From Grids to Clouds

Ian Fisk
CERN openlab Summer School
July 19, 2016

The phrase that pays

Follow the money

GRID
2000-2010

CLOUD
2010-2016

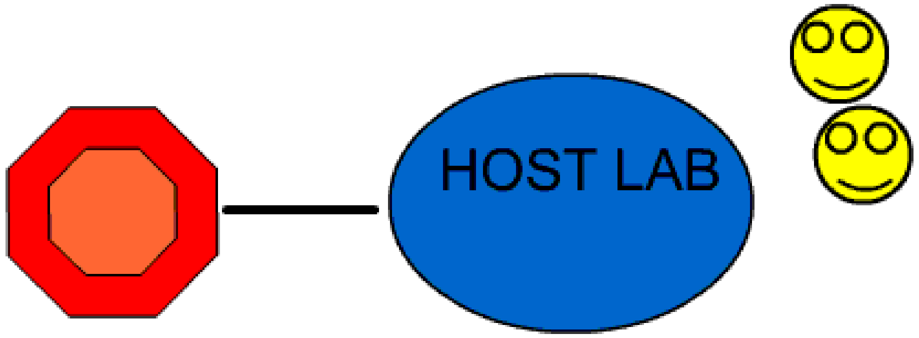
?

Open
Science
Cloud

"Technical" Choices

- A lot of the choices we make are motivated by non-technical reasons
  - What development can be supported at a particular moment in time
  - Where people choose to work and where people choose to invest
- Some choices are motivated by a need to scale at a determined or undetermined time in the future
- Some choices are designed to push R&D in distributed computing that might be generally beneficial
- As we discuss Grids and Clouds you will see that sometimes the simplest solution is not the one chosen

Beginning



In the beginning the computing was centralized

Experiments began to develop distributed computing models

- ➡ Two examples: Babar had Tier-As that users could connect to for access to the data and resources. CDF had distributed analysis centers
- ➡ Distributed centers tended to come later as other items were better understood

MONARC

All LHC Grid Computing Models are based on MONARC

- Introduced the idea of hierarchical tiers of computing centers
- Assumes poor networking on connectivity between sites

Motivated by investment

- Countries were more willing to invest in local computing and local infrastructure
- Rely on pool of distributed computing expertise

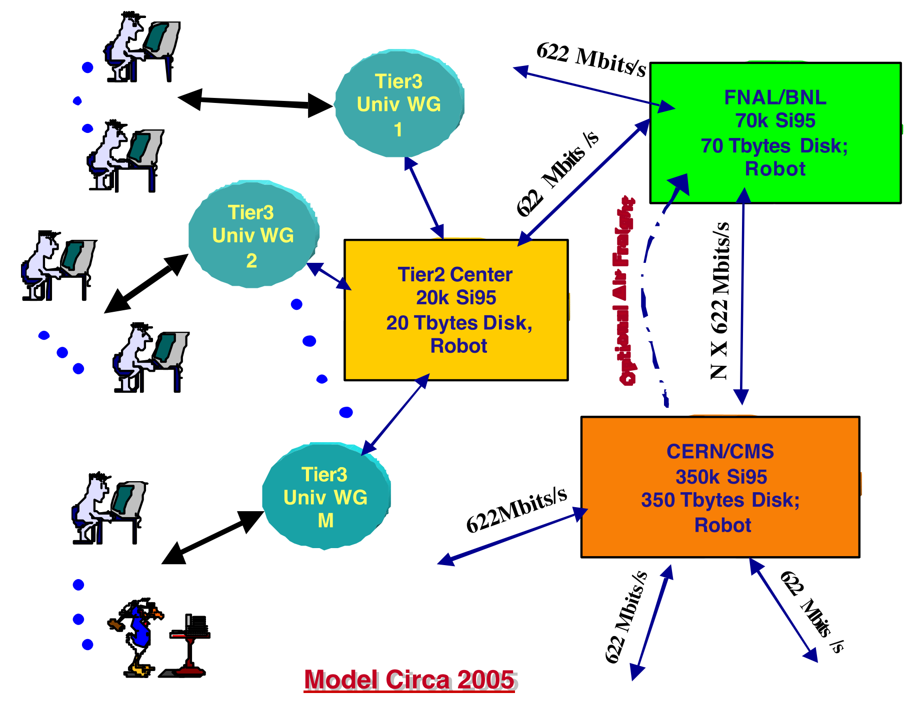
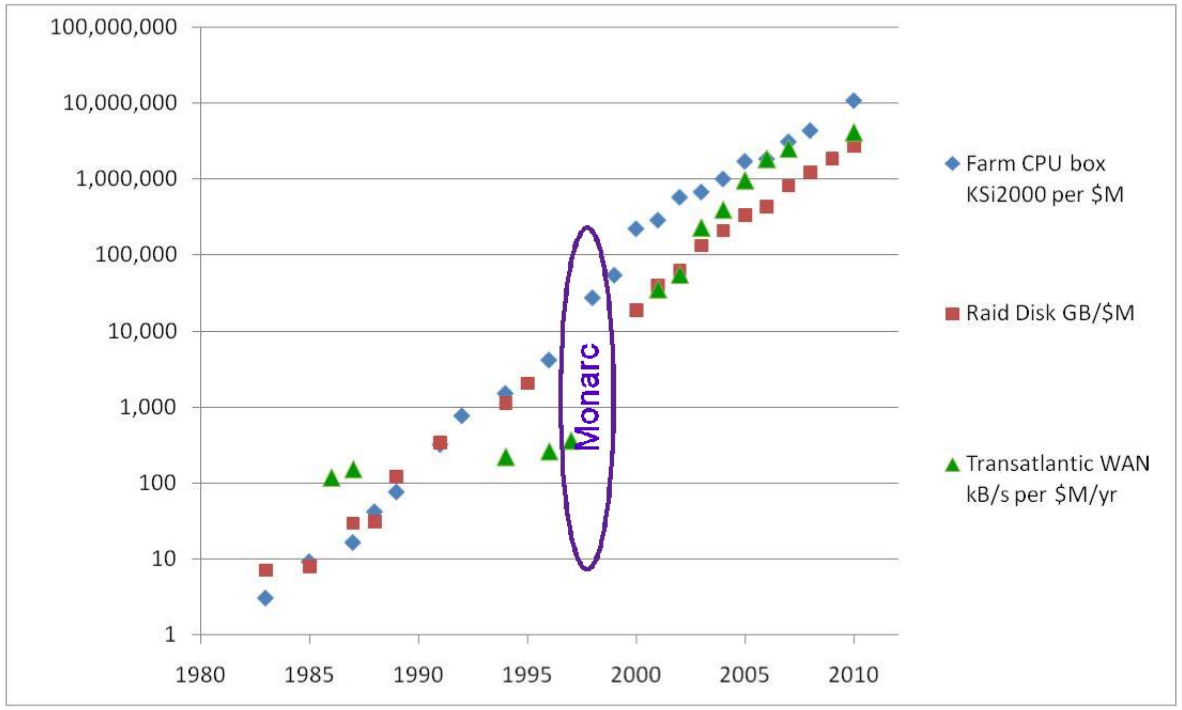


Fig. 4-1 Computing for an LHC Experiment Based on a Hierarchy of Computing Centers. Capacities for CPU and disk are representative and are provided to give an approximate scale).

- 16 -



LHC Computing Models

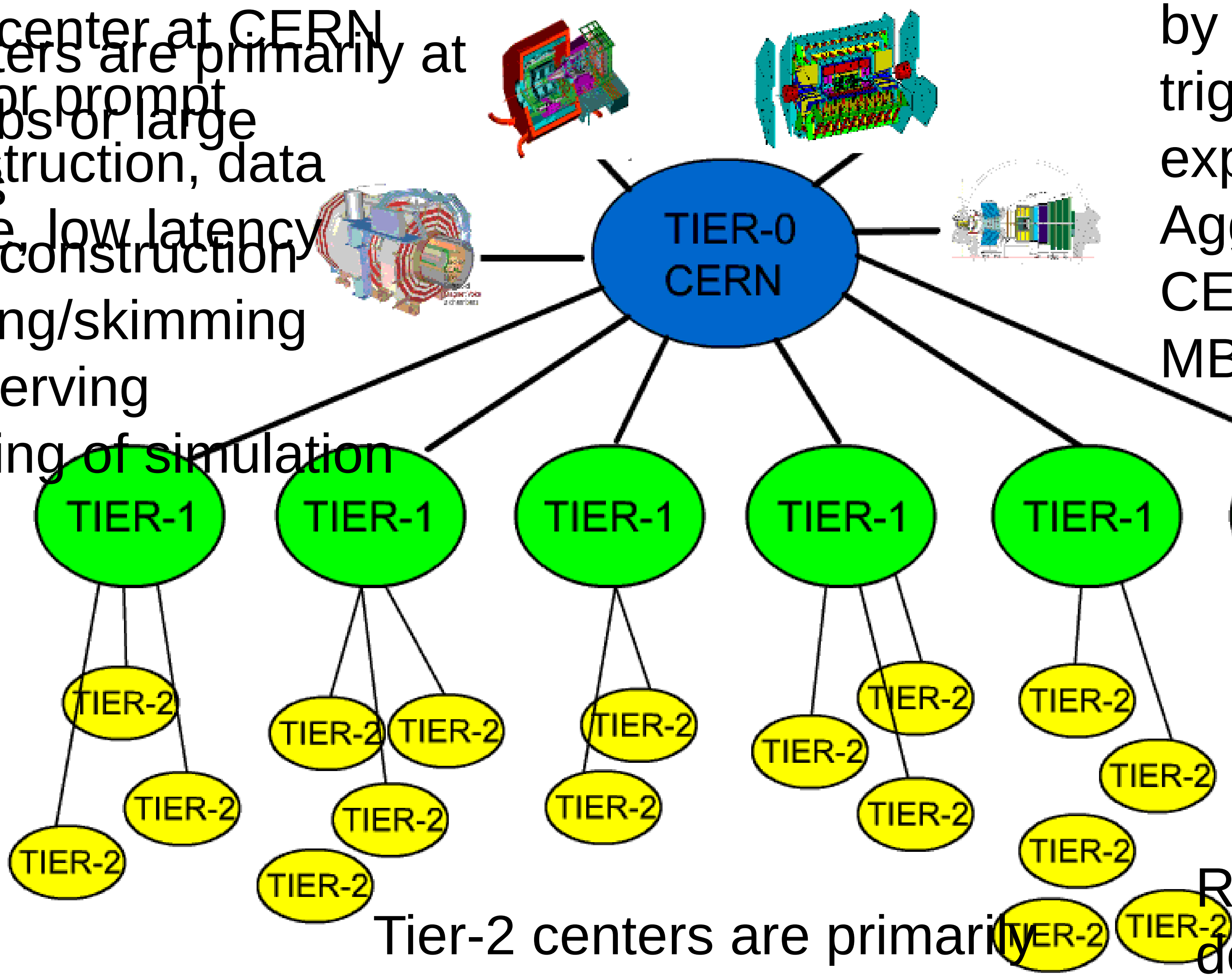
Tier-0 center at CERN used for prompt reconstruction, data archive, low latency work

Tier-1 centers are primarily at national labs or large universities

- Re-Reconstruction
- Stripping/skimming
- Data serving
- Archiving of simulation

Rate to Tier-1 varies by event size and trigger rate for each experiment.

Aggregate rate from CERN of hundreds of MB to nearly 1GB



Tier-2 centers are primarily at universities

- Simulation
- User Analysis

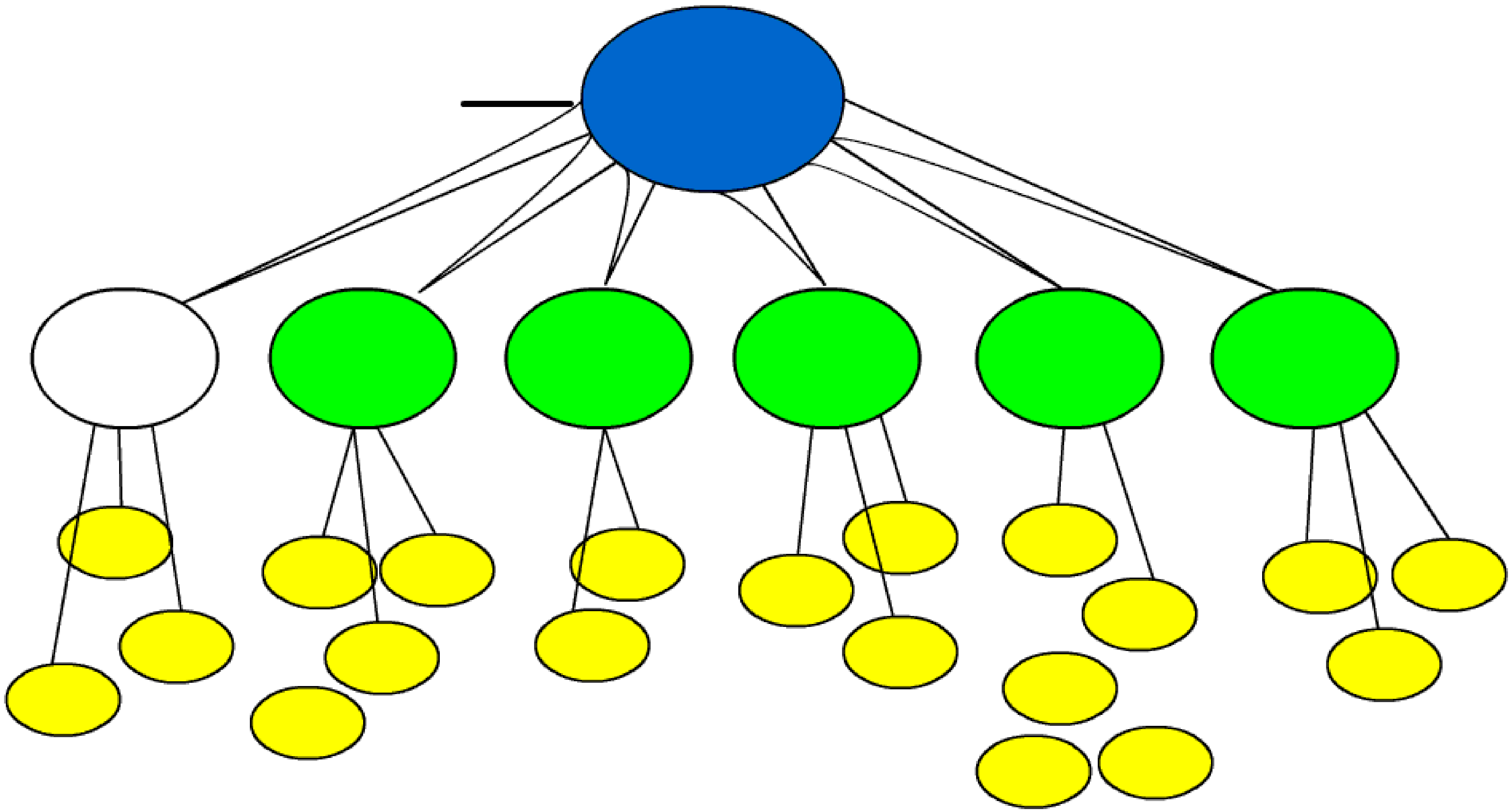
Rate to Tier-2s depends on the experiment and the expectations for updating storage

