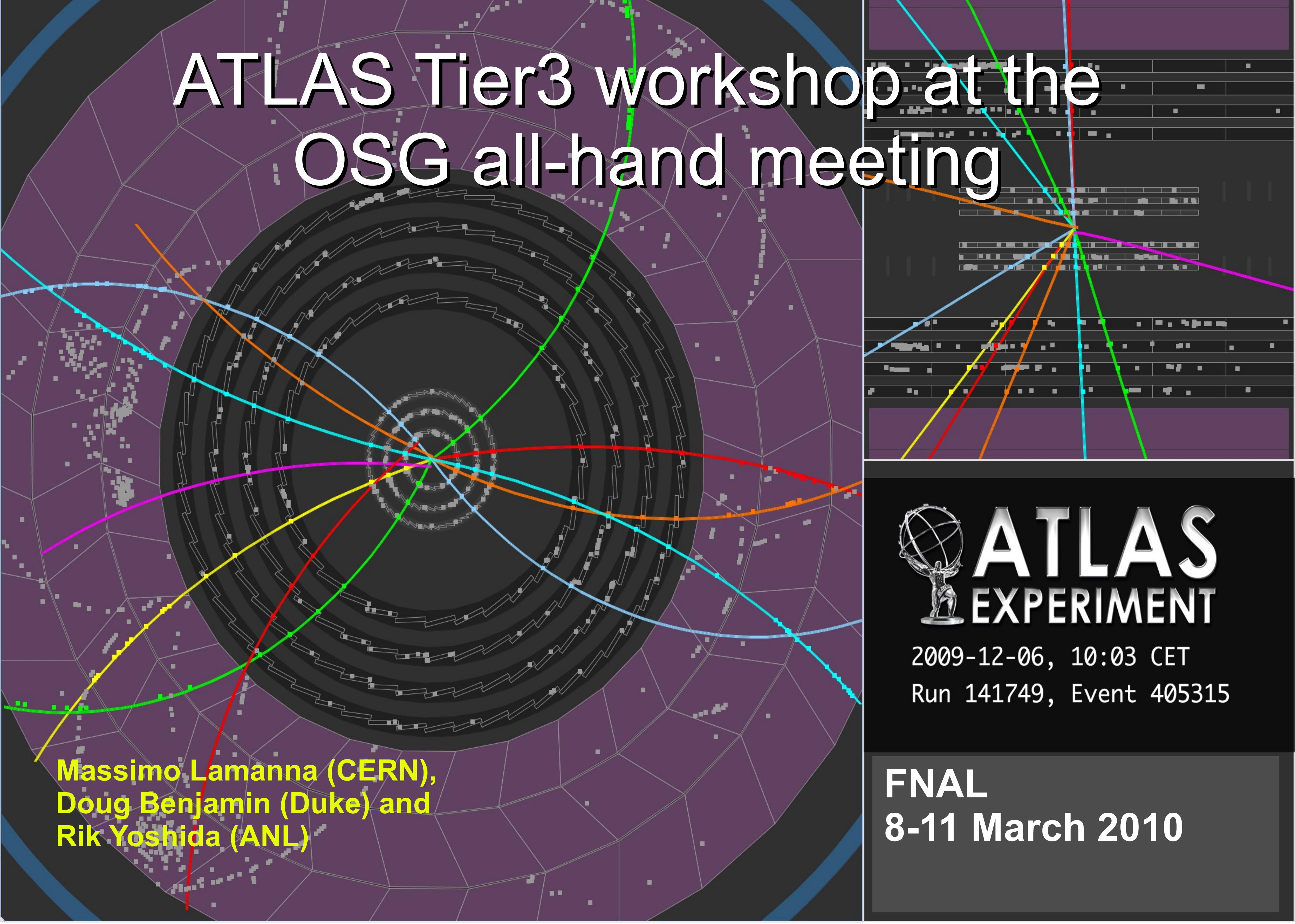
ATLAS Tier3 workshop at the OSG all-hand meeting



Massimo Lamanna (CERN),
Doug Benjamin (Duke) and
Rik Yoshida (ANL)

