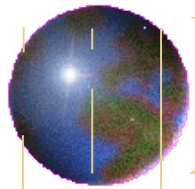


The ALICE Grid T2 in Mexico
First operation experiences

Luciano Diaz ICN-UNAM

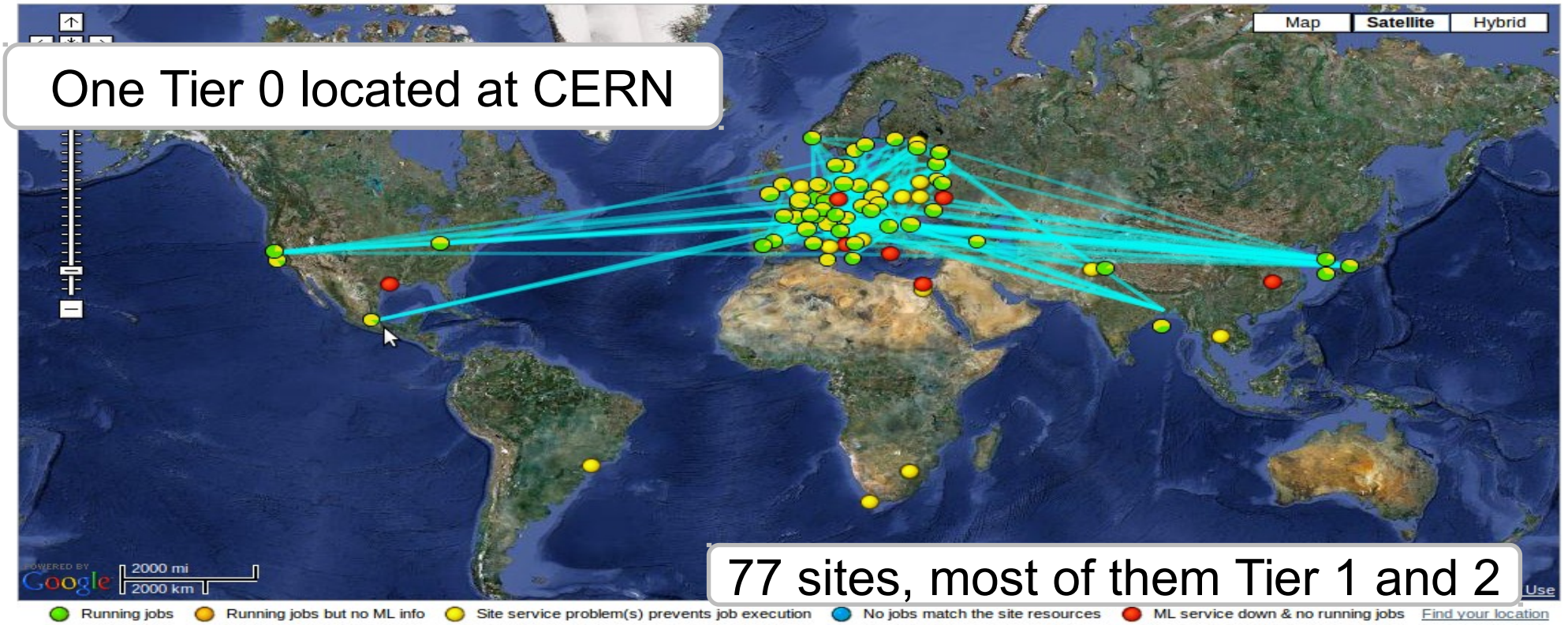
November 2014

Outline

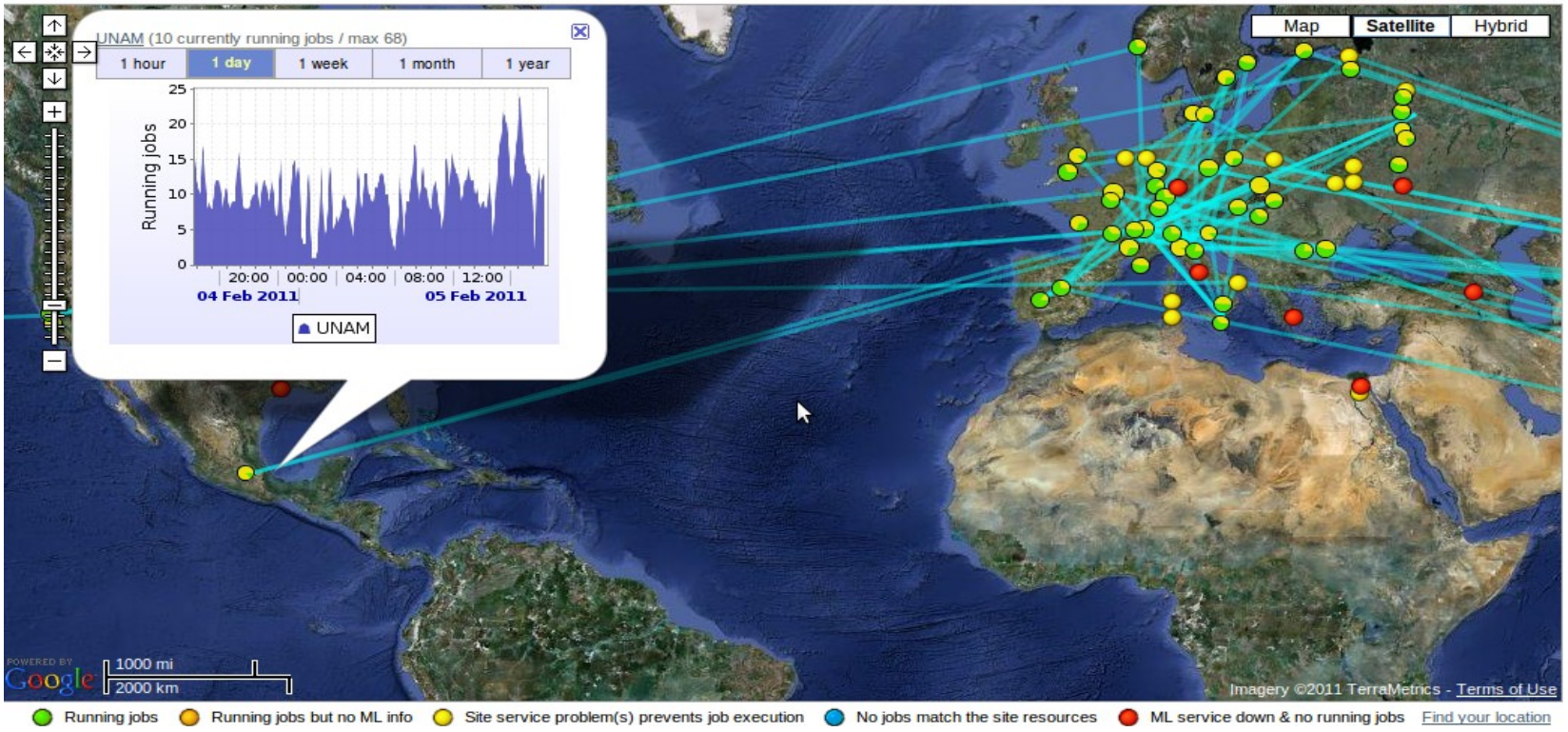
- Introduction
- History
  - 2006 - 2011
  - 2011 - 2014
- Installation T2
- Current status: site SUPERCOMPUTO-UNAM
- Conclusions

Introduction

The ALICE experiment has based their data and computing model on a GRID architecture based on several sites distributed around the world.



Mexico is part of ALICE Grid since 2007 when it joined the ALICE infrastructure through the site ICN-UNAM.



