Summary of Track 6 Facilities, Production Infrastructures, Networking, Collaborative Tools

Track 6 conveners, presented by
Helge Meinhard / CERN-IT
18 October 2013

Track 6 Conveners

- Brian Bockelman
- Ian Collier
- Alessandro De Salvo
- Maria Girone
- Steven Goldfarb
- Burt Holzman
- Helge Meinhard
- Ray Pasetes
- Wim Heubers (LOC)

Usual disclaimer

- Credit to presenters and session chairs
- All errors, omissions, … are mine
- Posters not covered – sorry

Statistics

- 83 abstracts; 81 accepted, 2 withdrawn
- 28 oral presentations, 53 posters

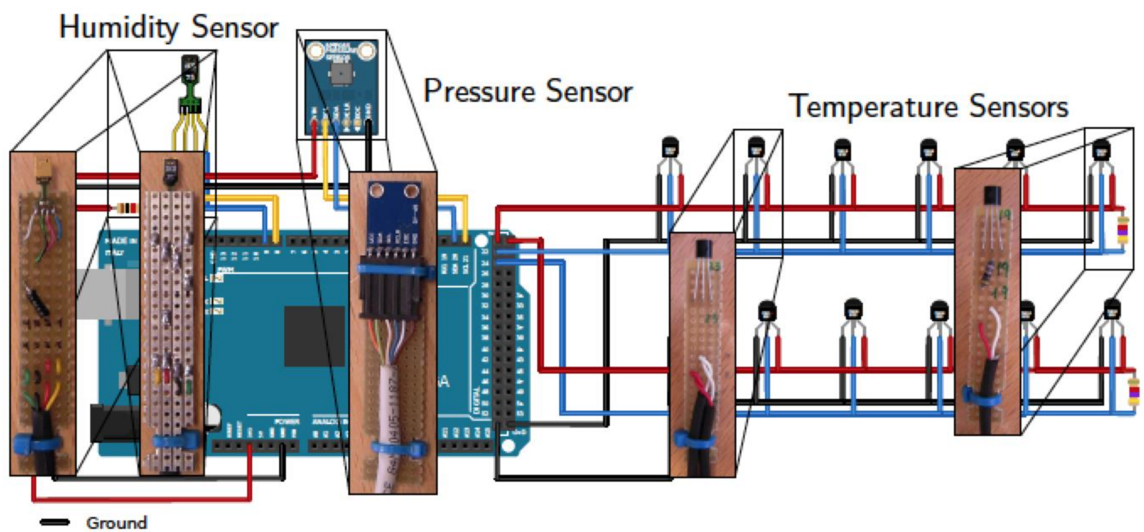
Topic	No contributions
Facilities	6
Production infrastructures	50
Networking	15
Collaborative tools	10

Facilities

Arduino and Nagios integration for monitoring (Victor Fernandez, U Santiago de Compostela)

- Aim: address monitoring needs of their compute farm
- Chose home-made integration of Arduino and Nagios for cost reasons

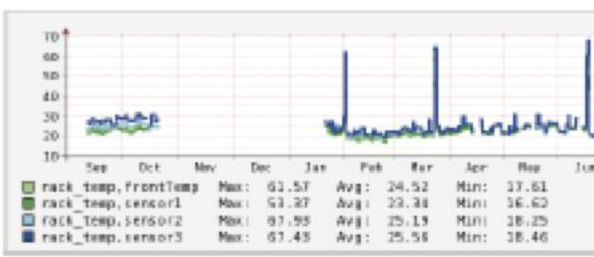
Fernandez, U Santiago de Compostela



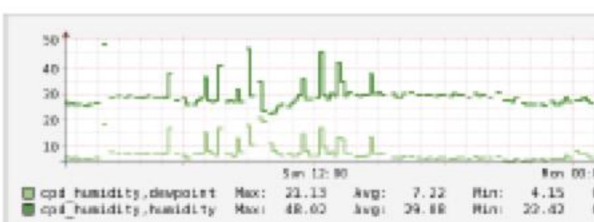
1	Arduino Mega 2560	46€
1	GSM Shield	97€
24	DS18B20 (Temperature)	198€
1	Sensirion SHT75 (Humidity)	40€
1	Bosch BMP085 (Pressure)	23€
2	Prototype boards with continuous strips	20€
	Total	424€

Week

Year



Daily Humidity



SynapSense wireless environmental monitoring system of RACF at BNL (Alexandr Zaytsev, BNL)

- Environmental monitoring needed that is easy to install – no complex wiring
- SynapSense: wireless sensors

Zaytsev, BNL

In the present configuration the system has 150+ base stations provided with 520+ low systematic temperature/humidity/pressure sensors reporting to the central servers every 5 minutes (0.27M readings per day)

The integral cost of the system is not exceeding the cost of 2 racks of equipment typical for RACF Linux farms

- There is a potential of extending the SynapSense™ monitoring system to include power consumption monitoring for all the CRAC and PDU devices in the facility and SynapSense™ ActiveControl™ features, thus providing real time estimated of the PUE of the data center and the means to optimize it

Operating dedicated data centres – is it cost-effective? (Tony Wong, BNL)

- Cost comparison of BNL facilities with commercial cloud offerings (EC2, GCE)

Wong, BNL

	USATLAS	RHIC
Server	$228/yr	$277/yr
Network	$28/yr	$26/yr
Software	$3/yr	$3/yr
Staff	$34/yr	$34/yr
Electrical	$12/yr	$16/yr
Space	$27/yr	$13/yr
Total	$332/yr ($0.038/hr)	$369/yr ($0.042/hr)

Includes 2009-2013 data

BNL-imposed overhead included

Amortize server and network over 4 or 6 (USATLAS/RHIC) years and use only physical cores

RACF Compute Cluster staffed by 4 FTE ($200k/FTE)

About 25-31% contribution from other-than-server

- Cost of computing/core at dedicated data centers compare favorably with cloud costs
  - $0.04/hr (RACF) vs. $0.12/hr (EC2)
  - Near-term trends
    - Hardware ↘
    - Infrastructure ↗
    - Staff ⇒
    - Data duplication ↗
- Data duplication requirements will raise costs and complexity – not a free ride

Hardware at remote hosting centre (Olof Barring, CERN)

- Wigner research centre in Hungary won open call for tender for extending CERN's computer centre capacity
- Issues around scalability and non-availability of physical access addressed

Barring, CERN

Preparation

CERN IT Department

- Review main processes
  - **Delivery requirements**
  - Hardware handling
  - Stock management
  - Inventory
  - **Network registration**
  - **Burn-in**
  - Production deployment
  - Remote console
  - **Onsite maintenance**

Barring, CERN

Conclusions

CERN IT Department

- Remote co-location is our way to scale beyond local power limitation
- Wigner contract awarded following competitive tender
- Preparation had positive impact also on local operation
  - Design workflows and automation with remote operation in mind
- Production service is up and running
  - But work still required to finalise operational procedures
- Started preparations for large scale (90%) deployment of new deliveries in 2014-15

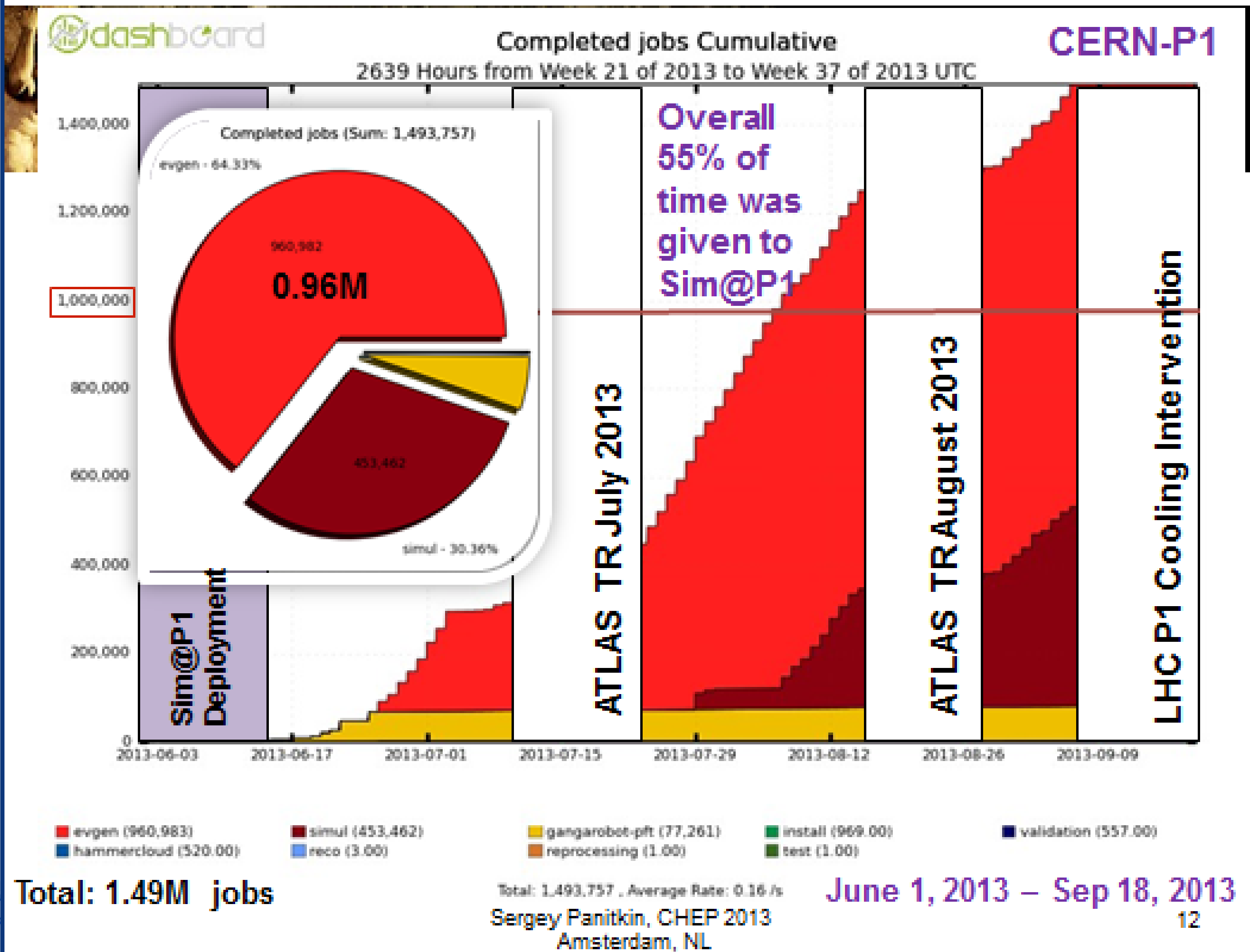
CHEP '1[illegible] AMSTERDA[illegible]

Production Infrastructures

ATLAS cloud computing R&D project (Randy Sobie, U Victoria)

- Private / academic clouds – HLT farm
- Public / hybrid clouds: Amazon EC2, Google compute engine
- CloudScheduler as “middleware” between HTCondor and cloud

Sobie, U Victoria



Fabric management (r)evolution at CERN (Gavin McCance, CERN)

- Agile Infrastructure project addressing
  - virtual server provisioning
  - configuration
  - monitoring

McCance, CERN

Agile Infrastructure "stack"

- Our current stack has been stable for one year now
  - See plenary talk at last CHEP (Tim Bell et al)
- Virtual server provisioning
  - Cloud "operating system": **OpenStack** -> (Belmiro, next)
- Configuration management
  - **Puppet** + ecosystem as configuration management system
  - **Foreman** as machine inventory tool and dashboard
- Monitoring improvements
  - **Flume** + **Elasticsearch** + **Kibana** -> (Pedro, next++)

