Journal of Physics: Conference Series **219** (2010) 022048 doi:10.1088/1742-6596/219/2/022048

System Administration of ATLAS TDAQ Computing Environment

A. Adeel-Ur-Rehman[1], F. Bujor[2], J.Benes[3], C. Caramarcu[4], M.Dobson[5], A. Dumitrescu[2], I. Dumitru[2], M. Leahu[2], L. Valsan[2], A. Oreshkin[6], D. Popov[7], G. Unel[5], A. Zaytsev[8, 9]

[1] National Centre for Physics, Islamabad, Pakistan
[2] Politehnica University of Bucharest, Romania
[3] Zapadoceska Univerzita v Plzni, Czech Republic
[4] National Institute of Physics and Nuclear Engineering, Romania
[5] University of California at Irvine, USA
[6] St. Petersburg Nuclear Physics Institute, Russia
[7] Max-Planck-Institut fuer Kernphysik, Heidelberg, Germany
[8] Budker Institute of Nuclear Physics, Novosibirsk, Russia

[9] E-mail: Alexandr.Zaytsev@cern.ch

Abstract. This contribution gives a thorough overview of the ATLAS TDAQ SysAdmin group activities which deals with administration of the TDAQ computing environment supporting High Level Trigger, Event Filter and other subsystems of the ATLAS detector operating on LHC collider at CERN. The current installation consists of approximately 1500 netbooted nodes managed by more than 60 dedicated servers, about 40 multi-screen user interface machines installed in the control rooms and various hardware and service monitoring machines as well. In the final configuration, the online computer farm will be capable of hosting tens of thousands applications running simultaneously. The software distribution requirements are matched by the two level NFS based solution. Hardware and network monitoring systems of ATLAS TDAQ are based on NAGIOS and MySQL cluster behind it for accounting and storing the monitoring data collected, IPMI tools, CERN LANDB and the dedicated tools developed by the group, e.g. ConfdbUI. The user management schema deployed in TDAQ environment is founded on the authentication and role management system based on LDAP. External access to the ATLAS online computing facilities is provided by means of the gateways supplied with an accounting system as well. Current activities of the group include deployment of the centralized storage system, testing and validating hardware solutions for future use within the ATLAS TDAQ environment including new multi-core blade servers, developing GUI tools for user authentication and roles management, testing and validating 64-bit OS, and upgrading the existing TDAQ hardware components, authentication servers and the gateways.

1. Introduction

The Trigger and Data Acquisition (TDAQ) System of the ATLAS experiment exploits a large online computing farm for the readout of the detector front-end data, the trigger decision farms (second and third level of trigger) and all the ancillary functions (monitoring, control, etc.). These systems are deployed in the vicinity of the detector cavern (USA15 underground area) and within the surface

facilities at the experimental area at the LHC intersection Point 1 (SDX1 area, ATLAS main and satellite control rooms, etc.). Two of these areas, which hold the majority of TDAQ equipment, are of particular importance:

- USA15 provided with 220 racks in total (deployed on 3 floors) which are 70% occupied by TDAQ and ATLAS sub-detectors equipment. The equipment in this area uses 1 MW of power at the present moment while 2.5 MW of cooling capacity is available for future upgrades.
- SDX1 provided with 120 racks in total (deployed on 2 floors) which are 50% occupied by TDAQ equipment The power consumption of this area is 0.5 MW at the moment and up to 1.5 MW of cooling capacity is available for the upgrades.

At the moment, the ATLAS TDAQ system exploits roughly 1200 computers and tens of thousands instances of various applications. These machines need to be administrated in a coherent and optimal way in order to maintain the computing farms at the highest level of availability, and minimize the downtimes. This is required for ATLAS to have a highly reliable and robust online computing environment which can make the best use of available luminosity provided by the LHC collider.

A dedicated group of system administrators (the ATLAS TDAQ SysAdmin Group) is dealing with these tasks and in addition with the support of ATLAS online computing users on 24x7 basis.

2. ATLAS TDAQ SysAdmin Group Activities

The group maintains multiple ATLAS TDAQ related computing areas across the CERN sites:

- ATLAS Point1: SDX1, USA15, ACR (Main Control Room), SCR (Satellite Control Room),
- Laboratories in Bat.4, Bat.32 and Bat.40 on the CERN Meyrin site.

