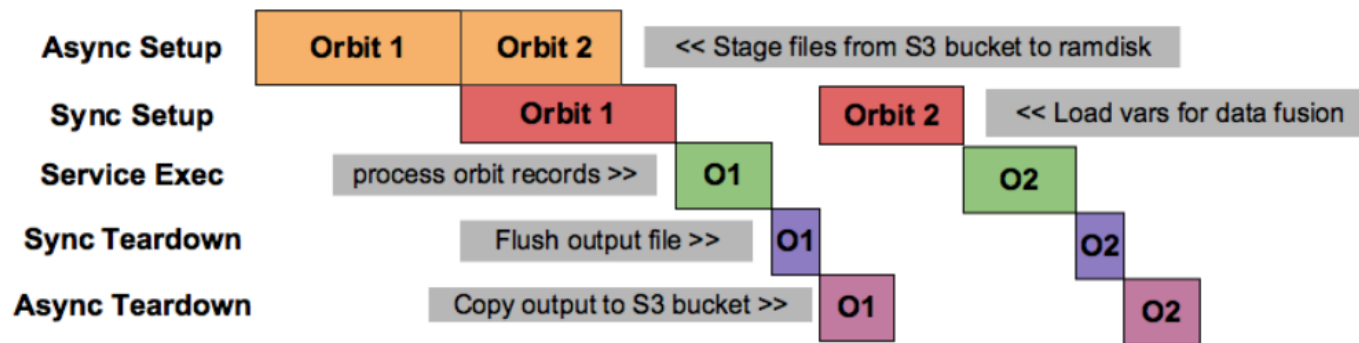
Multicore & Multinode ||-ization



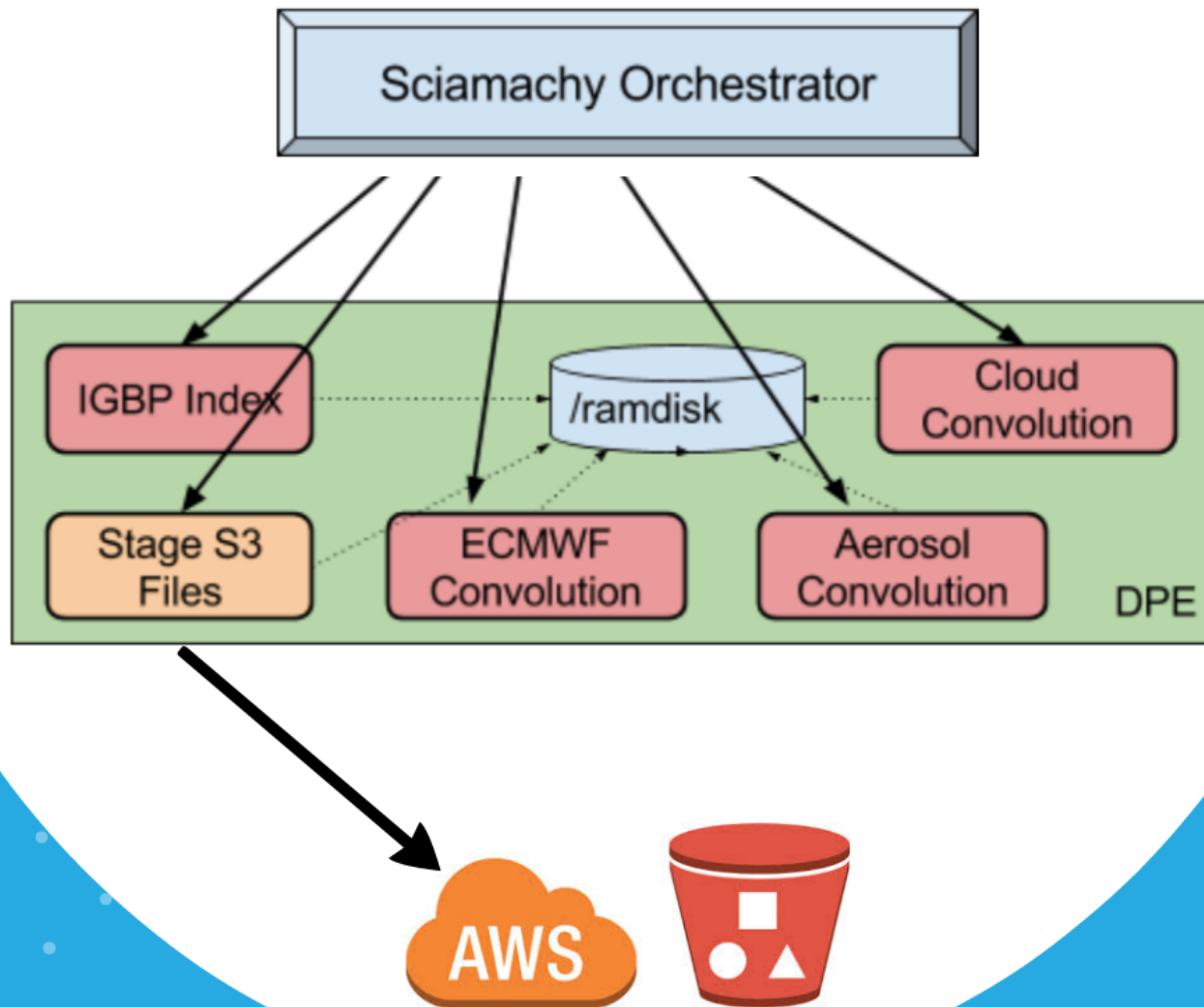