McCance, CERN

Community collaboration

- Traditionally one of HEPs strong points
- There's a large existing Puppet community with a good model - we can join it and open-source our modules
- New HEPiX working group being formed now
  - Engage with existing Puppet community
  - Advice on best practices
  - Common modules for HEP/Grid-specific software
  - https://twiki.cern.ch/twiki/bin/view/HEPIX/ConfigManagement
  - https://lists.desy.de/sympa/info/hepix-config-wg

McCance, CERN

Summary

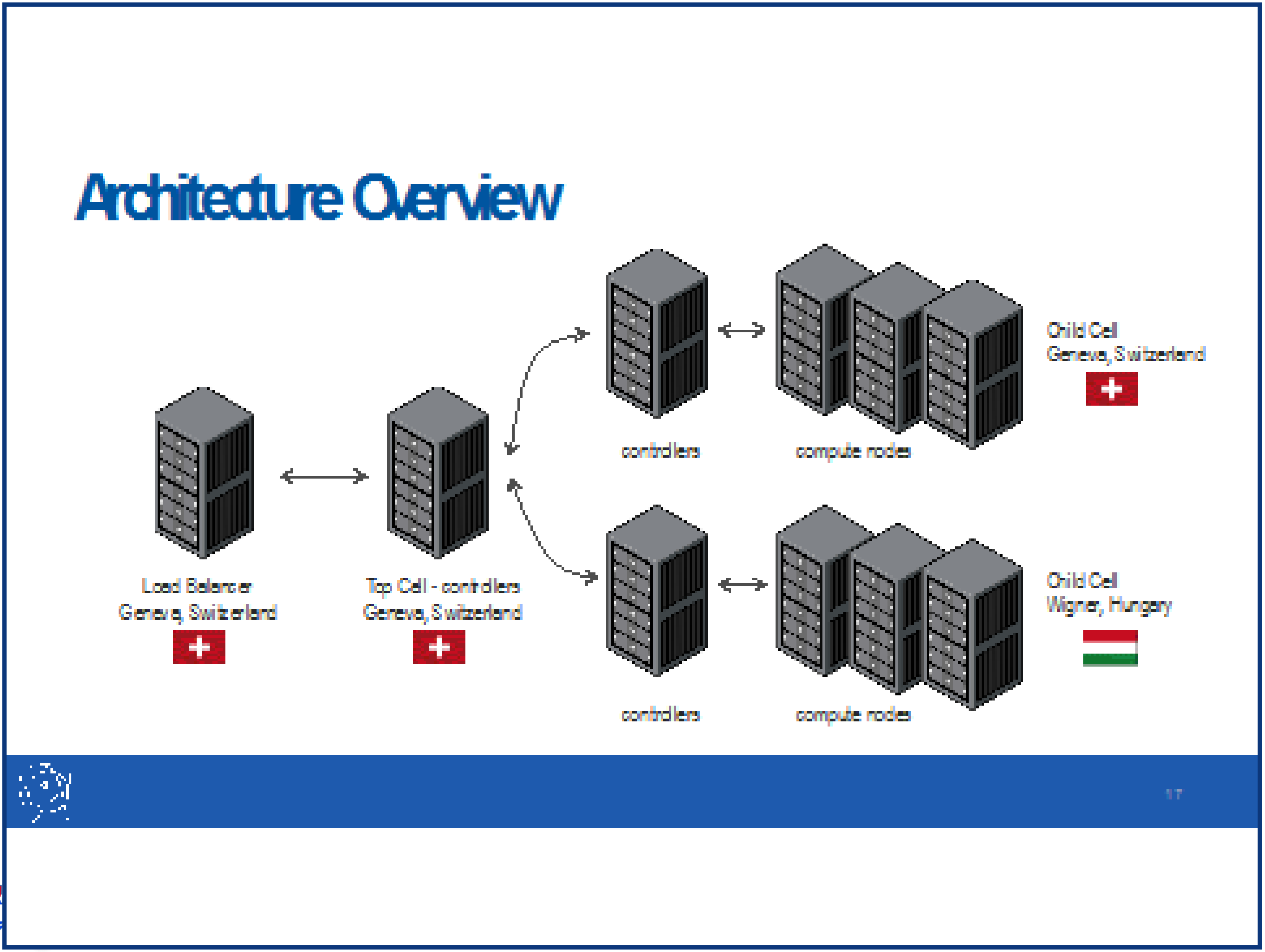
- The Puppet / Foreman / Git / Openstack model is working well for us
  - 4000 hosts in production, migration ongoing
- Key technical challenges are scaling and integration which are under control
- Main challenge now is people and process
  - How to maximise the utility of the tools
- The HEP and Puppet communities are both strong and we can benefit if we join them together

https://twiki.cern.ch/twiki/bin/view/HEPIX/ConfigManagement
http://github.com/cernops

Production large-scale cloud infrastructure experience at CERN (Belmiro Moreira, CERN)

- Motivation: Improve
  - operational efficiency
  - resource efficiency
  - responsiveness

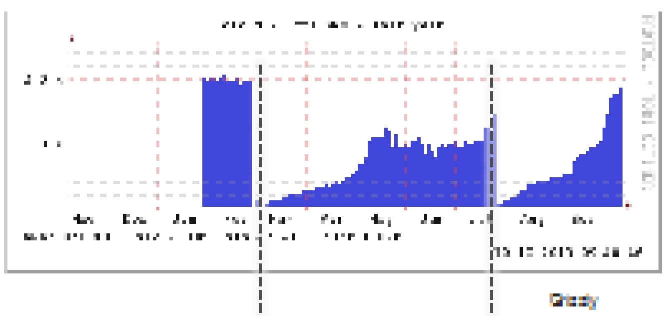
Moreira, CERN



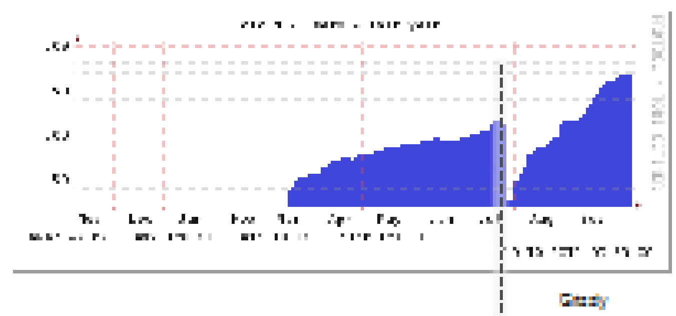
Moreira, CERN

CERN Cloud Infrastructure adoption

Number of VMs



Number of Users



Cell	nodes	Cores	mem (GB)	disk (TB)	VMs
Geneva	375	10976	21662	632	1438
Wigner	281	9312	18296	491	674
Total	656	20288	39958	1123	2112

(13/10/2013)

Moreira, CERN

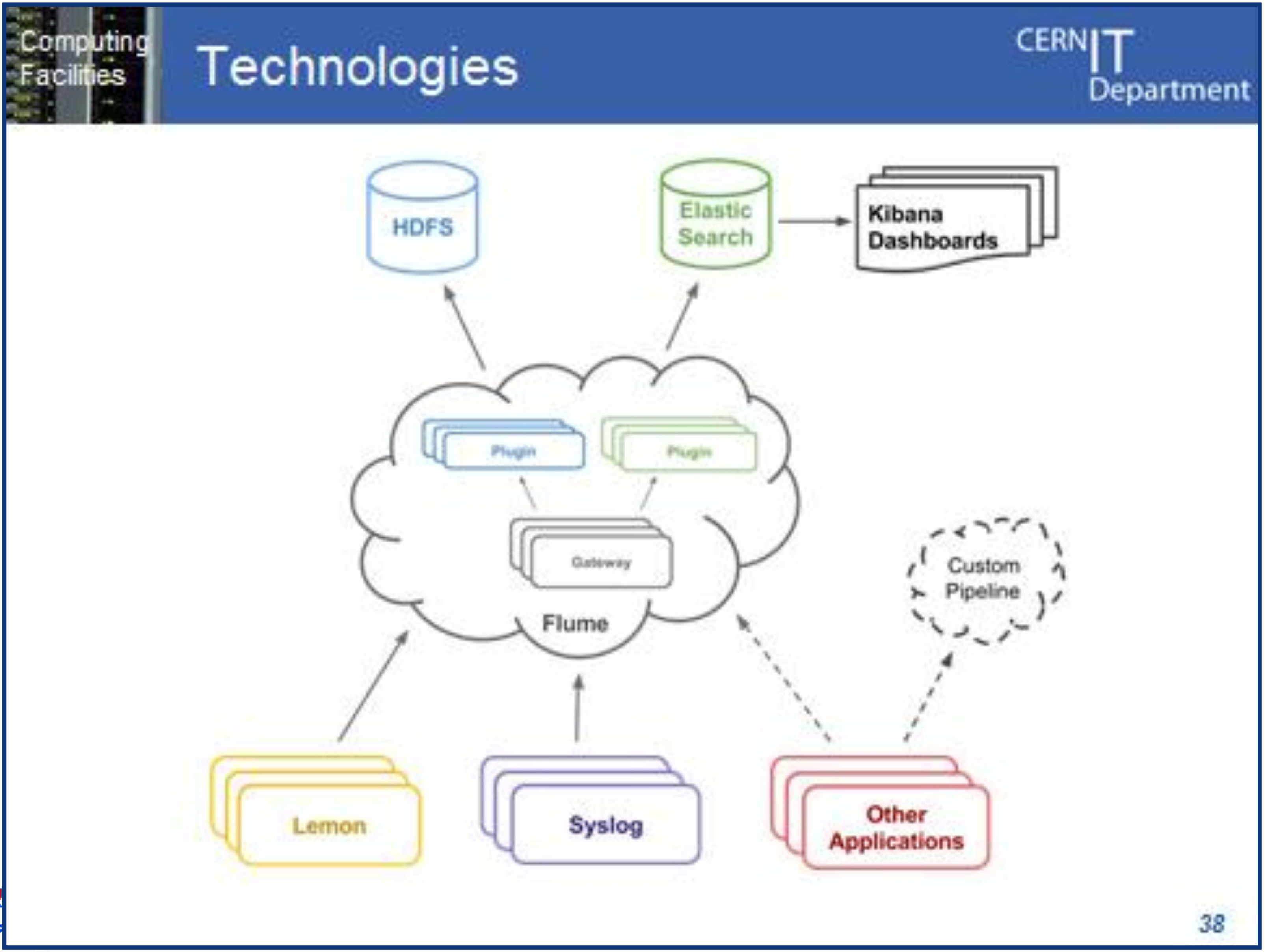
Next Challenges

- Growing the infrastructure
  - +100 compute nodes per week
  - 15000 servers – more than 300000 cores
- Migration from Grizzly to Havana
- nova-network deprecation
- Kerberos, X.509 user certificate authentication
- Keystone Domains