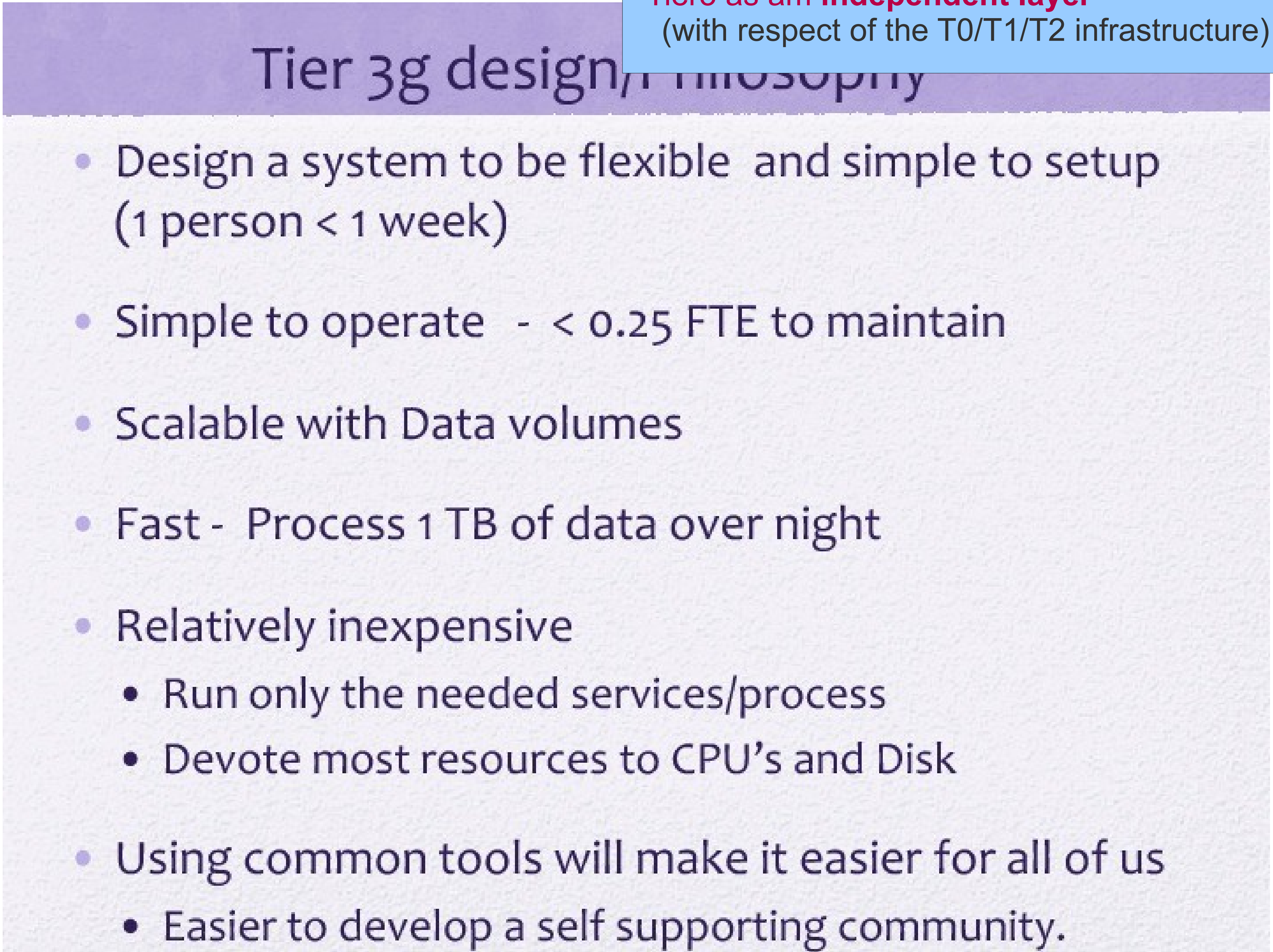
FNAL
8-11 March 2010

http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html

ATLAS Tier3

- ATLAS-wide model
  - Tier3: Analysis facility based on “non pledged” resources
  - US initiative restarted discussions within ATLAS (25-26/01/2010 at CERN)
    - http://indico.cern.ch/conferenceDisplay.py?confId=77057
    - 13 presentations only on plans/experience
      - more than 1 per cloud, here the granularity is more "country"
      - Typically T3 is a single experiment facility
      - Notable exceptions: **DESY** and **Lyon** analysis facilities (NAF and LAF)
- Another layer continuing the hierarchy after Tier0, Tier1s, Tier2s ?
    - Probably truly misleading...
    - Qualitative difference here:
      - **Final analysis vs simulation and reconstruction**
      - **Local control vs ATLAS central control**
      - **Operation load more on local resources (i.e. people) than on the central team (i.e. other people)**

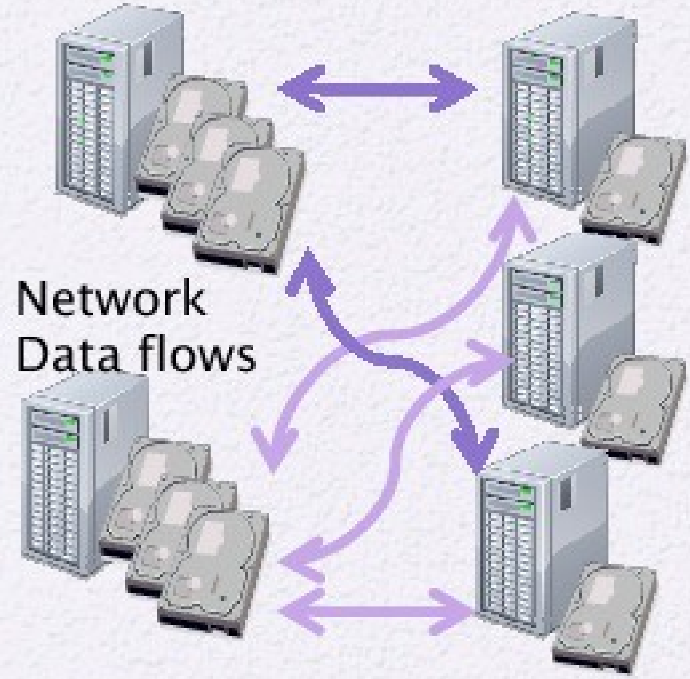
- Maintaining **grid services** vs using **grid clients**
- Tier3 as am **independent layer**
  (with respect of the T0/T1/T2 infrastructure)



Tier 3g design/Philosophy

- Design a system to be flexible and simple to setup (1 person < 1 week)
- Simple to operate - < 0.25 FTE to maintain
- Scalable with Data volumes
- Fast - Process 1 TB of data over night
- Relatively inexpensive
  - Run only the needed services/process
  - Devote most resources to CPU's and Disk
- Using common tools will make it easier for all of us
  - Easier to develop a self supporting community.