The everyday activity includes:

- Dealing with ATLAS Point 1 user and role management requests,
- Handling software areas synchronization requests,
- Addressing IT security issues and incidents within ATLAS online systems,
- 24x7 SysAdmin shifts (starting from Summer 2008),
- Providing on-call services,
- Hardware and software monitoring & maintenance of the computing infrastructure,
- New hardware items registration (upon request).

All these tasks are handled in close cooperation with other relevant groups (ATLAS Networking Team, ATLAS Technical Coordination, CERN IT Department) dealing with other aspects of maintenance and operation of ATLAS experiment. In addition the group carries out the development and validation of tools & solutions for automated user, software and hardware management, monitoring and control.

A variety of tools were developed within the group in order to automate the most frequently executed operations, for instance:

- ATLAS Point 1 user and role management scripts,
- Boot With Me (BWM) project components (control of netbooted nodes),
- Storage areas synchronization scripts,
- Tools for registering new entities in the CERN network database (LanDB),
- Scripts for batch command execution on a group of hosts,
- Tools for bulk firmware upgrade for the High Level Trigger Processor Unit (HLT PU) and Local File Server (LFS) nodes.

These tools are being intensively used and continuously improved.

3. Design and Organisation of the ATLAS TDAQ IT Infrastructure

3.1. Generic Design Overview

A major concern in any high availability computing farm is the mean time between failures. This is highly correlated to the mean time between failure of only a few of the components in the computers,

such as disks. For this reason, TDAQ decided to try to reduce the need for disks on the data acquisition system, by making nodes diskless. The diskless nodes are booted into Linux over the network from a server. This boot server approach has other advantages, such as ease of maintenance, reproducibility on a large scale, and the like. The BWM project was developed in order to respond to the need for a flexible system to build boot images and configure the booting of the diskless nodes.

Another major corner stone of the system is therefore the boot servers. These are designed to serve the DHCP requests and boot images, and to provide network mounted disks for the main part of the OS and TDAQ applications. In this kind of system, the servers are a single point of failure. To overcome this limitation, the system, which would anyway need many such servers to cover the large number of client machines involved, has been made more robust by sharing the responsibility of booting and providing network drives to a machine across two or more of the servers. This redundancy insures the high availability of the clients independently of that of individual servers. The redundancy and availability of servers and computers is dependent on a high performance and redundant network interconnecting the devices.

Another requirement for the system is the ability to run the experiment and take data for up to 24 hours while having lost the connection to the IT department and Tier0 centre (responsible for long term storage of the data, distribution of it to Tier1 centres, and also analysis of some of the data). The implications are that the system should replicate any services in IT which are vital such as DNS, NTP, user authentication; and should be able to buffer the event data. The latter is done using a few servers with large disk caches (12 TB), able to handle the incoming rate of selected events to the disks, and to simultaneously sustain twice the output rate from the disks to the permanent storage in IT (in order to catch up any connectivity loss).

3.2. Functional Layout and Review of Current Configuration

The functional layout of the ATLAS TDAQ IT infrastructure is shown on Fig. 1.

