



Tecniche di analisi dati a LHC




Tommaso Boccali (SNS Pisa)



Il Problema



- Dimensione del problema mai affrontata prima, sia dal punto di vista della necessita' di storage che di calcolo

	ALEPH 1995		CDF 2004		CMS 2007	
Dimensione dei dati raccolti	1 TB = 1000 GB		1 PB = 1000 TB x1000		~ 10 PB x10	
Capacita' di calcolo (SI2k)	<< 100k		1.4 M x50		> 25 M x20	

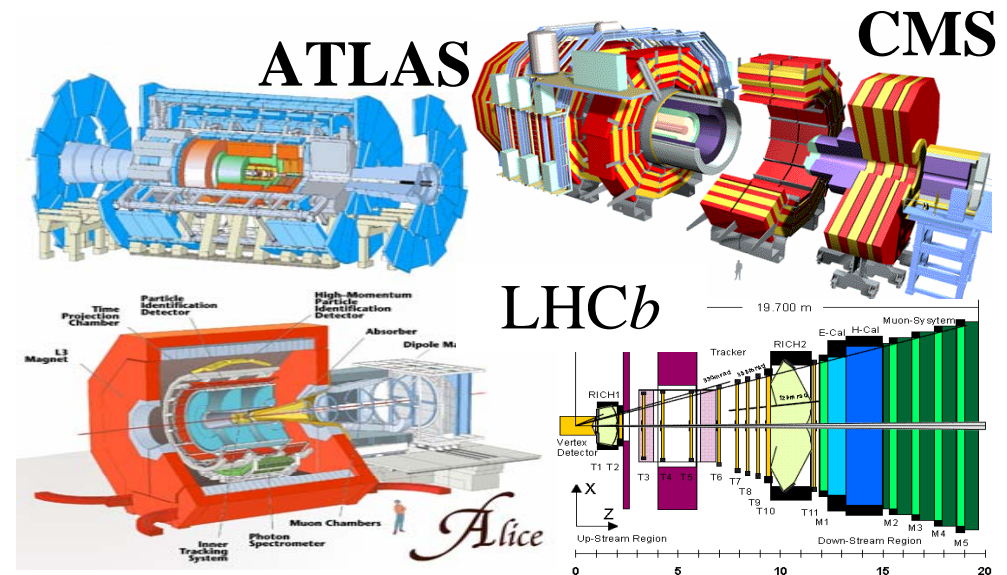


Ultime stime ... (LHCC)



■ Per i 4 esperimenti, in totale:

- **CPU ~ 100MSI2k**
 - ~ 100000 CPU attuali
 - ~ 25000 CPU 2007
- **Disco ~ 40 PB**
 - 100.000 HD da 400 GB
- **Nastro ~ 40 PB**



■ Quali soluzioni sono possibili per il calcolo a LHC?

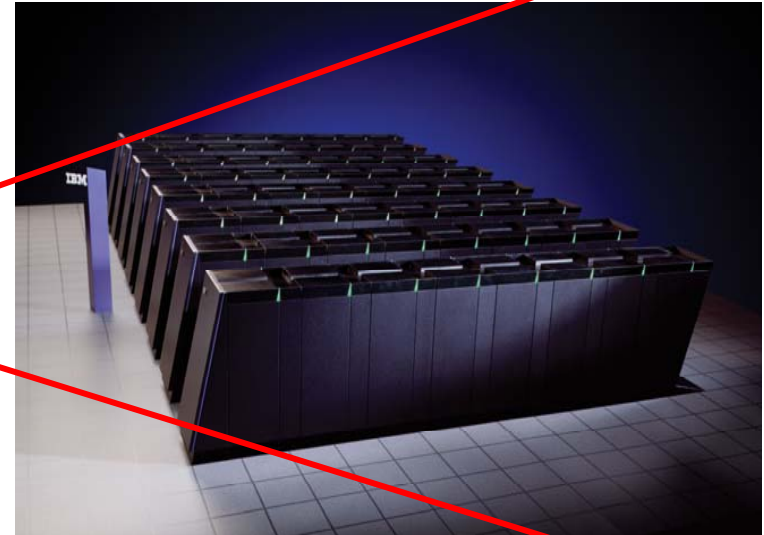


Soluzioni ?



1. SuperComputer

- **BlueGene** e' il supercomputer piu' potente (di cui sia stata svelata l'esistenza)
 - equivale a ~ 10000 CPU attuali
 - costa 250 M\$, se ne occupano in 100
 - ce ne servirebbero 10, costo simile all'acceleratore
 - e non da' soluzioni riguardo allo storage



2. Mega farm di PC

- approccio alla Google, farm piu' grande di cui si abbia conoscenza ~ 10000 PC in un singolo capannone





MegaFarm



- Naturalmente potrebbe funzionare, ma
 - richiederebbe di riunire tutte le persone che si occupano di calcolo a LHC al CERN
 - richiederebbe che le funding agencies fossero d'accordo a pagare una struttura CERN

- Altre soluzioni?