Orbit Processing

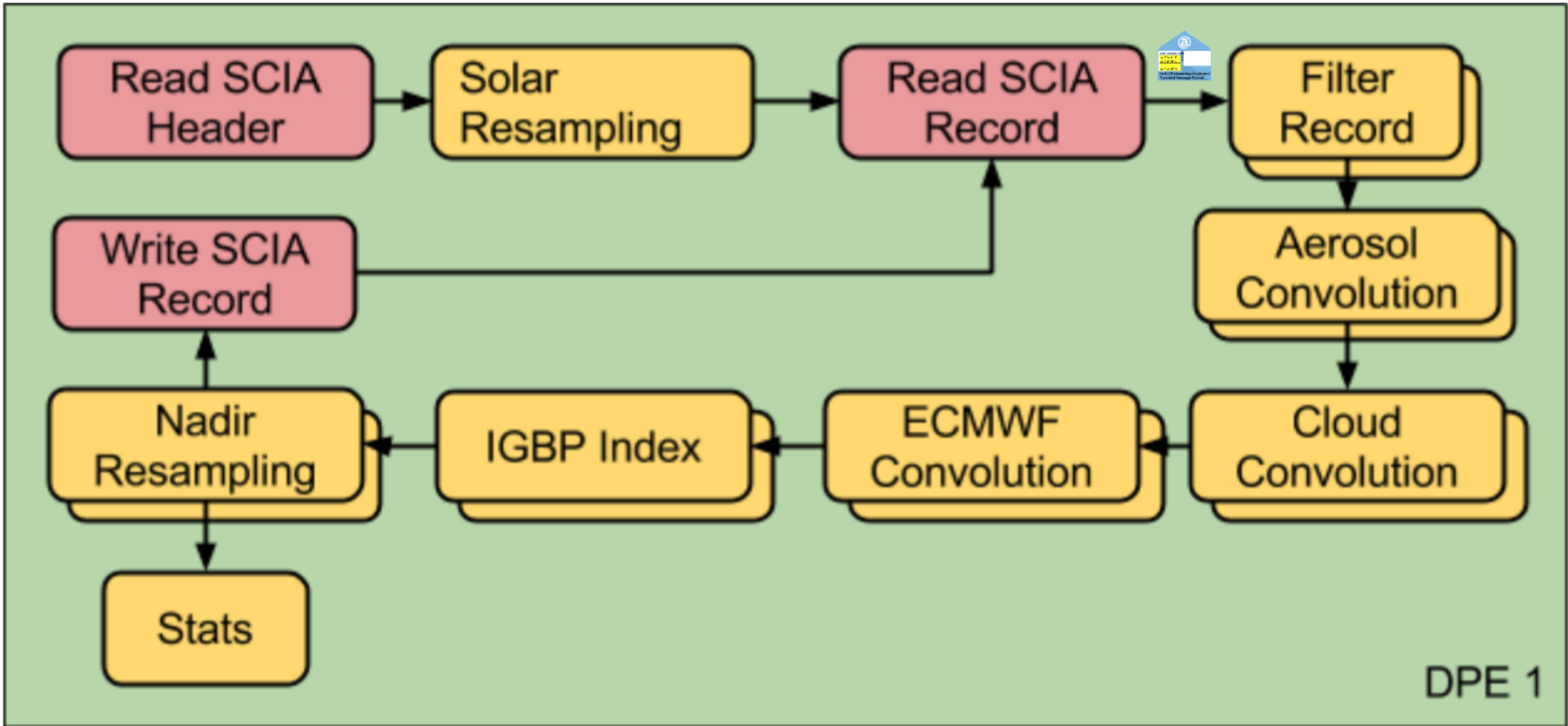
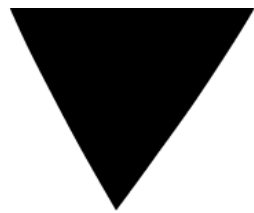


