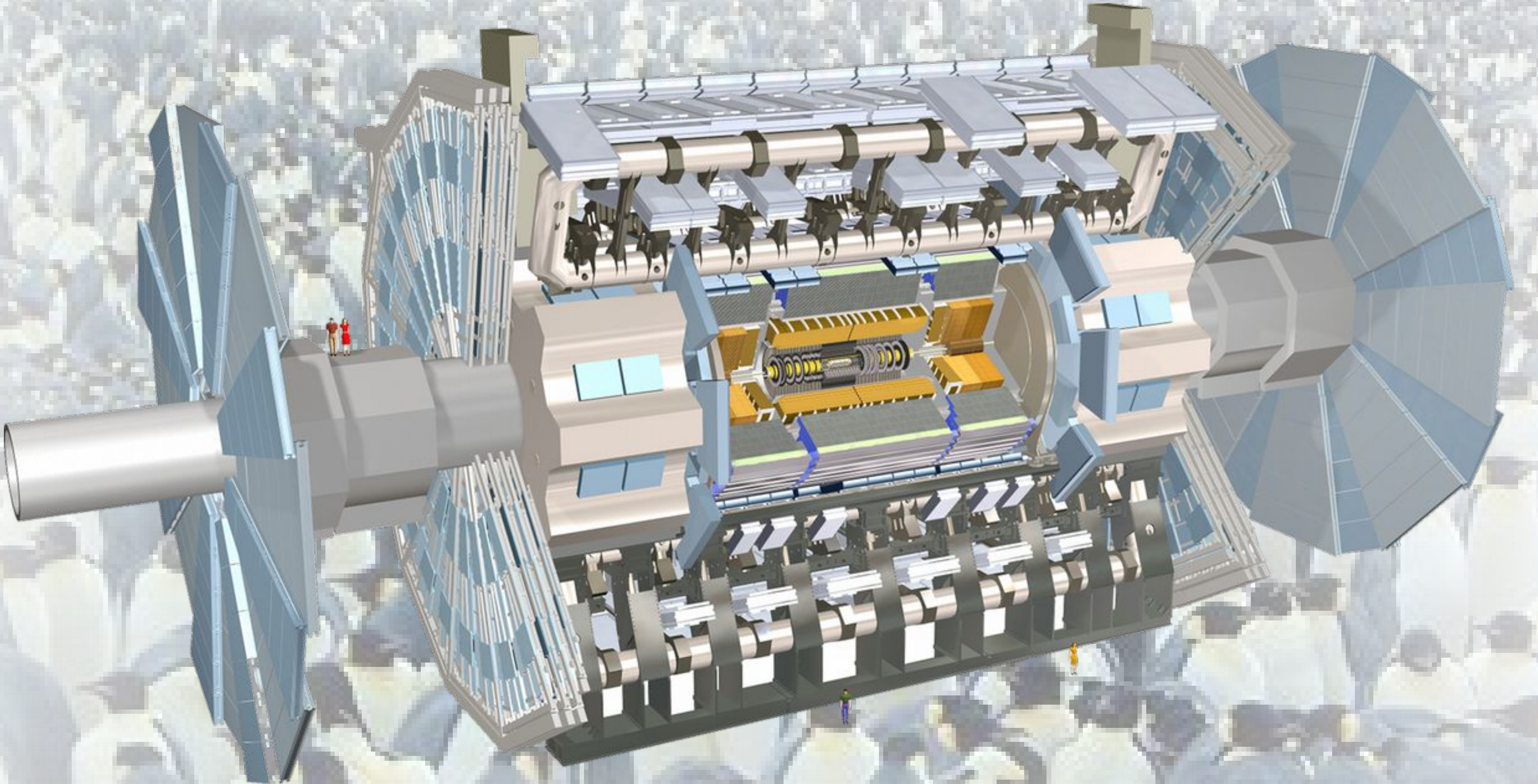
System administration work and opportunities in large Trigger and Data Acquisition systems

the ATLAS example



Sergio Ballestrero - *University of Johannesburg, South Africa*
for the
ATLAS TDAQ SysAdmins team.

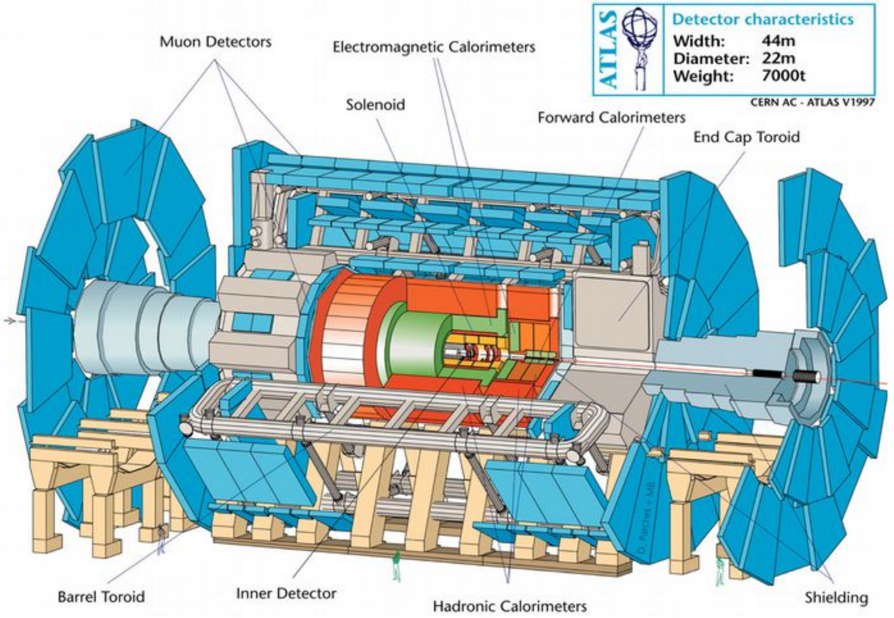
System Administrators' work

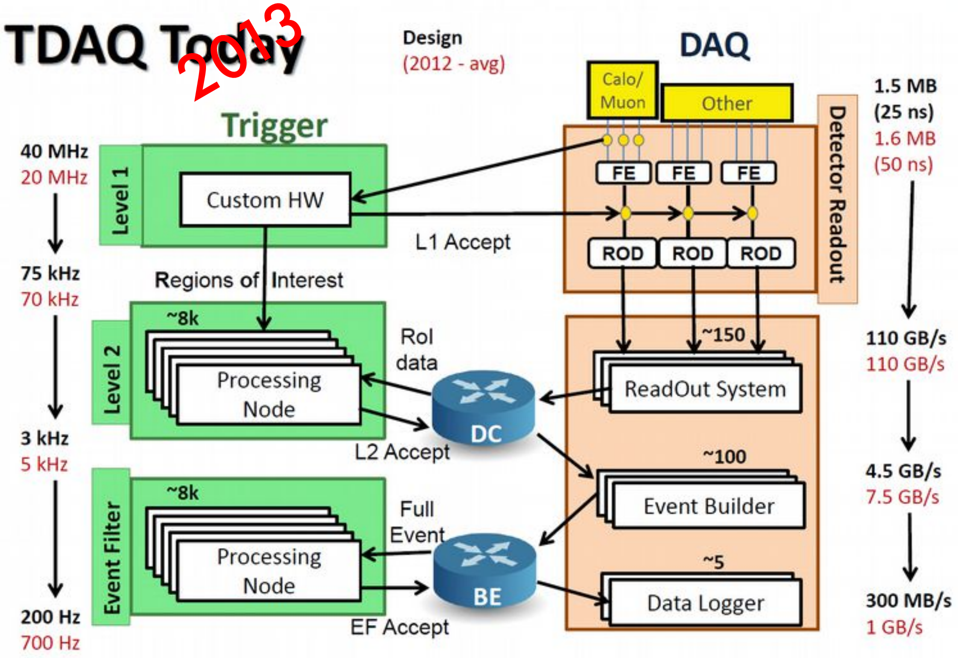
- Installing OS and software
- Configuring services
- Fixing hardware
- Creating user accounts
- Recovering files from backup...

Do you really think it's just that?

ATLAS Online Computing

- Front-ends
- Readout
- Event filters
- Buffer storage
- Detector controls
- Infrastructure
- Control Rooms
- Status publishing
- Access control



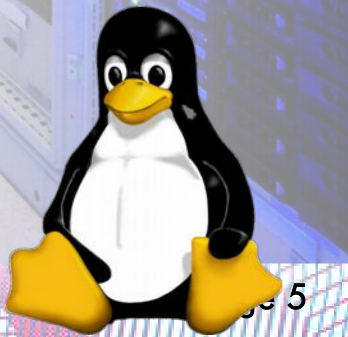


Many things - or not.

