CERN Infrastructure Evolution

Tim Bell

CHEP

24th May 2012

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

Agenda

- Problems to address
- Ongoing projects
  - Data centre expansion
  - Configuration management
  - Infrastructure as a Service
  - Monitoring
- Timelines
- Summary

Problems to address

- CERN data centre is reaching its limits
- IT staff numbers remain fixed but more computing capacity is needed
- Tools are high maintenance and becoming increasingly brittle
- Inefficiencies exist but root cause cannot be easily identified

Data centre extension history

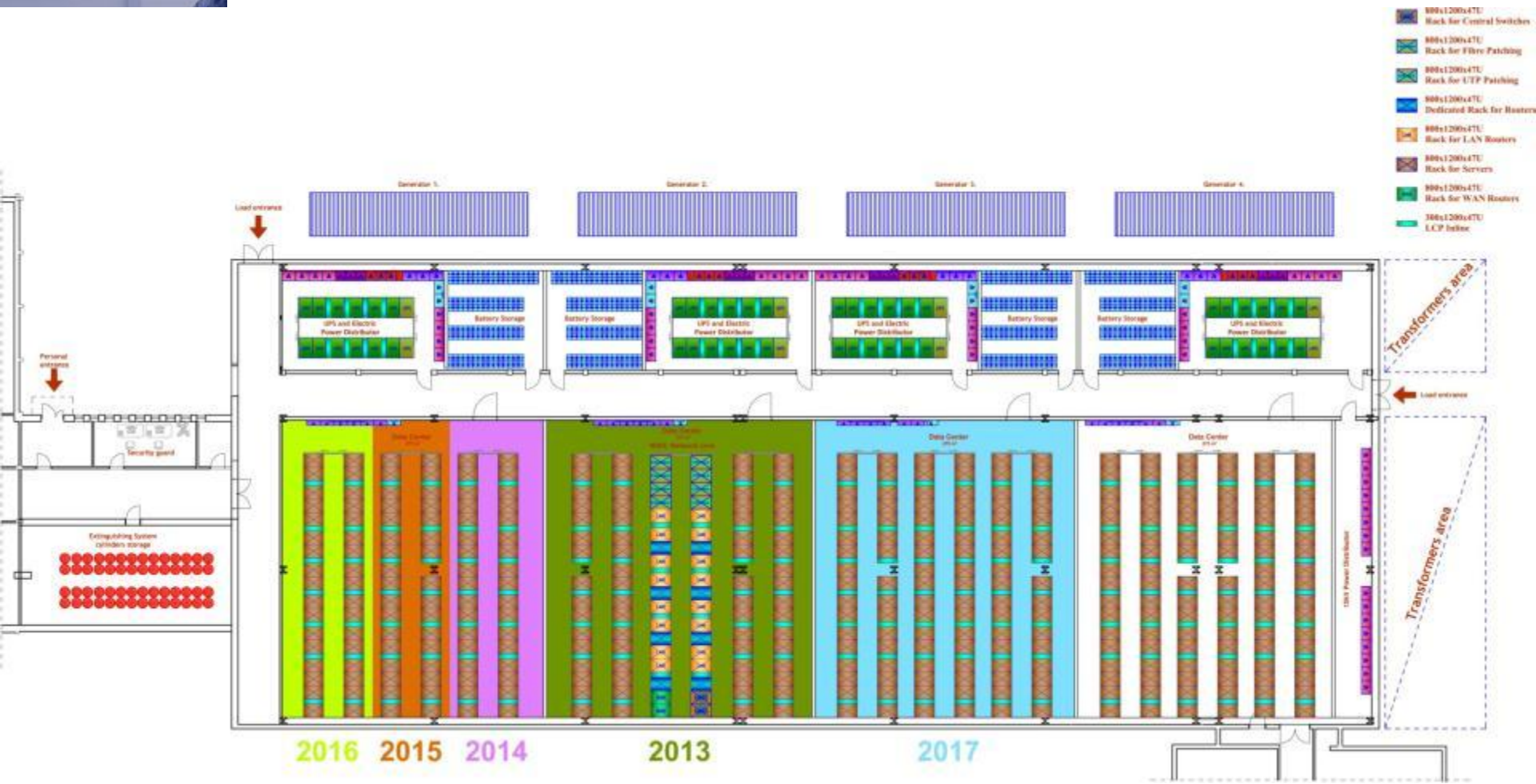
CERN IT Department

- Studies in 2008 into building a new computer centre on the CERN Prevessin site
  - Too costly
- In 2011, tender run across CERN member states for remote hosting
  - 16 bids
- In March 2012, Wigner Institute in Budapest, Hungary selected