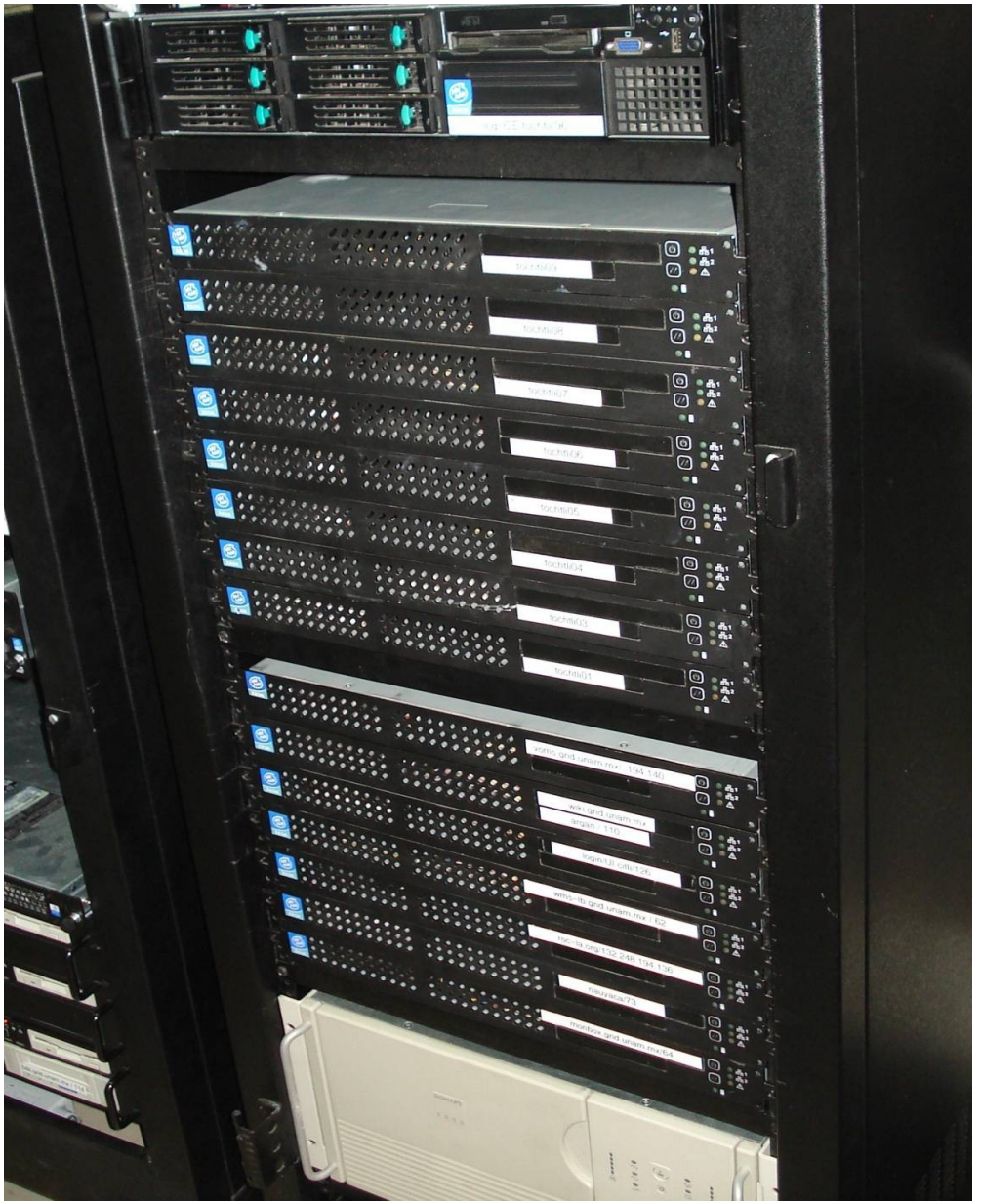
History

On **2006** the adventure began with tochtli, a cluster with 16 nodes:

- 32 Xeon 2.4GHz cores and 500MB of memory per core.
- A 30 Mbps CUDI connection used mainly for videoconferences.

Site ICN-UNAM on 2006

tochtli



2006

June 2006

First hardware upgrade of tochtli to run ALICE jobs:

- Upgrade of memory: 500 MB to 1.5 GB per core.
- 1TB of disk for a Storage Element required
- First steps to create a Resource Center with the name: site EELA-UNAM

September 2006

- Alice VoBox installed but with ResourceBrokers located at:

Brazil - UFRJ
Spain – CIEMAT

Network problem: latency up to 400 ms.

2006

At the end of 2006:

The main concern is the network:

- Mixed routing between CUDI and commodity links.
- Low bandwidth and high latency.

