

System Administration Basics: ATLAS

Florentin BUJOR
on behalf of ATLAS TDAQ SysAdmins



Outline

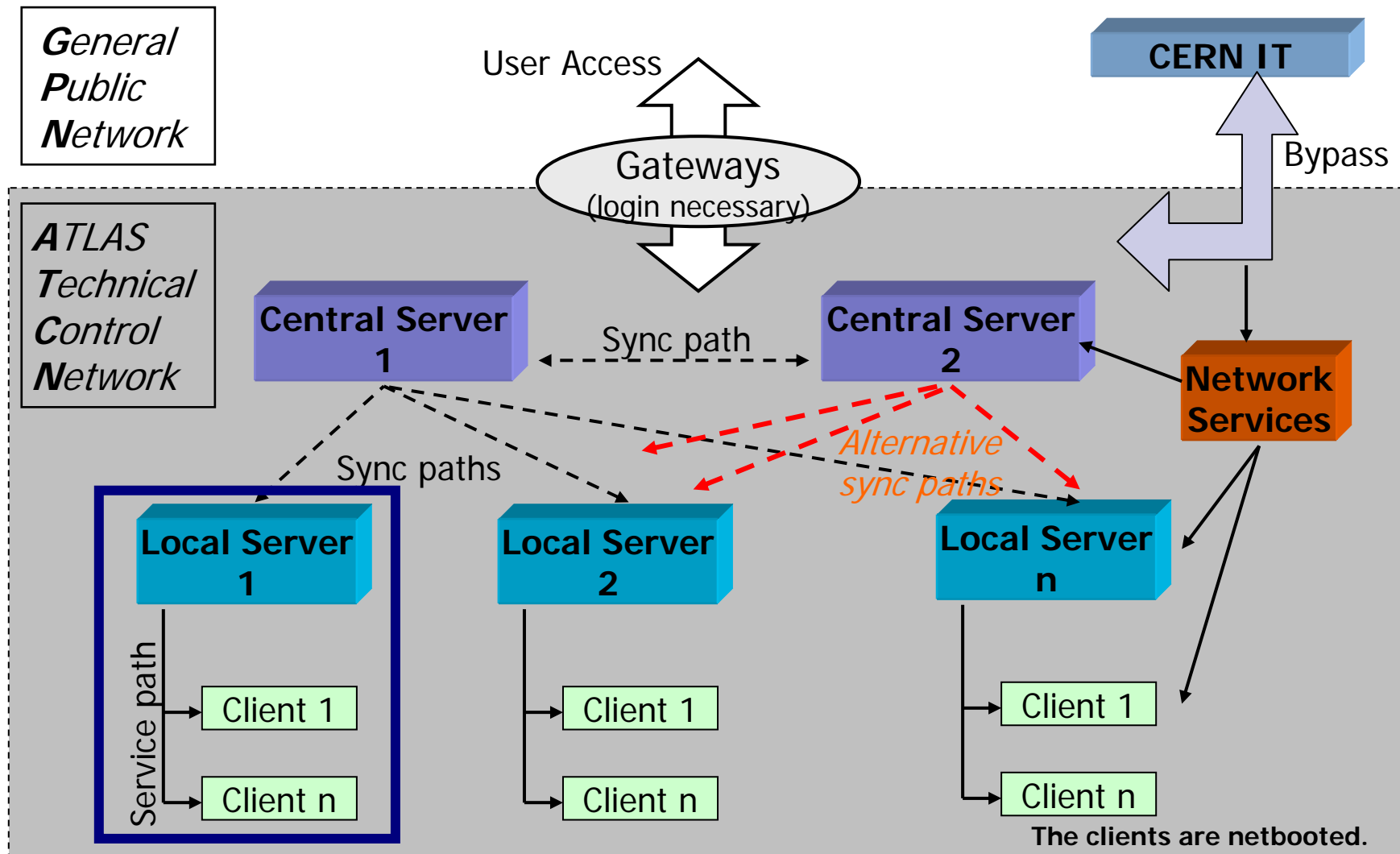
- SysAdmin Group Mandate
- Point 1 System Layout
- Provided services
- Developed tools