Tier 3g – Data storage options



Storage on worker nodes

XRootD can be used to manage either type of storage

ATLAS Tier3 Working groups

- **DDM-Tier3 link**
  - Simone Campana (CERN). *Presentation by Hironori Ito (BNL)*
- **Distributed storage (Lustre/Xrootd/GPFS)**
  - Rob Gardner (Chicago) and Santiago Gonzalez de la Hoz (Valencia)
- **Software / Conditions data Working Group**
  - Alessandro de Salvo (INFN Roma) and Asoka da Silva (TRIUM
- **PROOF Working Group**
  - Wolfgang Ehrenfeld (DESY) and Neng Xu (Wiscosin)
- **Tier 3 Support**
  - Dan van der Ster (CERN)
- **Virtualization working group**
  - Yushu Yao (LBL)

Status report
In this workshop

3-month time scale

Chaired by ATLAS persons

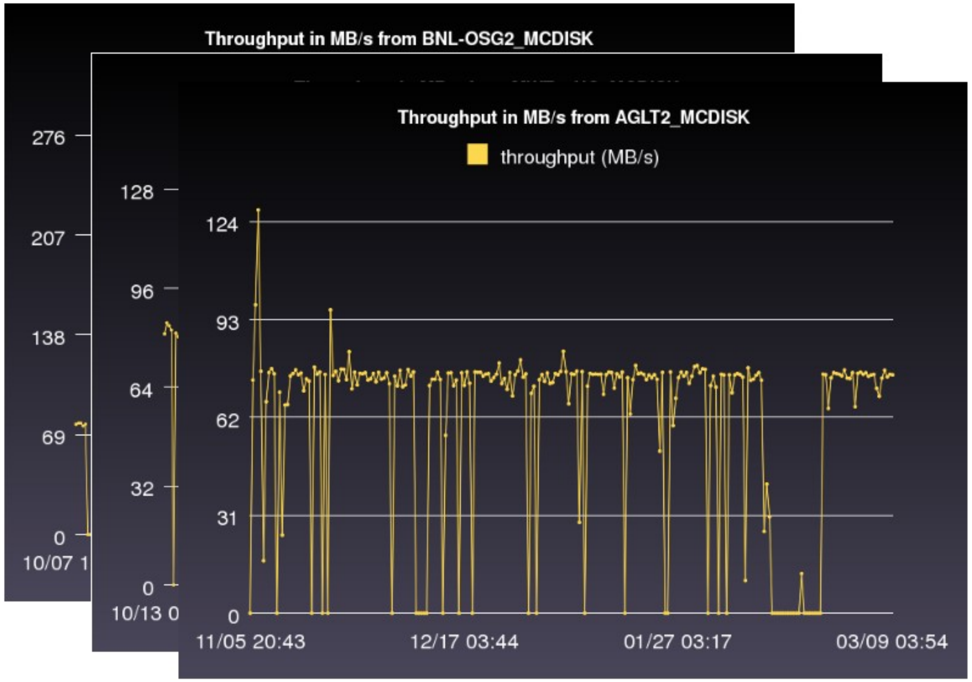
Open to experts (also from outside the collaboration)

Please note!

DDM-Tier3 Link

- How to populate your Tier3 with chosen datasets?
- Range of solutions exists:
  - dq2get (pure client)
  - hybrid solutions (gridFTP and FTS)
  - full fledged DDM subscription (centralised and asynchronous)

