CERN IT Department

CHEP 2010 Summary

19 November 2010

Ivan Fedorko
Joao Fernandes
Wojciech Lapka
Giuseppe Lo Presti
Alan Silverman

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Overview

- Some numbers
- Highlights
- Plenary talks
- Computing Fabrics and Networking
- Collaborative Tools
- Grid and Cloud Middleware
- Distributed Processing and Analysis
- The written report (IT-Note-2010-007) is at http://cdsweb.cern.ch/record/1305848/files/CHEP%202010%20Report.pdf

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CHEP 2010 in numbers

CERN IT Department

- 495 participants
  - Cf. Interlaken 516, Mumbai 450-480, Victoria 470, Prague 615
- 12 plenaries
  - 25% less than usual, higher quality?
- 256 oral presentations
  - 25% more than usual, one fewer parallel stream
- 197 posters scheduled
  - 15-20% no show, typical
- 2 11-12 course banquets (plus 1 for IAC,PC)
  - Too much!

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Highlights

- LHC and LHC computing works, from DAQ through to analysis
- This includes LCG so little controversy this time
- Virtualisation is everywhere
- So are clouds but no common definition and many home-built solutions. Some concrete production examples exist but beware of hype
- My tenth and last CHEP, glad it was a good one

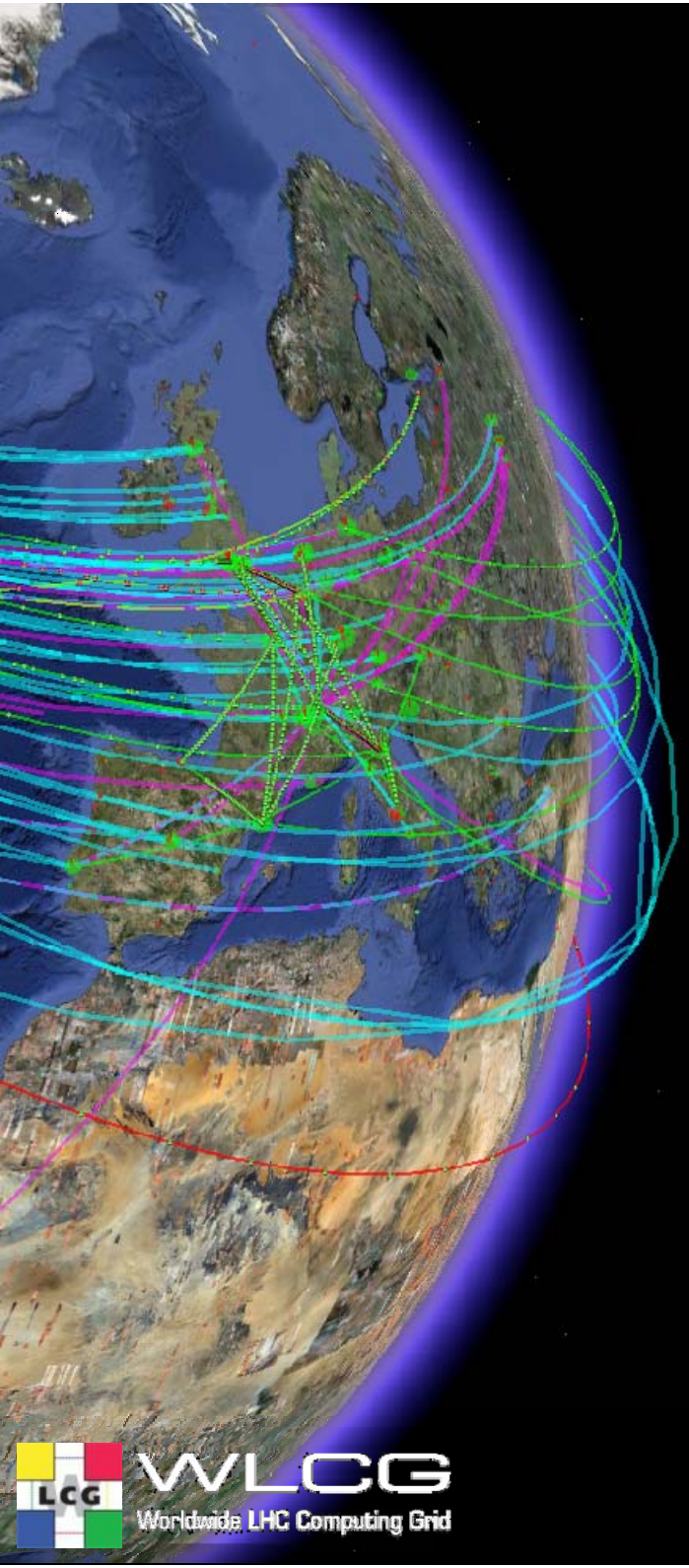
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Plenary – WLCG (Ian)

CERN IT Department

- Data transfer capability able to manage much higher bandwidths than expected/feared/planned
- The MoU has been an important tool in bringing services to an acceptable level
- Level of problems is still fairly significant: 5-6 per month require formal analysis
- Hardware is not reliable, no matter if it is commodity or not
- Deployment of upgrades/new services is very slow
- Have we (HEP) really understood how to use a distributed architecture?
- Areas for improvement - Grid Middleware, Global AAI, Fabrics. And, especially, Data Management

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it



Conclusions

- Distributed computing for LHC is a reality and enables physics output in a very short time
- Experience with real data and real users suggests areas for improvement –
  - The infrastructure of WLCG can support evolution of the underlying technology

Offline systems (Roger Jones)

CERN IT Department

- The first year of operations has been a great success
- The software for the experiments has been remarkably stable
- Tier 0 has worked very well (thanks CERN!)
- Data distribution worked very smoothly - CERN to OPN has exceeded 1GBps
- The Grid - The data side is still presenting challenges
- The Tier 3s are (by definition) not part of the central computing systems but by definition they are a vital part of the offline system

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Plenary topic - clouds

CERN IT Department

- Craig Lee (OGF) – many working examples
- Issues for HEP – security, cost, execution model, data access, SLA
- His expectation – private clouds will dominate
- We need standards
- Kate Keahey (Nimbus) – described Nimbus
- Compared benefits and challenges, Nimbus can help
- Clouds are here to stay, get on board

Multi-Cores and GPUs - Sverre

CERN IT Department

- Real consequence of Moore's law - We are being "drowned" in transistors, in order to profit we need to "think parallel"
- In the near future: Hundreds of CPU slots ! And, by the time new software is ready: Thousands !!
- GPU - Lots of interest in the HEP on-line community; e,g, Nvidia Fermi GPU has peak single precision floating point performance above 1 Tflop
- Software - We need forward scalability, we cannot afford to "rewrite" our software for every hardware change
- Options – rely on event parallelisation; forking (run through first event then fork N processes and rely on OS to copy on write); re-write as a multi-thread paradigm
- Control memory usage - We must not let memory limitations decide our ability to compute!
- Surround yourself with good software tools

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Networking – Harvey Newman

CERN IT Department

- HEP is a driver of mission-oriented networks; soon move to 40G and 100G
- Flattening of tier hierarchy (P2P) and a move to pull models (job requests data) imply greater reliance on network performance
- Creation of a Requirements Working Group to investigate future network requirements
- In future, an infrastructure of infrastructures, many players
- We must continue to address the Digital Divide in many world regions

Outreach – Lucas Taylor

CERN IT Department

- Risks of inviting media to LHC events, but do we have a choice?
- Strategy – open to media; high quality, consistent messages; exploit Web 2.0
- Work with story makers, e.g. UK BBC TV
- For younger generation, exploit new media, Twitter, Facebook, blogs, Youtube
- Live webcasts, ATLAS-Live, CMS TV
- Encourages everyone (?) to get involved

How well are we doing?

Language monitoring of online and print media

http://www.languagemonitor.com/news/top-words-of-2009/

Top Phrases of 2009

1. King of Pop
2. Obama-mania
3. Climate Change
4. Swine
5. Too Large to Fail
6. **Cloud Computing**
7. Public
8. Jai Ho!
9. Mayan Calendar
10. **God Particle**

Top Words of 2009

1. **Twitter**
2. Obama
3. H1N1
4. Stimulus
5. Vampire
6. **2.0 (next gen.)**
7. Deficit
8. **Hadron**
9. Healthcare
10. Transparency

Top Names of 2009

1. Barack Obama
2. Michael Jackson
3. Mobama
4. **Large Hadron Collider**
5. Neda Agha Sultan
6. Nancy Pelosi
7. M. Ahmadinejad
8. Hamid Karzai
9. Rahm Emmanuel
10. Sonia Sotomayor

