HEPiX Spring 2008

Monday 05 May 2008 - Friday 09 May 2008

CERN

Book of Abstracts

Contents

Applications and operating systems / 0

Scientific Linux Status Report

Author: Troy Dawson[1]

[1] *FERMILAB*

Corresponding Author: dawson@fnal.gov

Progress of Scientific Linux over the past 6 months.
What we are currently working on.
What we see in the future for Scientific Linux

Applications and operating systems / 1

SL Plenary Discussion

Author: Troy Dawson[1]

[1] *FERMILAB*

Corresponding Author: dawson@fnal.gov

Feedback to and input for the SL developers from the HEPiX community. This may influence upcoming decisions e.g. on distribution lifecycles.
One topic we will be gathering feedback on is the support lifetime of Scientific Linux 4.

Site reports / 2

GridKa Site Report

Authors: Andreas Heiss[1]; Manfred Alef[1]

[1] *FORSCHUNGSZENTRUM KARLSRUHE*

Corresponding Author: andreas.heiss@iwr.fzk.de

We present the status of CPU and storage installations at the German WLCG Tier-1
centre GridKa as well as our experiences with benchmarks as acceptance test of worker node deliveries.

HEPiX "bazaar and thinktank" / 3

Data management and offline computing for LCLS

Author: Alf Wachsmann[1]

[1] *SLAC*

Corresponding Author: alfw@slac.stanford.edu

With BaBar now switched off, SLAC is shifting its focus now to the Linac Coherent Light Source (LCLS).
I will present the plans for LCLS with its detectors and DAQ systems and then talk about the envisioned data management and the ideas for the necessary offline computing.

Site reports / **4**

LAL and GRIF Site Report

Author: Michel Jouvin[1]

[1] *LAL / IN2P3*

Corresponding Author: jouvin@lal.in2p3.fr

Site report about LAL and GRIF

Site reports / **5**

TRIUMF Site Report

Author: Steven McDonald[1]

[1] *TRIUMF*

Corresponding Author: mcdonald@triumf.ca

TRIUMF Site Report

Data centre management, availability and reliability / **6**

Updates and Experiences from the Genome Center's Data Center Construction Project

Author: Gary Stiehr[1]

[1] *The Genome Center at Washington University*

Corresponding Author: garystiehr@wustl.edu

Over the last couple of years, The Genome Center at Washington University in St. Louis has been involved with the planning and construction of a new data center. We will provide updates since our data center presentation at HEPiX Fall 2007 in St. Louis. In addition, we will share our experiences and lessons learned as we prepare to move into the new data center in May 2008.

HEPiX "bazaar and thinktank" / **7**

A redundant Servercluster using SL5.1 and Xen

Author: Klaus Steinberger[1]

[1] *LMU München*

Corresponding Author: klaus.steinberger@physik.uni-muenchen.de

Howto build a redundant (failover) Servercluster using Scientific Linux 5.1 with Xen. The services (filer, dns, dhcp, printing etc.) will be divided into individual virtual machines to isolate the services. On failure of a server, the services will be relocated automatically.

The system is using a redundant Fibrechannel storage as the shared storage, but same would be possible using other techniques. The talk shows the architecture of the system, as well as some problems we found both in the SL Clustersoftware and SL's Xen Implementation and how we avoid them.

Storage technology / 8

The unbearable slowness of tapes

Author: Charles Curran[1]

[1] *CERN*

Corresponding Author: charles.curran@cern.ch

It is still common for the lowest layer of data storage to be 'tape'.

While individual tape devices apparently offer very good performance, this is only achieved in a limited set of circumstances, and the overall throughput can easily become very poor very quickly.

How can this be avoided, if at all? Is it sufficient to just add more and more equipment? These difficulties will be discussed in the context of CERN's current CASTOR HSM version, and the need to 'carry LHC data forward' for many years.

Site reports / 9

NDGF Site Report

Author: Mattias Wadenstein[1]

[1] *NDGF*

Corresponding Author: mattias.wadenstein@cern.ch

An overview of the current state at NDGF along with some interesting events at the various computing centers hosting the distributed parts of NDGF.

Networking infrastructure and computer security / 10

IPv6 experience in a mixed unix environment

Author: Mattias Wadenstein[1]

[1] *NDGF / HPC2N / ACC*

Corresponding Author: mattias.wadenstein@cern.ch

We have been using the Academic Computer Club at Umeå University as our testbed for IPv6 deployment, this includes some popular public services like a rather big free software mirror as well as workstations and multi-user machines in a semi-production environment. This talk gives an overview of the various pitfalls as well as what "just works", with some pointers on what kind of systems and software need attention.

