

NCCAT Round Table April 2020:

How NCCAT is using AWS Cloud services for pre-processing and cross-training

Swapnil Bhatkar

AWS Solutions Architect

New York Structural Biology Center



Swapnil
Bhatkar



Micah
Rapp



Ed
Eng



TRANSFORMATIVE HIGH RESOLUTION CRYO-ELECTRON MICROSCOPY PROGRAM



<https://commonfund.nih.gov/CryoEM>



The program aims to broaden access to high-resolution cryoelectron microscopy (cryoEM) for biomedical researchers, by creating **national service centers**, and cultivating a skilled workforce, through the development and implementation of **cryoEM training material**.

NCCAT MISSION

The mission of NCCAT is twofold: to provide nationwide access to advanced cryoEM technical capabilities, and to assist users in the development of cryoEM skills needed for independent research.

NIH Common Fund Initiative

 In progress

 Completed



SIMONS ELECTRON MICROSCOPY CENTER
NATIONAL CENTER FOR CRYOEM ACCESS AND TRAINING





Apply now



Transformative High Resolution
Cryo-electron Microscopy



NIH Common Fund
program partners



National Center for
Cryo-EM Access
and Training

S²C² | Stanford-SLAC
Cryo-EM Center

PACIFIC NORTHWEST
Cryo-EM
Center

Nationwide **access** to **cryo-EM instrumentation**
and cross-**training** at **no cost**.

AGENDA



INTRODUCTION



CLOUD WORKSHOPS



CHALLENGES



ON-DEMAND WEB APP



TRANSITION TO THE CLOUD



PRICING



Q and A



SERVICES

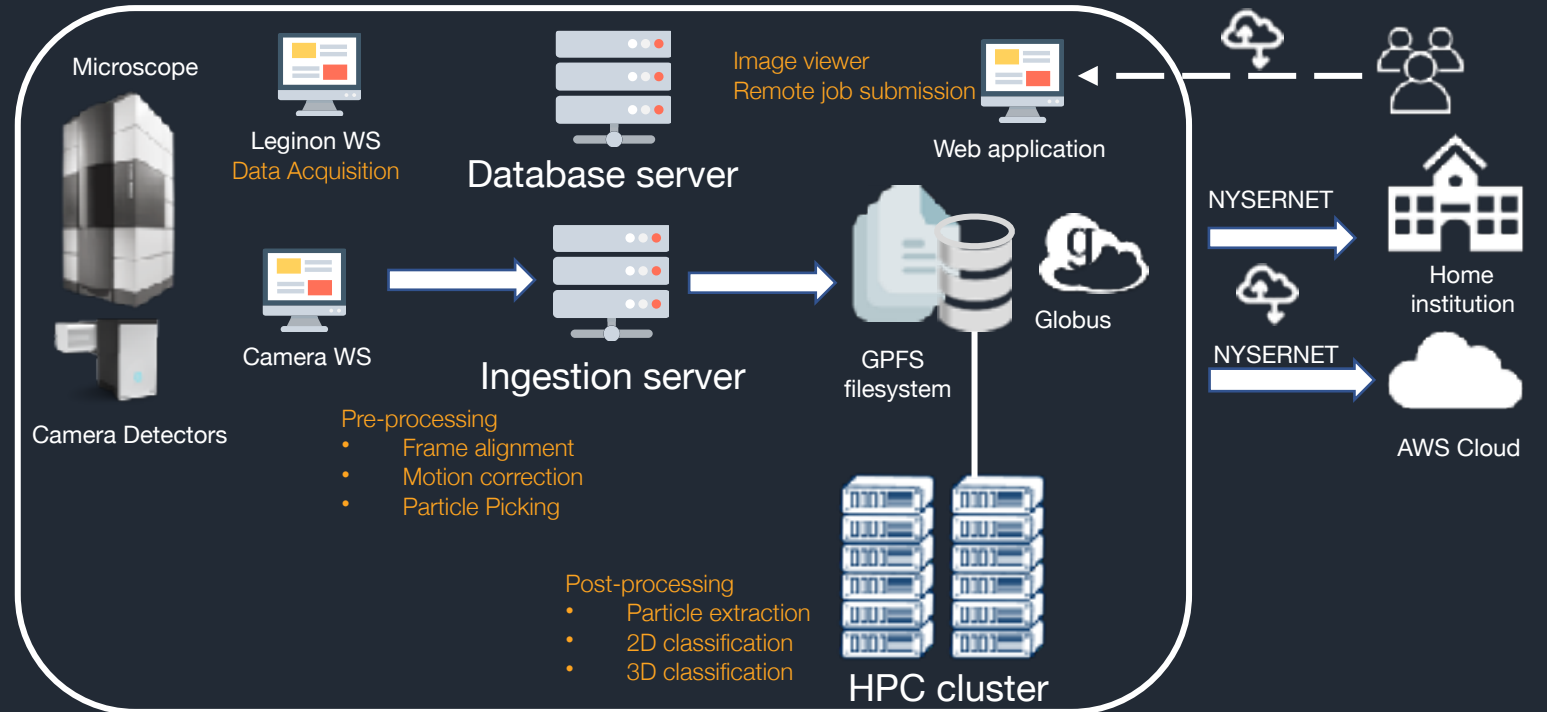
1 Sample evaluation

2 Data collection

3 Frame alignment

4 Computational training

local workflow:



