

Sistema automático de producción de ATLAS: *Generación masiva de sucesos MonteCarlo utilizando el GRID*

Santiago González de la Hoz¹,
Luis March²

¹ CERN, European Organization for Nuclear Research

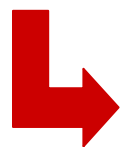
² IFIC, Instituto de Física Corpuscular

Presentado por: L. March



LHC

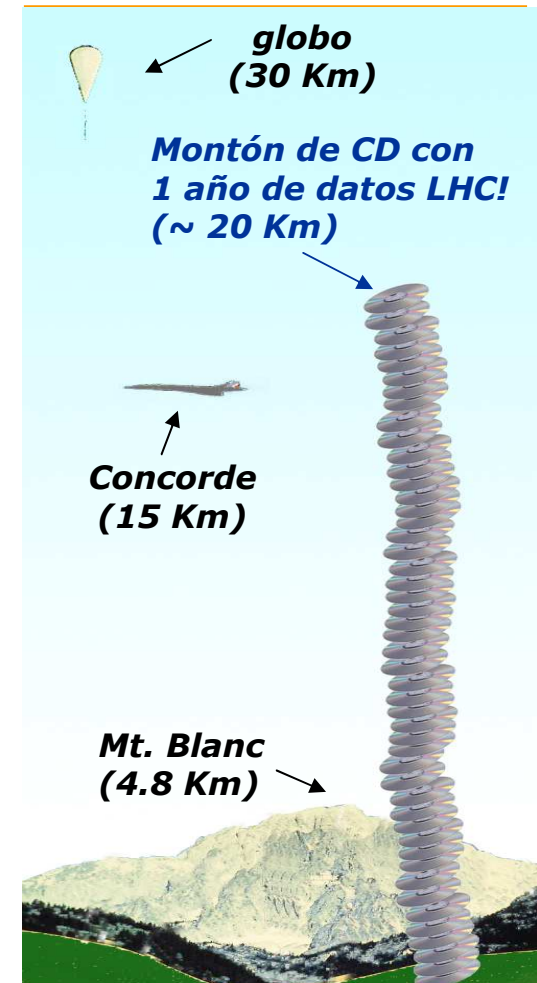
- En el LHC:
 - CPU estimada para procesar datos:
~ 100000 PC's actuales
 - Después del filtrado de datos:
100 sucesos/seg., siendo ~ 1 MB/suceso
→ grabación de 0.1 GB/s
 - ~ 10^{10} sucesos almacenados por año
→ 10 PB/año de datos
- Esto genera 3 problemas:
 - Almacenamiento de datos
 - Procesado
 - Usuarios dispersos



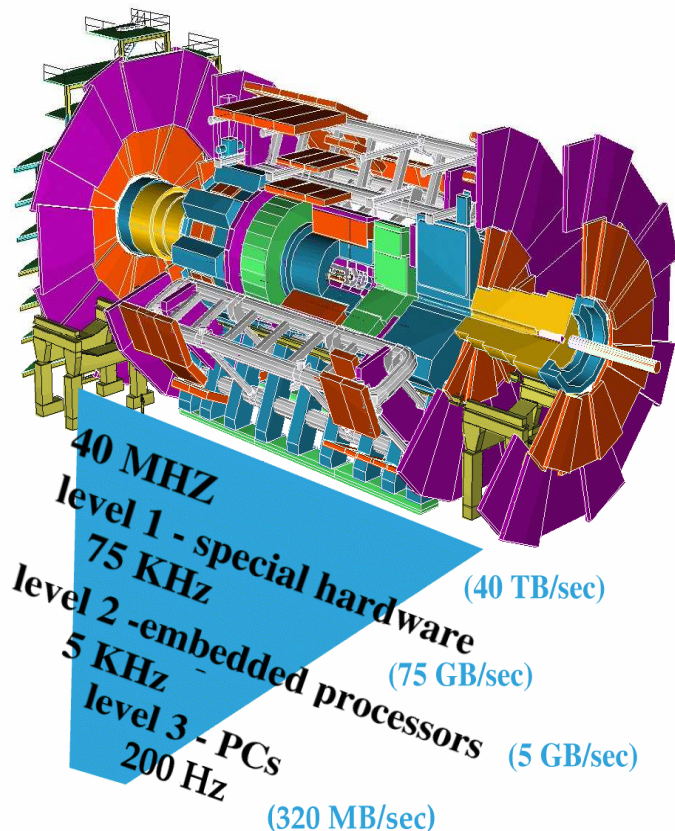
Una posible solución

Tecnologías GRID

Para ello, el CERN, ha desarrollado el proyecto LCG



ATLAS



Toma de datos empieza en 2007

Level 3 trigger: **200 sucesos/s**,

siendo el tamaño de **~ 1.6 MB/suceso**

(mientras que en **LEP** era **1 suceso/s**)

Volumen de datos: **~ 3 PB/año durante 10 años**



Ante estos retos computacionales, el experimento ATLAS elaboró un programa denominado **Data Challenges**

