



El grid de computo para ALICE en México

Tochtli, el primer nodo en América Latina

Lukas Nellen
13 de febrero de 2007



Introducción



- ¿Que es un grid?
- ¿Por que grids?
- **ALICE**
 - El experimento
 - El modelo de computo
- El cluster tochtli
 - El hardware
 - Configuración
 - El grid
- Resumen y planes



¿Que es un GRID?



- **Infraestructura avanzada de computo de alto rendimiento**

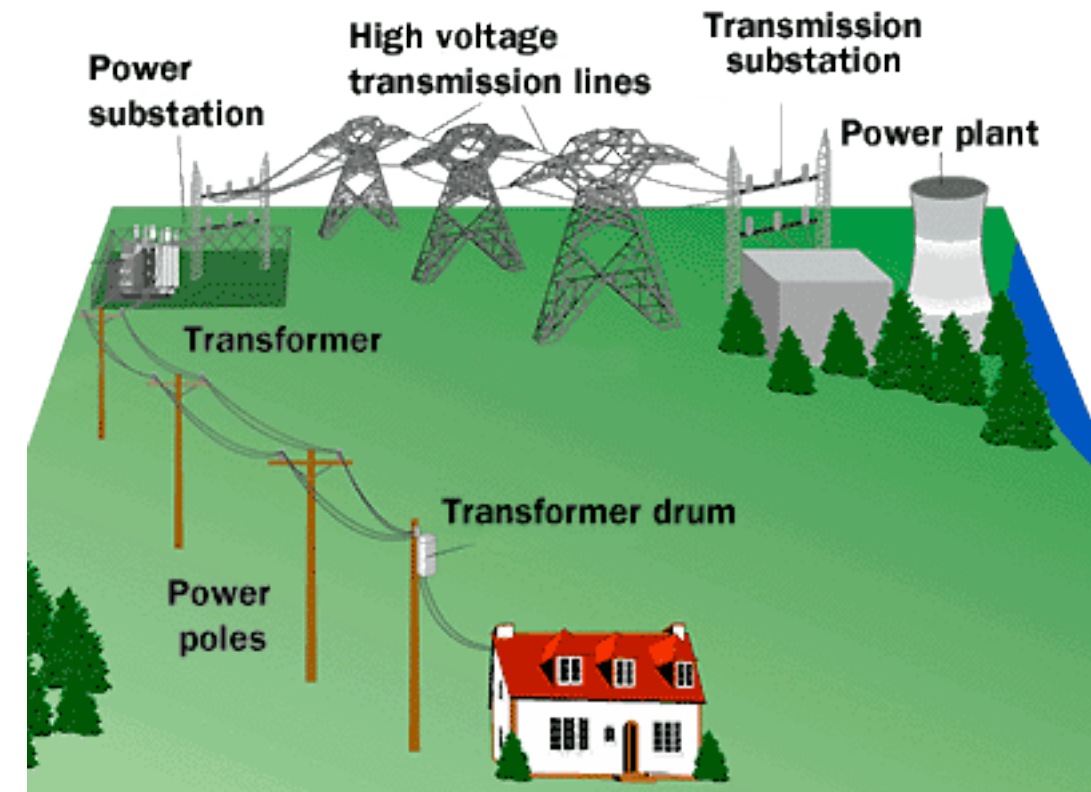
- **Basado en recursos compartidas:**

- contribuidas
- donadas

- **Transparente para el usuario**

- **Seguridad**

- autenticación
- autorización



Red de electricidad
"Power Grid"



Tipos de GRID



- **Formal, con organización central**
 - 📌 Grid europeo **EGEE** (<http://www.eu-egee.org/>)
 - 📌 Grid científico en EEUU: **Open Science Grid** (<http://www.opensciencegrid.org>)
- **Informal, peer-2-peer**
 - 📌 **XtremWeb** (<http://www.xtremweb.org/>)
 - 📌 **OurGrid** (<http://www.ourgrid.org/>)
- **Para proyectos específicos, recursos donados**
 - 📌 **BOINC** (<http://boinc.berkeley.edu/>)
 - * SETI@home, Einstein@home, ...



¿Porque GRID?



- **Permite juntar recursos de computo geográficamente distribuidas**
 - **manejo de datos**
 - **acceso a CPU**
- **Interfaces estándares**
- **Servicios de seguridad**
 - **Identificación de usuarios**
 - **Autorización por grupos o proyectos**



Aplicaciones en GRID



- ✓ **Ideal para trabajos seriales: sistema de colas a gran escala**
 - ☞ Física de partículas
 - ☞ Industria petrolera
 - ☞ Bioinformática
 - ☞ Sismología
- ✓ **Manejo de datos, replicas de datos**
- ✗ **Soporte limitado para aplicaciones paralelas**
 - ☞ No es fácil garantizar tiempos de respuesta o ancho de banda entre centros lejanos
 - ☞ Beneficia de configuraciones espaciales
- ✗ **Poco soporte para aplicaciones interactivas**



EL experimento ALICE



- **Objetivo científico principal**
 - estudio de colisiones de iones pesados
 - formación del plasma de quarks y gluones
- **Uno de los cuatro experimentos en el colisionador LHC en el laboratorio CERN**



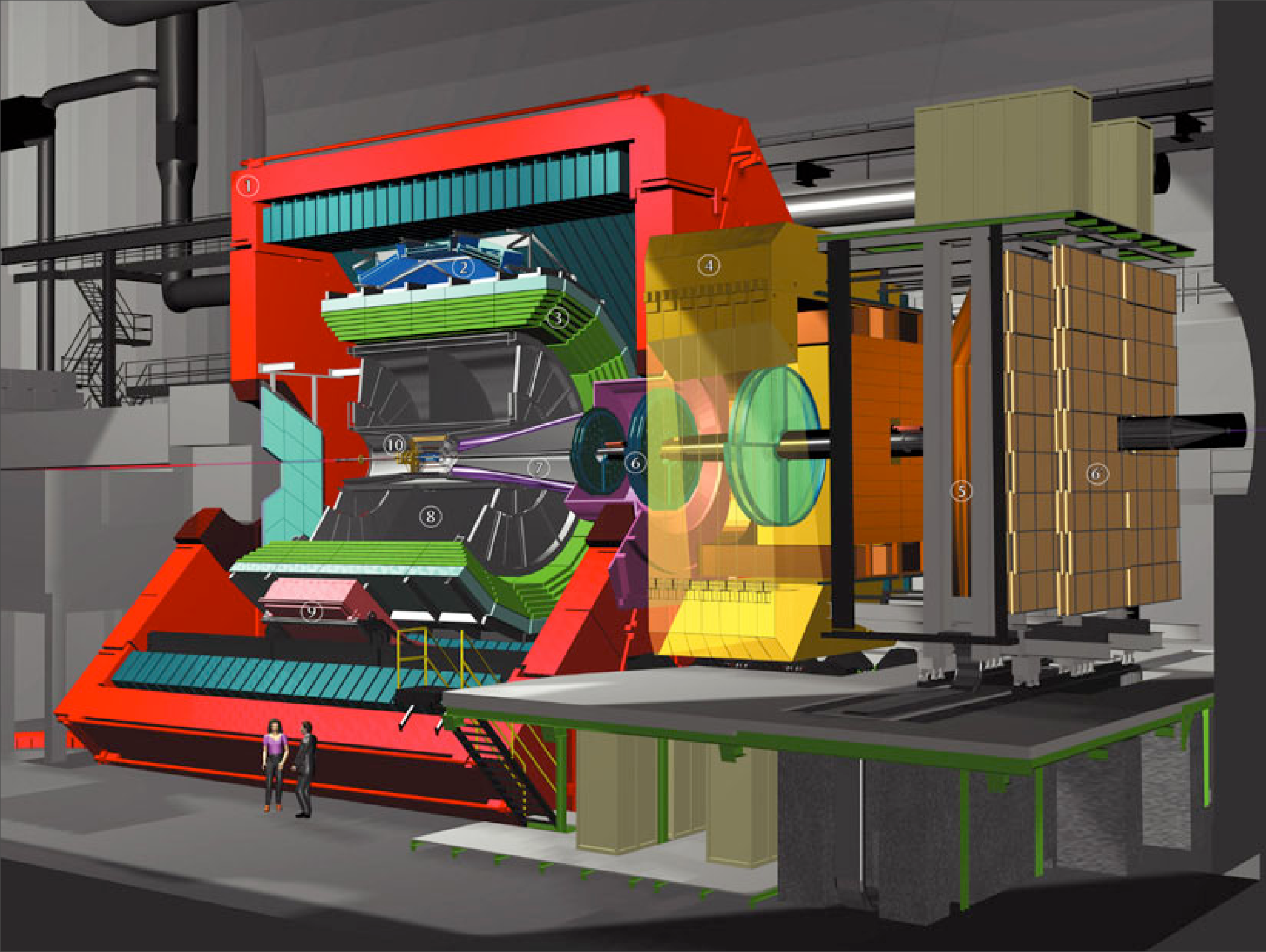
Large Heavy-ion Collider (LHC)



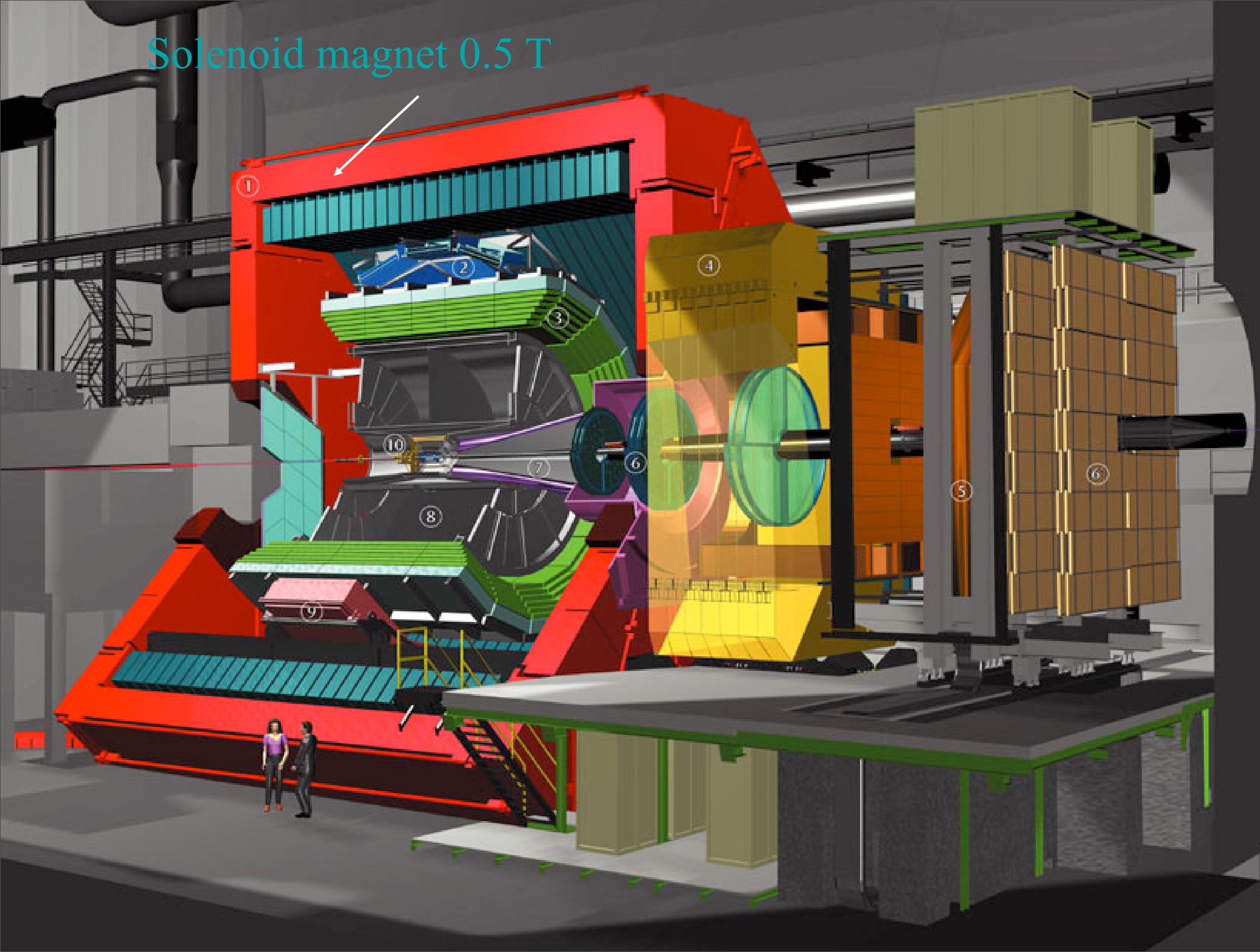


Large Heavy-ion Collider (LHC)