Processing Phases @ DPE (AWS S3)



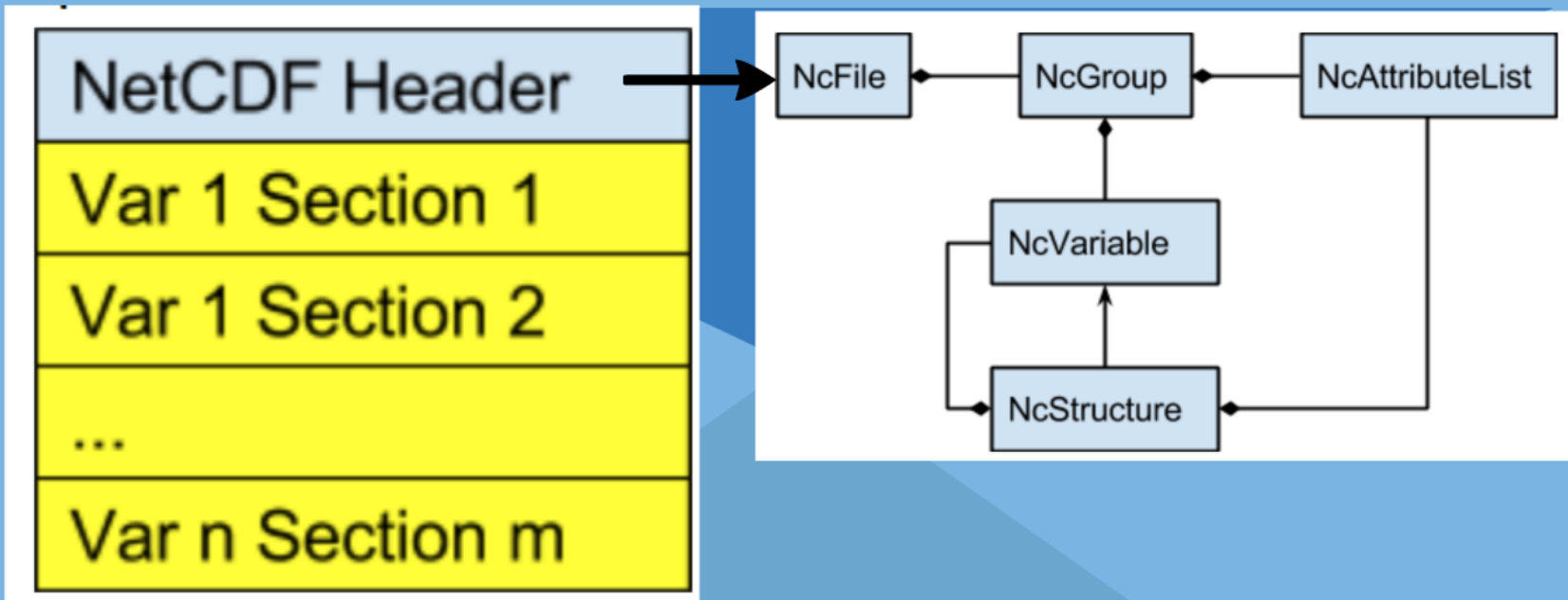
Service Configuration





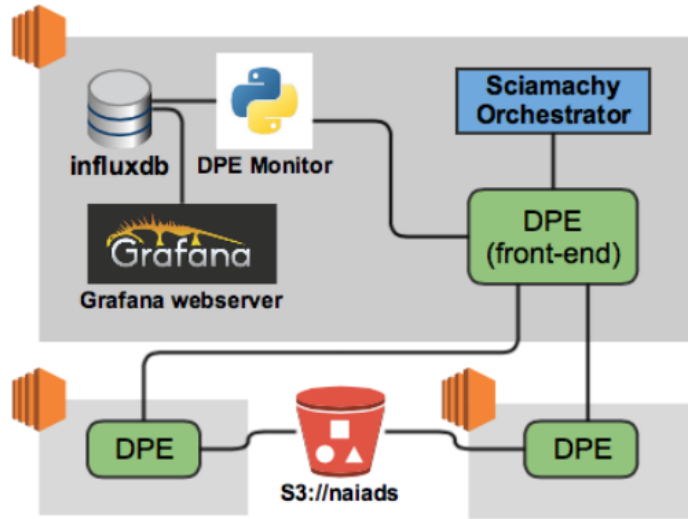
Worker DDF



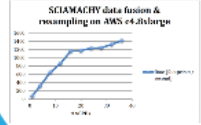


NetCDF-streaming (ncstream) Transient Message Format

SCIA Prototype (AWS)

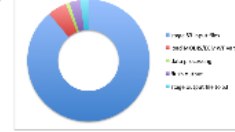


Vertical Scaling (AWS virtual CPUs)