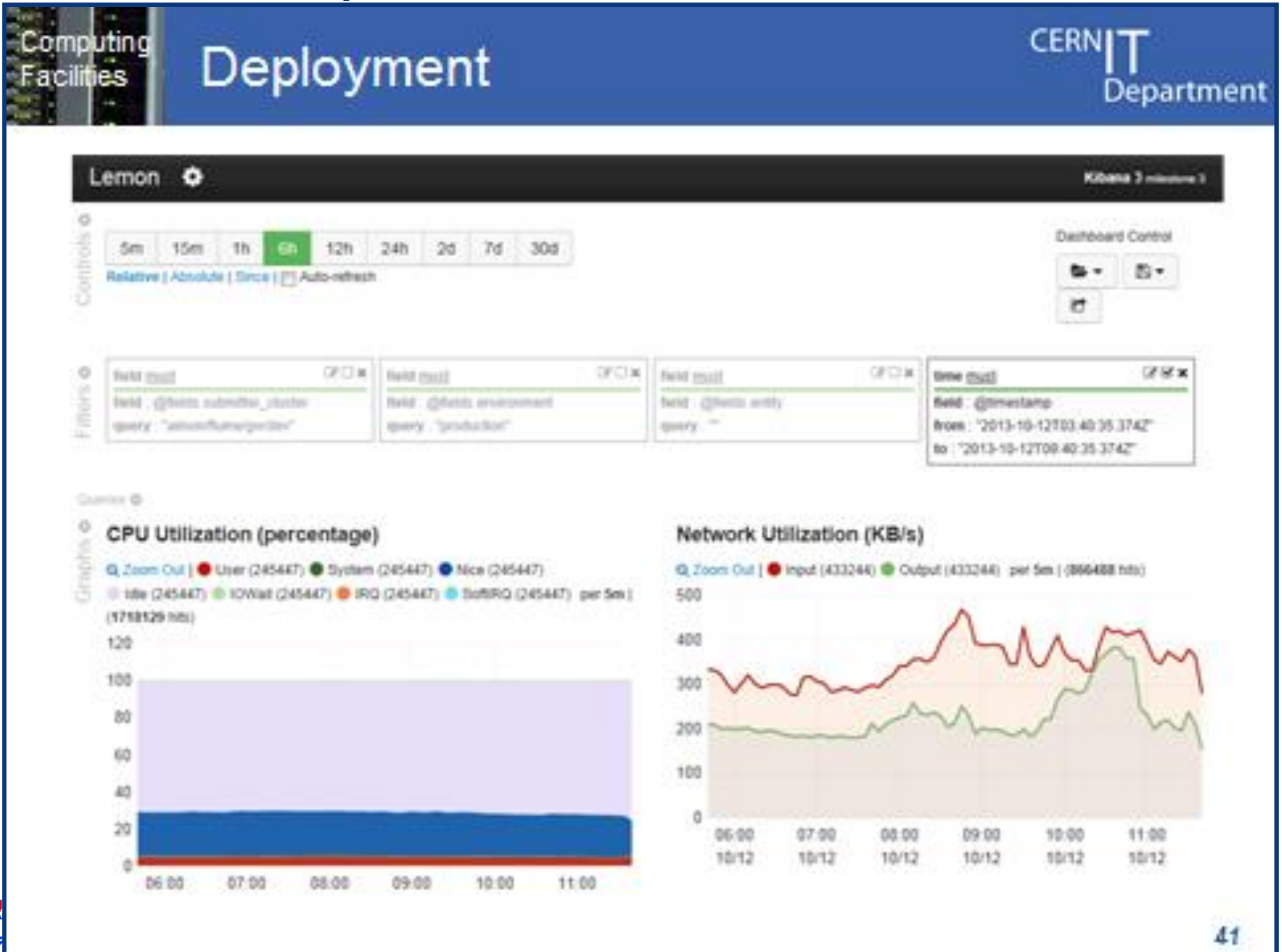
Agile Infrastructure monitoring (Pedro Andrade, CERN)

- Motivation
  - Several independent monitoring activities in CERN IT
  - Combination of data from different groups necessary
  - Understanding performance became more important
  - Move to a virtualised dynamic infrastructure
- Challenges
  - Implement a shared architecture and common tool-chain
  - Delivered under a common collaborative effort

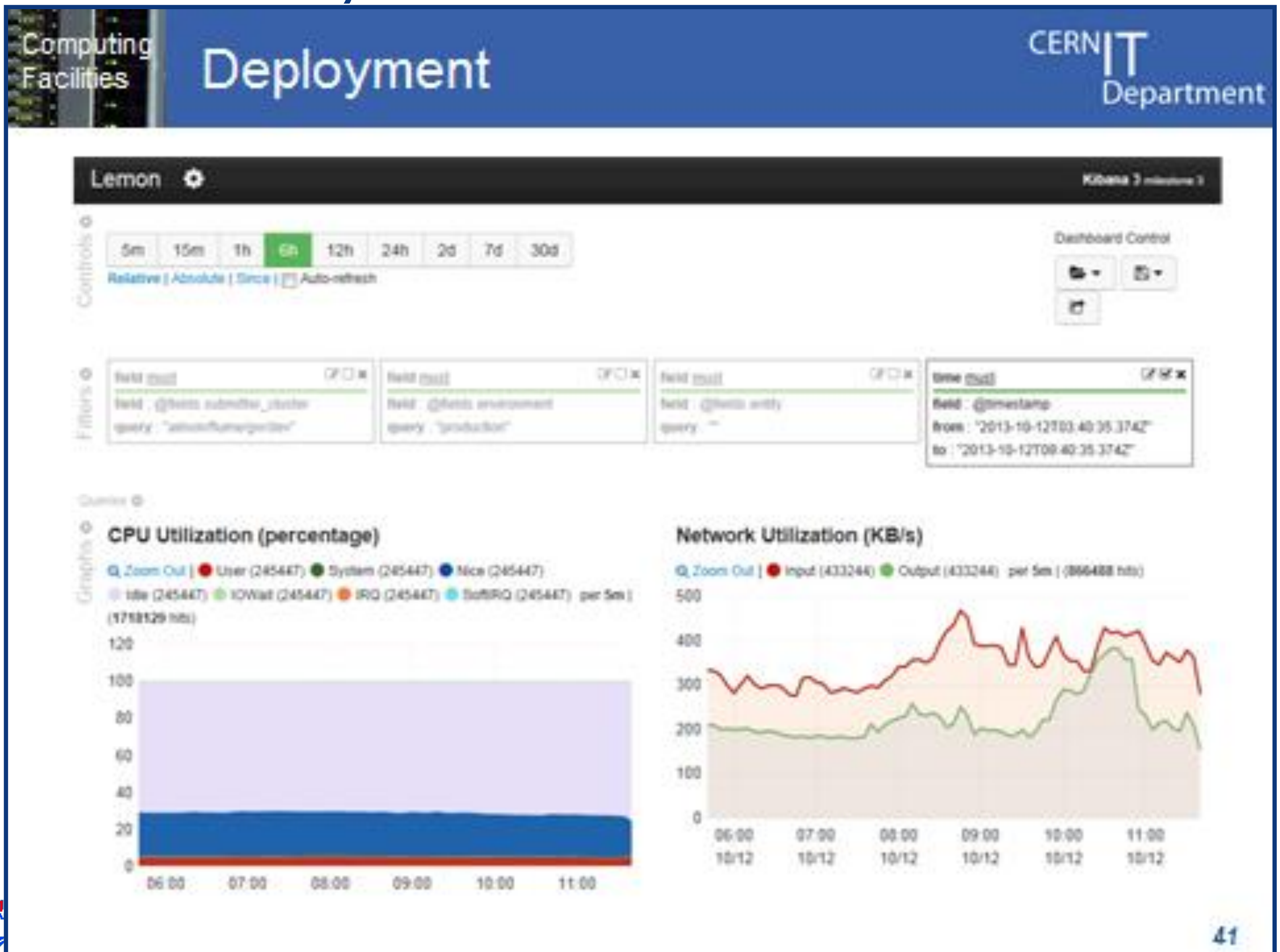
Andrade, CERN



Andrade, CERN

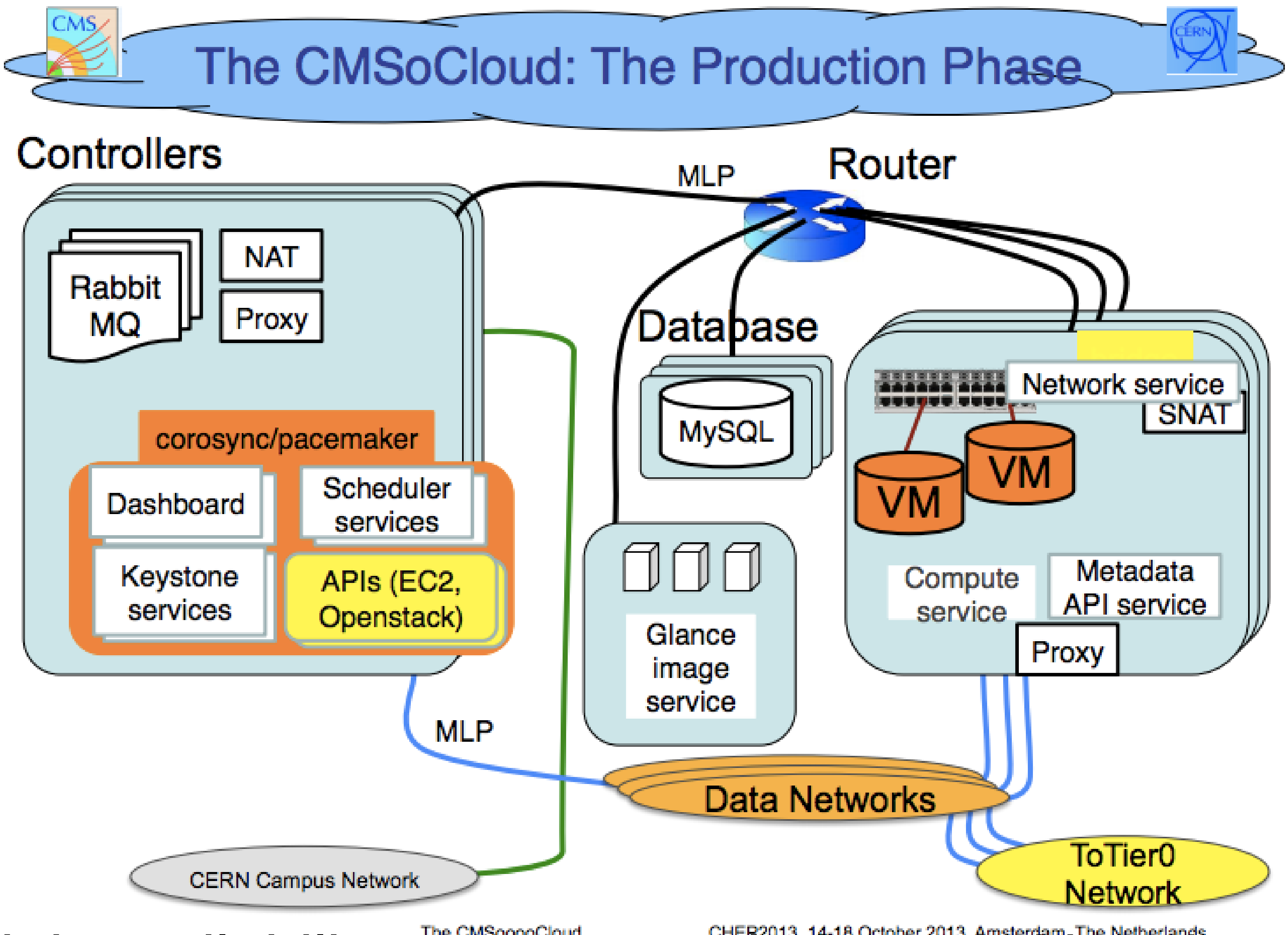


Andrade, CERN



The CMS openstack, opportunate, overlay, online-cluster cloud (Jose Antonio Coarasa, CERN)

- Idea: Reuse CMS Data Acquistion System as an opportunistic Open-Stack based cloud.
- A cloud of opportunity - when CMS is not taking data, give computing power of HLT to Offline. Online must be able to "take back" computing resources quickly.
- Overlays on top of existing cluster; OpenStack must deal with existing complex network configuration.
- Cloud has been running since January 2013.
- Has run up to 6,000 jobs at a time; a significant resource in CMS Offline.



