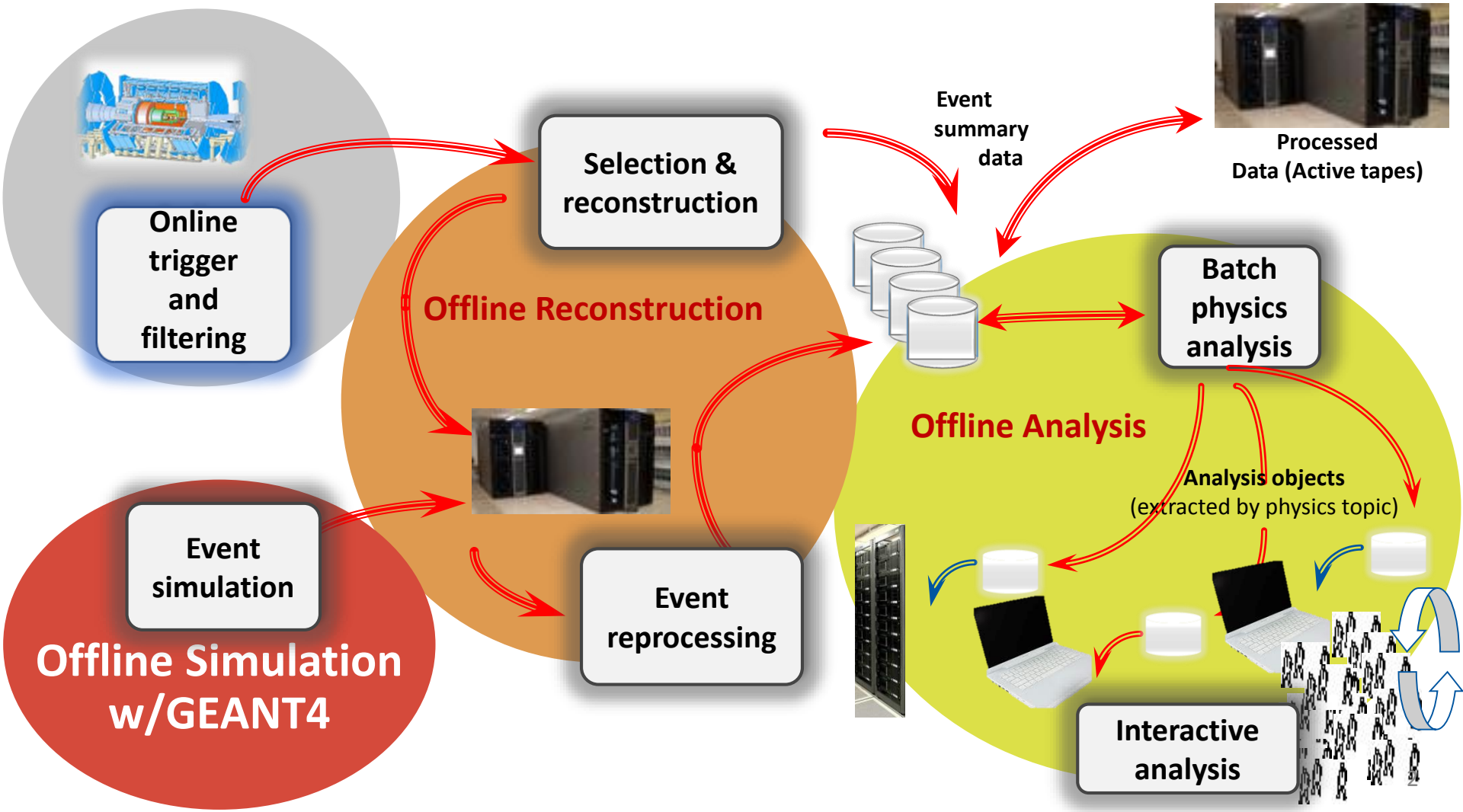