Not including AWS S3 staging
Not including Service config

Avg (AWS) DPE Performance

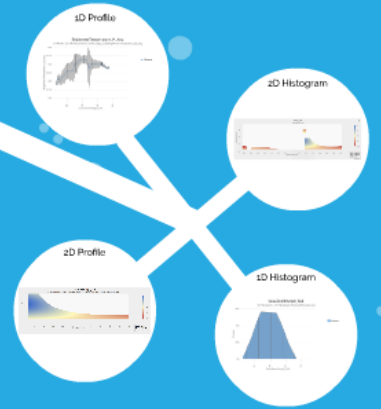


total time: ~82s
SCIA footprints: ~1447
MODIS files: ~62
Input data size: ~2G

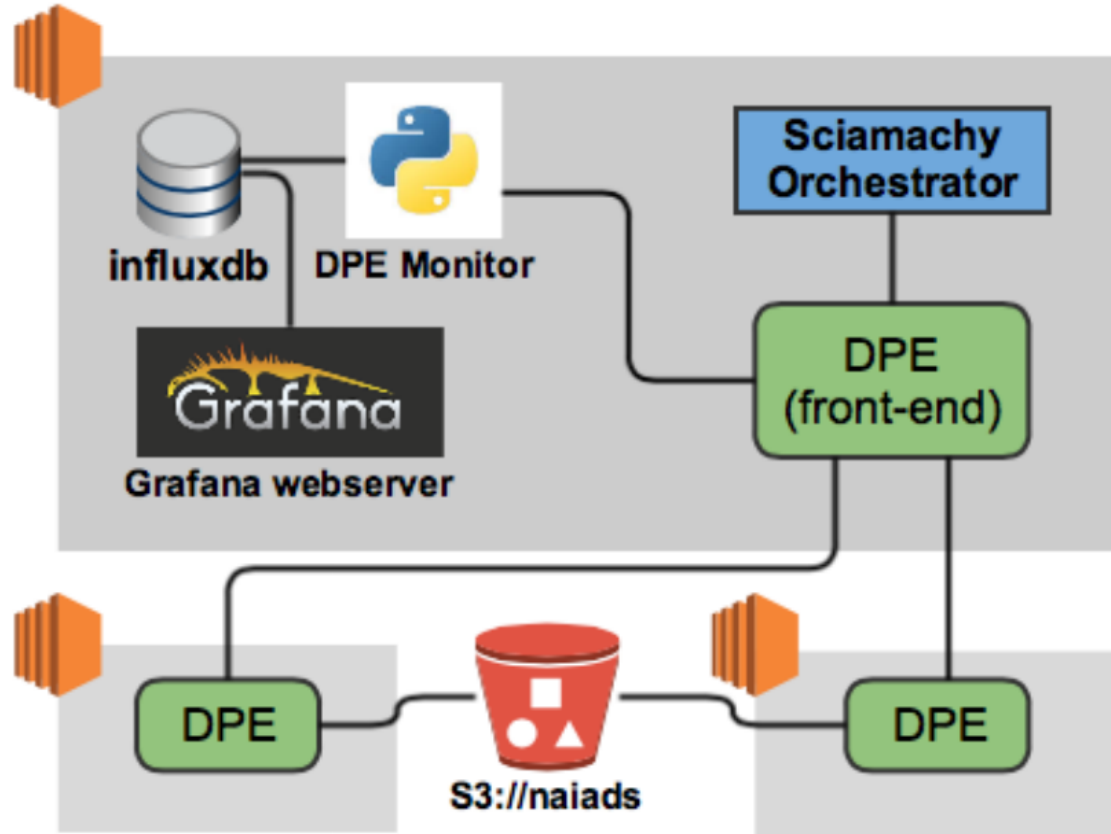
Horizontal Scaling (AWS c4.8xlarge)



Not including AWS S3 staging



SCIA Prototype (AWS)