with nothing at all LHC-related in 2008

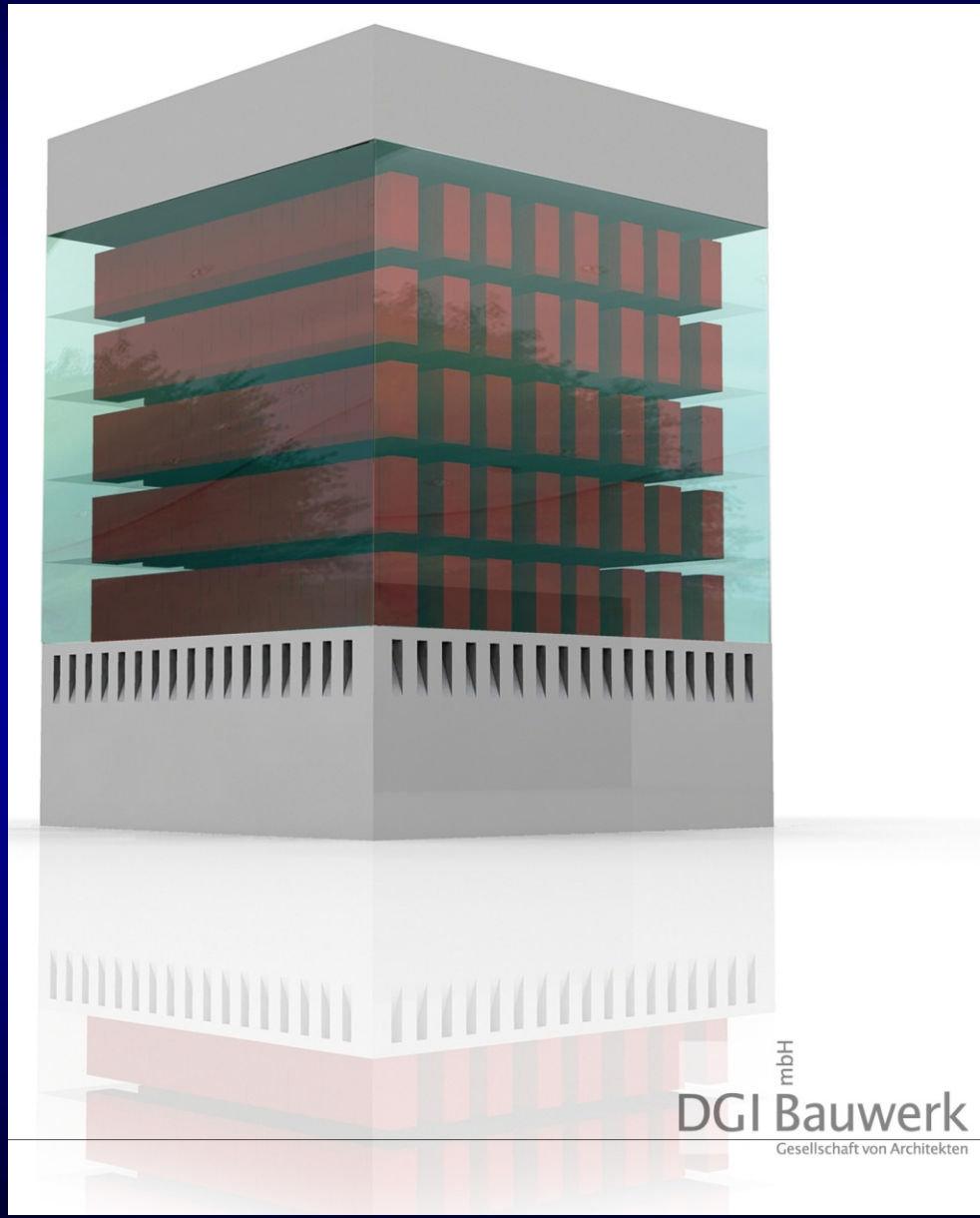
FAIR project at GSI

CERN IT Department

- Facility for Antiproton and Ion Research
- Data to be recorded in 2018: 1-10 x LHC
- Cost €1,027M, Experiments €220M, where is budget for computing? [Sound familiar?]
- DAQ – O(10^6 tracks), no hardware trigger, rely on event filtering, GPUs, high speed networks
- Datacentre plans – the Cube – 800 racks, 6MW cooling, 3D design

Plans for a new (dark) Datacenter

Minimal floor footprint
Space for 800 19“ racks
Planned cooling power 6 MW
(building supports more)
Internal temperature 30°
Minimal construction cost
Fast to build
Autonomous rack power mgnt.
Back door rack cooling
Smoke detection in every rack
Use of heat for building heating
Shortest cable lenghts
Water pressure below 1bar avoiding risk of spills
Use of any commercial 19“ architecture
Unmatched packing- and power density
Construction cost about 10 M€



Data Storage – Ian Fisk

CERN IT Department

- Technical Challenges – capacity, complexity, different technologies, protection of data, random access for analysis
- Greater access of data from disc, data placement becoming more important
- Use wide area file systems, e.g. Lustre?
- Cannot discuss data management without discussing networking and access patterns
- Modifications in the access and management might have big gains in efficiency

Data Preservation – D.South

CERN IT Department

- Most serious attempt yet to preserve an experiment
- Why? New phenomena appear, redo or perform new measurements
- Usually not part of initial planning
- Safeguarding data is not enough – missing background (why it was done that way), expertise (people retire or move on), working environment (can the software be run?)
- DPHEP – study group on data preservation
- Create a model for preservation
- Inspire – store doc, meta-data, data even

Experience (Rob Quick, FNAL)

- No Substitute
  - 20+ Years on Staff
  - Over 9000 Tickets Resolved
- Let the Experience Show
- Enjoy the Ride

"Some people believe football is a matter of life and death. I'm very disappointed with that attitude. I can assure you it is much, much more important than that."
Bill Shankly

CF

Computing Facilities

CERN IT Department

@ CHEP 2010

Ivan Fedorko

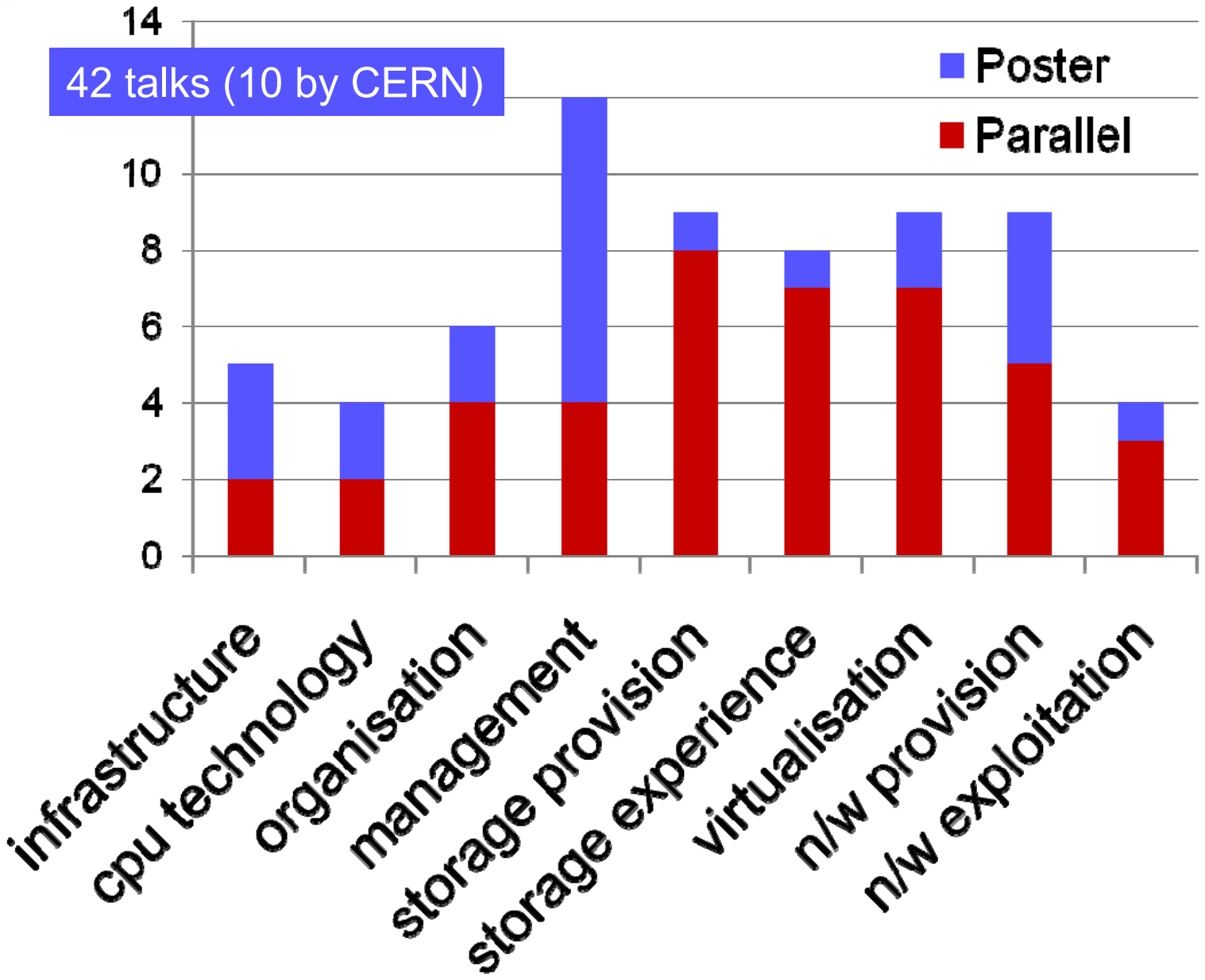
19/11/2010

- Session Summary
- Storage
- Networking
- Virtualization
- Fabrics

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

CERN

From Session Summary by Tony Cass



Session Message

From Session Summary by Tony Cass

- Fabrics working well!
- Many interesting presentations
  - No Luster/GPFS, IPv6
  - Puppet is hot, quattor not
- Well attended
- Virtualization topics split across 3 tracks
  - Dedicated track for CHEP '12?
    - or will it all be routine by then?
- We seem to be addressing many of concerns but…
  - wheels are often reinvented
  - developments sometimes occur in isolation
- Still scope for improved collaboration between sites and between different work areas.

Storage I - Castor

Storage Service Developments at CERN (PS35-5-304)

Exabyte Scale Storage at CERN (PS52-1-303)

First Step

LHC Experiments

- Use Castor for what it was designed for and for what it is good at

• Castor release 2.1.9 has performed well during 2010 and fixed SRM scalability issues

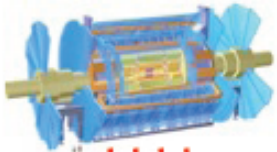
• The EOS disk pool demonstrator is currently being tested with experiments

RAL
TRIUMF

tape servers

October 2010

LHC Experiments



ANALYSIS

XROOT or mountable file system

ASGC
BNL
FNAL
FZK
IN2P3
CNAF
NDGF
NIKHEF
PIC
RAL
TRIUMF

Tier-1s data replication

Castor

Disk Pools

Managed repli

tape servers

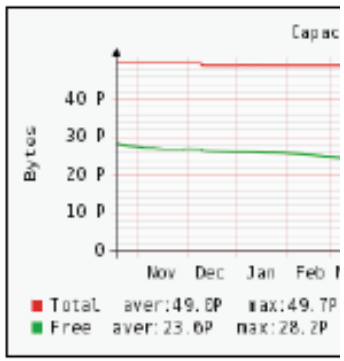
Scalable, secure, accountable, globally accessible, manageable

Allow to choose service level for
availability
reliability
performance
decoupled from HW

Storage II - Tapes

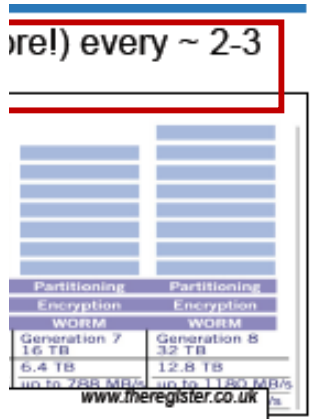
Tape Archive Challenges When Approaching Exabyte-scale (PS46-3-310)

- CERN Tape Arc
- Real rates are h
  - Oct 2009 to Apr
  - May 2010 to no
  - November (Hea



Conclusions

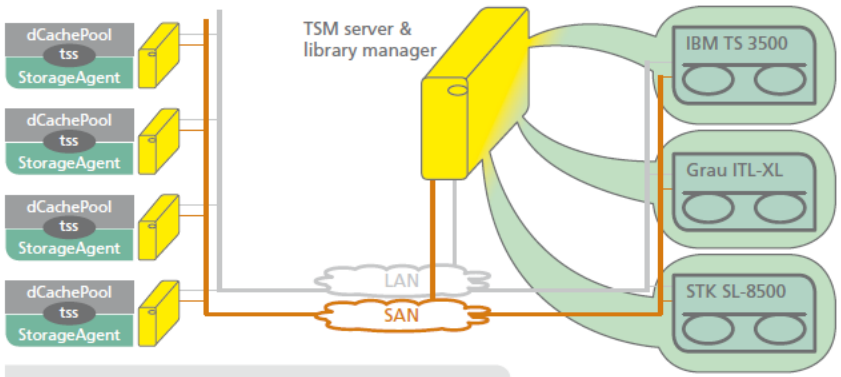
- Managing a near-Exabyte tape archive is an active task. The effort is proportional to the total archive size.
- A non-negligible fraction of resources need to be allocated for housekeeping such as migration and verification.
- Tape has a small *effective* lifetime requiring continuous media migration to new generations