High availability setup; complex networking due to required Online security!

Opportunistic resource usage in CMS (Peter Kreuzer, RWTH Aachen)

- CMS has a relatively flat funding budget for hardware.
  - CMS can keep its hardware fully occupied. Investment in people greater than investment in computing hardware. Must keep people productive!
  - Goal: Allow people to dynamically integrate shared resources.
- Three types of resource access considered
- Non-CMS grid site, opportunistic or Allocation-based cluster (no grid interface), or Virtualization-based resources (OpenStack, EC2).
- Operational issues - how does CMS integrate temporary resources into a system designed for permanent resources?
  - Either put all resources into a "fake" site or dedicated site for very large opportunistic resources.
  - Testing already done at large-scale; sustainable operations is the current challenge.

Operating the World-wide LHC computing grid (Andrea Sciaba, CERN)

- Dedicated effort as a follow-up from Technical Evolution groups in 2011/2012
- Activity resulted in a series of recommendations to be followed up by a new, dedicated coordination body

Sciaba, CERN

The WLCG operations coordination working group

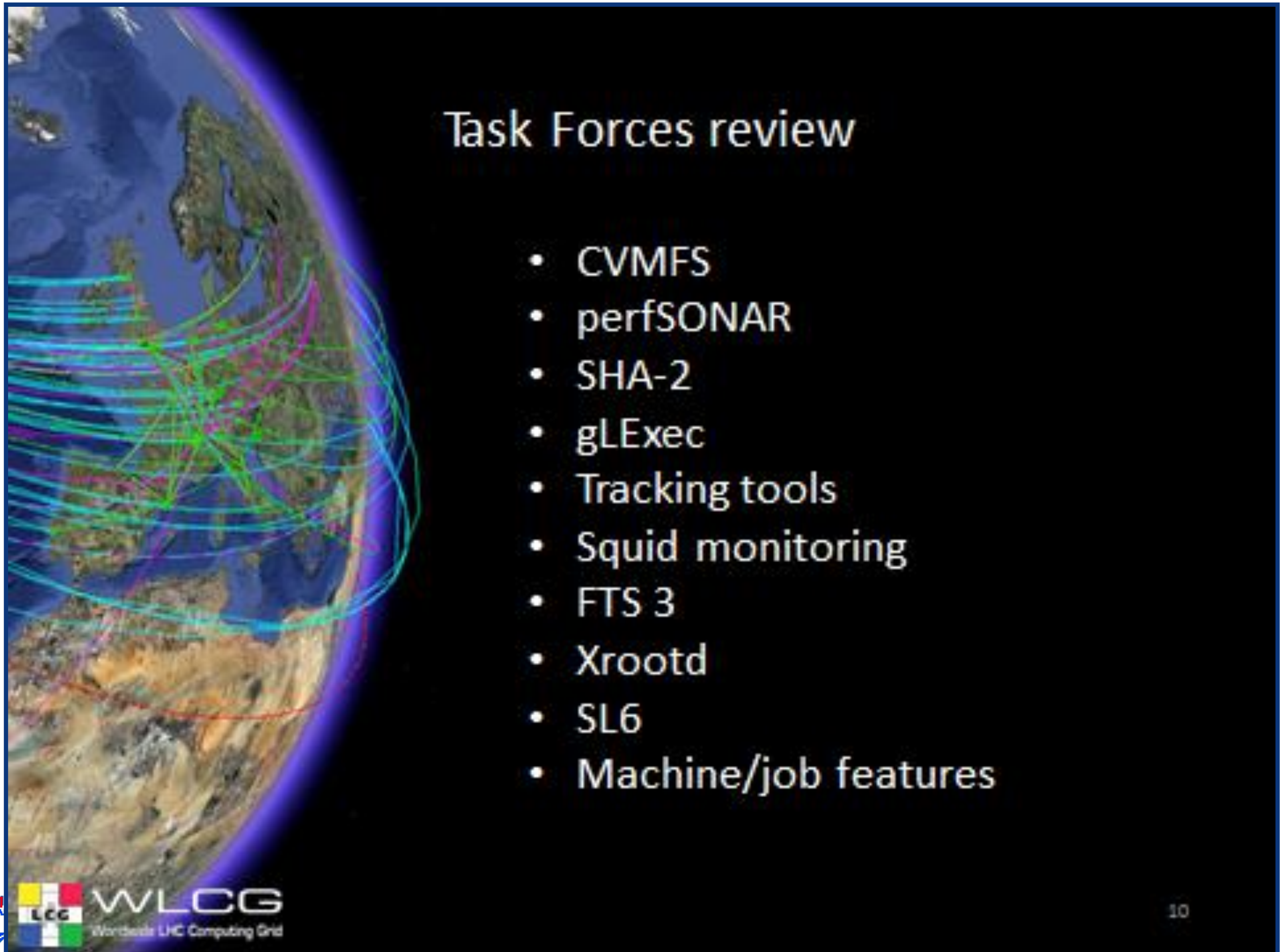
- Established in October 2012
- Acts as core operations and deployment coordination team
  - Manages operational issues, service deployment in synergy with EGI, OSG, NorduGrid
  - Discusses experiments plans and needs
  - Defines actions and work plans
  - Forms time-limited task forces on specific issues
  - Ensures communication among experiments, sites, projects
- All stakeholders are represented
  - LHC experiments, site regions, Tier-1's, Grid projects
  - Fortnightly meetings, quarterly planning meetings
  - Largely based on voluntary effort from the entire WLCG community

M. Girone, Operations Coordination Team, 11/07/2012, WLCG GDB

WLCG
Worldwide LHC Computing Grid

28

Sciaba, CERN



Task Forces review
• CVMFS
• perfSONAR
• SHA-2
• gLExec
• Tracking tools
• Squid monitoring
• FTS 3
• Xrootd
• SL6
• Machine/job features
WLCG
Worldwide LHC Computing Grid
10

Testing as a service with HammerCloud (Ramon Medrano Llamas, CERN)

- Large-scale flexible grid testing increasingly important and popular
- 50 M jobs / year
- Requires flexible infrastructure for rapid deployment

Medrano Llamas, CERN

New use cases

- Stress testing of sites
- Functional testing of sites
- AFT/PFT testing suite

} 12,000 test/year

- *Benchmarking testing* ***NEW!***
- *Cloud resource validation* ***NEW!***
- *Athena nightly build system* ***NEW!***
- *XRootD federation (FAX)* ***NEW!***
- *ROOT I/O and WAN tests* ***NEW!***

Medrano Llamas, CERN

Requirements

1. Elastic infrastructure (OpenStack)
2. Cloud orchestrator
3. Code sanitation (Gerrit)
4. Configuration Management (Puppet)
5. Deployment procedures

Medrano Llamas, CERN

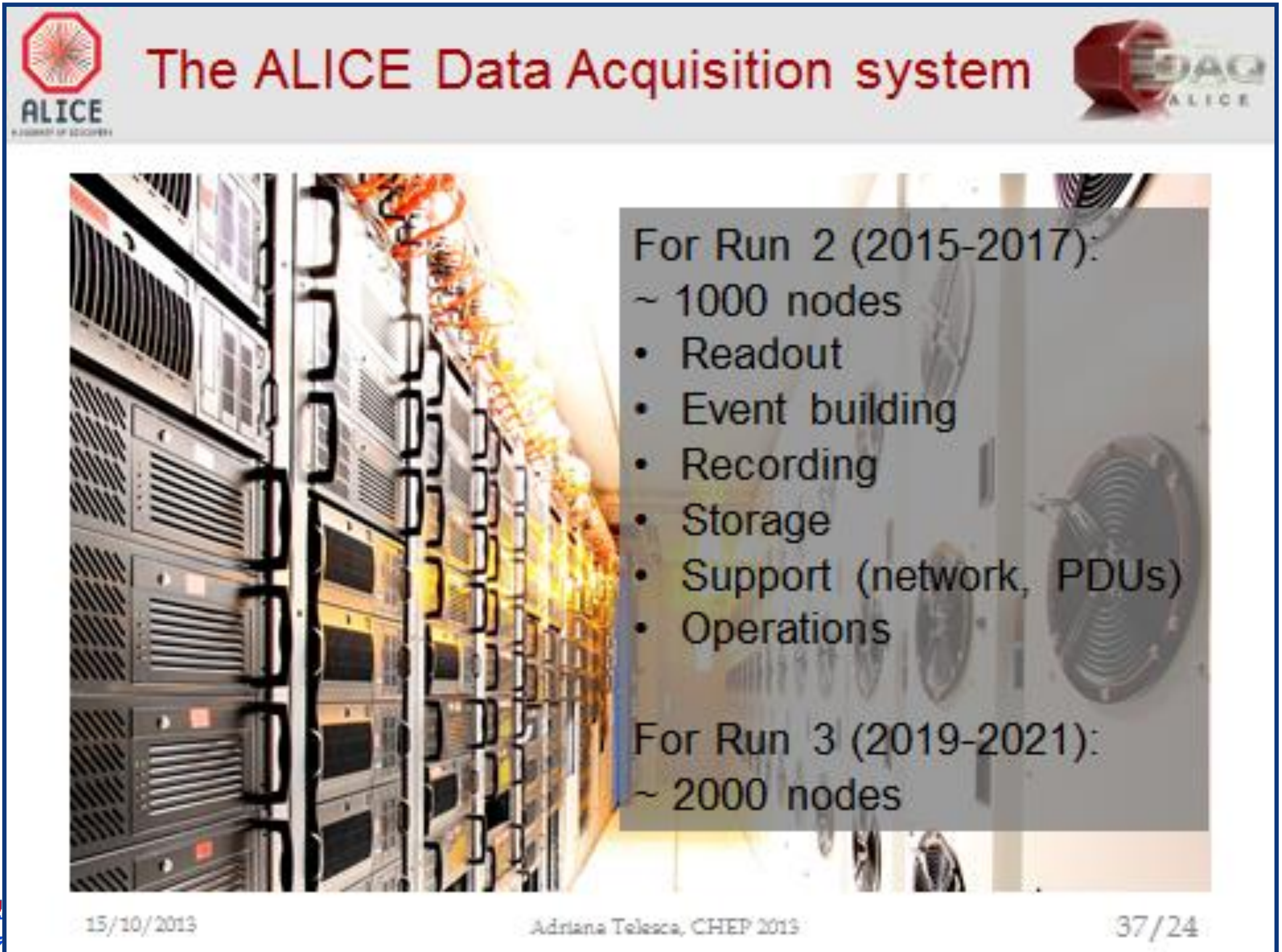
Dynamic testing scheduling

1. Test is requested (user, API, cron)
2. Creation of the VMs on demand
   - Isolation
   - Reliability
   - Elasticity
3. Configuration and startup of the VM
4. …test runs…
5. Cleanup and destroy

Performance monitoring of ALICE DAQ system with Zabbix (Adriana Telesca, CERN)

- Growing DAQ farm requires more flexible, powerful system monitoring
- Comprehensive study of candidate systems has resulted in choosing Zabbix