2007

- January 2007:
  First ALICE jobs received.

- March 2007:
  More UNAM resources added to the EELA Grid:
  IBT,CCG,FESC and **Supercomputo**

2007

- April 2007:

Network characterization, improvement of the network configuration: UNAM-CUDI-CLARA

  - TCP tuning on nodes
  - Symmetric routing
  - Avoid traffic to Europe from Sao Pablo

- September 2007:
  UNAM grid-CA ready, waiting for IGTF green light.

2008

February 2008:

UNAMGRID-CA: accredited by IGTF, start operations

March 2008:

Network:
Maintenance of the optical fiber from ICN to DGSCA (DGTIC today).

April 2008:

Middleware upgrade from gLite 3.0 to gLite 3.1

June 2008:

First 64 bits Xeon processors arrives to ICN

2008

December 2008:
Network improvement:
Second E3 to CUDI, duplicated the BW to 60Mbps.

2009

February 2009:
ICN resources under the Resource Center: **ICN-UNAM**.

September 2009:
First advice: The 32bits arch will be discontinued from ALICE production sites

ICN network improvement:
New optical fiber@1Gbps from ICN to DGSCA
Subnetwork with 64 IPs just for cluster access.
Direct link from ICN to I2 with jumbo frames support.

2009

September 30:
ROC-LA start operations

December 2009:
First 40 cores - 64 bits arch.

2010

January 2010:

UNAM site blocked from ALICE production due to obsolete 32 bits arch.

March 2010:

ICN-UNAM joint to ALICE production with 64 bits nodes.

Middleware upgrade to: gLite 3.2, including CREAM-CE and VOBOX

2010

June 2010:

New cluster, financed by CTIC arrives

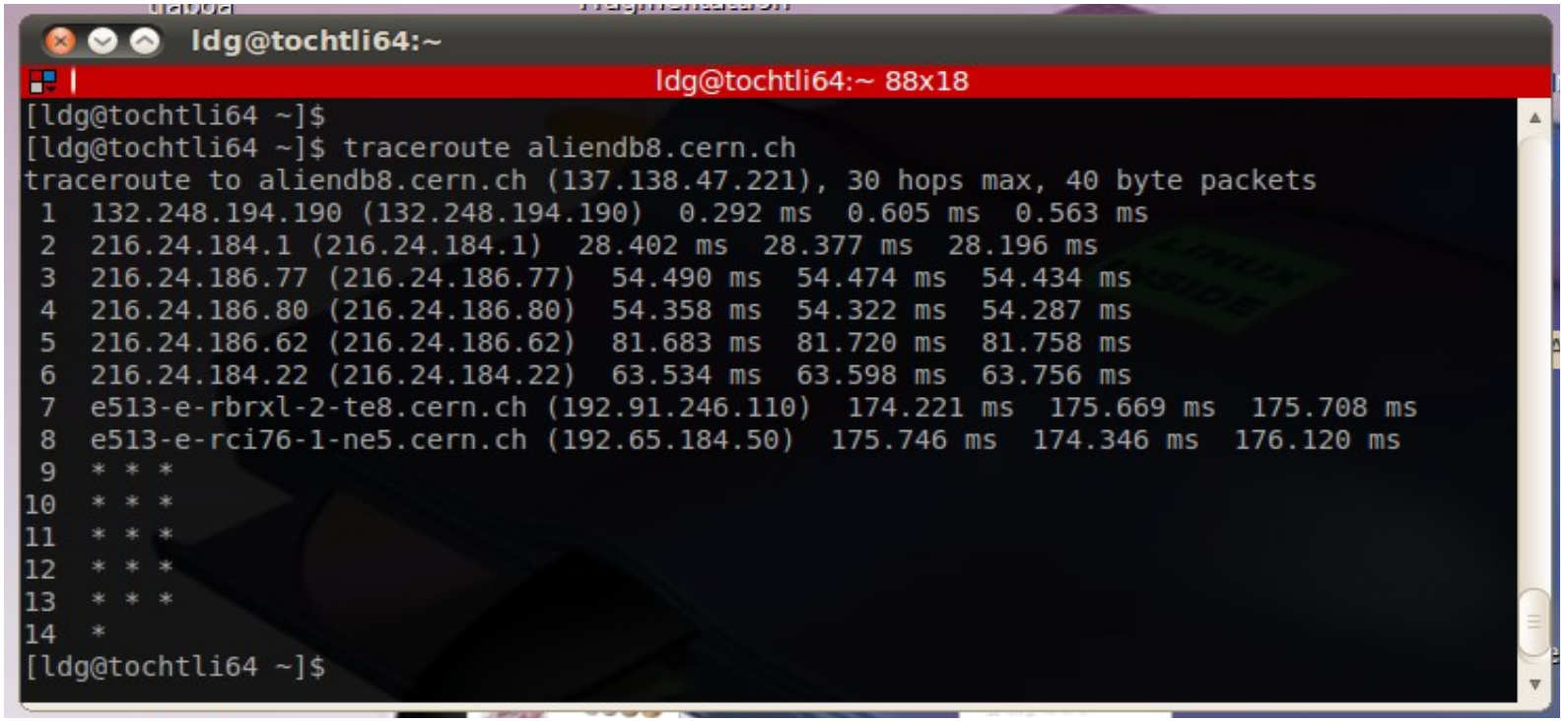
- all 64 bits, 104 cores
- added to tochtli64