The dCache-TSM connection without ERMM

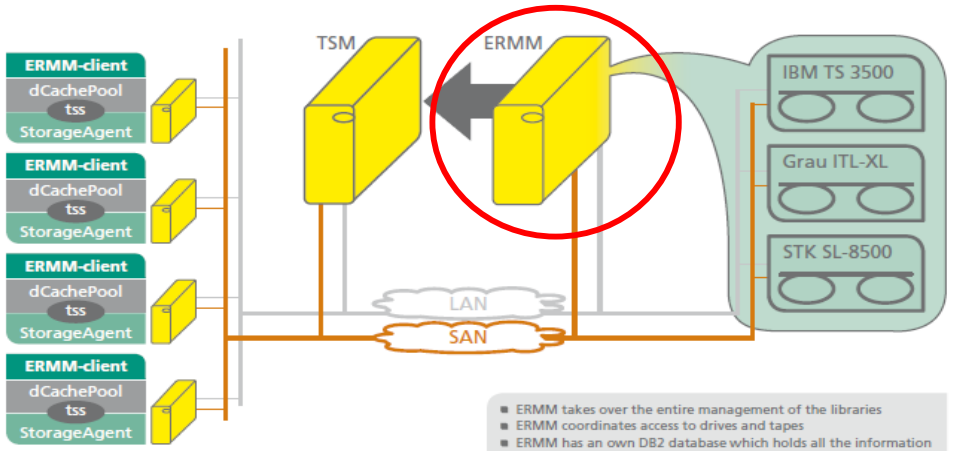
KIT Karlsruhe Institute of Technology

Tape library virtualization



The dCache-TSM connection with ERMM

KIT Karlsruhe Institute of Technology



Storage III – NFS

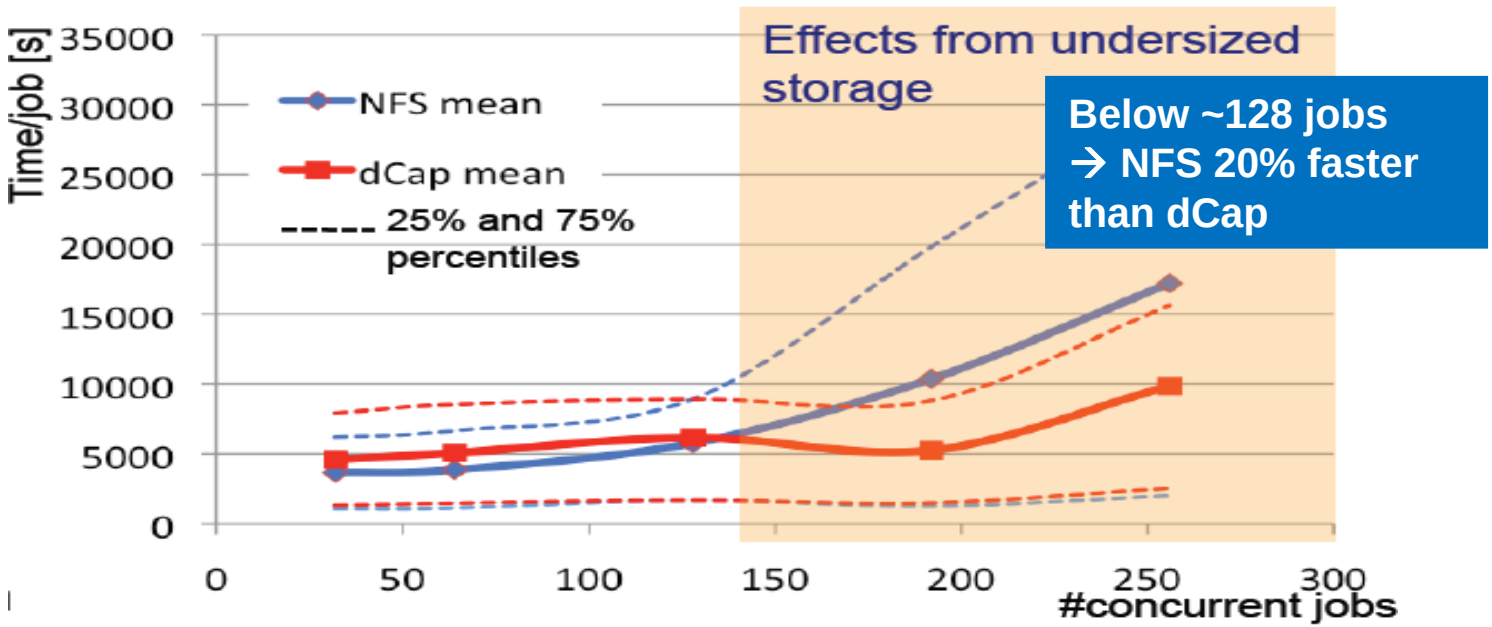
LHC Data Analysis Using NFSv4.1 (pNFS): A Detailed Evaluation. (PS35-4-289)

> Presented results

- **Synthetic**: Provide general performance and stability measurements of NFS 4.1/pNFS
- **ATLAS HammerCloud:** Stable and well-performing running over four days
- **CMS analysis:** See effects of FS cache, excellent behavior of NFS up to some point
- **ROOT files**: See effects of FS cache, better performance than dcap, even with most recent ROOT version and with TreeCache enabled

> **NFS 4.1/pNFS has advantages over traditional proprietary protocols**

> **We now know: Performance is one of them!**



RAID vs SSD

Establishing Applicability of SSDs to LHC Tier-2 Hardware Configuration (PS35-3-288)

Why not to use the faster SSD technology rather than HDDs?

Tested HW (only "affordable" solutions):

- 7200RPM 500GB SATA HDD
- Kingston SSDNow V-series 128GB SSD
- Intel X-25 G2 M 160GB SSD

Testing SW:

- blkparse, seekwatcher, HammerClouds

Measured:

- I/O, Seek Count, Throughput

RESULT: RAID solution is more efficient

Jobs:Cores	Storage	Efficiency	Throughput
Standard Node			
8:8	1×Kingston Value SSD	60%	4.5
8:8	1×SATA HDD	75%	5.5
8:8	1×Intel X25 SSD	80%	6
8:8	2×SATA HDD (RAID 1)	83%	6.6
8:8	2×SATA HDD (RAID 0)	90%	7

Virtualization

Batch virtualization

- CERN batch - based on golden quattor-mgmt node
- CernVM
- WNoDeS (INFN)

Virtualization (with iSCSI, for the LHCb etc.)

- Xen,KVM,Hyper-V, VMware: ~10 talks with message about evaluation
  - missing comparison methodology and benchmarking?
- Challenge of virtualization dedicated hardware (SCADA, CANBUS, PCI cards), license dependency (PVSS) etc.

Networking

- Providers are ready for 40GE & 100GE
- OTN (Optical Transport Network)
  - baseline of the next network transport layer
  - error recovery functionality → signal at 40Gbps over 1650km without regeneration
  - optical + packet networks = hybrid networks
- Network infrastructure as a service
  - bandwidth guarantees, traffic isolation (secure end-to-end connection), data caching etc.
- Move from static lightpaths to dynamic circuits (end-to-end circuits)
- 10GE well established
  - Now activities toward 100GE networks

Networking

FDT, SRM, GridFTP performance comparison

- Tests done on 7.2Gbps (900MB/s) route CERN-Caltech (measured mem-to-mem)
- Tested on 100 x 1GB file set, Hadoop – Hadoop transfer, measured execution time of the transfer command – all overhead (auth, etc) is included
- GridFTP (globus-url-copy)
  - -fast -parallel 35 (num. of streams): 74 mins, ~23MB/s
  - default buffer values
  - transfer files in parallel not supported
- SRM (srmcp)
  - -streams_num=35 -send_cksm=false: 80 – 120 mins, 14MB/s - 21MB/s
  - default buffer values
  - transfer files in parallel not supported
- FDT (fdtcp)
  - 35 streams, 1 file (no parallel files): 41min, ~42MB/s
  - 35 streams, 12 files in parallel: 8 - 12min, 142MB/s – 213MB/s
  - parallel files optimized given the distributed storage cluster size

CF - Fabric Management

Development:

- Cluman
- SysMES
  - Rule based tool
    - Monitoring
    - Problem recognition & solution trigger (rules)
    - Problem Solution (Tasks)
  - Heterogeneous HW/SQ @ Alice HLT

Single rule:
if E1.name = ntpd_down then restart ntpd
Complex Rule :
if Count(E1.name = ntpd_down)>3 and Δt < 5min
then { New event; Mail; SMS }

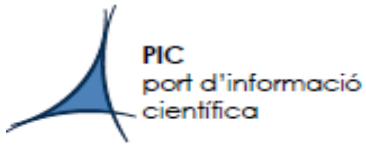
Puzzle:

- Fabric management using Open Source tools

CF - Configuration Management

Unified Fabric Management System with Open Source Tools (PS11-4-374)

Why Not Something Else?