MONARC

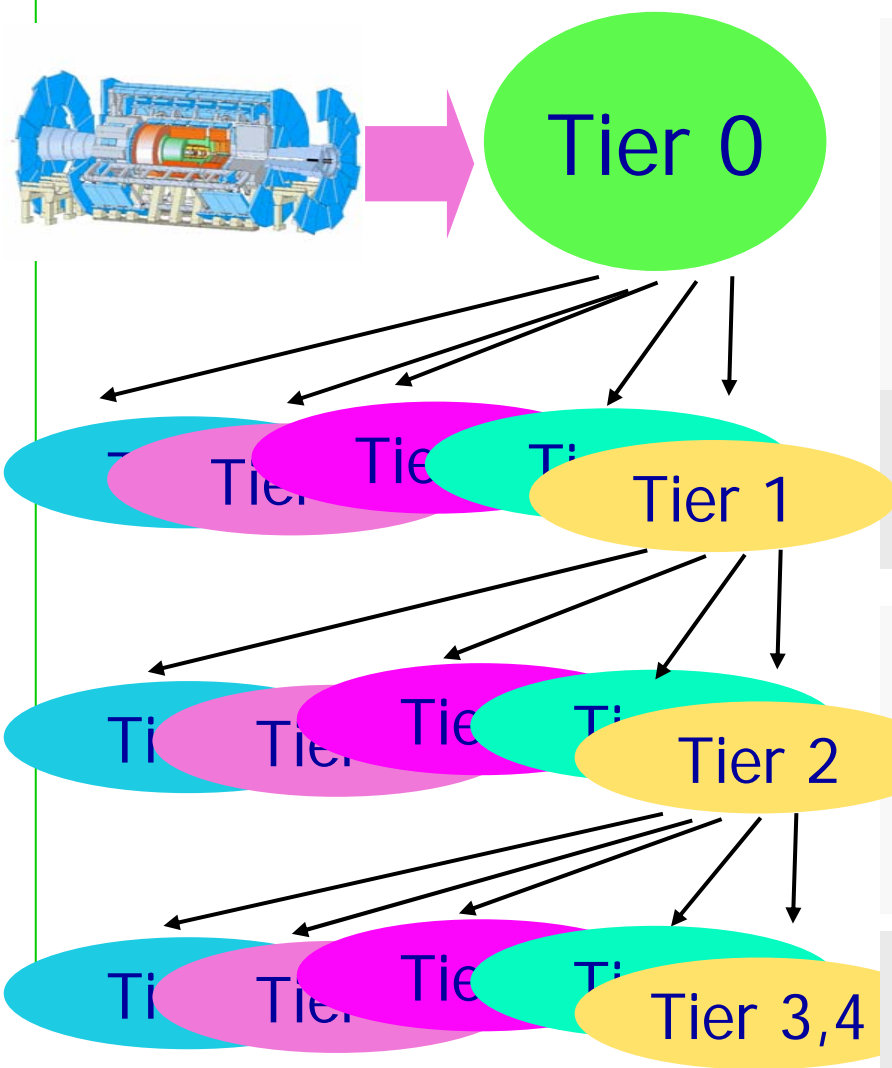


- Models of Networked Analysis at Regional Centres for LHC
 - definisce un modello di analisi basati su
 - calcolo distribuito
 - una gerarchia di centri di calcolo

- 2001 – Hoffman Review definisce *MONARC*
come baseline per il calcolo a LHC



RC



CERN

Storage della copia master dei dati RAW
calibrazioni "veloci" e monitoring
una prima ricostruzione

Si dividono una copia di "backup" dei RAW
ricostruzioni e calibrazioni piu' accurate

gran parte dell'attivita' di
analisi/simulazione avverra' qui
coprono I bisogni di una comunita' ~ 50
utenti

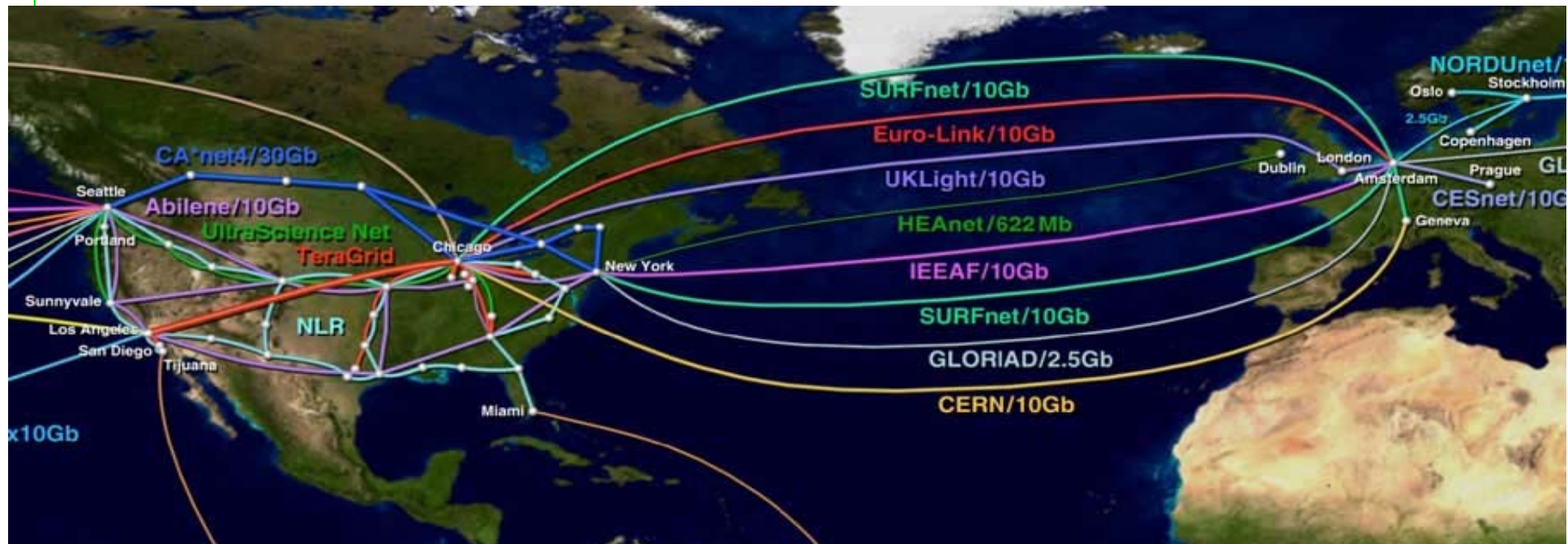
ogni altra risorsa da piccolo cluster
universitario, a macchina desktop, a
portatile ...



Limitandosi ai primi livelli ...



- Si stima una moltiplicazione x5 ad ogni livello
 - ~5 Tier1, per grandi nazioni e zone geografiche



■ bisogno di tanta rete, ma in modo gerarchico



Maggiori problemi di gestione?