Solenoid magnet 0.5 T



Solenoid magnet 0.5 T

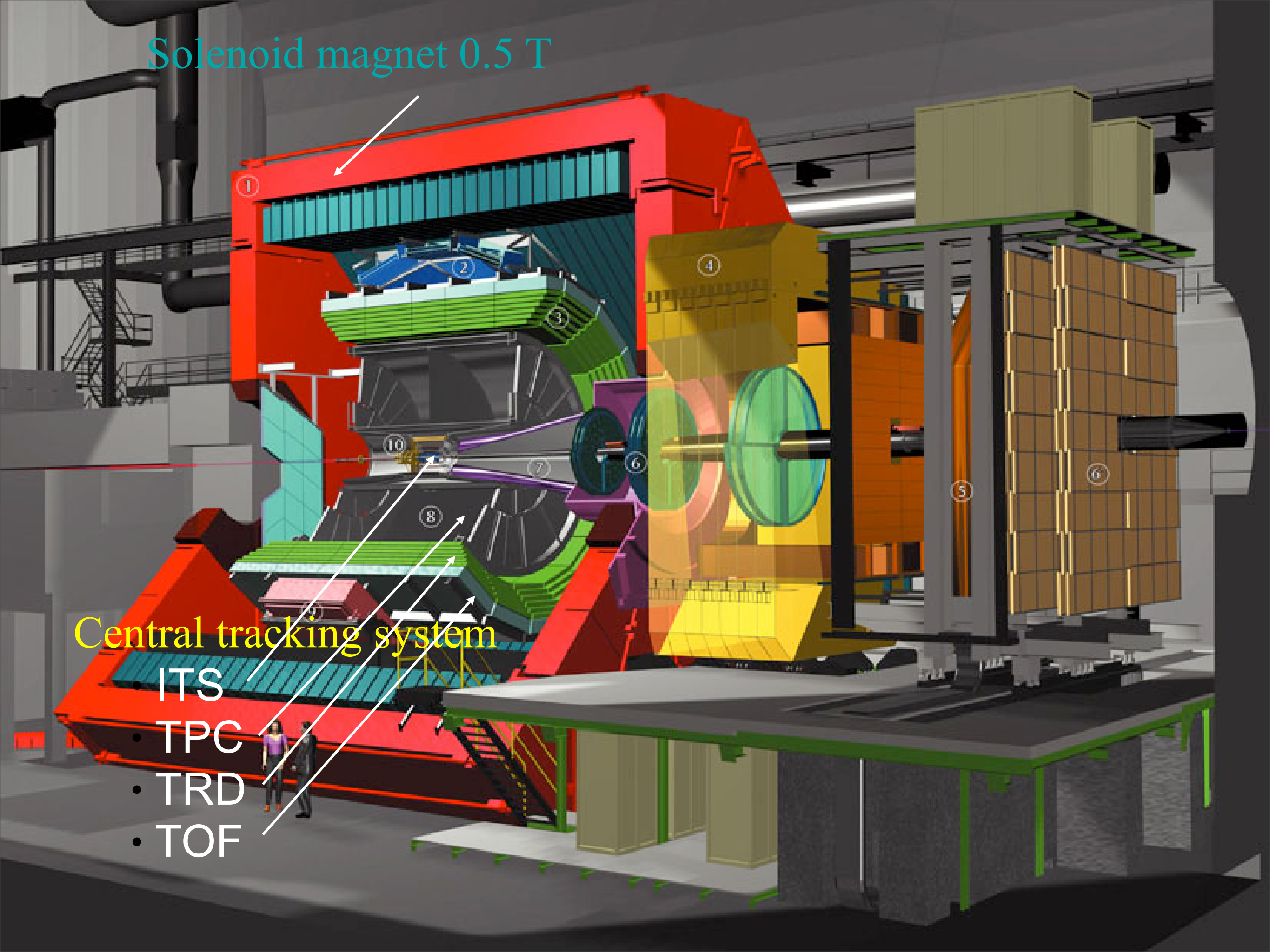
Central tracking system

ITS

• TPC

• TRD

• TOF



Solenoid magnet 0.5 T

Central tracking system

ITS

• TPC

• TRD

• TOF

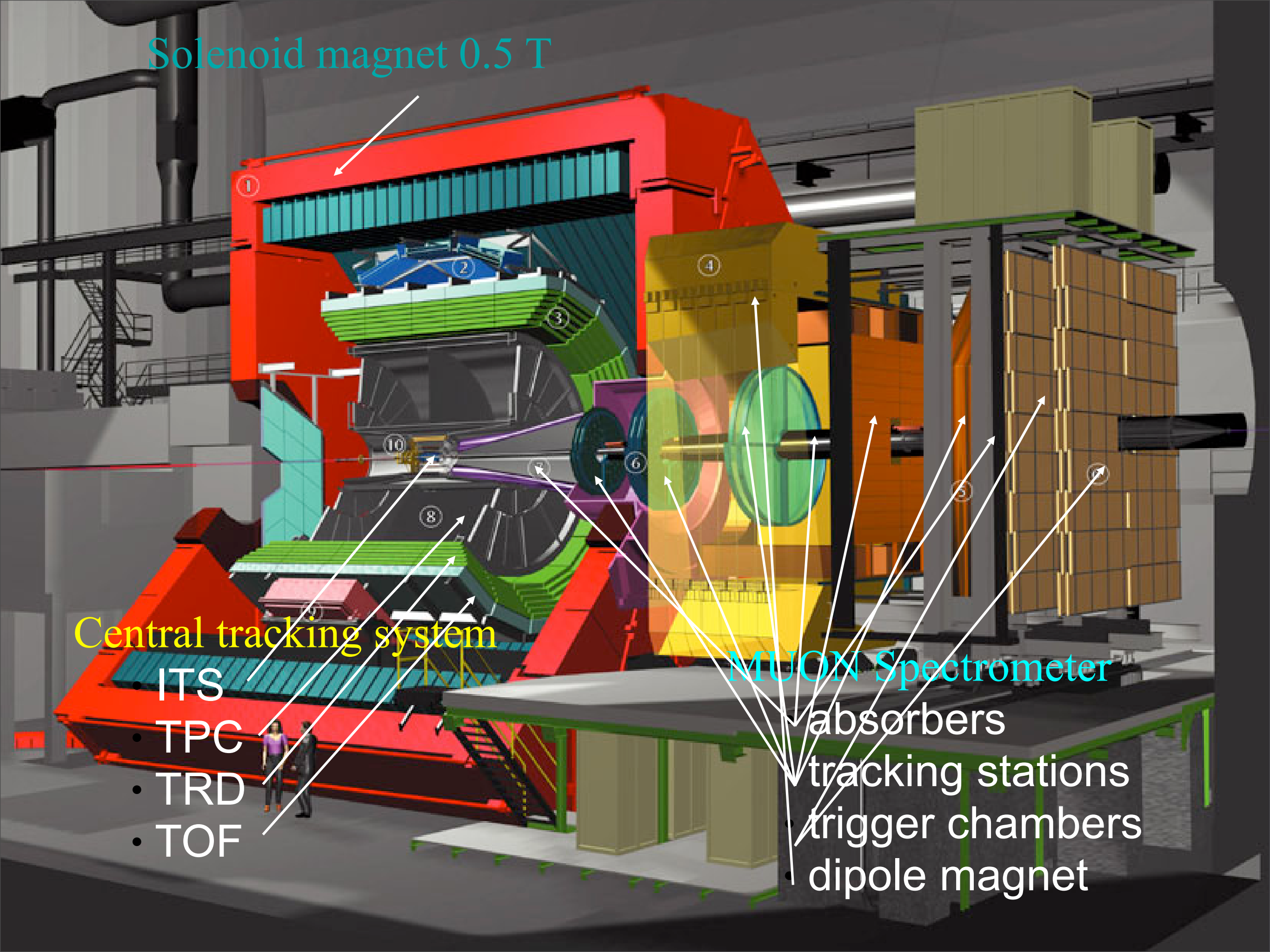
MUON Spectrometer

absorbers

tracking stations

trigger chambers

dipole magnet



Solenoid magnet 0.5 T

Specialized detectors

- HMPID

- PHOS

Central tracking system

- ITS

- TPC

- TRD

- TOF

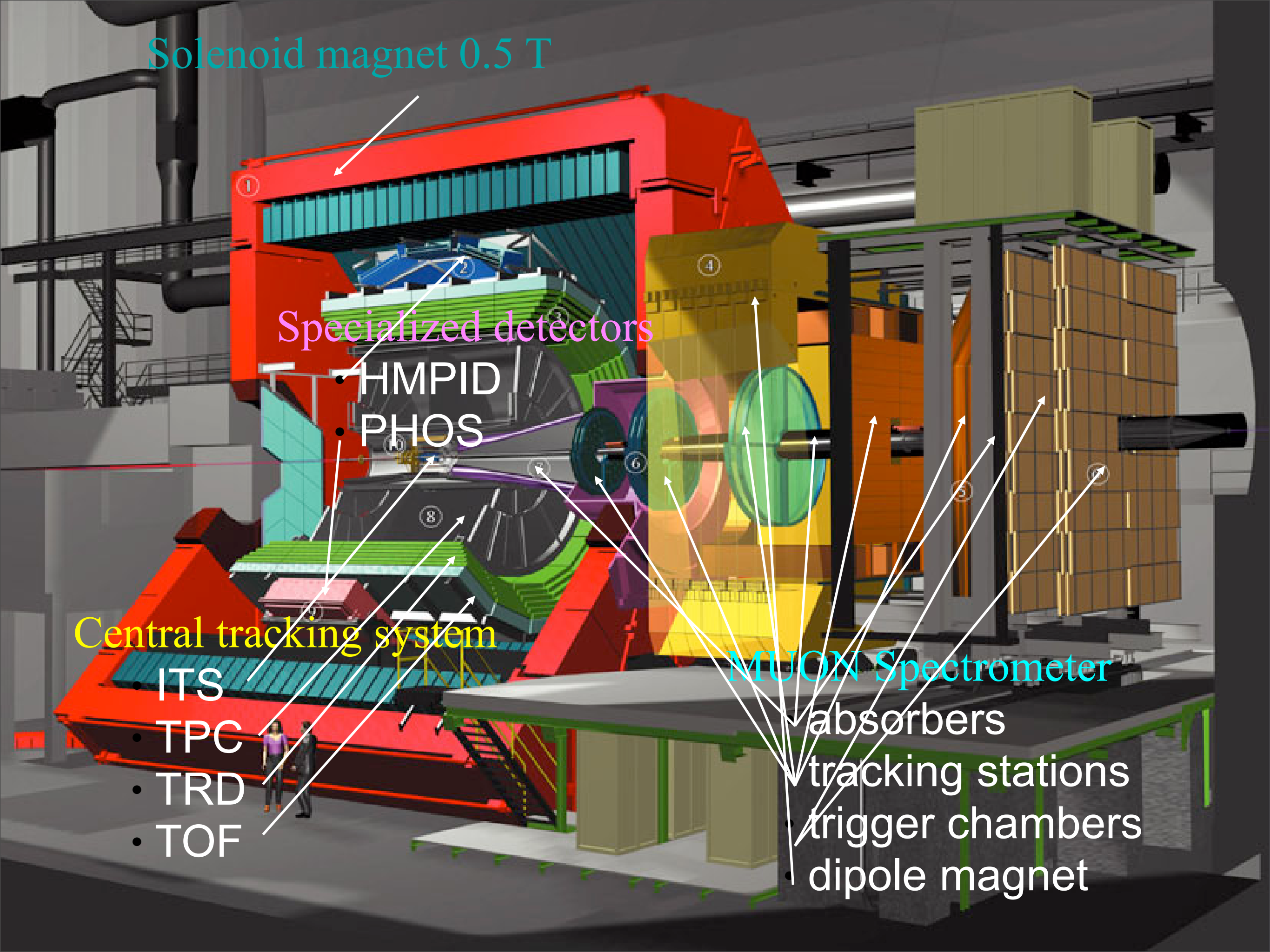
MUON Spectrometer

- absorbers

- tracking stations

- trigger chambers

- dipole magnet



Solenoid magnet 0.5 T

Forward detectors

- PMD
- FMD, T0, V0, ZDC

Specialized detectors