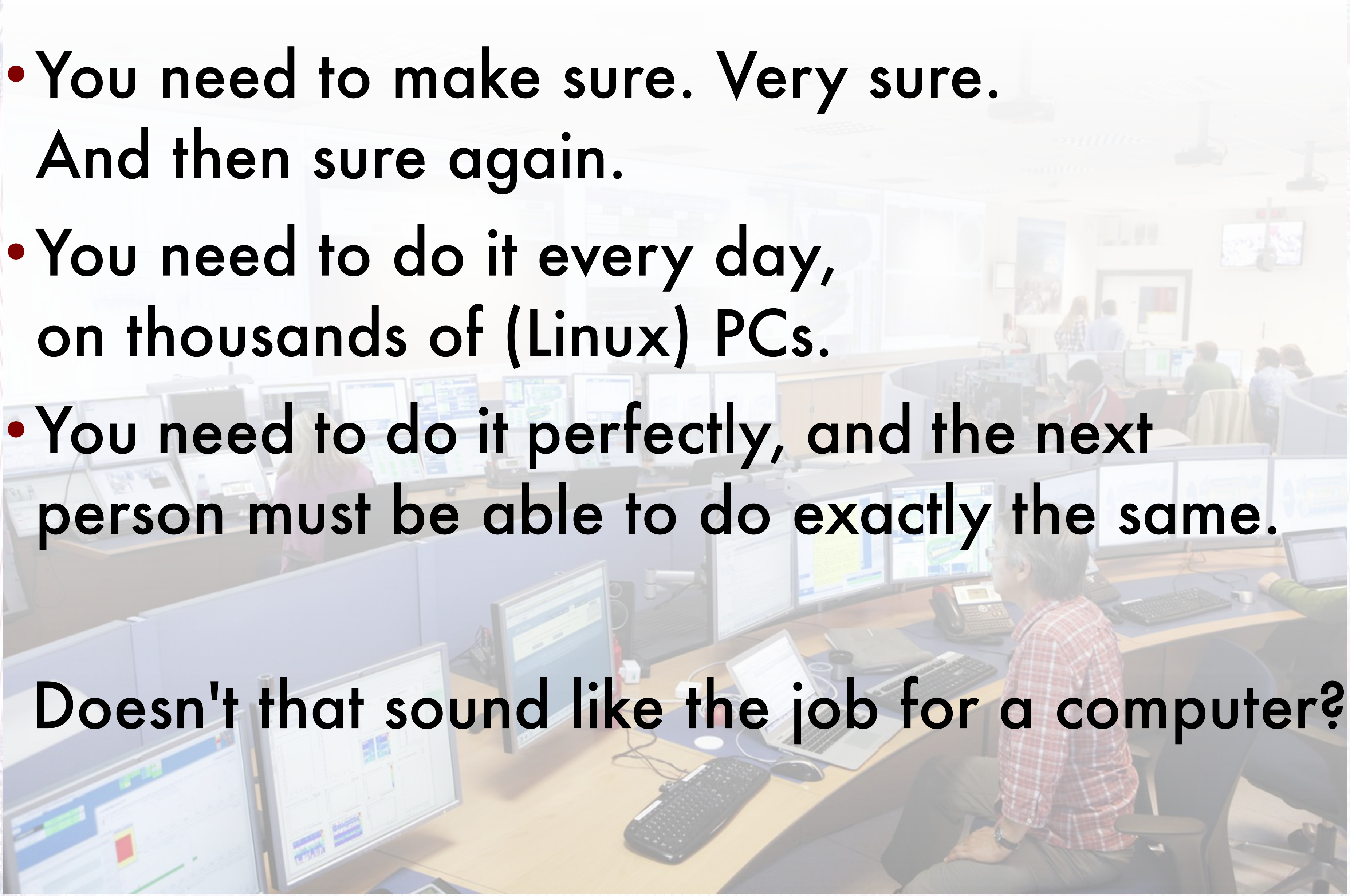
- Plenty of computers? Sure. But forget that.
- In the end, it's one big application, that happens to use one very custom piece of HW.
- That application is ATLAS Online. It consists of many pieces of custom code, different frameworks and goals, but it must work as one.
- That one custom HW is the whole of networks and computers near the ATLAS detector.
- If we just thought about one piece at a time, we could never make it work.

My computer is bigger...

- You probably know what it means to take care of one PC: installing OS, applications, configuring them, see it all works, remember to update, see it all works and the disk is not full...
- Now what if you have 10 PCs with 3 different configurations? "Remember" all the steps to make them the same..
- When you have 3000 with tens of different configs, that is not going to work.



Control Freaks



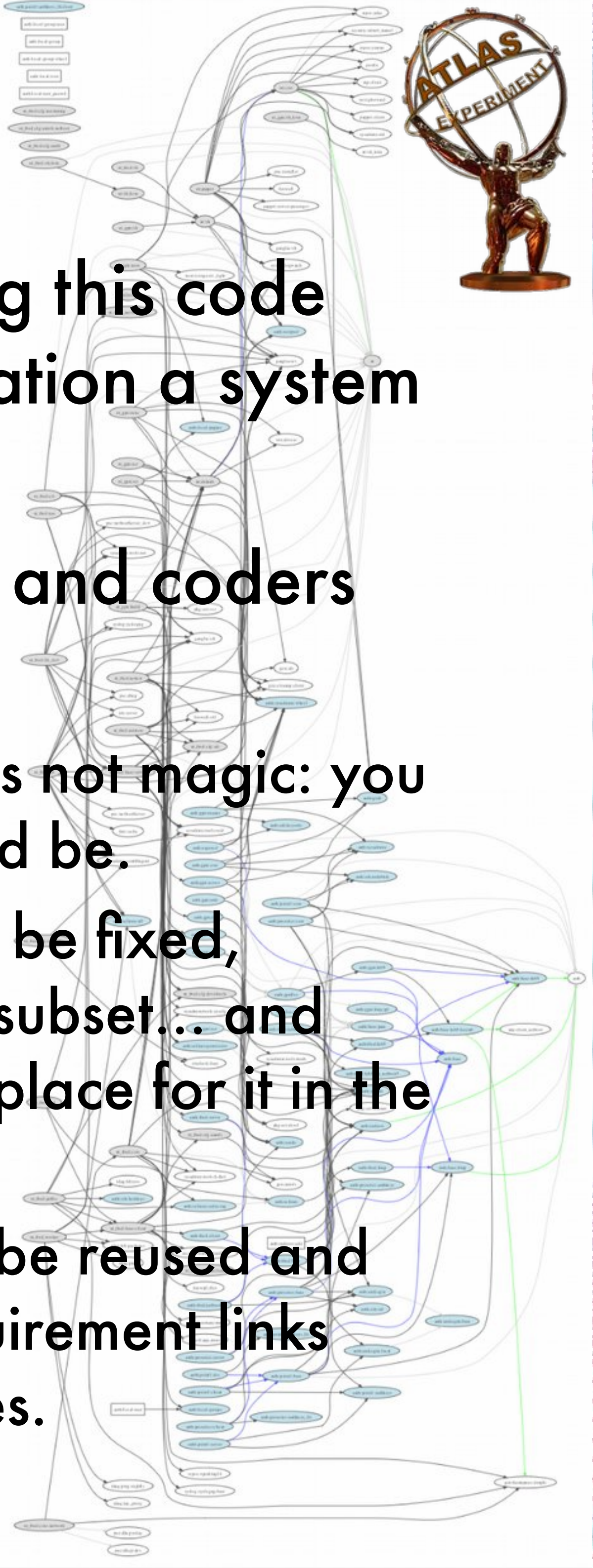
- You need to make sure. Very sure.
And then sure again.
- You need to do it every day,
on thousands of (Linux) PCs.
- You need to do it perfectly, and the next person must be able to do exactly the same.

Doesn't that sound like the job for a computer?

Puppet masters

- Turns out, softwares to do this exist. They are called **Central Configuration Management Systems**.
- They let you describe in detail the configuration of a PC, and take care of enforcing it systematically.
- These "descriptions" are themselves **code**.
  - Puppet uses a Domain Specific Language, that is (mostly) declarative and (kind of) object-oriented
  - ATLAS already has >~~15000~~ 19000 LOCs of Puppet code

Pulling the strings



- Our main job now is writing and maintaining this code that gives the one big ATLAS Online application a system it can run onto.
- We need admins who know (some) coding, and coders who know (some) system administration:
  - Puppet will put the configuration in place, but it's not magic: you still need to know what that configuration should be.
  - You also need to decide if the configuration will be fixed, parametric, applied to one host or all or which subset... and write your code accordingly, and find the right place for it in the class hierarchy.
  - And you need to design it so that the code can be reused and maintained, take care of making the actual requirement links explicit while avoiding a tangle of dependencies.

Look mom, no disk !

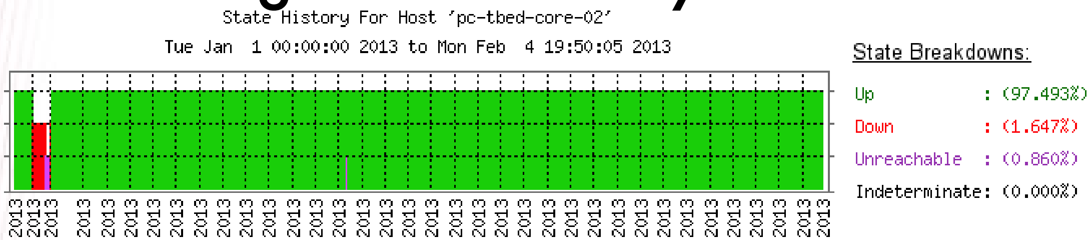
- In ATLAS we make extensive use of PCs with no operating system on disk, “netbooted” via PXE
  - Fastest way to provide single-image systems
  - Easy to switch between OS releases
  - Completely diskless systems have a few less points of failure (disk, controller, cables)
- It's not commonly done

  so it requires ad-hoc development and support
  - The boot configuration is provided by our ConfDB
  - After boot, the single image is specialised using Puppet.

ConfDB GUI
ADMINISTRATIVE INTERFACE

Obsessive Stalkers

- Say finally, everything works fine, now. And tomorrow?
  - Given a sufficient large time and population, everything that can possibly go wrong, will.
  - HW, systems and services need to be checked constantly, making sure that they're functional and healthy.



State History For Host 'pc-tbed-core-02'

Tue Jan 1 00:00:00 2013 to Mon Feb 4 19:50:05 2013

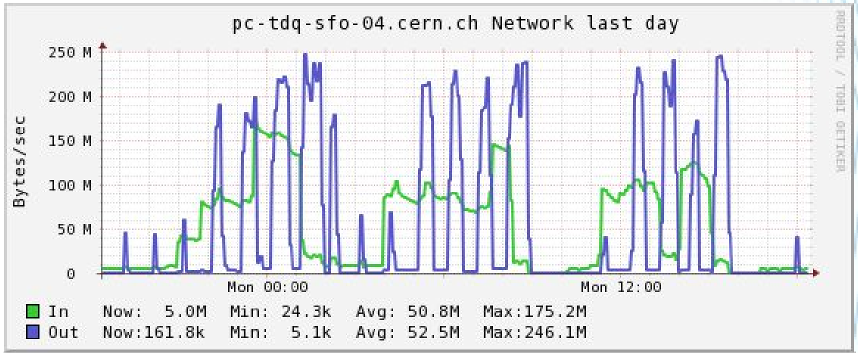
State Breakdowns:

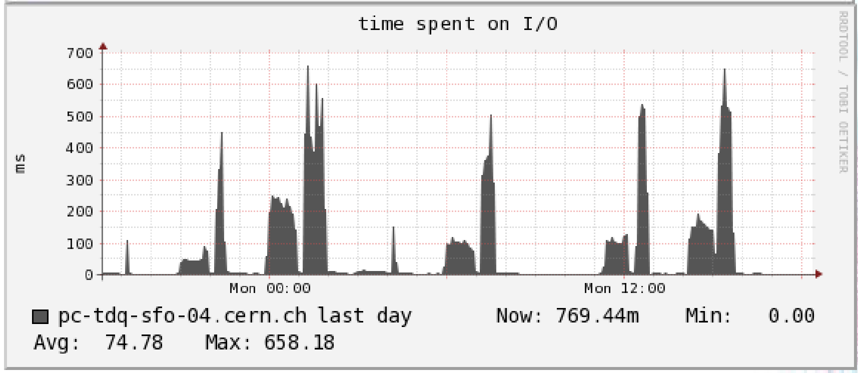
```
Up           : (97.493%)
Down         : (1.647%)
Unreachable  : (0.860%)
Indeterminate: (0.000%)
```

- Large systems need scalable monitoring and alerting tools, like Nagios, Icinga, Ganglia
  - These are good base tools but involve plenty of work on customization, integration, validation.