Sample Throughput Test results at T3



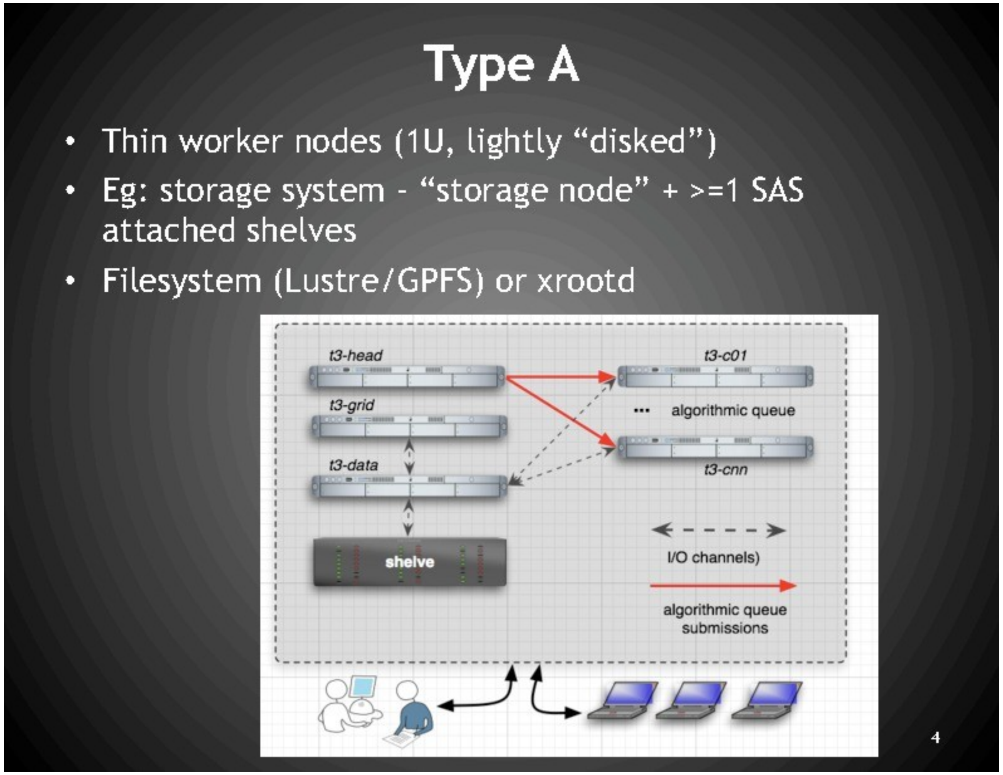
Hiro Ito (BNL)

Data access

- Inventory usage and best-practices for xrootd/Lustre/GPFS
  - Existing “polarisation”:
    - More xroot sites on the US side, Lustre (and GPFS) elsewhere
      - Notable exceptions exist
    - Similar use case (store and access data using filesystem(-like) namespace – use local protocols)
- Closely coupled with HW configuration (purchase guidelines)

Hardware models for Tier3

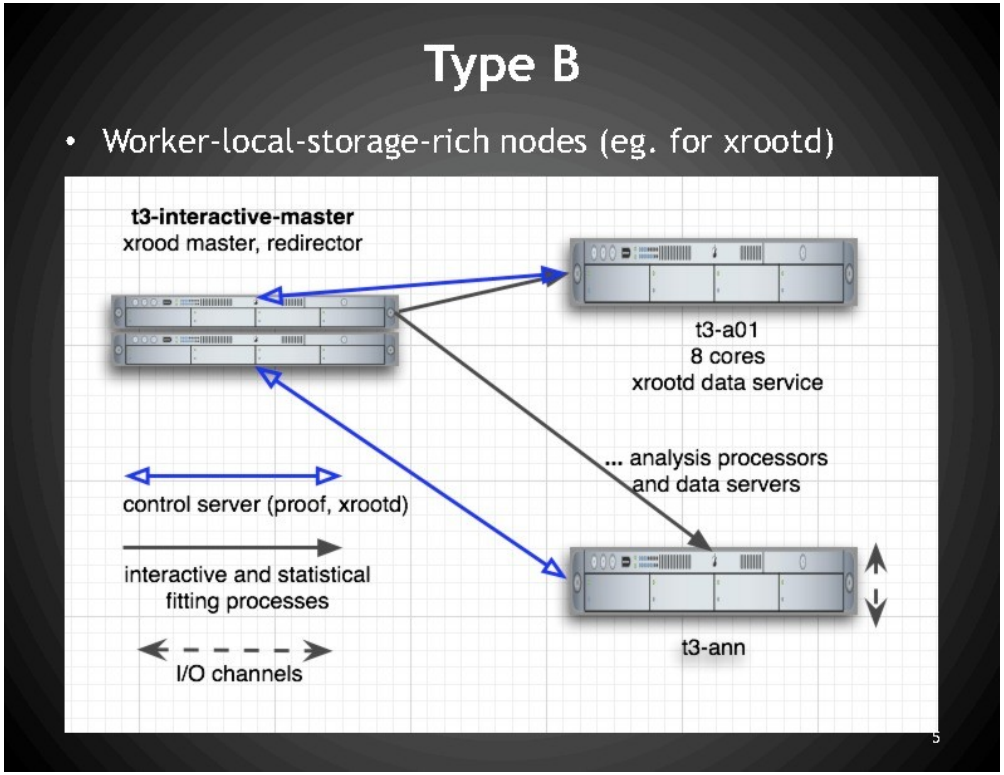
“xrootd” farm



Type A

- Thin worker nodes (1U, lightly “disked”)
- Eg: storage system - “storage node” + >=1 SAS attached shelves
- Filesystem (Lustre/GPFS) or xrootd