Data Centre Selection

- Wigner Institute in Budapest, Hungary

Data Centre Layout & Ramp-up



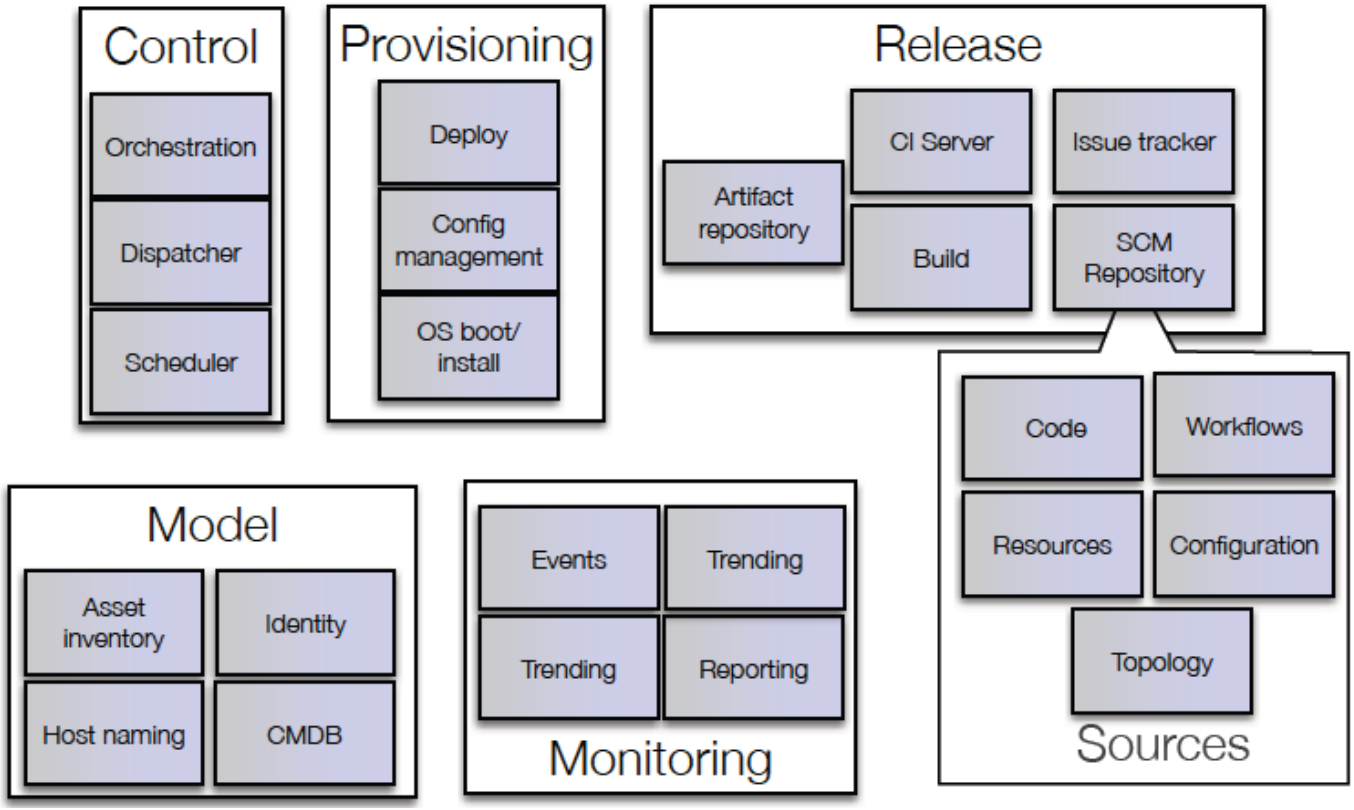
CERN Infrastructure Evolution

Infrastructure Tools Evolution

- We had to develop our own toolset in 2002
- Nowadays,
  - CERN compute capacity is no longer leading edge
  - Many options available for open source fabric management
  - We need to scale to meet the upcoming capacity increase
- If there is a requirement which is not available through an open source tool, we should question the need
  - If we are the first to need it, contribute it back to the open source tool

Where to Look?

- Large community out there taking the “tool chain” approach whose scaling needs match ours: O(100k) servers and many applications
- Become **standard** and join this community