Debido a la complejidad y magnitud del problema computacional planteado, durante la fase preparatoria los experimentos utilizan **datos simulados para desarrollar y optimizar el detector y el sistema de procesamiento de los datos**

ATLAS Data Challenges

La colaboración ATLAS en el año 2002 diseñó el programa
“Data Challenges (DC)”



Validación

- | | |
|-------------------------|-----------------------|
| ▶ Modelo de computación | ▶ Estructura de datos |
| ▶ Software | ▶ Tecnologías GRID |

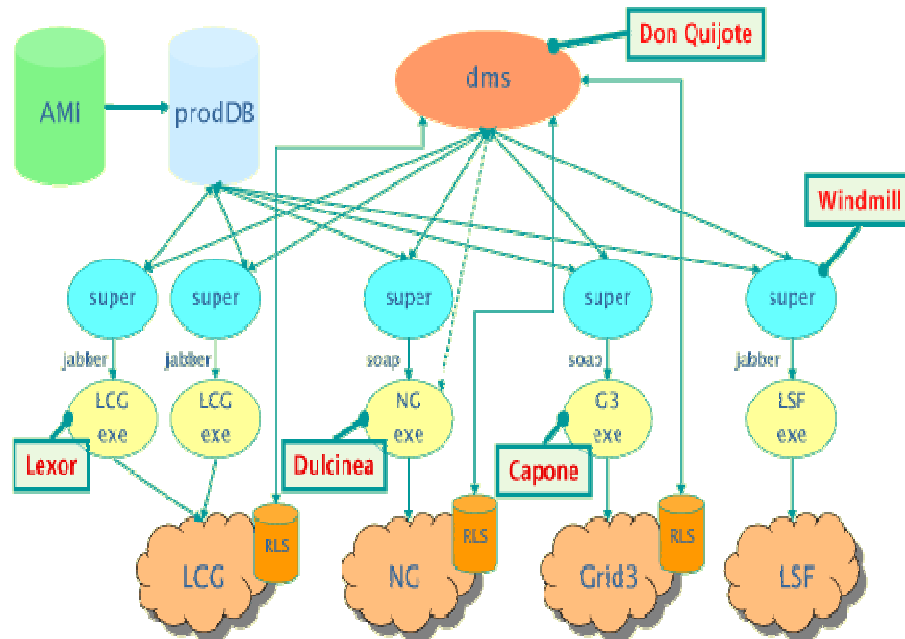
En su primera fase, **DC1**, se consiguió (aunque no se utilizó el GRID):

- Desarrollar y desplegar el software necesario para la producción de sucesos a gran escala, en la cual, participaron centros de todo el mundo

En su segunda fase, **DC2**, se ha conseguido producir sucesos a gran escala utilizando el Middleware GRID desarrollado en 3 proyectos (sabores de GRID):