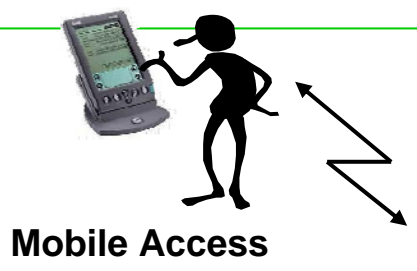
- Fortunatamente, nel frattempo si e' avuta un'evoluzione per gran parte inaspettata ai tempi di MONARC – la nascita delle GRID
- soluzione per il problema della complessita'
- ancora in parte da dimostrare, ma ... vediamo!



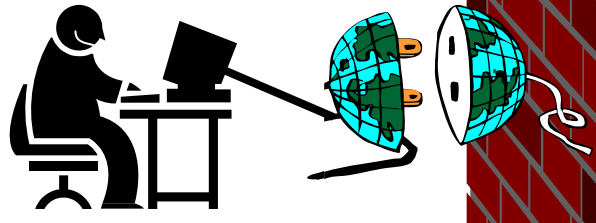
semplificazione della gestione



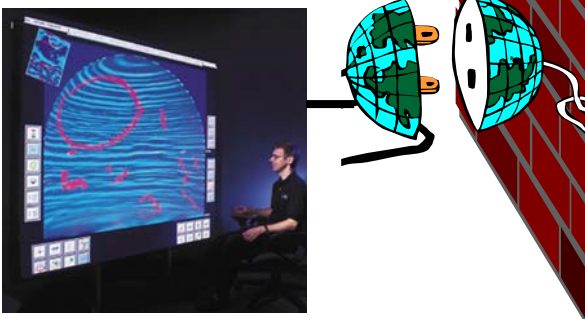
- La definizione di GRID viene dalla similitudine con la rete elettrica
 - quando attaccate un elettrodomestico, aspettate che funzioni senza che dobbiate specificare quale centrale elettrica usare, quale strada debba fare il segnale ecc
- il calcolo per nostra sfortuna al momento funziona in modo diverso:
 - dovete loggarvi su di una macchina che accetti di lavorare per voi, cercare il file da processare e copiarcelo, recuperare il risultato ..
- ma con la GRID ...