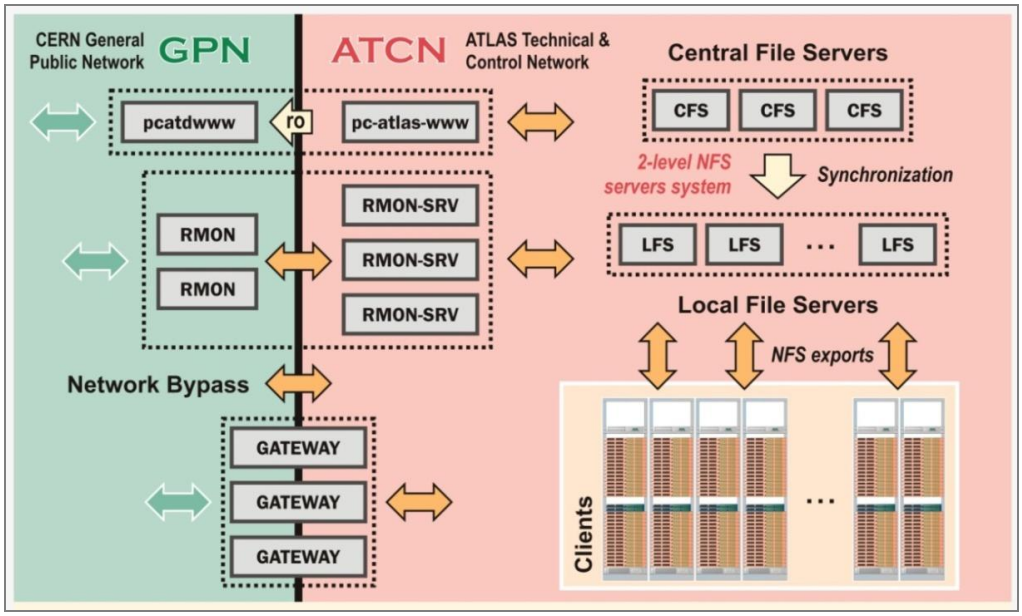

Figure 1. Functional layout of ATLAS TDAQ IT infrastructure.
Two level system of NFS servers exploited for software distribution among the clients is shown.
Most of the equipment deployed is hidden from CERN General Public Network (GPN)
behind the dedicated gateways, remote monitoring nodes and web/proxy servers.

Currently SDX1 & USA15 areas together host more than 200 fully occupied racks of ATLAS TDAQ and detector sub-systems equipment which can be subdivided into the following groups:

- 3 gateways, 2 DNS, 3 CFS (Central File Server), 60 LFS,
- More than 1500 netbooted nodes: 154 ROS (Read-Out System), 835 HLT PU (6680 CPU cores combined), 77 EB (Event Building nodes), 5 SFO (Sub-Farm Output buffer node to permanent storage), 64 online and monitoring nodes, 157 SBC (Single Board Computer), 109 test system nodes,
- Many locally installed machines: 21 ACR (standard Control Room machines, 4-screen each), 44 SCR machines (including 20 iMacs), Detector Control System (DCS) nodes, sub-detector PCs, public nodes (including the ones installed in ATLAS Visitors Area).

The number of HLT Processor Units (PU) deployed in SDX1 up to now represent only 50% of total SDX1 capacity. This is assumed to be sufficient for the initial phase of the accelerator programme and will be increased to full capacity over the next 15 months to meet demand.

3.3. Hardware Means of Reducing Unscheduled Downtimes

Since the complexity of IT infrastructure involved is really high it takes a significant amount of time to bring it to the production state after the complete shutdown so measures were taken to protect the TDAQ infrastructure from the power cuts of various origin. The whole facility is provided with two centralized UPS lines with diesel generators backup, plus the independent UPS lines are available in SDX1 for mission critical equipment. Nowadays approximately 5% of equipment deployed in SDX1 is on UPS lines or at least half on UPS and half on normal power (no UPS protection), and just a few machines like CFS nodes are half on the dedicated locally installed UPS and half on normal power.

In addition the key machines and services used in Point 1 are provided with the real time backup by means of Linux High Availability (Linux-HA implemented with Heartbeat) software [1]. In the present configuration the list of servers protected by means of Linux-HA includes two pairs of LFS nodes supporting the ACR and SCR machines.

3.4. Hardware and Service Monitoring

Host monitoring is currently being done using the NAGIOS software [2]. NAGIOS allows the monitoring of various services for the machines. For all machines, the following basic services are monitored:

- ping response & SSH connectivity, NTP synchronization,
- kernel version, BWM version, LFS name,
- temperature, HDD state (if present),
- auto-mount status, ramdisk usage (for netbooted nodes).

This list is not fixed, as NAGIOS allows the administrators to add and configure new features and services. On certain hosts, such as LFS or Application Gateways, advanced features are monitored, such as:

- NTP & DHCP daemon status,
- state of exported file systems,
- number of users, etc.

The NAGIOS graphs collected by the monitoring system are stored on disks in the form of more than 25800 RRD files of total size approximately 4.5 GB. Status information of all the nodes and NAGIOS graphs for various parameters is published automatically in the monitoring section of the private ATLAS Operations webserver (refreshed automatically each 90 seconds). The selected indicators which are of crucial importance for proper functioning of the ATLAS TDAQ infrastructure are provided with E-mail/SMS alerts. Normally an SMS alert is received within 2 minutes after the failure takes place (both latency of the monitoring system and message delivery delay are taken into account).