- LCG, GRID3, NorduGRID

Sistema de producción de ATLAS

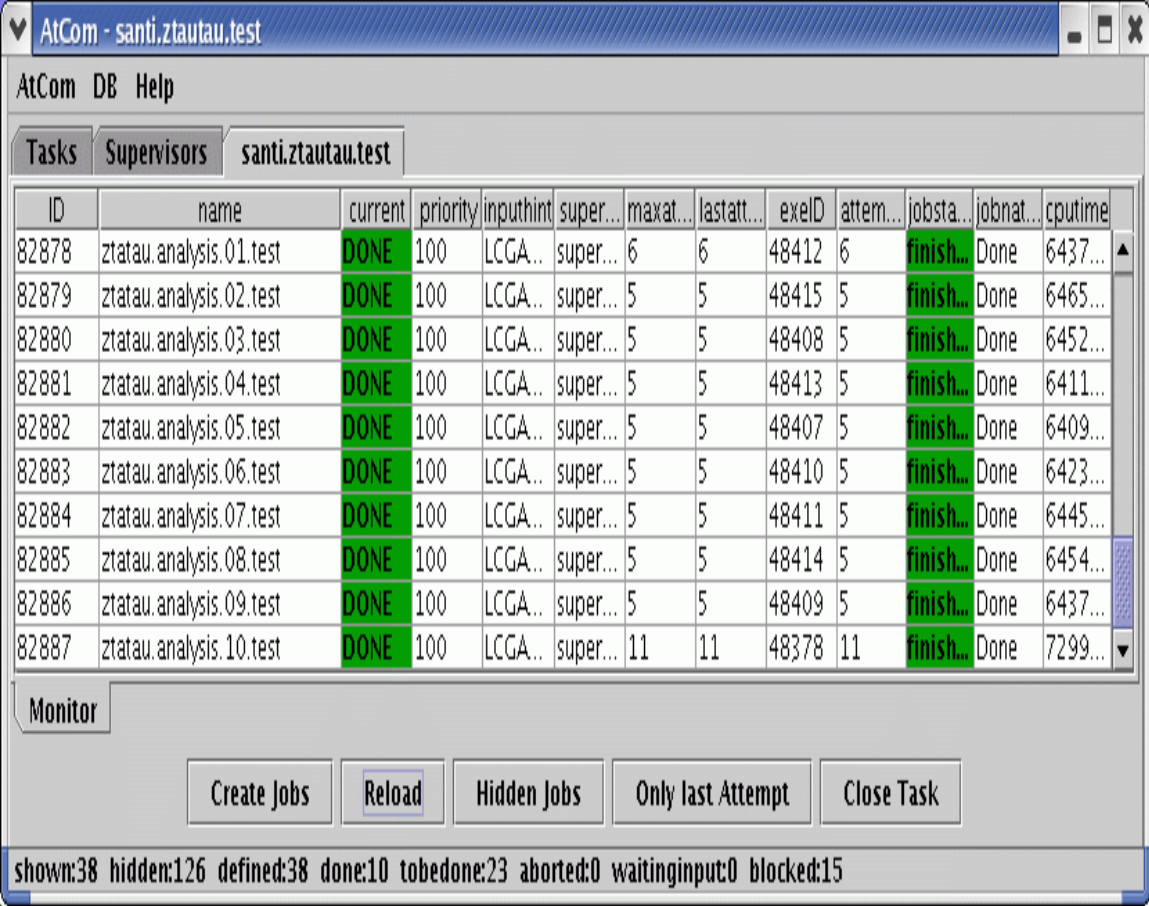


- ✓ Atlas diseñó un **sistema de producción automático** para poder llevar a cabo su producción masiva de datos en los **DC2**.
- ✓ El sistema de producción de ATLAS consiste en **4 componentes**:

- ✓ La **base de datos de producción**, la cual contiene la definición de los trabajos.
- ✓ El **supervisor** “windmill”, el cual consulta la base de datos para obtener la definición de los trabajos y se los pasa a los ejecutores de cada Grid en formato XML.
- ✓ Los **ejecutores**, hay uno por cada sabor Grid. Ellos reciben la definición del trabajo en formato XML y la convierten en el “job description language (JDL)” de cada “sabor” Grid.
- ✓ Don Quijote, **Atlas Data Management system**, el cual se encarga de:
 - ✓ Transferir ficheros desde su localización temporal hasta su destino final en un Storage Element (SE)
 - ✓ Registrarlos en el catálogo correspondiente a cada Grid (replica location service).

Base de datos de producción

- ✓ Definición de trabajos en la base de datos
- ✓ Los trabajos se crean en la base de datos y se les asigna automáticamente un ID
- ✓ También se monitoriza su estado.



The screenshot shows a window titled "AtCom - santi.ztautau.test" with a menu bar "AtCom DB Help". Below the menu bar are tabs for "Tasks", "Supervisors", and "santi.ztautau.test". The "Tasks" tab is active, displaying a table with columns: ID, name, current, priority, inputhint, super..., maxat..., lastatt..., exeID, attem..., jobsta..., jobnat..., and cputime. The table contains 10 rows of task data, all with a "current" status of "DONE". Below the table is a "Monitor" section with buttons: "Create Jobs", "Reload", "Hidden Jobs", "Only last Attempt", and "Close Task". At the bottom of the window, a status bar shows: "shown:38 hidden:126 defined:38 done:10 tobedone:23 aborted:0 waitinginput:0 blocked:15".

ID	name	current	priority	inputhint	super...	maxat...	lastatt...	exeID	attem...	jobsta...	jobnat...	cputime
82878	ztautau.analysis.01.test	DONE	100	LCCA...	super...	6	6	48412	6	finish...	Done	6437...
82879	ztautau.analysis.02.test	DONE	100	LCCA...	super...	5	5	48415	5	finish...	Done	6465...
82880	ztautau.analysis.03.test	DONE	100	LCCA...	super...	5	5	48408	5	finish...	Done	6452...
82881	ztautau.analysis.04.test	DONE	100	LCCA...	super...	5	5	48413	5	finish...	Done	6411...
82882	ztautau.analysis.05.test	DONE	100	LCCA...	super...	5	5	48407	5	finish...	Done	6409...
82883	ztautau.analysis.06.test	DONE	100	LCCA...	super...	5	5	48410	5	finish...	Done	6423...
82884	ztautau.analysis.07.test	DONE	100	LCCA...	super...	5	5	48411	5	finish...	Done	6445...
82885	ztautau.analysis.08.test	DONE	100	LCCA...	super...	5	5	48414	5	finish...	Done	6454...
82886	ztautau.analysis.09.test	DONE	100	LCCA...	super...	5	5	48409	5	finish...	Done	6437...
82887	ztautau.analysis.10.test	DONE	100	LCCA...	super...	11	11	48378	11	finish...	Done	7299...