Telesca, CERN



Telesca, CERN

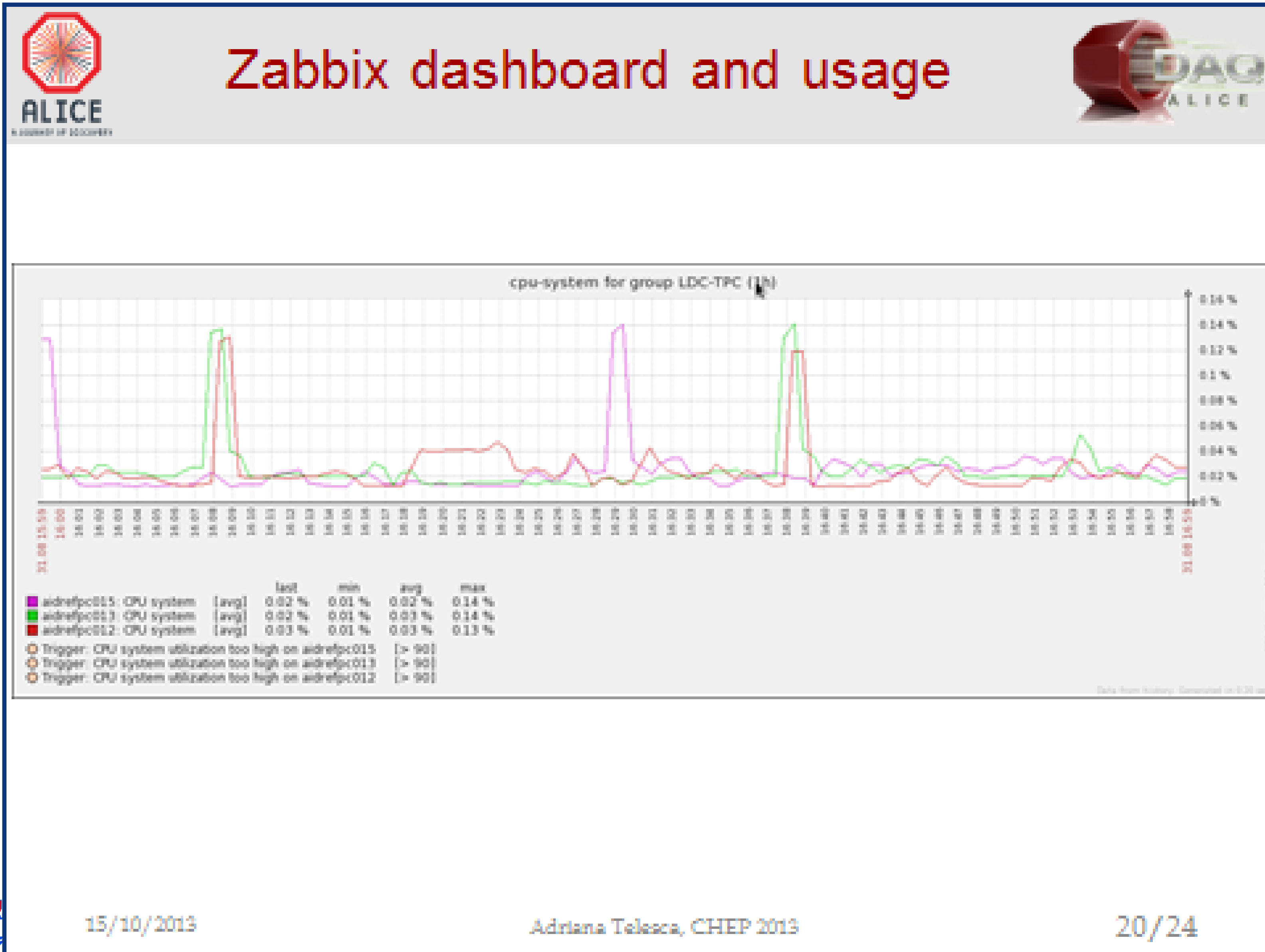
Tools comparison

Name	SNMP	Communit y	Granularity	Auto Discovery	Free	Total
Icinga	2	2	1 - 1 minute /metric	2	1	**12**
Cacti	2	2	1 - 1 minute / metric	1	1	**12**
Zenoss	1	1	1- 1 minute /collector	2	1	**11**
Zabbix	2	2	2 - No limit /metric	2	1	**16**
Splunk	2	2	2 - No limit / metric	2	0	**15**
MonALISA	2	1	1 - 1 minute /metric	2	1	**14**

0-1 Absent-Present

0-1-2 Absent - Present but not good - Good

Telesca, CERN



Beyond core count: new mainstream computing platforms for HEP workloads (Pawel Szostek, CERN)

- Improvements of performance and performance/watt by
  - Increasing core counts
  - Shrinking structure sizes
  - Introducing new microarchitectures

Szostek, CERN

Hardware setup for tests

Intel server CPUs

- "Sandy Bridge" E5-2690
- "Ivy Bridge" E5-2695 v2
- 2 sockets: 16 and 24 cores
- ➢ Shrink from 32 to 22nm
- ➢ Same cache, lower TDP

Intel workstation CPUs

- "Ivy Bridge" E3-1265L v2
- "Haswell" E3-1285L v3
- Single socket: 4 cores, 8 threads
- ➢ New micro-architecture
  - ➢ AVX2
  - ➢ Wider core (4th ALU, 3rd AGU, 2nd branch prediction unit)

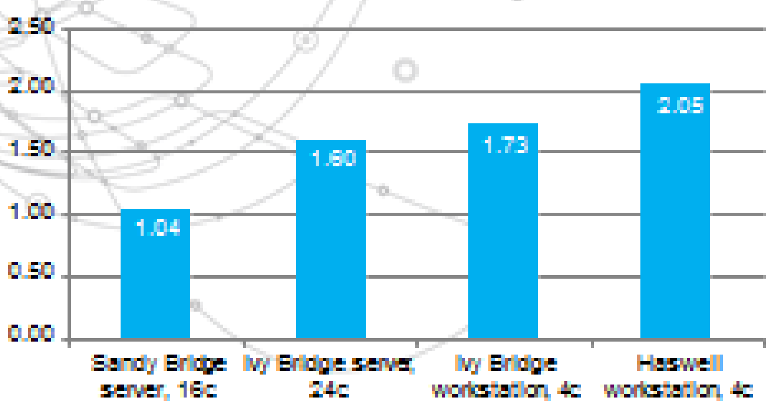
Szostek, CERN

CERNopenlab

HEPSPEC06 per Watt

	„Sandy Bridge" server E5-2690	„Ivy Bridge" server E5-2690 V2	„Ivy Bridge" workstation (E3-1265L V2)	„Haswell" workstation E3-1285L V3
HS06	381	463	94	115
Standard energy measurement	362	290	54	56
HS06 per Watt	1.04	1.60	1.73	2.06

HEPSPEC06/W (higher is better)



- From IVB to HSW: 20% improvement (incl. motherboard)
- From SNB-EP to IVB-EP: **54%**
- High values for desktops are due to manually optimized energy use: barebone systems

The effect of flashcache and bcache on I/O performance (Jason Alexander Smith, BNL)

- Flashcache, bcache: Linux kernel modules for block caching of disk data on fast devices (such as SSDs)
- Flashcache
  - Developed by Facebook in 2010
  - Not included in Linux kernel
- Bcache: different approach with similar goals
  - In Linux kernel as of 3.10
- Result: good for small records/files

Smith, BNL

Evaluation Hardware/Configuration (Cont.)

Software RAID0, Flashcache and Bcache Benchmarks
- Dell PowerEdge R620
- 2 8-core Xeon E5-2660@2.20 GHz CPUs (HT on: 32 logical cores total)
- 48 GB DDR3 1600 MHz RAM
- PERC H310 disk controller
- 64-bit Scientific Linux 6.4 (kernel 2.6.32-358.6.2.el6.x86_64, 3.11.1 for Bcache tests)
- 8 2.5" SATA hard drives in a software RAID0 configuration
  - Only 7 spindles used in the array for Flashcache and Bcache tests
  - Seagate ST9500620NS 2.5" drive
  - 500 GB, SATA 3.0 Gbps
  - 64 MB cache
  - 7200 RPM
  - Firmware release AA09

SSD TRIM ("discard" mount option) not enabled. EXT4 used in all tests

Tested both "clean" and "dirty" Flashcache and Bcache configurations
- Clean - no data written besides filesystem metadata before benchmark
- Dirty – benchmark run multiple times in succession before final test

Smith, BNL

Conclusions

- The single SSD tested provided excellent random I/O characteristics, particularly for small record sizes, but did not provide the performance of a multi-spindle software RAID0 configuration for larger record sizes
  - The software RAID0 configuration provided roughly double the random I/O performance compared to the SSD for large records and for parallel workloads
    - But it consisted of 8 times the number of drives
  - Single SSD random I/O performance was significantly better than a single SATA drive

- Flashcache and Bcache with an SSD cache generally augmented the I/O performance of a single SATA disk for files that fit within the cache
  - Generally true for both random and sequential I/O
  - Smaller gains, or performance losses, were typically seen when the cache was preloaded with dirty data during bonnie++ testing
  - Probably not suitable for scratch space utilization, since in this use case we're likely dealing with large files that are only written and/or read once
  - Would likely benefit database, webserver, or other applications where a set of relatively small files are repeatedly read/written

Challenging data and workload management in CMS computing with network-aware systems (Tony Wildish, Princeton)

- PhEDEx controls bulk data-flows in CMS.
  - Basic architecture is 10 years old. Retry algorithms are TCP-like (rapid backoff / gentle retries). No understanding of the underlying network activity.
  - Complex transfer mesh -- since it no longer follows the MONARC model, we no longer have an analytic model of CMS transfers. Why are datasets moved? Which movements are correlated?
- Working on long-term use cases and models for integrating network knowledge:
  - ANSE project working to integrate virtual network circuit control into PhEDEx. Explicitly control the networks.
  - Hope is that this will reduce latencies in PhEDEx.

Wildish, Princeton

Challenging CMS Computing with network-aware systems

Average rate last year	Production	Debug	Total
T0 -> T1	230 MB/sec	100 MB/sec	330 MB/sec
T2 -> T1	190	200	390
T1 -> T2	620	230	850
T2 -> T2	260	180	440
Total	1300	710	2010

Production instance is real data
Debug instance is for commissioning and link-tests
- 1/3 of total traffic is for knowledge of network state

Average rate ~ 2 GB/sec CMS-wide
- sustained over last 3 years.
- not b/w limited

Not currently bandwidth-limited, but preparing for the future!