Performance maniacs

- Beyond making sure that a system works, we often need to know that it works well, and identify possible performance issues.





- Look for bottlenecks, scalability, design errors, guide HW choice
- Help in complex debug

Paranoid worriers

- When a pretty harmless hack touched CERN, it was all over the news.
  We can't just think we're below the radar.
- Isolate the system, but keep remote access and usability; control incoming and outgoing, especially web sites
- Webservers, bastion/gateway hosts, role-based access control, short-lived authorizations, intrusion detection systems, security scanners, Single Sign-On authentication... plenty of fun!!

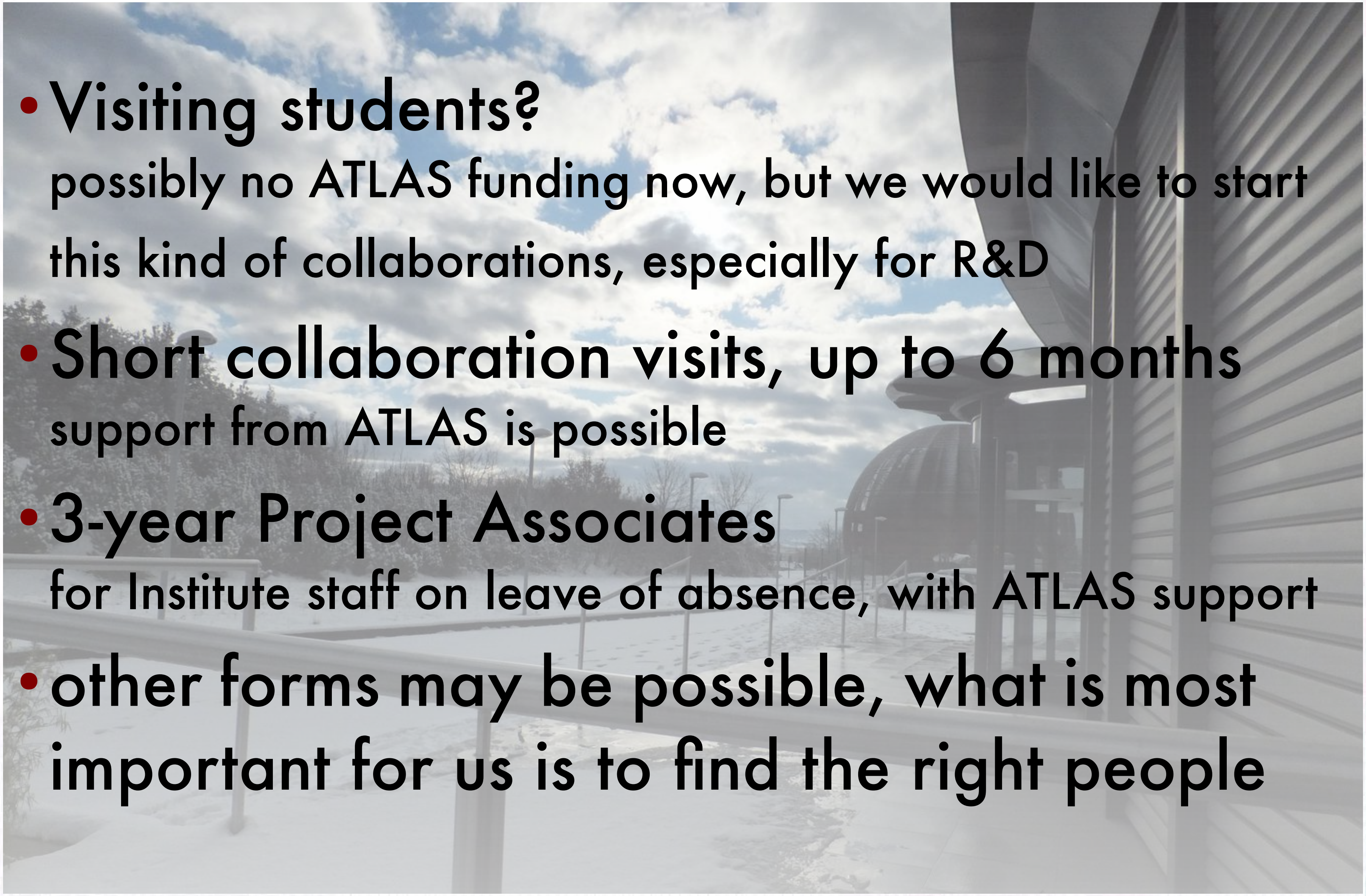
Looking ahead

- Technology changes, more variety - GPUs, many-core, ARM, which will give best performance/watt/space for which tasks?
- Distributed file systems, (maintainable) high-availability tools could allow drastic changes in the architecture
- Virtualization, cloud
  - from test and niche servers to full scale?
  - (ab)use the P1 computing power for other ATLAS tasks

Windows on VMs for DCS

Sim@P1 MonteCarlo production since July 2013

Opportunities @TDAQ SysAdm



- Visiting students?

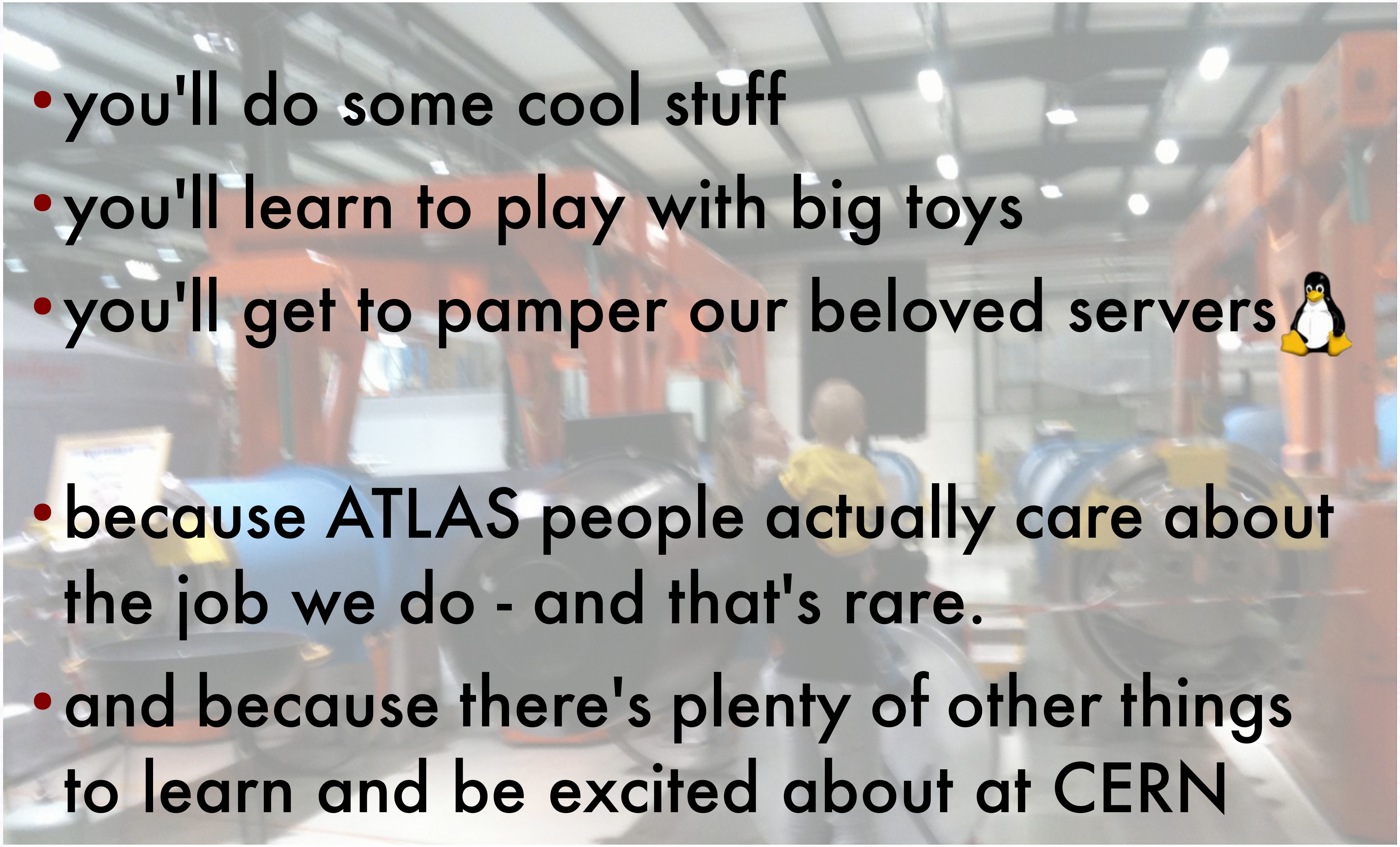
  possibly no ATLAS funding now, but we would like to start this kind of collaborations, especially for R&D
- Short collaboration visits, up to 6 months

  support from ATLAS is possible
- 3-year Project Associates

  for Institute staff on leave of absence, with ATLAS support
- other forms may be possible, what is most important for us is to find the right people

Contact: sergio.ballestrero@cern.ch

So why should you care?