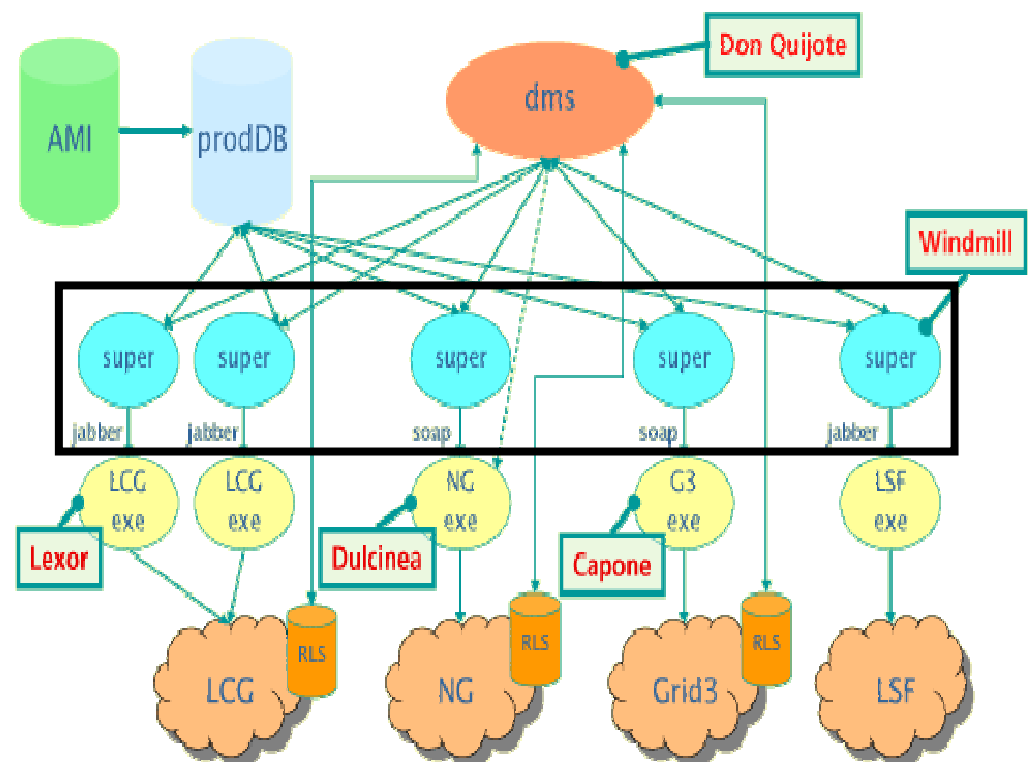
Monitor

Create Jobs Reload Hidden Jobs Only last Attempt Close Task

shown:38 hidden:126 defined:38 done:10 tobedone:23 aborted:0 waitinginput:0 blocked:15

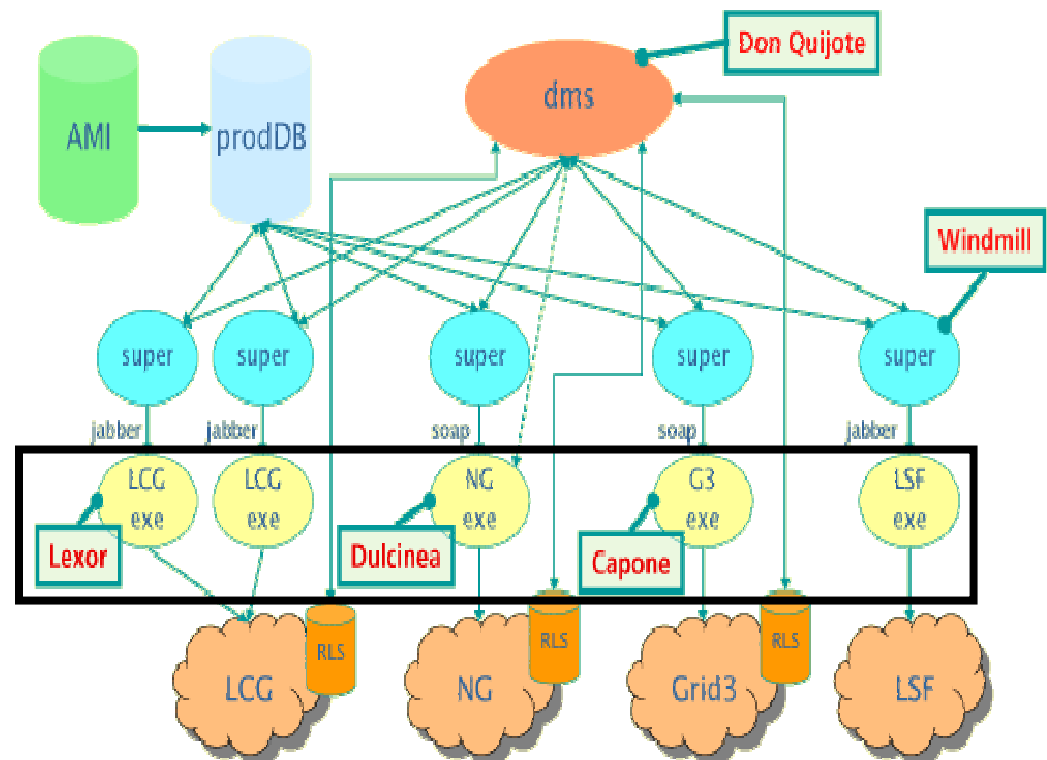
Supervisor

- ✓ **Supervisor: Windmill**
- ✓ Es una implementación del software “supervisor” para la gestión de la producción de ATLAS.
- ✓ Consulta la base de datos para obtener la definición de los trabajos.
- ✓ Interacciona con varios “ejecutores” para la distribución, ejecución y monitorización de los trabajos en los diferentes **Grids** (a través de XML, utilizando un jabber server,)