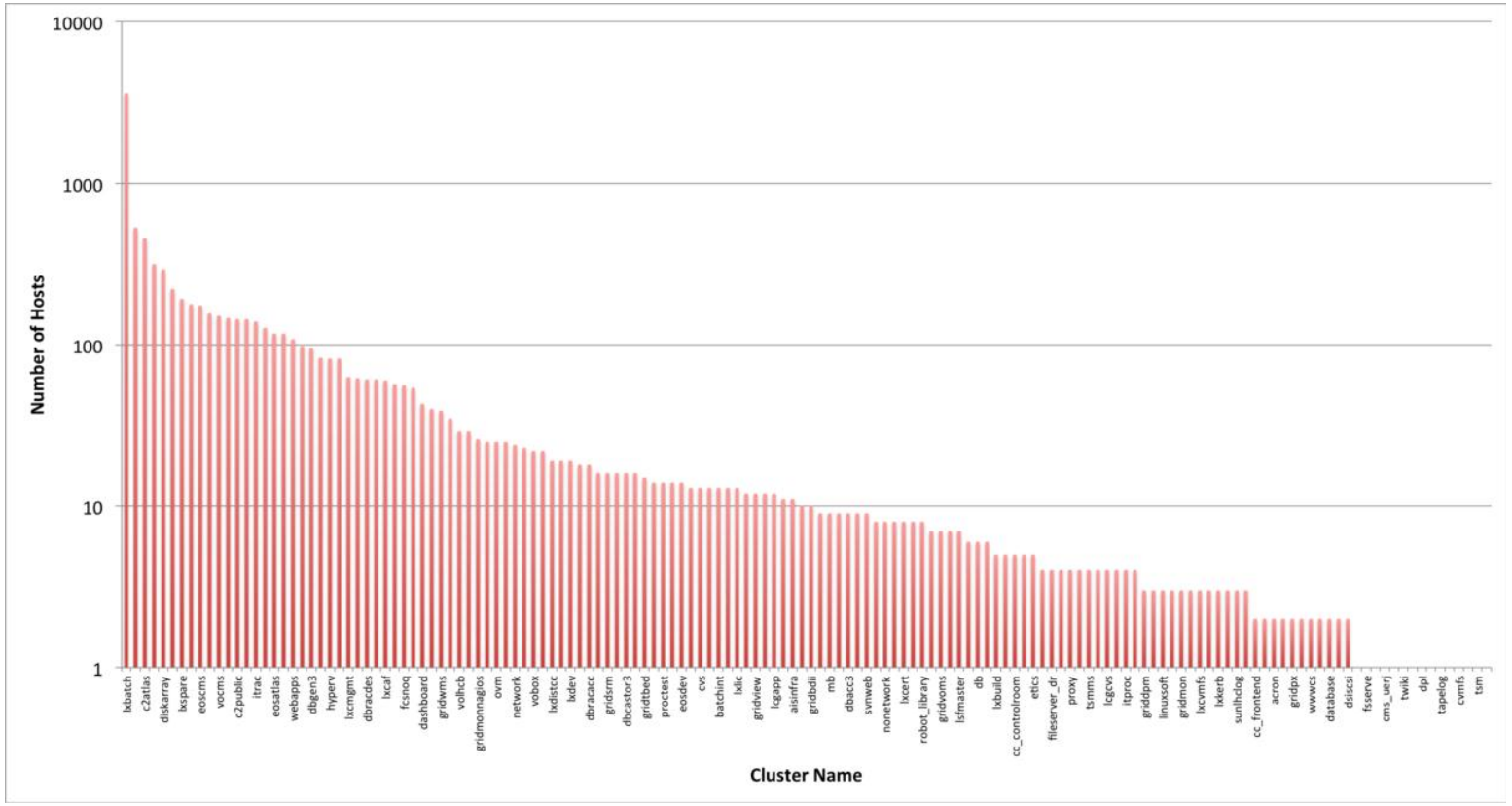


Bulk Hardware Deployment

- Adapt processes for minimal physical intervention
- New machines register themselves automatically into network databases and inventory
- Burn-in test uses standard Scientific Linux and production monitoring tools
- Detailed hardware inventory with serial number tracking to support accurate analysis of failures

Current Configurations

CERN IT Department



- Many, diverse applications (“clusters”) managed by different teams
- ..and 700+ other “unmanaged” Linux nodes in VMs that could benefit from a *simple* configuration system

Job Adverts from Indeed.com

Index of millions of worldwide job posts across thousands of job sites

These are the sort of posts our departing staff will be applying for.

Configuration Management

- Using off the shelf components
  - Puppet – configuration definition
  - Foreman – GUI and Data store
  - Git – version control
  - Mcollective – remote execution
- Integrated with
  - CERN Single Sign On
  - CERN Certificate Authority
  - Installation Server

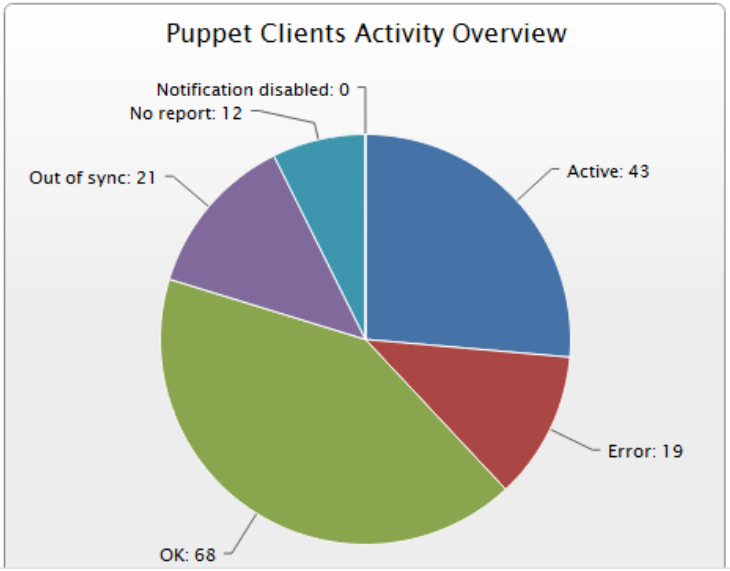
Status

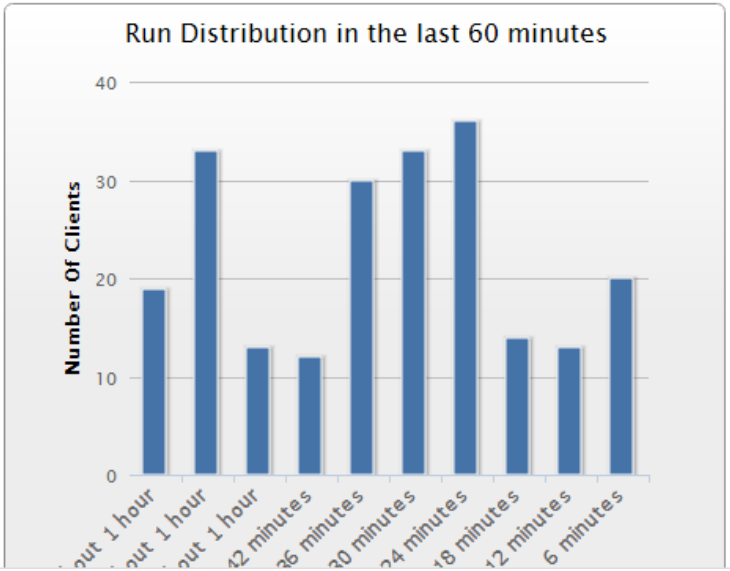
Foreman | Dashboard | Hosts | Reports | Facts | Audits | Statistics | More | Timbell

Generated at 18 May 15:21

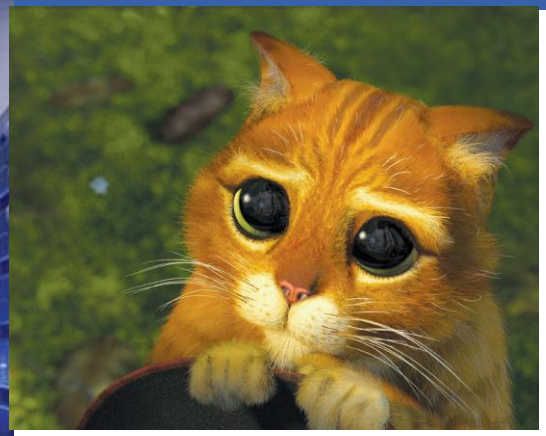
Description	Data
Hosts that had performed modifications	43
Hosts in Error State	19
Good Host Reports in the last 60 minutes	111 / 163 hosts (68%)
Out Of Sync Hosts	21
Hosts With No Reports	12
Hosts With Alerts Disabled	0