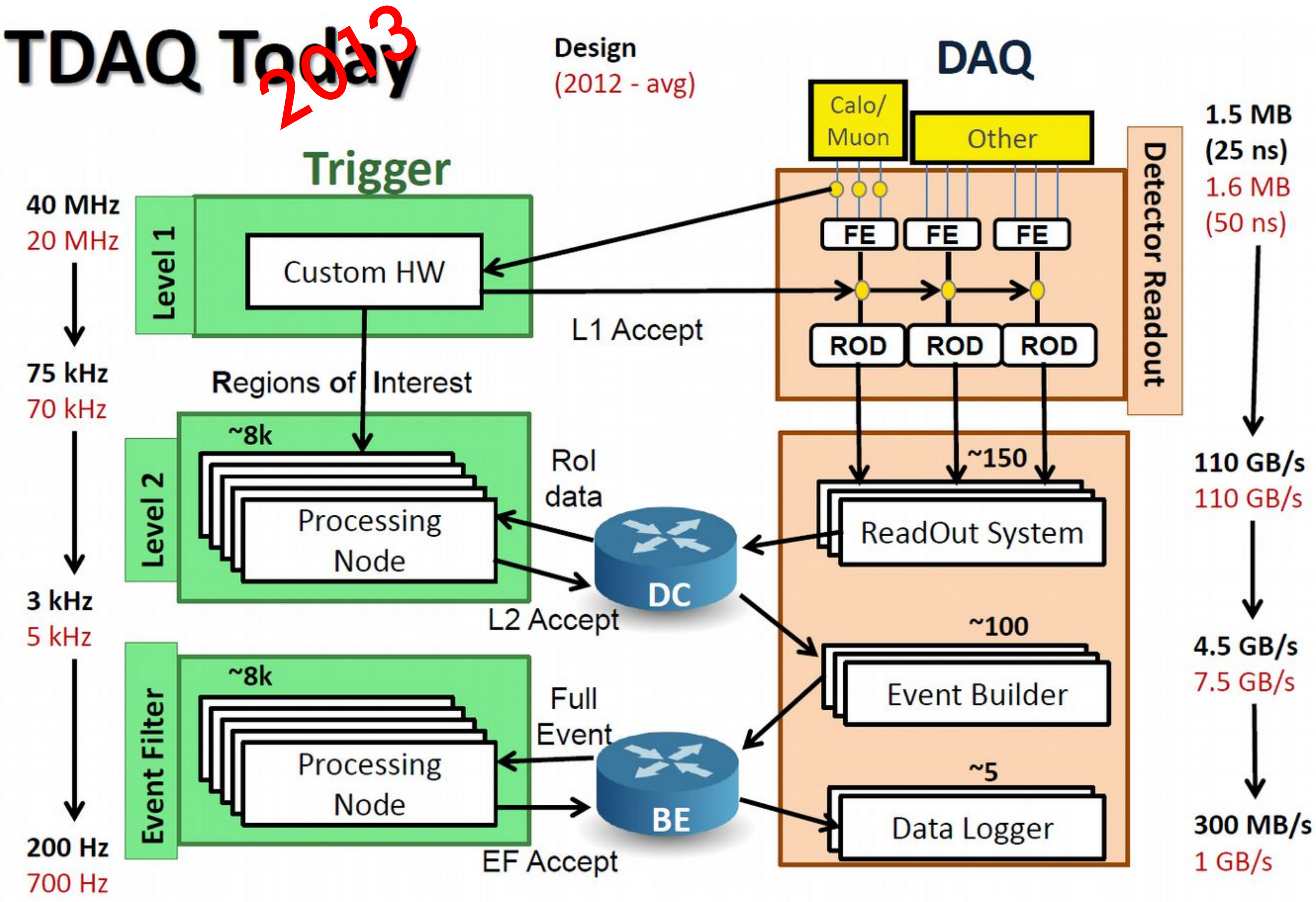
- you'll do some cool stuff
- you'll learn to play with big toys
- you'll get to pamper our beloved servers

- because ATLAS people actually care about the job we do - and that's rare.
- and because there's plenty of other things to learn and be excited about at CERN

More stuff

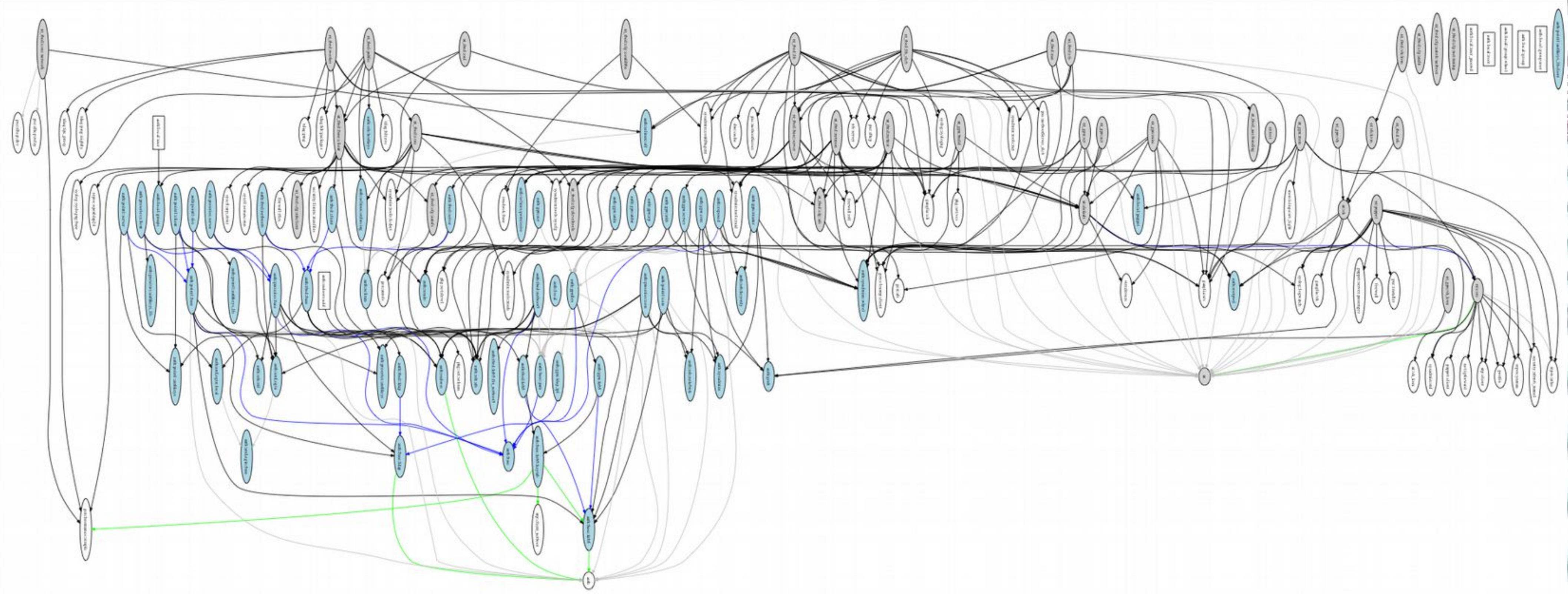
- TDAQ SysAdmins public website: http://atlas-tdaq-sysadmin.web.cern.ch/
  - some general info, publications/talks, links
- Following spare slides on a few aspects