- HMPID
- PHOS

Central tracking system

- ITS
- TPC
- TRD
- TOF

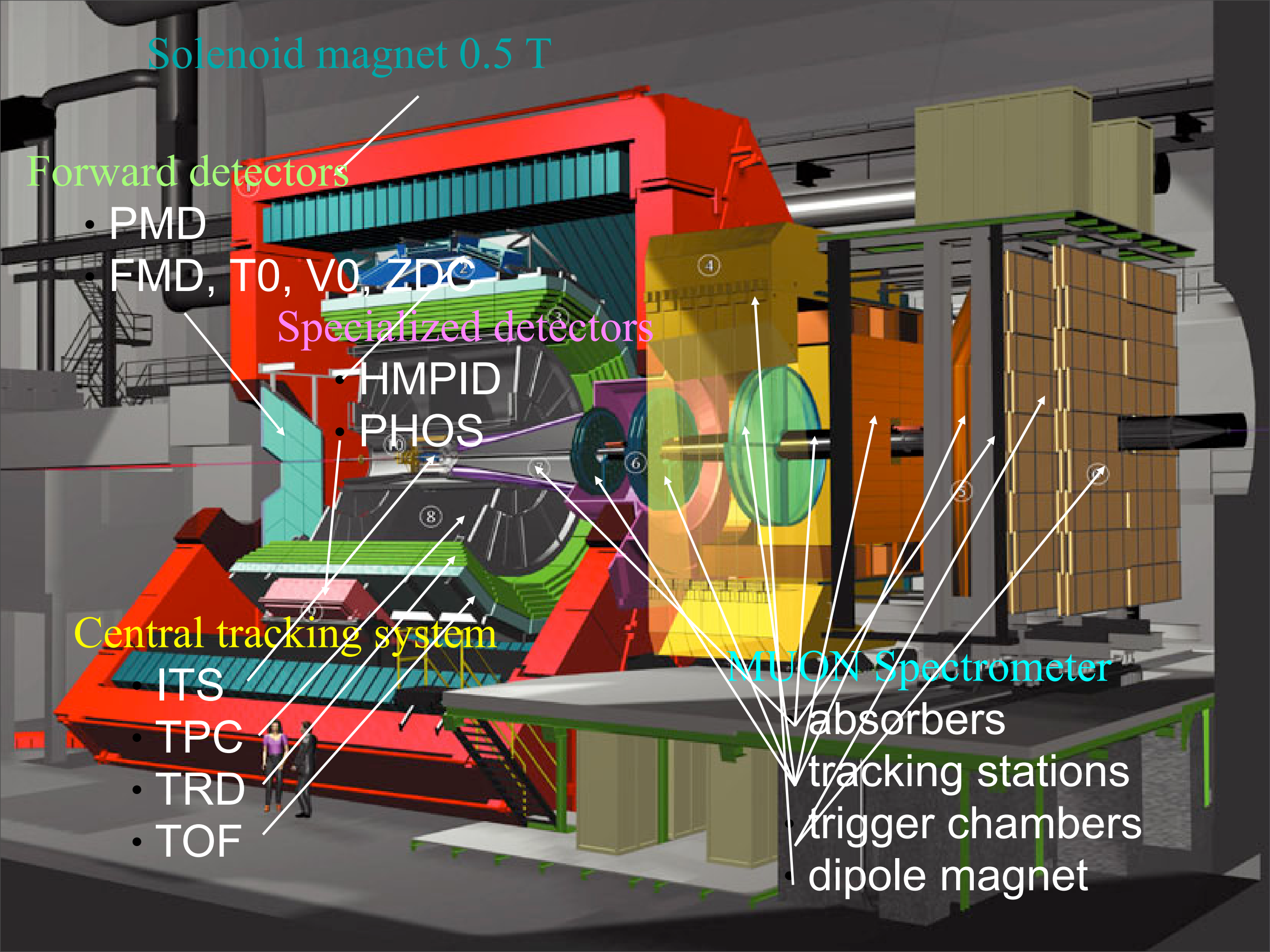
MUON Spectrometer

absorbers

tracking stations

trigger chambers

dipole magnet



Solenoid magnet 0.5 T

Cosmic-ray trigger

Forward detectors

- PMD
- FMD, T0, V0, ZDC

Specialized detectors

- HMPID
- PHOS

Central tracking system

- ITS
- TPC
- TRD
- TOF

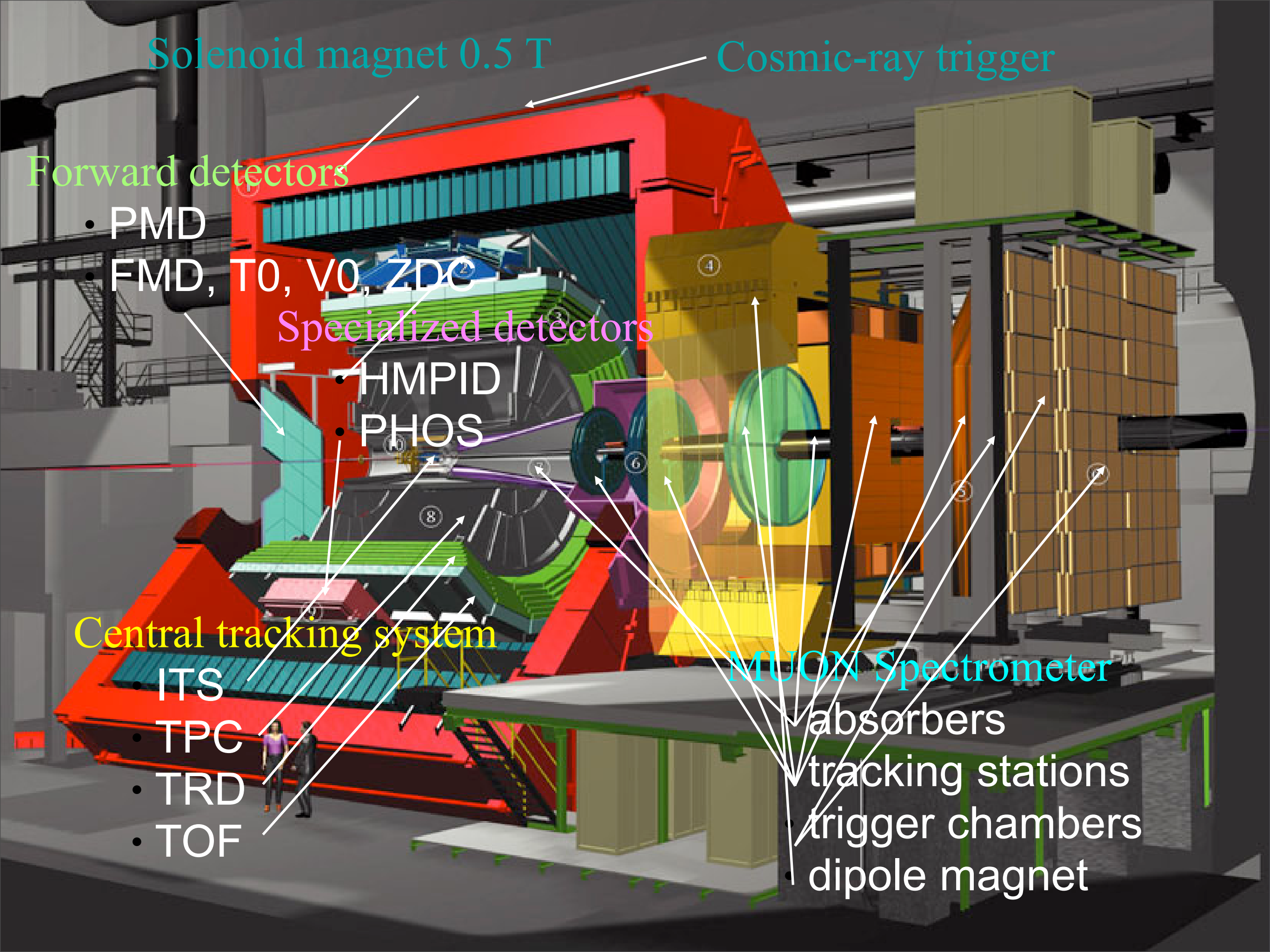
MUON Spectrometer

absorbers

tracking stations

trigger chambers

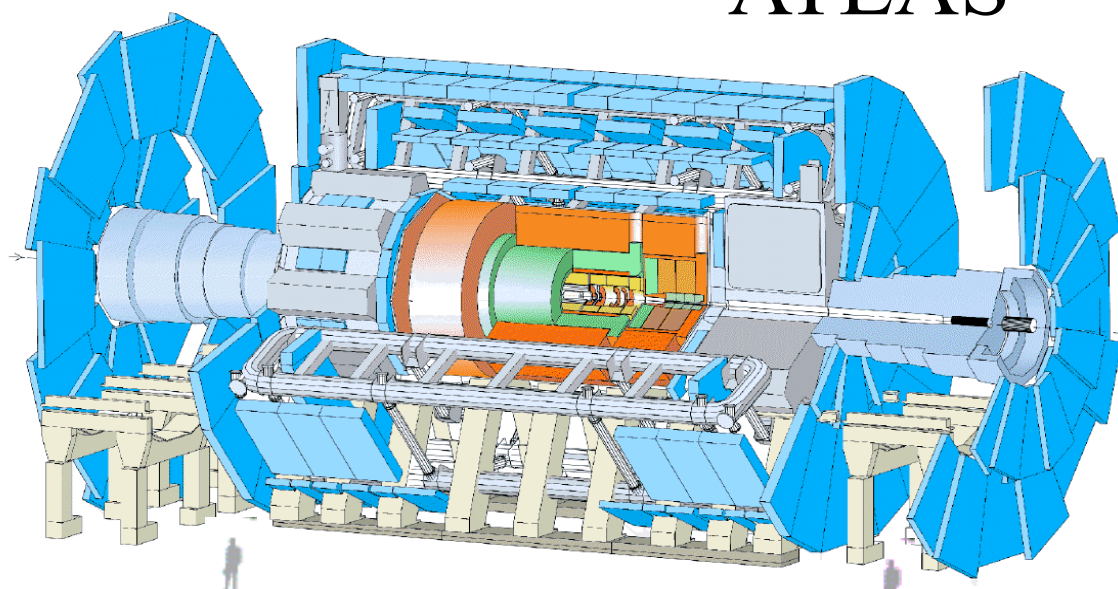
dipole magnet



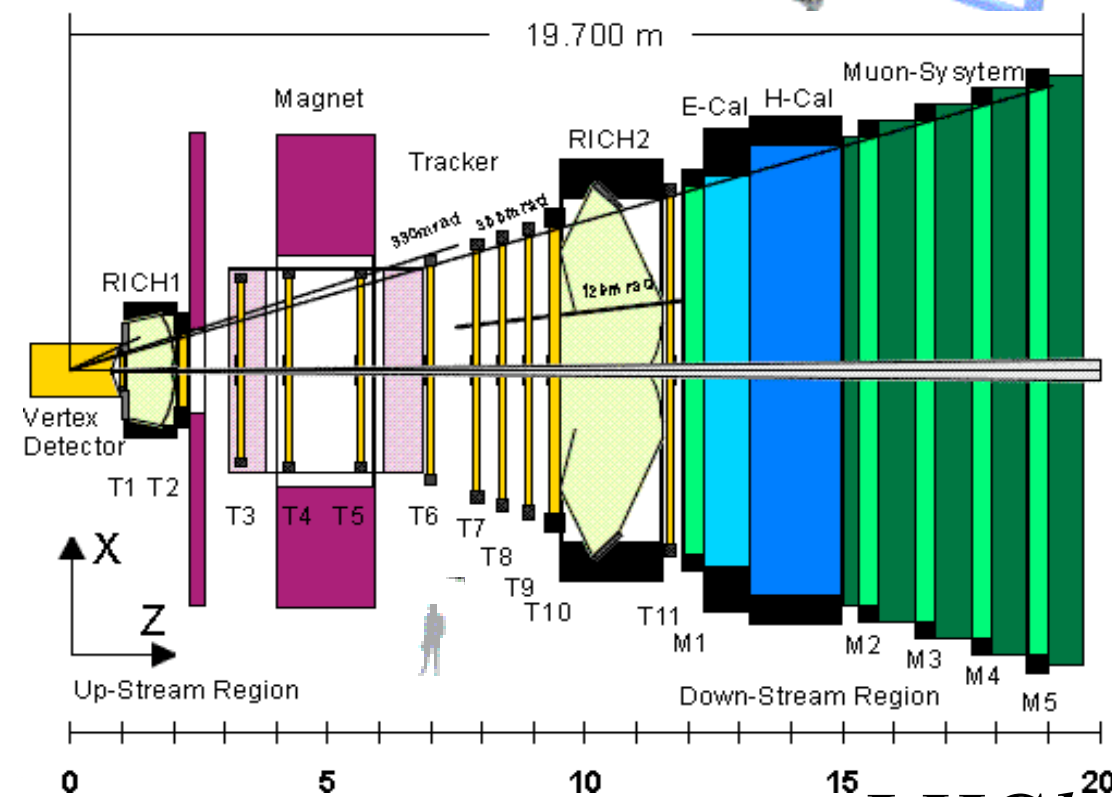
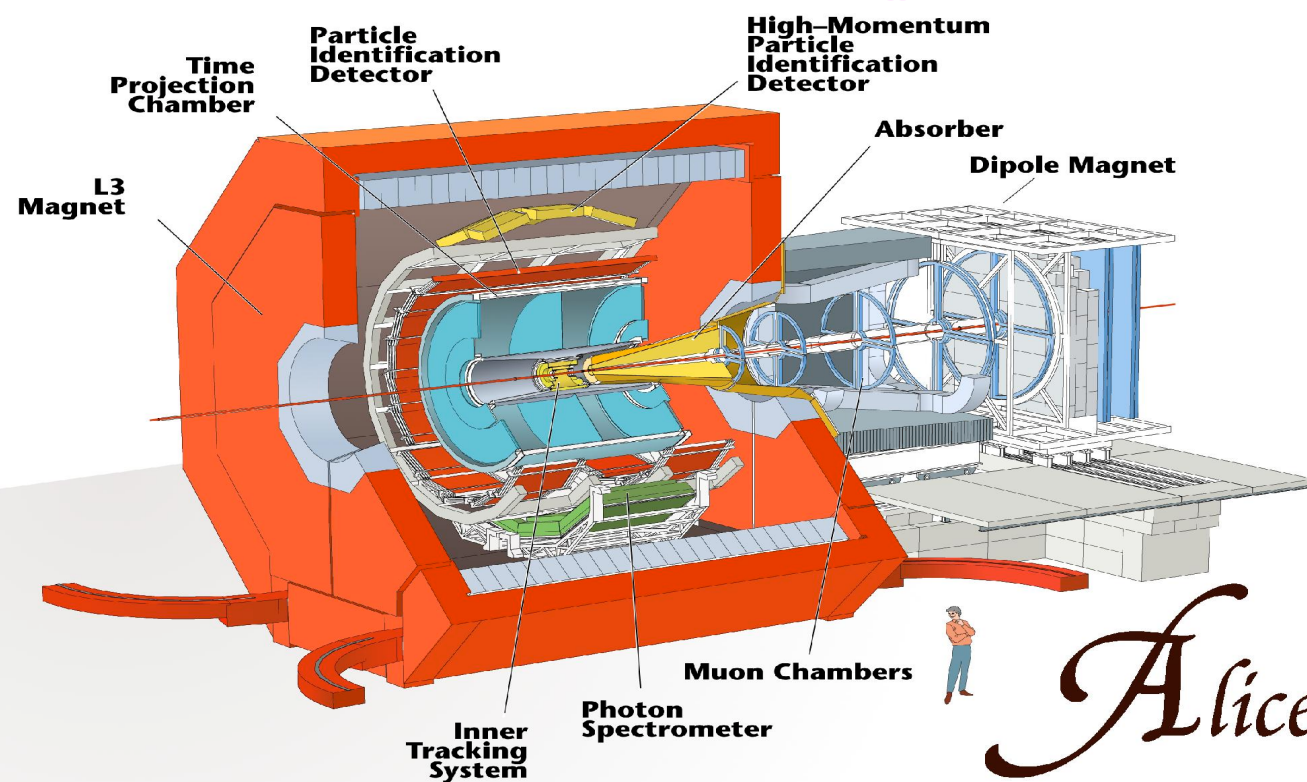
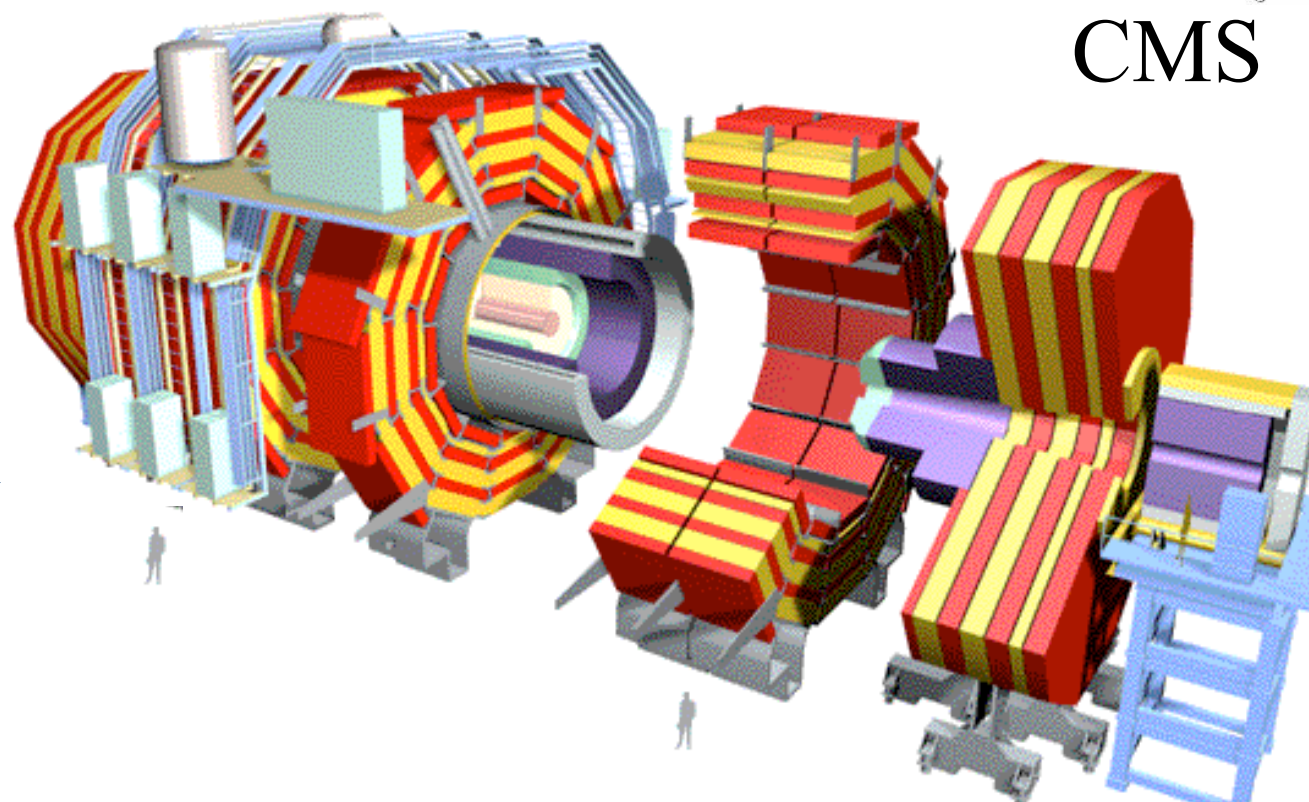