Puppet

-uniform

Around 600 nodes configured

Updating every 15 minutes

Around 30 different profiles used

- Some machines are servers which are not massively deployed, but it's still nice if you want to have rapid recovery

SVN to

We don't have all services at PIC running with puppet

- torque server is not using puppet, worker nodes are

We still rely on yaim for configuration of glite middleware

umans),

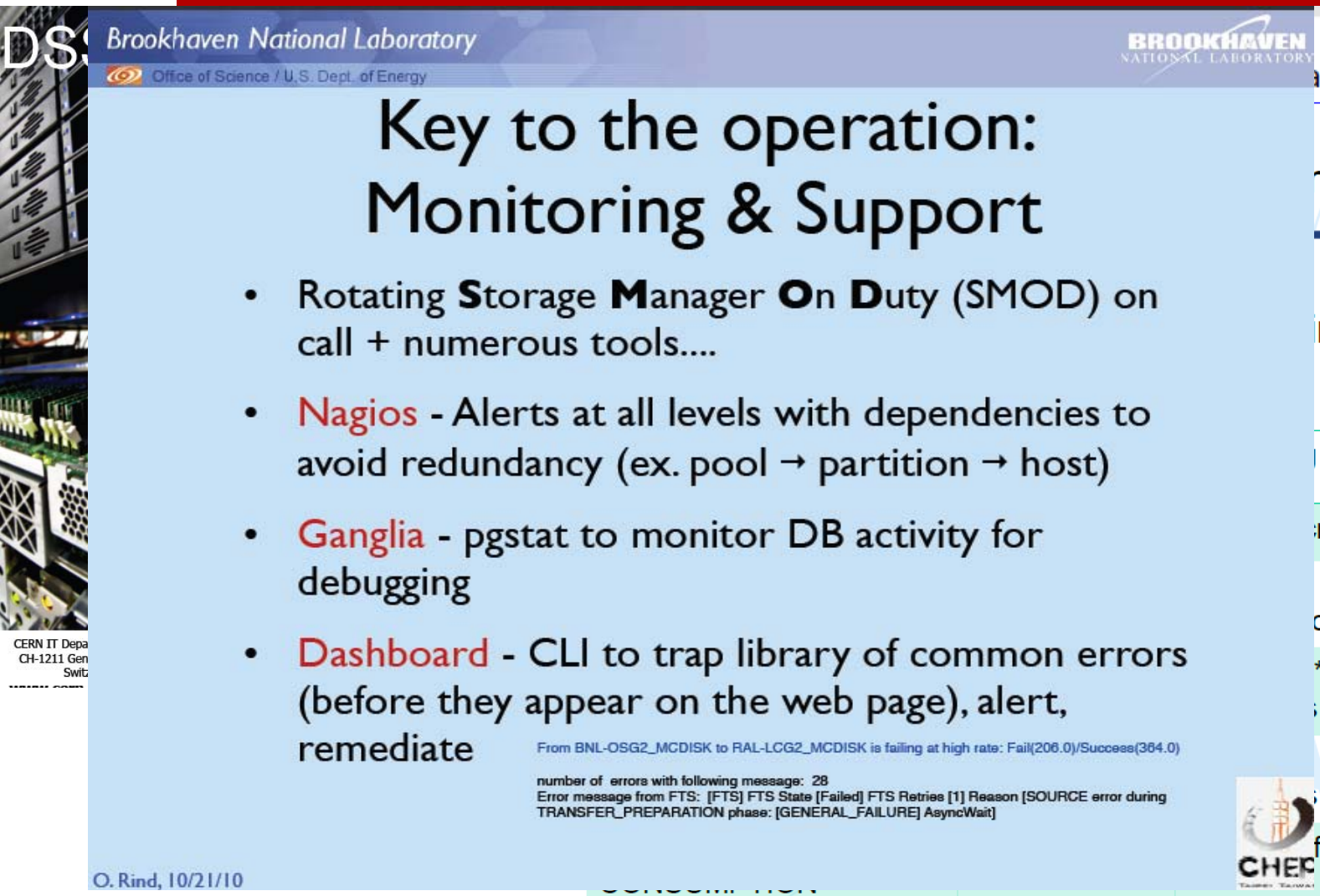
- It would be nice to eliminate the dependency but it's too much effort, specially regarding updates
- But puppet is the one running yaim

CF - Scale

Site	Nodes	Cores
NSC (Novosibirsk)		1280
BNL	2000	
CNAF		7000
CIS		2500
GSI (T2)	340	2700
Prague (T2)	336	2630
CERN	8000	57k (15k CPUs)

CF - Monitoring

No talk about monitoring but monitoring almost in every talk



Brookhaven National Laboratory

Office of Science / U.S. Dept. of Energy

BROOKHAVEN NATIONAL LABORATORY

Key to the operation: Monitoring & Support

- Rotating **S**torage **M**anager **O**n **D**uty (SMOD) on call + numerous tools....
- Nagios - Alerts at all levels with dependencies to avoid redundancy (ex. pool → partition → host)
- Ganglia - pgstat to monitor DB activity for debugging
- Dashboard - CLI to trap library of common errors (before they appear on the web page), alert, remediate

From BNL-OSG2_MCDISK to RAL-LCG2_MCDISK is failing at high rate: Fail(206.0)/Success(364.0)

number of errors with following message: 28
Error message from FTS: [FTS] FTS State [Failed] FTS Retries [1] Reason [SOURCE error during TRANSFER_PREPARATION phase: [GENERAL_FAILURE] AsyncWait]

O. Rind, 10/21/10

CERN IT Depa
CH-1211 Gen
Swit

atistics

KIT

nstrumentation

ngle management

TRAPS from CENTREON

ripts

programs translate
o TCP/IP

* designed by PIC and IFAE

designed by PIC and IFAE

formation from ILOM

CHEP

No simple solution for large scale

Collaborative Tools
Track Summary

Joao Fernandes
(CERN, IT-UDS)

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Collaborative Tools

CERN IT Department

- Track 7:
  - 25 contributions in total
  - 5 proposals rejected
  - 3 moved to other tracks
- 11 oral presentations
  - 6 posters but only 3 exposed
  - Talk on Experiments Outreach Activities (Plenary)
  - 2 parallel sessions
    - Parallel Session 1: on policies and new initiatives
    - Parallel Session 2: dedicated to Systems and Tools

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Overview

CERN IT Department

- Increasing areas in the CT field
  - HD videoconferencing systems, Outreach and Inreach activities, Rich Media Content, Information systems, etc.
- Representative examples covering the activities in the HEP community
  - 1st session dedicated to Policies and New initiatives
  - 2nd session dedicated to SW systems and Collaborative Tools
- Plenary Talk (Lucas Taylor, FNAL/CMS)
  - Overview about the importance of the outreach activities for the HEP community
    - Contract with the Society
    - Need of a defined strategy: HQ messages, defined relation with the Media and use of latest multimedia technologies
    - Everyone needs to be involved

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Session 1 (Wednesday)

CERN IT Department

- **HEP Outreach Inreach and Web 2.0**
  - **Steven Goldfarb, University of Michigan, ATLAS**
    - Overview about the Web 2.0 initiatives in Internal/External communications in ATLAS
      - Forced by the fast information updates for multiple users
      - Public portals, info streams, social networks, sites blogs
      - Usage foreseen to continue and need to optimize it:
        - » Focus on Effective and High Visibility sites, move to Open Source Content Management System (Drupal)
- **CMS Worldwide: a New Collaborative Infrastructure**
  - **Lucas Taylor, Fermilab, CMS**
    - Update about the project (presented in CHEP'09)
      - Set of standard CMS remote control centers with almost permanent video presence, status displays, detector quality monitoring allowing remote shifts
        - » 1st was set FNAL in 2007 followed by CMS CC in 2008
        - » From 16 centers in 2009 to 55 centers in 2010
        - » Typical cost: 12kCHF for a generic standard installation

Session 1 (Wednesday)

CERN IT Department

- **ATLAS Live: Collaboration Information Streams**
  - **Steven Goldfarb, University of Michigan, ATLAS**
    - •Project launched in January
    - •Problem: collaborations having difficulties to find or maintain important information in the web and in disseminating it externally
    - •Installation of monitors all over CERN
    - •CERN Infrastructure of webcast and streaming
    - •Content is created and disseminated at CERN and elsewhere
    - •integrated automatically with the existing information repositories (CDS, Indico, Picasa, etc.)

Session 1 (Wednesday)

CERN IT Department

- **Physicists get Inspired: Inspire Project and Grid Applications**
  - **Jukka Klem, CERN**
    - Invenio (Digital Library SW) + Spires (former HEP information system) started in 2007
    - Next generation information source HEP open access full text articles, experimental notes, etc.
    - Gathers info from 4 labs: CERN, FNAL, SLAC and Desy
    - D4Sience initiative
      - links Inspire and other communities to build a knowledge ecosystem
      - Main uses cases: document OCR, Full text editing and bibliometrics (using Hirsch Index)
- **Perspective of User Support for the CMS collaboration**
  - **Sudhir Malik, FNAL/University of Nebraska, CMS**
    - CMS: 40 countries, 200 institutes and 3500 people.
      - Project with lifetime of 30 years
    - Challenges to the User support: Users of different background
    - Needs Organization: cycle between people where the collaborative effort is the key of success

Session 2 (Thursday)

CERN IT Department

- **Glance Information system for ATLAS**
  - **Laura Moraes, UFRJ, ATLAS**
    - General framework to access various existing applications to support experiment management and members
      - technologies (models and inventories), members, talks, papers, appointments, alarms, etc.
    - interacts with current repositories
      - Oracle/MS SQL MySQL
      - Import and export of different formats
    - Goal: decentralize ATLAS activities

The Glance Applications

CERN IT Department

ATLAS TCn Applications

The Glance Project	ATLAS Membership	ATLAS Appointment	DSS alarms viewer	ATLAS Traceability	MTF Database
more info	more info	more info	more info	more info	more info

Cable Database	ATLASeditor3D	ATLAS Survey	Activation Studies	Muon Spare	Muon Equipment
more info	more info	more info	more info	more info	more info

TDAQ Equipment	Rack Wizard	ATLAS SCAB	Analysis - Papers	Analysis - Conference Notes	ATLAS Thesis
more info	more info	more info	more info	more info	more info

Related Projects or Other TCn Applications

LHCb Traceability	CAD Web Navigator
more info	more info

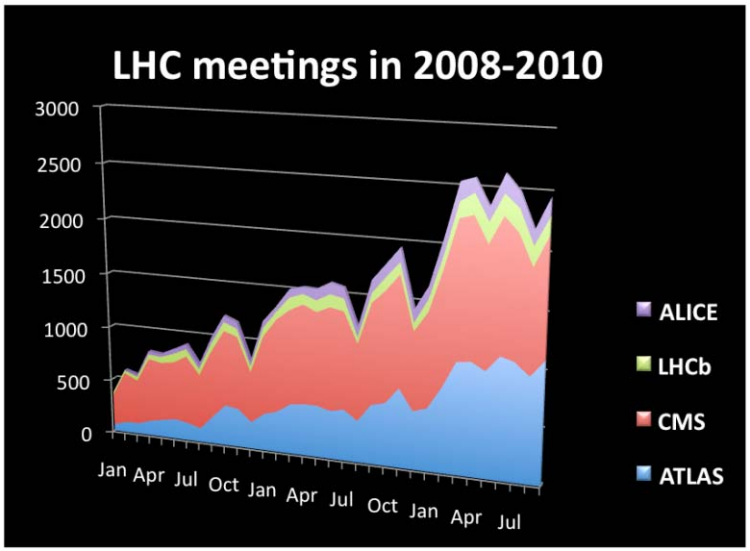
Contact Us

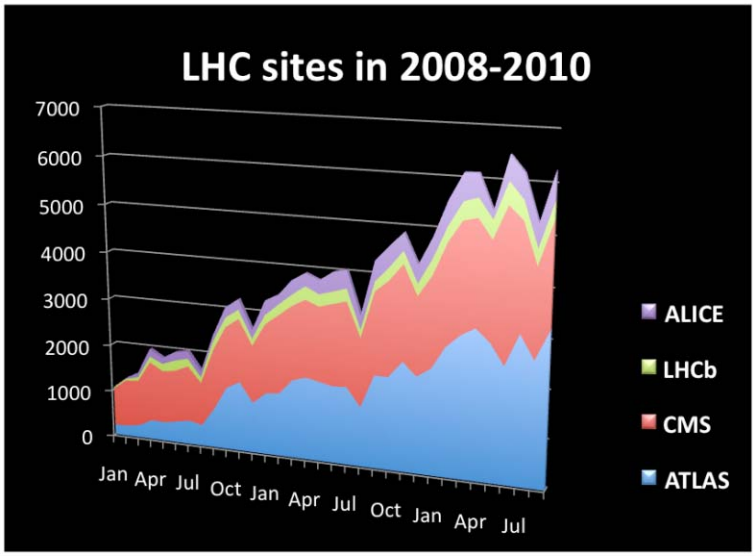
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Session 2 (Thursday)