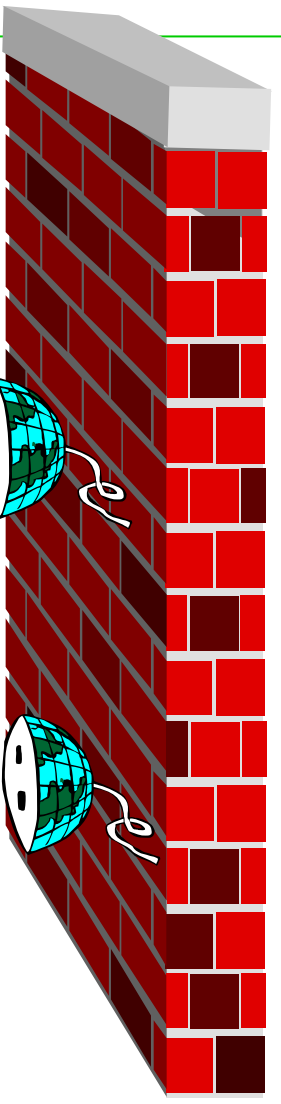
Mobile Access



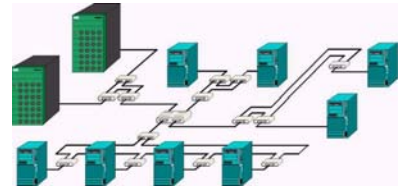
Workstation



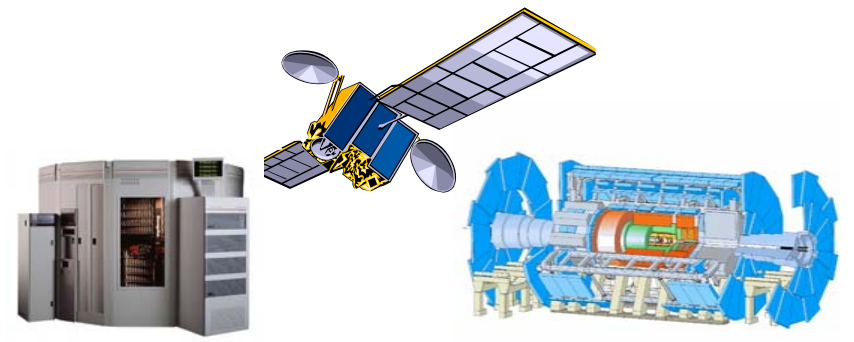
Visualising



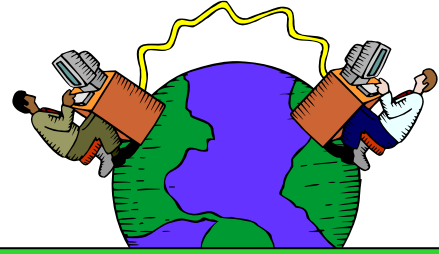
GRID MIDDLEWARE



Supercomputer, PC-Cluster



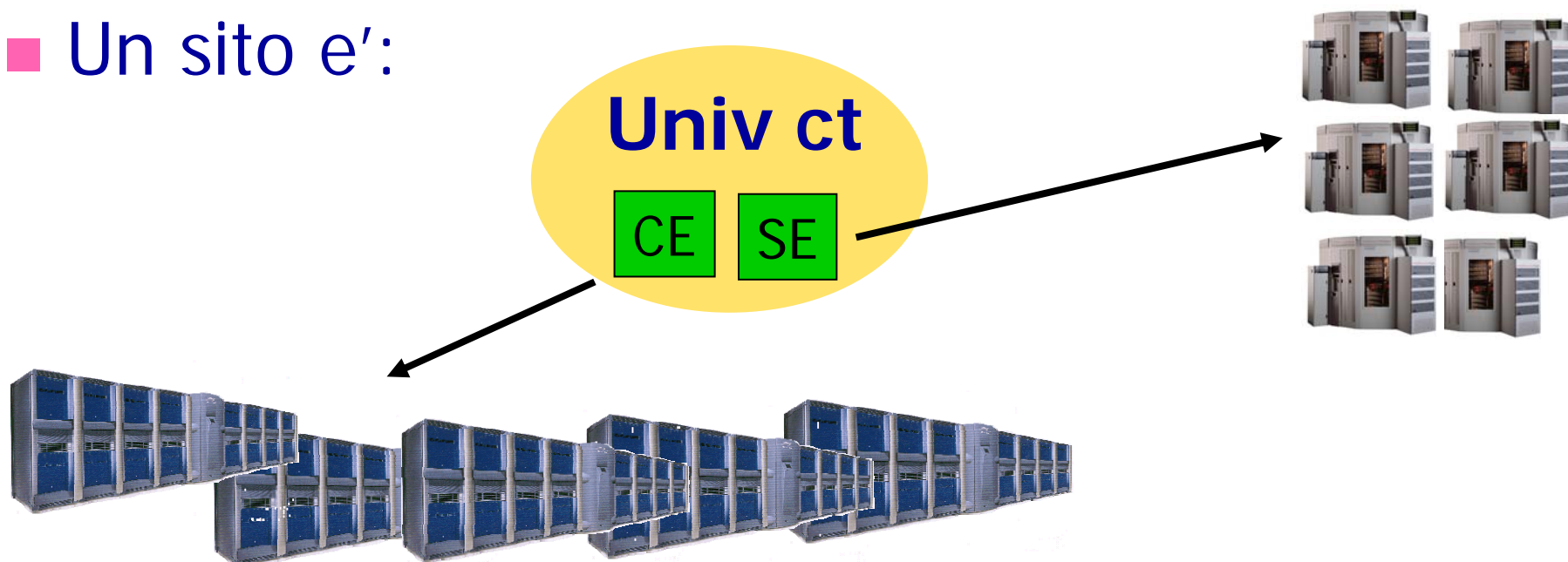
Data-storage, Sensors, Experiments



Internet, networks



- La GRID ben si adatta al modello proposto da MONARC, in quanto permette accesso trasparente a risorse sia di CALCOLO (farm) che di STORAGE (i famosi PB di cui prima)
- Un sito e':

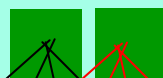




Su scala geografica

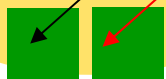


GRID



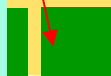
- il fisico non deve/vuole sapere:
 - dove va a finire il suo job
 - dove sono i dati
- vuole solo il risultato in tempi piu' piccoli possibile

Univ PI

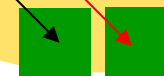


decisione automatica se spostare i job dai dati o i dati dove ci sono CPU

Univ NA



Univ MI





Cosa Esiste?

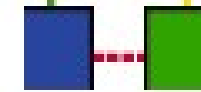
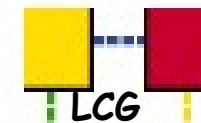


■ Premessa:

- parlo solo della rete LCG,
 - LHC Computing Grid
 - pensata per esperimenti LHC
- ci sono altre realta' altrettanto interessante (alien, grid3, nordugrid...)
- gLite sostituirà a breve LCG



CrossGrid



Enabling Grids for E-science in Europe



Condor High Throughput Computing



■ Stato Attuale:



Grid LCG



ba.infn.it	4	352	16	95%	78	34	83%	216K	25	50	88%	313.4 Gb	701.3 Gb	58%
bo.infn.it	8	84	48	8%	10	0	17%	81K	17	34	26%	1.8 Tb	2.6 Tb	82%
ca.infn.it	4	56	0	100%	14	28	100%	48K	9	14	100%	132 Gb	142.1 Gb	7%
cnaf.infn.it	17	40	0	100%	10	33	100%	48K	5	10	80%	1.5 Tb	1.7 Tb	10%
cr.cnaf.infn.it	13	8752	5921	82%	442	203	7%	2M	177	354	0%	56.1 Tb	87.8 Tb	86%
ct.infn.it	4	224	88	61%	26	15	23%	131K	13	26	12%	1.4 Tb	2 Tb	82%
dma.unina.it	4	8	8	0%	0	0	0%	6K	2	2	1%	-	-	-
egee.unile.it	4	8	0	100%	2	25	100%	-	-	-	-	-	-	-
esrin.esa.int	2	24	24	0%	0	0	0%	-	-	-	-	436.8 Gb	436.8 Gb	0%
fe.infn.it	4	56	0	100%	14	10	100%	53K	7	14	0%	3.5 Gb	25.6 Gb	86%
fi.infn.it	4	144	132	8%	3	4	2%	107K	9	18	12%	720.9 Gb	722.1 Gb	0%
le.infn.it	4	8	8	0%	0	0	0%	3K	1	2	0%	15.6 Gb	17.4 Gb	10%
lnf.infn.it	4	24	0	100%	6	63	100%	34K	3	6	100%	440.2 Gb	1.1 Tb	62%
lnl.infn.it	6	1082	56	95%	172	43	75%	967K	98	196	86%	187.3 Gb	1.3 Tb	86%
mi.infn.it	5	270	245	9%	5	0	2%	294K	32	64	3%	1.5 Tb	3 Tb	4%
na.infn.it	8	144	24	83%	30	197	84%	182K	18	36	81%	337 Gb	870.3 Gb	61%
oat.ts.astro.it	4	12	12	0%	0	4	0%	5K	2	1	5%	28.9 Gb	34.3 Gb	14%
pd.infn.it	6	584	20	97%	92	259	82%	496K	50	102	86%	9.6 Tb	10.1 Tb	5%
pg.infn.it	4	104	104	0%	0	0	0%	103K	13	26	0%	214.6 Gb	216.5 Gb	1%
pi.infn.it	11	240	0	100%	58	42	100%	105K	17	34	88%	105.9 Gb	123.8 Gb	15%
roma1.infn.it	8	208	0	100%	52	48	100%	261K	26	52	96%	518.7 Gb	973.2 Gb	47%
roma2.infn.it	4	32	0	100%	8	21	100%	67K	4	14	68%	1.1 Tb	1.1 Tb	3%
sns.it	4	40	40	0%	0	0	0%	11K	2	2	0%	31.4 Gb	34.3 Gb	9%
to.infn.it	6	280	5	98%	55	25	92%	297K	28	56	5%	28.7 Gb	1.9 Tb	99%
ts.infn.it	4	8	0	100%	2	28	100%	4K	1	2	100%	28 Gb	35.1 Gb	20%