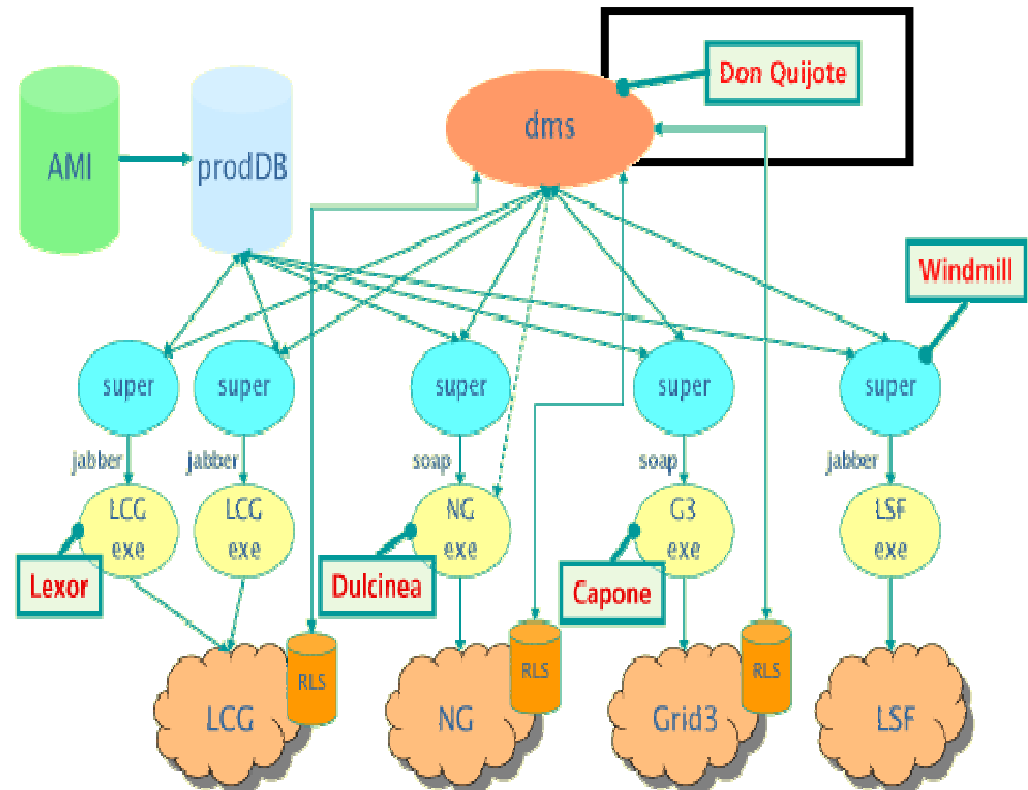
Ejecutores

- ✓ **Ejecutores: Lector, Dulcinea y Capone**
- ✓ **Por diseño son una interfaz entre el supervisor (que es único) y los tres “sabores” Grid.**
- ✓ **La información que reciben del supervisor en XML, la convierten en el JDL (Job Description Language) de que “sabor” Grid**
- ✓ **Envían los trabajos a los correspondientes Computing Elements asociados con dicho sabor de Grid.**