TDAQ Dataflow



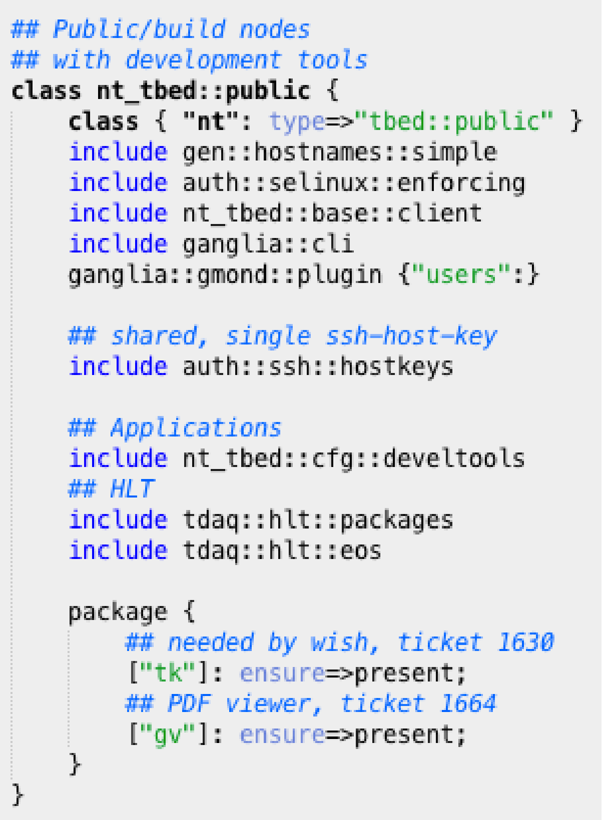
Drawing from N.Garelli

Puppet dependency-graph



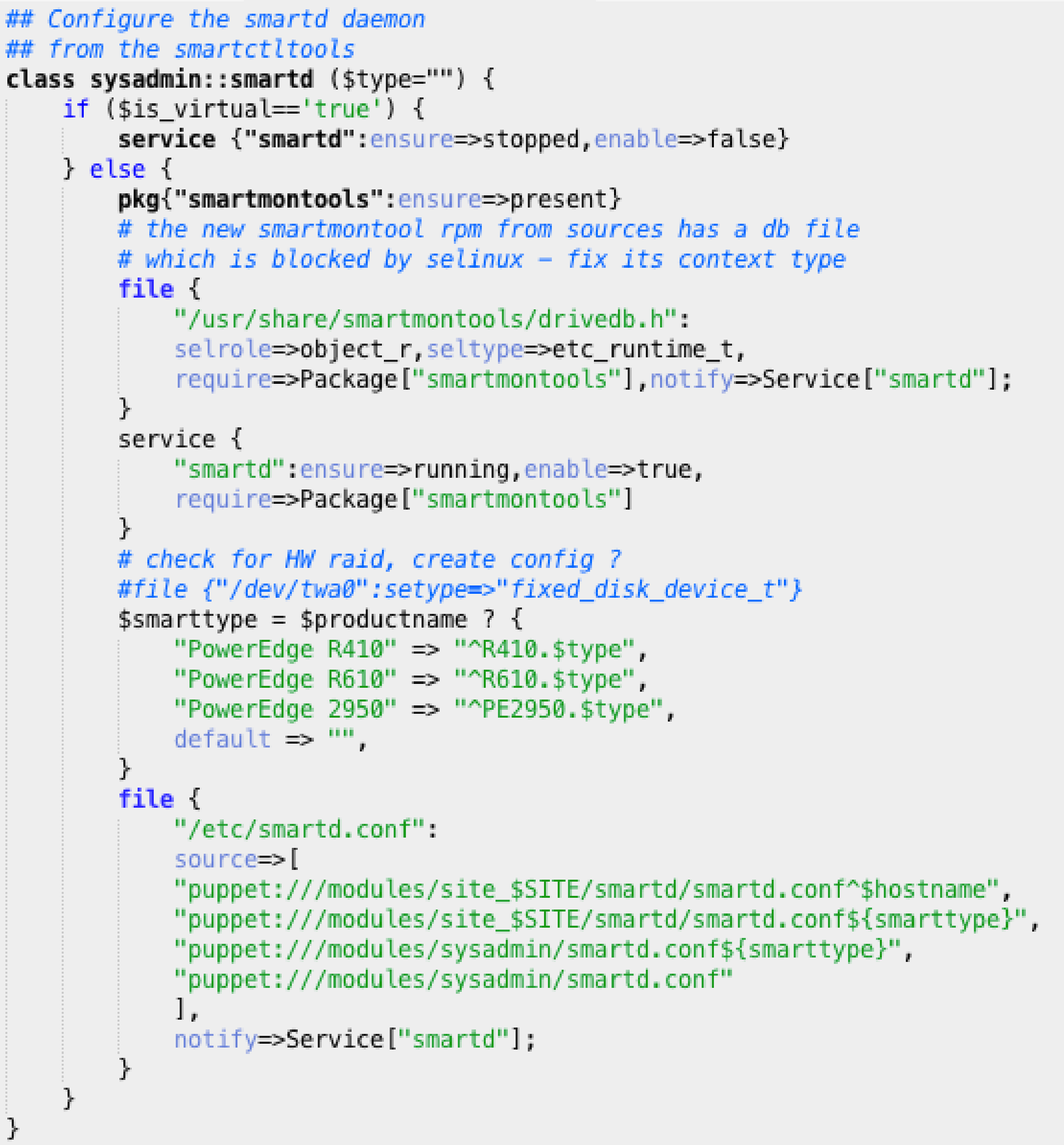
Puppet code samples

Node “type” definition



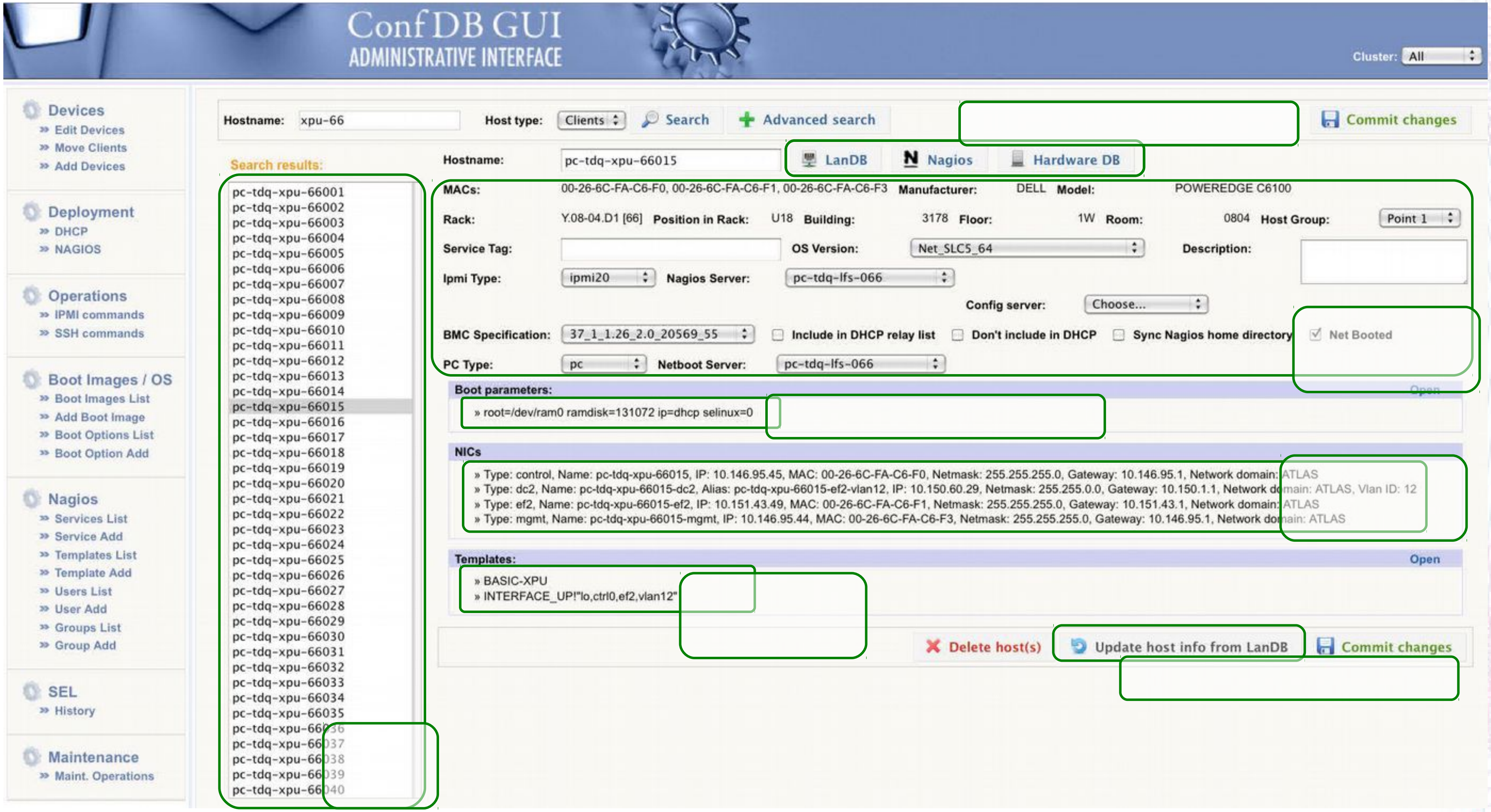
```
## Public/build nodes
## with development tools
class nt_tbed::public {
    class { "nt": type=>"tbed::public" }
    include gen::hostnames::simple
    include auth::selinux::enforcing
    include nt_tbed::base::client
    include ganglia::cli
    ganglia::gmond::plugin {"users":}

    ## shared, single ssh-host-key
    include auth::ssh::hostkeys

    ## Applications
    include nt_tbed::cfg::develtools
    ## HLT
    include tdaq::hlt::packages
    include tdaq::hlt::eos

    package {
        ## needed by wish, ticket 1630
        ["tk"]: ensure=>present;
        ## PDF viewer, ticket 1664
        ["gv"]: ensure=>present;
    }
}
```

Service definition and configuration



```
## Configure the smartd daemon
## from the smartctltools
class sysadmin::smartd ($type="") {
    if ($is_virtual=='true') {
        service {"smartd":ensure=>stopped,enable=>false}
    } else {
        pkg{"smartmontools":ensure=>present}
        # the new smartmontool rpm from sources has a db file
        # which is blocked by selinux - fix its context type
        file {
            "/usr/share/smartmontools/drivedb.h":
            selrole=>object_r,seltype=>etc_runtime_t,
            require=>Package["smartmontools"],notify=>Service["smartd"];
        }
        service {
            "smartd":ensure=>running,enable=>true,
            require=>Package["smartmontools"]
        }
        # check for HW raid, create config ?
        #file {"/dev/twa0":setype=>"fixed_disk_device_t"}
        $smarttype = $productname ? {
            "PowerEdge R410" => "^R410.$type",
            "PowerEdge R610" => "^R610.$type",
            "PowerEdge 2950" => "^PE2950.$type",
            default => "",
        }
        file {
            "/etc/smartd.conf":
            source=>[
            "puppet:///modules/site_$SITE/smartd/smartd.conf^$hostname",
            "puppet:///modules/site_$SITE/smartd/smartd.conf${smarttype}",
            "puppet:///modules/sysadmin/smartd.conf${smarttype}",
            "puppet:///modules/sysadmin/smartd.conf"
            ],
            notify=>Service["smartd"];
        }
    }
}
```

ConfDB GUI



Hostname: xpu-66   Host type: Clients

Search results:

pc-tdq-xpu-66001
pc-tdq-xpu-66002
pc-tdq-xpu-66003
pc-tdq-xpu-66004
pc-tdq-xpu-66005
pc-tdq-xpu-66006
pc-tdq-xpu-66007
pc-tdq-xpu-66008
pc-tdq-xpu-66009
pc-tdq-xpu-66010
pc-tdq-xpu-66011
pc-tdq-xpu-66012
pc-tdq-xpu-66013
pc-tdq-xpu-66014
pc-tdq-xpu-66015
pc-tdq-xpu-66016
pc-tdq-xpu-66017
pc-tdq-xpu-66018
pc-tdq-xpu-66019
pc-tdq-xpu-66020
pc-tdq-xpu-66021
pc-tdq-xpu-66022
pc-tdq-xpu-66023
pc-tdq-xpu-66024
pc-tdq-xpu-66025
pc-tdq-xpu-66026
pc-tdq-xpu-66027
pc-tdq-xpu-66028
pc-tdq-xpu-66029
pc-tdq-xpu-66030
pc-tdq-xpu-66031
pc-tdq-xpu-66032
pc-tdq-xpu-66033
pc-tdq-xpu-66034
pc-tdq-xpu-66035
pc-tdq-xpu-66036
pc-tdq-xpu-66037
pc-tdq-xpu-66038
pc-tdq-xpu-66039
pc-tdq-xpu-66040

Hostname: pc-tdq-xpu-66015

MACs: 00-26-6C-FA-C6-F0, 00-26-6C-FA-C6-F1, 00-26-6C-FA-C6-F3   Manufacturer: DELL   Model: POWEREDGE C6100

Rack: Y.08-04.D1 [66]   Position in Rack: U18   Building: 3178   Floor: 1W   Room: 0804   Host Group: Point 1

Service Tag: OS Version: Net_SLC5_64   Description:

Ipmi Type: ipmi20   Nagios Server: pc-tdq-lfs-066

Config server: Choose...

BMC Specification: 37_1_1.26_2.0_20569_55   □ Include in DHCP relay list   □ Don't include in DHCP   □ Sync Nagios home directory   ☑ Net Booted

PC Type: pc   Netboot Server: pc-tdq-lfs-066

Boot parameters:

» root=/dev/ram0 ramdisk=131072 ip=dhcp selinux=0

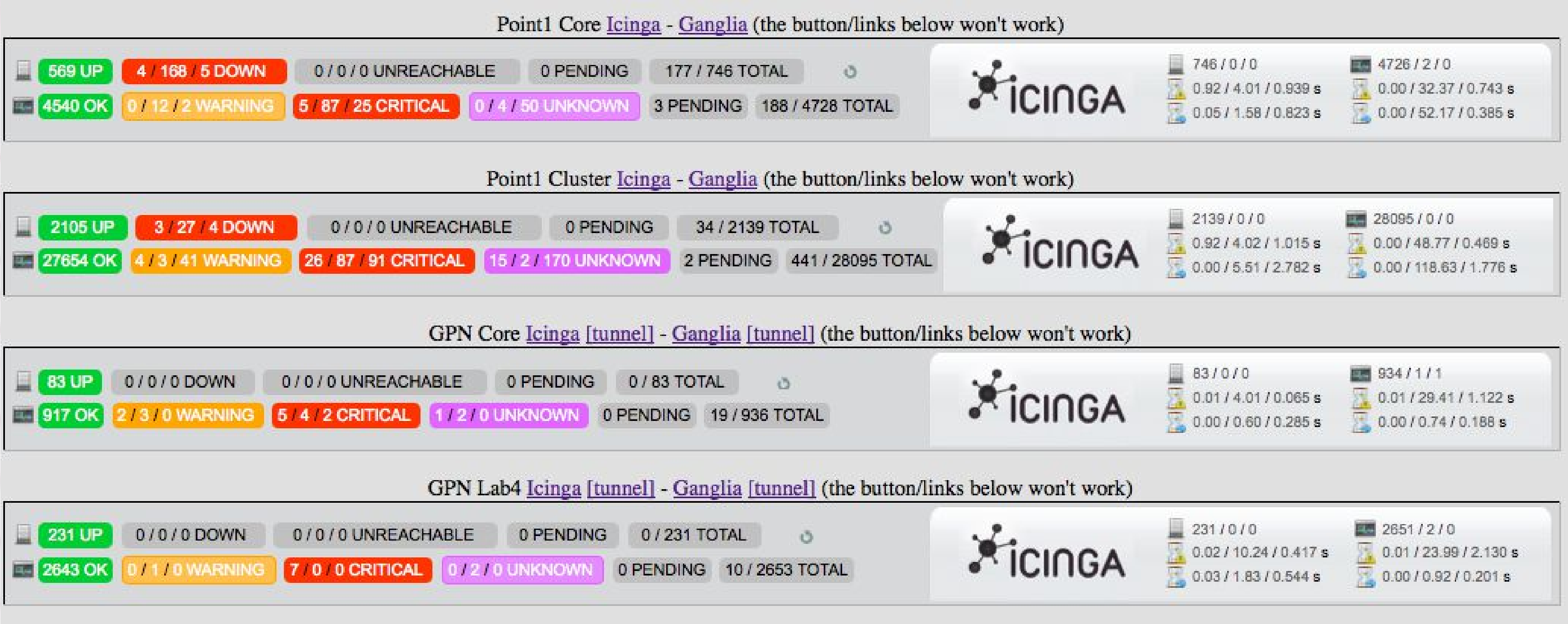
NICs

» Type: control, Name: pc-tdq-xpu-66015, IP: 10.146.95.45, MAC: 00-26-6C-FA-C6-F0, Netmask: 255.255.255.0, Gateway: 10.146.95.1, Network domain: ATLAS
» Type: dc2, Name: pc-tdq-xpu-66015-dc2, Alias: pc-tdq-xpu-66015-ef2-vlan12, IP: 10.150.60.29, Netmask: 255.255.0.0, Gateway: 10.150.1.1, Network domain: ATLAS, Vlan ID: 12
» Type: ef2, Name: pc-tdq-xpu-66015-ef2, IP: 10.151.43.49, MAC: 00-26-6C-FA-C6-F1, Netmask: 255.255.255.0, Gateway: 10.151.43.1, Network domain: ATLAS
» Type: mgmt, Name: pc-tdq-xpu-66015-mgmt, IP: 10.146.95.44, MAC: 00-26-6C-FA-C6-F3, Netmask: 255.255.255.0, Gateway: 10.146.95.1, Network domain: ATLAS

Templates:

» BASIC-XPU
» INTERFACE_UP!"lo,ctrl0,ef2,vlan12"

Icinga: Overall view



Point1 Core Icinga - Ganglia (the button/links below won't work)

569 UP | 4 / 168 / 5 DOWN | 0 / 0 / 0 UNREACHABLE | 0 PENDING | 177 / 746 TOTAL

4540 OK | 0 / 12 / 2 WARNING | 5 / 87 / 25 CRITICAL | 0 / 4 / 50 UNKNOWN | 3 PENDING | 188 / 4728 TOTAL

ICINGA

746 / 0 / 0 | 4726 / 2 / 0
0.92 / 4.01 / 0.939 s | 0.00 / 32.37 / 0.743 s
0.05 / 1.58 / 0.823 s | 0.00 / 52.17 / 0.385 s

Point1 Cluster Icinga - Ganglia (the button/links below won't work)

2105 UP | 3 / 27 / 4 DOWN | 0 / 0 / 0 UNREACHABLE | 0 PENDING | 34 / 2139 TOTAL

27654 OK | 4 / 3 / 41 WARNING | 26 / 87 / 91 CRITICAL | 15 / 2 / 170 UNKNOWN | 2 PENDING | 441 / 28095 TOTAL

ICINGA

2139 / 0 / 0 | 28095 / 0 / 0
0.92 / 4.02 / 1.015 s | 0.00 / 48.77 / 0.469 s
0.00 / 5.51 / 2.782 s | 0.00 / 118.63 / 1.776 s

GPN Core Icinga [tunnel] - Ganglia [tunnel] (the button/links below won't work)

83 UP | 0 / 0 / 0 DOWN | 0 / 0 / 0 UNREACHABLE | 0 PENDING | 0 / 83 TOTAL

917 OK | 2 / 3 / 0 WARNING | 5 / 4 / 2 CRITICAL | 1 / 2 / 0 UNKNOWN | 0 PENDING | 19 / 936 TOTAL

ICINGA

83 / 0 / 0 | 934 / 1 / 1
0.01 / 4.01 / 0.065 s | 0.01 / 29.41 / 1.122 s
0.00 / 0.60 / 0.285 s | 0.00 / 0.74 / 0.188 s

GPN Lab4 Icinga [tunnel] - Ganglia [tunnel] (the button/links below won't work)

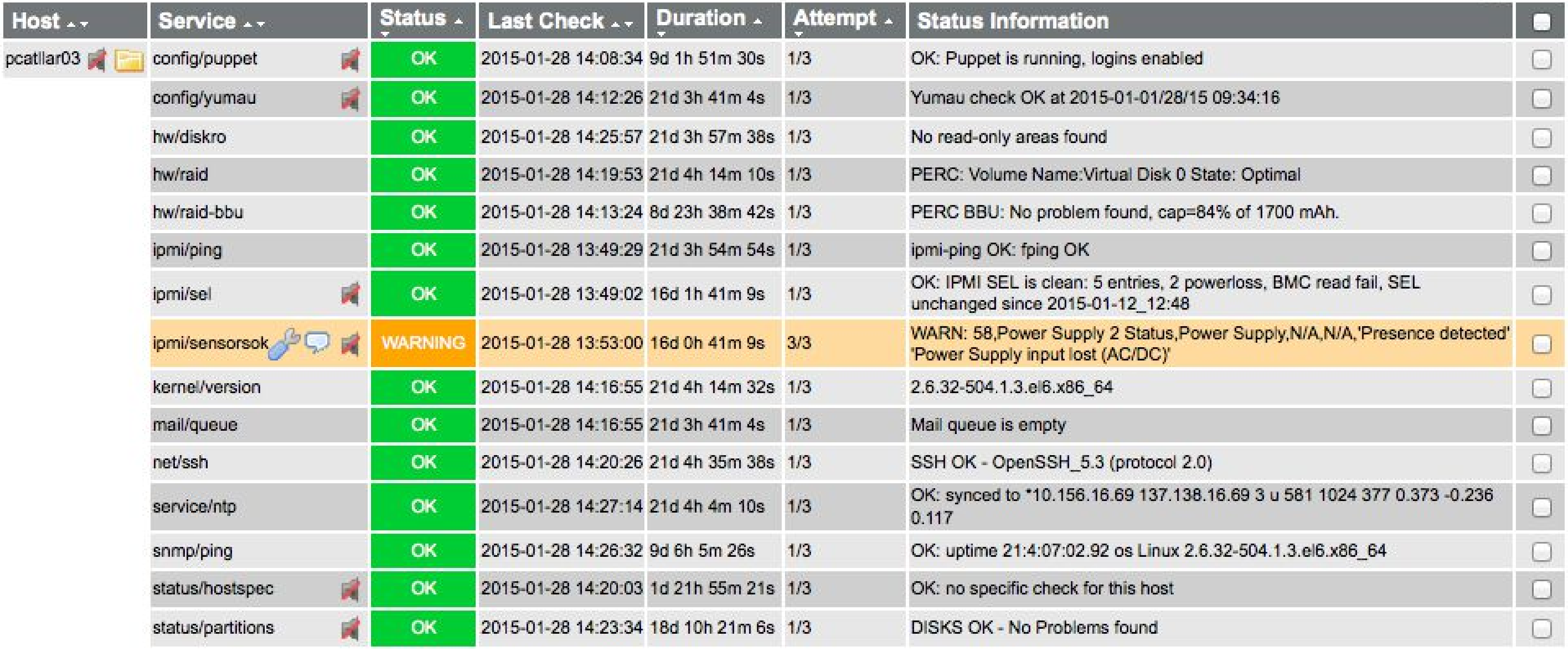
231 UP | 0 / 0 / 0 DOWN | 0 / 0 / 0 UNREACHABLE | 0 PENDING | 0 / 231 TOTAL

2643 OK | 0 / 1 / 0 WARNING | 7 / 0 / 0 CRITICAL | 0 / 2 / 0 UNKNOWN | 0 PENDING | 10 / 2653 TOTAL

ICINGA

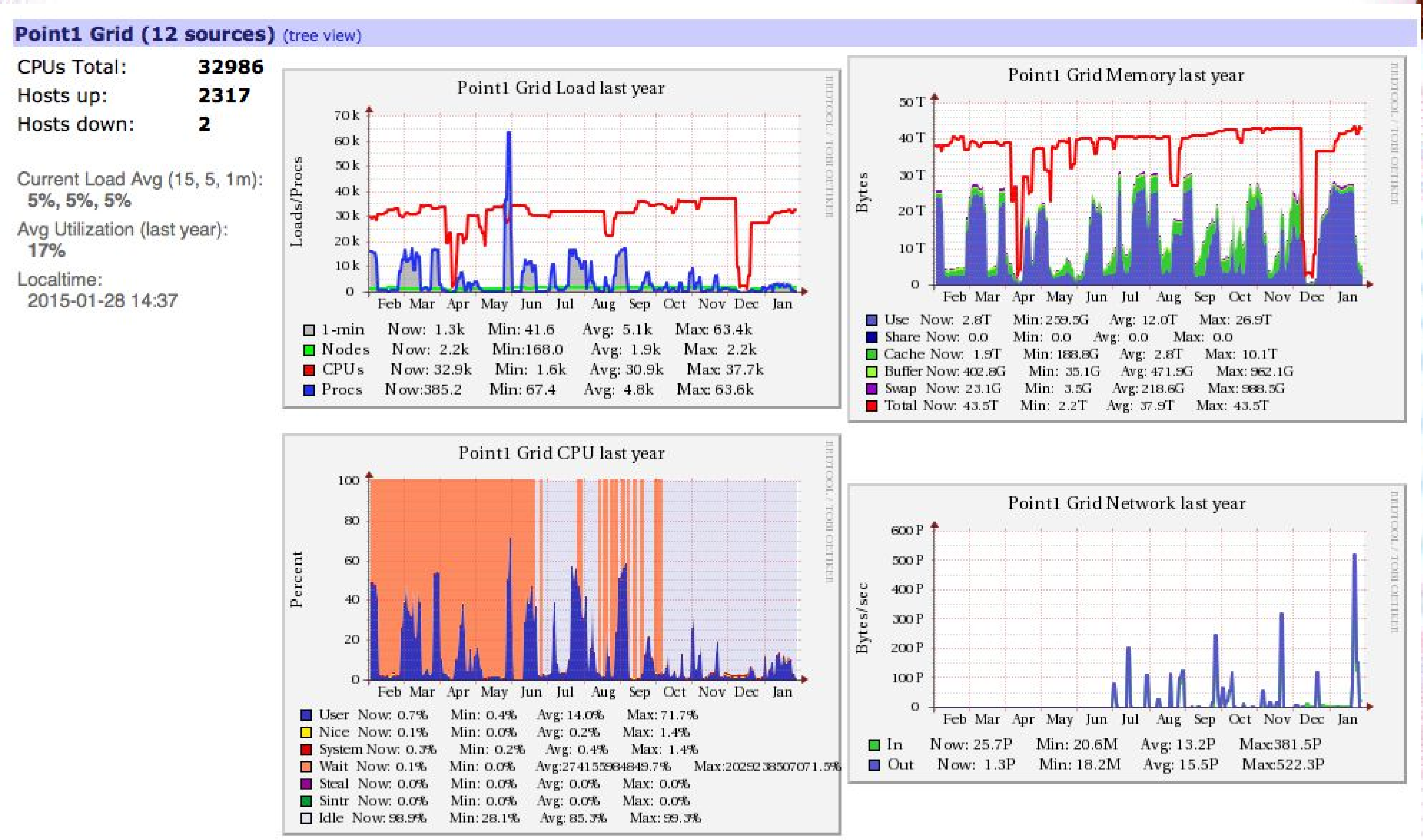
231 / 0 / 0 | 2651 / 2 / 0
0.02 / 10.24 / 0.417 s | 0.01 / 23.99 / 2.130 s
0.03 / 1.83 / 0.544 s | 0.00 / 0.92 / 0.201 s