Site reports / **11**

Status Report Finnish CMS-T2 HIP/CSC

Author: Christof Hanke[1]

[1] *CSC ltd.*

Corresponding Author: hanke@csc.fi

A report about implementation of dCache at CSC as a storage element for the CMS-T2 centre. Technical details about network, disk, monitoring and so on will be given. CSC provides for HIP some 100TB of disk storage for CMS-T2 and about 70 TB of disks with a tape-backend for the NDGF-Alice T1-centre. Internal and external network traffic is split on the network-level. The CE and the middleware situation is presented in less detail.

12

Status report on LHC and the Experiments

Author: Jos Engelen[1]

[1] *CERN*

Corresponding Author: jos.engelen@cern.ch

13

A review of the current technical activities in the CERN openlab

Author: Sverre Jarp[1]

[1] *CERN*

Corresponding Author: sverre.jarp@cern.ch

14

LHC Networking to the Tier-1's, Present and Future

Author: David Foster[1]

[1] *CERN*

Corresponding Author: david.foster@cern.ch

15

The (WLCG) Common Computing Readiness Challenge (CCRC'08)

Author: Jamie Shiers[1]

[1] *CERN*

Corresponding Author: jamie.shiers@cern.ch

Summary:

Since several years the four LHC experiments have been running periodic stress tests of their planned computing operations when LHC data arrives.
<P>
They have written design documents which describe their models for how they will acquire, distribute and process their data. They differ in details but have a common base in a model of levels or 'Tiers' of resources.
<P>
A weakness of all the tests done is that they have never overlapped all the experiments together so never reached the full rates expected at CERN and the Tier sites and never had to cope with the lack of homogeneity of having the 4 different experiment data management and batch processing models being busy at the same time. The CCRC08 (Common Computing Readiness Challenge) tests address this by deliberately exercising all experiments computing at the target rates at the same time.

16

LCG Update

Author: Ian Bird[1]

[1] *CERN*

Corresponding Author: ian.bird@cern.ch

Storage technology / 17

Setting up a simple Lustre Filesystem

Author: Stephan Wiesand[1]

[1] *DESY*

Corresponding Author: stephan.wiesand@desy.de

We recently set up a small (40TB) Lustre built from HP commodity components. It was surprisingly simple, and we achieved reasonable performance without much tuning. We'd like to share this experience.

HEPiX "bazaar and thinktank" / 18

Implementation of the National Analysis Facility at DESY

Author: Stephan Wiesand[1]

[1] *DESY*

Corresponding Author: stephan.wiesand@desy.de

A common analysis facility for german physicists working on LHC and ILC topics is currently being built, starting
at DESY's two sites in Hamburg and Zeuthen. We present the current technical implementation of this distributed facility.

Storage technology / 19

Experience and Lessons learnt from running high availability databases on Network Attached Storage

Authors: Juan Manuel Guijarro[1]; Nilo Segura Chinchilla[1]; Ruben Gaspar[1]

[1] *CERN*

Corresponding Author: nilo.segura.chinchilla@cern.ch

The Database and Engineering Services Group of CERN's Information Technology Department provides the Oracle based Central Data Base services used in many activities at CERN. In order to provide High Availability and ease management for those services, a NAS (Network Attached Storage) based infrastructure has been set up. It runs several instances of the Oracle RAC (Real Application Cluster) using NFS as share disk space for RAC purposes and Data hosting. It is composed of two private LAN's to provide access to the NAS file servers and Oracle RAC interconnect, both using network bonding. NAS nodes are configured in partnership to prevent having single points of failure and to provide automatic NAS fail-over. This presentation describes that infrastructure and gives some advice on how to automate its management and setup using a Fabric Management framework such as Quattor. It also covers aspects related with NAS Performance and Monitoring as well Data Backup and Archive of such facility using already existing infrastructure at CERN.

Summary:

Experience and Lessons learnt from running high availability databases on Network Attached Storage.

Site reports / 20

RAL Site Report

Author: Martin Bly[1]

[1] *STFC/RAL*

Corresponding Author: m.j.bly@rl.ac.uk

A report on developments at RAL since the last HEPiX.

CPU technology / 21

Benchmarking Multi-core Processors in the NERSC Production Environment.

Author: Iwona Sakrejda[1]

[1] *LBNL/NERSC*

Corresponding Author: isakrejda@lbl.gov

Results of benchmarking based on HENP production codes from single, dual and quad - core processors
will be presented. We will look at ratios of CPU/wall clock, memory consumption and I/O performance of those codes on the latest additions to the PDSF production systems.