Sample screenshots of the main NAGIOS monitoring web page and a derived web page dedicated for a particular group of hosts being monitored are presented on Fig. 2 & 3.

3.5. Hardware Control and Software Deployment Tools

Due to the large number of netbooted nodes being managed, nontrivial levels automation are required to reduce the amount of time consumed by performing various routine operations, like rebooting a group of machines, assigning clients to the LFS servers, etc. A dedicated set of tools based on a MySQL database and the GUI front end to them called ConfdbUI, were developed to cover these requirements. Most of the machines are currently being migrated from SLC 4.x to SLC 5.x (both x86 and x86_64 architectures). The standard solution for the remote hardware management in ATLAS Point 1 is based exclusively on IPMI [3, 4].

ConfdbUI is now widely used within the group and capable of handing the following tasks:

- registering new netbooted clients on basis of the data extracted from CERN LanDB,
- configuring netbooted clients (including Nagios monitored services),
- client-to-LFS assignment schema and fast client migration,
- deploying DHCP and NAGIOS configurations on the Point 1 nodes,
- issuing IPMI and system commands on multiple nodes in parallel.

A sample screenshot of ConfdbUI web interface is shown on Fig. 4.

3.6. Remote Access Subsystems

The access to the ATLAS Technical and Control Network (ATCN) from outside of the Point 1 is highly restricted and only allowed via one of the following gateways systems:

- ATLAS Point 1 gateways allowing expert users to access a particular set of machines within the Point 1 via SSH and SCP protocols,
- ATLAS Remote Monitoring System providing the graphical terminal services required for organizing the remote participation in ATLAS sub-detector monitoring shifts,
- ATLAS DCS Windows Terminal Servers allowing DCS experts to access the DCS SCADA system used in ATLAS.

All the gateways are provided with both host based and network based accounting, security monitoring and intrusion prevention systems.

The migration of existing gateway system to the new 8 CPU core servers (based on SLC 5.x 64-bit OS provided with XEN virtualization) is being finalized as of the writing of this note.

3.7. Centralized and Distributed Storage Systems

Multiple task-specific storage systems are currently used within the ATLAS TDAQ environment including:

- Storage areas on CFS nodes (user home directories, DAQ software distribution area, NAGIOS RRD files aggregation area, CIFS/NFS exports for DCS, etc.),
- Netbooted nodes configuration storage areas,
- DAQ configuration areas,
- Exports from the private webserver to the public one,
- Dedicated storage area shared among the gateways.

The existing multiplicity of storage systems is to be replaced later in 2009 by a high performance storage solution matching the following TDAQ specific requirements:

- Initial capacity 4-5 TB, scalability up to 10-20 TB,
- Peak performance up to 100 kIOPs,
- NFS/iSCSI/CIFS exports support,
- Serving up to 2500 client machines simultaneously without degradation of performance,
- Multiple levels of hardware/software redundancy (1+1 or N+1 redundancy schemas).

This new system will ensure the scalability of the centralized storage system with the growing amount of computing power in the TDAQ environment over the coming years.

