Reusing Distributed Computing Software and Patterns for Midscale Collaborations

Benedikt Riedel
On behalf of the OSG Science Collaboration Support Team

Open Science Grid

CHEP 2019
Adelaide
November 5, 2019

Motivation & Background

In Open Science Grid we support science at all scales

- Large scale science
  - Significant software and computing teams
  - Experiment-specific workload and data management systems
- Individual investigators across many domains
  - Usually can be supported one-on-one to facilitate distributed high-throughput computing

2

Midscale challenge

- A number of collaborations fall in between - "Midscale"
  - In terms of scale - dozens of researchers and institutions, multiples of petabytes,
  - "Small" scale continuous production with CPU needs driven by episodic campaigns
- But typically have less personpower - Scientists rather than professionals
- Existing distributed infrastructure and services often present too high a barrier for their intermediate scale and immediate use

Example mid-scale collaborations

VERITAS:
4 12m Cerenkov telescopes for gamma ray astronomy: Arizona, USA

- 1st ton-scale experiment
- 3.2t of LXe, 2t in TPC
- All systems commissioned since Fall 2016
- Calibration and science data taking now ongoing

XENON1T:
Dark matter detector at Gran Sasso National Laboratory, Italy

South Pole Telescope:
Microwave-millimeter telescope

Reusing Cyberinfrastructure
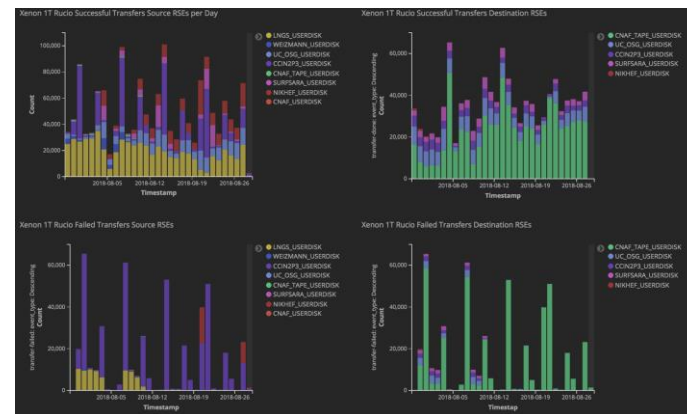
- **DATA - Rucio and FTS**
  - XENON currently has **9 storage endpoints defined in the US, Europe and Israel**
  - One of the first production deployments outside of ATLAS
  - Catalyst for the Rucio open source community (c.f. Mario's plenary)
- **Software - CVMFS**
- **Job Workload - HTCondor, Pegasus**
  - tools for OSG and EGI submission

XENON1t Experiment
Searching for dark matter candidates

PATTERN REUSE: data intensive
processing, simulation, analysis

- 1st ton-scale experiment
- 3.2t of LXe, 2t in TPC
- All systems commissioned since Fall 2016
- Calibration and science data taking now ongoing