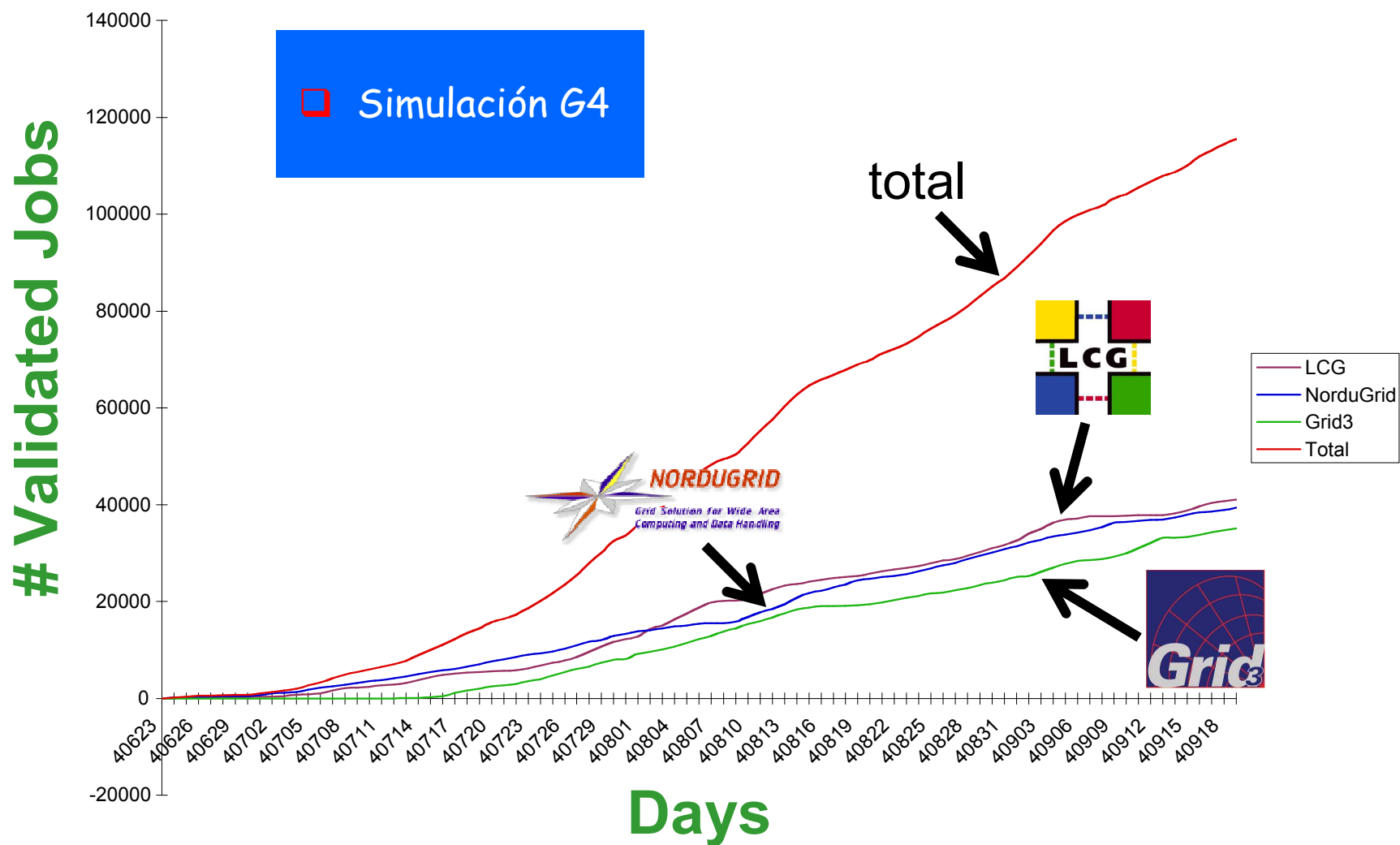


Sistema de gestión de datos

- ✓ **Data Management System: Don Quijote (DQ)**
- ✓ El objetivo de DQ es integrar todos los servicios de gestión de datos (data management) usados por los tres Grids y por lo tanto por el Experimento ATLAS.
- ✓ De esta forma los **datos** se pueden **replicar y registrar** de forma transparente entre los **diferentes catálogos de los tres “sabores” grids**.
- ✓ Realmente hay tres servidores DQ, uno para cada Grid, pero para el usuario final, esto es totalmente transparente. Como si solo hubiera uno

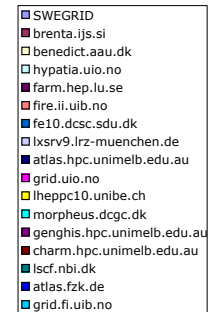
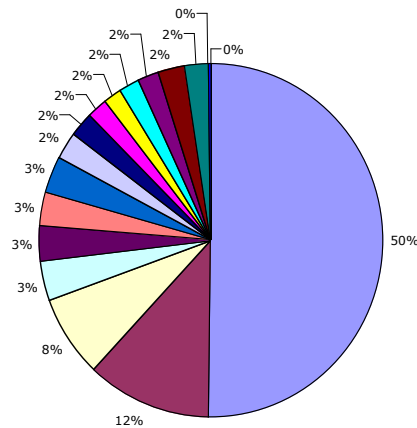
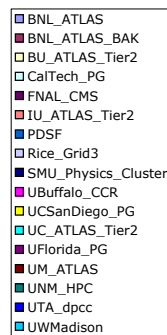
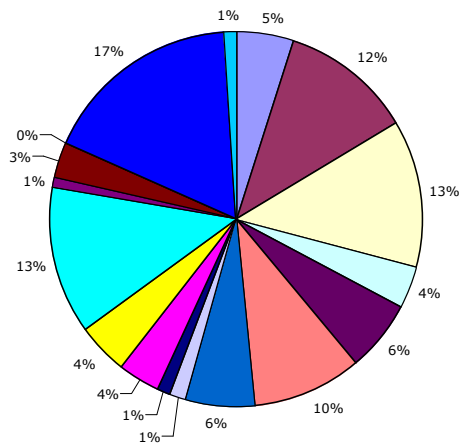
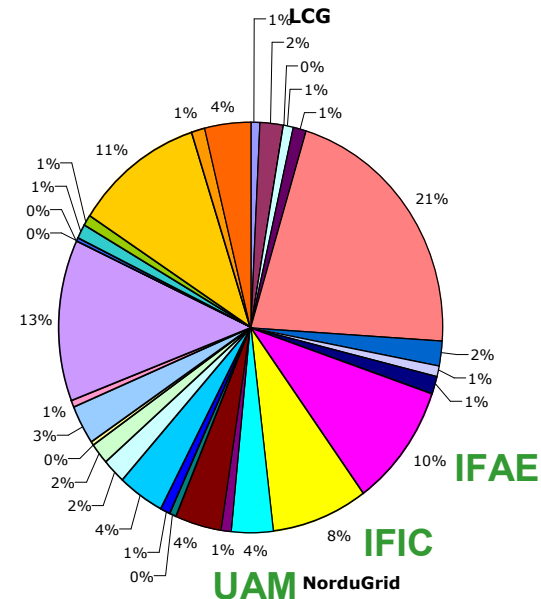
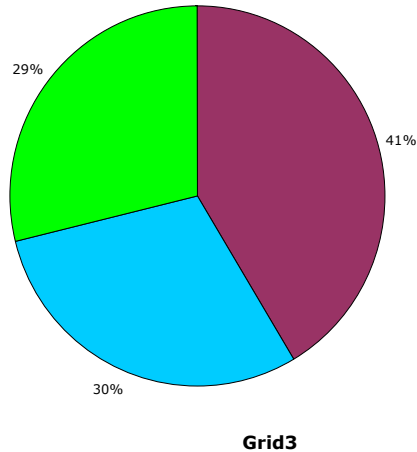