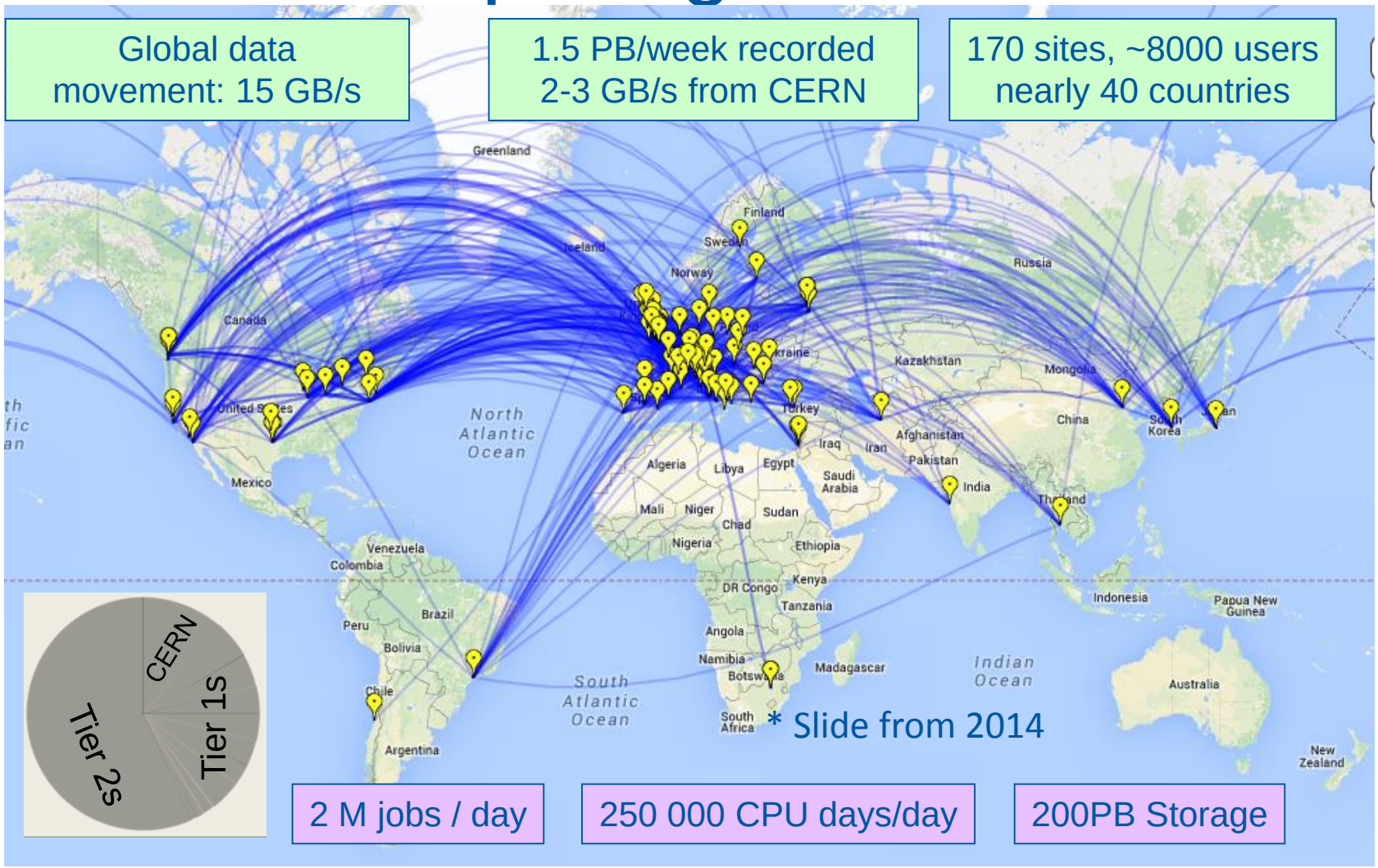
HTCondor Integration