Site reports / 22

GSI Site Report

Author: Walter Schön[1]

[1] *GSI*

Summary:

Site report from GSI

Storage technology / 23

Lustre cluster in production at GSI

Author: Walter Schön[1]

[1] *GSI*

Summary:

Since a few months, the Lustre cluster at GSI is in production and available from almost 1400 CPUs. First experience with “real” users and “real” files will be presented.

HEPiX ”bazaar and thinktank” / 24

High Availability and Backup Strategies for the lustre MDS Server

Author: Karin Miers[1]

[1] *GSI*

Summary:

The MDS Server is the heart of the Lustre file system. It manages the cluster and contains the meta data of the file system. Standard Linux tools are used to make this service highly available. Moreover, a special backup strategy was developed to allow a fast desaster recovery in case of failure.

Data centre management, availability and reliability / **25**

Remote Administration via Service Modules

Author: Stefan Haller[1]

[1] *GSI*

Summary:

Decentralized server environments require modern solutions for administration. The standard IPMI (Intelligent Platform Management Interface) has been created for monitoring system health and manage the system from remote. Corresponding hardware modules allow administration of servers independent from the operating system even if the underlying server has been powered down. This talk covers basics like conceptional considerations for a management network up to a practical demonstration.

Site reports / **26**

NIKHEF site report

Author: Paul Kuipers[1]

[1] *NIKHEF*

Site report from NIKHEF

Data centre management, availability and reliability / **27**

Amsterdam Tier1 Data Centre Infrastructure

Author: Wim Heubers[1]

[1] *NIKHEF*

Corresponding Author: wimh@nikhef.nl

Extension of the NIKHEF/SARA data centre

Storage technology / **28**

Final Report from File Systems Working Group

Author: Andrei Maslennikov[1]

[1] *CASPUR*

Corresponding Author: andrei.maslennikov@caspur.it

Final Report from File Systems Working Group

Storage technology / **29**

Options for medium-/long-term improvements to LHC mass storage and data management

Author: Dirk Duellmann[1]

[1] *CERN-IT*

Corresponding Author: dirk.duellmann@cern.ch

Options for medium-/long-term improvements to LHC mass storage and data management

Storage technology / **30**

Storage elements at BNL

Author: Robert Petkus[1]

[1] *Brookhaven National Laboratory*

Corresponding Author: rpetkus@bnl.gov

Storage technology / **32**

Towards the new data management solution at CNAF

Author: Vladimir Sapunenko[1]

[1] *CNAF*

Towards the new data management solution at CNAF

Data centre management, availability and reliability / **33**

AFS Monitoring: The CERN AFS Console

Author: Arne Wiebalck[1]

[1] *CERN*

Corresponding Author: arne.wiebalck@cern.ch

CERN's AFS installation serves between 1 and 2 billion accesses per day to its around 20'000 users. Keeping track of the system's overall status and trying to find problems before the users do is not a trivial task, esp. as the installation is growing in almost all aspects. This talk will present CERN's AFS Console, a Lemon- and web-based monitoring tool used by the AFS administrators at CERN to quickly identify problematic entities (servers, partitions, volumes etc.) and to assist them in solving the issues found.

Storage technology / **34**

CASTOR Status and Plans

Author: Sebastien Ponce[1]

[1] *CERN*

Corresponding Author: sebastien.ponce@cern.ch

This presentation will cover the current status of the CASTOR mass storage system, both in terms of software status and deployment and performance. Near future developement plans will also be presented while the longer term view will be addressed in a separate presentation, by Dirk Duellmann.

Data centre management, availability and reliability / **35**

New Data Center at BNL – Status Update

Author: Tony Chan[1]

Co-author: Michael Ernst [1]

[1] *Brookhaven National Laboratory*

Corresponding Author: awchan@bnl.gov

This presentation provides an update on the status of the new Data Center to support the ATLAS Tier 1 Center and RHIC Computing at Brookhaven. A brief discussion provides details of the new facility to Brookhaven, as well as timelines for availability to both the ATLAS and RHIC programs. Some of our experiences described in this presentation will also be beneficial to other sites who are considering expansion of their own facilities.

Data centre management, availability and reliability / **36**

A Service-Based SLA Model

Authors: Jason Smith[1]; Tony Chan[1]; Tristan Ziska[1]

[1] *Brookhaven National Laboratory*

Corresponding Author: awchan@bnl.gov

The RACF provides computing support to a broad spectrum of programs at Brookhaven. The growth of the facility, the varying needs of the scientific programs and the necessity for distributed computing requires the RACF to change from a system to a service-based SLA with our end users. This presentation describes the adjustments made by the RACF to transition to a service-based SLA, including changes to its monitoring, alarm notification and problem resolution policies at the facility.