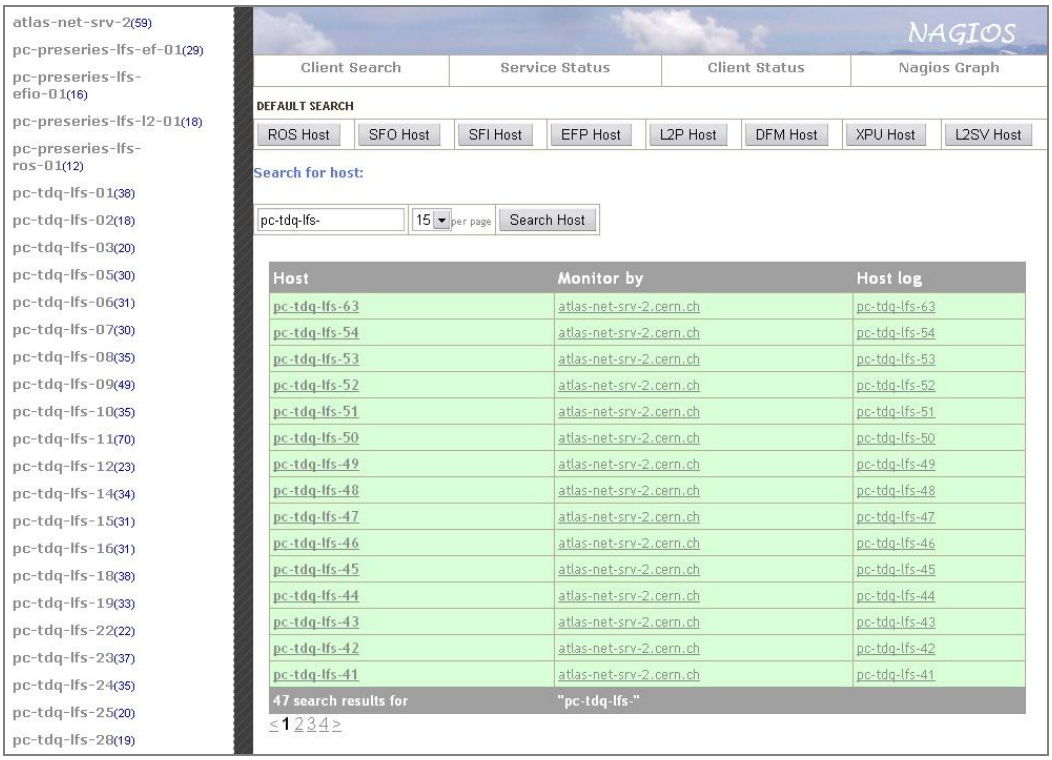

Figure 2. Main web page of NAGIOS monitoring system with the list of LFS nodes displayed.

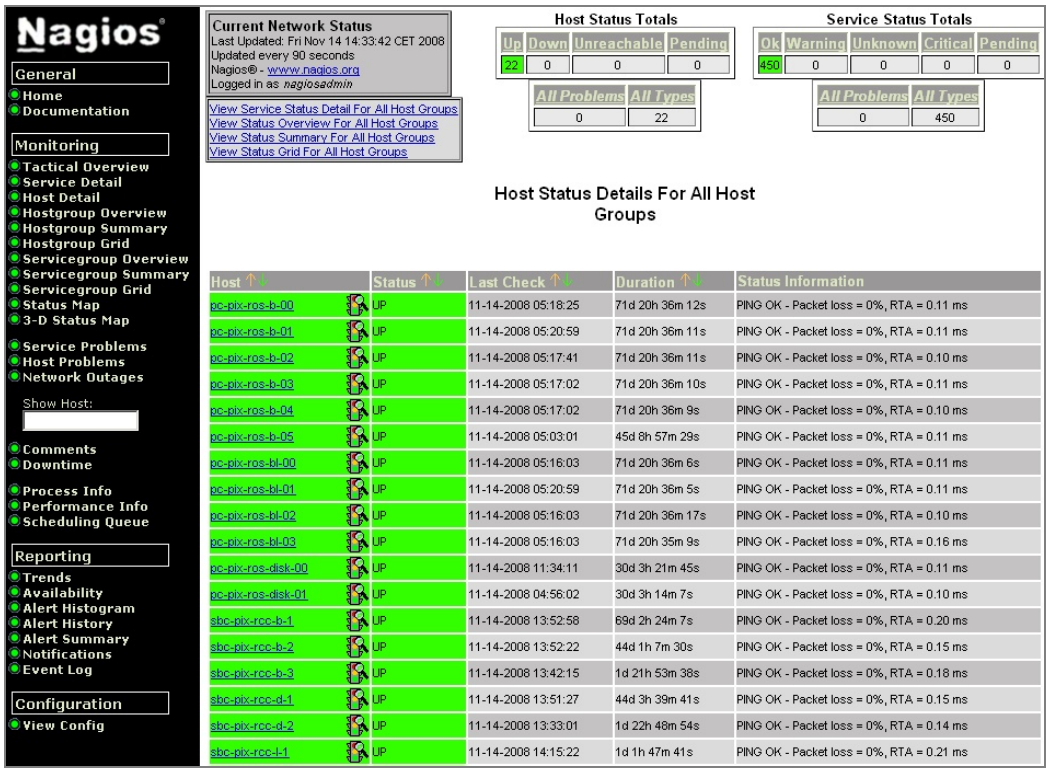

Figure 3. NAGIOS monitoring system status web page sample for a particular group of hosts.