CERN IT Department

- **EVO (Enabling Virtual Organizations)**
  - **Philippe Galvez, Caltech, CMS**





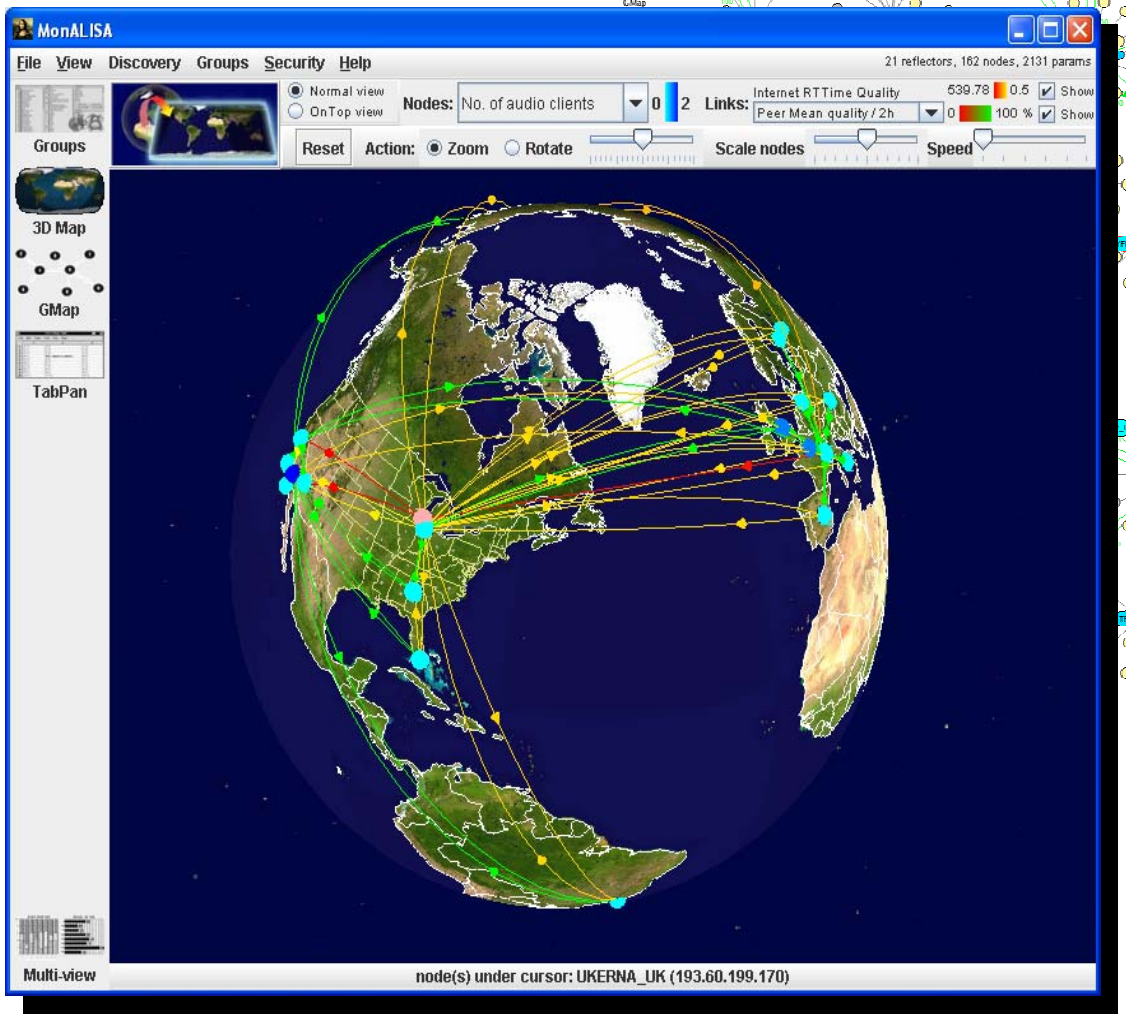
September 2010

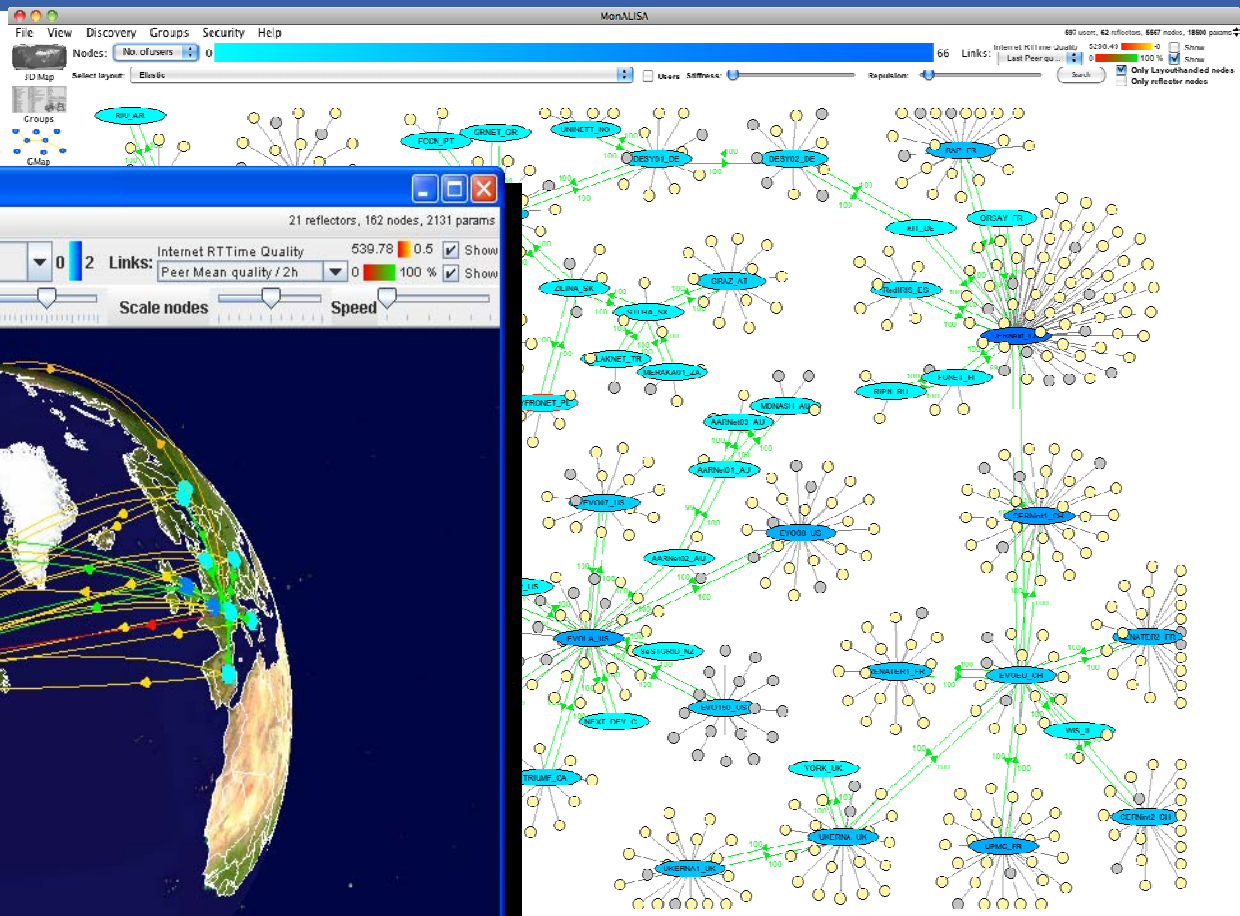
Experiments	Meetings	Participants	Phone	Skype
CMS	1046	1713	385	112
ATLAS	1097	1959	1262	110
ALICE	158	387	188	19
LHCB	153	292	65	8

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

EVO worldwide

CERN IT Department





- 60 servers worldwide
- Usage continues to grow at 60-80% per year
- Current simultaneous sites ~850

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Session 2 (Thursday)

CERN IT Department

- **University of Michigan & CERN Lecture Archiving System**
  - **Jeremy Herr, University of Michigan, CERN**
    - Goal: to have a Lecture Archiving system available for CERN users integrated in the CERN environment
    - CERN and UM lecture archiving agreement
    - Comprehensive lecture archiving based on U-M's system:
      - recording
      - processing
      - archiving
      - publishing
      - monitoring
      - Analytics
    - Recording manager developed in Indico
    - Micala (http://micala.sourceforge.net) contains all the monitoring info

Session 2 (Thursday)

CERN IT Department

- **Towards a New PDG Computing System**
  - **Juerg Beringer, Lawrence Berkeley National Laboratory**
    - international collaboration charged with summarizing Particle Physics, as well as related areas of Cosmology and Astrophysics
    - 176 authors from 21 countries and 108 institutions and 700 consultants in the particle physics community, coordinated by the PDG group at LBNL
    - New Computing model presented to face new challenges
- **AbiWord and AbiCollab – Real Time Collaborative Document Creation**
  - **Martin SEVIOR, University of Melbourne and Abisource B.V.**
    - AbiCollab allows real time collaboration between arbitrary numbers of AbiWord sessions.
    - allows real-time document creation
    - operates in a decentralized peer to peer network
    - Provides abicollab.net webservice – central document repository
    - A company has been created to commercialize abicollab

Posters Presented

CERN IT Department

- LeaRN: A Collaborative Learning-Research Network for a WLCG Tier-3 Centre
  - PEREZ CALLE, Elio (University of Science and Technology of China)
- Visualization via HD Videoconferencing
  - HLADKA, Eva (CESNET)
- Planning and Organization of an E-learning Training Program on the Analysis Software in CMS
  - LASSILA-PERINI, Kati (Helsinki Institute of Physics)

Attendance

CERN IT Department

- Auditorium AC with good audience on Wednesday
  - More than 30 attendees, Thursday lower
- Extensive discussion after each presentation
  - Usually 1 to 4 questions
- General discussions at the end of the sessions
- Presentations were generally of high quality generating a lot of interest

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Conclusion

CERN IT Department

- The track seems to have an increasing interest
- General opinion that some of the topics covered are of critical importance for the HEP community

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Grid and Cloud Middleware
Track Summary

Wojciech Lapka
(CERN, IT-GT-TOM)

Clouds were everywhere...