Site reports / **37**

PSI Site Report

Author: Heiner Billich[1]

[1] *Paul Scherrer Institut*

Corresponding Author: heiner.billich@psi.ch

Site Report for PSI
Applications, HPC, Scientific Linux, Storage and Network

Networking infrastructure and computer security / **38**

Web applications security

Author: Romain Wartel[1]

[1] *CERN*

Corresponding Author: romain.wartel@cern.ch

This talk presents common security issues with Web applications and provides some practical recommendations to users, developers, and Web services managers to mitigate the risks linked to these applications.

Networking infrastructure and computer security / **39**

Operational security in a grid environment

Author: Romain Wartel[1]

[1] *CERN*

Corresponding Author: romain.wartel@cern.ch

This talk presents the main goals of computer security in a grid environment, by using a FAQ approach. It details the evolution of the risks in the recent years, likely objectives for attackers and the progress made by the malware toolkits and frameworks. Finally, recommendations to deal with these threats are proposed.

Networking infrastructure and computer security / **40**

Advanced Monitoring Techniques for the Atlas TDAQ Data Network

Authors: Ali Al-Shabibi[1]; Brian Martin[2]; Lucian Leahu[3]; Matei Ciobotaru[4]; Silvia Batraneanu[3]; Stefan Stancu[4]

Co-authors: Georgiana Darlea [5]; Mihail Ivanovici [6]

[1] *CERN, University of Heidelberg*

[2] *CERN*

[3] *CERN, Politehnica Bucharest*

[4] *CERN, UC Irvine, Politehnica Bucharest*

[5] *CERN, Polytech' Savoie, Politehnica Bucharest*

[6] *University of Brasov, Romania*

Corresponding Authors: matei.ciobotaru@cern.ch, brian.martin@cern.ch

We describe the methods used to monitor and measure the performance of the Atlas TDAQ data network.
The network consists of four distinct Ethernet networks interconnecting over 4000 ports using up to 200 edge switches and five multi-blade chassis. The edge networks run at 1Gbps and 10Gb/s are used for the detectors raw data flow as well as at the cores of the data flow networks. The networks feed event data to farms of up to 3000 processors. Trigger applications running on these processors examine each event for acceptability and assemble the accepted events ready for storage and further processing on Grid linked data centers. We report in detail on the monitoring and measurement techniques deployed and developed.

Summary:

Monitoring and measurement must go beyond merely assuring that the installed system network is functional and that its performance meets the design specifications. The system will regularly be operated above its design point as applications are refined to optimally consume all available computing and network resources. Even before that stage users will typically reduce the filtering criteria to accept an increasing number of events until the point that something fails. This could be for example saturation due to insufficient computing power or network overload and subsequent backpressure. Thus for such a network it is the norm is to be overloaded, links running continuously at 100% are to be expected and packet discards in the thousands per second will happen. During all this, the network monitoring system must still be able to distinguish between real and self inflicted problems as well as identify traffic anomalies that may be preventing the system from going even faster. While redundancy is built in there will still be performance loss caused by equipment failure and identification of the appropriate repairs must be made rapidly and accurately. Experience has shown that analysis of individual, or aggregate traffic flows is extremely useful for system diagnosis and performance measurement but the sheer size of the system imposes issues of scaling for any of the traditional measurement techniques. We describe the approach taken to meet these challenges.

Networking infrastructure and computer security / **41**

Cybersecurity Update

Author: Denise Heagerty[1]

[1] *CERN*

Corresponding Authors: lionel.cons@cern.ch, denise.heagerty@cern.ch

An update on recent security issues and vulnerabilities affecting Windows, Linux and Mac platforms. This talk is based on contributions and input from a range of colleagues both within and outside CERN. It covers clients, servers and control systems.

Storage technology / **42**

FZK storage news

Author: Silke Halstenberg[1]

[1] *FZK*

Corresponding Author: silke.halstenberg@iwr.fzk.de

FZK storage news

Storage technology / **43**

Handling large datasets at Google: Current systems and future directions

Author: Sascha Brawer[1]

[1] *Google*

Handling large datasets at Google: Current systems and future directions

CPU technology / **44**

Status report from the Benchmarking working group

Author: Helge Meinhard[1]

[1] *CERN*

Corresponding Author: helge.meinhard@cern.ch

Status report from the Benchmarking working group

Applications and operating systems / **45**

Virtualisation with Windows at CERN

Author: Juraj Sucik[1]