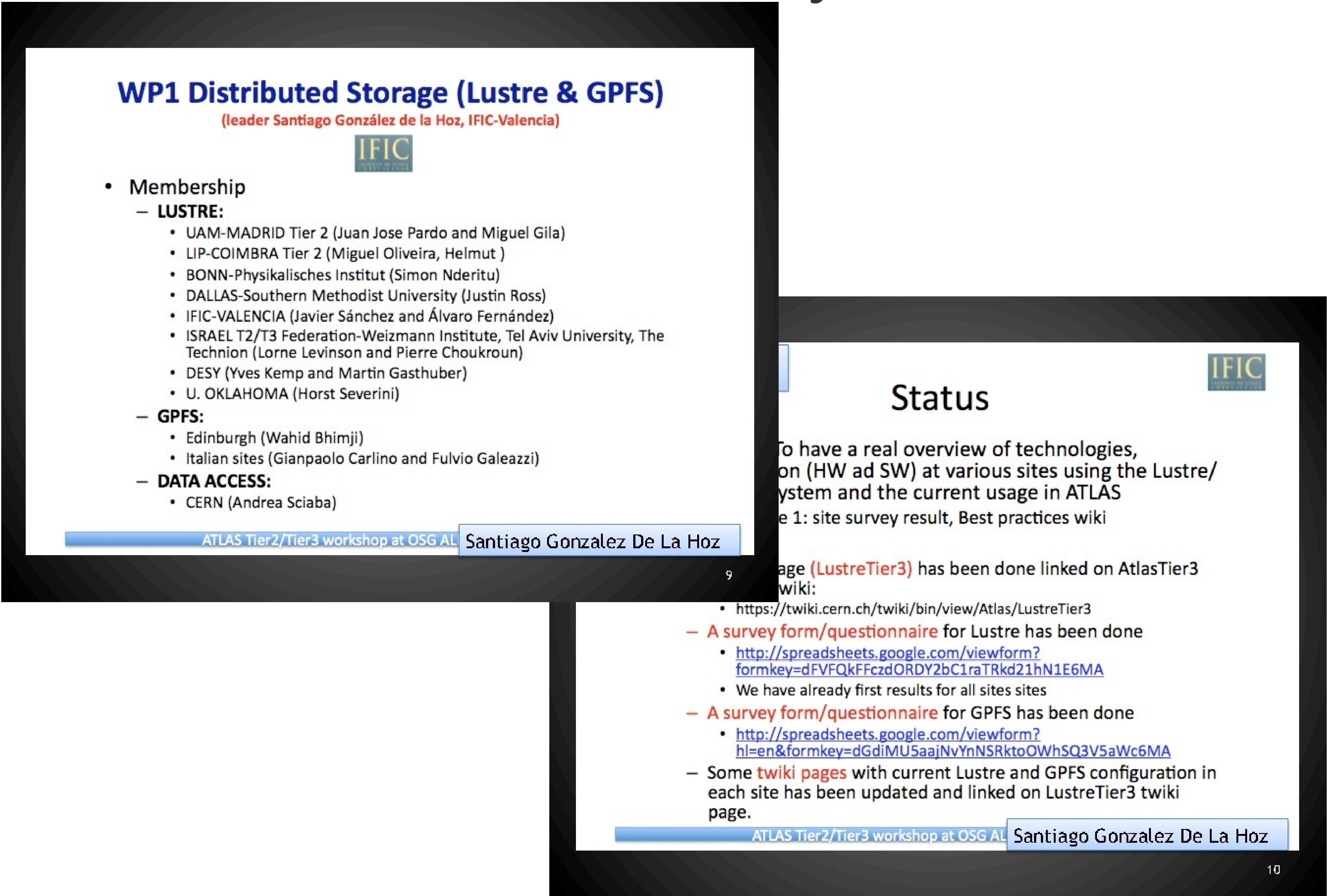
“proof” farm



Type B

- Worker-local-storage-rich nodes (eg. for xrootd)

Site survey



WP1 Distributed Storage (Lustre & GPFS)

(leader Santiago González de la Hoz, IFIC-Valencia)

IFIC

- Membership
  - LUSTRE:
    - UAM-MADRID Tier 2 (Juan Jose Pardo and Miguel Gila)
    - LIP-COIMBRA Tier 2 (Miguel Oliveira, Helmut)
    - BONN-Physikalisches Institut (Simon Nderitu)
    - DALLAS-Southern Methodist University (Justin Ross)
    - IFIC-VALENCIA (Javier Sánchez and Álvaro Fernández)
    - ISRAEL T2/T3 Federation-Weizmann Institute, Tel Aviv University, The Technion (Lorne Levinson and Pierre Choukroun)
    - DESY (Yves Kemp and Martin Gasthuber)
    - U. OKLAHOMA (Horst Severini)
  - GPFS:
    - Edinburgh (Wahid Bhimji)
    - Italian sites (Gianpaolo Carlino and Fulvio Galeazzi)
  - DATA ACCESS:
    - CERN (Andrea Sciaba)

ATLAS Tier2/Tier3 workshop at OSG AL Santiago Gonzalez De La Hoz

9

IFIC

Status

o have a real overview of technologies,
on (HW ad SW) at various sites using the Lustre/
ystem and the current usage in ATLAS
e 1: site survey result, Best practices wiki