The LHC Detectors

ATLAS



CMS



LHCb

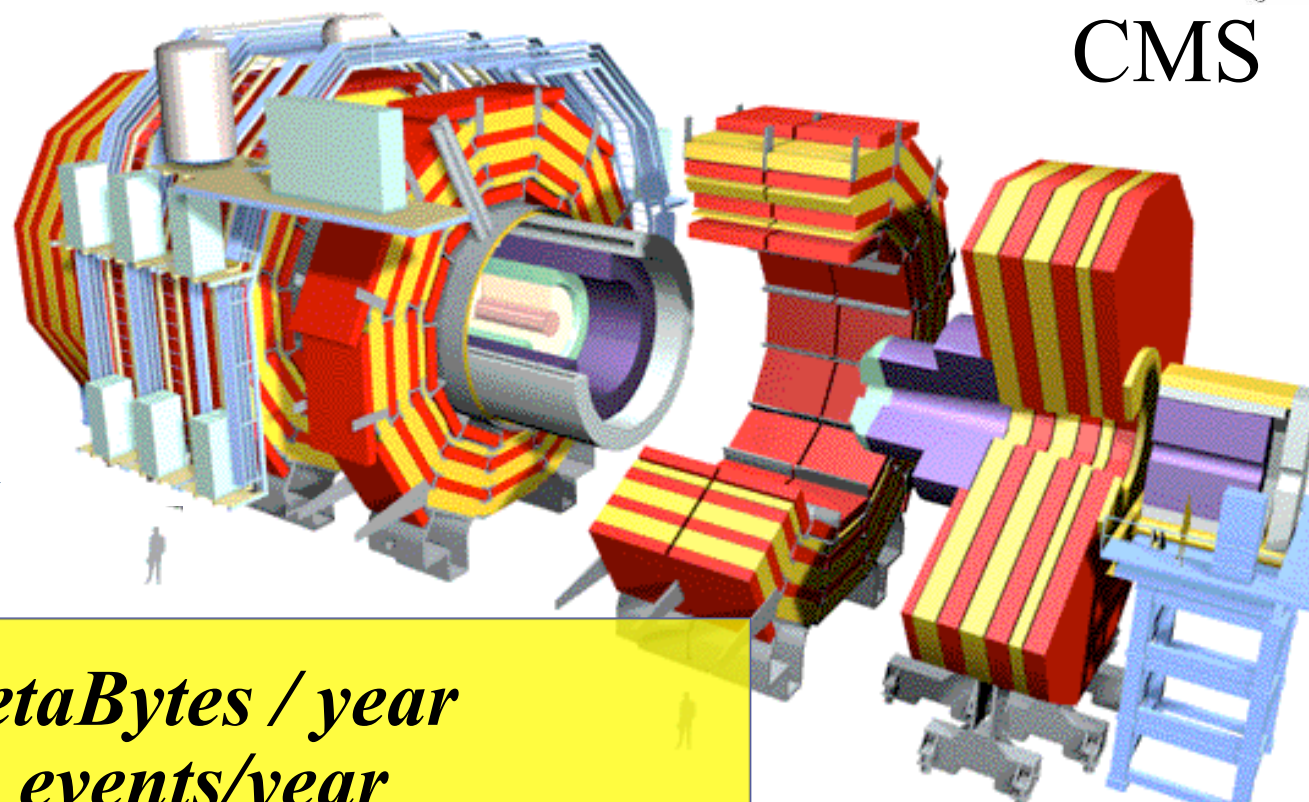
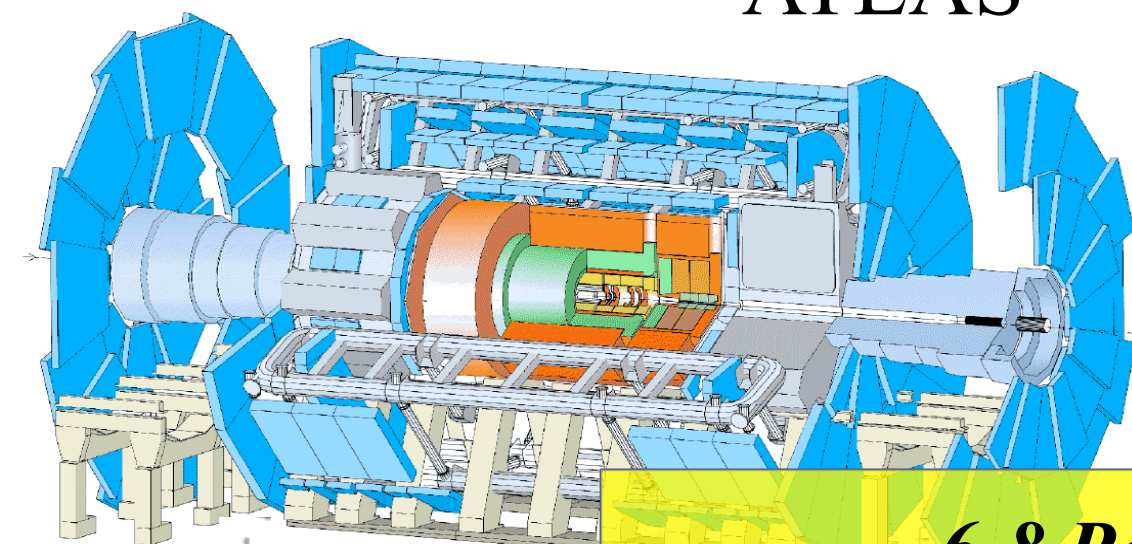




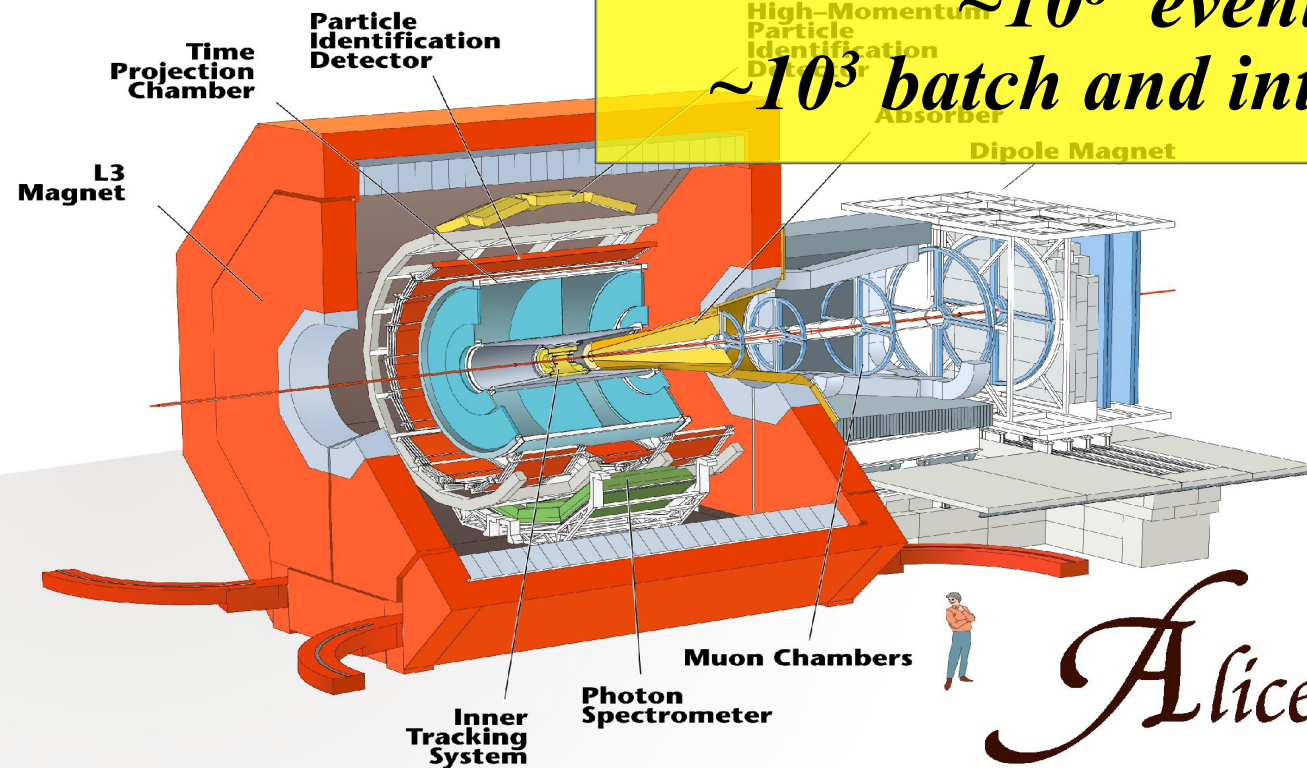
The LHC Detectors

ATLAS

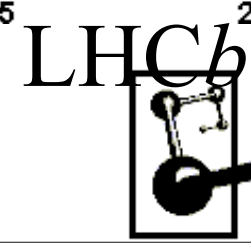
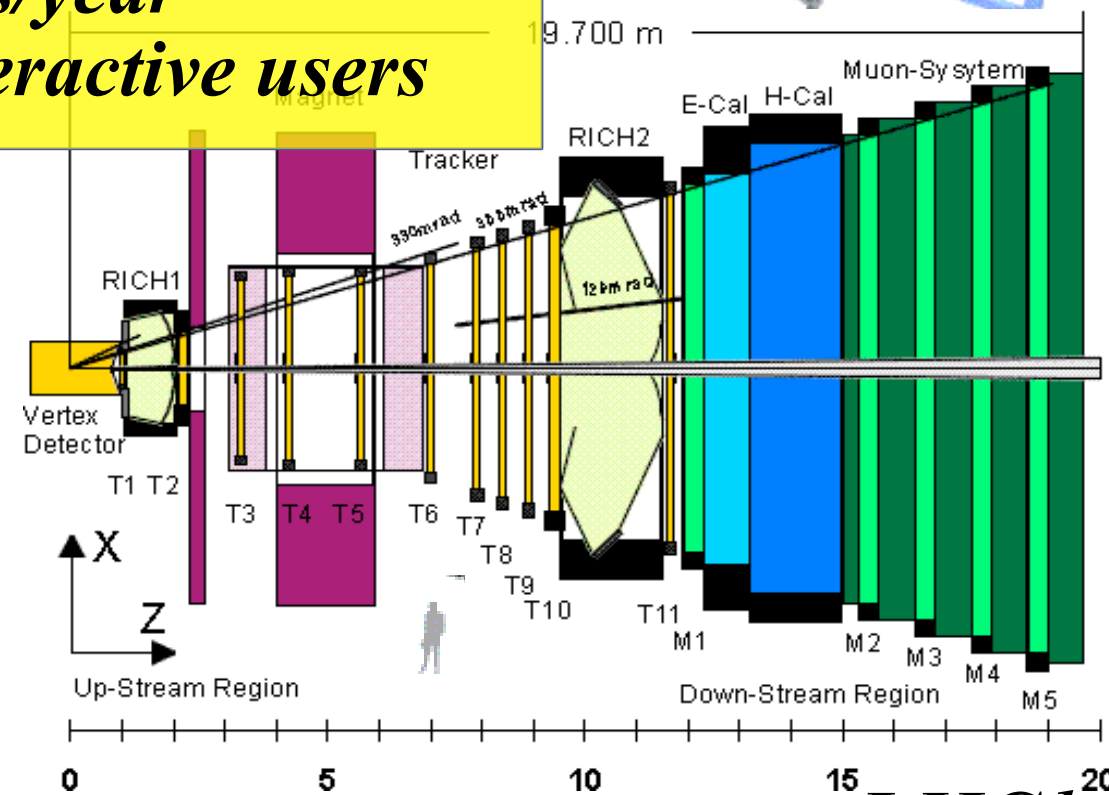
CMS