Laurence Field
CERN IT

Computing Workflow



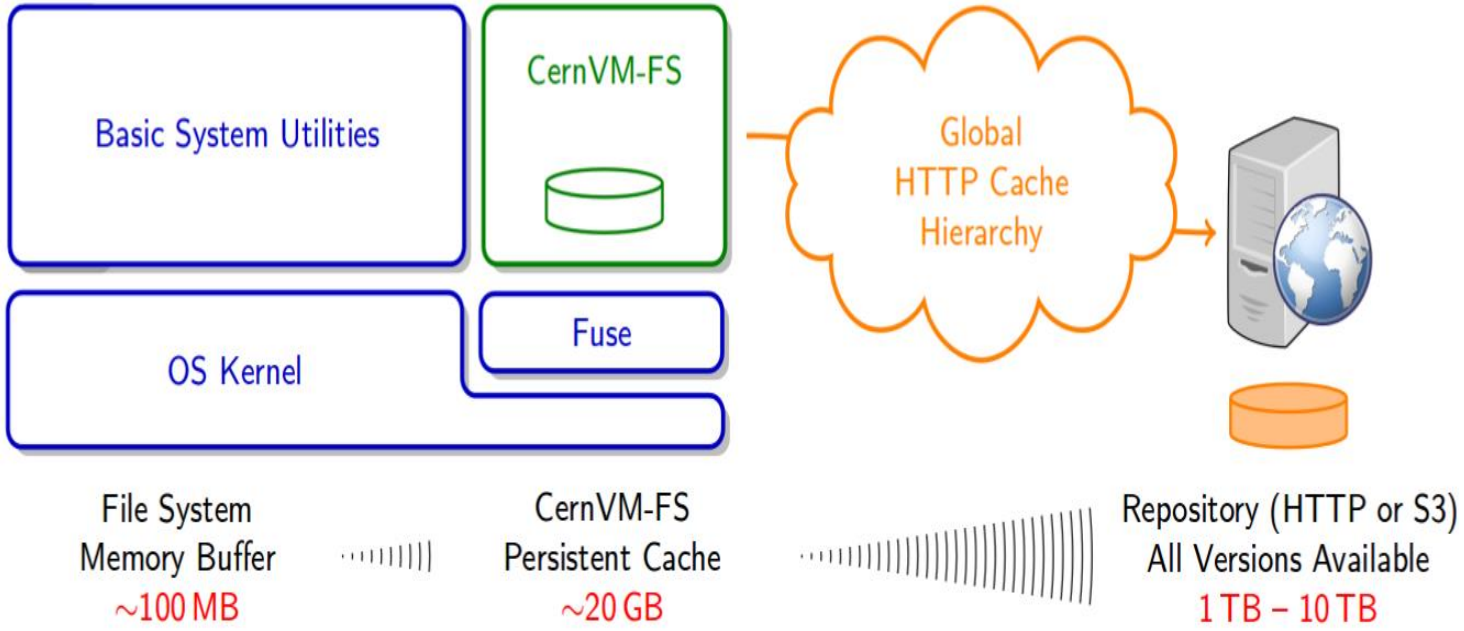
WLCG Computing Infrastructure



The Challenges

- Running HEP Software on Windows
  - ~85% of the volunteers
- Seamlessly integrating with the existing workflow
  - Experiment specific infrastructure and services
  - Trusted and untrusted environments
  - Low prioritization
    - Data taking and analysis is always higher
    - Focus on resources at stake rather than unfulfilled potential
  - Reduce the overall operations cost
    - Build upon existing tooling and support structures