Puppet Clients Activity Overview





Different Service Models



- Pets are given names like lsfmaster.cern.ch
- They are unique, lovingly hand raised and cared for
- When they get ill, you nurse them back to health

- Cattle are given numbers like vm0042.cern.ch
- They are almost identical to other cattle
- When they get ill, you get another one

- Future application architectures tend towards Cattle
- .. But users now need support for both modes of working

Infrastructure as a Service

- Goals
  - Improve repair processes with virtualisation
  - More efficient use of our hardware
  - Better tracking of usage
  - Enable remote management for new data centre
  - Support potential new use cases (PaaS, Cloud)
  - Sustainable support model
- At scale for 2015
  - 15,000 servers
  - 90% of hardware virtualized
  - 300,000 VMs needed

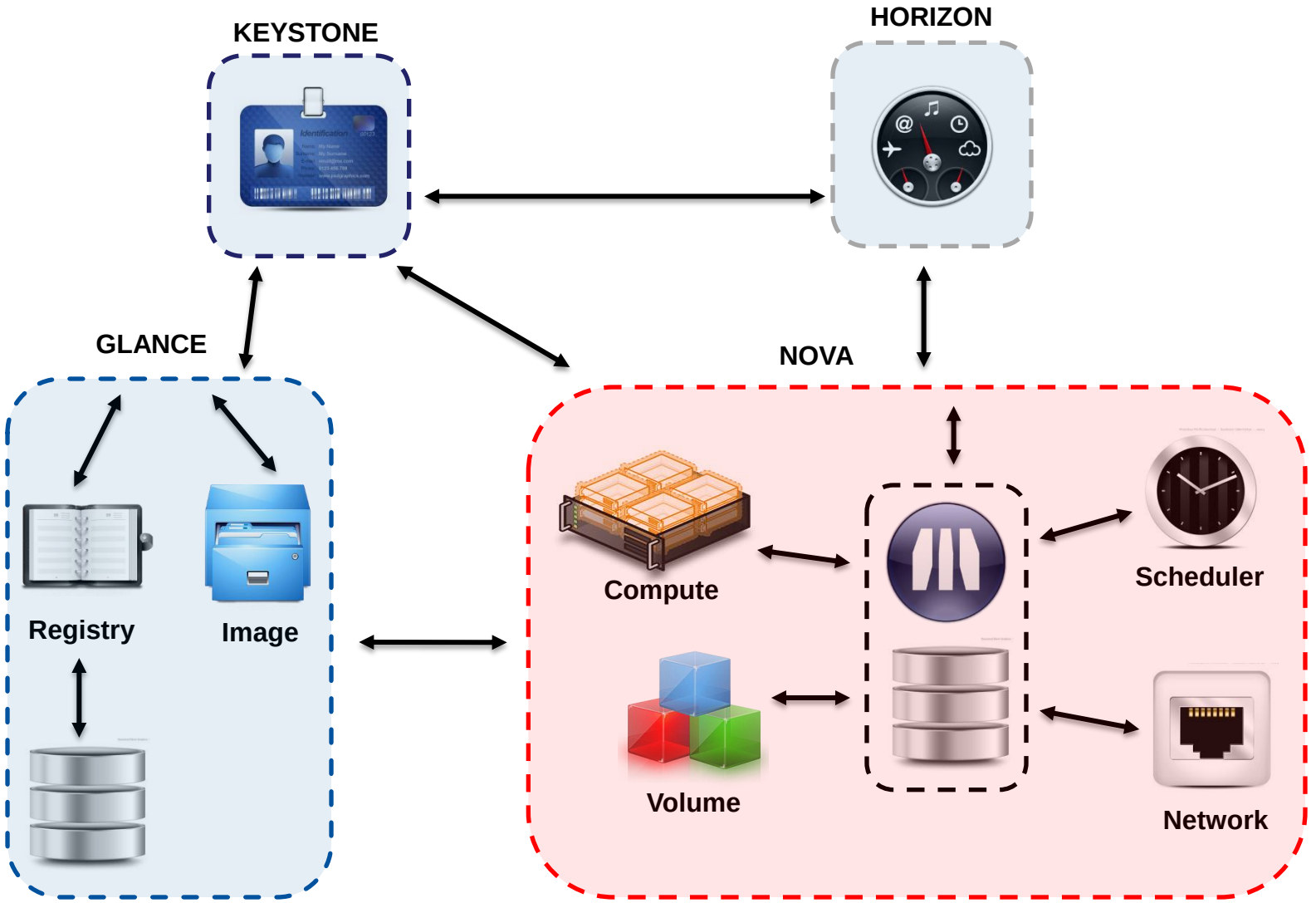
OpenStack

- Open source cloud software
- Supported by 173 companies including IBM, RedHat, Rackspace, HP, Cisco, AT&T, …
- Vibrant development community and ecosystem
- Infrastructure as a Service to our scale
- Started in 2010 but maturing rapidly