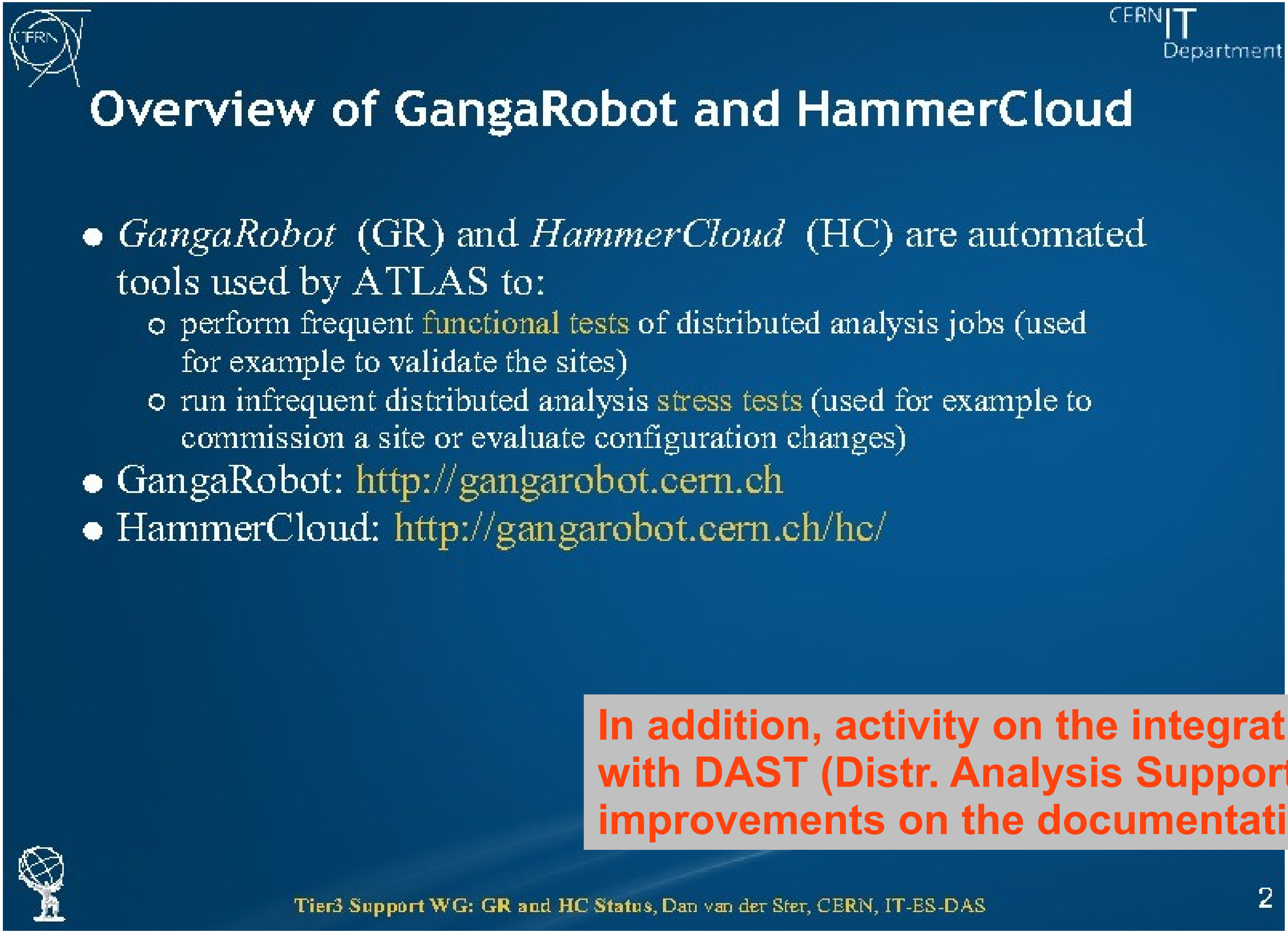
age (LustreTier3) has been done linked on AtlasTier3
wiki:
- https://twiki.cern.ch/twiki/bin/view/Atlas/LustreTier3

- A survey form/questionnaire for Lustre has been done
  - http://spreadsheets.google.com/viewform?formkey=dFVFQkFFczdORDY2bC1raTRkd21hN1E6MA
  - We have already first results for all sites sites
- A survey form/questionnaire for GPFS has been done
  - http://spreadsheets.google.com/viewform?hl=en&formkey=dGdiMU5aajNvYnNSRktoOWhSQ3V5aWc6MA
- Some twiki pages with current Lustre and GPFS configuration in each site has been updated and linked on LustreTier3 twiki page.