Experiment located at the Laboratori Nazionali del Gran Sasso (LNGS), Italy
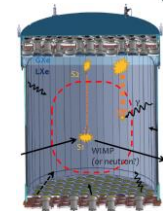
165 scientists, 25 institutions, 11 countries
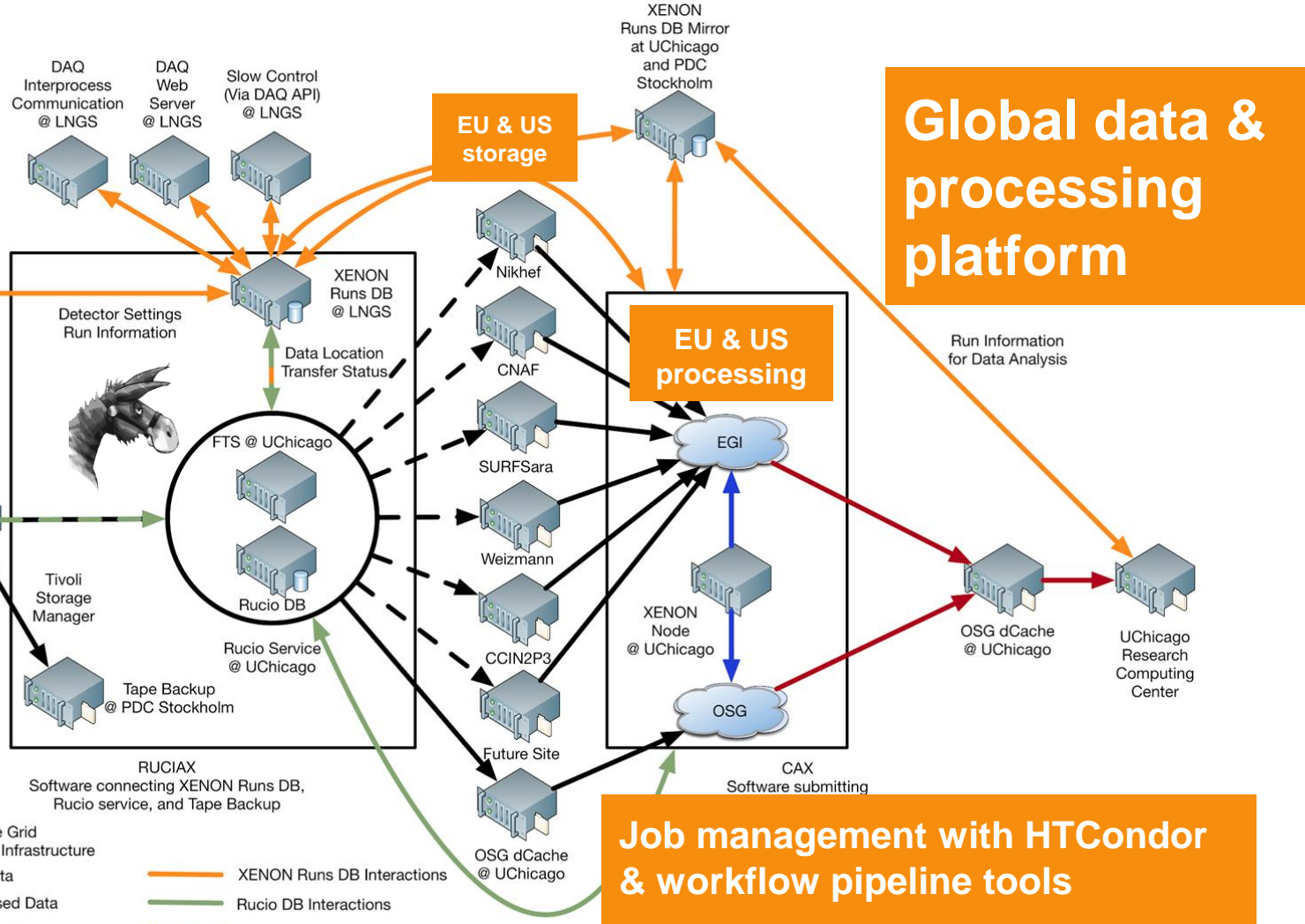
Experiment located at the Laboratori Nazionali del Gran Sasso (LNGS), Italy