for local and grid use.

2010

October 2010:

Second network improvement:
1 Gbps link from UNAM to USA, but it will need more configuration to be useful for Grid purposes.



Status

- Site ICN-UNAM on ALICE and AUGER production infrastructure
- VO's supported: AUGER, EELA, BIOMED
- Certified site on EGI infrastructure through ROC-LA
- Support of the GISELA applications through EELA VO
- Up to 146 cores with 64 bits
- Direct connection from ICN to I2 at 1Gbps with Jumbo frames support

Site ICN-UNAM on 2011

tochtli64

2011

January 2011:
Workshop: “Grid Computer Center of the Americas”
Plans to install the First T1 for ALICE in America.

Goal:
1000 cores
1 PB storage facility

2011

Also it will need:

- Infrastructure (space, cooling and electric power / backup)
- Network lot of improvement (local and external)
- CA operation
- Human resources

ICN and DGTIC will work together to achieve the T1.

The ICN-UNAM was the first ALICE node in Latin America until 2010 and is interested to be the first Tier-1 for ALICE in the region.

2011

Febrary 2011:
Proposal to create a testbed at DGTIC with:
100~150 cores.
2~5 TB storage.

March 2011:
Registration of the site SUPERCOMPUTO-UNAM in GOCDB
Basic resource center:
Site-BDII,CREAM-CE,VOBox
Waiting for CPU resources to be released. Expected: 136 cores.

2011

May 2011:
The number of cores is reduced to 32 cores due to high CPU demand. Waiting to be released.

June 2011:
DGTIC purchase plan for: T1, HPC, data warehouse, cloud, ..

October 2011:
Network problem:
SUPERCOMPUTO-UNAM BW:
To ICN: 125 Mbps
Outside UNAM: 25 Mbps

2011-2012

Due to the size of the new UNAM supercomputer the process to buy the cluster took almost 1 year since the purchase.

2012

July 30, 2012:
Miztli: New supercomputer at DGTIC arrives to DGTIC.

August 2012:
SUPERCOMPUTO-UNAM testbed removed from GOCDB, due to new hardware installation.

2013

April 2013:
Network bottleneck solved:
1Gbps link UNAM- USA /San Antonio fully working at least for ICN.

Notification: Miztli CPUs and storage ready, but not yet for ALICE.

Resume meetings for T1.

2013

May 2013:
In a meeting with Paolo Giubellino was decided to start with a T2 and later upgrade to a T1.

June 2013:
Discussion about configuration of Miztli for the ALICE T2.

September 2013:
T2 won be part of Miztli , decision to buy new hardware.

Decision about the configuration of the cluster:
Intel or AMD, Memory, disk space for CVMFS cache.

2013

November 2013:
Network problem:
Asymmetric route between ICN and SUPERCOMPUTO-UNAM.
SUPERCOMPUTO-UNAM up to 100 Mbps.
The storage hardware will be provided for HP.
Selection of the vendor for the T2 nodes.