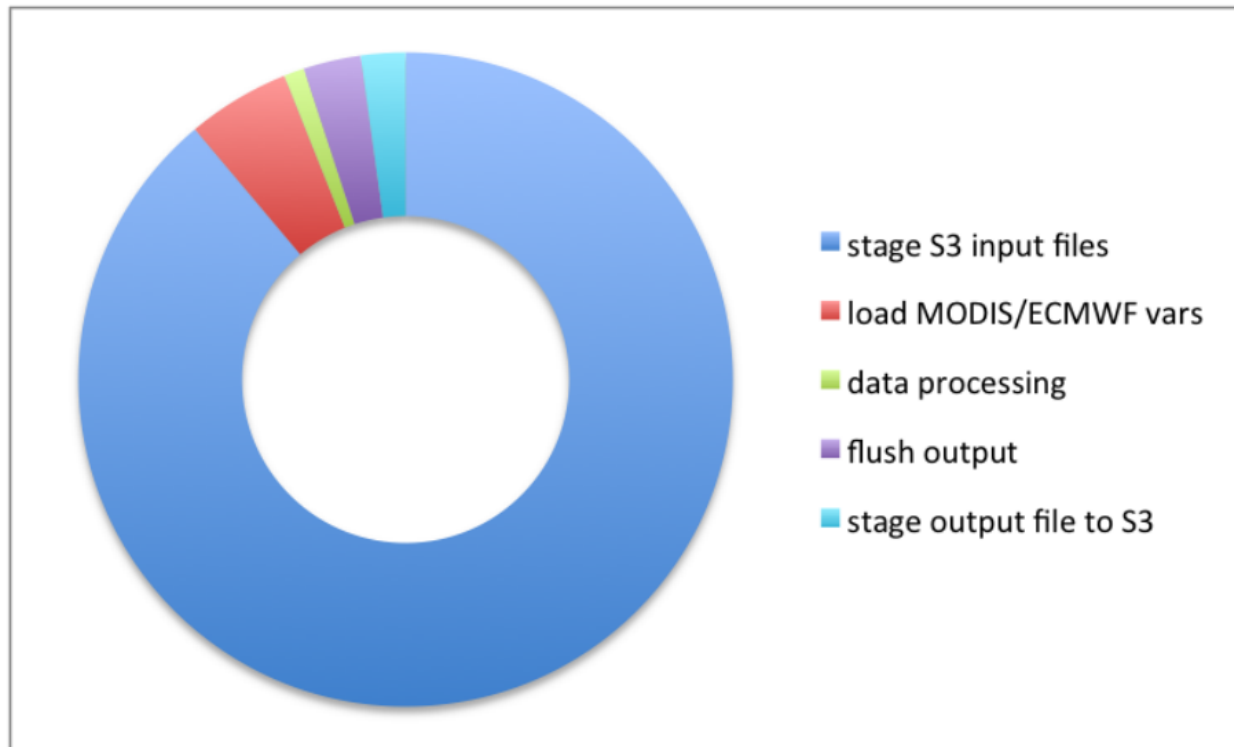
DPE Perf



total time: ~
SCIA footprint
MODIS files
Input data s

- Not including AWS S3 staging
- Not including Service config