Networking

Deployment of PerfSONAR-PS networking monitoring in WLCG
(Simone Campana, CERN)

- Introduction to PerfSONAR and PerfSONAR-PS
- Deployment plan for WLCG
- Status

Campana, CERN

perfSONAR and perfSONAR-PS

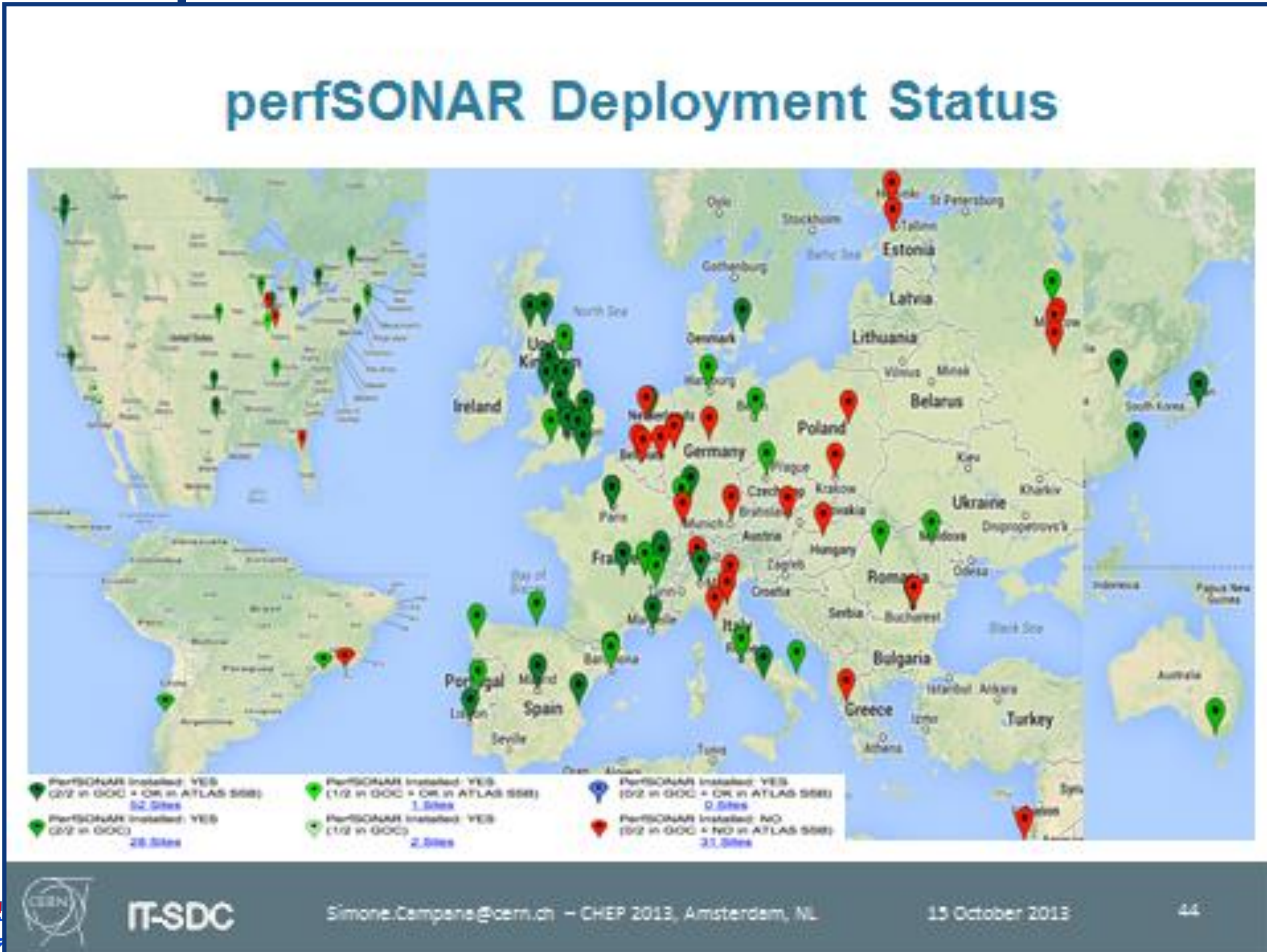
- **perfSONAR** is an infrastructure for network performance monitoring
  - Organized as consortium of organizations
    - building an interoperable network monitoring middle-ware
  - Defines the service types and a protocol for them to communicate
  - Develops the software packages to implement the services
- **perfSONAR-PS** is an open source development effort based on perfSONAR
  - targeted at creating an easy-to-deploy and easy-to-use set of perfSONAR services
  - Comes with all-in-one solution (CD or USB) or single packages for CentOS 5 and 6

Campana, CERN

WLCG deployment plan

- WLCG choose to deploy perfSONAR-PS at all sites worldwide
  - A dedicated WLCG Operations Task-Force was started in Fall 2012
- Sites are organized in regions
  - Based on geographical locations and experiments computing models
  - All sites are expected to deploy a bandwidth host and a latency host
- Regular testing is setup using a centralized ("mesh") configuration
  - Bandwidth tests: 30 seconds tests
    - every 6 hours intra-region, 12 hours for T2-T1 inter-region, 1 week elsewhere
  - Latency tests; 10 Hz of packets to each WLCG site
  - Traceroute tests between all WLCG sites each hour
  - Ping(ER) tests between all site every 20 minutes

Campana, CERN



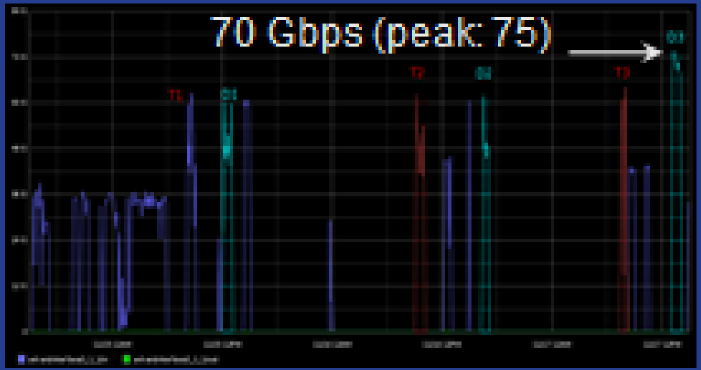
Big data over a 100G network at Fermilab (Gabriele Garzoglio, FNAL)

- One of our remote presentations
- Goal: verify whole stack of software and services end-to-end for effectiveness at 100G across participating labs
- Results on GridFTP/SRM/GlobusOnline, xrootd, squid/Frontier

Garzoglio, FNAL

100G High Throughput Data Program

- 2011: Advanced Network Initiative (ANI) Long Island MAN (LIMAN) testbed.
  - GO / GridFTP over 3x10GE.
- 2011-2012: Super Computing '11
  - Fast access to ~30TB of CMS data in 1h from NERSC to ANL using GridFTP.
  - 15 srv / 28 clnt – 4 gFTP / core; 2 strms; TCP Win. 2MB



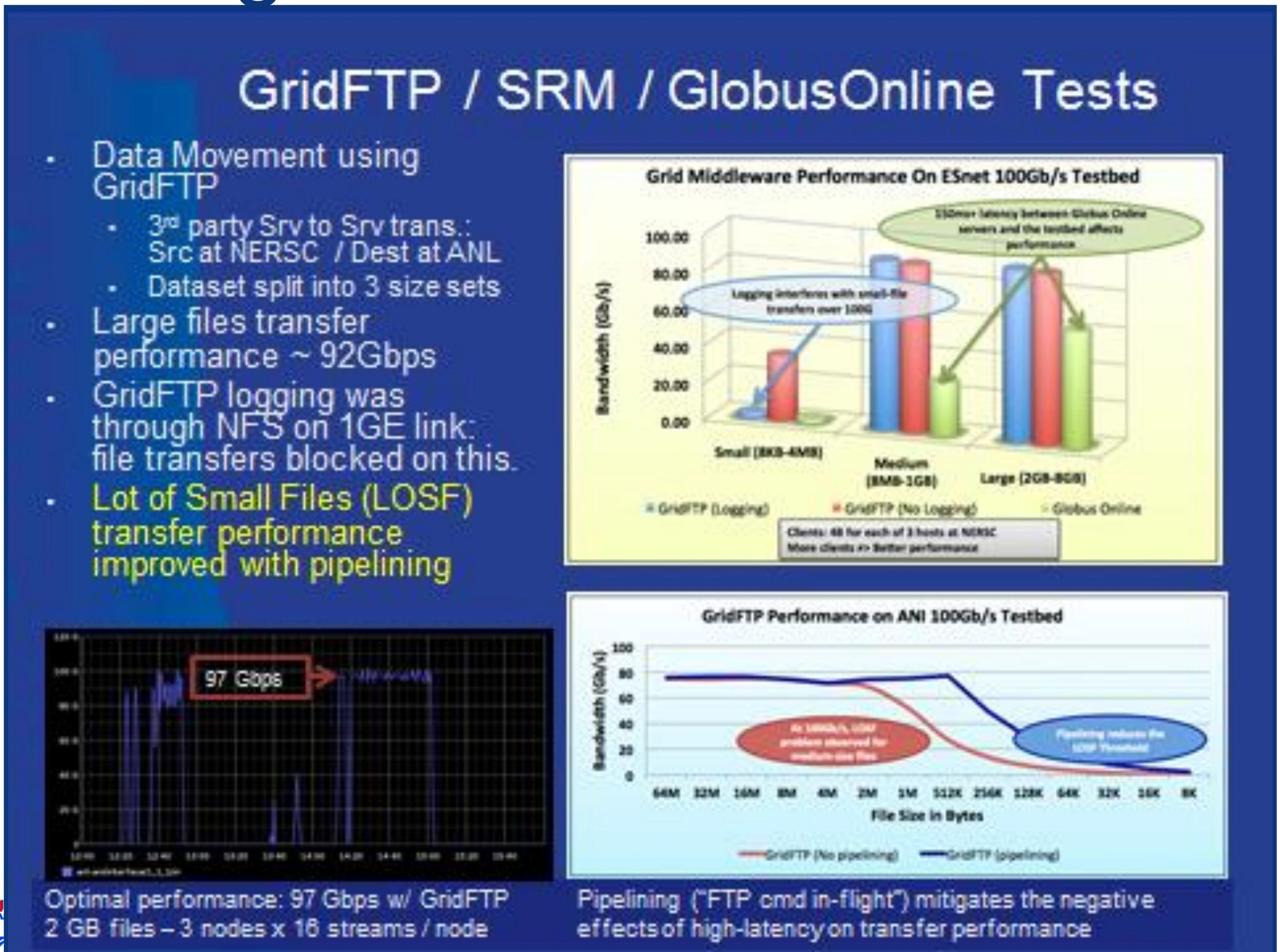
- 2012-2013: ESnet 100G testbed
  - Tuning parameters of middleware for data movement: xrootd, GridFTP, SRM, Globus Online, Squid. Achieved ~97Gbps
    - Rapid turn around on the testbed thank to custom boot images
  - Commissioning Fermilab Network R&D facility: 12 nodes at 8.5 Gbps per 10G node
  - Test NFS v4 over 100G using dCache (collab. w/ IBM research)
- Fall 2013 / Winter 2014: 100G Endpoint at Fermilab
  - Validate hardware link w/ data transfer apps with UFL and others. Talk to me if you want to participate in these activities.
  - Demonstrate storage-to-storage 100G rates

U.S. DEPARTMENT OF ENERGY

Fermilab

50

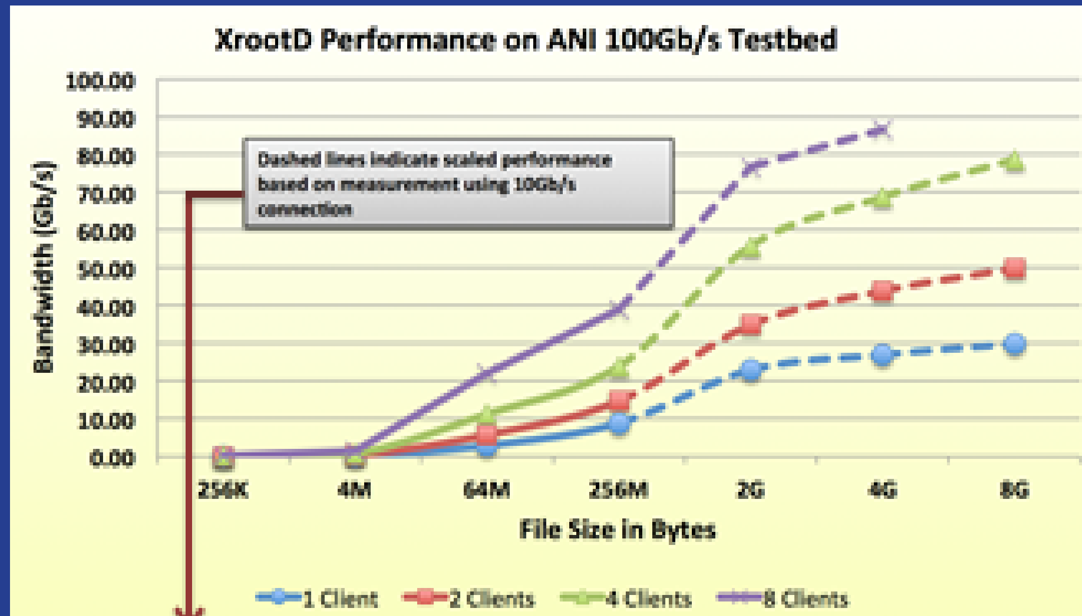
Garzoglio, FNAL



Garzoglio, FNAL

XRootD Tests

- Data Movement over XRootD, testing LHC experiment (CMS / Atlas) analysis use cases.
  - Clients at NERSC / Servers at ANL
  - Using RAMDisk as storage area on the server side
- Challenges
  - Tests limited by the size of RAMDisk
  - Little control over xrootd client / server tuning parameters