December 2013:
xrootd installed for Storage system but will be replaced later with EOS.

2014

March 2014:
New hardware arrives:
1024 cores (with HT) Intel and 4 GB memory

Another network problem detected:
Intrusion Prevention Systems (IPS) found – Limit the BW to 300 Mbps.

April: 2014:
All the Hardware is connected: Nodes, storage and network.

2014

June-July 2014:
During the visit to CERN
• Configuration of the nodes
1024 cores finally added to ALICE production.
• EOS installation in the storage system
580 TB storage system in 5 data servers.

The 1 Gbps interface in each data servers is saturated, it will need at least 10 Gbps for each data server: Plans to upgrade the cluster network.

2014

August 2014:
Network improvement:
Cluster out of IPS – BW now to 1 Gbps.
1Gbps link UNAM- USA /San Antonio fully working for SUPERCOMPUTO-UNAM too.

Hardware errors: Memory faulty in all the nodes.

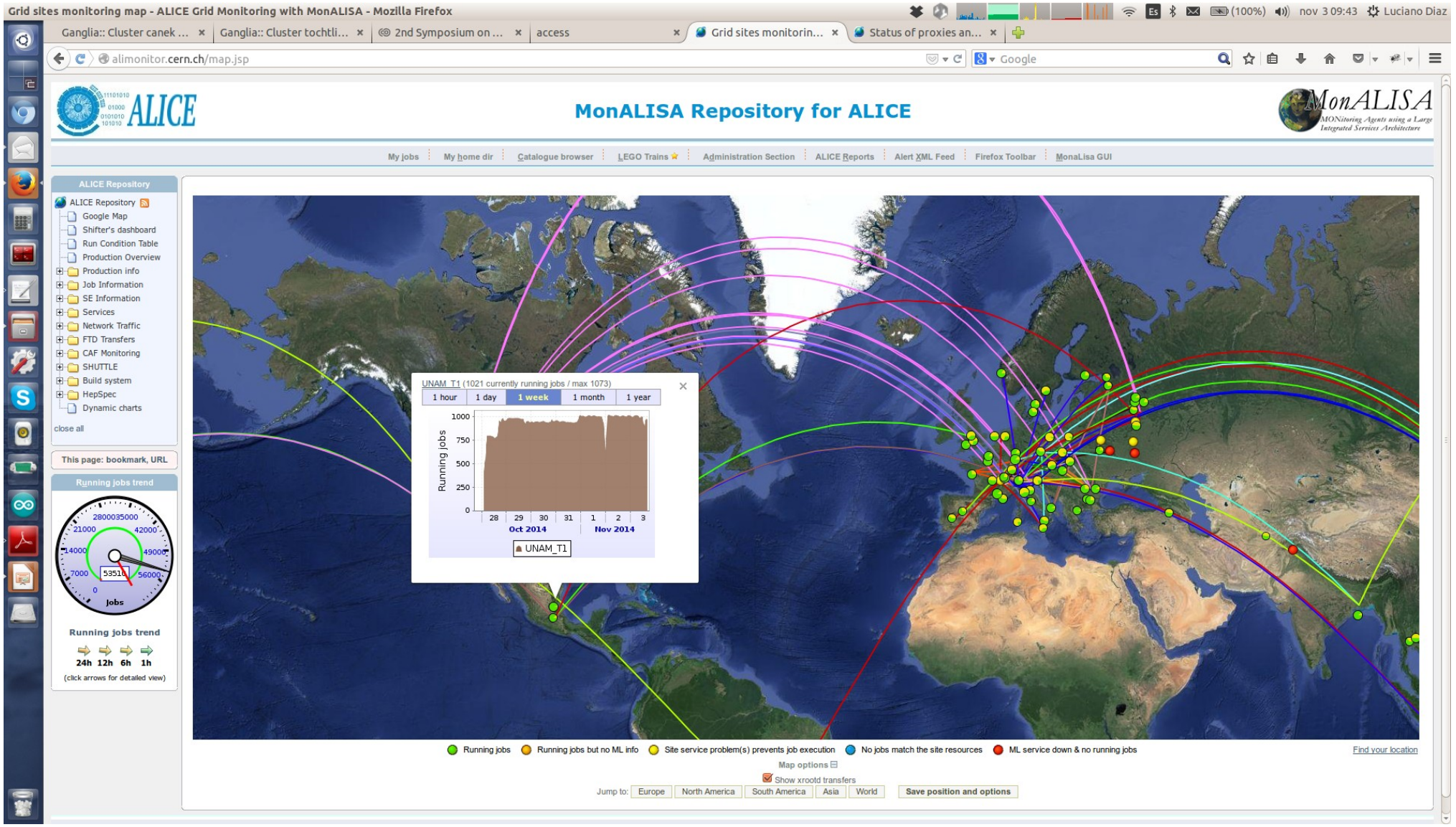
October 2014:
256 Memory modules replaced (All nodes memory).
Cluster working at full capacity.
New switch for the cluster network already requested, must arrive in a couple of weeks.