ATLAS Tier2/Tier3 workshop at OSG AL Santiago Gonzalez De La Hoz

10

Software distribution



Overview of GangaRobot and HammerCloud

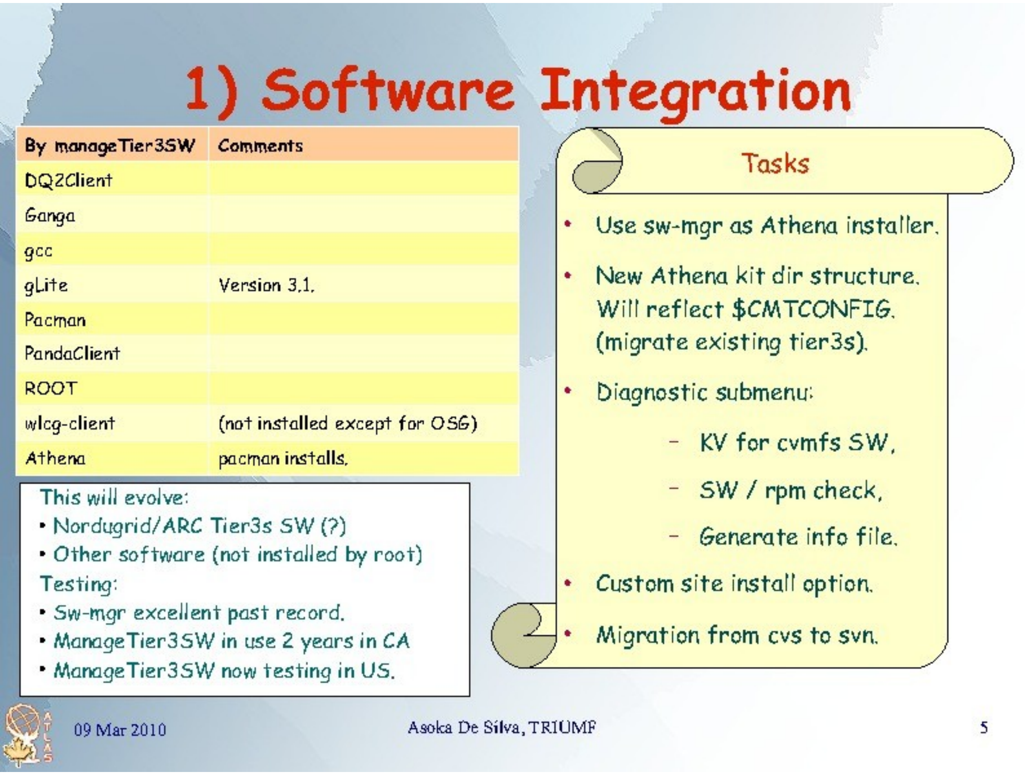
- *GangaRobot* (GR) and *HammerCloud* (HC) are automated tools used by ATLAS to:
  - perform frequent functional tests of distributed analysis jobs (used for example to validate the sites)
  - run infrequent distributed analysis stress tests (used for example to commission a site or evaluate configuration changes)
- GangaRobot: http://gangarobot.cern.ch
- HammerCloud: http://gangarobot.cern.ch/hc/

Tier3 Support WG: GR and HC Status, Dan van der Ster, CERN, IT-ES-DAS

2

In addition, activity on the integration with DAST (Distr. Analysis Support) and improvements on the documentation

SW installation WG



1) Software Integration

By manageTier3SW	Comments
DQ2Client	
Ganga	
gcc	
gLite	Version 3.1.
Pacman	
PandaClient	
ROOT	
wlcg-client	(not installed except for OSG)
Athena	pacman installs.

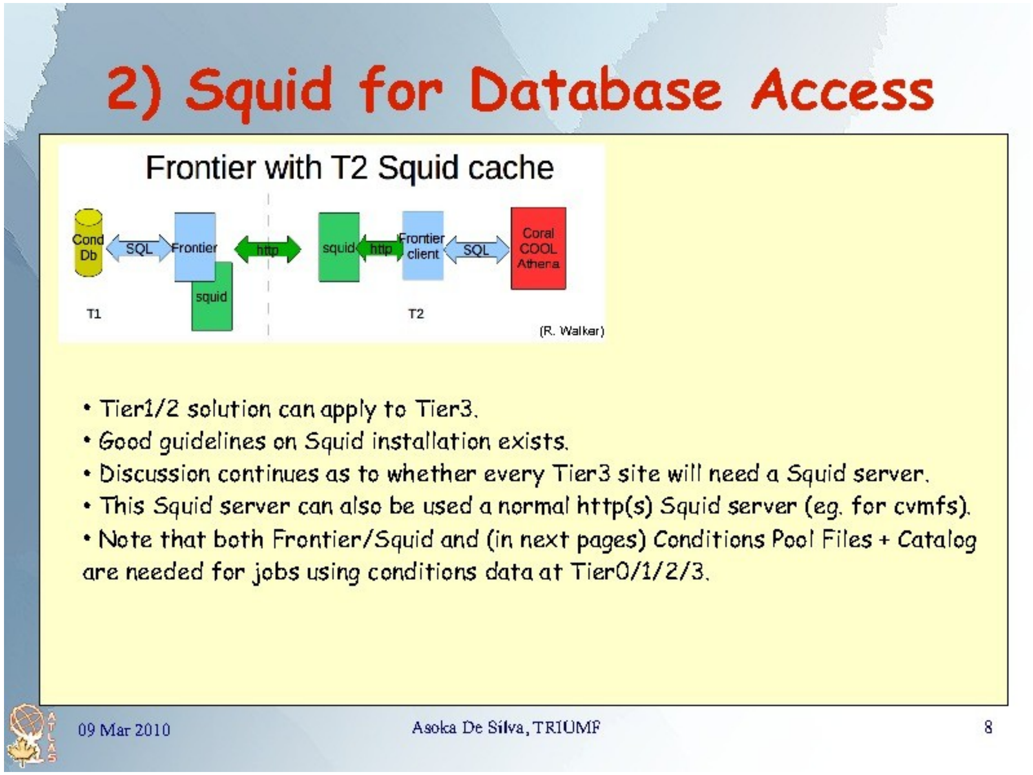
This will evolve:
- Nordugrid/ARC Tier3s SW (?)
- Other software (not installed by root)

Testing:
- Sw-mgr excellent past record.
- ManageTier3SW in use 2 years in CA
- ManageTier3SW now testing in US.

