

# CDF Computing Status



Donatella Lucchesi

University and INFN of Padova  
for CDF Italian Computing Group

## Outline

- CDF computing status
- CDF usage of Tier1 and Grid.it
- Requests for 2008

# Status of data processing

---

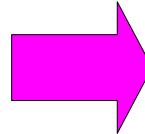
- Almost all the analysis presented at the LP07 used  $\sim 2\text{fb}^{-1}$  of data.
- All data collected are processed, but that with silicon problems.
- Method developed to include the trigger prescale factor in the event format at production time.
- Ntuples data available for analysis right away after production for several physics group
- Production farm almost merged in the CDFCAF, first step toward running production on Grid

# Infrastructures upgrades: disk space

---

- No change in the model, use of new fileservers which are BlueArc, high performances NFS servers
- Migration of data disks almost done, transparent to the users
- Migration of university fileservers that have to be retired this fall including Italian fileservers
- New space will be assigned to physics group

# Infrastructures upgrades: farms



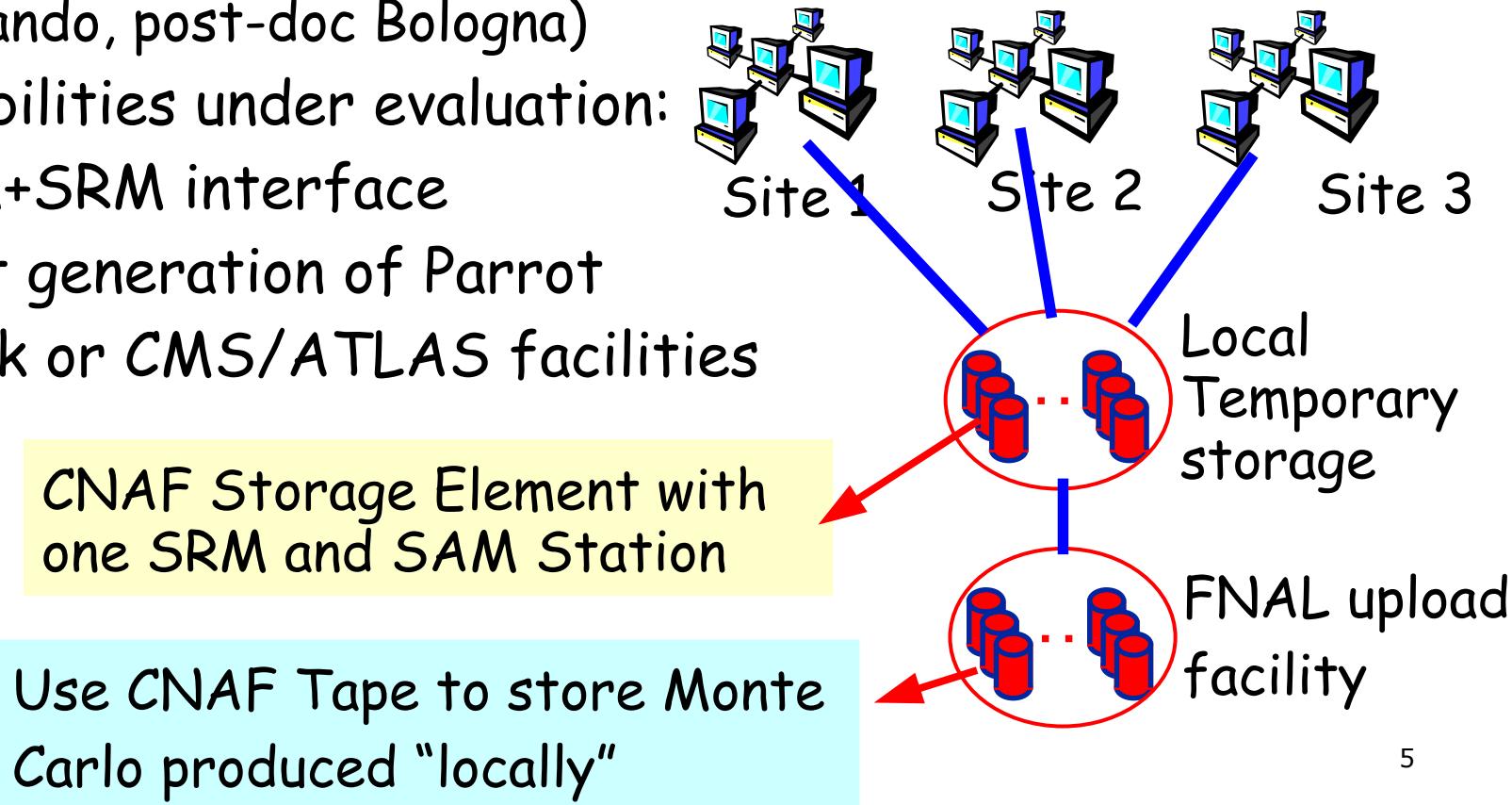
Move toward a 2 or 3 submission points:

CDF Grid transition task force charged to migrate all the farm adopting standard grid middleware where possible.

F. Wuerthein, I. Sfiligoi, D. Benjaming, F. Moscato,  
M. Zvada +CD experts, DL, G. Compostella, S. Pagan Griso

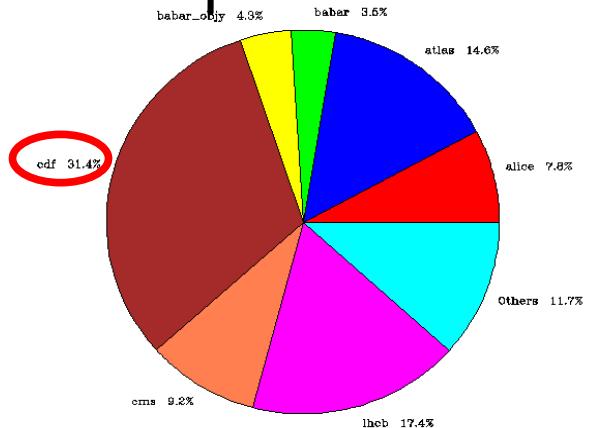
# Infrastructures upgrades: Data/MC movement

- Monte Carlo data produced on Grid wn copied (rcp) to FNAL
- Bottleneck: fileserver FNAL ⇒ need a facility to move data quickly from wn to release cpu, new project started (laureando, post-doc Bologna)
- Possibilities under evaluation:
  - ✗ SAM+SRM interface
  - ✗ Next generation of Parrot
  - ✗ Stork or CMS/ATLAS facilities