openstack™

Openstack @ CERN

CERN IT Department



Scheduling and Accounting

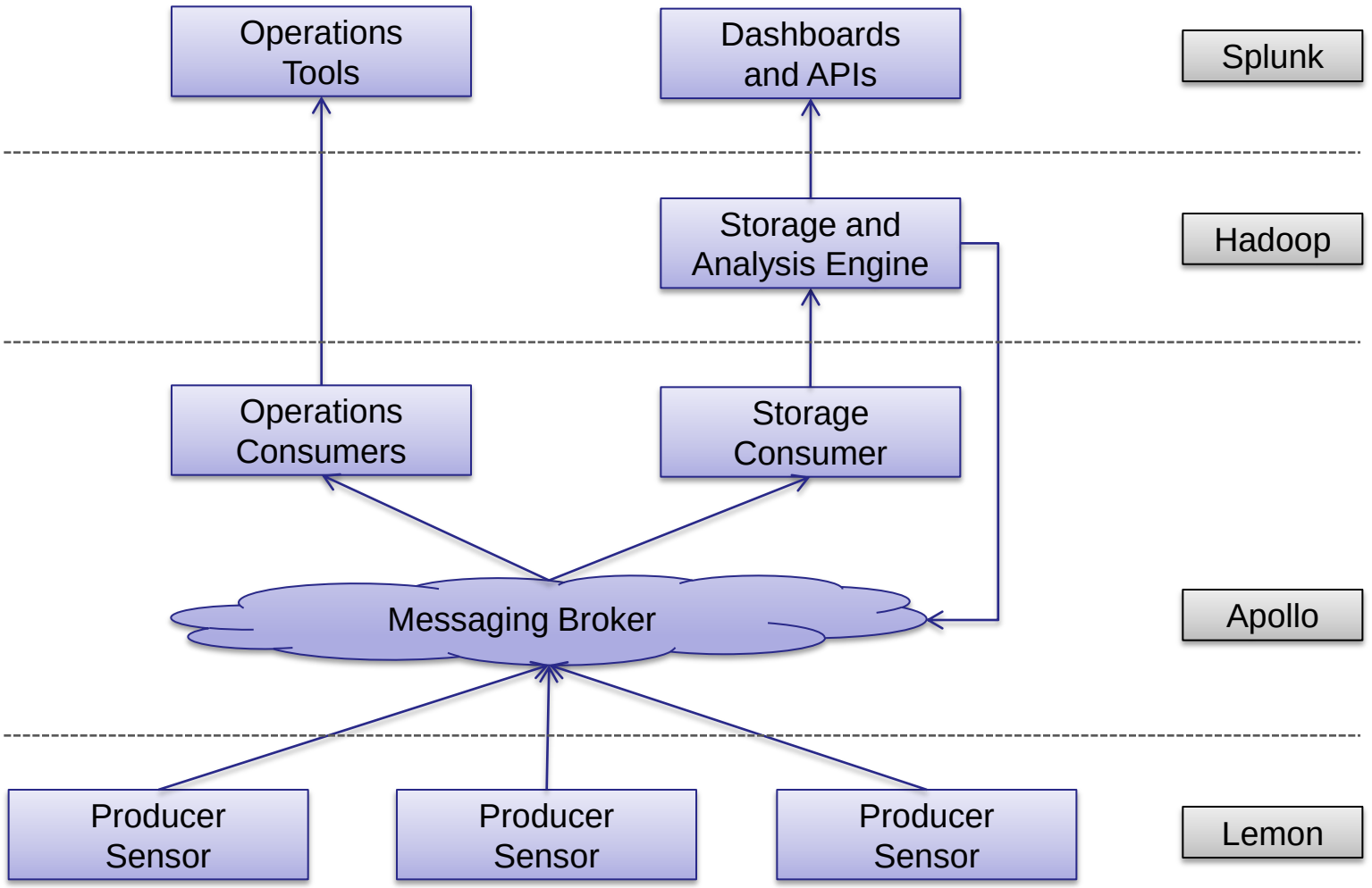
- Multiple uses of IaaS
  - Server consolidation
  - Classic batch (single or multi-core)
  - Cloud VMs such as CERNVM
- Scheduling options
  - Availability zones for disaster recovery
  - Quality of service options to improve efficiency such as build machines, public login services
  - Batch system scalability is likely to be an issue
- Accounting
  - Use underlying services of IaaS and Hypervisors for reporting and quotas

Monitoring in IT

- >30 monitoring applications
  - Number of producers: ~40k
  - Input data volume: ~280 GB per day
- Covering a wide range of different resources
  - Hardware, OS, applications, files, jobs, etc.
- Application-specific monitoring solutions
  - Using different technologies (including commercial tools)
  - Sharing similar needs: aggregate metrics, get alarms, etc
- Limited sharing of monitoring data
  - Hard to implement complex monitoring queries