ATLAS TDAQ SysAdmin mandate

- Provide support for:
 - Deployment of ATLAS online farms in Point 1
 - ~ 2500 machines (gateways, service servers, central and local servers, netbooted machines, public machines, laptops, etc)
 - Test beds
 - ~ 100 nodes netbooted identical to Point 1, with uniform environment across all test beds
 - Differences with respect to Point 1:
 - Accessible directly from CERN Public Network (i.e. no gateway)
 - AFS available
- User support
 - FAQ web site
 - Trouble ticket system

Point 1 System Layout





Point 1 System : current status (1/2)

- 2 of n Gateway machines + Windows Terminal Service
- 2 Network service machines
- 2 of 4 Central File Servers
- 11 of ~100 Local File Servers
 - +8 to install for final system
 - Base config for each rack: 1 LFS and ~30 clients
- ~450 of ~2500 client nodes
 - 72 in preseries (prototype cluster)
 - the rest are detectors and final system
- 8 TDAQ laptops as mobile control stations

Point 1 Network Layout: status (2/2)

- 11 public machines
- Nodes for ATLAS Control Room
 - 6 PCs with 4 monitors each installed with SLC4
- UPS for essential services
 - Network starpoints
 - CFSs, gateways, network services
 - Essential DCS/DSS





Provided services (1/7)

- Gateway & Firewall services
- IP numbers & names using IT services (LanDB)
- Network time service, DNS cache
- Net boot : both PXE and BOOTP technology
- Post-Boot configuration of the nodes
- File servers
 - Space for data and ATLAS software
 - Home directories (backed up by Tivoli)
- Link to IT services (CASTOR, CVS, DataBases ...)
- Cluster monitoring with Nagios
 - Status of monitored services (new plugins)
 - Alarms via email/SMS
 - Time series graphs: voltages, temperatures, fans speed, etc...

Provided services: Nagios (2/7)

Current Network Status
 Last Updated: Fri Oct 7 17:10:36 CEST 2005
 Updated every 90 seconds
 Nagios@ - www.nagios.org
 Logged in as *nagiosadmin*

Host Status Totals

Up	Down	Unreachable	Pending
118	37	0	0

Service Status Totals

Ok	Warning	Unknown	Critical	Pending
504	212	17	79	0

All Problems | **All Types**

37	155
----	-----

All Problems | **All Types**

308	812
-----	-----

Status Summary For All Host Groups

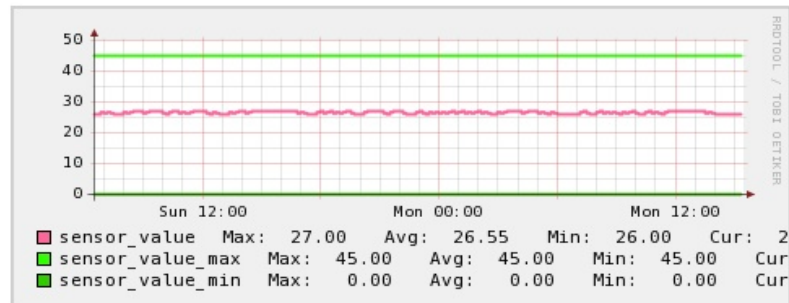
Host Group	Host Status Totals	Service Status Totals
ROS pROS L2PUs L2SVs DFM SFI (atlas-hosts_B32)	58 UP 5 DOWN	290 OK 126 WARNING 7 UNKNOWN 18 CRITICAL
ATLAS Hosts (Bdq_40) (atlas-hosts_B40)	8 UP	32 OK 16 WARNING 8 UNKNOWN
ATLAS Hosts (Bdq_513) (atlas-hosts_B513)	5 UP 2 DOWN	10 OK 4 CRITICAL
ATLAS Hosts (Bdq_513) netbooted (atlas-hosts_B513_nb)	28 UP 2 DOWN	135 OK 70 WARNING 5 CRITICAL
Cluster Hosts (cluster-hosts)	8 UP 14 DOWN	16 OK 2 UNKNOWN 26 CRITICAL
MAGNI Hosts (magni-hosts)	11 UP 11 DOWN	21 OK 23 CRITICAL
Switches (switches)	3 DOWN	3 CRITICAL