~6-8 PetaBytes / year
~10⁸ events/year
~10³ batch and interactive users



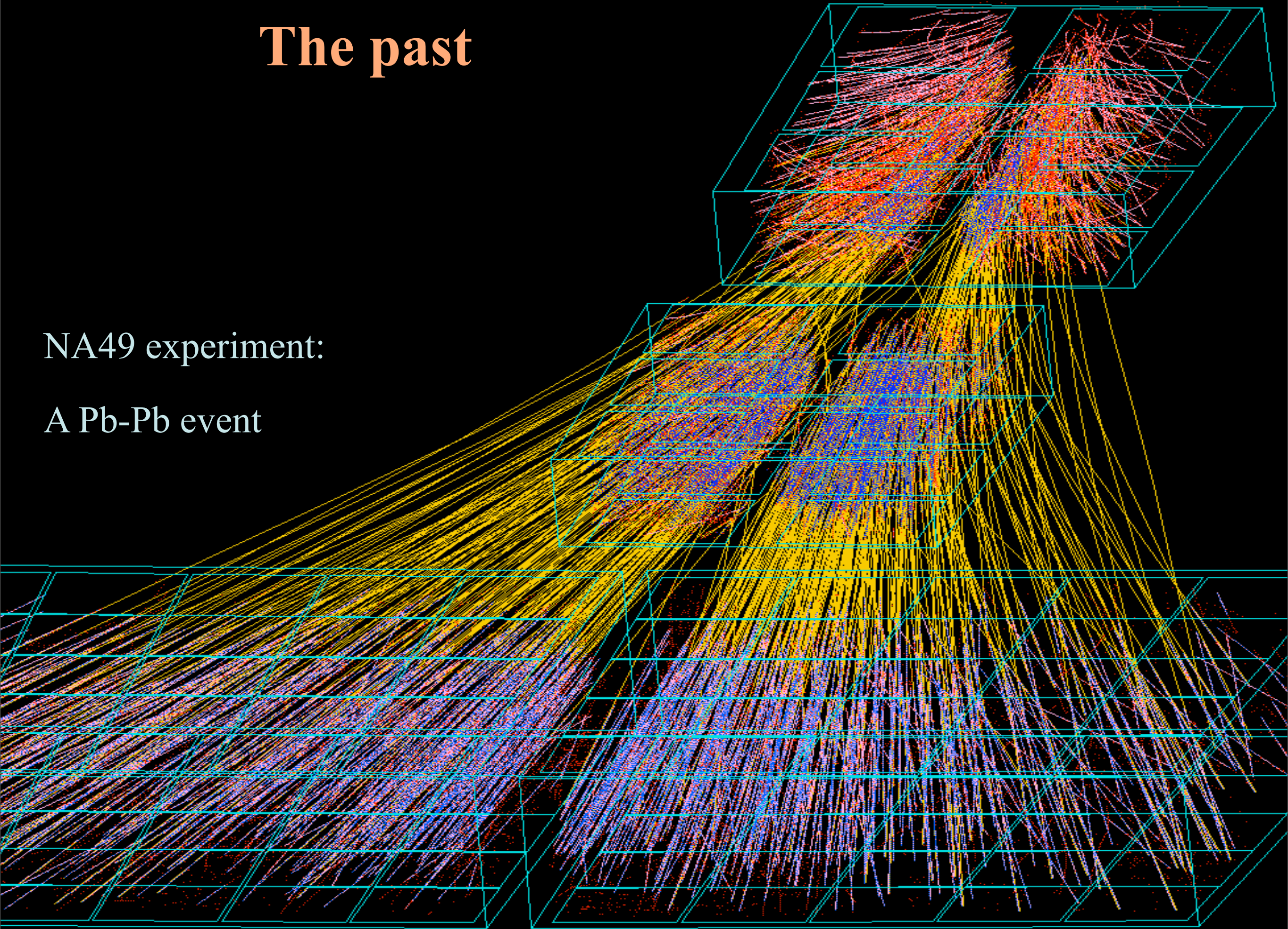
Alice



The past

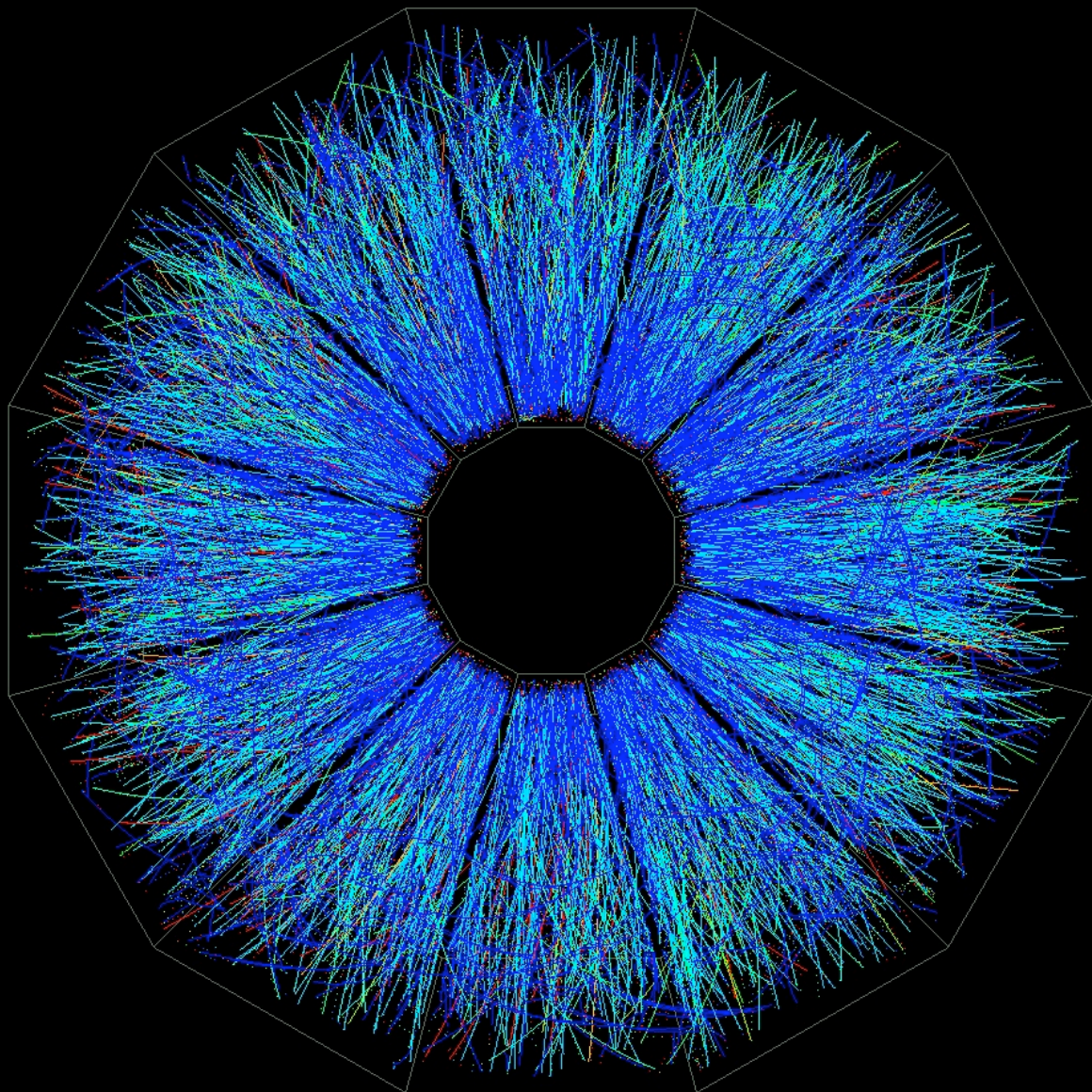
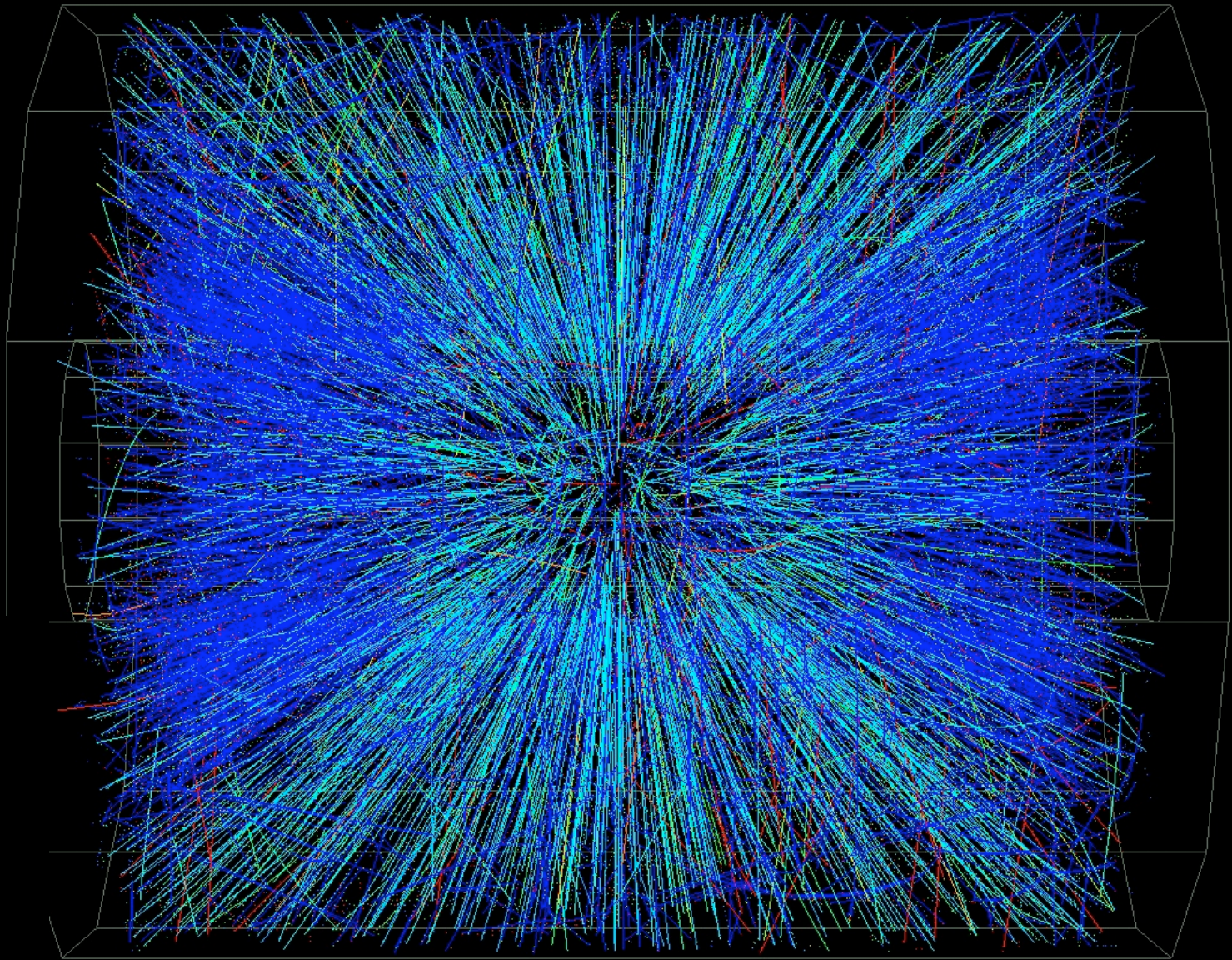
NA49 experiment:

A Pb-Pb event





today



A central Au-Au event
@ ~ 130 GeV/nucleon

... future

Alice event: 0, Run:0

cles = 36276 Nhits = 19431047

Top View

Side View

Front View

All Views

OpenGL

X3D

..
ROOT
ALICE
...

Pick

Zoom

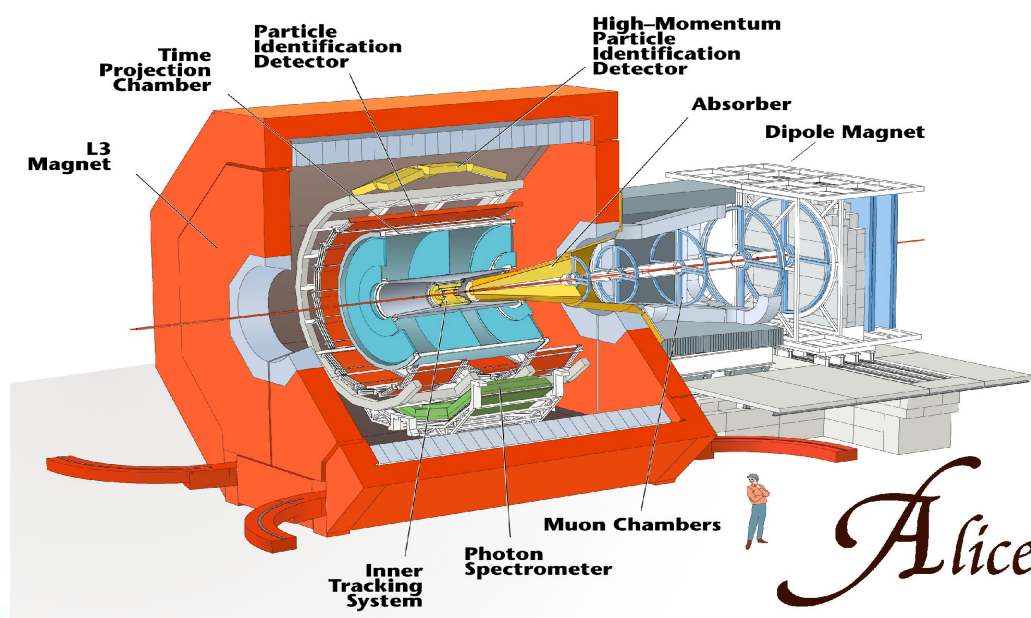
UnZoom

ALICE Pb-Pb central
event

$$N_{\text{ch}}(-0.5 < \eta < 0.5) = 8000$$



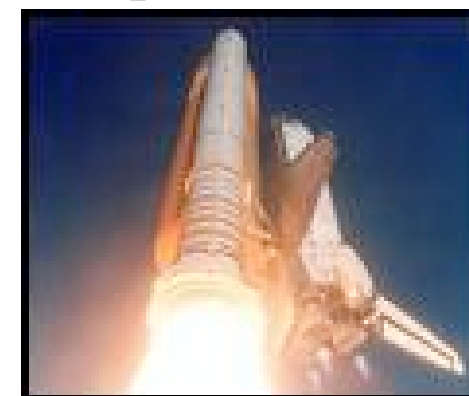
Generación de datos en Alice



Alice detector

Total weight	10,000t
Overall diameter	16.00m
Overall length	25m
Magnetic Field	0.4Tesla

Compare to Shuttle



Total weight	2,000t
Overall length	17.3m

8 kHz (160 GB/sec)

level 0 - special hardware

level 1 - embedded processors

level 2 - PCs

30 Hz
(1.25 GB/sec)

data recording &
offline analysis

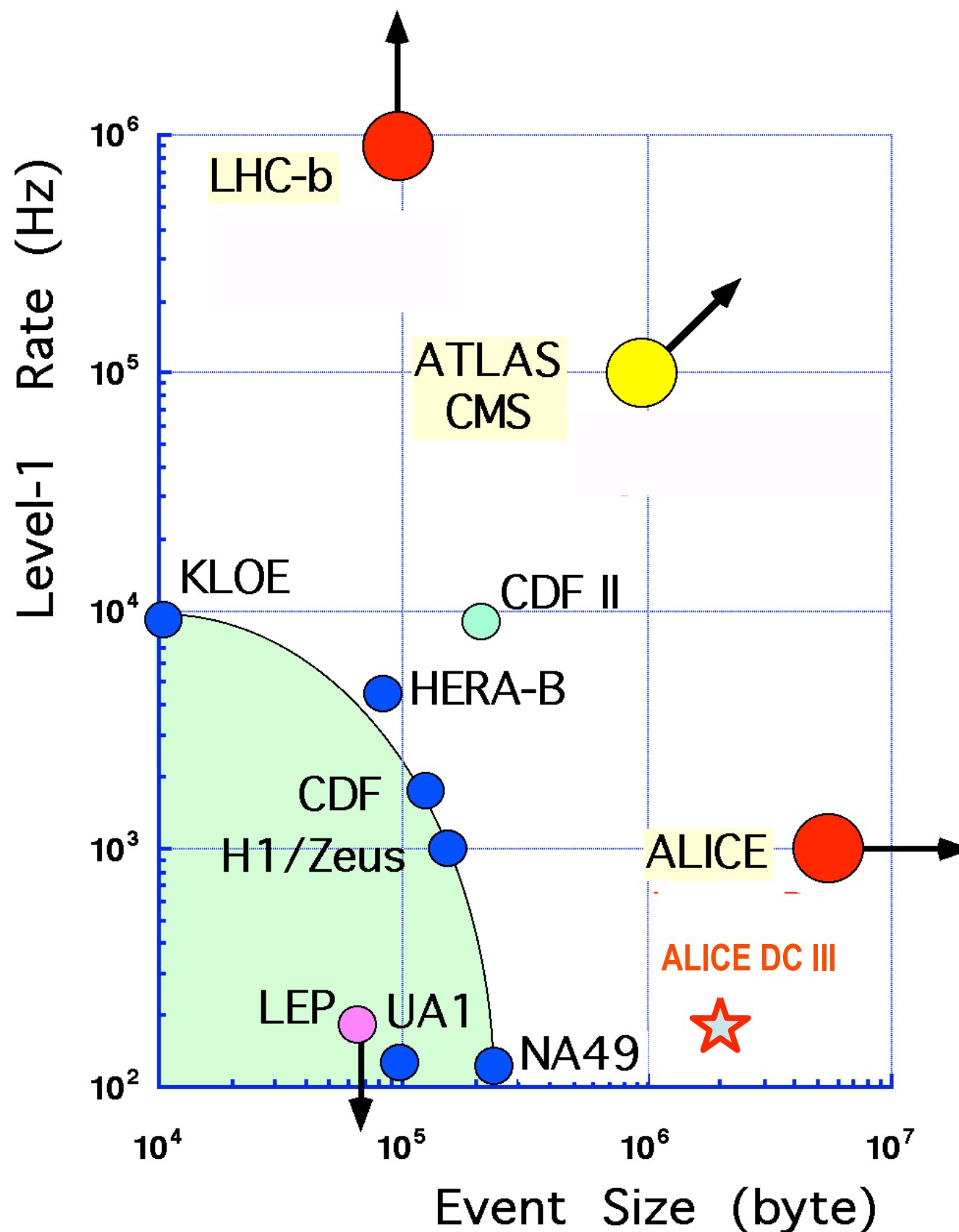
online system

multi-level trigger
filter out background
reduce data volume

More than 3 PB/
year!



Complejidad computacional



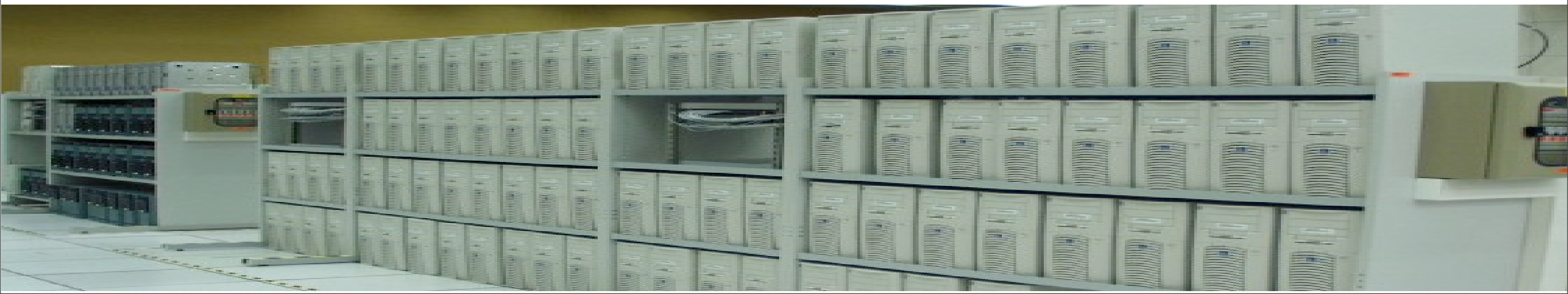


Computo central en CERN



- **Computo de alto rendimiento en base de componentes comerciales estándares y confiables**
- **Más que 1500 PC's dual CPU**
 - 5000 en 2007
- **Más que 3 PB de datos en disco y cinta**
 - > 15 PB en 2007

iNo es suficiente!