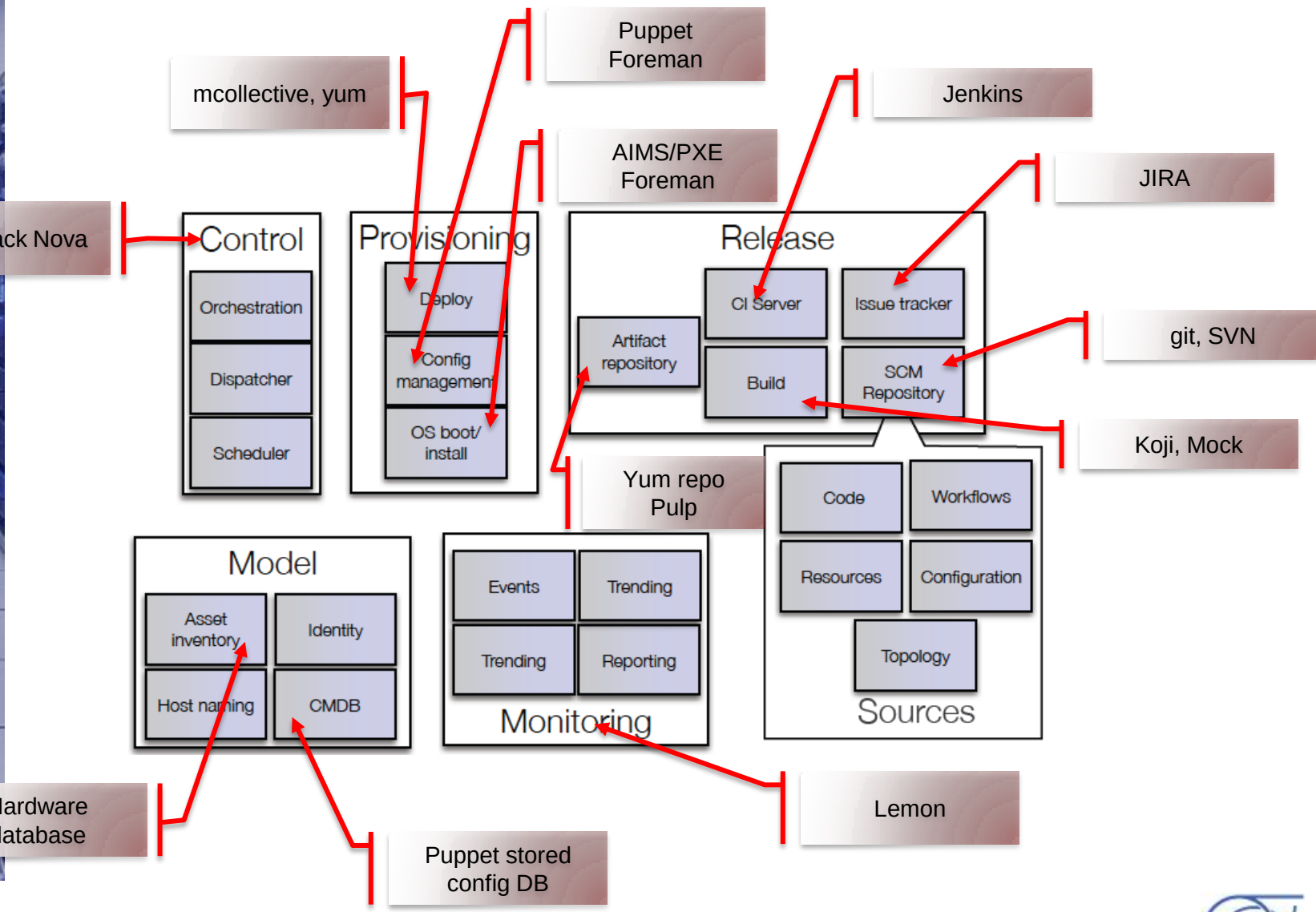
Monitoring Architecture

CERN IT Department



Current Tool Snapshot (Liable to Change)

CERN IT Department



Timelines

CERN IT Department

Year	What	Actions
2012		Prepare formal project plan Establish IaaS in CERN Data Centre Monitoring Implementation as per WG Migrate lxcloud users Early adopters to use new tools
2013	LS 1 New Data Centre	Extend IaaS to remote Data Centre Business Continuity Migrate CVI users General migration to new tools with SLC6 and Windows 8
2014	LS 1 (to November)	Phase out legacy tools such as Quattor

Conclusions

- New data centre provides opportunity to expand the Tier-0 capacity
- New tools are required to manage these systems and improve processes given fixed staff levels
- Integrated monitoring is key to identify inefficiencies, bottlenecks and usage
- Moving to an open source tool chain has allowed a rapid proof of concept and will be a more sustainable solution

Background Information

- HEPiX Agile Talks
  - http://cern.ch/go/99Ck
- Tier-0 Upgrade
  - http://cern.ch/go/NN98