- Can burst with activity

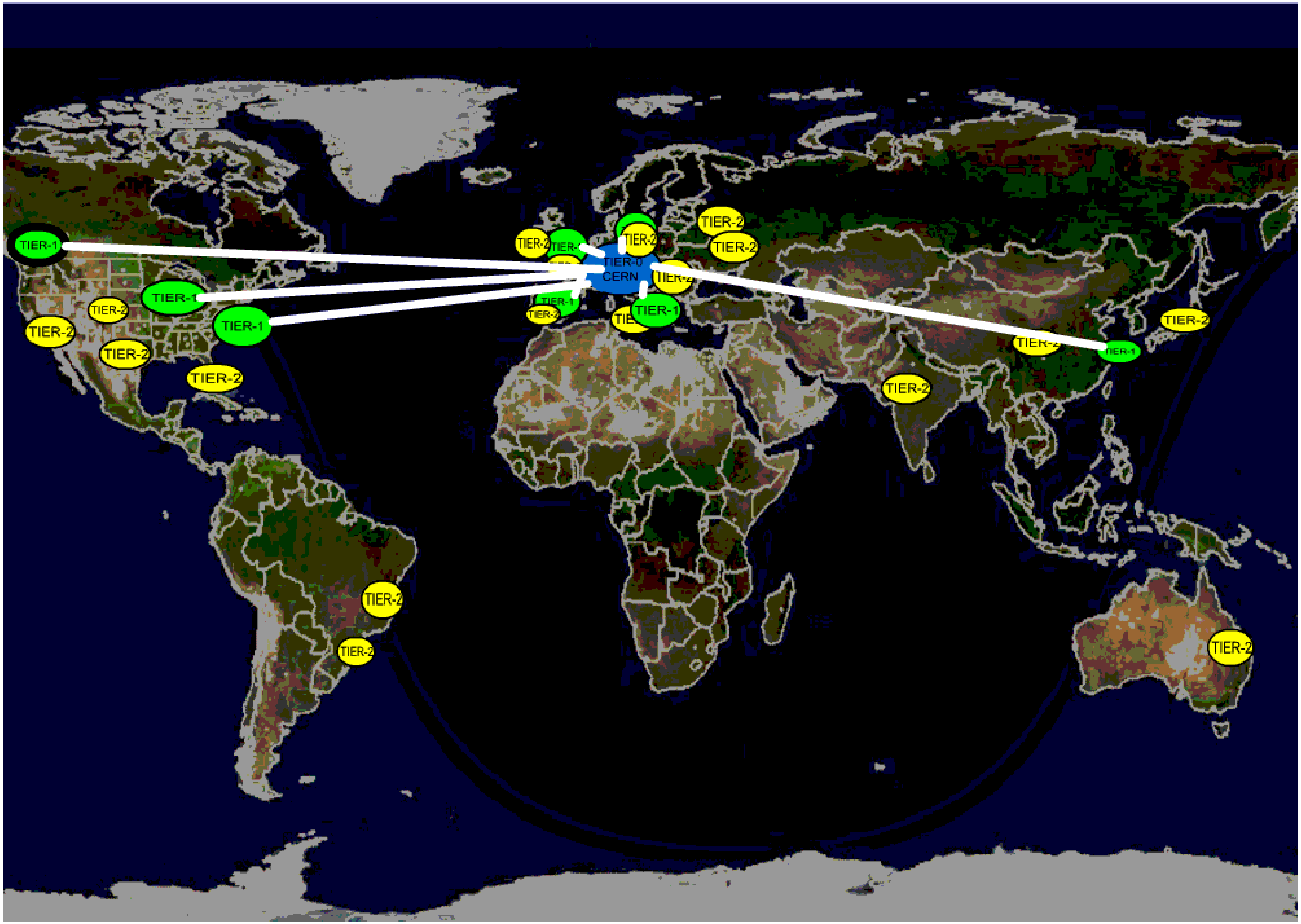
MONARC Tiered computing model came in the late 90's

➡ Level of distribution motivated by the desire to empower and leverage resources and to share load, infrastructure, and funding

LHC Computing Models



Networking

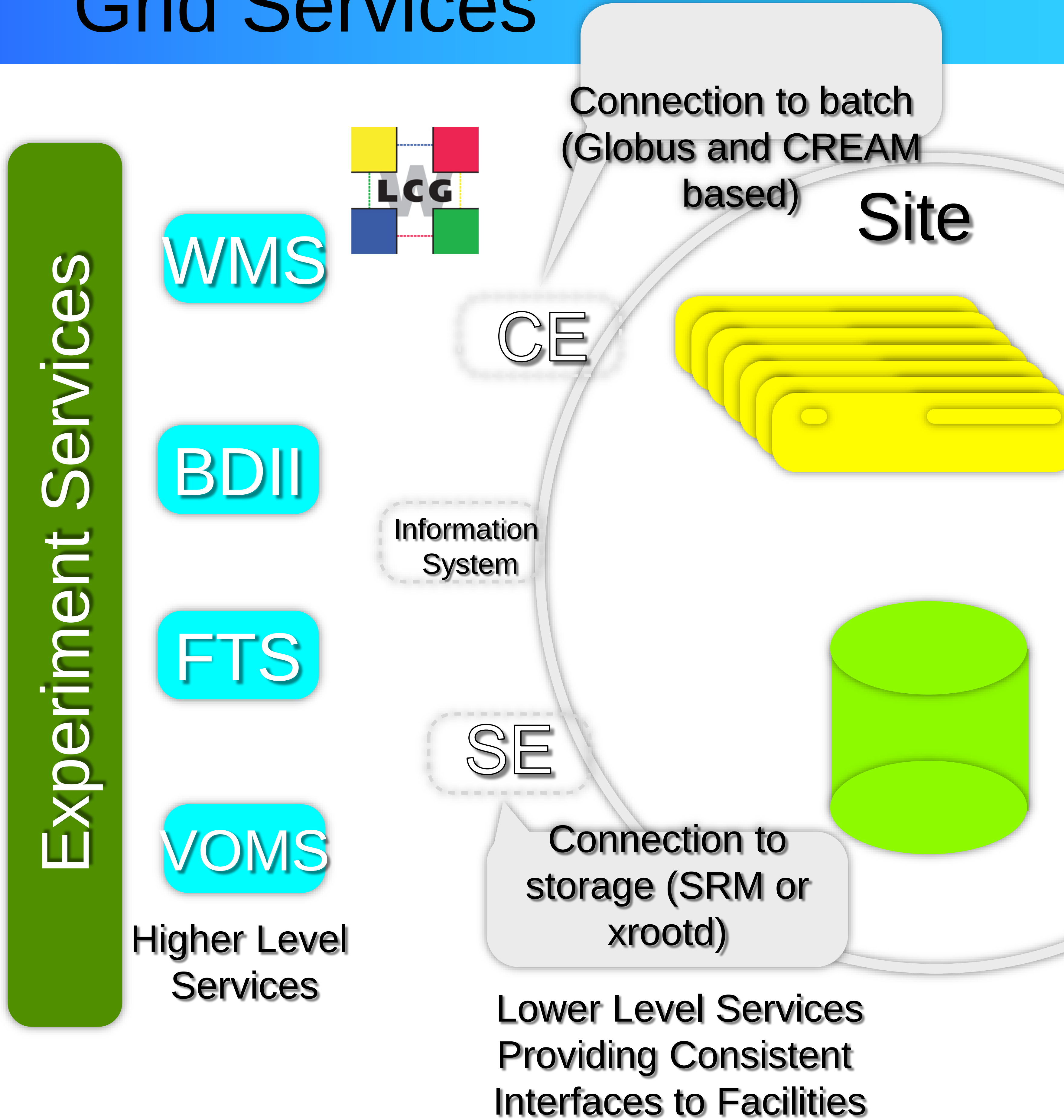


Optical Private Network (OPN) connects CERN and Tier-1. Other connections handled by shared networks

Grid Services

During the evolution the low level services are largely the same

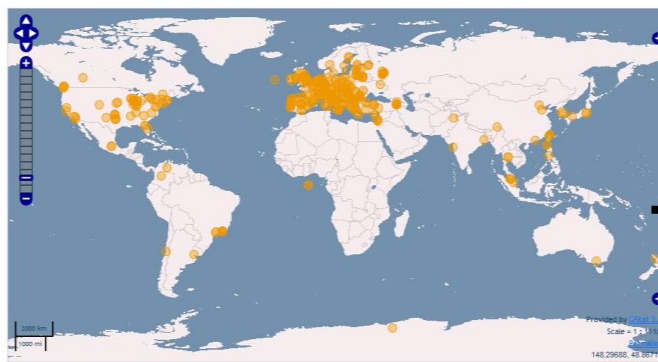
Most of the changes come from the actions and expectations of the experiments



Problems with the Grid

A lot of services have to function to successfully execute a job

Much of the development effort has been to shield this complexity from the user



→ bsub

- Local Environment Packaged
- Choice of site(s) with the desired data files or resources
- Success submission through site grid interface
- Arrival at Batch farm
- Authenticate User and VO
- Find local environment
- Discovery and open local data file or remote file
- Communicate job status and write results somewhere

Reliability and Robustness

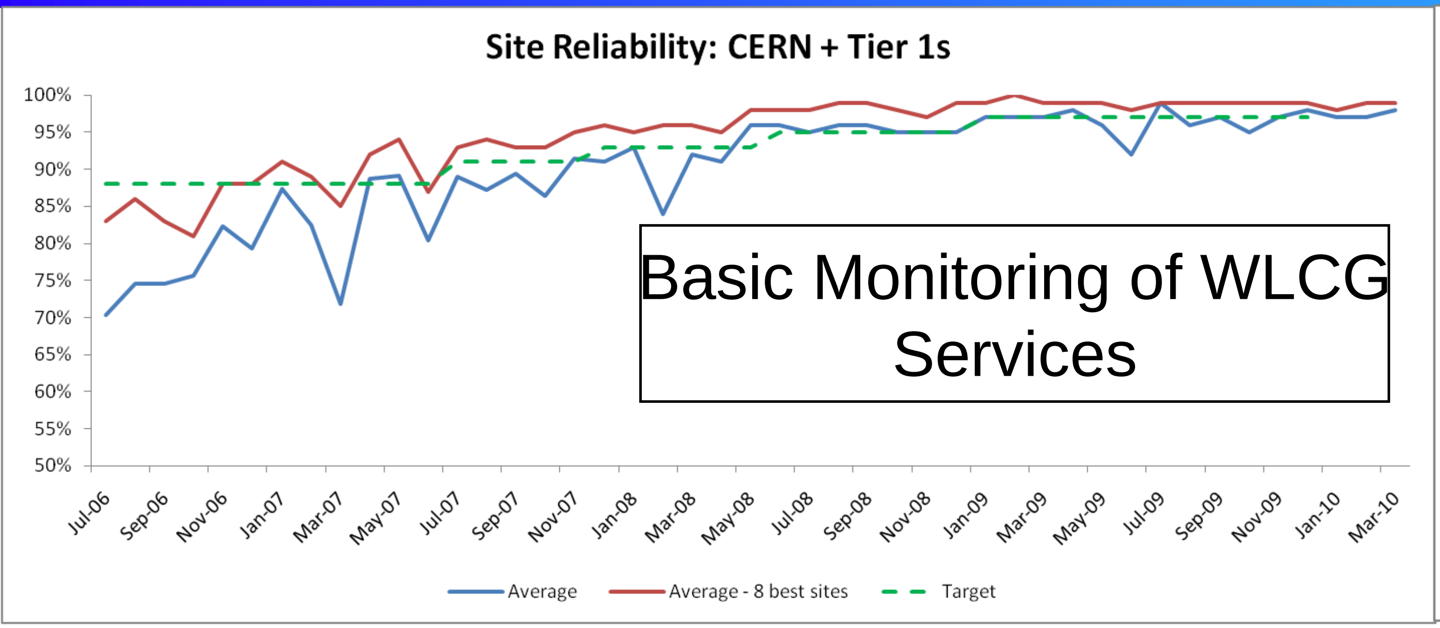
The level of distribution and the number of services requires an advanced system to check the health of the globally distributed system

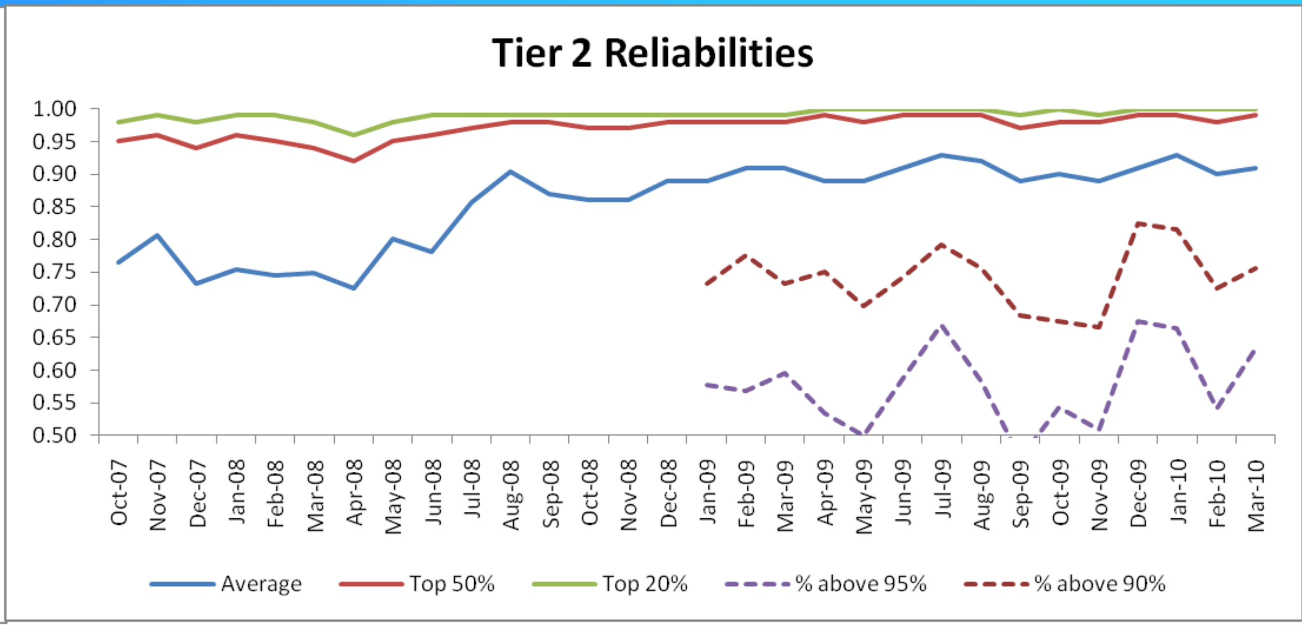
- ➡ WLCG has developed a series of Site Availability Monitors (SAM) tests
- ➡ Series of automatically submitted and tracked tests
  - Validate the processing services all the way down to worker nodes
  - Validate storage services
  - Information systems
- ➡ Tests run every few hours and results are tracked and published