Today

Success:
1024 cores with 4 GB of memory is in full load running ALICE jobs.

570 TB of storage only for ALICE. One of the firsts ALICE sites using EOS instead of xrootd.

Connected to the 1 Gbps link, configured and tunned for Grid operations.


Grid sites monitoring map - ALICE Grid Monitoring with MonALISA - Mozilla Firefox
Ganglia:: Cluster canek ...
Ganglia:: Cluster tochtli...
2nd Symposium on ...
access
Grid sites monitorin...
Status of proxies an...
alimonitor.cern.ch/map.jsp
ALICE
MonALISA Repository for ALICE
MonALISA
My jobs
My home dir
Catalogue browser
LEGO Trains
Administration Section
ALICE Reports
Alert XML Feed
Firefox Toolbar
MonaLisa GUI
ALICE Repository
ALICE Repository
Google Map
Shifter's dashboard
Run Condition Table
Production Overview
Production info
Job Information
SE Information
Services
Network Traffic
FTD Transfers
CAF Monitoring
SHUTTLE
Build system
HepSpec
Dynamic charts
close all
This page: bookmark, URL
Running jobs trend
Running jobs trend
24h 12h 6h 1h
(click arrows for detailed view)
UNAM_T1 (1021 currently running jobs / max 1073)
1 hour
1 day
1 week
1 month
1 year
Running jobs
Oct 2014
Nov 2014
UNAM_T1
Running jobs
Running jobs but no ML info
Site service problem(s) prevents job execution
No jobs match the site resources
ML service down & no running jobs
Find your location
Map options
Show xrootd transfers
Jump to:
Europe
North America
South America
Asia
World
Save position and options

2014

Without success yet:

Lack of human resources for technical operation

Only 2 people worked partial (spare) time for the T2 setup, but nobody is assigned full time for the daily operation and maintenance.

8 years summary

Hardware
CPU
32 → 1024 cores
Storage
0 → 580 TB

Network
35 Mbps → 60 Mbps → 1 Gbps
Several improvements:
Cluster network.
Campus network.
Academic network.
- Routing
- Tunning

8 years summary

Personnel for Resource Center operation.
1 → 2 people.

but academic people
1 → 1

Working CA

Middlware migrations : gLite 3.0, 3.1, 3.2, EMI-1, EMI-2, EMI-3

The T1 project has survived thanks to the really big effort from the UNAM to improve and provide all the necessary infrastructure for the ALICE experiment and the invaluable support from the ALICE Grid experts.

Thanks for all the support provided:

ICN

Dirección : **Dr. Miguel Alcubierre**

Departamento de Altas Energías: **Dr. Guy Paic, Dr. Lukas Nellen**

Unidad de Cómputo y Seguridad Informática

DGTIC

Dirección: **Dr. Felipe Bracho**

Dirección de sistemas y servicios institucionales: **Fabián Romo**

Supercomputo: **Jose Luis Gordillo, Arión Perez**

Seguridad de la Información

Firma Electronica Avanzada

Dirección de telecomunicaciones

CERN-ALICE: Paolo Giubellino, Latchezar Betev, Yves Schutz

CUDI

CONACYT