ppm.marcus.it	2	20	79	30	88%	99%	-	-	-	-	-	494.6 Gb	517.9 Gb	4%
prd.hp.com	1	5	0	0	0%	0%	-	-	-	-	-	29.8 Gb	34.2 Gb	13%
psn.ru	1	3	4	0	6%	11%	-	-	-	-	-	170.2 Gb	172.3 Gb	1%
pv.infn.it	1	3	0	0	0%	0%	8K	3	6	0%	-	-	-	-
roma1.infn.it	2	8	52	48	100%	100%	261K	26	52	98%	1.5 Tb	2.8 Tb	7%	
roma2.infn.it	1	4	8	20	100%	100%	67K	4	14	96%	1.1 Tb	1.1 Tb	3%	
sara.nl	-	-	-	-	-	-	-	-	-	-	-	3 Tb	6.4 Tb	5%
saske.sk	1	3	0	0	0%	0%	-	-	-	-	-	582.6 Gb	1.2 Tb	5%
scai.fraunhofer.de	1	5	6	0	4%	13%	-	-	-	-	-	229.5 Gb	249.5 Gb	8%
scotgrid.ac.uk	1	6	0	0	-	-	-	-	-	-	-	-	-	-
shef.ac.uk	-	-	-	-	-	-	-	-	-	-	-	2.1 Tb	2.2 Tb	1%
sinp.msu.ru	1	6	1	0	0%	2%	-	-	-	-	-	1.8 Tb	2 Tb	9%
sns.it	1	4	0	0	0%	0%	11K	2	2	0%	31.3 Gb	34.3 Gb	9%	
tier2.hep.man.ac.uk	2	6	3	1	1%	1%	-	-	-	-	-	62 Gb	90.6 Gb	32%
tifgrid.res.in	1	3	2	0	3%	8%	-	-	-	-	-	30.6 Gb	34.3 Gb	11%
to.infn.it	1	5	55	30	92%	98%	285K	28	54	8%	28.7 Gb	1.9 Tb	99%	
triumf.ca	2	7	174	202	95%	97%	-	-	-	-	-	548.7 Gb	1.1 Tb	5%
ts.infn.it	1	4	2	28	100%	100%	4K	1	2	100%	28 Gb	35.1 Gb	20%	
uibk.ac.at	1	2	20	83	100%	100%	-	-	-	-	-	547.4 Gb	1.6 Tb	67%
ui.savba.sk	1	4	0	0	0%	0%	-	-	-	-	-	25.7 Gb	107 Gb	76%
usatlas.bnl.gov	1	5	0	0	-	-	-	-	-	-	-	4.8 Gb	8 Gb	0%
usc.cesga.es	1	3	1	0	1%	2%	10K	1	2	0%	52.3 Gb	90.6 Gb	2%	
weizmann.ac.il	1	5	0	0	0%	0%	-	-	-	-	-	-	-	-
westgrid.ca	-	-	-	-	-	-	-	-	-	-	-	2.3 Tb	10.1 Tb	78%
zam.kfa-juelich.de	-	-	-	-	-	-	-	-	-	-	-	-	-	-
zib.de	-	-	-	-	-	-	-	-	-	-	-	43.1 Gb	45 Gb	4%

Italia:

~ 1500 CPU Online

A livello mondiale:

> 6000 CPU

IN PRODUZIONE!



Ma ... funziona?



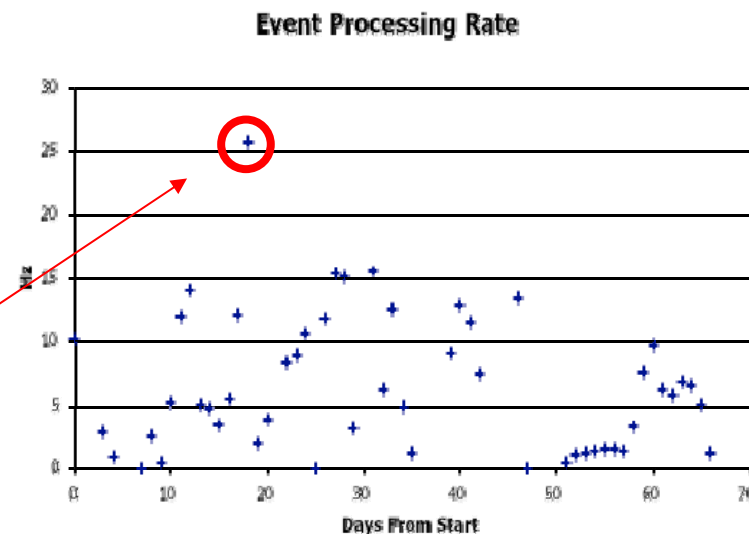
- Naturalmente nessuno vuole andare in presa dati nel 200x e poi accorgersi del tracollo
- Il 2004 e' stato per tutti gli esperimenti LHC anno di Data Challenge:
 - **“simulazione delle operazioni giornaliere dopo il 2007, su scala ridotta”**