STATISTICS AT A GLANCE



2 PB
Storage



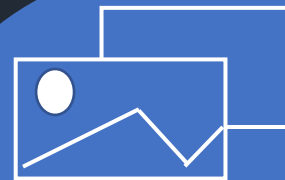
600
Users served



175
Publications
acknowledged



5 TB
Data generated
per day



4000
Images
collected per day



1M+
Database
queries per day



As of 2019

SIMONS ELECTRON MICROSCOPY CENTER
NATIONAL CENTER FOR CRYOEM ACCESS AND TRAINING



WORKSHOPS AND CREDITS

The purpose of conducting free workshops on AWS Cloud is to train the researchers to process data

\$1200

Free credits acquired to conduct workshops

6

Workshops conducted in a year so far

80%

Cost savings as compared to buying a dedicated hardware

\$110

Approx cost of 16 GPU AWS instance with SSD storage for 7 hours

*“User feedback is most valuable to us.
With every workshop we identify our bottlenecks and get better and better”*



CHALLENGES

FOR NYSBC

1.

Scalability

2.

Storage gets full

3.

Physical disruptions

FOR RESEARCHERS

1.

Lack of high-end
computational resources

2.

Limited IT support

3.

Access management
issues



TRANSITION TO THE CLOUD

ON-PREMISE

Spec out requirements
Talk to different vendors
Send quotes
Negotiations
100's of emails
Hardware setup
Configure environment
2 – 3 months



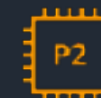
CLOUD

IAM Policies
Setup VPC
Setup S3
Setup Key Value Pairs
Enable Cloud logs
3 – 5 hours



IAC

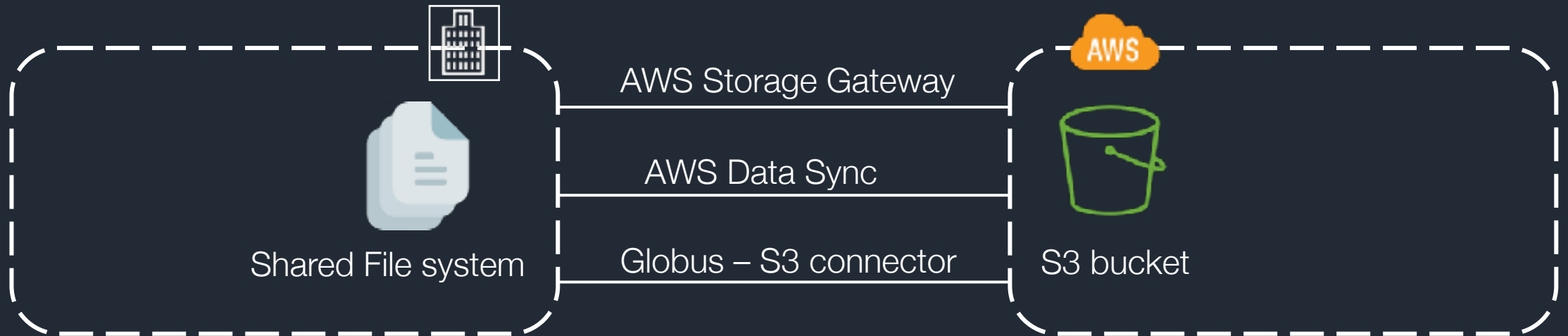
Cloud Development Kit



3 – 5 minutes



DATA TRANSFER METHODS



DATA TRANSFER: GLOBUS - S3 CONNECTOR

The screenshot displays the Globus File Manager interface. On the left is a navigation sidebar with options like 'File Manager', 'RECENTLY USED', 'BOOKMARKS', 'Activity', 'Endpoints', 'Groups', 'Console', 'Account', 'Logout', and 'Help'. The main area is titled 'File Manager' and shows two collections: 'NYSBC - SEMC' and 'globuss3test'. The 'NYSBC - SEMC' collection is selected, and its path is '/gpts/'. The 'globuss3test' collection is also selected, and its path is '/~/sbhatkar/'. Both collections show a list of files and folders. A red box highlights a blue 'Start' button at the bottom of the interface, which is used to initiate the data transfer.

Advantages:

- Lightning fast data transfers
- Simple and intuitive interface



DEMO: CryoEM ON DEMAND



Q & A



SIMONS ELECTRON MICROSCOPY CENTER
NATIONAL CENTER FOR CRYOEM ACCESS AND TRAINING