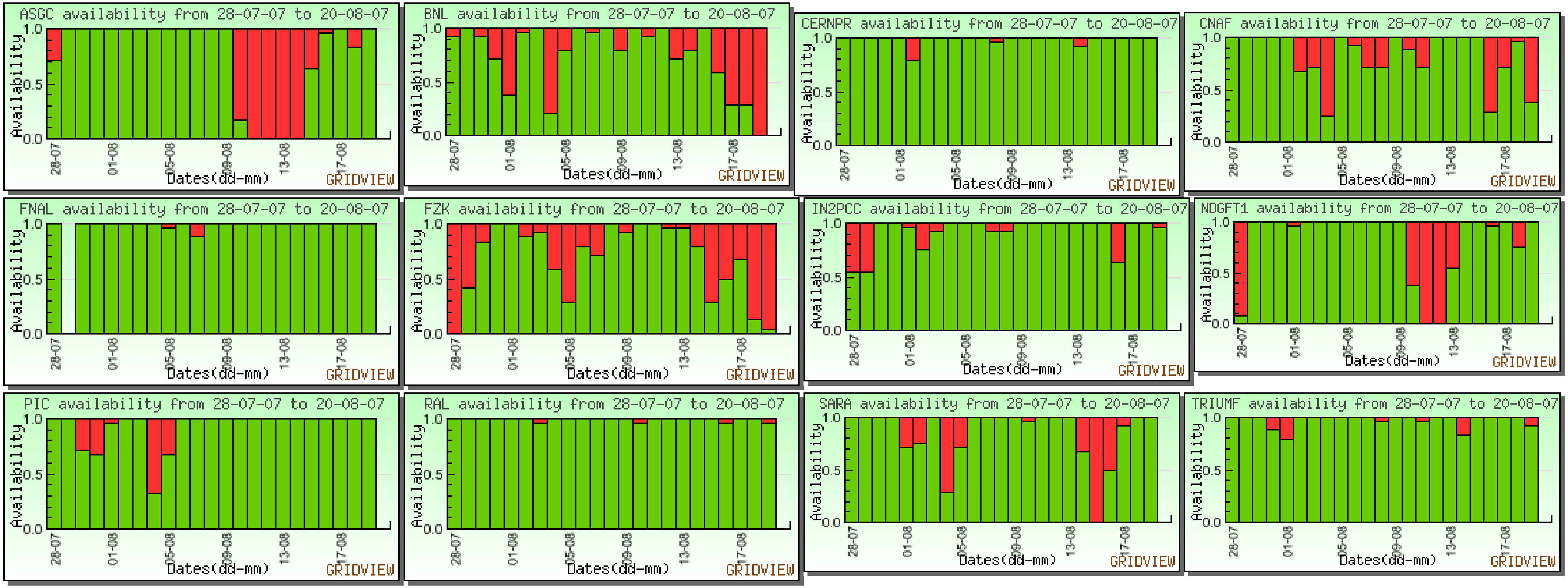
Experiments (VOs) also introduced their own tests

- ➡ Verify the experiment workflows within the SAM framework
- ➡ Utilize the experiment submissions systems to update the SAM tests

Results


Site Reliability: CERN + Tier 1s
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
Jul-06 Sep-06 Nov-06 Jan-07 Mar-07 May-07 Jul-07 Sep-07 Nov-07 Jan-08 Mar-08 May-08 Jul-08 Sep-08 Nov-08 Jan-09 Mar-09 May-09 Jul-09 Sep-09 Nov-09 Jan-10 Mar-10
Average
Average - 8 best sites
Target
Basic Monitoring of WLCG Services


Tier 2 Reliabilities
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
Oct-07 Nov-07 Dec-07 Jan-08 Feb-08 Mar-08 Apr-08 May-08 Jun-08 Jul-08 Aug-08 Sep-08 Oct-08 Nov-08 Dec-08 Jan-09 Feb-09 Mar-09 Apr-09 May-09 Jun-09 Jul-09 Aug-09 Sep-09 Oct-09 Nov-09 Dec-09 Jan-10 Feb-10 Mar-10
Average
Top 50%
Top 20%
% above 95%
% above 90%


ASGC availability from 28-07-07 to 20-08-07
BNL availability from 28-07-07 to 20-08-07
CERNPR availability from 28-07-07 to 20-08-07
CNAF availability from 28-07-07 to 20-08-07
FNAL availability from 28-07-07 to 20-08-07
FZK availability from 28-07-07 to 20-08-07
IN2PCC availability from 28-07-07 to 20-08-07
NDGFT1 availability from 28-07-07 to 20-08-07
PIC availability from 28-07-07 to 20-08-07
RAL availability from 28-07-07 to 20-08-07
SARA availability from 28-07-07 to 20-08-07
TRIUMF availability from 28-07-07 to 20-08-07
Availability
Dates(dd-mm)
GRIDVIEW

Now what?

So now you have a consistent set of sites with a consistent way to communicate with them

- You still need
  - A way to distribute the software environment
  - A way to get common information like conditions
  - A way to track and manage the input and output data

Distributing the Software Environment

At the start of Run 1 there were more solutions for software environment deployment than experiments

- ➡ Some used grid jobs to deploy the environment
- ➡ Site admins installed the software locally to NFS at some sites

BitTorrent used by ALICE

AFS used as a local file system and regionally between sites

Many of the solutions were seen as non-scalable, operationally intensive, and/or with high-latency

A better solution was sought

Courtesy Maria Girone, CHEP 2015

HEP Software Distribution and CernVM-FS

Developed (outside the Grid) for Cern Virtual Machines

Ideal for replicating the software environment to sites

- Minimization of file transfers
- Aggressive caching
- Deduplication and optimal identification of changes
  - Only 10% of new files between releases
- Optimized encapsulation of metadata to offload to clients expensive operations (e.g. ls, stat)

CernVM-FS (gradually) adopted by the Grid

- ATLAS was an early adopter

In 2012, the WLCG Operations Technical Evolution Group recommended it

Summary of Recommendations

Name	Description	Effort	Impact
R3.2	Software deployment via CVMFS	Moderate	Significant

CVMFS:
http://cernvm.cern.ch/portal/filesystem

M. Girone and J. Templon, Final Report on the Operations and Tools TEG http://wlcg.web.cern.ch/news/teg-reports

Courtesy Maria Girone, CHEP 2015

CVMFS Architecture

Central publication point (Stratum-0) R/W

Minimal transfer protocol requirements (HTTP)

Aggressive hierarchical cache strategy for scalability

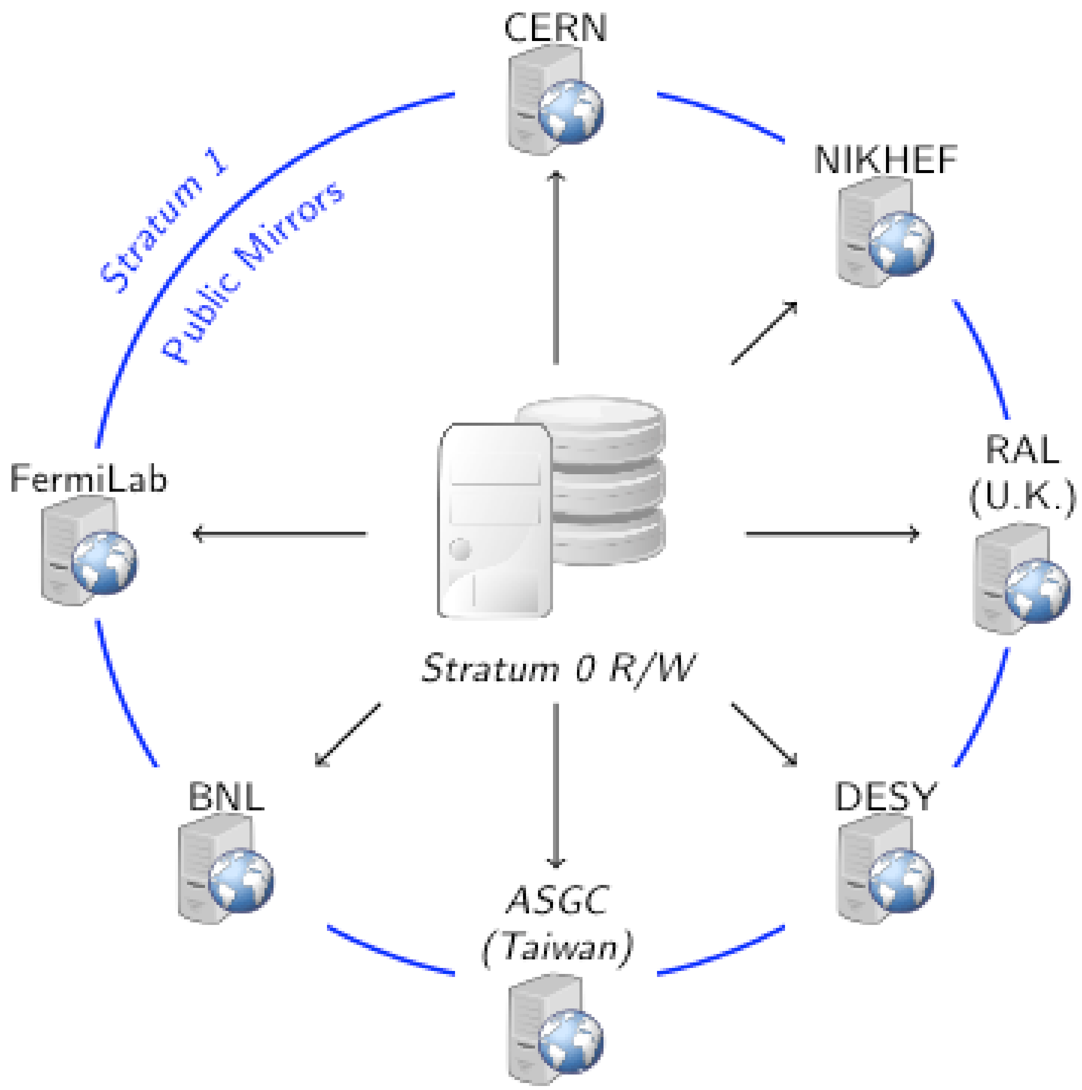
- ➡ Stratum-1, squid at local sites, read-only

POSIX mount point on clients

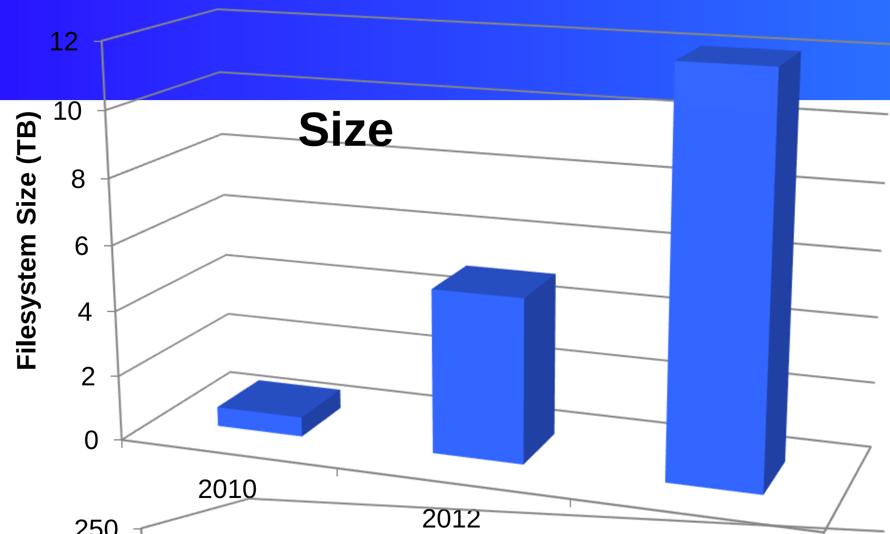
- ➡ FUSE, local NFS share, Parrot

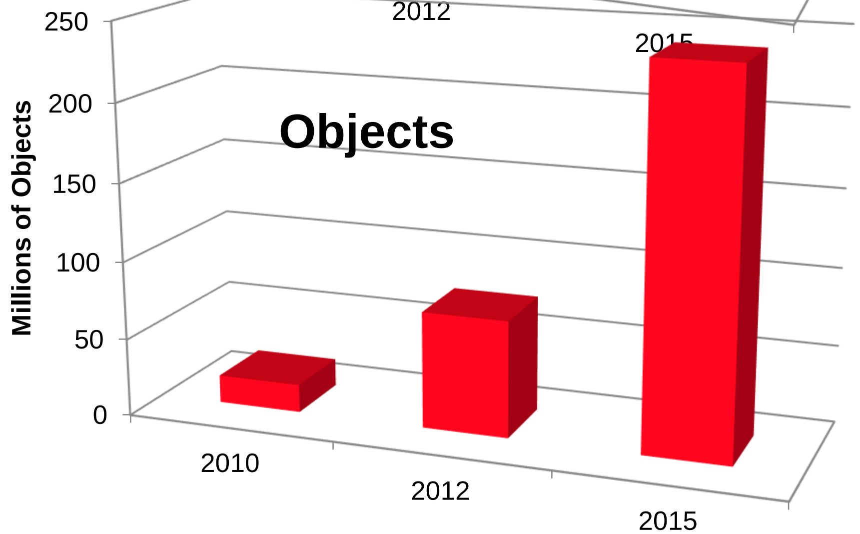
Automatic versioning

- ➡ "Time-machine" for experiment software
- ➡ E.g. Impact on data preservation



CVMFS Scale





For 5 years the contents of CVMFS have grew linearly

Number of experiments using the system continuously increasing

- CERN and EGI stratum-0 host more than 30 repositories, including non-HEP experiments