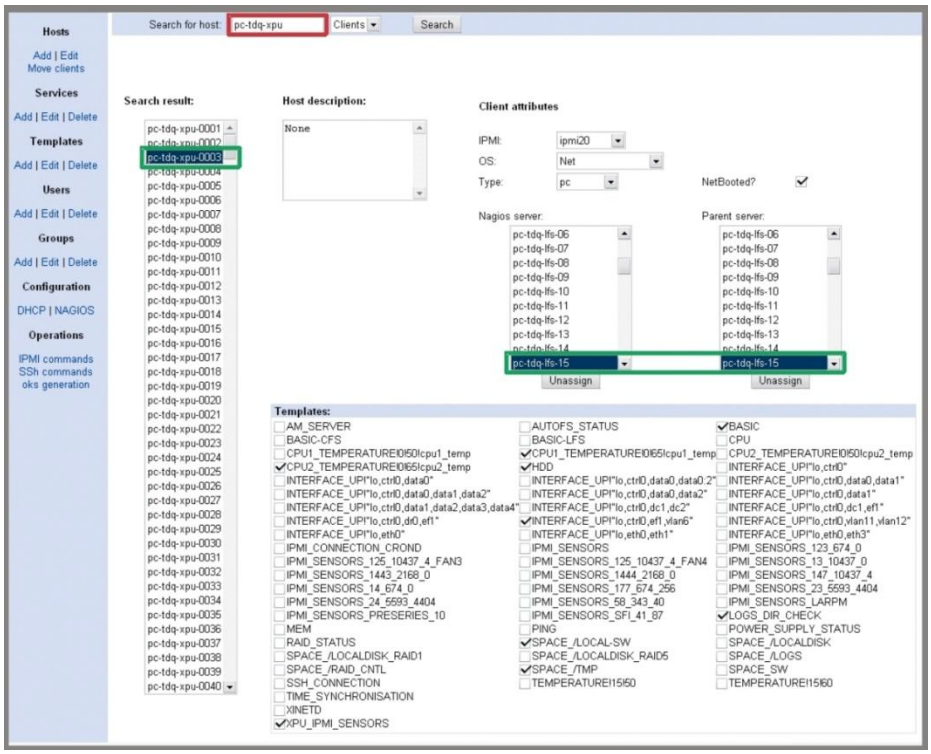

Figure 4. Sample screenshot of the ConfdbUI web interface: hardware monitoring configuration page for a particular group of hosts. Search results for a particular pattern applied for the hostname ('pc-tdq-xpu') and the properties of the particular host ('pc-tdq-xpu-0003'), including the list of configuration templates activated for it are shown.

3.8. Ongoing Activities

The list of major milestones which are to be encountered in 2009 contains the following items:

- Hardware maintenance pipeline with manufacture is established for all the XPU nodes,
- Gateways upgrade,
- CFS nodes upgrade,
- LDAP cluster upgrade,
- Web cluster deployment,
- Centralized storage system installation in SDX1.

4. ATLAS Point 1 User and Role Management Mechanisms

Following the requirement to be independent from IT, and to allow more flexibility, the experimental area has its own user database in the form of an LDAP server based on OpenLDAP software [5, 6]. The system is standalone but for consistency it is synchronized with IT for the usernames and user IDs. Contrary to the way some of the previous experiments have been run, it has been decided to have user based authentication and not group based authentication, in order to have accountability and traceability of actions, as well as increased security.

For authentication, the consistency is maintained by having a slave Windows Domain Controller [7, 8] using the CERN NICE credentials of a user. The only exceptions are the local service accounts which hold their authentication (passwords) in the LDAP server.

The reasons for wanting to use group accounts in past experiments comes from the natural splitting of users into categories of people and tasks which they are allowed to do, for example some users are shifters, detector experts, TDAQ experts. In order to address this categorization and still have the user

authentication for accountability and traceability, it has been decided in TDAQ to implement a Role Based Access Control (RBAC) authorization system [9].

Currently ATLAS Point1 is provided with the dedicated role based access control and authorization system currently holding more than 150 unique roles in a hierarchical structure. The total number of users registered is more than 1600, but only a small fraction of them would be allowed to access Point1 remotely during the data taking period (experts which are on shift or on call).