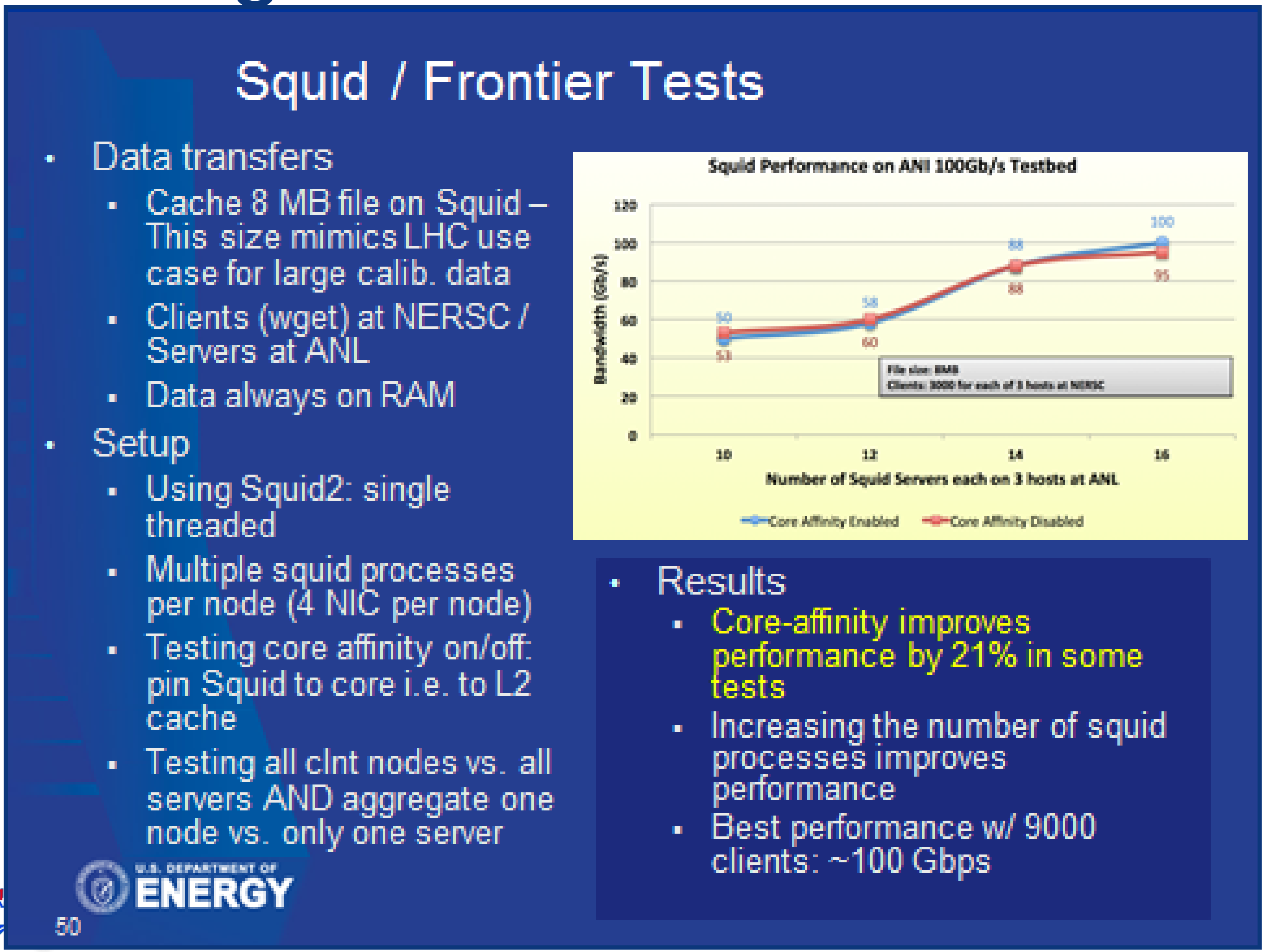
Dataset (GB)	1 NIC measurements (Gb/s)	Aggregate Measurements (12 NIC) (Gb/s)	Scale Factor per NIC	Aggregate estimate (12 NIC) (Gb/s)
0.512	4.5	46.9	0.87	–
1	6.2	62.4	0.83	–
4	8.7 (8 clients)	–	0.83	86.7
8	7.9 (4 clients)	–	0.83	78.7

Calculation of the scaling factor between 1 NIC and an aggregated 12 NIC for datasets too large to fit on the RAM disk

U.S. DEPARTMENT OF ENERGY

50

Garzoglio, FNAL



Network architecture and IPv6 deployment at CERN (David Gutierrez Rueda, CERN)

- Core network interconnecting all infrastructure, including Wigner, is IPv6 ready
  - Non-blocking 1 Tbps