Icinga: ServiceCheck for a host



Host	Service	Status	Last Check	Duration	Attempt	Status Information	□
pcatllar03	config/puppet	OK	2015-01-28 14:08:34	9d 1h 51m 30s	1/3	OK: Puppet is running, logins enabled	□
	config/yumau	OK	2015-01-28 14:12:26	21d 3h 41m 4s	1/3	Yumau check OK at 2015-01-01/28/15 09:34:16	□
	hw/diskro	OK	2015-01-28 14:25:57	21d 3h 57m 38s	1/3	No read-only areas found	□
	hw/raid	OK	2015-01-28 14:19:53	21d 4h 14m 10s	1/3	PERC: Volume Name:Virtual Disk 0 State: Optimal	□
	hw/raid-bbu	OK	2015-01-28 14:13:24	8d 23h 38m 42s	1/3	PERC BBU: No problem found, cap=84% of 1700 mAh.	□
	ipmi/ping	OK	2015-01-28 13:49:29	21d 3h 54m 54s	1/3	ipmi-ping OK: fping OK	□
	ipmi/sel	OK	2015-01-28 13:49:02	16d 1h 41m 9s	1/3	OK: IPMI SEL is clean: 5 entries, 2 powerloss, BMC read fail, SEL unchanged since 2015-01-12_12:48	□
	ipmi/sensorsok	WARNING	2015-01-28 13:53:00	16d 0h 41m 9s	3/3	WARN: 58,Power Supply 2 Status,Power Supply,N/A,N/A,'Presence detected' 'Power Supply input lost (AC/DC)'	□
	kernel/version	OK	2015-01-28 14:16:55	21d 4h 14m 32s	1/3	2.6.32-504.1.3.el6.x86_64	□
	mail/queue	OK	2015-01-28 14:16:55	21d 3h 41m 4s	1/3	Mail queue is empty	□
	net/ssh	OK	2015-01-28 14:20:26	21d 4h 35m 38s	1/3	SSH OK - OpenSSH_5.3 (protocol 2.0)	□
	service/ntp	OK	2015-01-28 14:27:14	21d 4h 4m 10s	1/3	OK: synced to *10.156.16.69 137.138.16.69 3 u 581 1024 377 0.373 -0.236 0.117	□
	snmp/ping	OK	2015-01-28 14:26:32	9d 6h 5m 26s	1/3	OK: uptime 21:4:07:02.92 os Linux 2.6.32-504.1.3.el6.x86_64	□
	status/hostspec	OK	2015-01-28 14:20:03	1d 21h 55m 21s	1/3	OK: no specific check for this host	□
	status/partitions	OK	2015-01-28 14:23:34	18d 10h 21m 6s	1/3	DISKS OK - No Problems found	□

Ganglia: all ATLAS Point1



Point1 Grid (12 sources) (tree view)

CPUs Total: **32986**
Hosts up: **2317**
Hosts down: **2**

Current Load Avg (15, 5, 1m):
5%, 5%, 5%
Avg Utilization (last year):
17%
Localtime:
2015-01-28 14:37

Point1 Grid Load last year

```
□ 1-min   Now:  1.3k   Min: 41.6    Avg:  5.1k   Max: 63.4k
□ Nodes   Now:  2.2k   Min:168.0    Avg:  1.9k   Max:  2.2k
□ CPUs    Now: 32.9k   Min:  1.6k   Avg: 30.9k   Max: 37.7k
□ Procs   Now:385.2    Min: 67.4    Avg:  4.8k   Max: 63.6k
```

Point1 Grid Memory last year

```
□ Use    Now:  2.8T    Min: 259.5G   Avg:  12.0T   Max:  26.9T
□ Share  Now:  0.0     Min:   0.0    Avg:   0.0    Max:   0.0
□ Cache  Now:  1.9T    Min: 188.8G   Avg:   2.8T   Max:  10.1T
□ Buffer Now: 402.8G   Min:  35.1G   Avg: 471.9G   Max: 962.1G
□ Swap   Now:  23.1G   Min:   3.5G   Avg: 218.6G   Max: 988.5G
□ Total  Now:  43.5T   Min:   2.2T   Avg:  37.9T   Max:  43.5T
```

Point1 Grid CPU last year

```
□ User   Now: 0.7%   Min: 0.4%    Avg: 14.0%   Max: 71.7%
□ Nice   Now: 0.1%   Min: 0.0%    Avg:  0.2%   Max:  1.4%
□ System Now: 0.3%   Min: 0.2%    Avg:  0.4%   Max:  1.4%
□ Wait   Now: 0.1%   Min: 0.0%    Avg:27415584849.7%   Max:20292385070071.5%
□ Steal  Now: 0.0%   Min: 0.0%    Avg:  0.0%   Max:  0.0%
□ Sintr  Now: 0.0%   Min: 0.0%    Avg:  0.0%   Max:  0.0%
□ Idle   Now:98.9%   Min:28.1%    Avg: 85.3%   Max: 99.3%
```

Point1 Grid Network last year

```
□ In    Now: 25.7P   Min: 20.6M   Avg: 13.2P   Max:381.5P
□ Out   Now:  1.3P   Min: 18.2M   Avg: 15.5P   Max:522.3P
```

LocalBoot SLC5

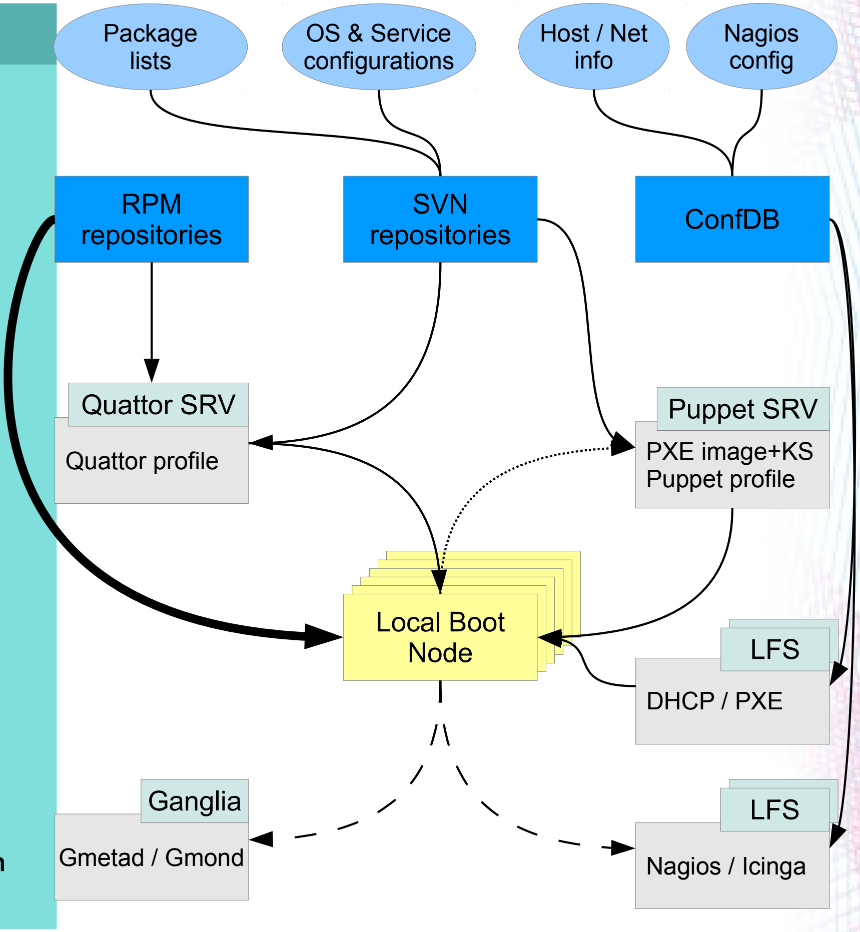
Local Boot nodes

Provisioning by PXE + KickStart
- DHCP+PXE provided by an LFS, from ConfDB info
- Kickstart files generated by template-based system

Quattor
- CERN standard Configuration Management Tool
- Production system, managing 237 hosts in the Online Farm
- Tight control on installed packages
- Lack of flexibility for complex configuration/service dependencies
- Multiple languages for implementing modules

Puppet
- Widespread industry adoption, active development
- Full features, high flexibility
- Gentler learning curve
- Focus on consistency and idempotence
- In production, manages exclusively 25 complex servers and complements Quattor on the remaining 237
- Planned to completely replace Quattor on SLC6

For both Quattor and Puppet the configuration code is maintained in a Revision Control System (Subversion).