Other information, get in touch

Tim.Bell@cern.ch

Backup Slides

CERN Data Centre Numbers

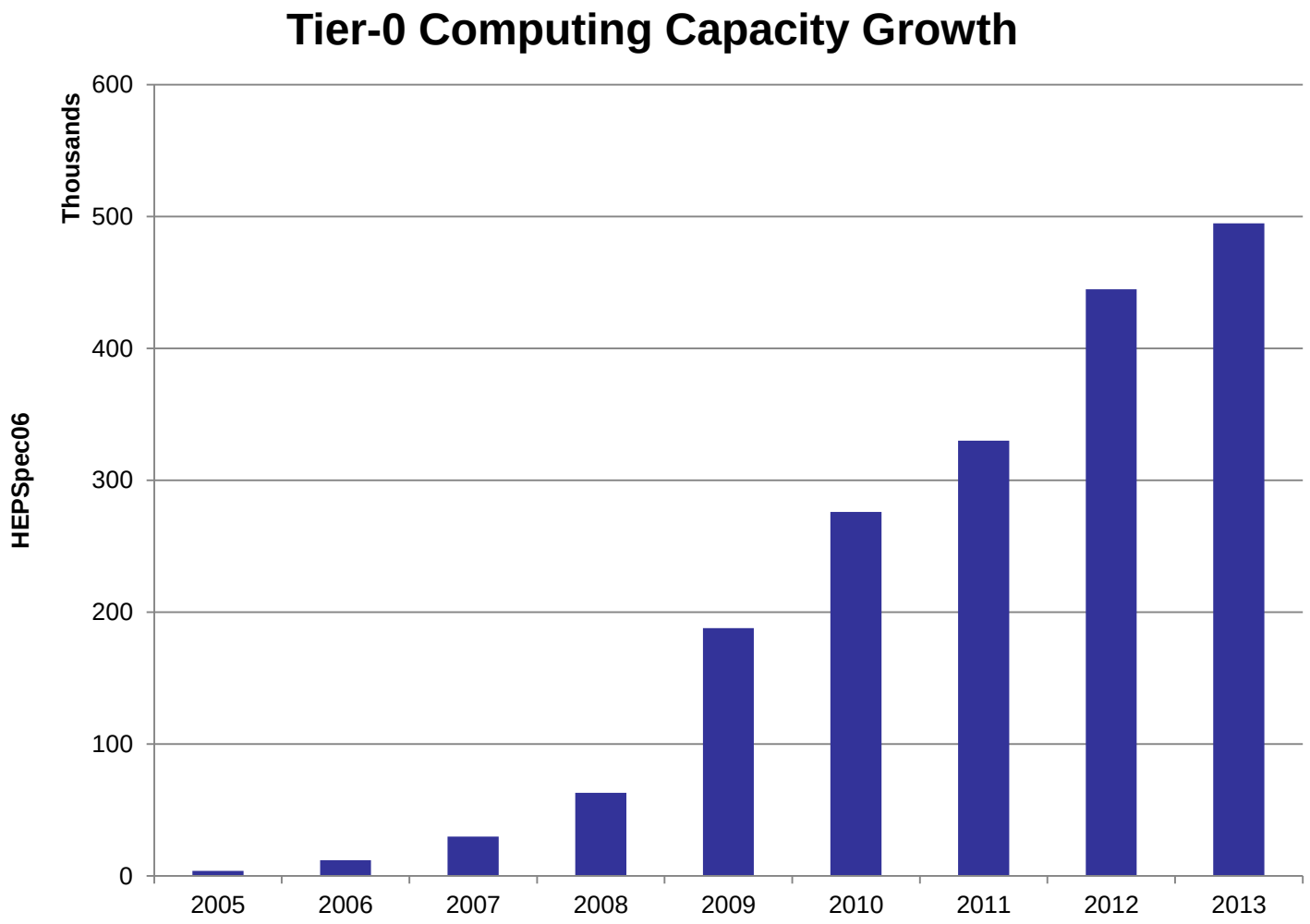
CERN IT Department

Systems	7,899	Hard disks	62,023
Processors	14,972	Raw disk capacity (TiB)	62,660
Cores	64,623	Tape capacity (PiB)	47
Memory (TiB)	165	Ethernet 1Gb ports	16,773
Racks	1,070	Ethernet 10Gb ports	622
Power Consumption (KiW)	2,345		