Producción de ATLAS en DC2



Producción de ATLAS en DC2

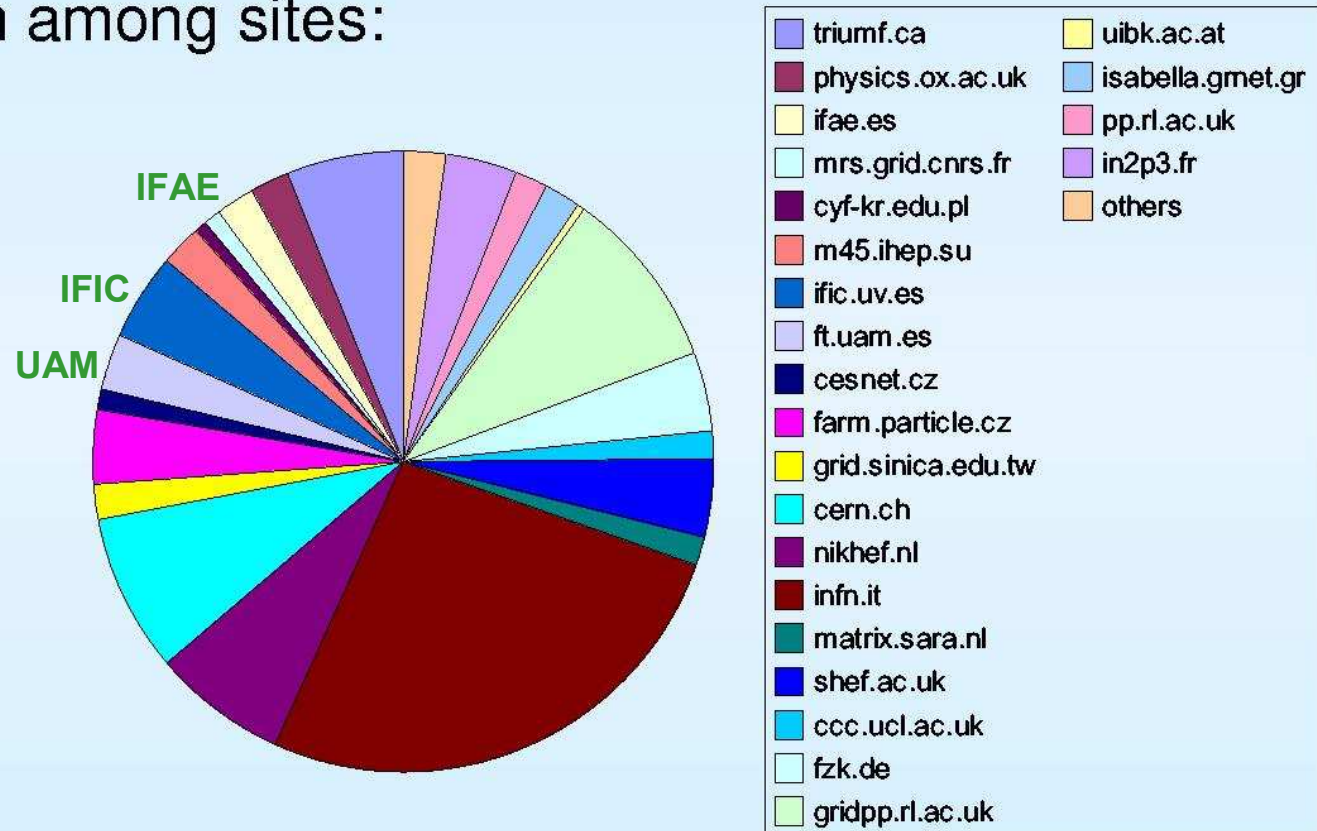
ATLAS DC2 - CPU usage



Producción de ATLAS para Roma

Producción para el **Workshop de Física de ATLAS** en Roma (6-11 Junio 2005)