[1] *CERN*

Corresponding Author: juraj.sucik@cern.ch

Virtualisation with Windows at CERN

Applications and operating systems / **46**

Experience with Windows Vista at CERN

Author: Rafal Otto[1]

[1] *CERN*

Corresponding Author: rafal.otto@cern.ch

Experience with Windows Vista at CERN

Applications and operating systems / **47**

Lifecycle management of Windows desktop applications

Author: Sebastien Dellabella[1]

[1] *CERN*

Corresponding Author: sebastien.dellabella@cern.ch

Lifecycle management of Windows desktop applications

HEPiX ”bazaar and thinktank” / **48**

The problem of managing 236 million user accounts

Authors: Emmanuel Ormancey[1]; Geerd Dietger hoffmann[1]

[1] *CERN*

Corresponding Authors: geerd-dietger.hoffmann@cern.ch, emmanuel.ormancey@cern.ch

The problem of managing 236 million user accounts

Data centre management, availability and reliability / **49**

CluMan

Author: Sebastian Lopienski[1]

[1] *CERN*

Corresponding Author: sebastian.lopienski@cern.ch

Managing large clusters that host complex services has particular challenges. Operations like checking configuration consistency, running some actions on node or nodes, moving them between clusters etc. are very frequent. When scaling up to running thousands of CPU and STORAGE nodes in order to meet LHC requirements some of these challenges are becoming more evident. These scaling challenges are the basis for CluMan, a new cluster management tool being designed and developed at CERN.

HEPiX "bazaar and thinktank" / 50

Inspire

Author: Tibor Simko[1]

[1] *CERN*

Corresponding Author: tibor.simko@cern.ch

Inspire is the project name of a new High Energy Physics information system which will integrate present databases and repositories to host the entire corpus of the HEP literature and become the reference HEP scientific information platform worldwide. It is a common project between CERN, DESY, FERMILAB and SLAC. It will empower scientists with new tools to discover and access the results most relevant to their research; enable novel text- and data-mining applications; deploy new metrics to assess the impact of articles and authors. In addition, it will introduce the Web2.0 paradigm of user-enriched content in the domain of sciences with community-based approaches to the peer-review process.

Data centre management, availability and reliability / 51

Problem tracking at CERN

Author: Eric Grancher[1]

[1] *CERN*

Corresponding Author: eric.grancher@cern.ch

Problem tracking at CERN

CPU technology / 52

CPU-level performance monitoring with Perfmon

Author: Andreas Hirstius[1]

[1] *CERN*

Corresponding Author: andreas.hirstius@cern.ch

CPU-level performance monitoring with Perfmon

CPU technology / **53**

The CERN benchmarking cluster

Author: Alex Iribarren[1]

[1] *CERN*

The CERN benchmarking cluster

CPU technology / **54**

Benchmarking ATLAS applications

Authors: Alessandro de Salvo[1]; Franco Brasolin[1]

[1] *INFN*

Benchmarking ATLAS applications

CPU technology / **55**

Benchmarking CMS applications

Author: Gabriele Benelli[1]

[1] *CERN*

Corresponding Author: gabriele.benelli@cern.ch

Benchmarking CMS applications

CPU technology / **56**

Benchmarking ALICE applications

Author: Peter Hristov[1]

[1] *CERN*

Corresponding Author: peter.hristov@cern.ch

Benchmarking ALICE applications

CPU technology / **57**

Benchmarking LHCb applications

Author: Hubert Degaudenzi[1]

[1] *CERN*

Corresponding Author: hubert.degaudenzi@cern.ch

Benchmarking LHCb applications

HEPiX ”bazaar and thinktank” / 58

EGEE status and plans

Author: Bob Jones[1]

[1] *CERN*

EGEE status and plans

Site reports / 59

BNL Site Report

Author: Robert Petkus[1]

[1] *Brookhaven National Laboratory*

Corresponding Author: rpetkus@bnl.gov

RHIC and USATLAS Computing Facility site report

HEPiX ”bazaar and thinktank” / 60

Indico at DESY

Author: Alexander Krieg[1]

[1] *DESY*

Corresponding Author: alexander.krieg@desy.de

The Indico computer system at DESY consist of loadbalanced application servers, a central database server and a media archive on high availability storage. Currently, more than 700 events and users are organised in Indico at DESY. This talk will give an overview about the
setup, the experiences we had and how this service is organised and supported at DESY.

Site reports / 61

SLAC site report

Author: Chuck Boeheim[1]

[1] *SLAC*

SLAC site report

62

Organisational matters

Author: Helge Meinhard[1]

[1] *CERN*

Site reports / 63

CNAF site report

Author: Vladimir Sapunenko[1]

[1] *INFN*

CNAF site report