CernVM and CVMFS

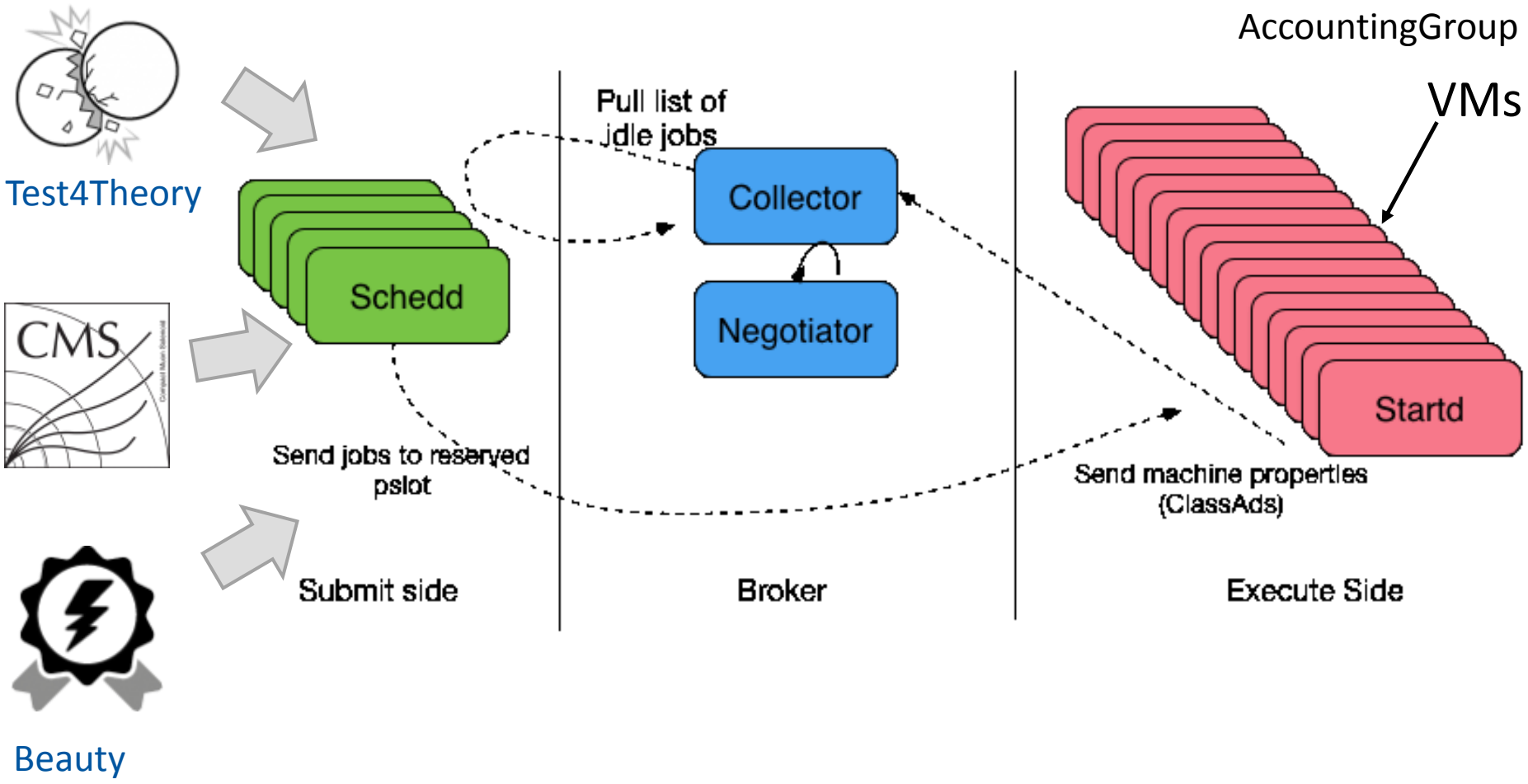


- Small image size but need to “bake” the images to reduce unnecessary downloads

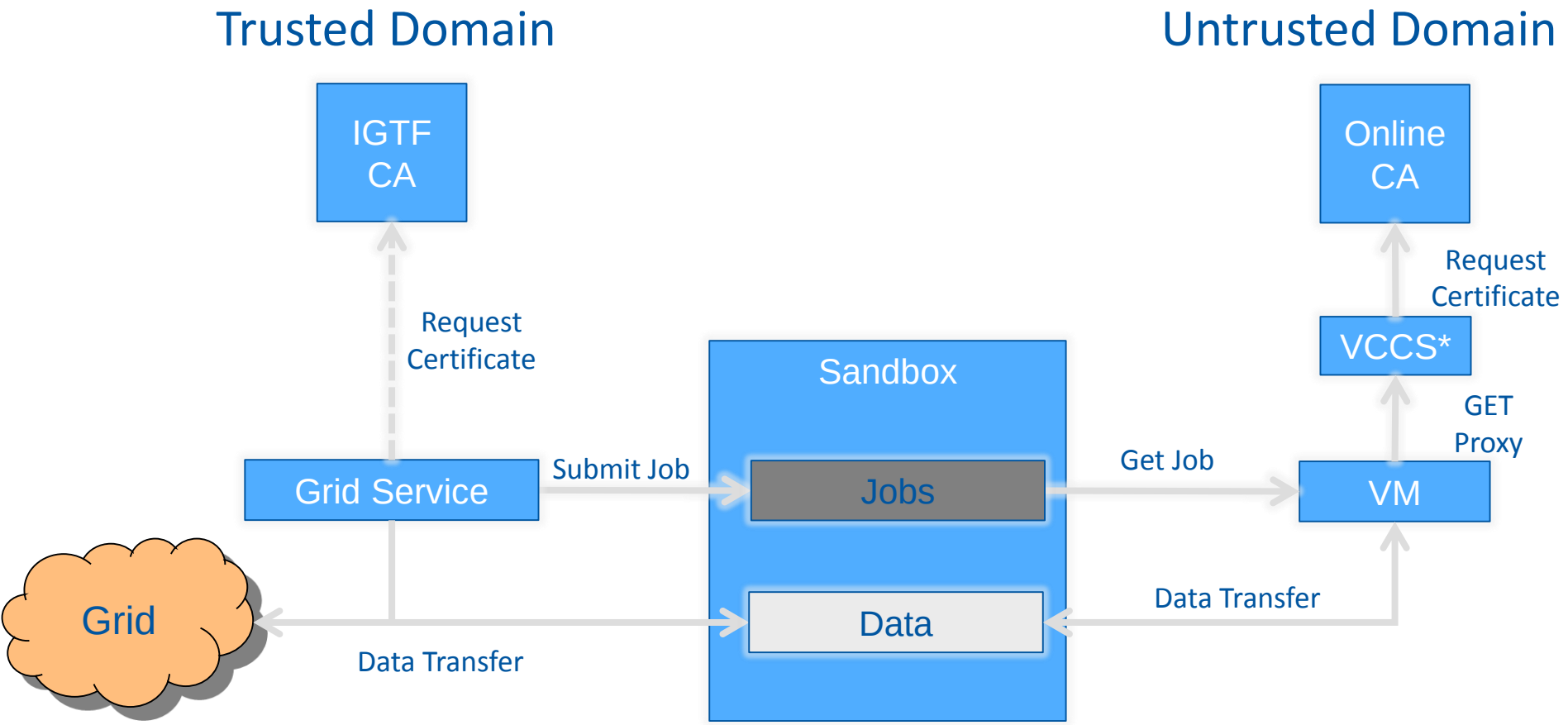
HTCondor

- Open Source batch system from the University of Wisconsin
  - Focus on High Throughput Computing
- Symmetric matching of job requests to resources
  - Using ClassAds of job requirements and machine resources
- Long history in HEP and elsewhere
  - Used extensively in OSG
  - Also for the CMS global pool (160K+ cores)
  - CERN currently migrating from LSF to HTCondor
    - Build upon existing expertise and operational support infrastructure
- Use HTCondor with BOINC to implement the Vacuum model
  - Overlay a batch system upon autonomous elastic resources