Computo para ALICE

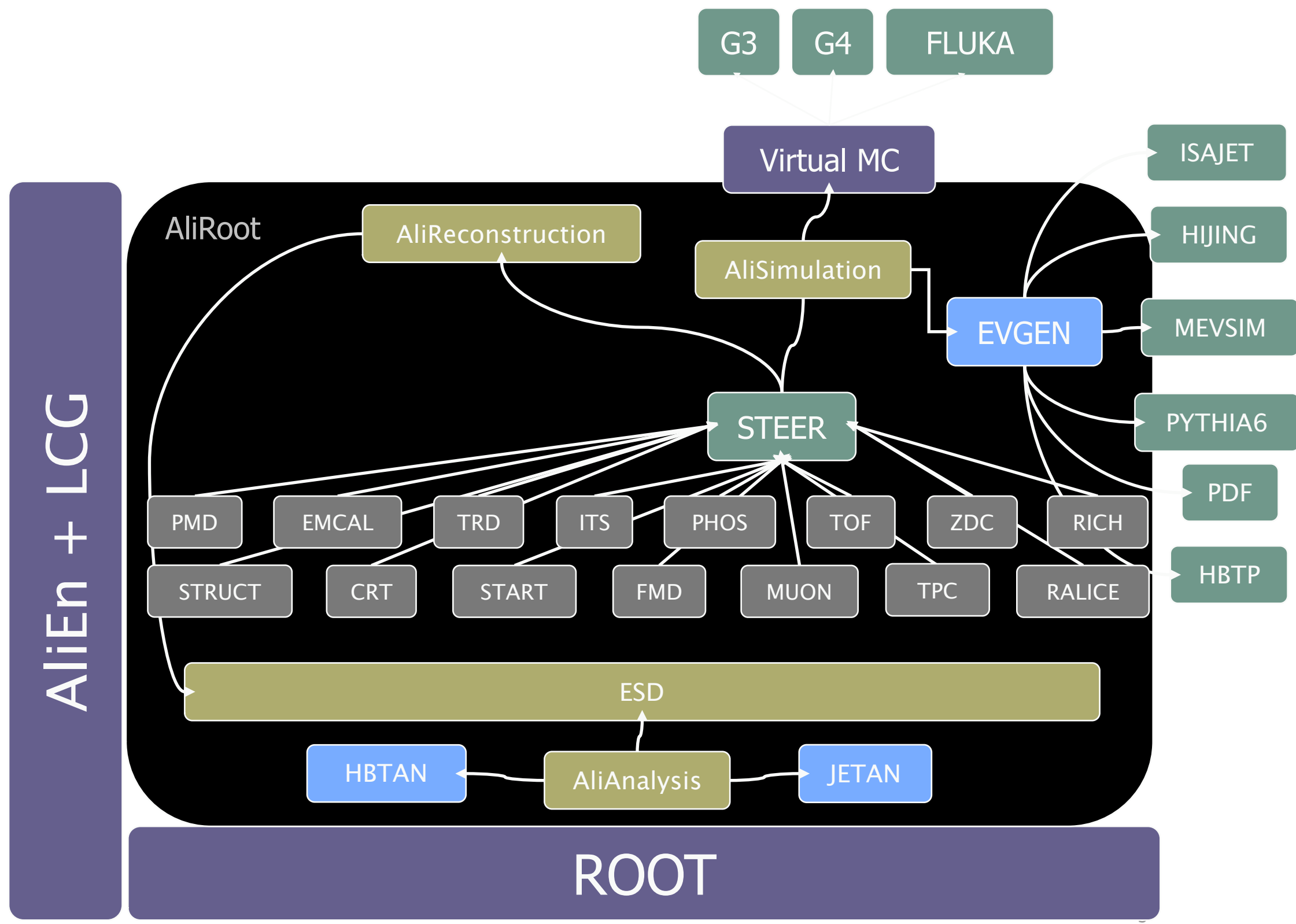
y el resto de los
experimentos
en LHC



- **Computo distribuido**
 - ⌚ Eventos son independientes
 - ⌚ Instalaciones locales con fondos locales
- **Usando tecnologías GRID**
 - ⌚ Servicios estandares
 - ⌚ Mantenido por su propio comunidad
 - ⌚ Integración de recursos en todo el mundo
- **Coordinación**
 - ⌚ Central: LHC computing
 - ⌚ En ALICE



AliROOT





AliEn: entorno para GRID (1)



- Based on Open Source components (95% imported code)
- Offers for ALICE users a single interface into the heterogeneous and fast-evolving GRID reality
 - ⌚ More than 130 registered AliEn users
- Whenever possible, uses common services
 - ⌚ LCG/gLite CE + RB, gLite FTS for scheduled file transfers
- ALICE is taking active part in the definition and testing of these components



AliEn: entorno para GRID (2)

- **The services provided by AliEn are**
 - **ALICE job database and related distributed tools and services**
 - **ALICE file catalogue and related distributed tools and services**
- **ALICE specific job reporting services**
- **Their (high-level) functionality is ALICE-specific and not found elsewhere**



Data Challenge



- **Verificación de la infraestructura**
 - **Operación**
 - **Funcionalidad**
 - **Estabilidad**
- **Generación de simulaciones para otros estudios**





Casi 700,000 jobs en el ultimo año: El sistema trabaja





Sitios de GRID para ALICE





Computo para ALICE en México



- **Participación en el GRID para ALICE**
 - Recursos de computo
 - Acceso a datos
- **Instalación central**
 - Cluster **tochtli**
- **Acceso para todos los colaboradores**
 - Por medio de equipo central
 - **User Interface** local



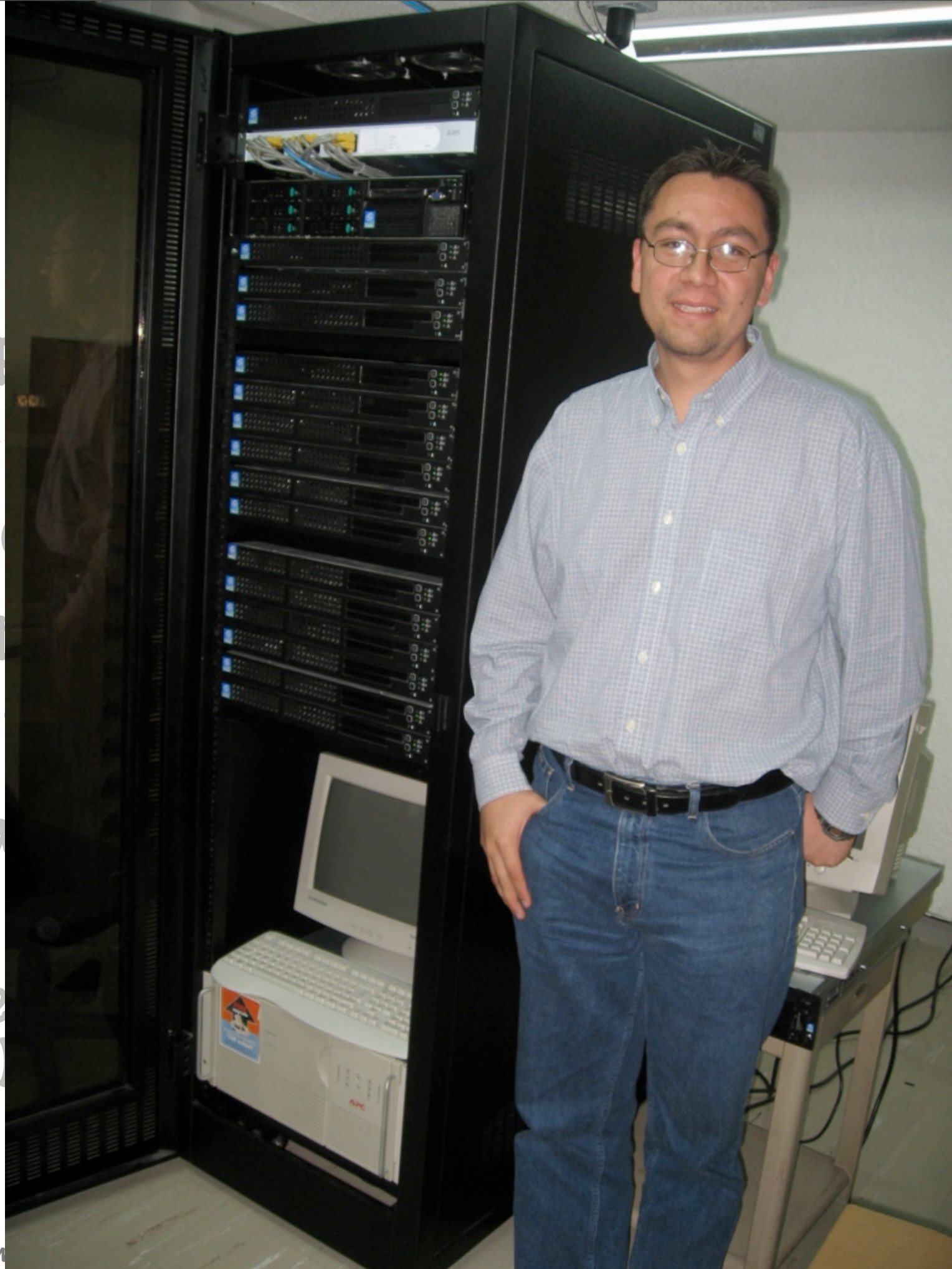
tochtli



- **Master node (CE):**
tochtli.nucleares.unam.mx
- **14 Worker Nodes (WN)**
- **Storage Element (SE)**
tlapiacalli.nucleares.unam.mx
(almacen)
~900GB disk space
- **Alice VO box**
calpolli.nucleares.unam.mx
(vecindad, tribu)



- Master node (CPU)
tochtli.nucleare
- 14 Worker Nodes
- Storage Element
tlapiacalli.nucleare
(almacen)
~900GB disk space
- Alice VO box
calpolli.nucleare
(vecindad, tribu)





Nodo maestro

- Master node (CPU)
tochtli.nucleare
- 14 Worker Nodes
- Storage Element
tlapiacalli.nucleare
(almacen)
~900GB disk space
- Alice VO box
calpolli.nucleare
(vecindad, tribu)





Nodo maestro

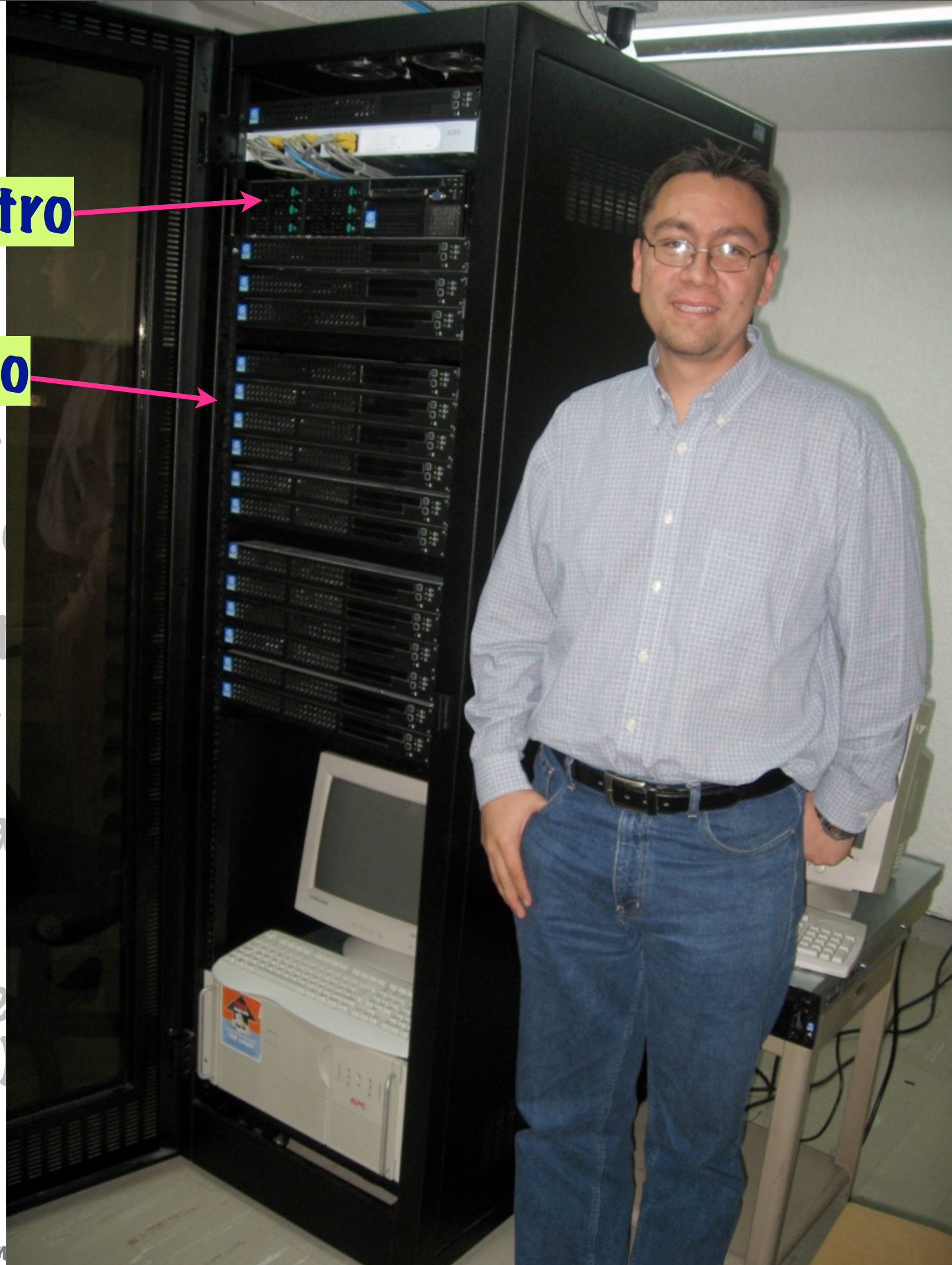
Nodos de trabajo

tochtli.nucleare

14 Worker Nodes

Storage Element
tlapiacalli.nucleare
(almacen)
~900GB disk space

Alice VO box
calpolli.nucleare
(vecindad, tribu)