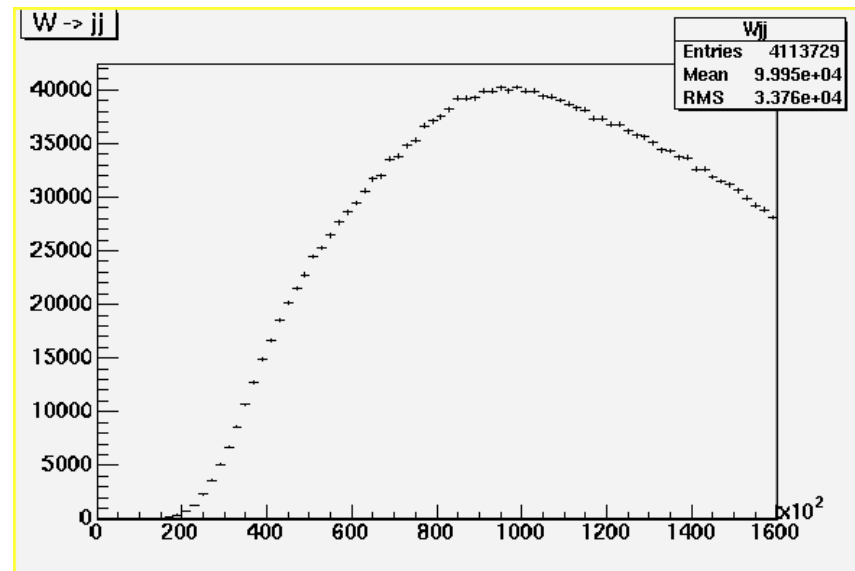
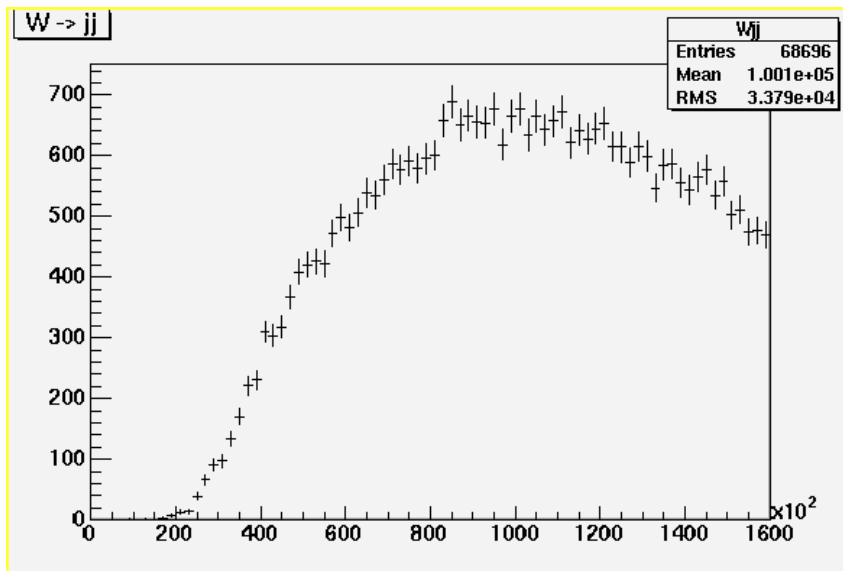
- Distribution among sites:



✓ Número total de trabajos validados: 305410

Análisis distribuido

Se está examinando este sistema de producción para emplearlo del mismo modo en **análisis distribuido**



- ✓ **Reconstrucción de la masa invariante del W desintegrándose a jet-jet (W->jj):**
 - ✓ A la izquierda uno de los 70 histogramas producidos a través del sistema de producción de ATLAS mediante Grid
 - ✓ A la derecha una vez hecho el encadenamiento de los 70 histogramas
 - ✓ Hemos analizado **350000 sucesos!!!**

Conclusiones

- ✓ **ATLAS** ha sido el primer experimento LHC que ha utilizado las tecnologías GRID para llevar a cabo una producción masiva de datos simulados a la totalidad del ejercicio.
 - ✓ Se han utilizado 3 “sabores” de GRID.
 - ✓ La producción ha sido llevada a cabo de una forma coherente y automática.
- ✓ Los **Data Challenges 2** nos han enseñado que todos los elementos involucrados en el sistema de producción de ATLAS necesitan mejoras
 - ✓ En tres meses (**Julio a Septiembre 2004**), se enviaron mas de **235k trabajos**, se consumió mas de **1.5 Mspecint2000/mes de CPU** y se produjeron más de **30 TB de datos**.
- ✓ En la producción de Roma se enviaron **380k trabajos**, almacenándose **1.4M ficheros** correspondiendo a una cantidad de **45 TB de datos**, siendo una clara mejoría respecto a los DC2. Aún así, se necesita mejorar su estabilidad y rendimiento.
- ✓ El modelo de análisis propuesto por ATLAS se probó con éxito en el sistema de producción.
 - ✓ En menos de 4 horas **350000 sucesos** fueron “analizados”. Dichos sucesos estaban almacenados en diferentes lugares por todo el mundo y los trabajos de análisis se ejecutaron en un gran abanico de centros.
 - ✓ Si los trabajos corren donde los datos están almacenados, los sucesos se analizan **en menos de 30 minutos**.