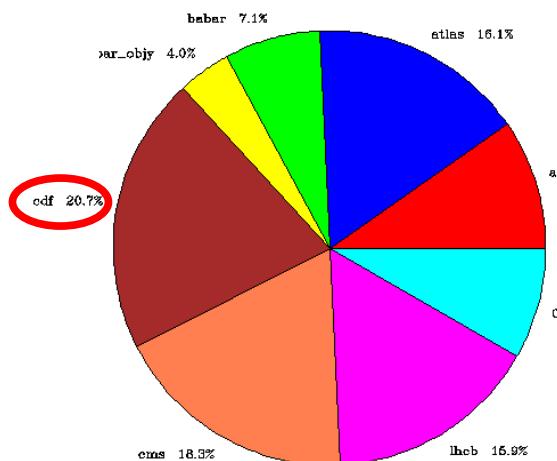
# CDF usage of CNAF

Since April 2005

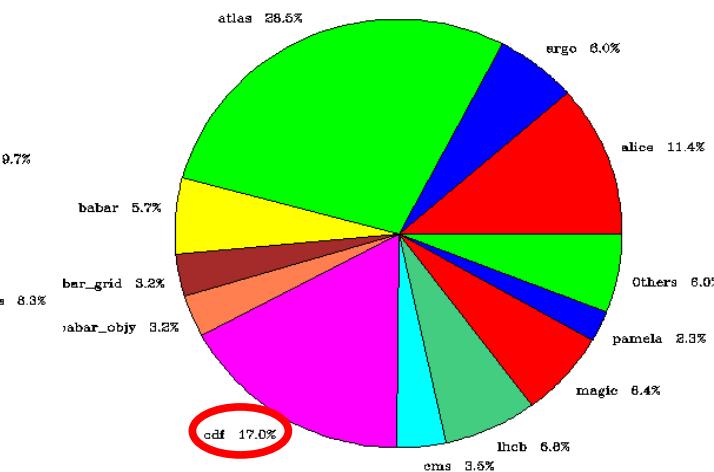


CDF uses ~20% of the CNAF resources

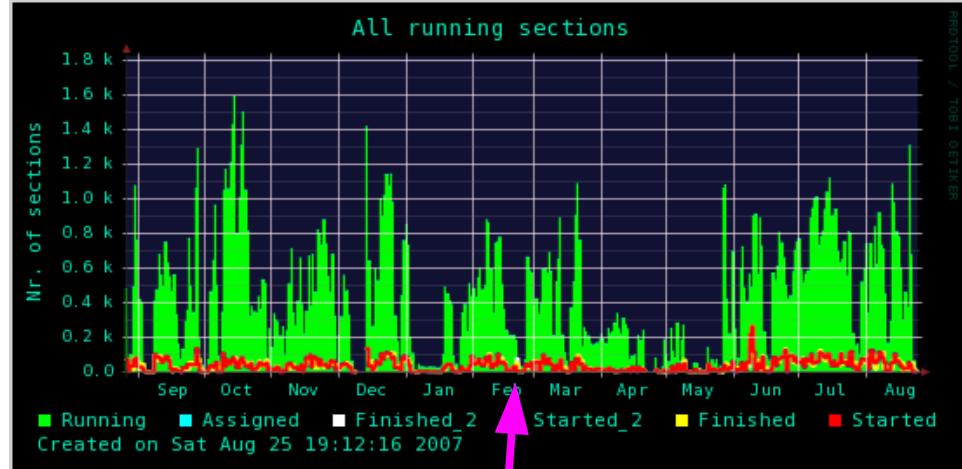
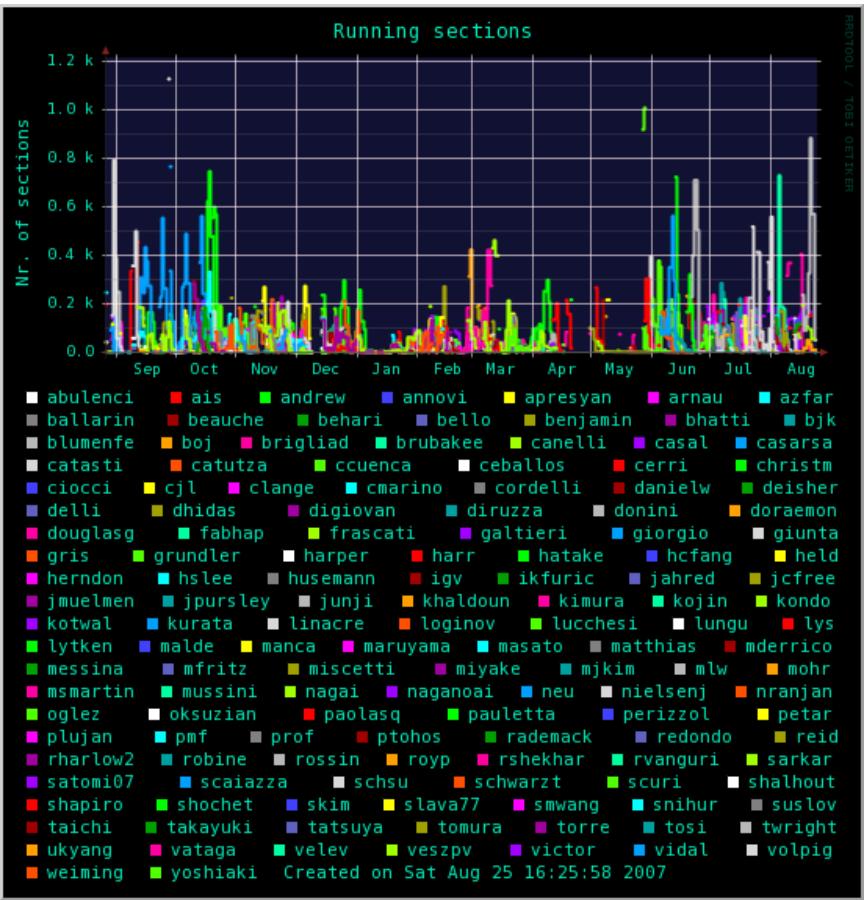
First trim. 2007