HTCondor Model

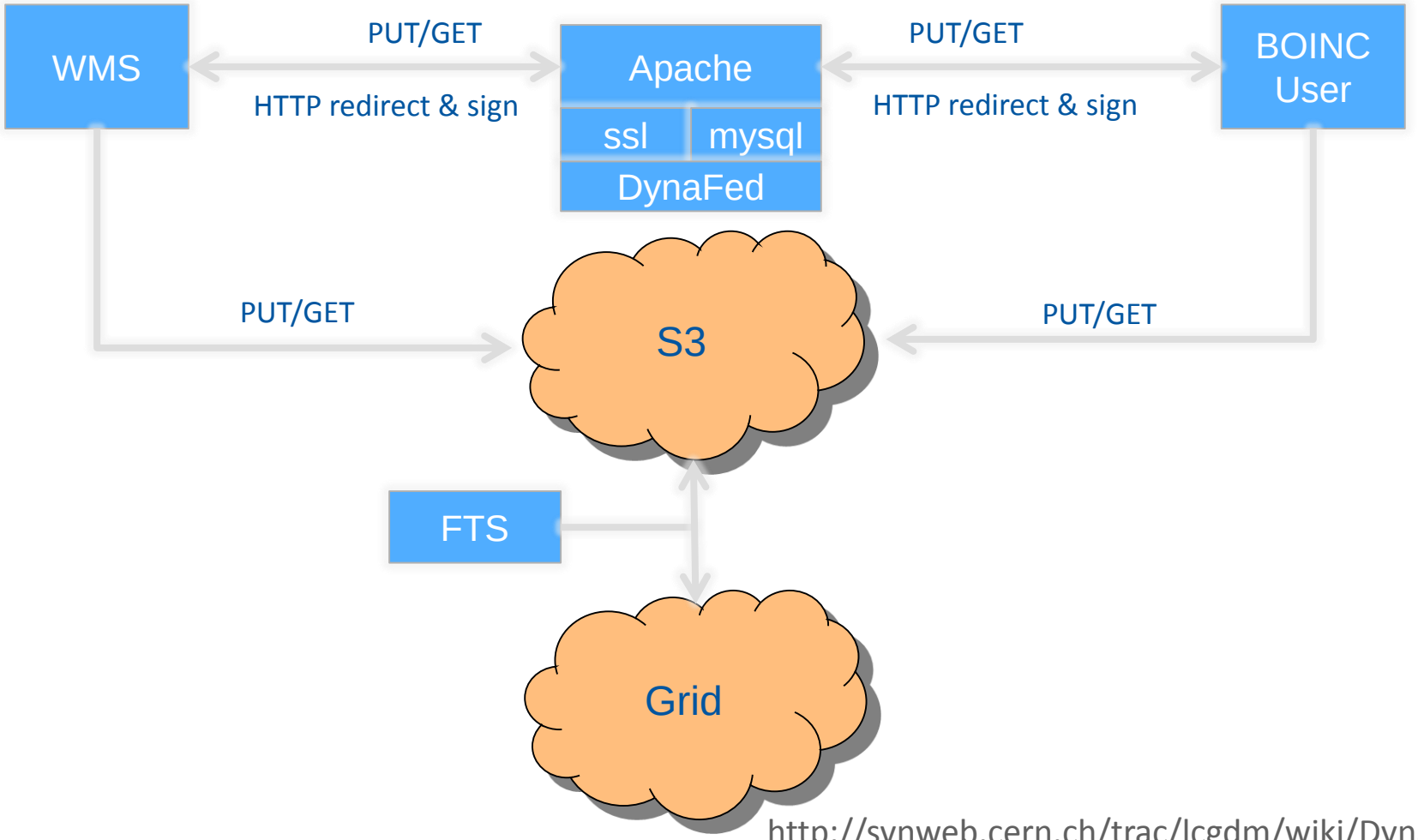


Sandboxing and Authentication



*Volunteer Computing Credential Service

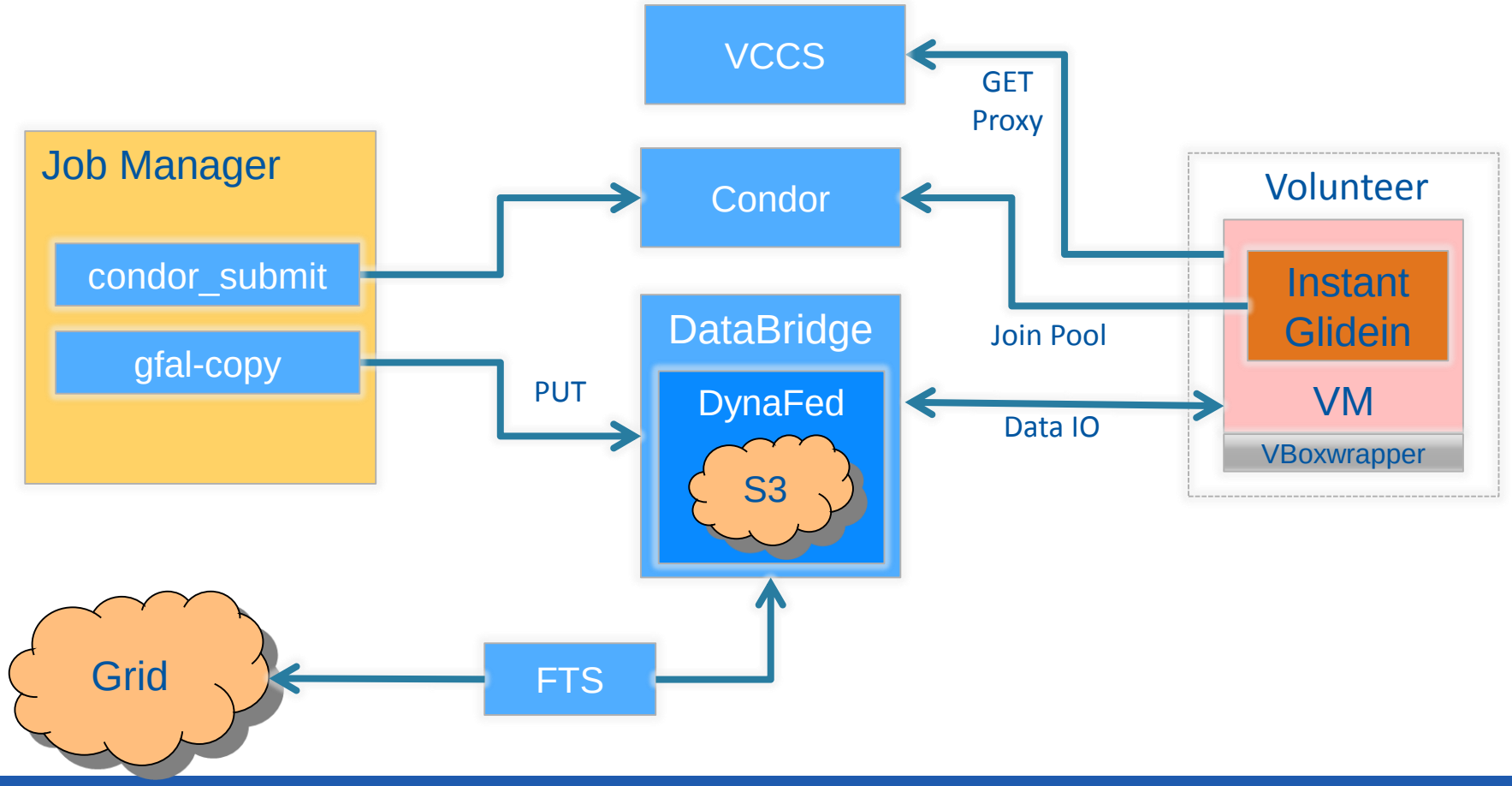
The DataBridge



http://svnweb.cern.ch/trac/lcgdm/wiki/Dynafeds

The Solution

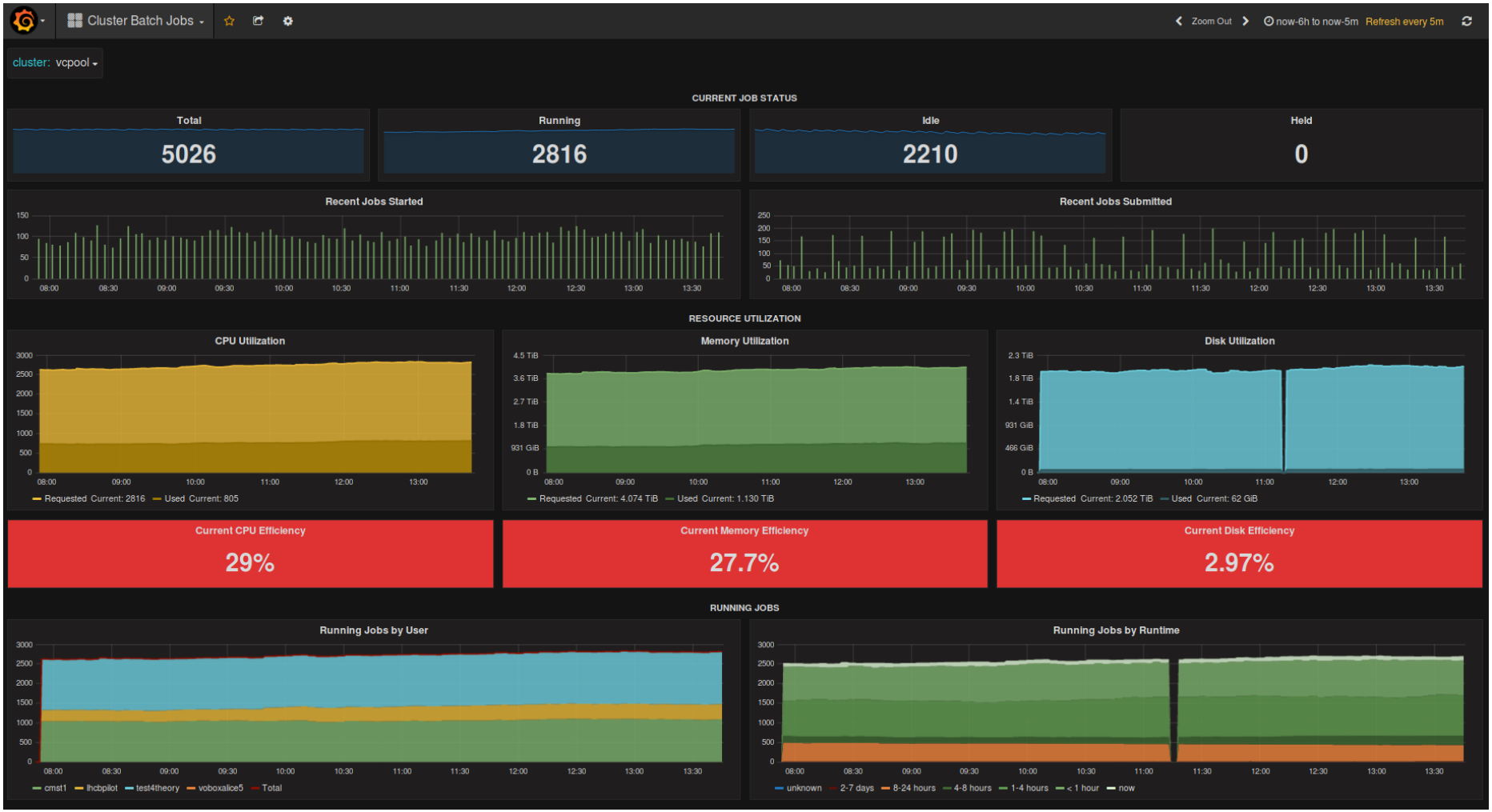
Application Server | Common Infrastructure | Volunteer's machine



Direct Submission For Sixtrack

- Simplify for the project scientists
  - HTCondor interface for both batch and BOINC
  - Reduce BOINC specific knowledge
  - DAGs to seamless use both together
- All jobs going through the same system
  - Common monitoring and accounting
- Simplify for the service managers
  - Same skills set required