Provided services: Nagios (3/7)

Nagiosgraph

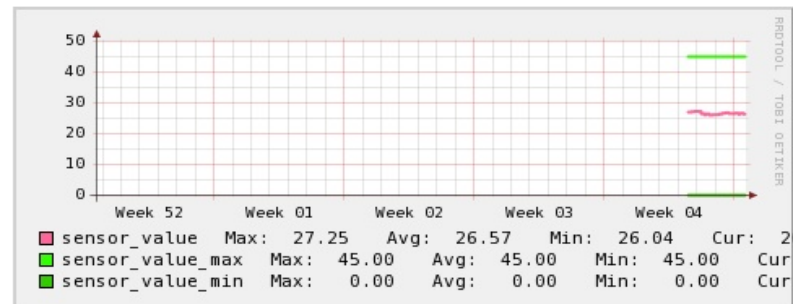
Performance data for **Host:** pc-lar-ros-emba-00 · **Service:** CPU1 temperature

Daily



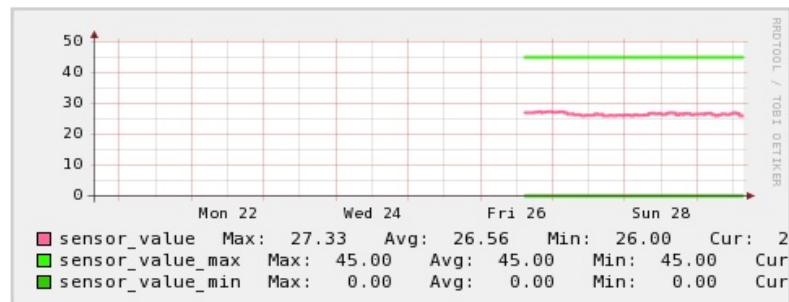
sensor_value

Monthly



sensor_value

Weekly



sensor_value

Yearly



sensor_value



Provided services (4/7)

- LDAP server for:
 - Users accounts (aim to use it for both Windows and Linux systems)
 - Configurations of autofs, sudo, samba
- Users management and Role Based Access Control
 - detailed presentation made at *1st CNIC Information Exchange on Computer-Based Access Control*
<http://indico.cern.ch/materialDisplay.py?contribId=3&materialId=slides&confId=6506>
- Windows specific (DCS)
 - Windows Terminal Server: gateway for Windows PCs, and platform to run distributed PVSS projects
 - Samba export of users' home directories & detector areas
- Remote management
 - Automated power down/up of the clusters



Remote management (5/7)

- Need to minimize time to power up the clusters automatically (eg after power cut) and to control each machine without need to go to technical cavern
- Decided to use IPMI (Intelligent Platform Management Interface).
- Why IPMI? Industry standard widely available on server motherboards.
- Problem is variety of versions (pure 1.5, 1.5 with 2.0 back-ported features, 2.0) and implementations: Hide differences in tool developed by us.
- IPMI is either part of the motherboard or an add-on card (most often the case). It either shares the onboard NIC (used by OS) with the same or different IP, or has a separate NIC onboard or off-board.