CERN IT Department

Grid and Cloud Middleware

- Grid/Cloud middleware and monitoring tools
- Grid/Cloud middleware interoperability
- Grid/Cloud reliability
- Grid /Cloud security
- Evolution of Grids and Clouds
- Global usage and management of resources
- Experiment-specific middleware applications

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

General

- 38 talks (9 from CERN)
- Very good attendance
  - Many active discussions
- Main Topics
  - Operational Experience
  - Operations and Monitoring
  - Messaging
  - Data Management
  - Workflow Management
  - Security
  - Clouds and Virtualization

Lessons Learnt

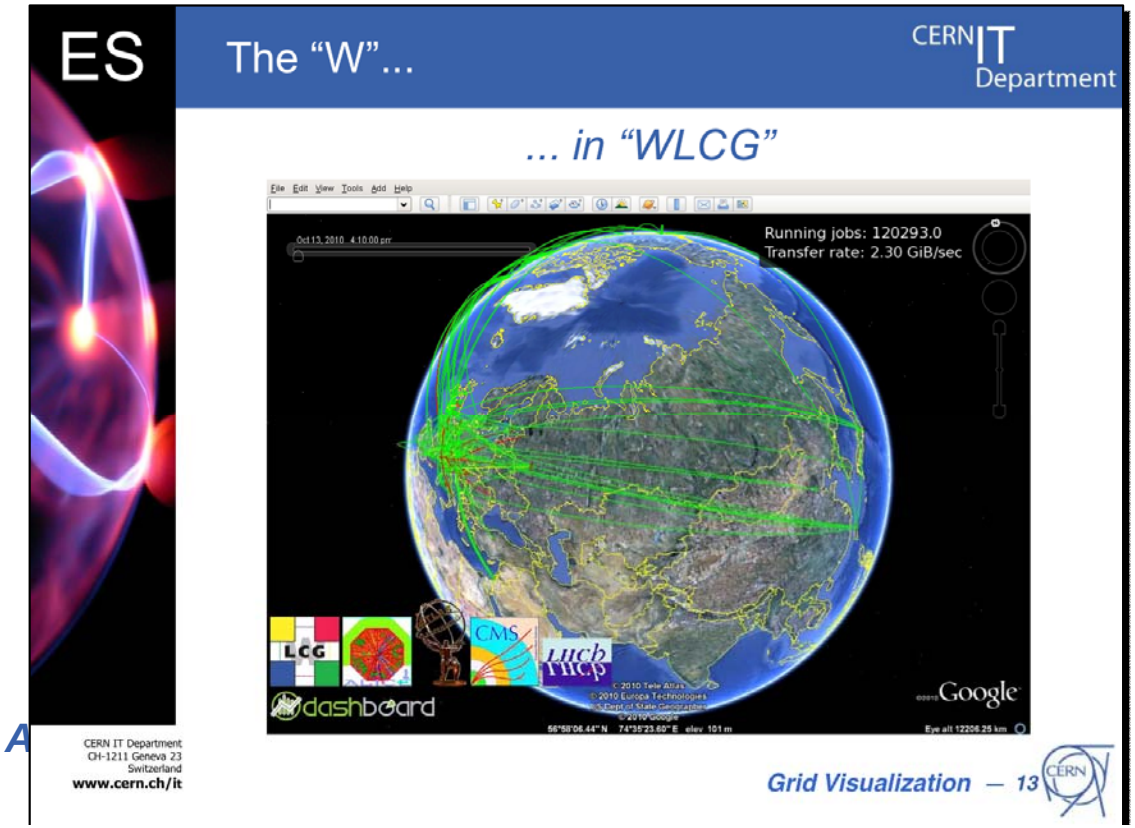
- "Ten Years of European Grids: What Have We Learnt?" by Stephen Burke (RAL)
  - Current middleware has less functionality than it was planned at the beginning
  - Complex functionality moved into the experiments' software
- Only one safe prediction for the next ten years
  - Things will change

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Monitoring

CERN IT Department

- "Visualization of the LHC Computing Activities on the WLCG Infrastructure", David Tuckett (IT-ES)
  - Promoting the WLCG
  - A tool for WLCG experts.



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Monitoring

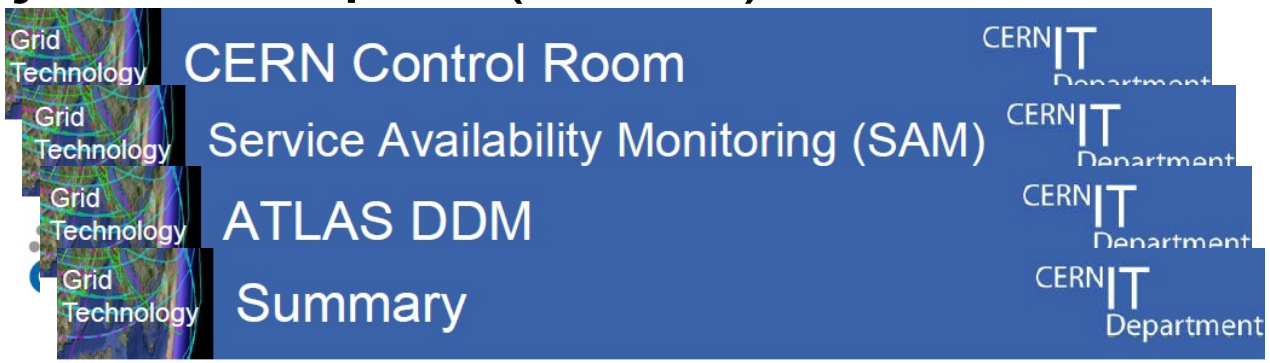
- "Flexible Availability Calculation Engine for WLCG", Wojciech Lapka (IT-GT) & GridView Team (BARC)

Summary

ACE satisifies the requirements of the LHC experiments

Messaging

- "A Messaging Infrastructure for WLCG", Wojciech Lapka (IT-GT)

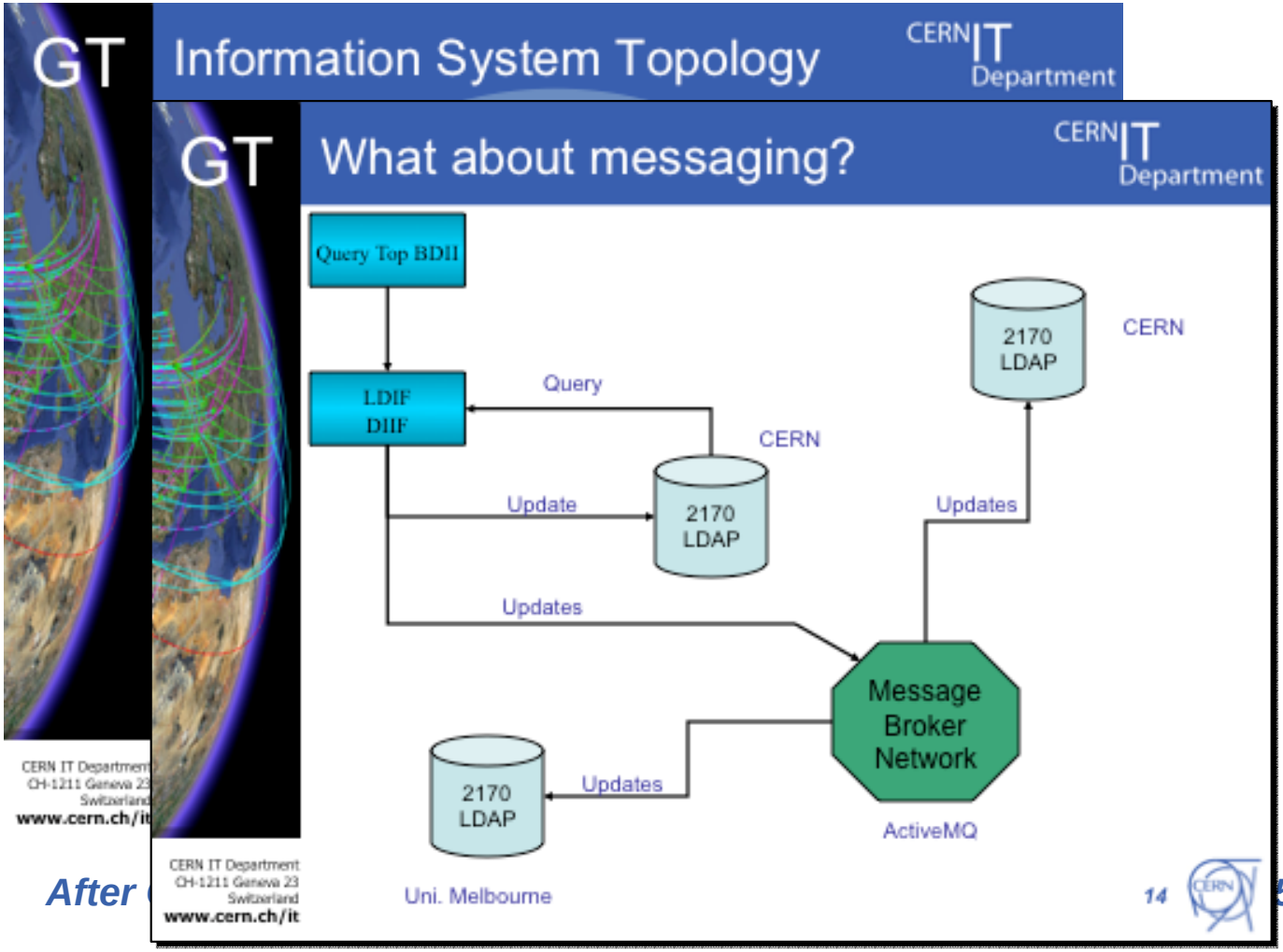


Messaging is a key technology for WLCG

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Information Systems

- "Designing the Next Generation Grid Information Systems" by Laurence Field (IT-GT)