DC nel 2004 - CMS

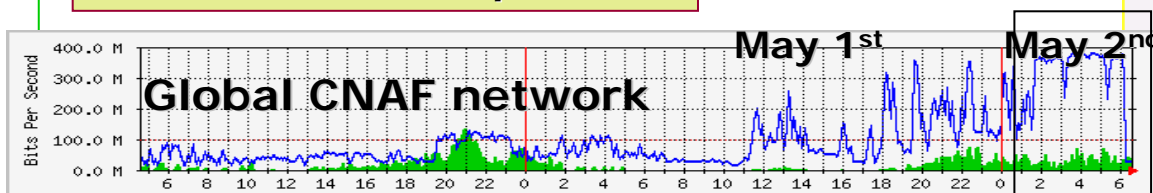
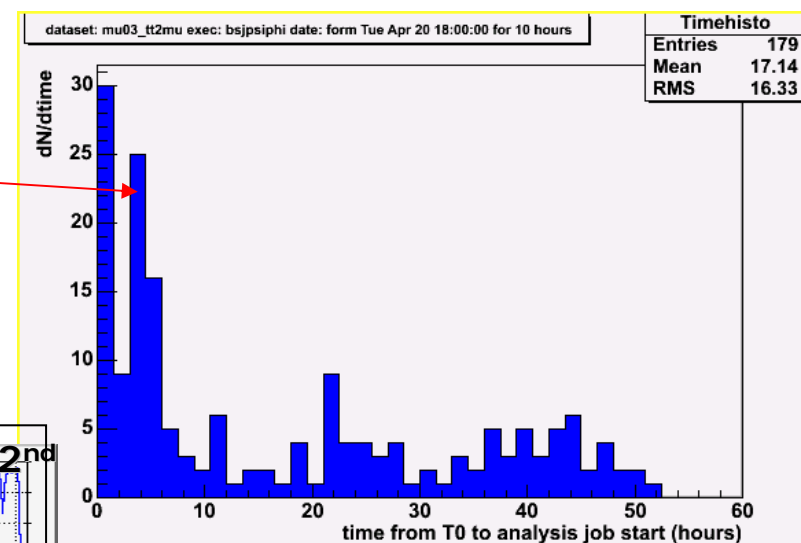


- DC04 ha testato la ricostruzione al **Tier 0** (produzione di dati ricostruiti) **al 25% del rate 2007**
 - 25 Hz sostenuti per almeno 24 ore
 - 80M eventi prodotti
- Spostamento dati **T0**→**T1** via sistema proprietario ma con autenticazione da GRID
- Analisi ai **T2** in tempo reale!



~340 Mbps (>42 MB/s)
sustained for ~5 hours(max
was **383.8 Mbps**)

~4 TB/d

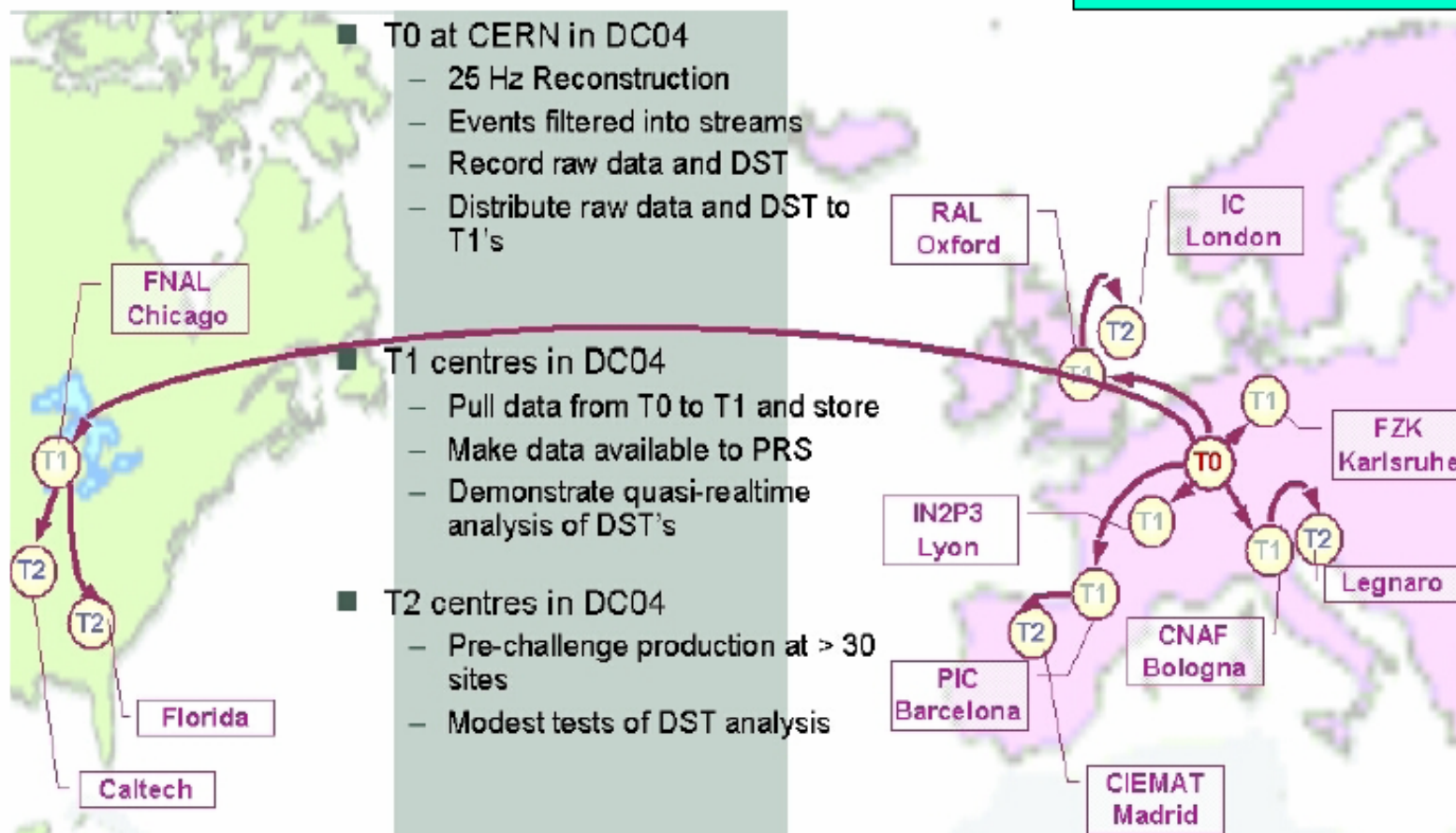




LHCC Review - 22 Nov. 2004

CMS: DC04 (2)