Avg (AWS) DPE Performance



total time: ~82s

SCIA footprints: ~1447

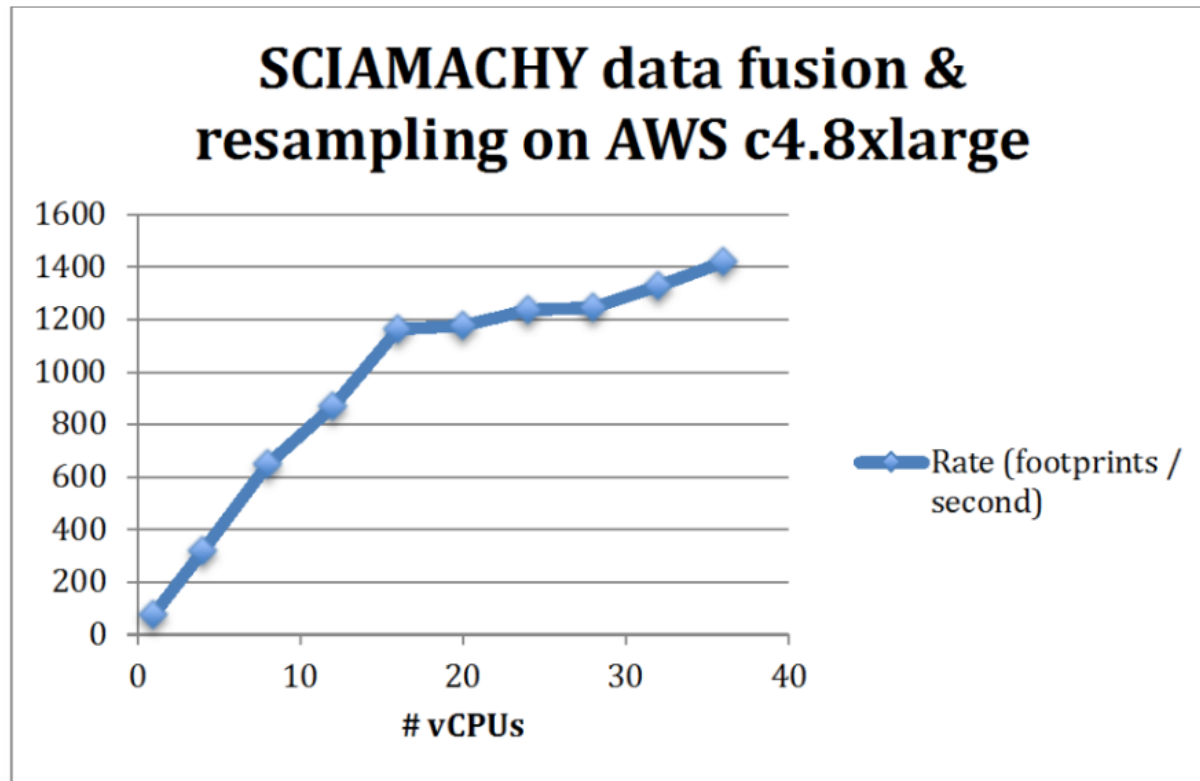
MODIS files: ~62

Input data size: ~2G

Vertical Scaling (AWS virtual CPUs)