Storage element

Nodo maestro

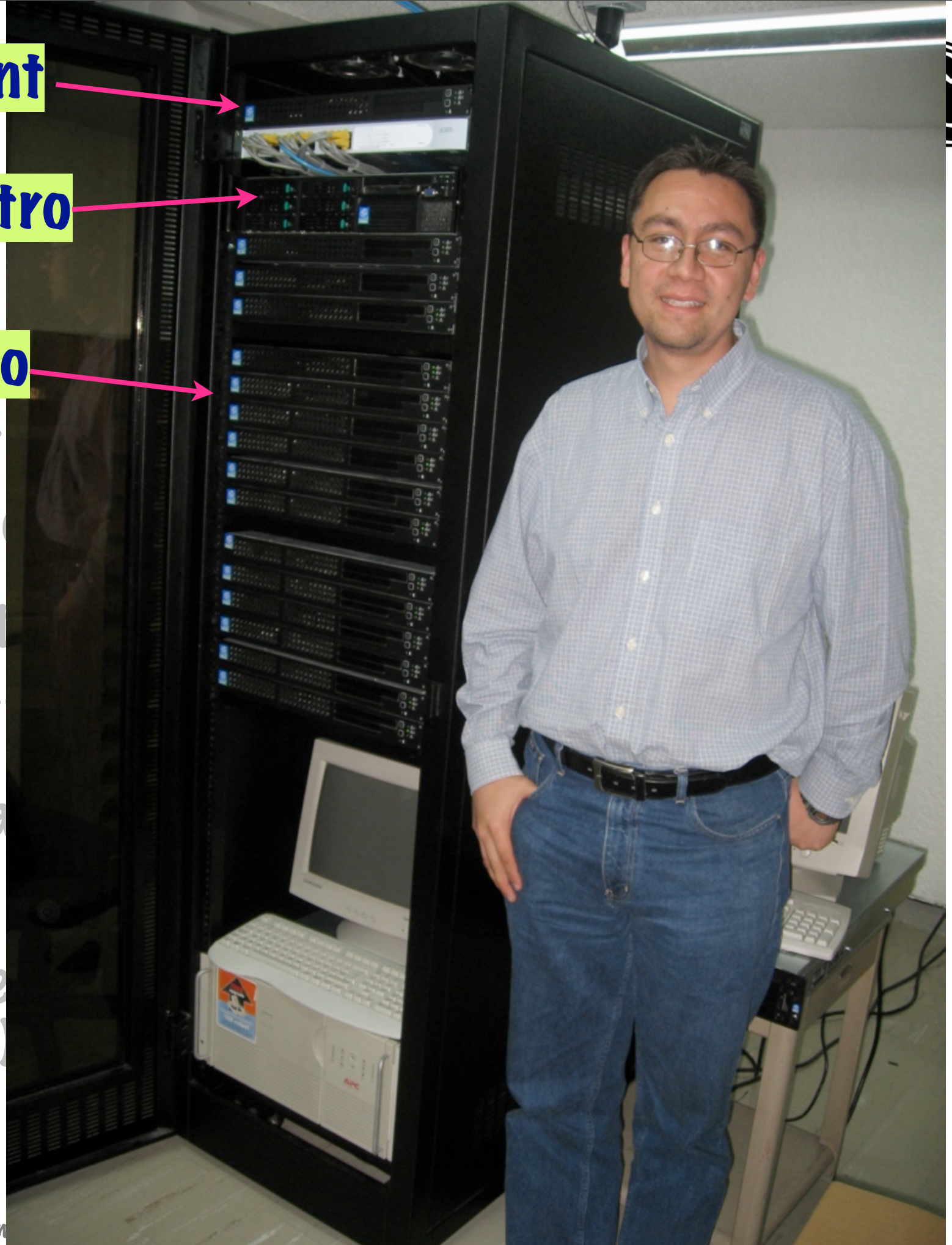
Nodos de trabajo

tochtli.nucleare

14 Worker Nodes

Storage Element
tlapiacalli.nucle
(almacen)
~900GB disk space

Alice VO box
calpolli.nucleare
(vecindad, tribu)





Storage element

Nodo maestro

Nodos de trabajo

tochtli.nucleare

14 Worker Node

Storage Element
tlapiacalli.nucle

(almacen)
~900GB disk spa

Alice VO box
calpolli.nucleare
(vecindad, tribu)

**El entrenador:
Luciano Díaz**





Configuración



- **Sistema de colas**
 - **Acceso local**
 - **Trabajos de GRID**
- **Middleware de GRID**
- **Software de ALICE**
- **Abierto para usuarios locales**



Estructura del middleware



- **Dos capas**
 - **Middleware LCG/g-lite**
 - **AliEn**
- **Servicio locales son distribuidos**
 - **algunos equipos dedicados**

Programa de usuario

GRID

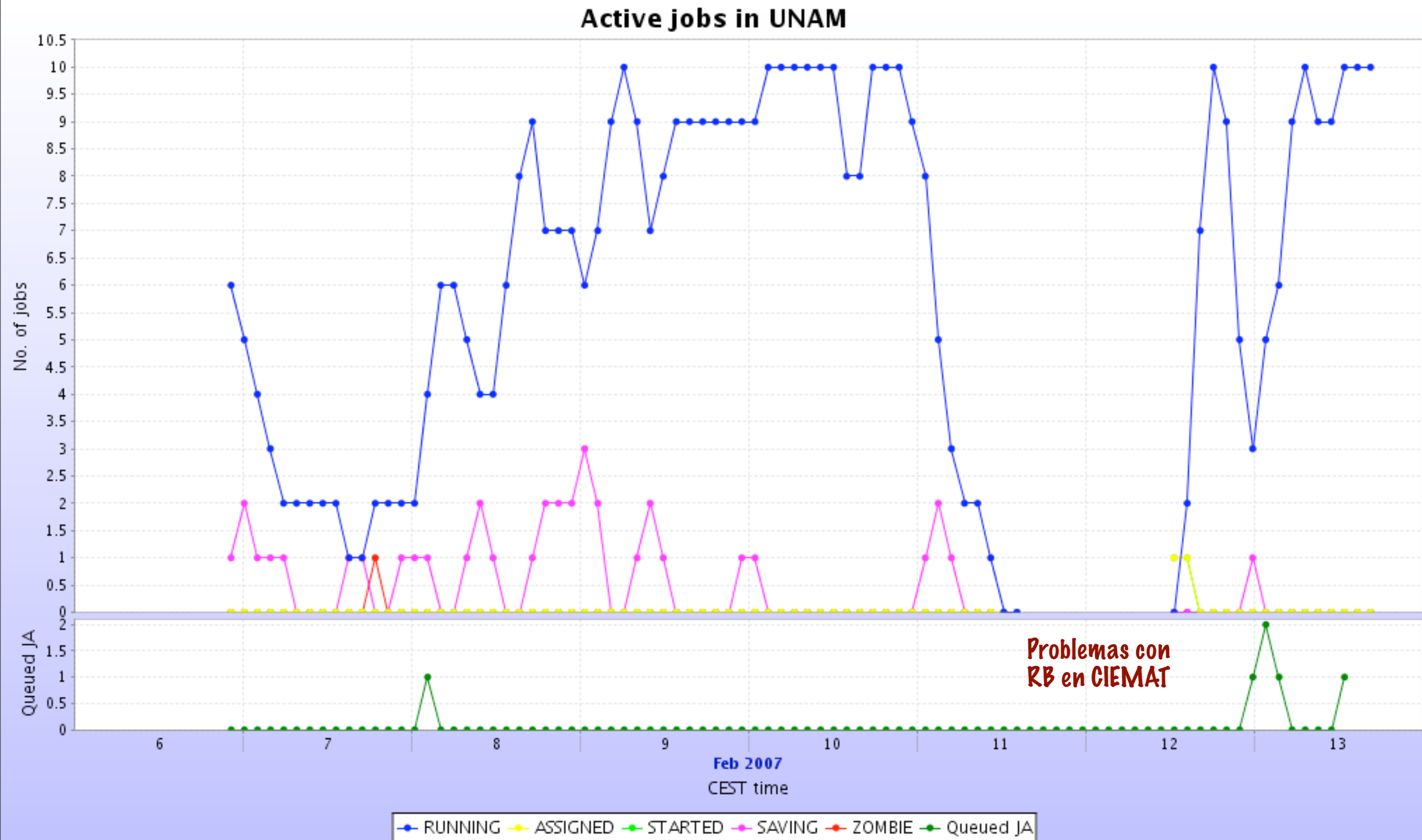
AliEn:
Grid para ALICE

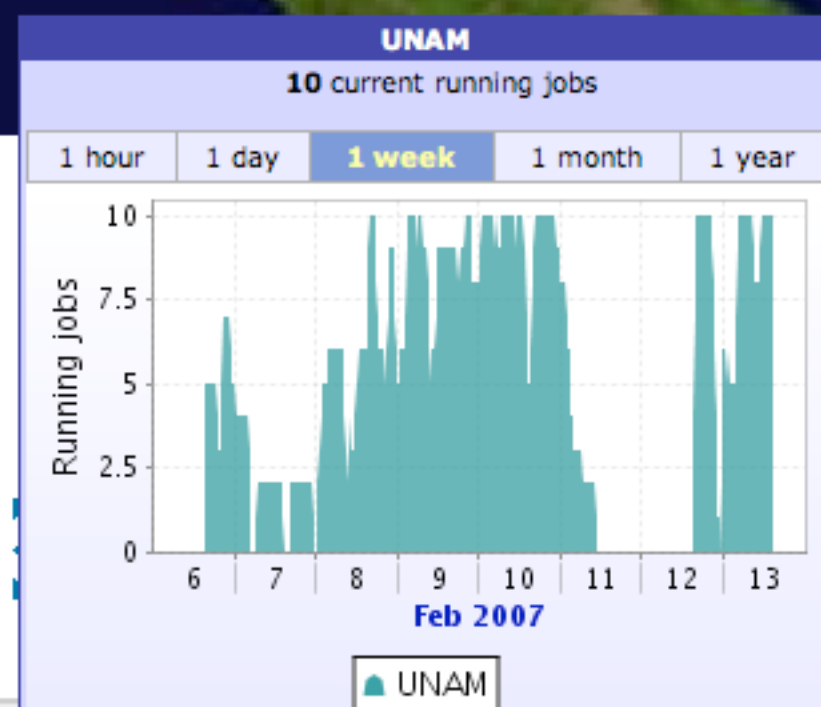
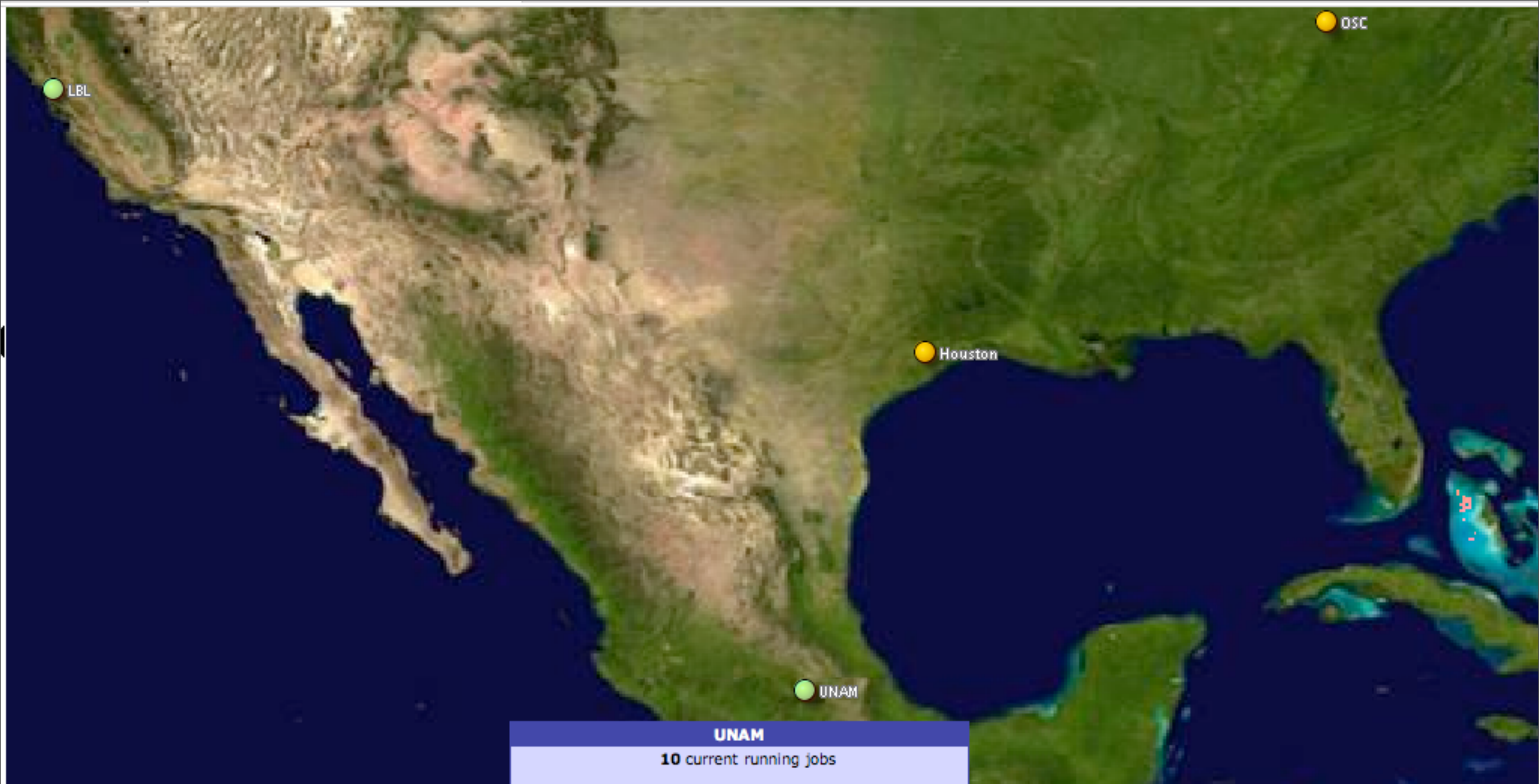
LCG/g-lite
Grid genérico

Sistema Operativo



GRID jobs en tochtli





ive Jobs ■ ML Service Down

Show site relations

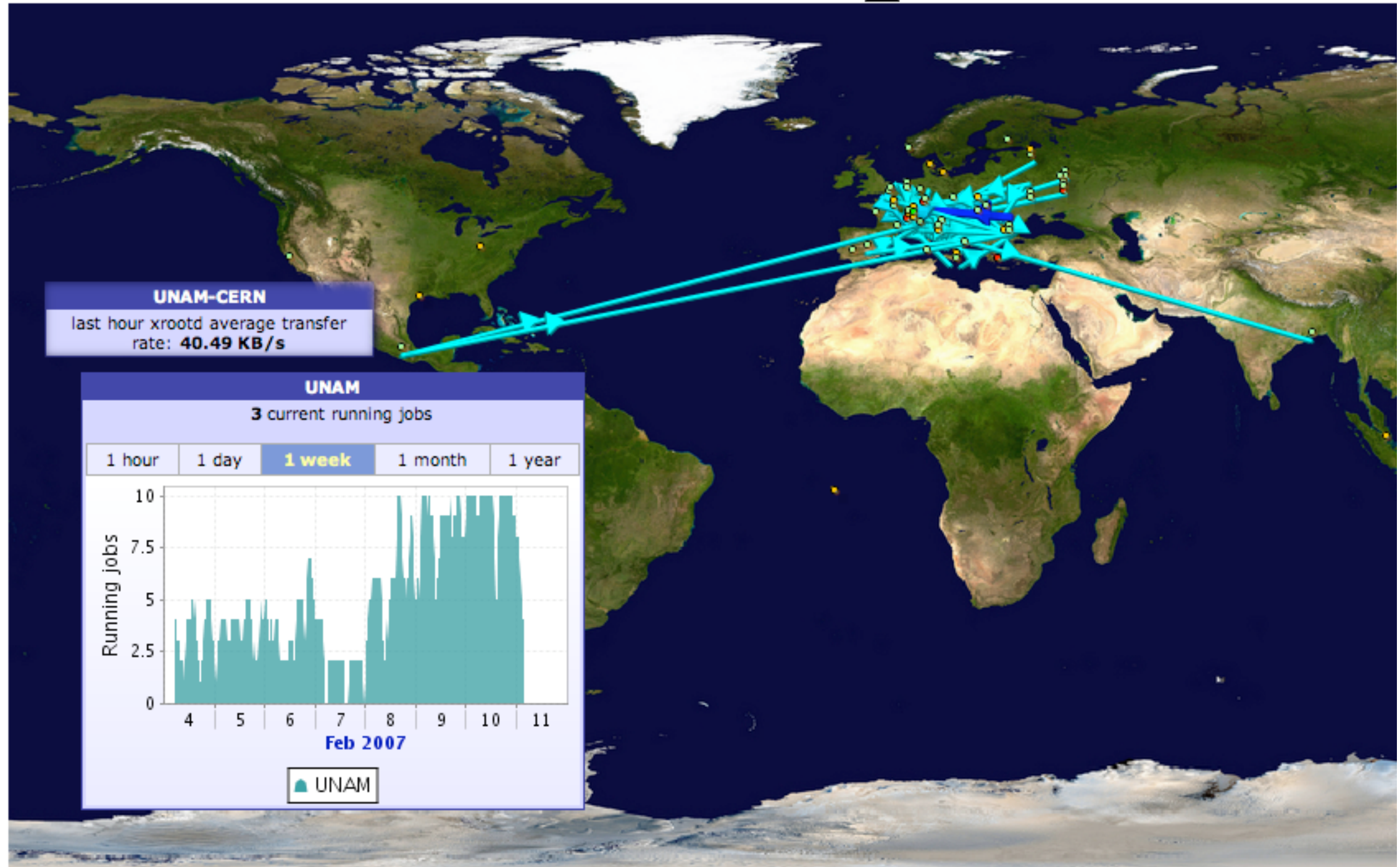
Out (Center on the mouse position)