75 M Events
425 kSI2k-years
96 TB in POOL





ATLAS



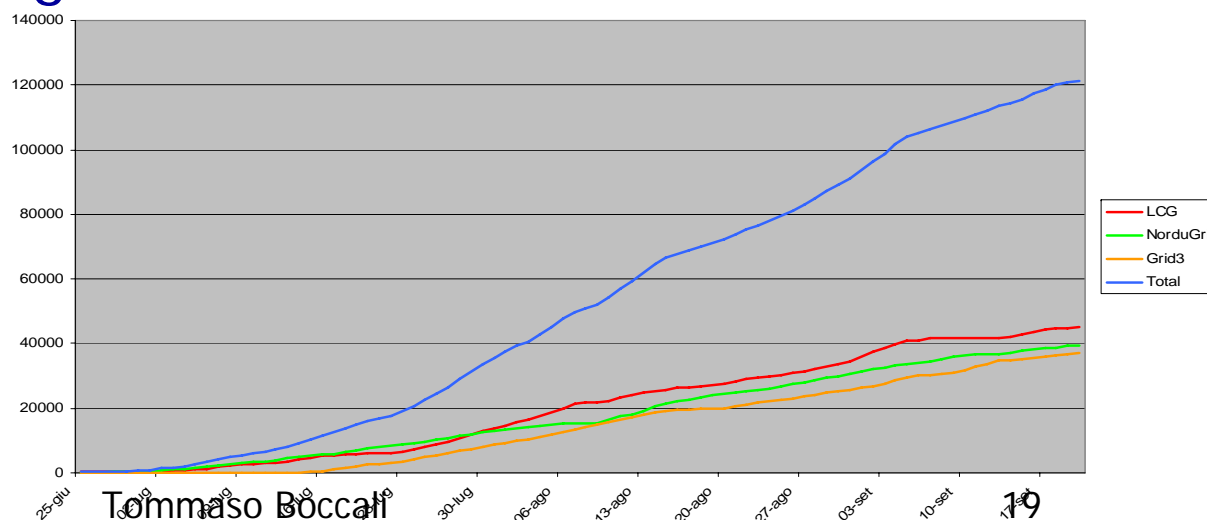
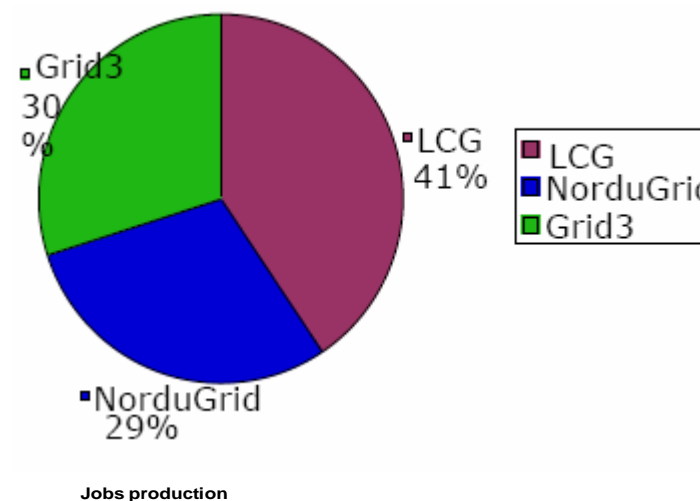
- Simulazione effettuata su 3 grid diverse

- ok
- 10M eventi + altri per calibrazioni
- 35 TB

- test di operazioni T0 (ricostruzione dai raw) al 10% del rate

- analisi distribuita su grid

- da ultimare



30/3/2005

Tommaso Boccali

19



ALICE



molto ambizioso: produrre e analizzare in modo realistico il 10% dei dati presi nel primo anno

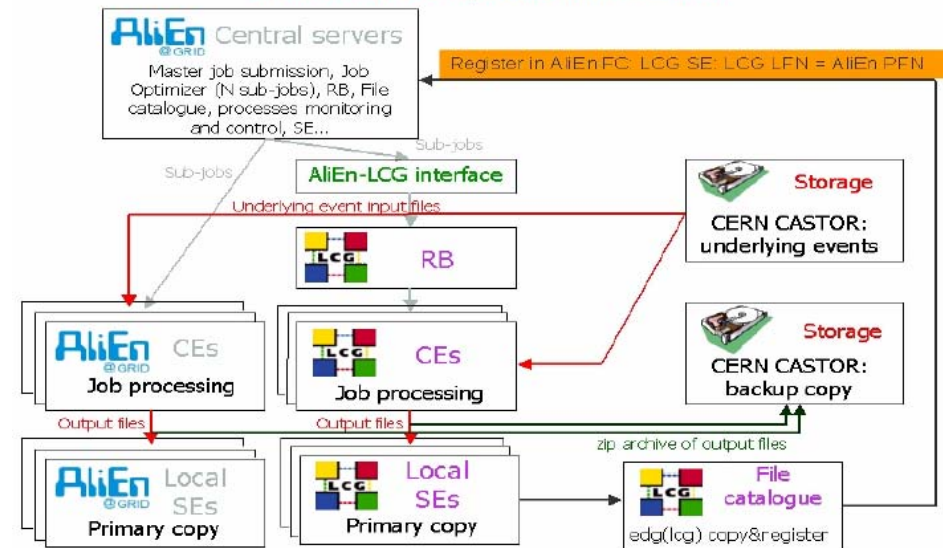
Tre fasi:

1. produzione distribuita delle collisioni Pb-Pb/p-p underlying ~ 20 M (ma complessi!) ~ 200 TB
2. produzione dei segnali per di trigger
3. analisi distribuita