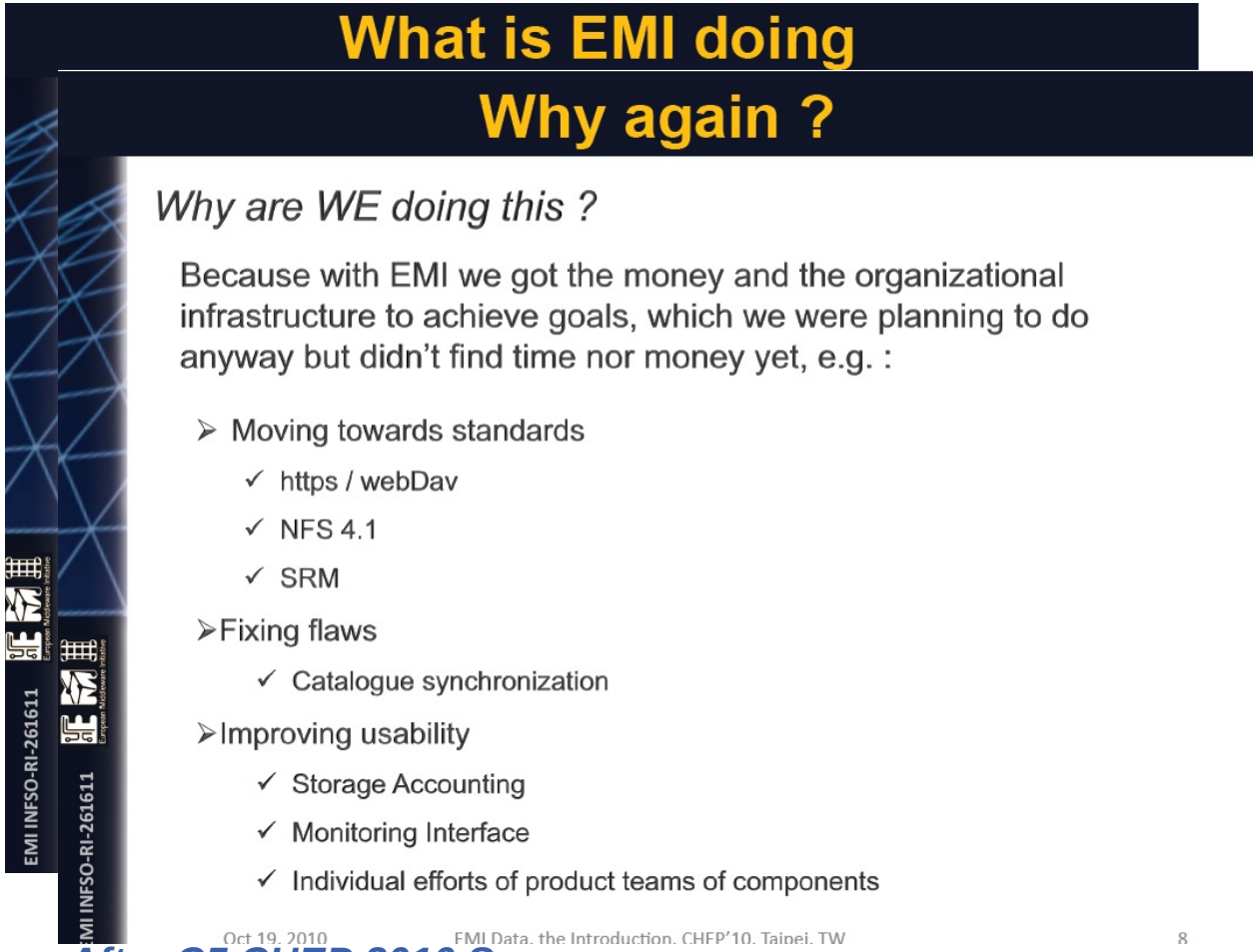


After (

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

EMI (European Middleware Initiative)

- "EMI, the Future of the European Data Management Middleware" by Patrick Fuhrmann (DESY)



After C5 CHEP 2010 Summary

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Data Management (1/2)

CERN IT Department

- "Services for Grid Data Management", Oliver Keeble (IT-GT)

- The main plans are:
  - Improve performance (clients, services..)
  - Move towards standards (NFS-4.1,..)
  - DPM/LFC (Replication, Monitoring, Accounting, ...)
  - FTS (less hierarchical structure)
  - Usage of Messaging Technologies

Data Management (2/2)

CERN IT Department

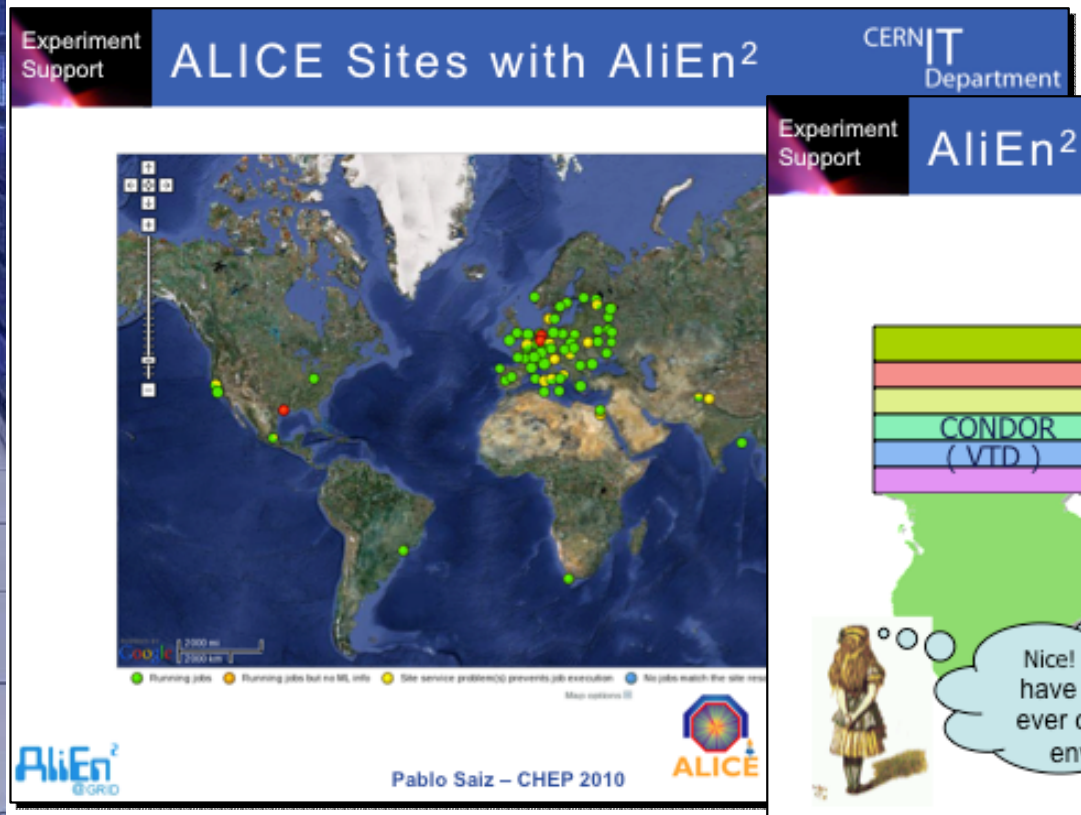
- "Standard Protocols in DPM" by Ricardo Rocha (CERN, IT-GT)
  - DPM: lightweight solution for disk storage management
  - Used in over 200 sites
  - Largest deployment: 1.5 PB

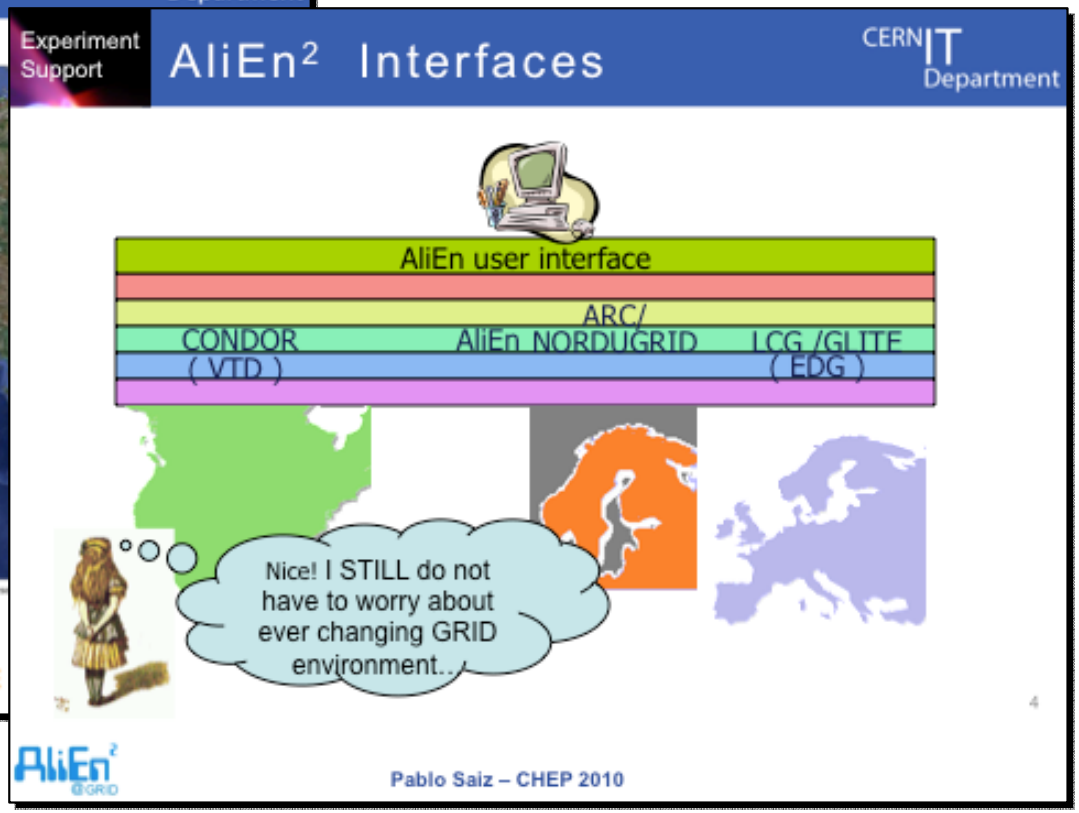
- With HTTP/WebDAV and NFS4.1, DPM provides standard based solutions for all its use cases
- Benefits exist for both clients and system administrators (and even developers)
- DPM will continue its work on improving the status of grid data storage and access

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

AliEn2

- "AliEn2: The ALICE Grid Framework", Steffen Schreiner





CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CEs and Workflow

CERN IT Department

- ARC-CE, CREAM
  - Well established systems
  - Gradual improvements
  - Integration in common environment (WLCG & EMI)
- CREAM – rollout takes lots of time
  - October 2008 – first production release
  - Today > 160 instances, but still not clear when the transition is complete
- WMS
  - Working towards better support for Pilots

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Security

CERN IT Department

- "Status and Challenges of Security in Distributed Computing", Stefan Lueders
  - Patching
  - SSH attacks
  - Virtualization
  - Complexity
  - Commercial Clouds

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Clouds (1/3)

CERN IT Department

- "CernVM CoPilot: (...)", Artem Harutyunyan
  - framework for the execution of LHC experiments' pilot jobs on a different computing resources, e.g.:
    - enterprise computing clouds (e.g. Amazon EC2),
    - scientific computing clouds (e.g. Nimbus)
    - volunteer computing clouds (Boinc)
      - CHEP talk: "Volunteer Clouds and Citizen Cyberscience for LHC Physics", Artem Harutyunyan

citizen cyberscience centre

Clouds (2/3)

- STAR actively uses Cloud infrastructure, Jerome Lauret (BNL)
  - Aggregating usage of grid and cloud resources
  - Test results are promising but there is still lots of work to be done
- H1 Collaboration Data Preservation Model, Bogdan Lobodzinski (DESY)
  - Tests with Cloud Computing Model (Eucalyptus 2.0) and distributed Petabyte File System
  - The concepts look promising but not ready for production mode

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Clouds (3/3)

CERN IT Department

- "StratusLab: Cloud-like Resource Delivery for Production Grids", Michel Jouvin (IN2P3)
  - 2 year project, 6 partners, 3.3MEuro
  - provide coherent, open-source private cloud distribution for grid and cluster computing
  - first release in November
- "Integration of Cloud, Grid and Local Cluster Resources with DIRAC", Tom Fifield
  - Belle ran 1/3 of their MC production on EC2

Summary

- Grid Middleware – slow evolution
  - Standard technologies (e.g. messaging)
  - Standard protocols
- Clouds used in production (but don't replace grids)
  - Technology developed for grid scheduling is used to link grids and clouds
- Data Management is a complex problem

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Clouds were everywhere...