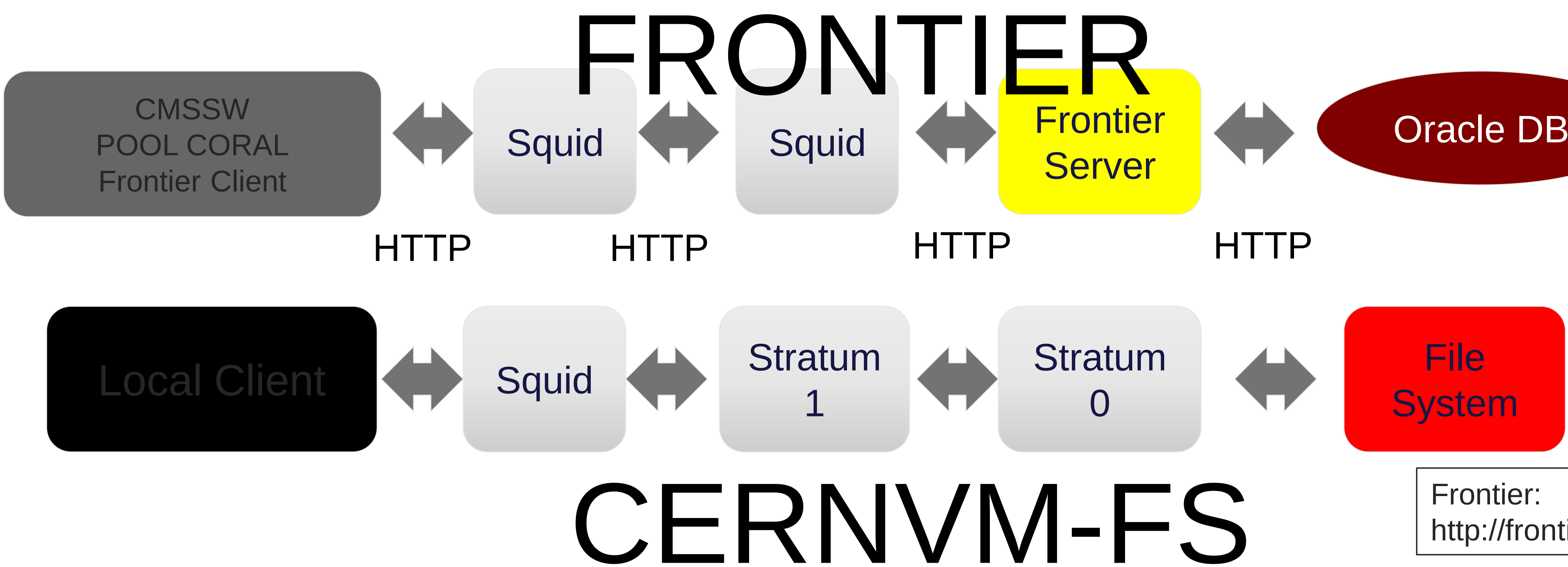
CVMFS has spread to 5 continents and is used on all WLCG resources

- There are at least 64k nodes at 160 sites
- Is now a critical service in WLCG

Frontier

Frontier is an earlier example of introducing independent services (Distributed Database Cache as a Service)

Before Frontier many Tier-1 sites operated databases for the local processing needs



Frontier:
http://frontier.cern.ch

Architecture of Frontier and CVMFS are similar

Maria Girone, CHEP 2015

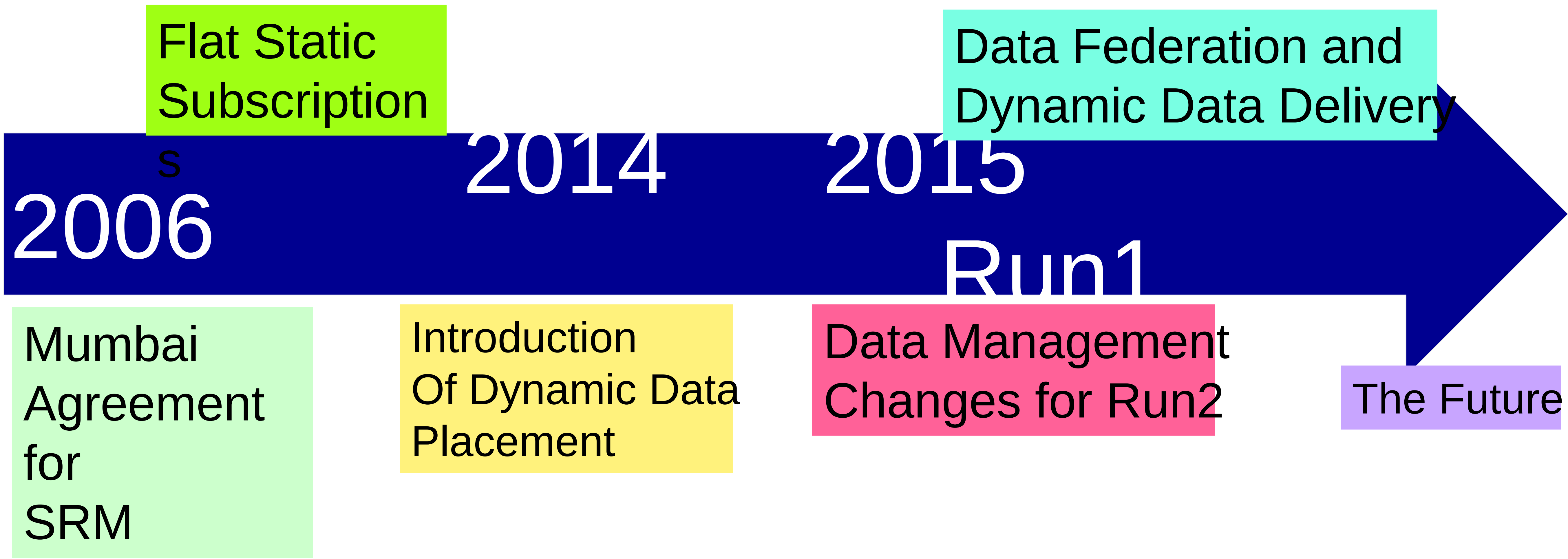
Software Distribution vs Data Management

	Software Distribution	Data Management
Size of samples	~10TB	~100PB
Level of Replication	All sites	Average sample replication factor 2-3
Latency	Full synchronization in 1 hour	Completing a replica can take a week
Update rate	Packages are updated frequently (incl. nightly)	New datasets are created less frequently

Evolution of LHC Data Management

Key stages marking the path to evolution of Data Management

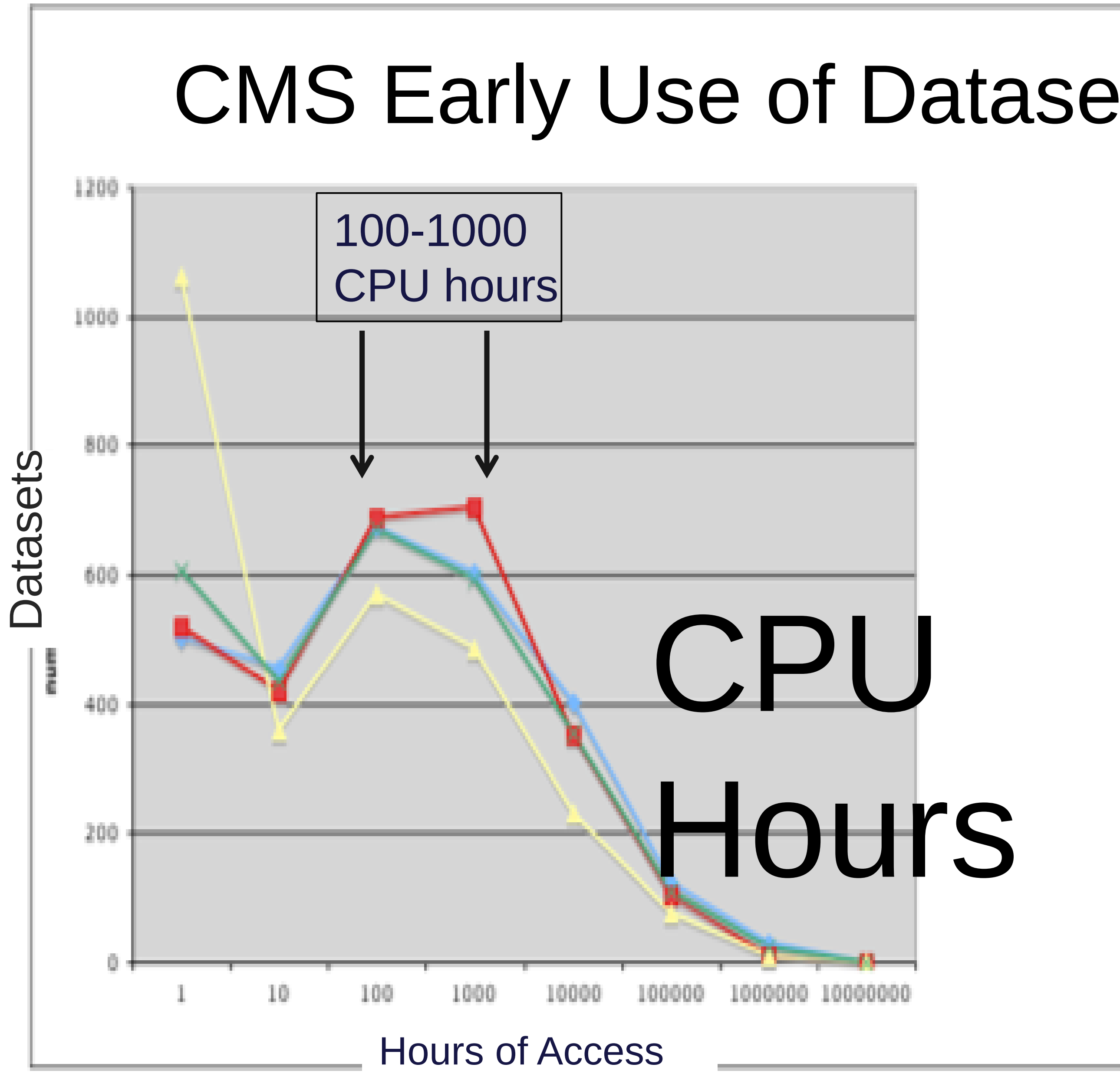
Starting from tight services and static models, moving towards **decoupling** and **dynamism**



Flat Static Subscriptions

The primary method for pushing data to sites is by subscription

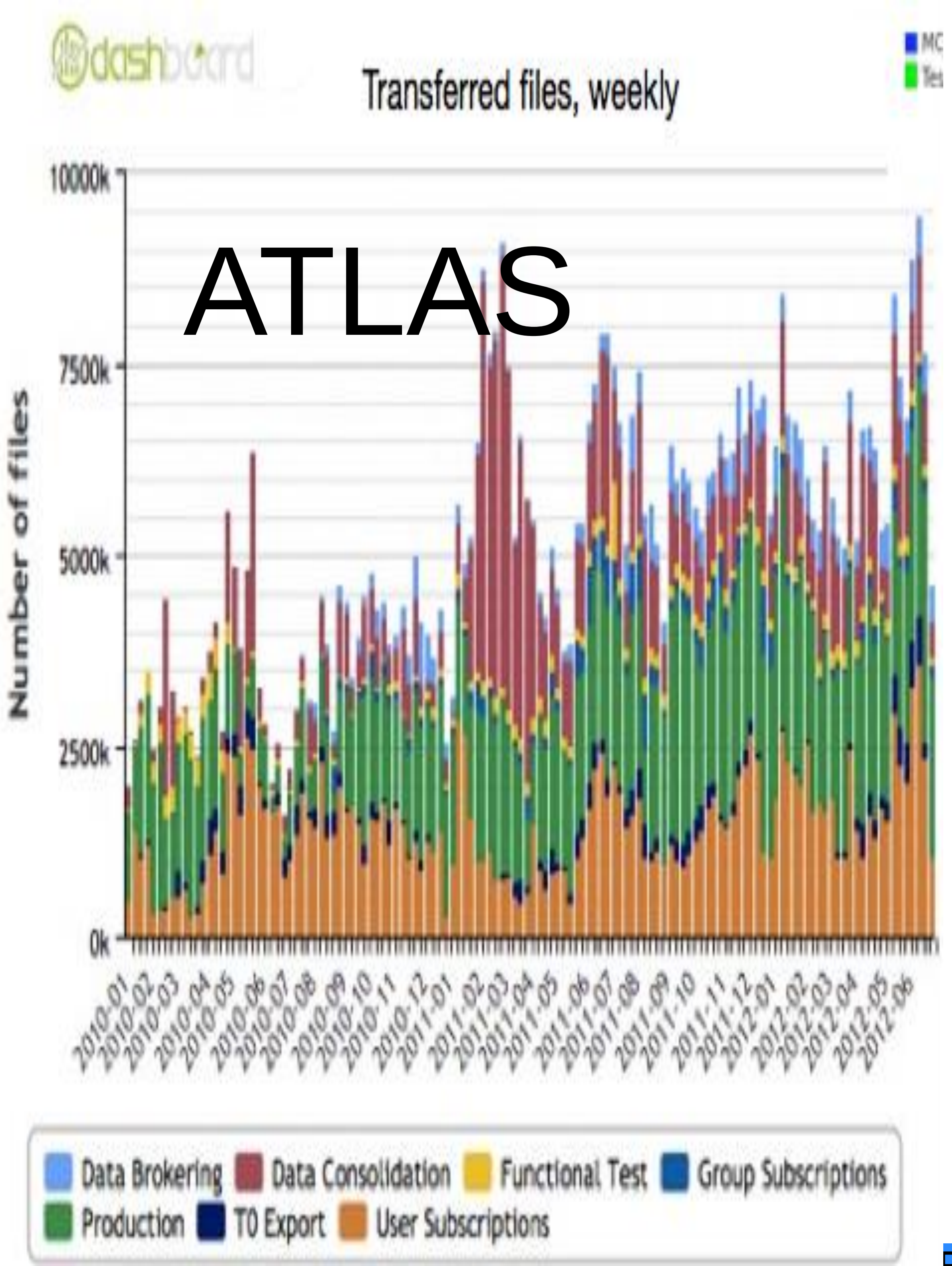
- ➡ Processing and storage are coupled and only data available locally is visible



Flat static subscriptions assume that most samples have a similar number of access, which unfortunately is wrong

Maria Girone, CHEP 2015

Introduction of Dynamic Data Placement



ALICE and ATLAS developed the Dynamic Data Placement that deploys samples in response to changing processing demands

- The system is still based on subscriptions
  - made when needed and removed when finished

ATLAS

- Re-brokering allows jobs to move to another site if the first one is underperforming