* Not including
AWS S3 storage

* Not including
Service

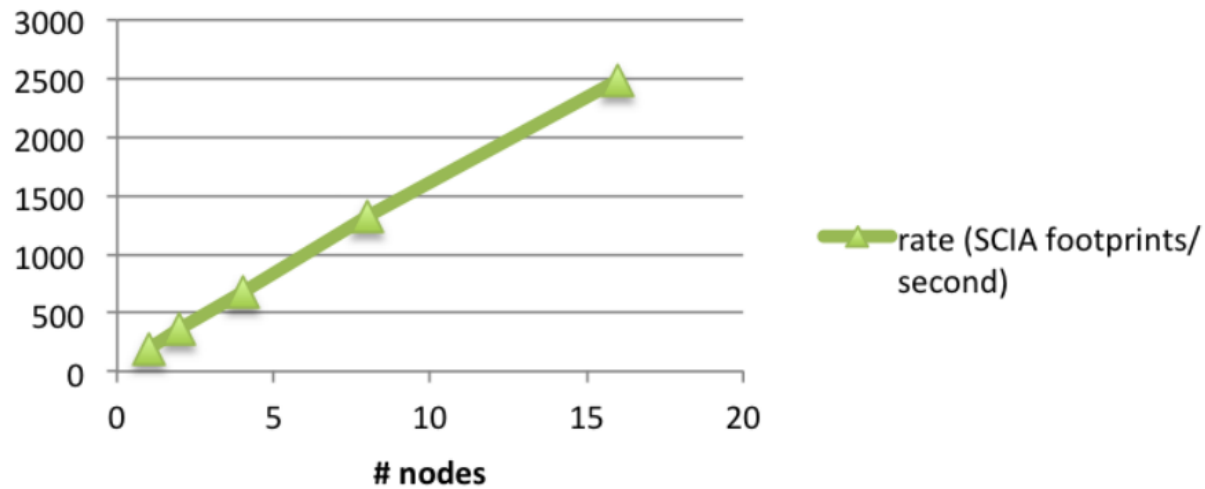


cluding
3 staging

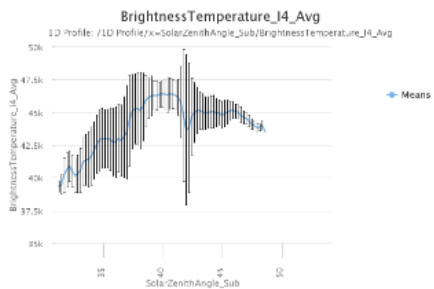


Horizontal Scaling (AWS c4.8xlarge)

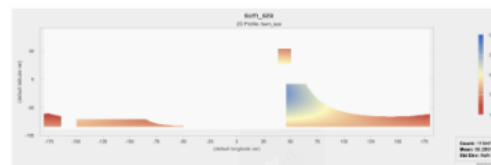
SCIAMACHY data fusion & resampling
on AWS c4.8xlarge