FifeMon


Cluster Batch Jobs
Zoom Out
now-6h to now-5m
Refresh every 5m
cluster: vcpool
CURRENT JOB STATUS
Total
5026
Running
2816
Idle
2210
Held
0
Recent Jobs Started
Recent Jobs Submitted
RESOURCE UTILIZATION
CPU Utilization
Requested Current: 2816
Used Current: 805
Memory Utilization
Requested Current: 4.074 TiB
Used Current: 1.130 TiB
Disk Utilization
Requested Current: 2.052 TiB
Used Current: 62 GiB
Current CPU Efficiency
29%
Current Memory Efficiency
27.7%
Current Disk Efficiency
2.97%
RUNNING JOBS
Running Jobs by User
cmst1
lhcbpilot
test4theory
voboxalice5
Total
Running Jobs by Runtime
unknown
2-7 days
8-24 hours
4-8 hours
1-4 hours
< 1 hour
now

What Do We Care About?

- BOINC submission RPC
  - Surprised this is not the default approach
- HTCondor integration
  - Both approaches
  - RPC and GAHP scalability
- A platform rather than a project
  - Not all scientists are computer scientists
    - Not all computer scientists are computer engineers
- boinc_gahp
- vboxwrapper

Summary

- HTCondor plays a key role in HEP
  - Seamlessly integration
  - Common skill set
- Provided a few complementary solutions
  - CernVM and CVMFS
    - The original vboxwrapper
  - VCCS
    - Create an x509 credential from a BOINC credential
  - DataBridge
    - Authentication overlay for CEPH
- Using two different approaches
  - An elastic batch system
  - Job routing