Remote management (6/7)

- For large farms, distribute IPMI commands in parallel to minimize time: NILE (see paper at Chep06) was used for its parallel and hierarchical command distribution
- Results of command distribution on the cluster:
 - Preseries (~75 machines): power status command ~ 60s (use of SSH for some nodes increases time compared to below)
 - Final PCs (~154 machines): power down/up/status ~ 20s
- Scripts being updated to include full error catching and recovery procedures. One command will bring set of machine into a well defined state (e.g. booted and fully functional or error log)

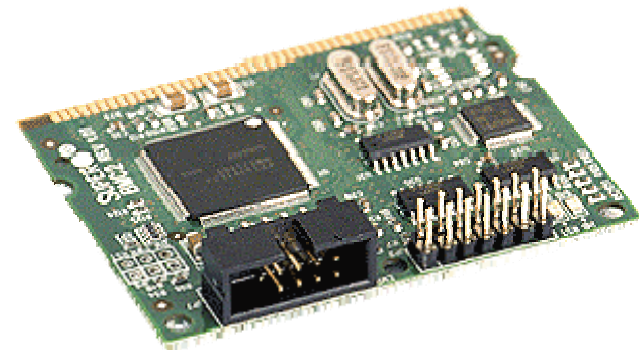
Remote management (7/7)

- Example IPMI devices

Functions as a separate NIC



Shared NIC with Motherboard



- SBCs: power on and reset via is available via PVSS and ELMB connection to the racks. Not yet deployed to all SBCs.
- All others PCs (ie no IPMI, not many machines): no solution.



Link to Windows World (DCS)

- We provide all Standard net services (DNS, Time, registration and network infrastructure)
- DAQ servers/services mainly aimed at DAQ and Linux based, however: export home directories and detector areas via Samba
- PVSS is now also available with same GUI technology on Linux. Being tested for use in control room etc...
- A Windows Terminal Server (see previous slides)



Tools developed internally

- “Boot With Me” (BWM) –aim to get an uniform OS environment
 - builds boot images for net booted machines
 - specially designed for heterogeneous systems like ATLAS
- File system synchronization scripts using Nile and rsync (sw, data...)
- Tool for generation of DHCP configuration files
- Plugins for Nagios
- Knowledge databases (internal usage)
 - SysAdmin FAQ – collects the experience for the installation and maintenance of the ATLAS system (109 entries so far)
- Utilities for LDAP

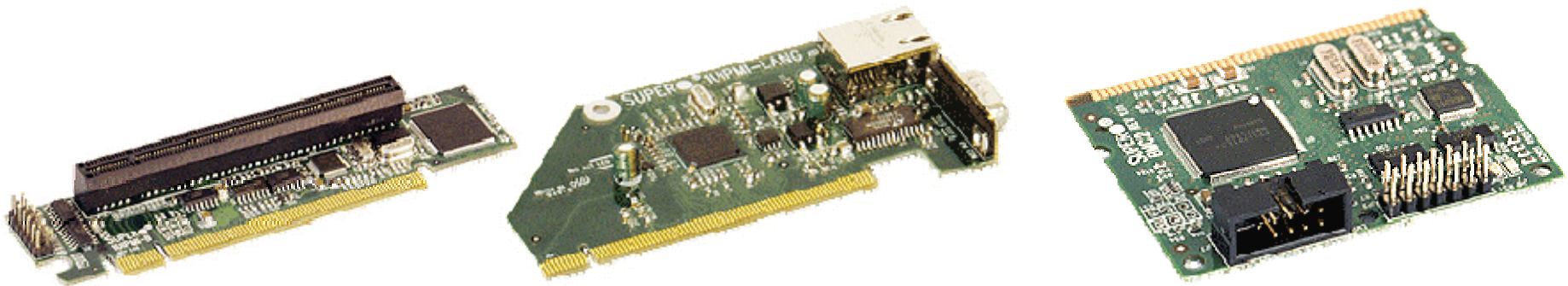


Useful links

- Web site: user guides howto's, connection form request
 - <http://cern.ch/atlas-tdaq-sysadmin/>
- FAQ website:
 - <http://pcatdwww.cern.ch/FAQ/point1>
- NILE website:
 - <http://nile.ifae.es/documentation.html>

IPMIs

- Shares LAN controller with main board



- Functions as a 3rd LAN port