CERN IT Department

Grid and Cloud

- Grid/Cloud
- Grid/Cloud
- Grid/Cloud
- Grid /Cloud
- Evolution of
- Global usage
- Experiment-

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Distributed Processing and Analysis

Track Summary

Giuseppe Lo Presti
(CERN, IT-DSS)

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Distributed Processing & Analysis

CERN IT Department

- Statistics
  - 48 oral presentations, ~20 posters
  - Wide variety of topics
- Main Areas
  - Experiments' status reports
    - LHC and others
  - "Hot" topics
    - E.g. Data Preservation, use of WebDAV, Parallel MC

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Experiments' status reports

CERN IT Department

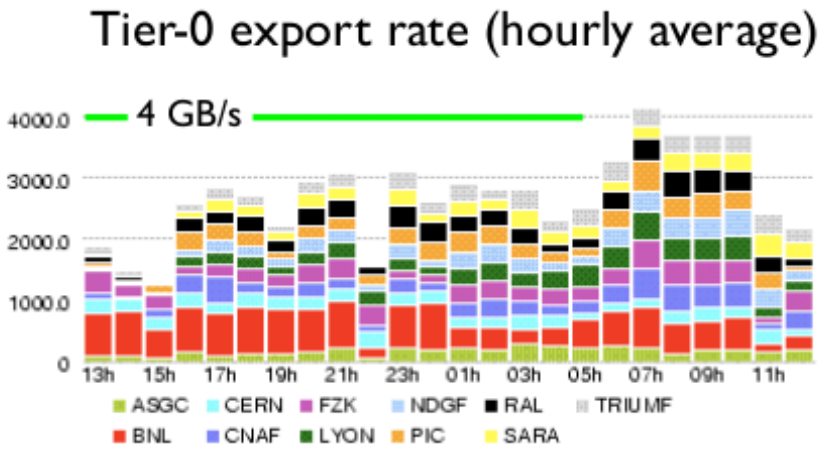
- Many contributions about the first year of physics with LHC
  - Overall message: smooth operations, lots of positive results
  - However, LHC not yet delivering as much beam time as expected
  - However, some paradigms are changing
- Some contributions from non-LHC experiments, with longer experience on data analysis

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

ATLAS [1-2-108, 41-5-296]

CERN IT Department

- Focus on Data Management operations experience
  - Exceeding targets: need to throttle export as reaching the physical limit for the Tier-0
  - Data distribution slightly different from expected
    - This is the 'first' data at all



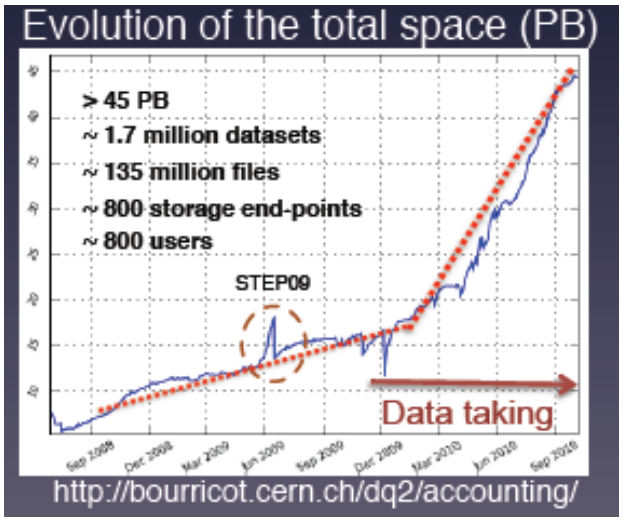
- Distr. Data Management (DDM) update
  - System to enforce ATLAS computing model
    - Based on Oracle backend, Apache frontend, CLI for data movements that is dataset oriented
  - Features include tracer service to monitor dataset usage over time
  - Consolidation, with an eye on future storage evolutions

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CMS [1-3-176, 42-2-202]

CERN IT Department

- Focus on Data Management experience
  - Performances (again) exceeding targets
    - But resource utilization not yet at target level
  - Sites' readiness defined as an AND of many tests, and monitored over time
    - E.g. Jobs success rate. Typical figure: 80% of the jobs successfully complete at first round, almost all after some automatic retries
  - Data movements across all sites: commissioning of the T2-T2 full mesh almost completed
  - Analysis: 800 unique users/day and counting
- Overall smooth operations, but 'stay tuned'

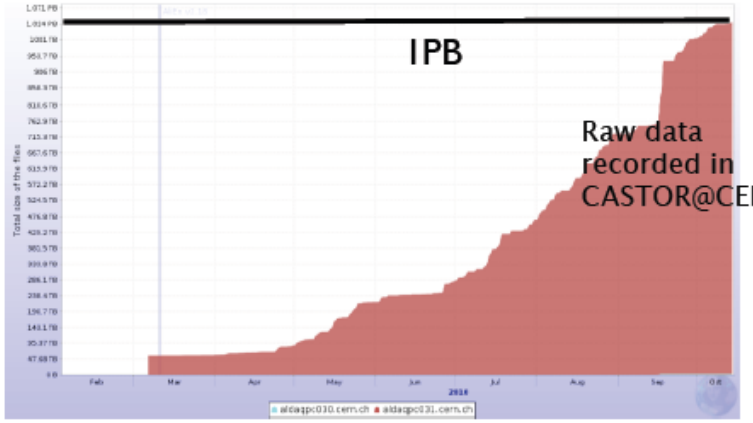


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

ALICE

CERN IT Department

- Experience with AliEn, targeting simplification of operations at sites
  - Data movements done via xroot 3rd party copy
  - Getting ready for the Heavy-Ion run...
  - Chaotic analysis run in the whole Grid with high stability
    - Foreseeing no differences between T1 and T2 sites
  - Grid operations is now routine



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

LHCb

CERN IT Department

- Computing experience
  - Small (~35 kB) event size in theory, yet high trigger rate
    - Reconstruction also at T1s
  - 2010 run: higher pile-up (and larger events) than expected
    - Adapted the computing model, yet suffering from the data access being the main weakness of the Grid
  - Analysis: very little at T2s
    - Moving to a concept of 'Analysis centre' and 'Reconstruction centre' (not necessarily matching T2s and T1s...)
- Full reprocessing to start in Nov 2010

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

PHENIX [25-5-491]

CERN IT Department

- Detector for Au-Au collisions at BNL
- Typical issue: inefficient access to tape when running a cycle of analysis
  - 3 weeks, ~200k files retrieved from tape
- Addressed by adopting the 'analysis taxi' paradigm:
  - Closed access to general users
    - Only taxi drivers (i.e. production) can run
  - Users are allowed to jump in once a week provided they run resource constrained (and valgrind-validated!) jobs

CDF [31-2-280]

CERN IT Department

- The Collider Detector at Fermilab
  - ppbar collisions, 8 fb^{-1} of data on tape
- Moved from a dedicated computing facility to using the WLCG
  - Developed an extra middleware layer to enable existing software to submit jobs with gLite
- But jobs are composed by subparts
  - Time for completion often very large
  - Partially addressed by resubmitting jobs that don't complete after 1h

Gamma-Ray Space Telescope

CERN IT Department

- Key requirement from NASA: results need to be public within 24 hours
- A ‘pipeline system’ software is in place for the entire analysis process
  - Includes [web] interfaces to monitor the process, the data catalogue, etc.
  - Entirely based on standard technologies
    - Oracle, Java Stored Procedures, Jython, ...
  - Currently being evaluated for upcoming projects like the next version of the Hubble Space Telescope

[48-5-502]

Hot Topics: WebDAV [19-4-353]

CERN IT Department

- Experiences at PIC, motivated by moving towards standard protocols
  - Today mostly dcap/dCache
- Promising performance tests on bulk transfers
  - Little throughput overhead w.r.t. native xroot (xrdcp) on large (1 GB) files, and better throughput on small (2 MB) files
- Not so performing on random access
  - Envisaging more investigations with NFSv4.1

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Hot Topics: Data Preservation

CERN IT Department

- HERA at DESY [41-2-357]
  - Proposal to develop a storage solution at DESY to target data preservation, including periodical recall and reprocessing
  - Data is in ROOT format
- However,
  - Acknowledged that typically no budget is allocated specifically for this task
  - Documentation preservation is as critical

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Hot Topics: Parallel MC [31-3-272]

CERN IT Department

- A framework to run MPI-like parallel applications on the Grid
  - Based on Ganga, called GaMPI
- Some promising preliminary results
  - It compares well with pure MPI
  - A number of issues to solve, typically on submission of parallel jobs

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Hot Topics: MyOSG [42-3-316]

CERN IT Department

- The problem: there is a variety of operational tools for the grids
  - developed by different teams with different spirits
- The goal: have a common visual web-based interface
  - In particular homogenizing the interaction with OSG and gLite grids
  - Output provided also in ‘novel’ popular formats
    - e.g. iGoogle, content for mobile devices
  - However, risk that this is just yet another interface...

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Other talks

- xroot @ GridKa [19-3-002]
  - SE for ALICE based on GPFS over SAN storage
  - Main advantage: hardware failures don't result in service unavailability
  - No performance tests
  - Second largest SE for ALICE with 1.3 PB
- Proxy caches with xroot [48-3-305]
  - Report about the usage of proxy caches for analysis jobs for ATLAS and CMS
  - Preliminary tests run with BNL suggest remote access looks feasible once cache gets warmed up
  - Demonstrator to deliver some results by Jan 2011

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Outlook

CERN IT Department

- Lots of positive results
- Computing models are successfully being applied by experiments
- Still room for some developments and consolidation
- Quoting Daniele's conclusions [Experience with CMS computing]: *'stay tuned'*

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

And a final summary …

CERN IT Department

Entrance, Academia Sinica Guest House

Photo by G. Lo Presti