LocalBoot SLC6

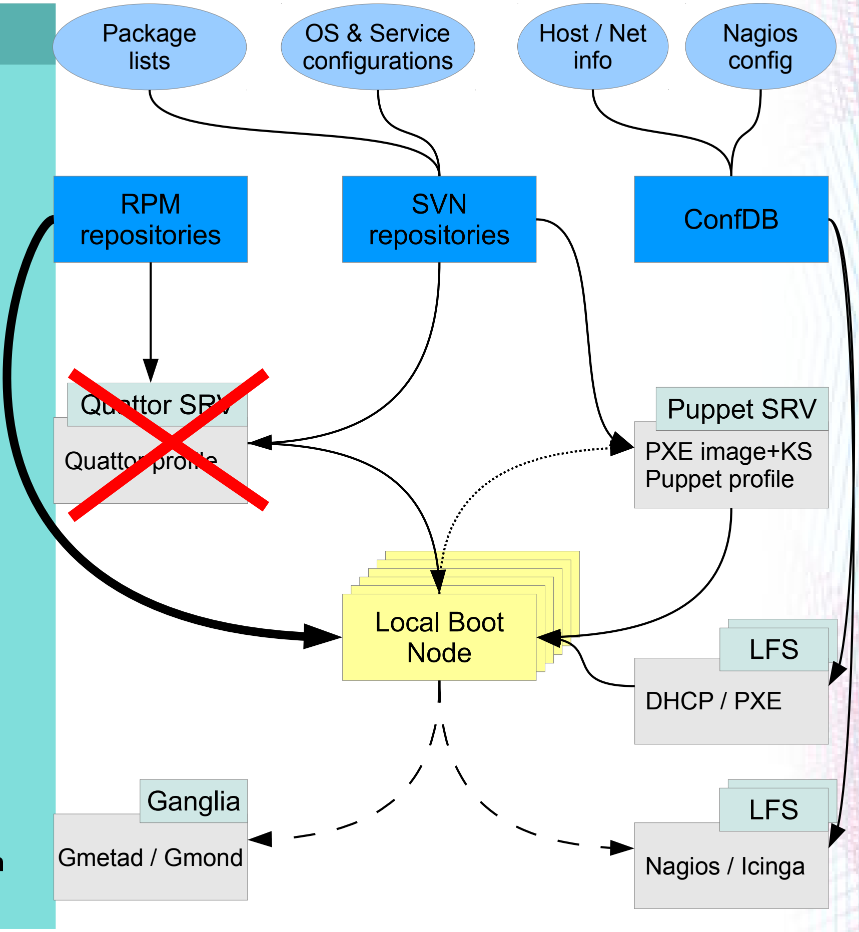
Local Boot nodes

Provisioning by PXE + KickStart
- DHCP+PXE provided by an LFS, from ConfDB info
- Kickstart files generated by template-based system

Quattor
- ~~CERN standard Configuration Management Tool~~
- ~~Production system, managing 237 hosts in the Online Farm~~
- ~~Tight control on installed packages~~
- ~~Lack of flexibility for complex configuration/service dependencies~~
- ~~Multiple languages for implementing modules~~

Puppet
- Widespread industry adoption, active development
- Full features, high flexibility
- Gentler learning curve
- Focus on consistency and idempotence
- In production, manages exclusively ~~25 complex servers and complements Quattor on the remaining 237~~
- ~~Planned to~~ completely replace Quattor on SLC6

For both Quattor and Puppet the configuration code is maintained in a Revision Control System (Subversion).



NetBoot SLC5

NetBooted Nodes

- ~2350 nodes boot the Scientific Linux CERN 5 OS via PXE
- ~80 Local File Server (LFS) hosts provide DHCP, PXE, TFTP for booting, `/usr` read-only directory via NFS.
- Configuration of DHCP, PXE and boot parameters provided by ConfDB, our CMT for NetBoot nodes which is described in a separate poster
  (Centralized configuration system for a large scale farm of network booted computers)

Boot With Me tool

- Generates PXE boot images (kernel + RAMdisk root) and `/usr`
- Uses a reference SLC5 VM image as source

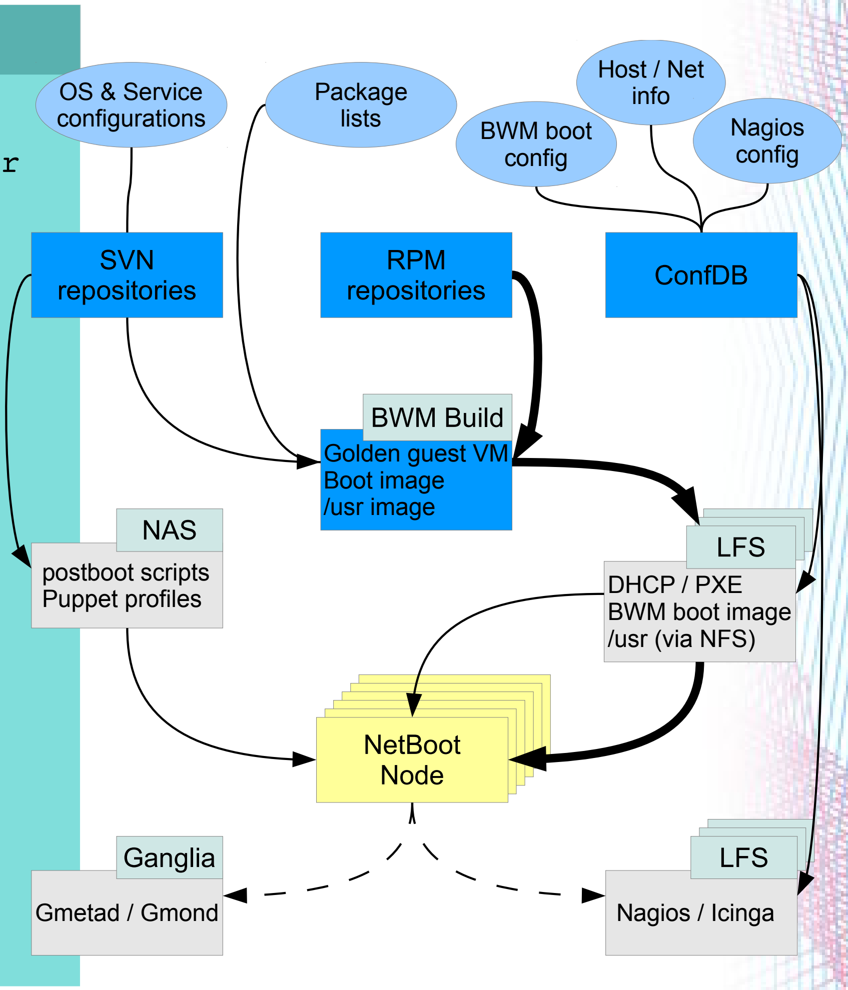
BWM post-boot script system

- Hierarchy of shell scripts, configures services, disk and NFS mounts etc
- Uses the standardised hostname to decide which sequence of scripts
- Stored on central Network-Attached Storage, executed by the client

BWM puppet

- Start to introduce Puppet profiles to replace BWM scripts
- Improve consistency and maintainability
- Serverless configuration, for scalability, using the NAS as storage

A Subversion repository is used to track changes of the BWM image creation configuration, of the post-boot scripts and Puppet profiles.



NetBoot SLC6

NetBooted Nodes

- ~2350 nodes boot the Scientific Linux CERN ~~5~~6 OS via PXE
- ~80 Local File Server (LFS) hosts provide DHCP, PXE, TFTP for booting, `/usr` read-only directory via NFS.
- Configuration of DHCP, PXE and boot parameters provided by ConfDB, our CMT for NetBoot nodes which is described in a separate poster
  (Centralized configuration system for a large scale farm of network booted computers)

~~Boot With Me tool~~

- ~~Generates PXE boot images (kernel + RAMdisk root) and `/usr`~~
- ~~Uses a reference SLC5 VM image as source~~

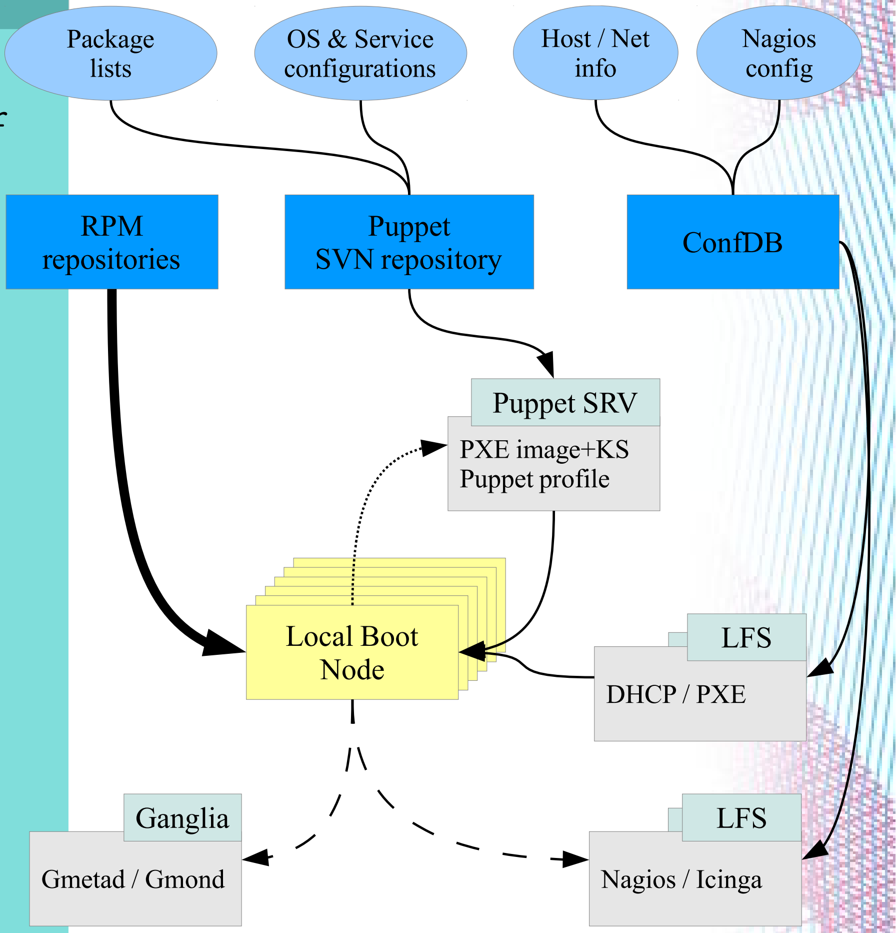
BWM post-boot script system

- ~~Hierarchy of shell scripts, configures services, disk and NFS mounts etc~~
- ~~Uses the standardised hostname to decide which sequence of scripts~~
- ~~Stored on central Network-Attached Storage, executed by the client~~

BWM puppet

- ~~Start to introduce~~ Puppet profiles to replace BWM scripts
- Improve consistency and maintainability
- Serverless configuration, for scalability, using the ~~NAS~~ CFS as storage

A Subversion repository is used to track changes of the ~~BWM image creation configuration, of the post-boot scripts and~~ Puppet profiles.


Package lists
OS & Service configurations
Host / Net info
Nagios config
RPM repositories
Puppet SVN repository
ConfDB
Puppet SRV
PXE image+KS
Puppet profile
Local Boot Node
LFS
DHCP / PXE
Ganglia
Gmetad / Gmond
LFS
Nagios / Icinga