Multi-institution
International scope
Resource providers from two grids

Columbia
Col

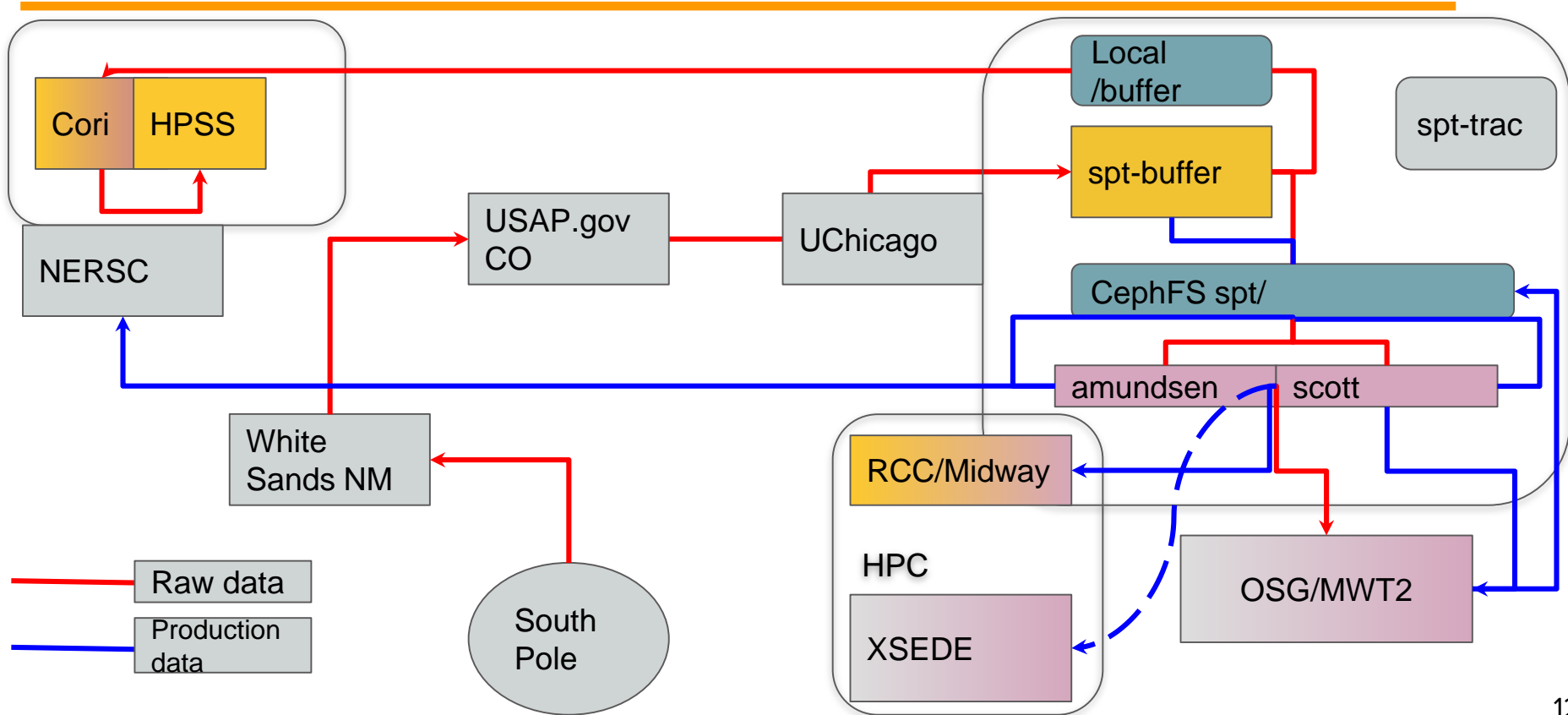CHICAGO
Ch

UCSa

Rice

PURDUE
UNIVERSITY

Purdue   Coimbra   Subatech   LPNHE   LAL   Bologna LNGS Torino Napoli   Weizmann   NYUAD

Subatech

LPNHE
PARIS

LABORATOIRE
DE L'ACCÉLÉRATEUR
LINÉAIRE

INFN

NYU ABU DHABI

Kobe

Global data & processing platform

EU & US storage

EU & US processing

Job management with HTCondor & workflow pipeline tools

XENON Runs DB Mirror at UChicago and PDC Stockholm

DAQ Interprocess Communication @ LNGS

DAQ Web Server @ LNGS

Slow Control (Via DAQ API) @ LNGS

XENON DAQ @ LNGS

Detector Settings Run Information

XENON Runs DB @ LNGS

Data Location Transfer Status

Data Storage @ LNGS

Tivoli Storage Manager

FTS @ UChicago

Rucio DB

Rucio Service @ UChicago

Tape Backup @ PDC Stockholm

RUCIAX
Software connecting XENON Runs DB, Rucio service, and Tape Backup

Nikhef

CNAF

SURFSara

Weizmann

CCIN2P3

Future Site

OSG dCache @ UChicago

EGI

XENON Node @ UChicago

OSG

CAX
Software submitting

Run Information for Data Analysis

OSG dCache @ UChicago

UChicago Research Computing Center

OSG = Open Science Grid
EGI = European Grid Infrastructure

|	
Raw Data	XENON Runs DB Interactions
Processed Data	Rucio DB Interactions
Random Selected Element	Jobs

10

South Pole Telescope

- Job submission infrastructure with OSG tools
- Managed data transfer and management
  - South Pole, Chicago, and archival in Berkeley
- Software distribution via CVMFS
- OSG flocking