Tasks

- Use sw-mgr as Athena installer.
- New Athena kit dir structure. Will reflect $CMTCONFIG. (migrate existing tier3s).
- Diagnostic submenu:
  - KV for cvmfs SW,
  - SW / rpm check,
  - Generate info file.
- Custom site install option.
- Migration from cvs to svn.

09 Mar 2010 Asoka De Silva, TRIUMF 5



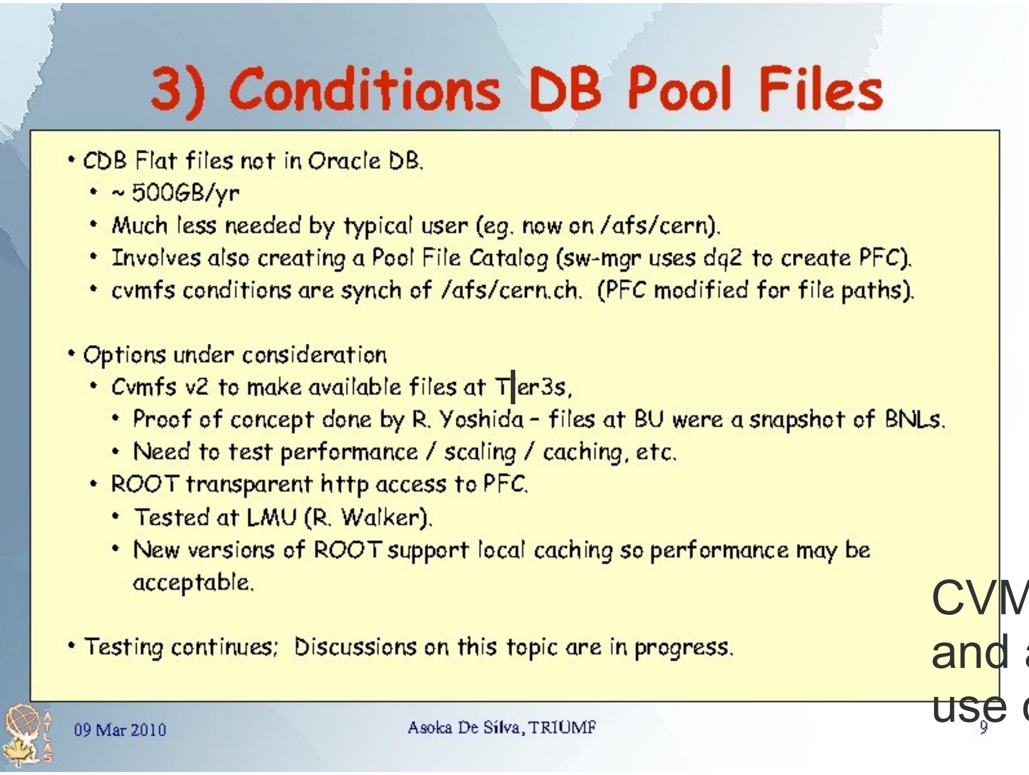
2) Squid for Database Access

Frontier with T2 Squid cache

Cond Db
SQL
Frontier
squid
http
squid
http
Frontier client
SQL
Coral COOL Athena
T1
T2
(R. Walker)

- Tier1/2 solution can apply to Tier3.
- Good guidelines on Squid installation exists.
- Discussion continues as to whether every Tier3 site will need a Squid server.
- This Squid server can also be used a normal http(s) Squid server (eg. for cvmfs).
- Note that both Frontier/Squid and (in next pages) Conditions Pool Files + Catalog are needed for jobs using conditions data at Tier0/1/2/3.

09 Mar 2010 Asoka De Silva, TRIUMF 8



3) Conditions DB Pool Files

- CDB Flat files not in Oracle DB.
  - ~ 500GB/yr
  - Much less needed by typical user (eg. now on /afs/cern).
  - Involves also creating a Pool File Catalog (sw-mgr uses dq2 to create PFC).
  - cvmfs conditions are synch of /afs/cern.ch. (PFC modified for file paths).
- Options under consideration
  - Cvmfs v2 to make available files at Tier3s,
    - Proof of concept done by R. Yoshida – files at BU were a snapshot of BNLs.
    - Need to test performance / scaling / caching, etc.
  - ROOT transparent http access to PFC.
    - Tested at LMU (R. Walker).
    - New versions of ROOT support local caching so performance may be acceptable.
- Testing continues; Discussions on this topic are in progress.

09 Mar 2010 Asoka De Silva, TRIUMF 9

CVMFS interesting for conditions data and as distributed file system. Different use cases under investigation

The ATLAS Tier3 VM Workgroup

- March - May 2010
- Not only observations and recommendations, but also Tests / Developments
- Members:
  - Torre Wenaus, Massimo Lamanna, Doug Benjamin, Sergey Panitkin, Amir Farbin, Waruna Fernando, Harris Kagan
- This surely will not cover all the existing work inside ATLAS.
- Please feel free to contact me for any suggestions / contributions

Immediate needs (for theTier3s).
Longer term perspective (clouds...)

Next stop?

- ATLAS SW week: CERN (April 19-23)
- Interest to exchange ideas/plans with other colleagues (notably CMS)