ALICE

- Goes to nearest replica based on network information

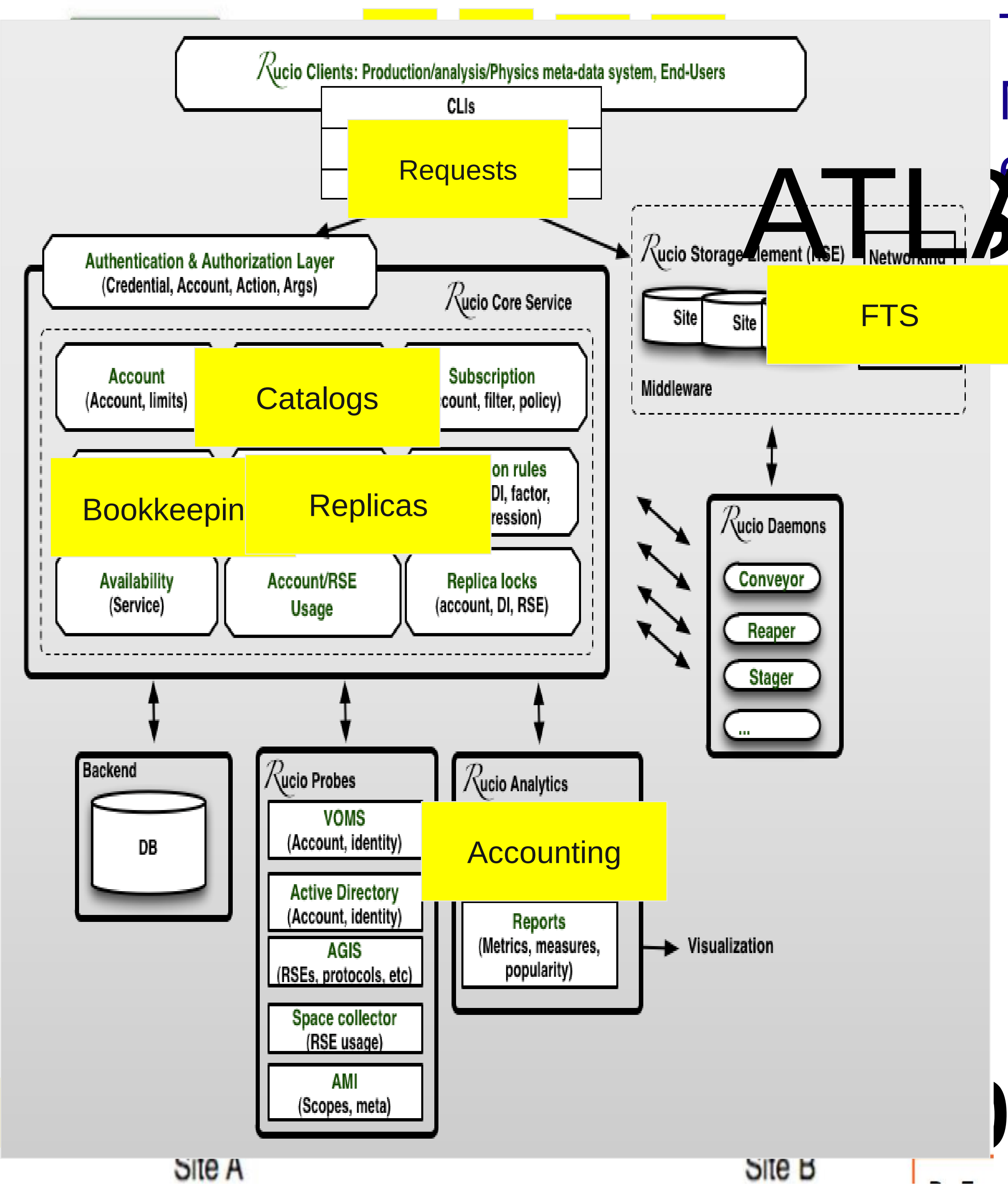
Data Management Commonalities

Each LHC experiment has developed a data management solution

There is a lot commonality in the underlying services and design elements

Service	ALICE	ATLAS	CMS	LHCb
File Transfer Tool	Xrootd	FTS/ SRM	FTS/ SRM	FTS/ SRM
Technology for Catalogs	MySQL	Central Oracle	Central Oracle	Central Oracle
Information System	ALIEN	AGIS	SiteDB	Dirac from BDII
Primary File Access • Local Access	✓Xrootd	✓Misc.	✓Misc.	✓Xrootd
• Copy to disk		✓Misc.		✓SRM
• Served Remotely	✓Xrootd	✓Xrootd	✓Xrootd	✓Xrootd

Data Management in Run 2



ATLAS

The functional elements of the Data Management system are similar for the 4 experiments

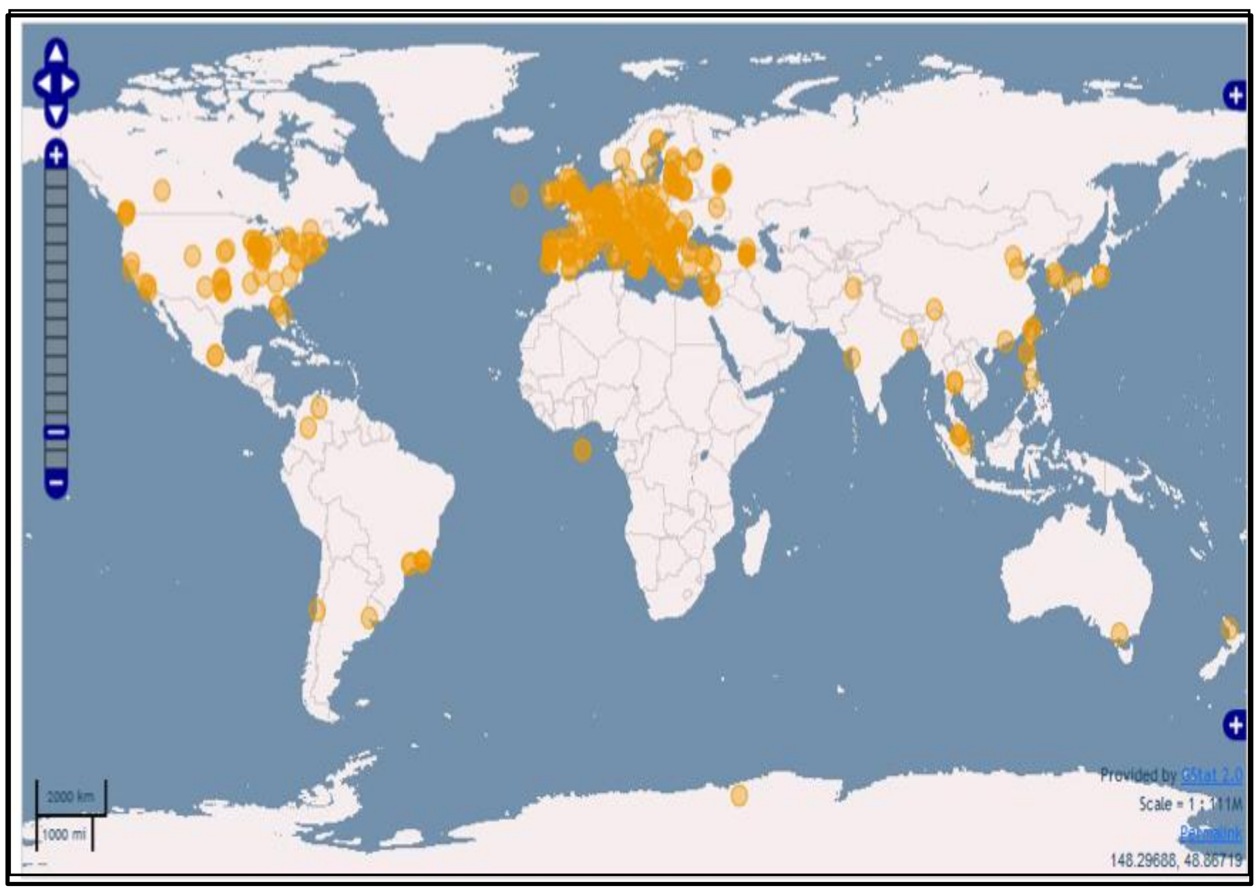
- Bookkeeping – how files relate to each other and what samples are contained (metadata)
- File catalogs – the list of files
  - The independent file catalog LFC has largely been replaced by replica services integrated in the experiments DM system
- Replica catalogs – if a dataset is subscribed in multiple locations
- FTS – The file transfer system for moving files
- Requests - Interface to request subscriptions and file movement
- Accounting – system for generating reports of usage

The Data Management Problem

There are close to 200 sites in WLCG

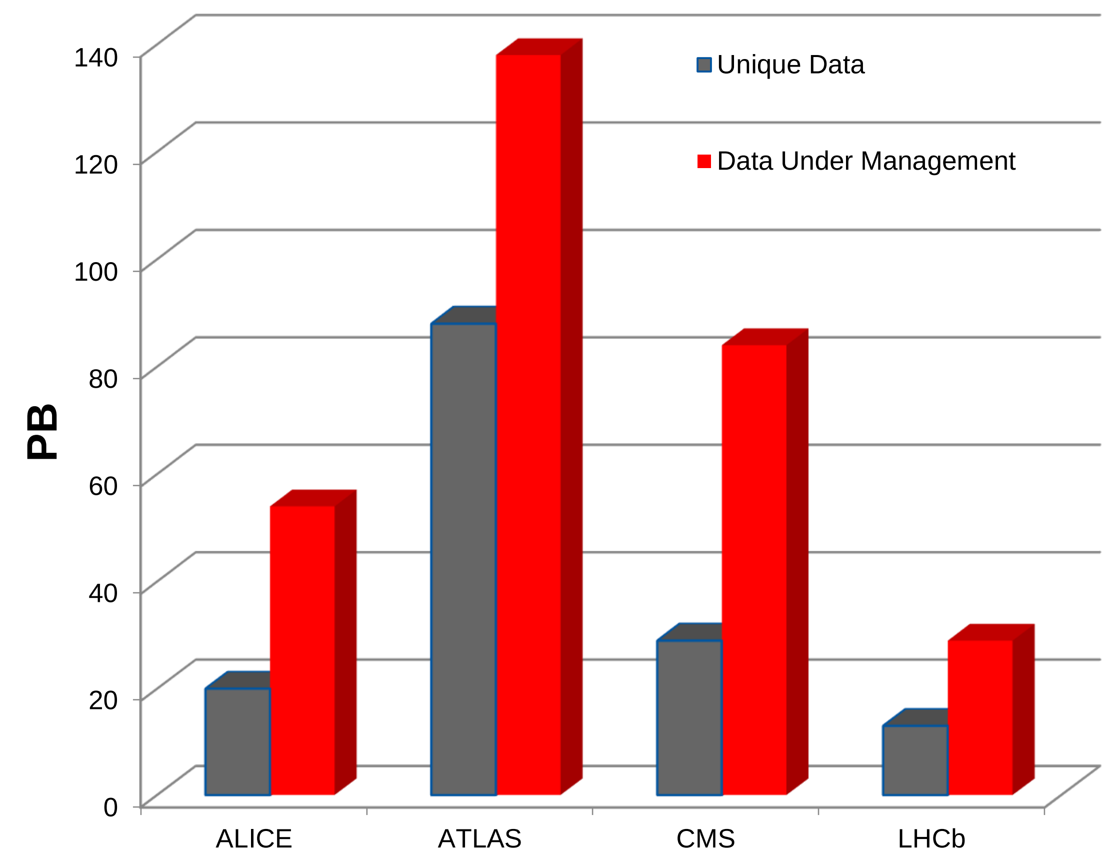
246 PB of disk

267 PB of tape

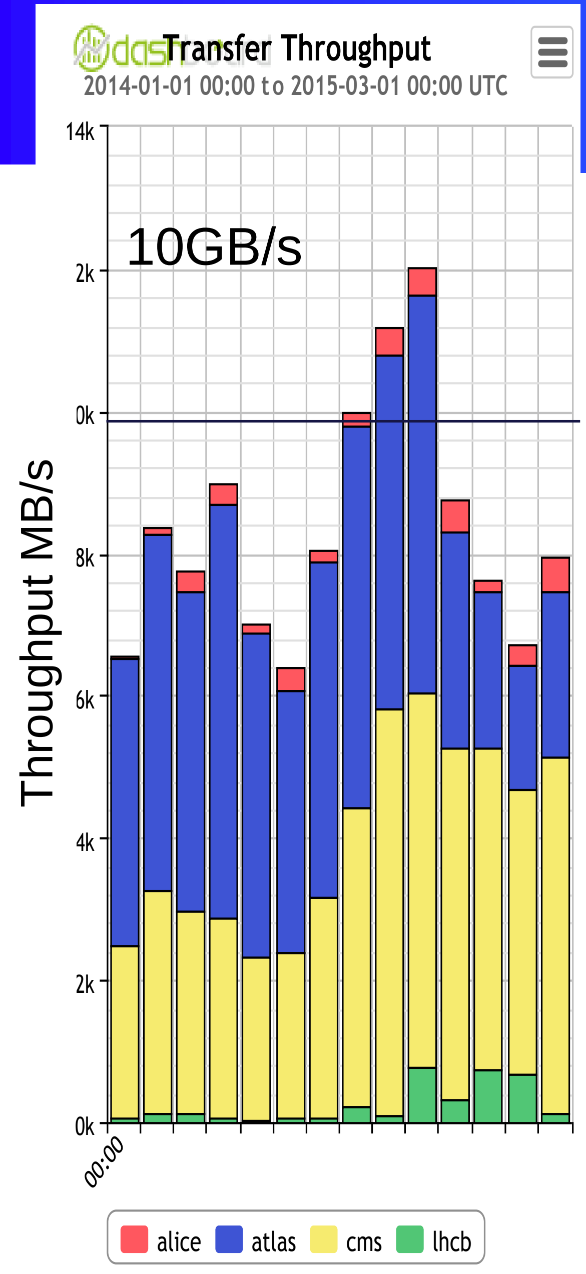


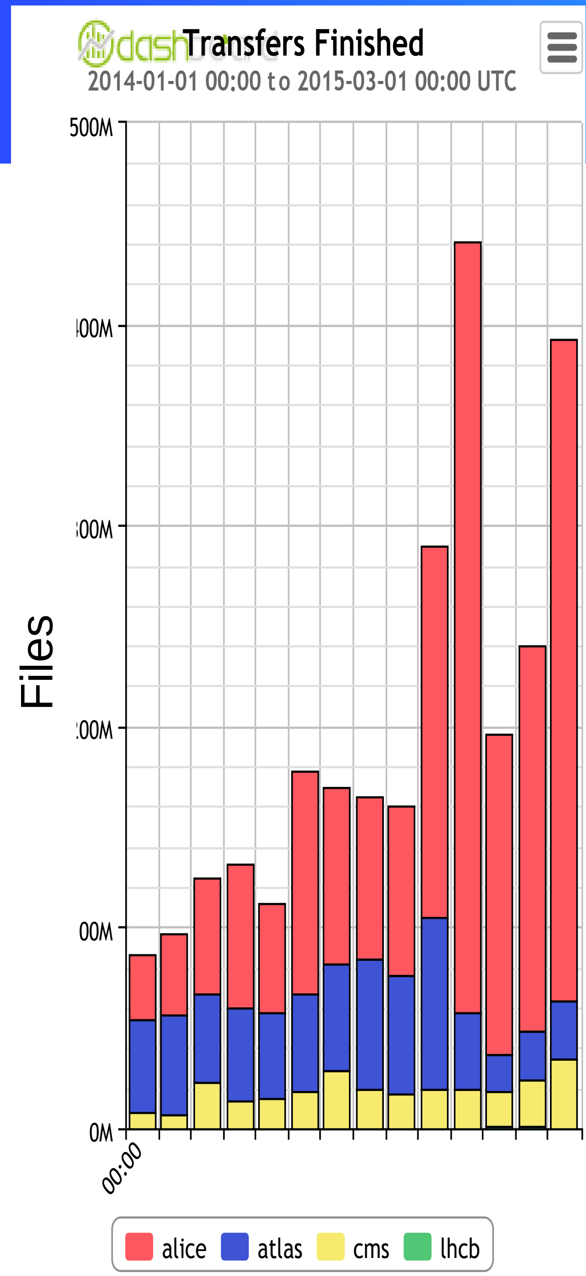
WLCG has 140PB of unique data and 280PB under management