SPT Data flow

Cori | HPSS

NERSC

Local /buffer

spt-trac

spt-buffer

USAP.gov CO

UChicago

CephFS spt/

amundsen | scott

White Sands NM

RCC/Midway

HPC

XSEDE

OSG/MWT2

Raw data

Production data

South Pole

12

VERITAS

- Job submission infrastructure based on Pegasus WMS
- CVMFS for software distribution
- Ceph storage and pre/post job processing
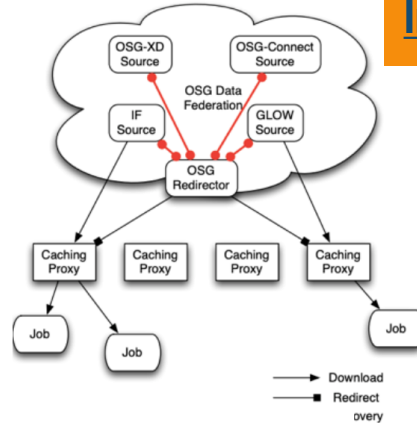- RADOS gateway as an interface to AWS S3 RESTful API

Wall Hours by Site per 1M
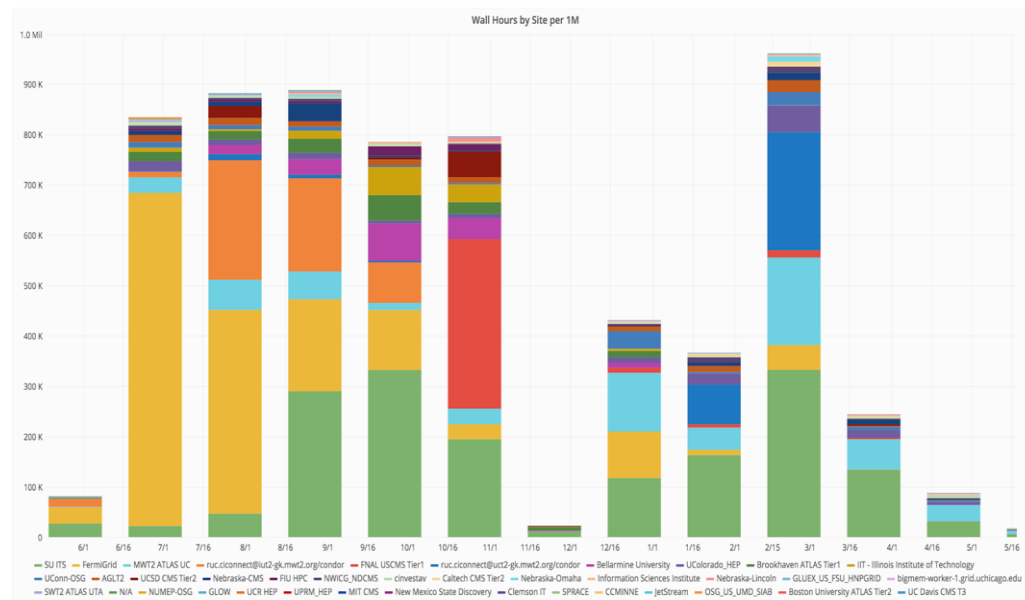
Another successful effort - Creating Science Caching Networks

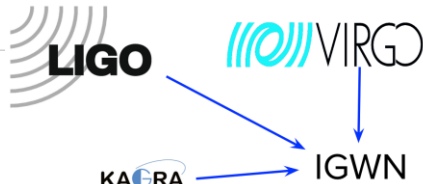## Introduction to Stashcache

- Caching infrastructure based on SLAC XRootD server & XRootD protocol.
- Cache servers are placed at several strategic cache locations across the OSG.
- Jobs utilize GeoIP to determine the nearest cache
- Job talks to the cache using HTTP(S) via CVMFS

Powered by:

XRootD

Image taken from Brian's slide

LIGO, Virgo, and Kagra are now in a collaboration-of-collaborations, and are working towards a data analysis common computing infrastructure called IGWN (for the International Gravitational-Wave Observatory Network)

c.f. "Creating a content delivery network for general science on the backbone of the Internet using XCache(s)"   I.Sfligoi

Mr Worldwide

- Cache at institution
- Cache in the backbone

Given the worldwide location of the IGWN computing resources a need rise to expand beyond the US. This lead to a worldwide network of caches.

Summary & going forward

- Re-use of software infrastructure, services and tools built for the LHC can be repurposed
  - there is significant demand for this
  - comes with particular challenges - technical and cultural
- CVMFS, Rucio, FTS, XrootD/Caching, OSG, EGI infrastructure re-used
- Going forward - keeping multi-domain & reuse can pay dividend down the road
- "Midscale" experiments a new focus in US