From http://sls.cern.ch/sls/service.php?id=CCBYNUM

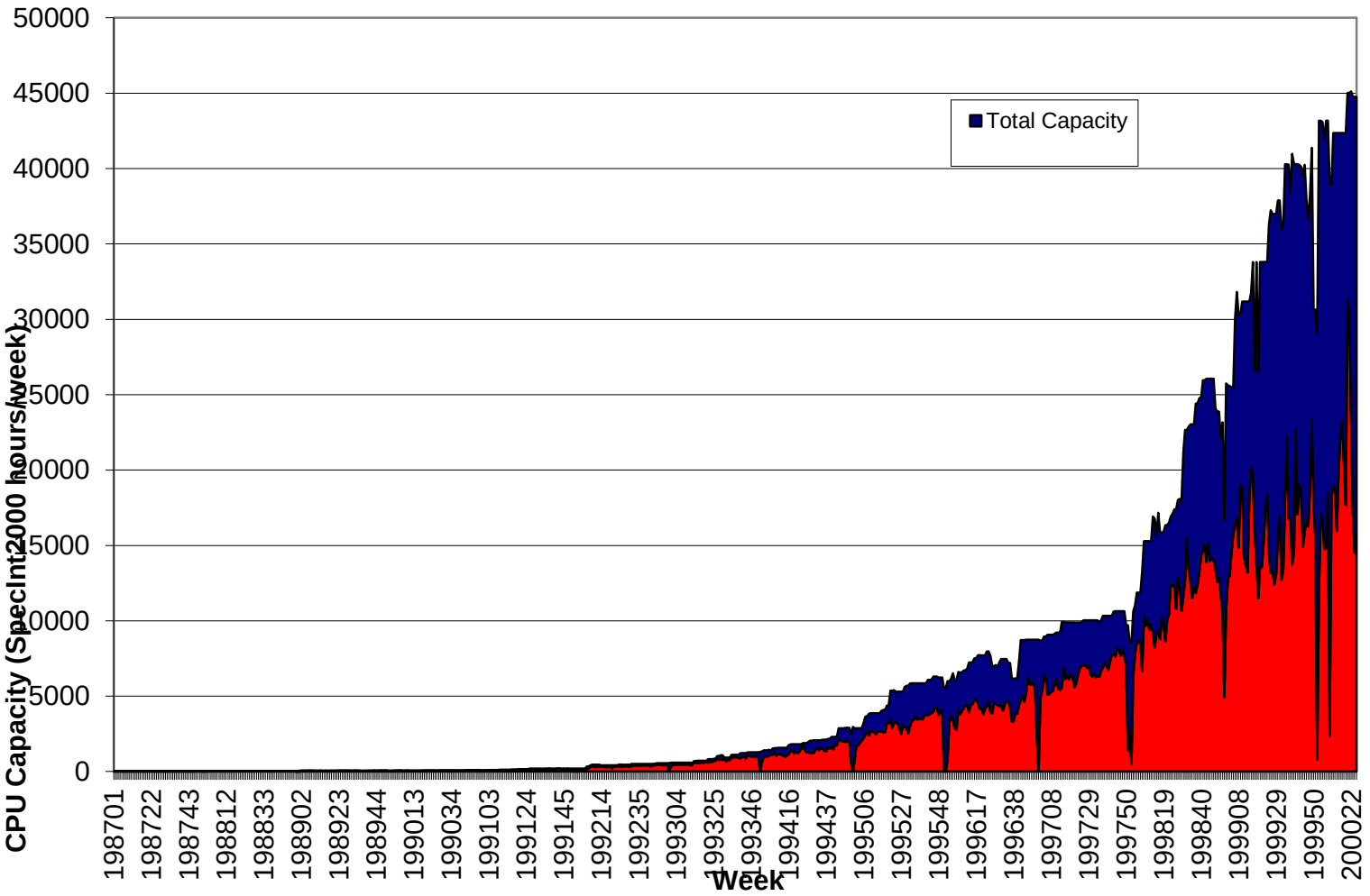
Computer Centre Capacity

CERN IT Department

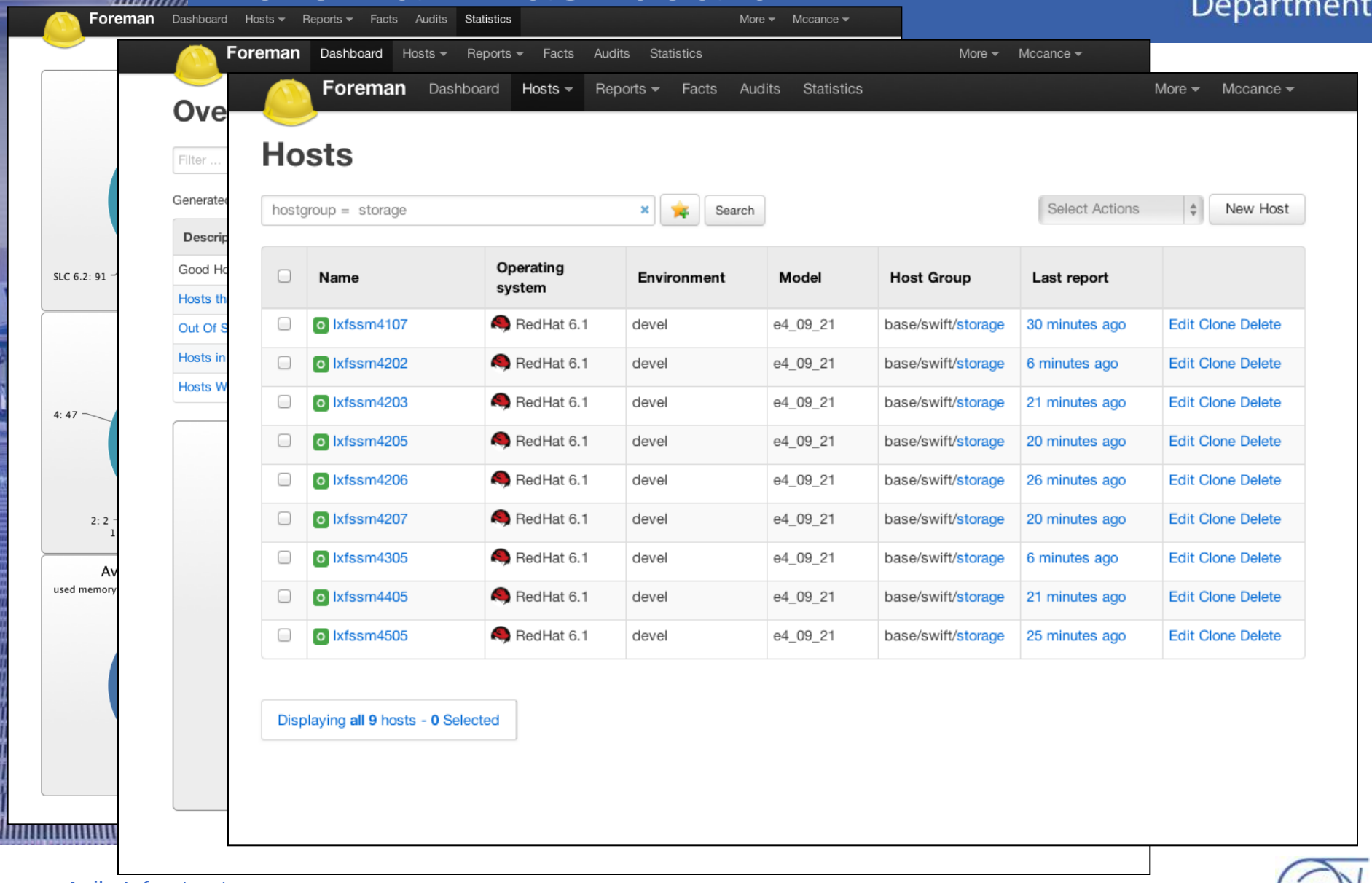


Capacity to end of LEP

CERN IT Department



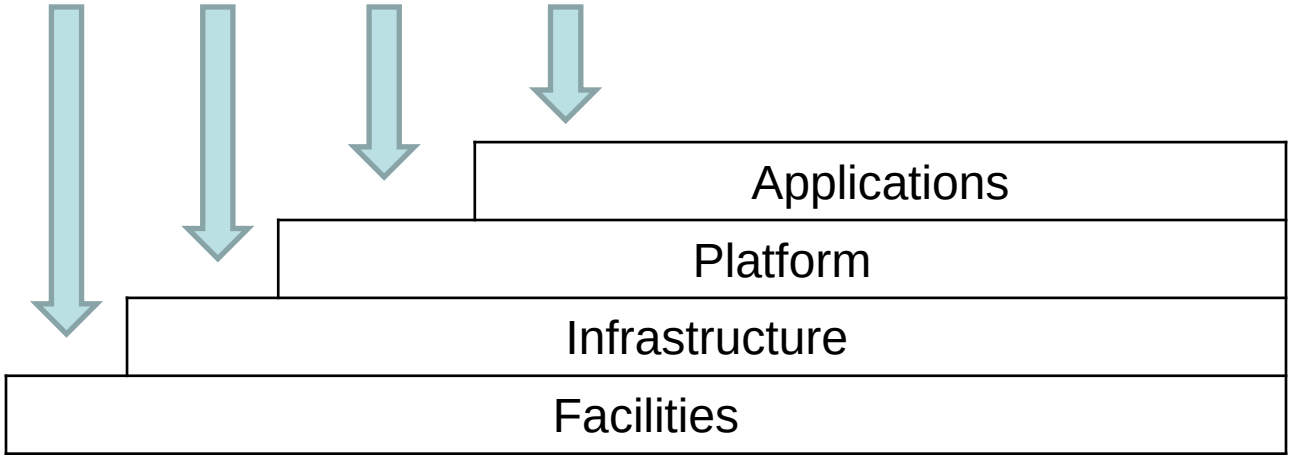
Foreman Dashboard



Industry Context - Cloud

CERN IT Department

- Cloud computing models are now standardising
  - Facilities as a Service – such as Equinix, Safehost
  - Infrastructure as a Service - Amazon EC2, CVI or lxcloud
  - Platform as a Service - Microsoft Azure or CERN Web Services
  - Software as a Service – Salesforce, Google Mail, Service-Now, Indico
- Different customers want access to different layers
  - Both our users and the IT Service Managers



SLC 4 to SLC 5 migration

CERN IT Department