- ➡ More than 1B files
- ➡ Average file size 0.2GB to 2.5GB

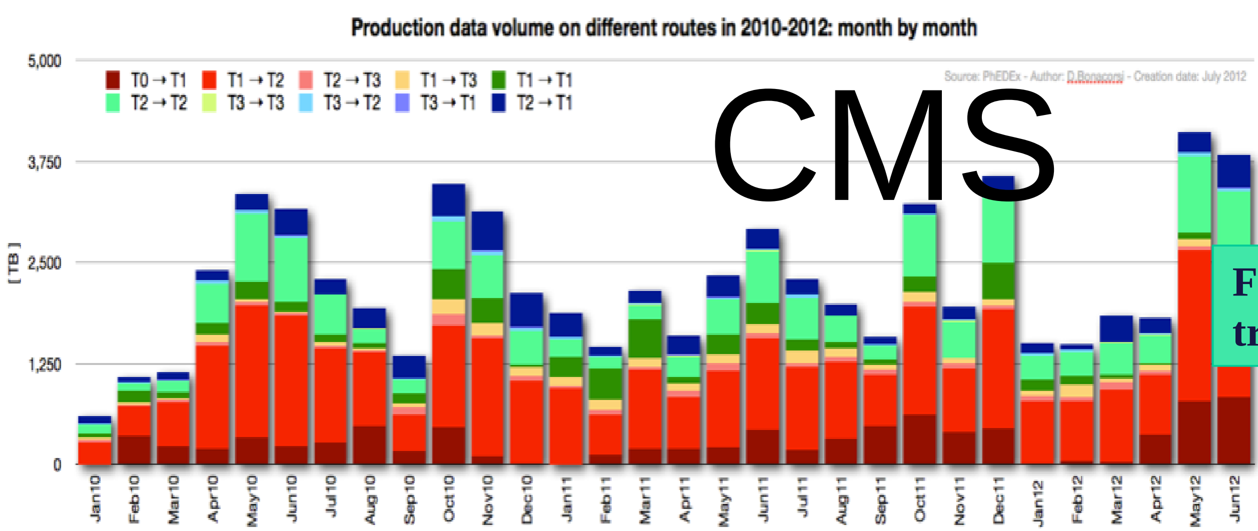


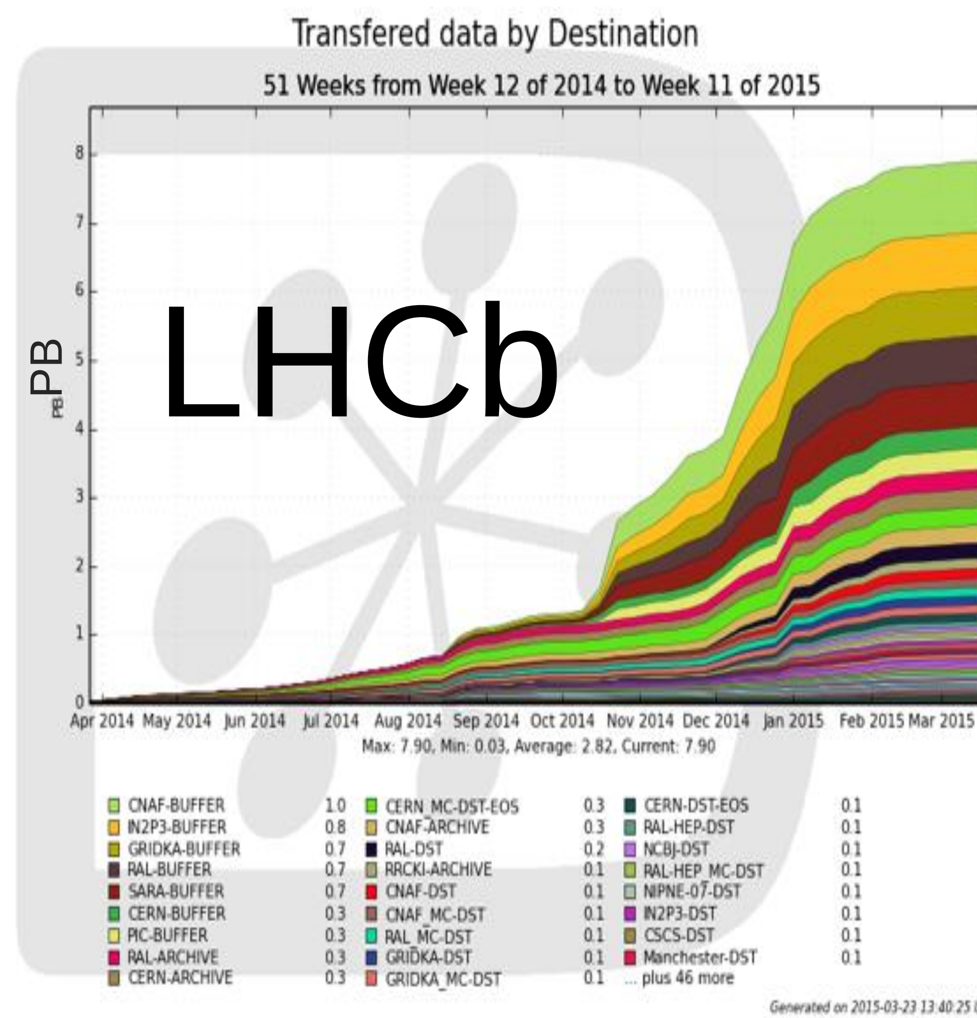
Scale of Movement





- Over all of LS1 the LHC experiments (mostly ATLAS and CMS) have been moving more than **0.5PB/day**
- In total, **1 EB** over the long shutdown

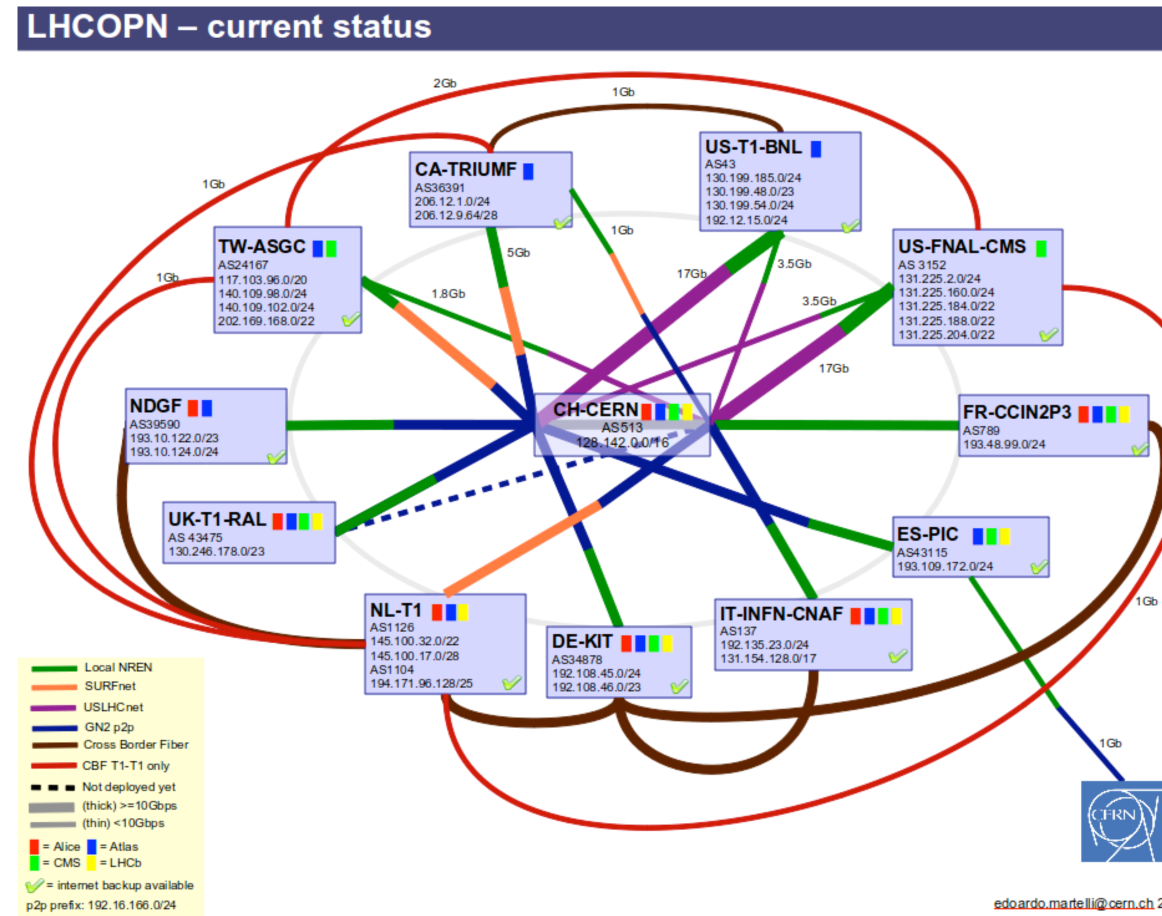


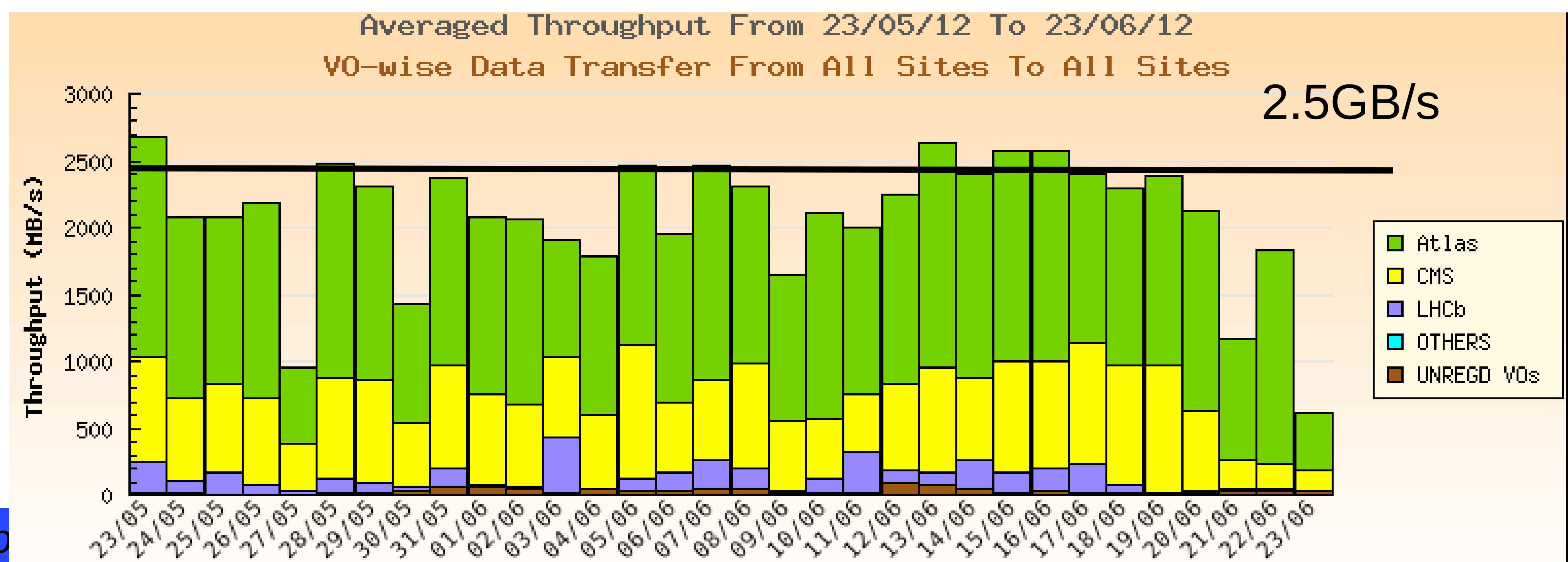


Networking

Wide Area networks allow us to move the data to remote sites for archiving and processing

- A dedicated network to for initial distribution
- Much shared use R&E networking to analysis centers

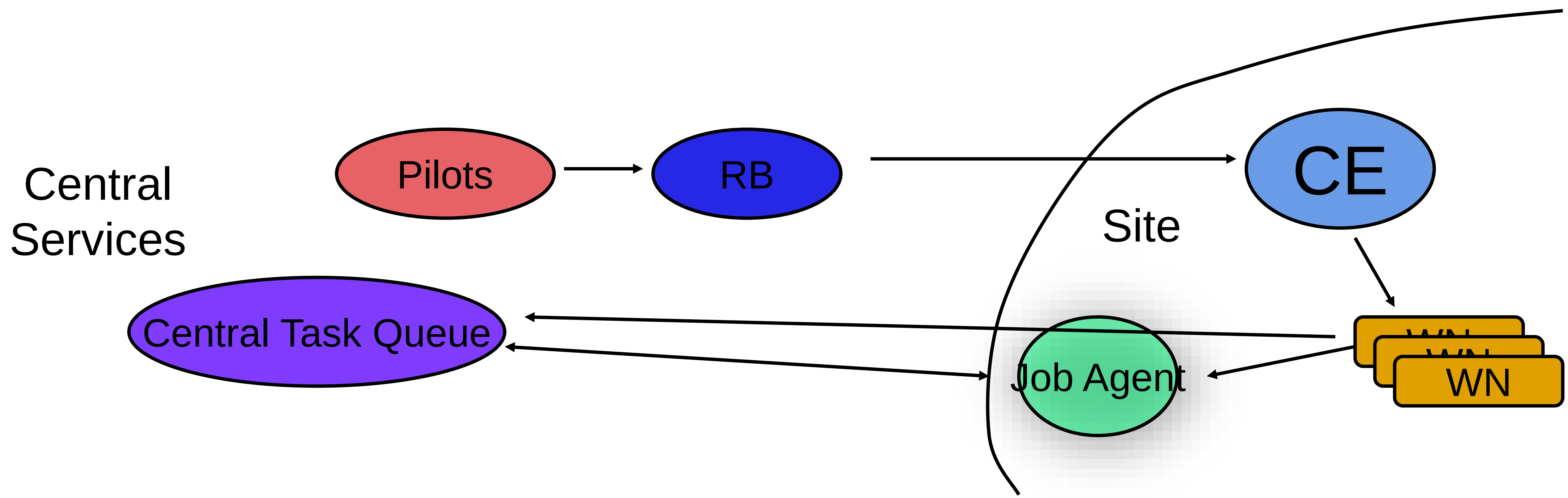




Submission Techniques

Both ALICE and LHCb have developed pull based job submission systems for both Production and Analysis