2

ed Map: North America



Transferencia de resultados

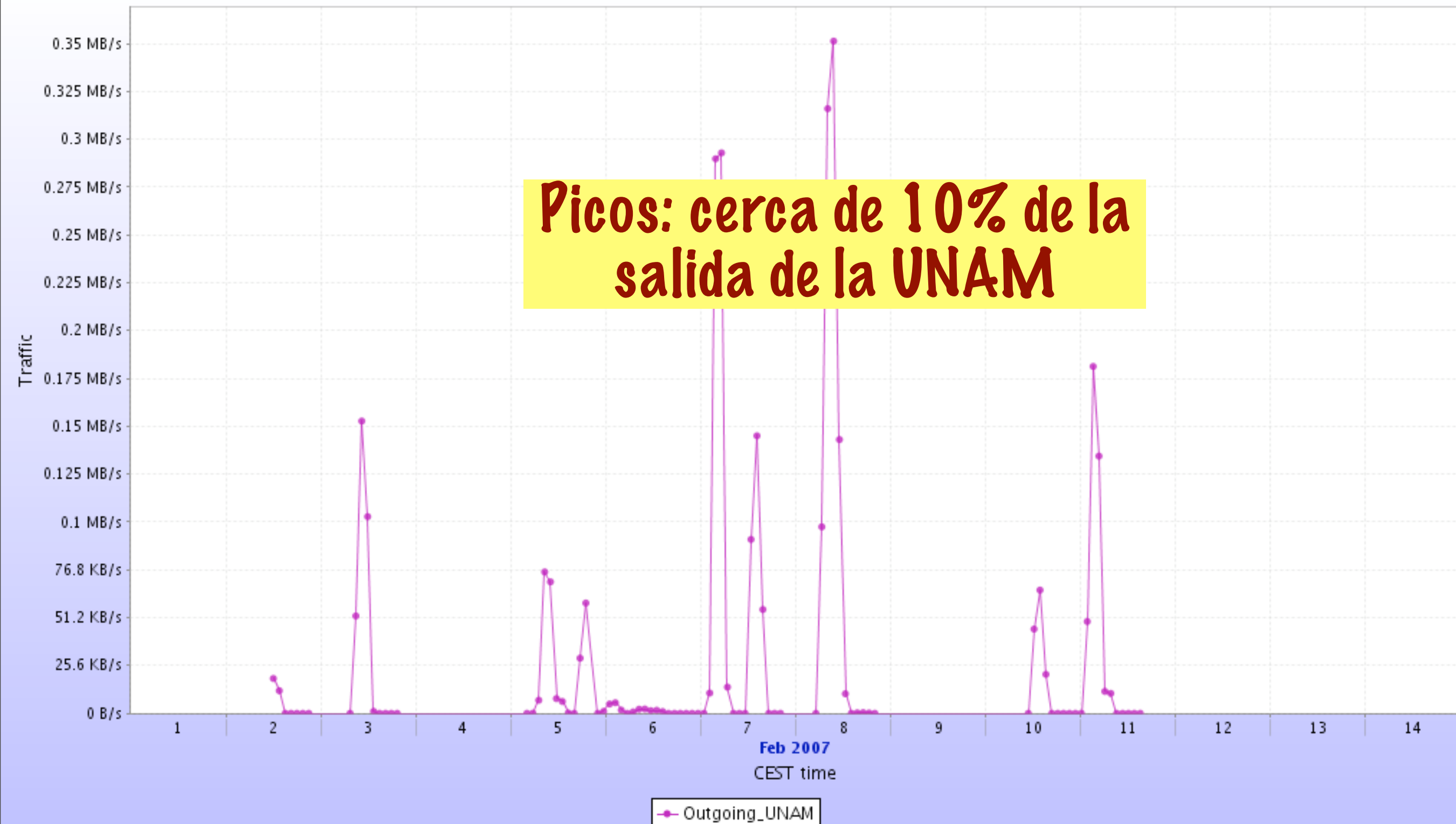




Trafico generado



Site Outgoing Traffic





Usuarios locales



Firefox File Edit View History Bookmarks Tools Window Help

ALICE Grid Monitoring with MonALISA

http://pcalimonitor.cern.ch:8889/job_events.jsp

tochtli - Google Search INFN - GridICE - Grid Moni... ALICE Grid Monitoring with ... alice grid - Google Search GRID.ORG™ - Project Over...

MonALISA Repository ALICE

MonALISA Client

Usage info

Client

ALICE Repository

- Interactive Map
- Running trend
- Job Information
 - Site views
 - Summary plot
 - Job states
 - Jobs per site
 - Jobs per site t
 - Job agents
 - Resource usag
 - User views
 - Summary plot
 - Jobs status
 - Task queue
 - Task queue su
 - Jobs in TQ tab
 - Jobs details
- SE Information
- Services
- Network Traffic
- FTD Transfers
- CAF Monitoring
- SHUTTLE
- Original

Job parameters				Application software			Details		
Run#	PID	Owner	Events	ROOT	ALIROOT	GEANT	Date	Output dir	Type of job
		podesta (15)					last year		
7	1711466	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/007/	test production events
6	1711465	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/006/	test production events
5							19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/005/	test production events
4							19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/004/	test production events
3							19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/003/	test production events
3	1797427	podesta		v5-13-02	v4-04-Rev-08	v1-6-1	25.01.2007 22:06	/alice/cern.ch/user/p/podesta/pp_900GeV/03/	MC simulation of pp min.bias
2	1711421	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/002/	test production events
2	1794911	podesta		v5-13-02	v4-04-Rev-08	v1-6-1	25.01.2007 20:06	/alice/cern.ch/user/p/podesta/pp_900GeV/02/	MC simulation of pp min.bias
1	673395	podesta					06.09.2006 21:08	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/001/	test production events
1	1711410	podesta					19.01.2007 03:10	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/001/	test production events
1	1794908	podesta		v5-13-02	v4-04-Rev-08	v1-6-1	25.01.2007 20:06	/alice/cern.ch/user/p/podesta/pp_900GeV/1/	MC simulation of pp min.bias
	1769912	podesta					23.01.2007 20:37	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/test/	test events
	1771607	podesta					23.01.2007 21:36	/alice/cern.ch/user/p/podesta/demo/flow/output_nc/test/	test events
	1771617	podesta					23.01.2007 21:36	/alice/cern.ch/user/p/podesta/demo/	test events
	1771691	podesta					23.01.2007 23:39	/alice/cern.ch/user/p/podesta/demo/	test events
TOTAL : 15 jobs			0					Export folders	



Equipos de trabajo



- **ALICE México: Instalación de User Interface en Puebla**
- **Colaboración con la DGSCA**
 - Instalación de equipo
 - Entrenamiento
- **En América Latina y Europa**
 - Expertos de ALICE
 - Proyecto EELA



EELA: GRID en general



- **Certificación de usuarios y de equipos**
- **Entrenamiento**
 - Tutorial en México en septiembre de 2006
- **Colaboración con expertos**
- **Formación de equipos de trabajo**
- **Centro de operación regional**
 - monitoreo
 - asistencia a emergencias

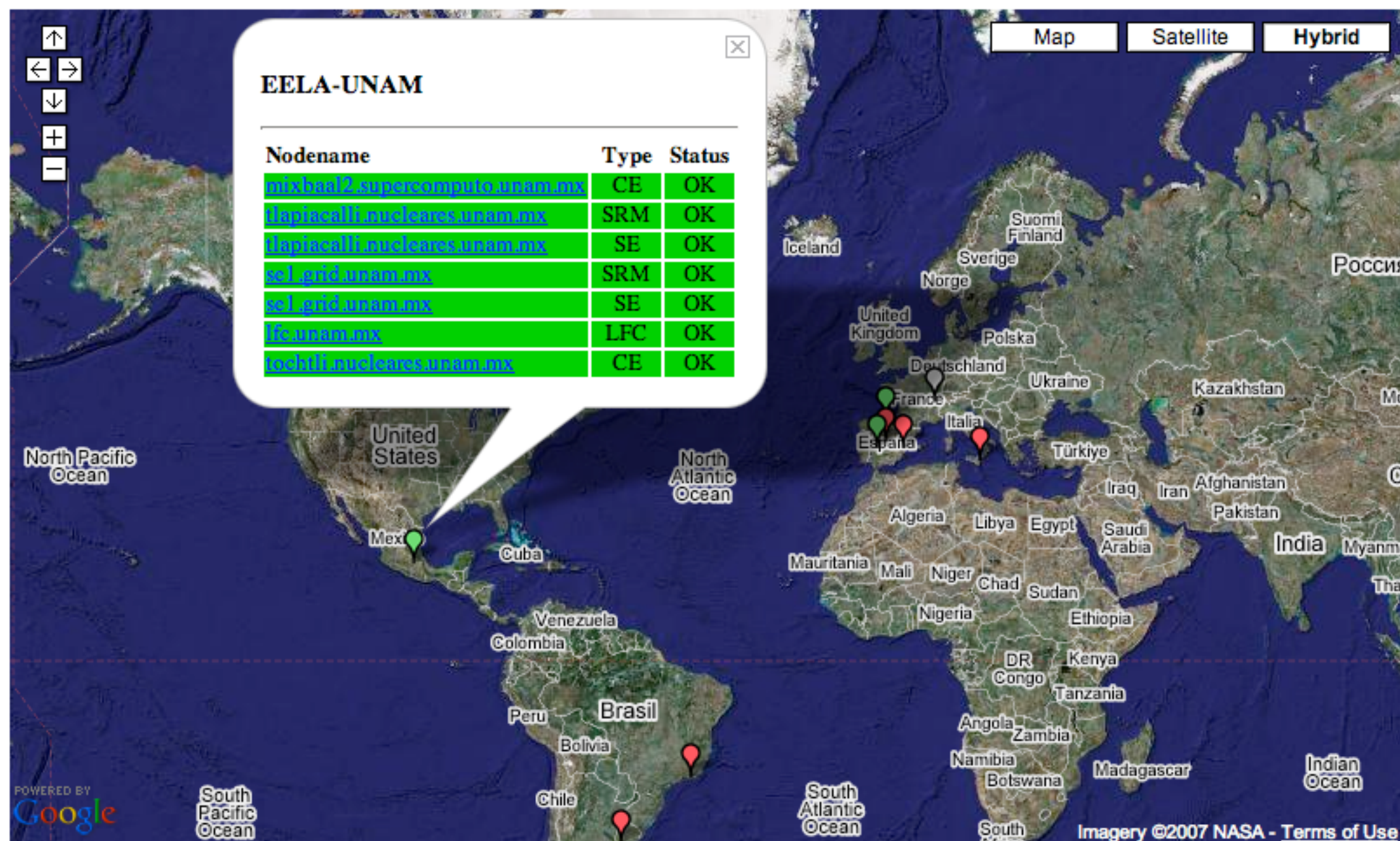


Estatus de los nodos de EELA



Last update: Sun Feb 11 05:01:34 UTC 2007

Updated hourly based on the Service Availability Monitoring (SAM) results.



- Site Ok
- Site with problems.
- Test results not available.



Estatus de los nodos de EELA



Last update: Sun Feb 11 05:01:34 UTC 2007

Updated hourly based on the Service Availability Monitoring (SAM) results.



- Site Ok
- Site with problems.
- Test results not available.



Monitoreo



is monitoring



Geo view

Site view

VO view

Help

About

GridICE >> Site::ALL

General

Gris

Host

Job

Charts

Network

Site ▼		Computing Resources										Storage Resources		
		GK#	Q#	RunJob	WaitJob	SlotLoad	MH#	Power	WN#	CPU#	CPUload	Available	Total	%
EELA-CETA		1	3	4	0	8%	38	550K	33	132	0%	37.3 TB	37.3 TB	0%
EELA-CIEMAT		1	10	16	0	9%	73	2M	67	268	3%	264.6 GB	354.9 GB	25%
EELA-UC		1	2	0	0	29%	8	36K	6	6	33%	92.6 GB	99.5 GB	7%
EELA-UFF		1	8	0	0	0%	3	10K	1	2	0%	26.9 GB	33.8 GB	20%
EELA-UFRJ		1	10	0	0	0%	18	186K	13	46	0%	991.8 GB	1 TB	5%
EELA-UNAM		2	10	7	0	72%	22	99K	17	24	82%	576 GB	1.3 TB	58%
EELA-UNLP		1	6	0	0	0%	–	–	–	–	–	–	–	–
EELA-UTFSM		1	6	0	0	86%	–	–	–	–	–	–	–	–
INFN-CATANIA		1	6	231	150	100%	96	1M	89	256	22%	18.7 TB	21.2 TB	12%
ULA-MERIDA		1	7	0	0	0%	–	–	–	–	–	486.4 GB	512.4 GB	5%
UPV-GRyCAP		1	6	0	0	0%	–	–	–	–	–	60.9 GB	68.7 GB	11%
[cuautitlan2.unam.mx]	–	1	5	0	22220	0%	–	–	–	–	–	–	–	–
TOTAL	#12	13	79	258	22370	25%	258	4M	226	734	20%	58.5 TB	61.9 TB	16%

Generated: Tue, 13 Feb 2007 23:11:54 +0100

GridICE Homepage



Logros



- Equipo **tochtli** instalado y operacional en el GRID de ALICE.
- Transferencia de tecnología grid de EGEE
- Formación de equipos de trabajo en GRID
 - Nacional
 - Internacional
- Usuarios
 - ALICE
 - Bioinformatica



Retos



- **Incrementar nuestros recursos en GRID**
 - **Más CPUs**
 - **Almacenamiento**
 - **Incorporar KanBalam y otros centros**
- **Generar interés en otras comunidades**

¿¿ GRID Nacional ?!?