The RBAC system now undergoes the major upgrade implying the following operations:

- Creating Role/POSIX group synchronization mechanisms,
- Role hierarchy and groups optimization,
- Implementing user account management interfaces (GUI),
- Creating the role assignment delegation mechanisms for future use by ATLAS shift leaders,
- Installing the dedicated read optimized LDAP cluster based of multiple redundant servers.

A sample screenshot of the LDAP interface involved in all these operations is shown on Fig. 5.

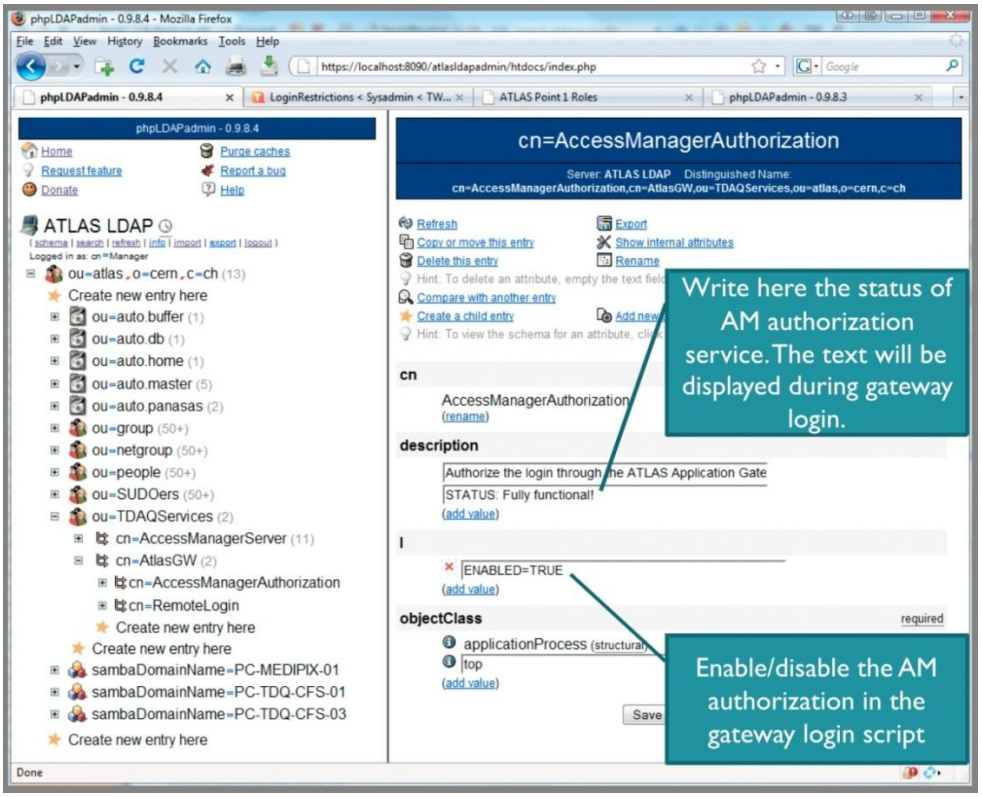

Figure 5. Screenshot of the web interface to ATLAS LDAP which is involved in user role, group membership and access rights management for ATLAS TDAQ environment.

5. Conclusion

The ATLAS TDAQ system is fully operational and is being constantly tested in the ongoing cosmic and test data taking runs with the ATLAS sub-detectors. The design of the system and the supporting computer infrastructure has been validated. However some fundamental aspects are being finalised to allow the increase of the computing power required by the experiment to be able to take advantage of the increased energy and luminosity of the accelerator over the next 15 months.

6. References

[1] Linux-HA (Heartbeat) Project Web Site: http://www.linux-ha.org/HeartbeatProgram
[2] NAGIOS Web Site: http://www.nagios.org
[3] Intelligent Platform Management Interface (IPMI): http://en.wikipedia.org/wiki/Intelligent_Platform_Management_Interface
[4] OpenIPMI Project Web Site: http://sourceforge.net/projects/openipmi
[5] Lightweight Directory Access Protocol: http://en.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol
[6] OpenLDAP Project Web Site: http://www.openldap.org
[7] Windows Domain Controller: http://en.wikipedia.org/wiki/Domain_controller
[8] Microsoft Active Directory Domain Services: http://www.microsoft.com/windowsserver2008/en/us/active-directory.aspx
[9] Role Based Access Control (RBAC) Authentication System: http://csrc.nist.gov/rbac