- ➡ Eventually all experiements did



Processing Data

Most of what we do is process files of groups of files in embarrassing parallel high throughput computing (HTC)

With data it's important to process every file

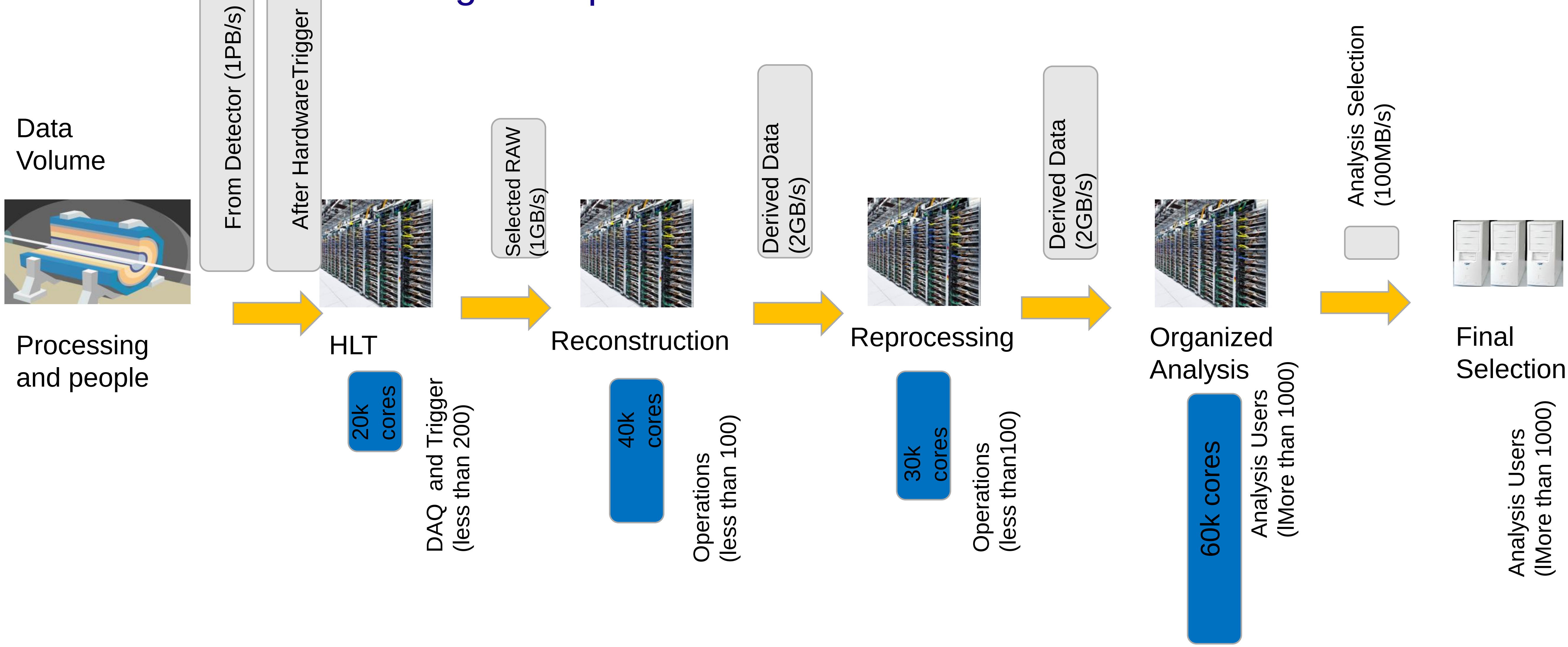
- Important not to have systematic failures in the processing system

All the experiments have some sort of a DB that keeps track of the pieces of split workflows

- Oracle, Couch. MySQL are all used

Data Path Through LHC

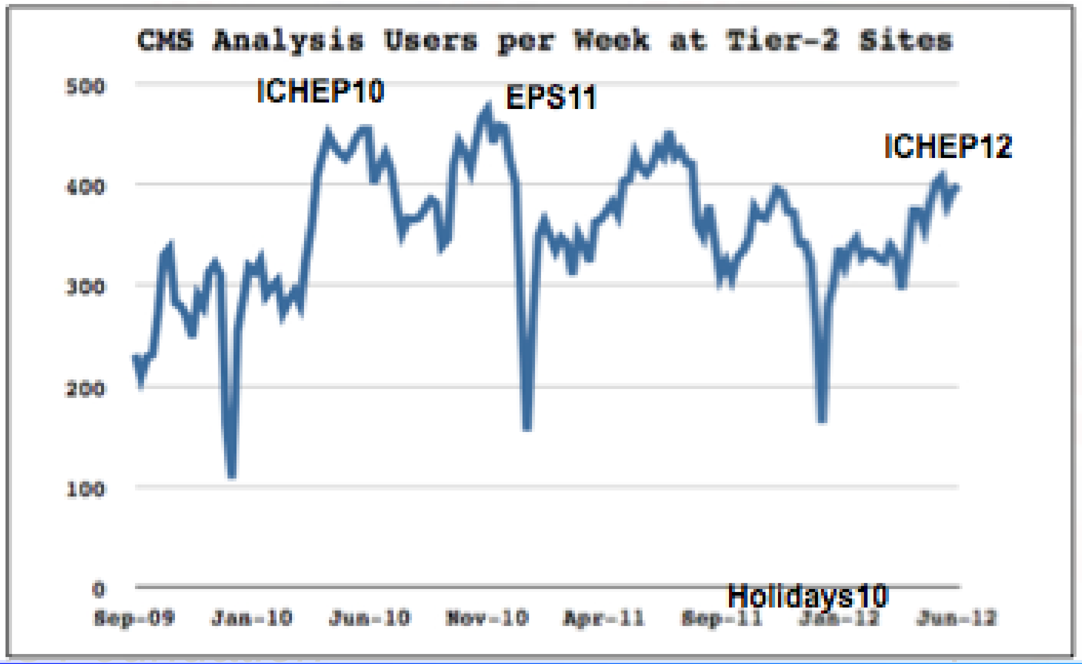
Data Path through LHC mostly servers to reduce the data to more manageable pieces

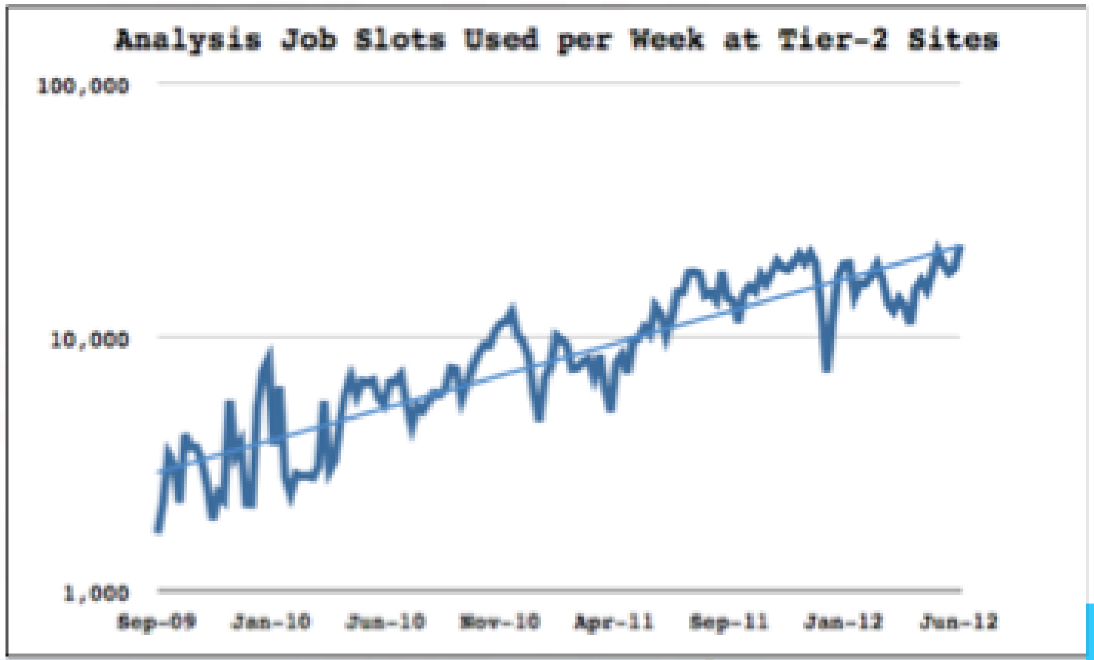


Analysis Centers

When the WLCG started there was a lot of concern about the viability of the Tier-2 Program

➡A university based grid of often small sites

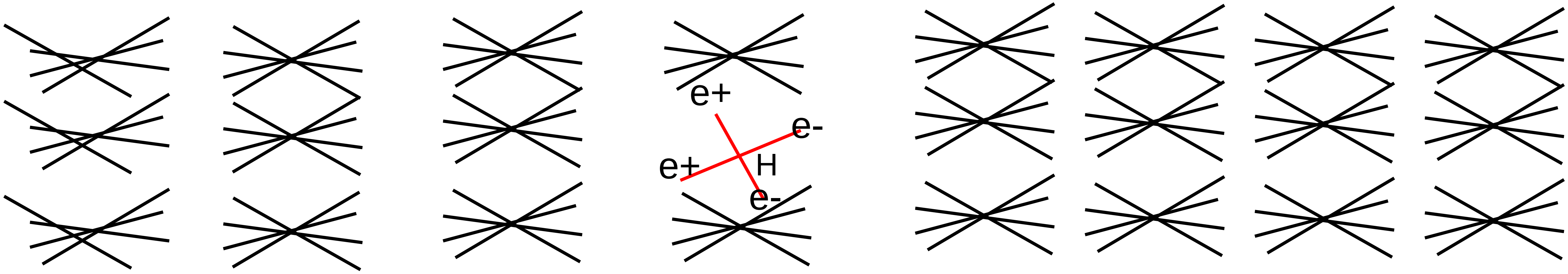




Simulation Challenges

➡ Beams contain many particles and beam collisions are frequent

- For every signal event at High Luminosity there are 35 minimum bias events from that crossing.   The calorimeters are sensitive to the preceding 10 and following 5 crossings
    - For every event we simulate we provide 100MB of minimum bias events



Scale of the final system

Progress in distributed computing and evolution of computing capacity

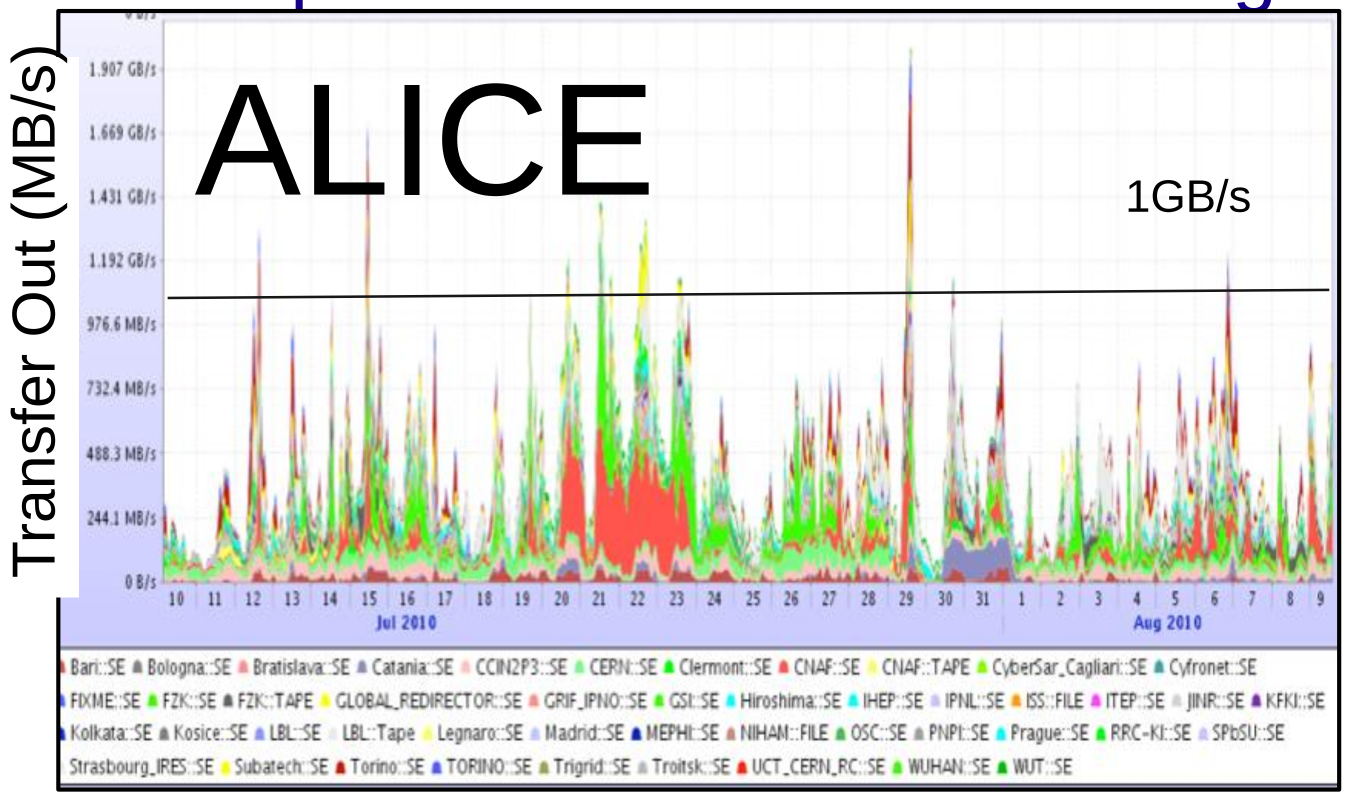
- ➡WLCG processes ~4M jobs on the grid per day
- ➡Disk and tape combined are now close to an Exabyte of storage