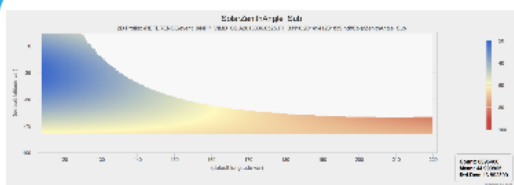
1D Profile



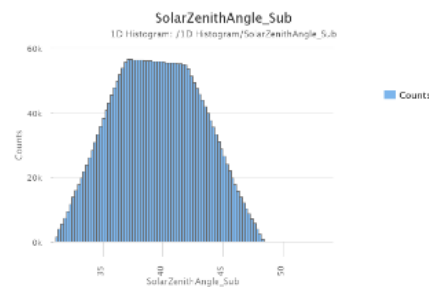
2D Histogram

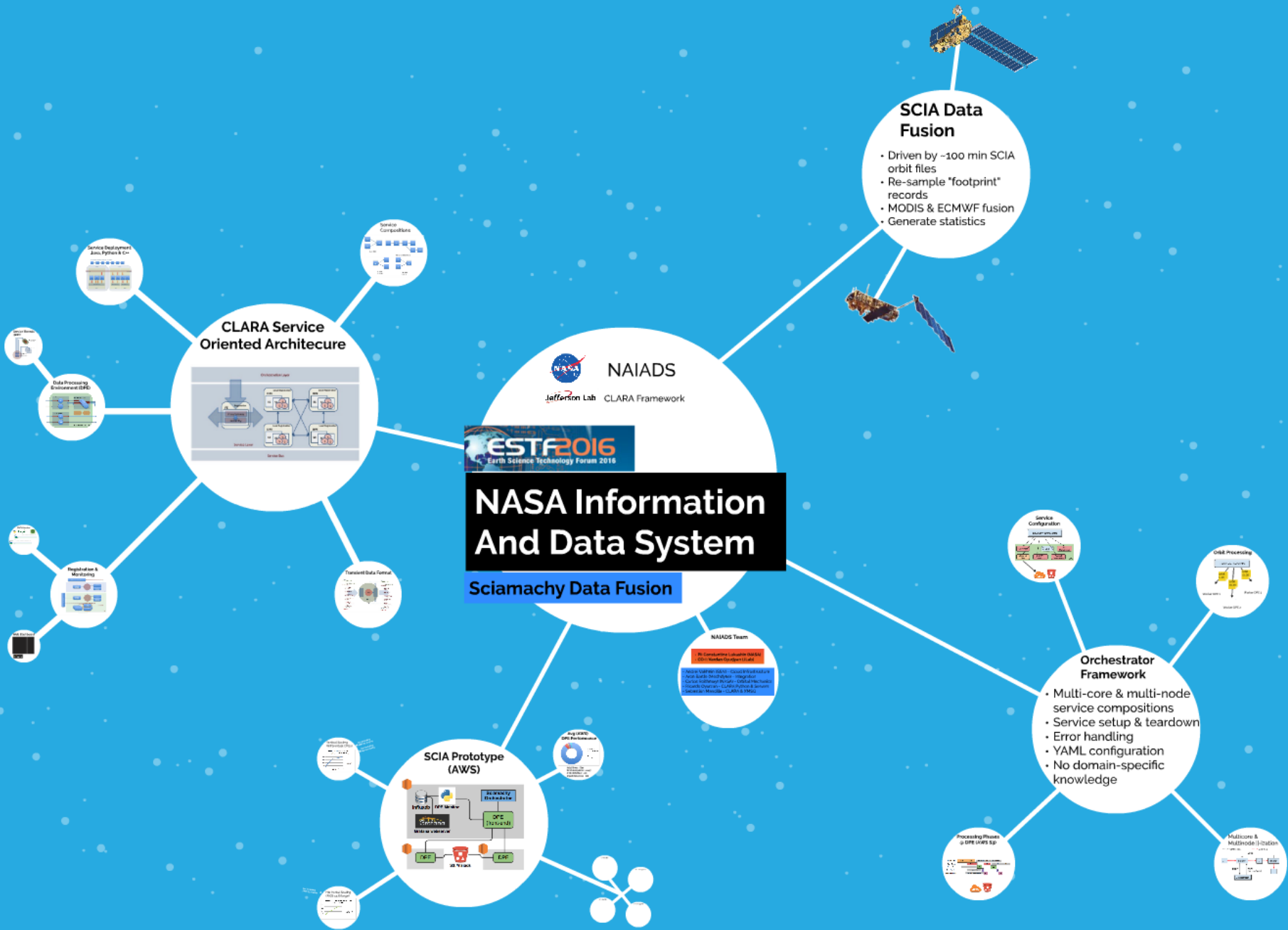


2D Profile



1D Histogram





NAIADS Team

- **PI: Constantine Lukashin (NASA)**
- **CO-I: Vardan Gyurjyan (JLab)**

- Andrei Vakhnin (SSAI) - Cloud Infrastructure
- Aron Bartle (Mechdyne) - Integration
- Carlos Roithmayr (NASA) - Orbital Mechanics
- Ricardo Oyarzun - CLARA Python & Servers
- Sebastián Mancilla - CLARA & XMSG