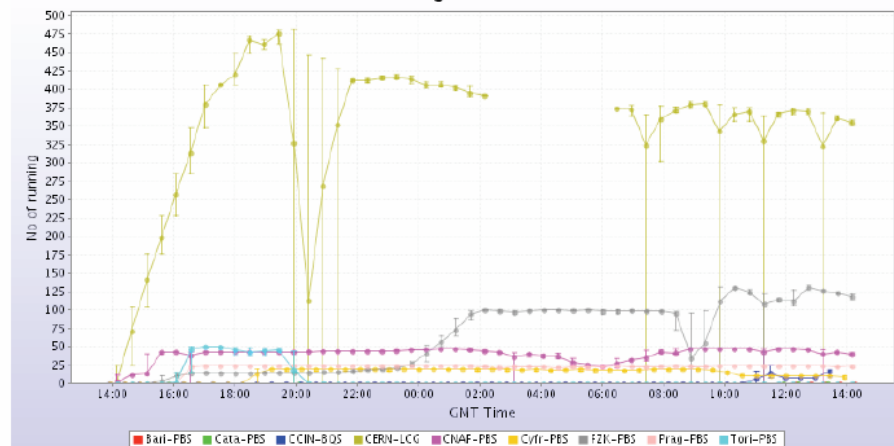
Alien integrato con LCG

LCG viene visto come un unico sito da Alien

Structure of event production in Phase II



Running Processes

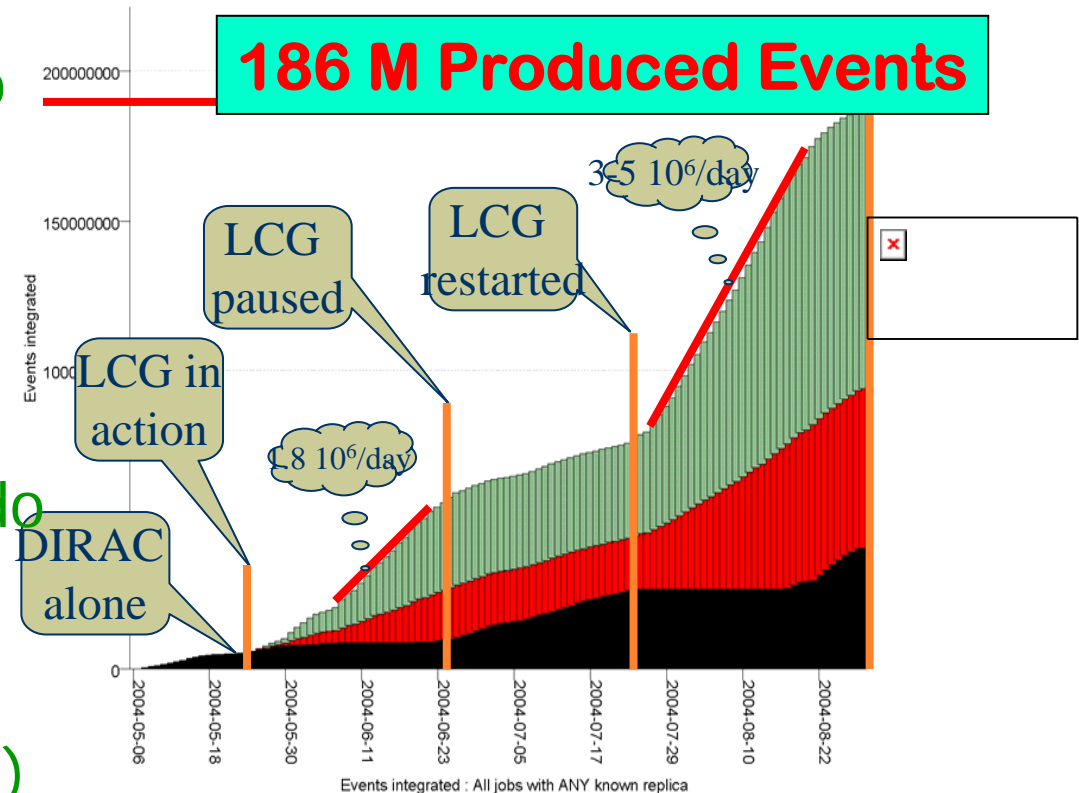




LHCb



- Goal di Fisica:
 - studi di Trigger di Alto Livello
 - studi di background
- Produzione su grid di
 - 30M eventi di segnale
 - 140M di eventi di fondo
- Anche qui tre fasi
 - simulazione (fatta)
 - ricostruzione (in corso)
 - analisi (da fare)





Commenti ...



- Il fatto che **4/4 DC** su Grid alla fine abbiano funzionato (con se problemi medio/piccoli) e' un risultato straordinario visto che non era solamente un test
 - **Prodotti ~ $O(10^9)$ eventi**
 - **CPU in gioco ~ 10000**
 - **Generate simulazioni per $O(500TB)$**

- **Problemi?**
 - Rete e CPU hanno funzionato bene, esiste qualche preoccupazione in piu' per quello che riguarda lo storage



Conclusioni



- Ci sono pochi dubbi che il modello MONARC integrato da (una o piu' delle) GRID sara' il modello di analisi all'inizio di LHC
- I Data Challenge effettuati fino a questo momento sembrano sostanzialmente validare le soluzioni scelte per rate fino ad un quarto di quello iniziale aspettato
 - il fattore quattro mancante dovrebbe farlo la tecnologia
- Se dal lato delle reti e della potenza di calcolo non sembrano esserci sostanziali problemi, forse qualcosa di piu' e di meglio si potrebbe farlo con lo storage.
- Domanda: deve davvero essere tutto cosi' complicato?
 - al momento questo e' lo stato dell'arte del calcolo distribuito; soluzioni piu' semplici sono comunque in vista/in fase di progetto



Internet Explorer 13.19
 https://grid003.ct.infn.it/pda/infnpd

genius

Job Queue

#	4
Job ID	bYqRSTNJoW6xOuWzVkrG
JDL Name	/home/barbera/hostname.j
Last Update	Thu Oct 9 07:52:14 2003
Destination	testbed001.cnaf.infn.it:211 pbs-medium
Status	Done
Exit Code	0
Action	Get Output

Visual. Strum.

particles travelling in the LHC

10⁻¹⁸ seconds
10⁻²² seconds

Why do particle physicists need accelerators? Recreating the conditions of the early universe

LHC@home is a screen-saver program that uses your computer's idle time to simulate particle travelling in CERN's new accelerator

Help scientists to build the LHC!

NEWS

Welcome to the LHC@home Project!

We are currently in beta-test mode, correcting some bugs and improving our numerical accuracy. Our present user base of about 6000 users is sufficient, so for now we are not accepting new users. When we re-open for production running in a while, we will announce it and then welcome a bigger user base.

Thank you for your interest!

News is available as an [RSS feed](#).

Server Status:
UP
Out of work