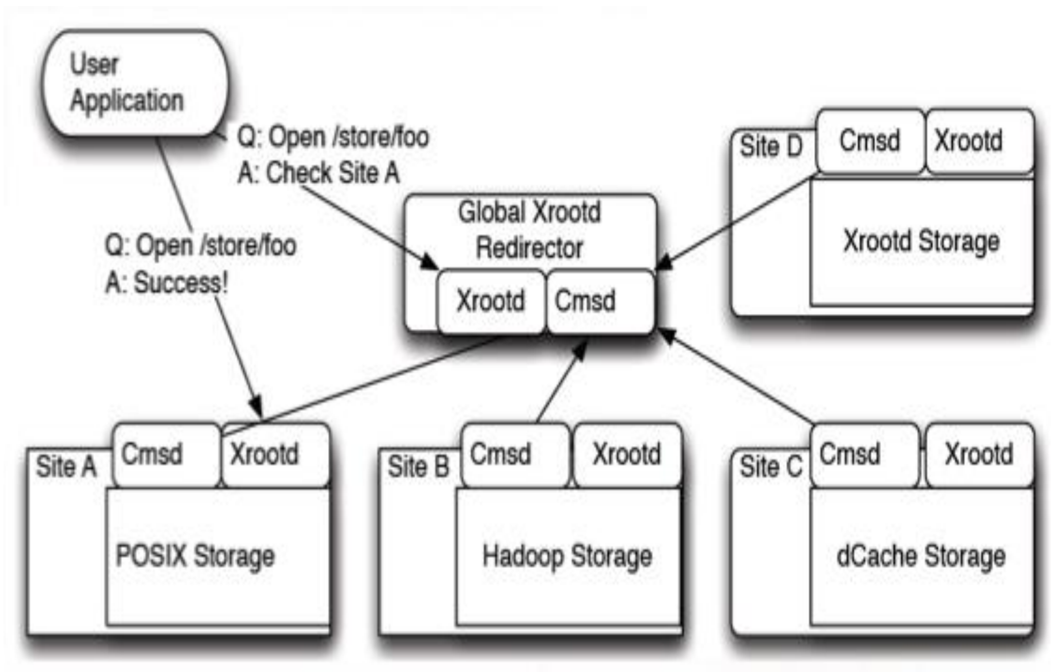
Essentially a leadership class super computer distributed over 5 continents

Introduction to Federation

From the beginning ALICE based their data management on Xrootd

- ➡ Other experiments have subsequently been deploying data federations and similar techniques
  - ALICE and LHCb use experiments catalogs to identify the file location and mainly open files locally
  - ATLAS and CMS have data federations fully based on Xrootd and separate from the data management and transfer systems





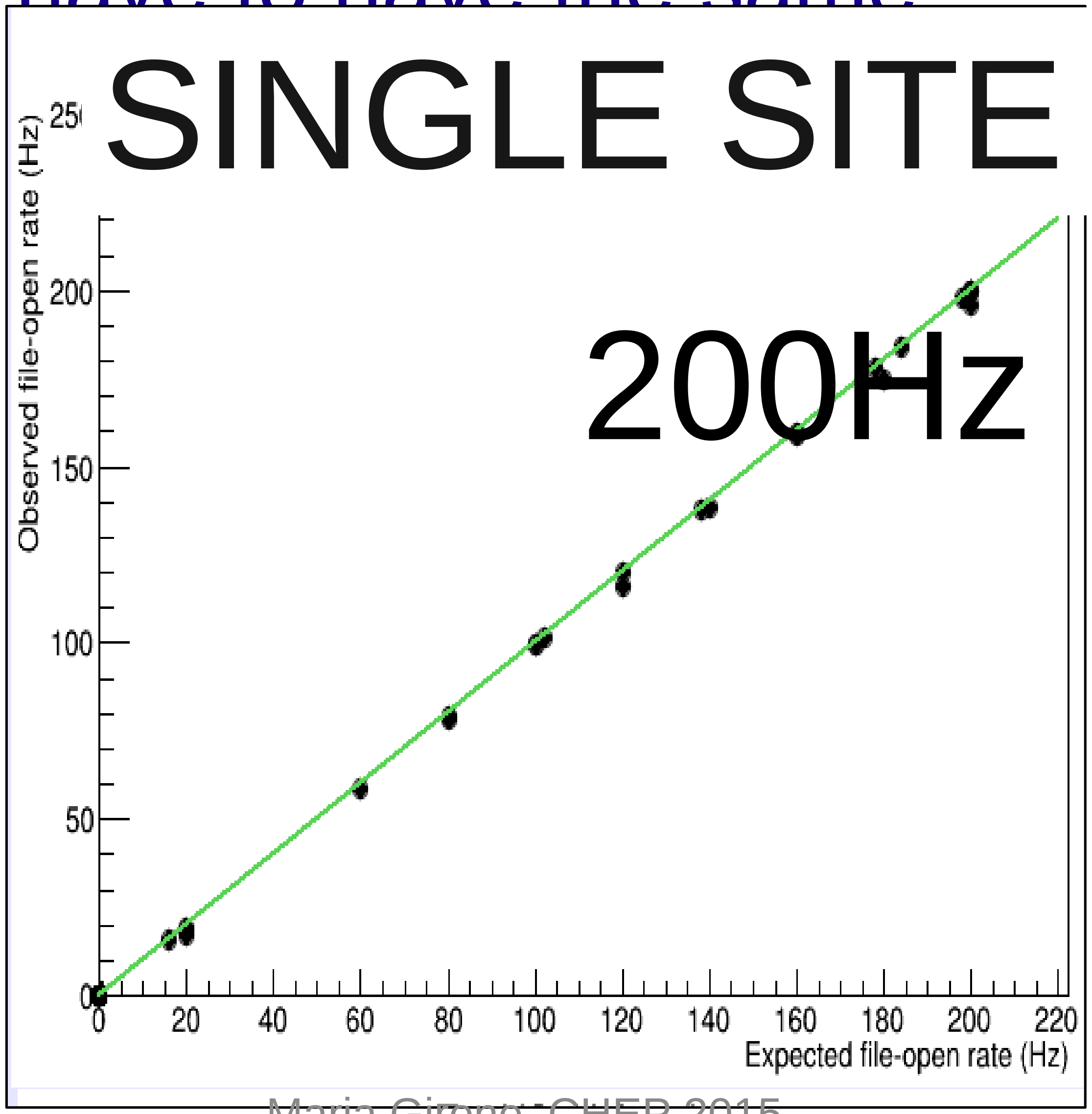
Xrootd as a Distributed File System

The way Xrootd maintains a file system is simple and clever

All servers have the same name space, though they don't have the same contents

Site 1
/data/items/files/file1
Site 2
Site 3
/data/items/files/file1
file3
file4

SINGLE SITE SCALING T



Maria Girone, CHEP 2015

Files can be opened at a rate of hundreds of Hz

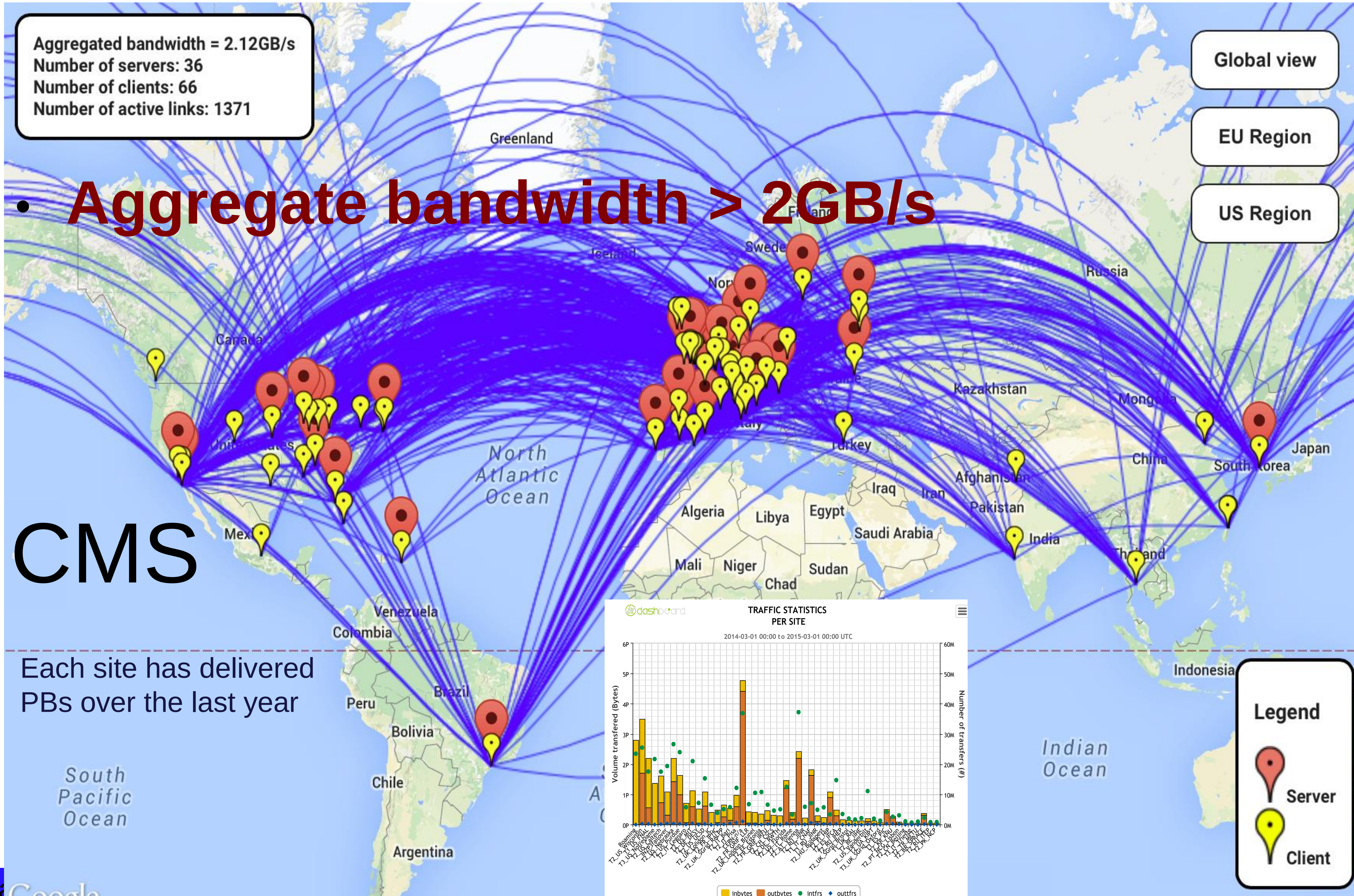
Successes in Connectivity

xxx

- **Aggregate bandwidth > 2GB/s**

CMS

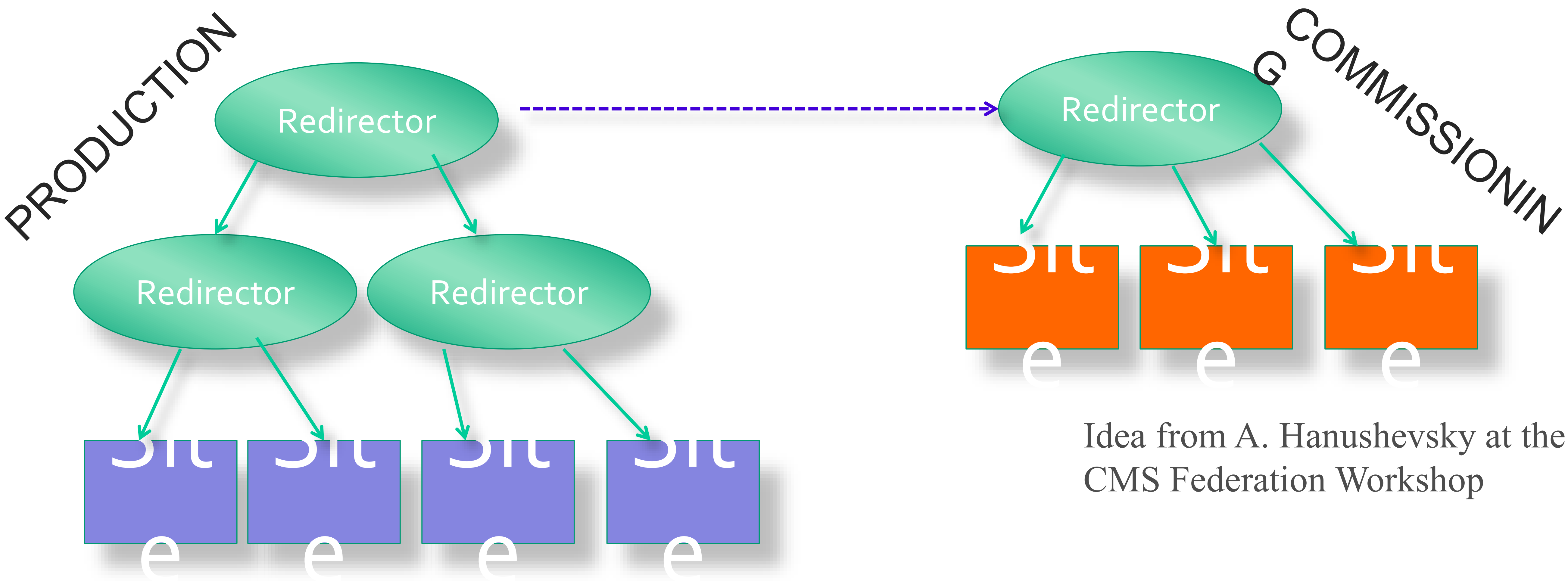
Each site has delivered PBs over the last year



Transitional Federation

The use of data federation adds enormous functionality but also complexity

- ➡ Now there is another site that has to successfully perform an action and not all sites are equal



Idea from A. Hanushevsky at the CMS Federation Workshop

Maria Girone, CHEP 2015

To recap

On the positive side:

- We now have a system where we can utilize a set of globally distributed computing centers
- We have reached a very high scale
- We can distribute a software environment and conditions
- We can move data, discover data, and for a portion of the access even serve over the WAN

On the negative:

- A lot has to go right for work to get done
  - There are a lot of expectations of the resources when you arrive on a site
    - Operating systems, configurations, and services
    - Limits the resources that can be used
    - Makes the resources more difficult to share
    - Places a reasonably heavy load on site administrators
    - The system remains mostly homogenous
    - OS, hardware profiles, interfaces all need to stay in lock stepMore difficult to share resources with other communities
- We have coupled the processing and the storage
  - Systems with very different time scales are tied together

Thursday we talk about Clouds

Clouds vs Grids

Grids offer primarily standard services with agreed protocols

- ➡Designed to be as generic as possible, but execute a particular task

Clouds offer the ability to build custom services and functions

- ➡More flexible, but also more work