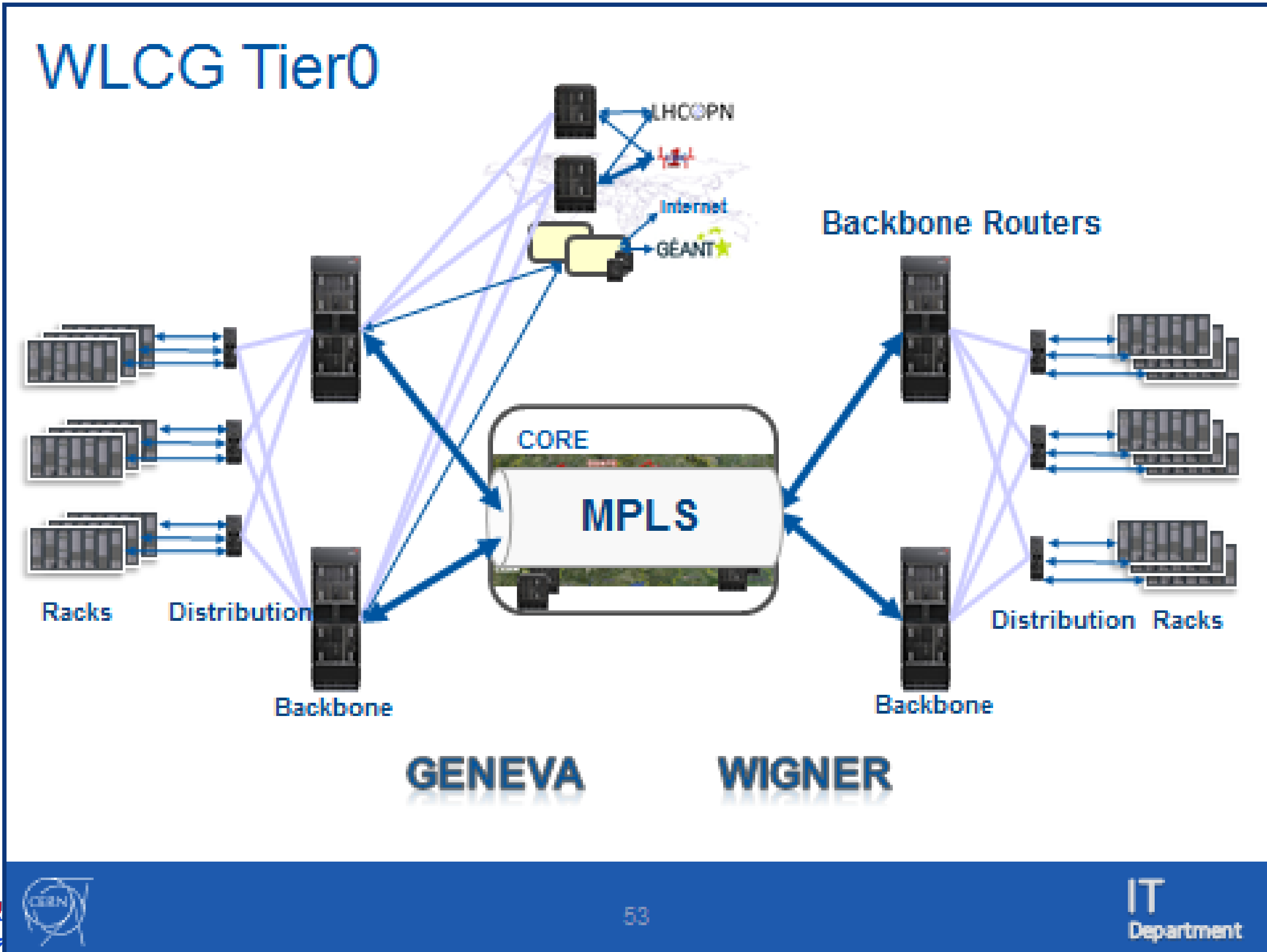
Gutierrez Rueda, CERN

External Network

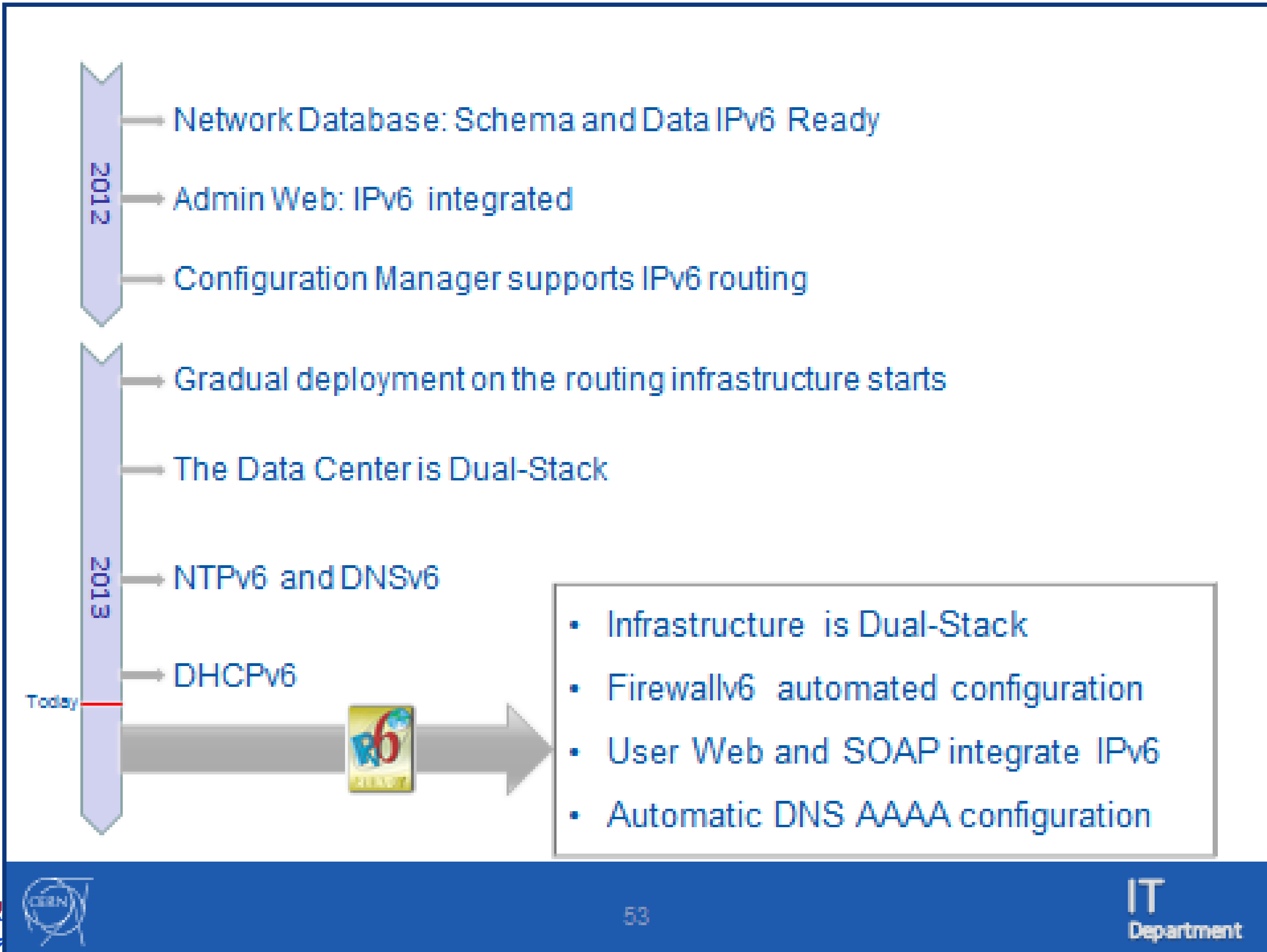
- Public general purpose connections
  - Full BGP Internet routing table
  - Geant, CIXP, ISPs
- Private WLCG
  - LHCOPN
    - 70Gbps peaks to T1
  - LHCONE

Brocade Routers	8
BGP Peerings	86
Aggregated BW	232 Gbps
IPv4/IPv6 Dual Stack	YES

Gutierrez Rueda, CERN



Gutierrez Rueda, CERN



Application performance evaluation and recommendations for the DYNES instrument (Shawn McKee, U Michigan)

- DYNES is a “distributed instrument” in the US: has networking infrastructure at ~40 universities for creating virtual circuits.
- Solving a mystery: When creating circuits 1Gbps, they were getting 200Mbps performance.
  - Traditional network debugging techniques yielded nothing.
  - Solution: Using the Linux outgoing packet queue management layer to pace packets on the host at less than the circuit speed. Yielded >800 Mbps.
  - Belief the issue is QoS in the internal implementation of one hop in the circuit is at fault.
- Lesson: Virtual circuits still depend heavily on the underlying hardware implementation. The “virtualization” is perhaps not a complete extraction. You must know your circuit!

McKee, U Michigan

TC Test Results

```
[dynes@fdt-wisc ~]$ nuttcp -T 30 -i 1 -p 5679 -P 5678 10.10.200.10
    2.1875 MB /   1.00 sec =   18.3486 Mbps     0 retrans
    8.3125 MB /   1.00 sec =   69.7281 Mbps     1 retrans
   28.3125 MB /   1.00 sec =  237.5170 Mbps     0 retrans
   99.1875 MB /   1.00 sec =  832.0559 Mbps     0 retrans
  108.5000 MB /   1.00 sec =  910.1831 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6078 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6706 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6215 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5747 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1354 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1363 Mbps     0 retrans
  108.2500 MB /   1.00 sec =  908.0605 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1218 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5911 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5902 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6133 Mbps     0 retrans
  108.5000 MB /   1.00 sec =  910.1731 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6533 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1199 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6388 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1154 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6406 Mbps     0 retrans
  108.3750 MB /   1.00 sec =  909.1154 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5911 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6388 Mbps     0 retrans
  108.5000 MB /   1.00 sec =  910.1640 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5593 Mbps     0 retrans
  108.5000 MB /   1.00 sec =  910.1967 Mbps     0 retrans
  108.4375 MB /   1.00 sec =  909.6397 Mbps     0 retrans
  108.3125 MB /   1.00 sec =  908.5911 Mbps     0 retrans

 2965.6678 MB /  30.12 sec =  825.9052 Mbps 3 %TX 8 %RX 1 retrans 36.73 msRTT
```

- **Very close to the 900 Mbps shaped request. This works much better. Retry with 1000 Mbps TC config next**

WLCG security: a trust framework for security collaboration among infrastructures (David Kelsey, STFC-RAL)

- All about trust of infrastructures
- Building on experience with EDG/EGEE/EGI, OSG, WLCG

Kelsey, STFC-RAL (1)

Security for Collaborating Infrastructures (SCI)

- A collaborative activity of information security officers from large-scale infrastructures
  - EGI, OSG, PRACE, EUDAT, CHAIN, WLCG, XSEDE, …
- Developed out of EGEE – started end of 2011
- We are developing a *Trust framework*
  - Enable interoperation (security teams)
  - Manage cross-infrastructure security risks
  - Develop policy standards
  - Especially where not able to share identical security policies

GPU-based network traffic monitoring and analysis tools (Phil DeMar, FNAL)

- Another remote presentation
- 10G common in servers, 40G and 100G coming on backbones
- Current flow- and traffic-based tools will break down

DeMar, FNAL (1)

Packet-Based Analysis

- Our preferred choice for 40/100GE traffic analysis:
  - Flow data limitations (sampled) constrain flow-based analysis
- Characteristics of packet-based network monitoring & analysis applications
  - Time constraints on packet processing.
  - Highly compute and I/O throughput-intensive
  - High levels of data parallelism.
    - Each packet can be processed independently
  - Extremely poor temporal locality for data
    - Typically, data processed once in sequence; rarely reused

63 Fermilab

US LHC Tier-1 WAN data movement security architectures (Phil DeMar, FNAL)

- Remote again…
- Both FNAL and BNL chose to separate science data movements from general network traffic

DeMar, FNAL (2)

Summary

- Separating science data movement from general network traffic has worked well at US-LHC Tier-1s
  - Enabled us to meet needs of both LHC stakeholders & general users, but not at each other's expense
  - Science DMZ architectures based around PBR for LHC traffic:
    - Avoids performance issues with overloading perimeter security tools
- Our implementations work well for us because:
  - We are dealing with established traffic characteristics
  - Our stakeholders are well-organized & long-lived
  - May not translate well to other disciplines
- Looking toward OpenFlow as a more standard approach to separate out our science data movement

Fermilab

WLCG and IPv6: The HEPiX IPv6 working group (David Kelsey, STFC-RAL)

- IPv4 address depletion coming soon…
- Network infrastructures increasingly ready for IPv6
- Many services not yet tested, much work to be done still

Kelsey, STFC-RAL (2)

HEPiX

Timetable WLCG IPv6 transition

- In 2012 we said:
  - Support for IPv6-only clients *not before* Jan 2014
- Still true!
  - And likely to be much later
  - Needs MANY sites to support IPv6
- Sysadmins, Security staff, Monitoring and Operations
  - Training required
  - New operational procedures
- The WG will continue to test Use Cases
- Aim for dual-stack on most/many services – when?
- Aiming for an IPv6 workshop at CERN in Spring/Summer 2014

Collaborative Tools

Indico 1.0+ (Jose Benito Gonzalez Lopez, CERN)

- Remarkable growth, very popular service
- Added user dashboard, version optimised for mobile devices, …
- Coming: rich abstract editor, configurable registration form, e-ticket, off-line web site

Gonzalez Lopez, CERN

Gonzalez Lopez, CERN

GLOBAL INDICO SERVICE

New service

No restrictions

Open to the whole research **world**

Hosted by IT department at CERN

Benefit from effort on scalability, virtualization

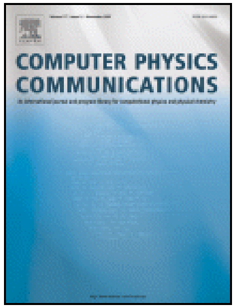
Vidyo for the LHC (Thomas Baron, CERN)

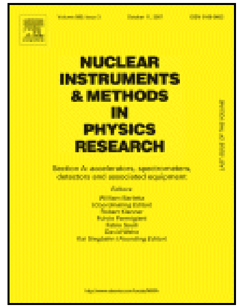
- In production since December 2011
- Strong points for CERN
  - Multiplatform capabilities
    - Not only desktops but extension to mobiles and tablets
  - Integration with H323/SIP protocols
  - Extensible (several hundreds in a single meeting)
  - Natural interactions
    - very low latency and excellent lip sync
  - Good A/V quality and resilience/adaptability to poor network conditions
  - Simple interface
  - Good integration possibilities

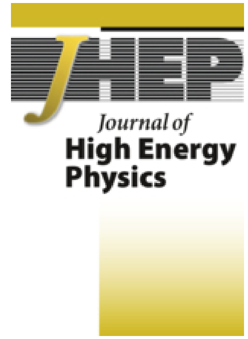
Baron, CERN

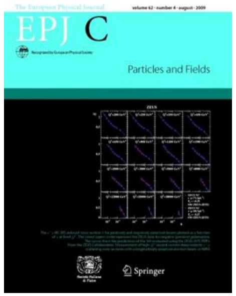
CERN Vidyo Service Scale

- 18000 users
- 800-1600 simultaneous connections
- Up to 168 simultaneous H323/SIP
- 11 Phone access points worldwide
- 12 simultaneous recordings
- 2 VidyoPanoramas (CERN site)









Scholarly literature and the press: scientific impact and social perception of physics computing

M. G. Pia[1], T. Basaglia[2], Z. W. Bell[3], P. V. Dressendorfer[4]

[1]*INFN Genova, Genova, Italy*
[2]*CERN, Geneva, Switzerland*
[3]*ORNL, Oak Ridge, TN, USA*
[4]*IEEE, Piscataway, NJ, USA*

CHEP 2013
Amsterdam

IEEE NSS 2013
Seoul, Korea

Pia / INFN Genova

Conclusions

"You have to shout to be heard when it comes to getting science into the media and people to listen," said Professor Lythgoe.

THE TIMES

Hannah Devlin Science Editor
Published at 12:01AM, June 7 2013

How loud to get
HEP software into the media and
HEP management
to listen?

Setting up collaborative tools for a 1000-member community (Dirk Hoffmann, CPPM)

- Cta: Collaboration without support of a strong institute
- Had to set up basic services themselve

Hoffmann, CPPM

Summary

- **CTA member administration and collaborative workflows implemented on four (principal) feet:**
  SharePoint2010 – InDiCo – Mailman + LDAP
- **InDiCo still without alternative**
- **Good universal "group registry" not for free**
- **SharePoint is expensive, but valuable framework, if costs are shared (a bit like Oracle).**
  - Work (temporary staff, trainee) on ShPt highly qualifying
  - Standard tools and options satisfactory for most cases
  - Development and interoperability not easy, but feasible

CHEP 2013 – Dirk Hoffmann, October 17th, 2013

Final Words from Track 6

THANK YOU

- to all speakers and poster presenters
  for many interesting contributions
- to all attendees to the sessions
  for their interest and for lively discussions
- to my fellow conveners
  for a smooth sailing of the track,
  and for their input to this summary
- to the organisers
  for a great CHEP 2013!

SEE YOU IN …