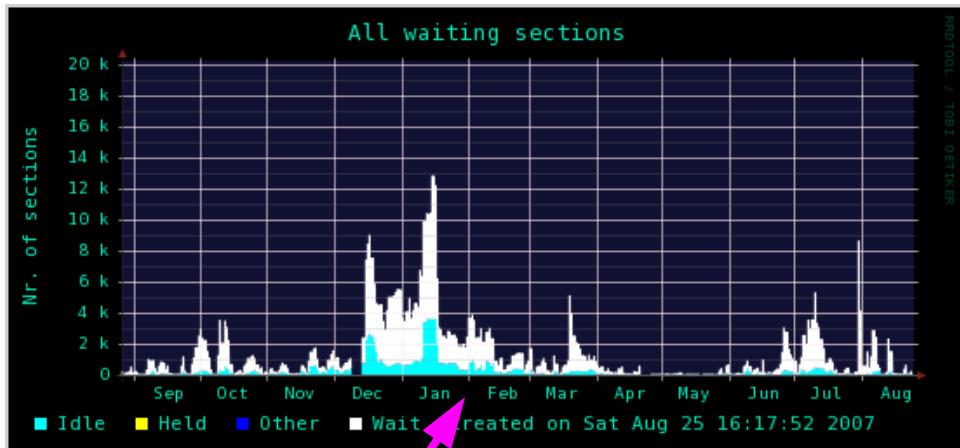
Second trim. 2007



# GlideCAF @CNAF



A lot of running jobs



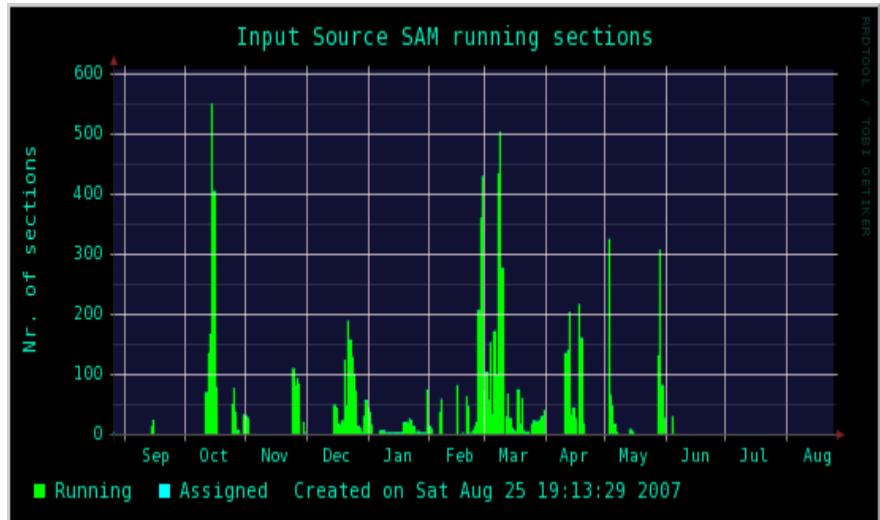
Always waiting jobs

More then 150 active users  
in the last year.

Aug. 28, 2007

# GlideCAF @CNAF

Used also for data



Datasets available:

- B physics skimmed data
- Top, higgs, W ntuples being imported
- Official Monte Carlo ntuples

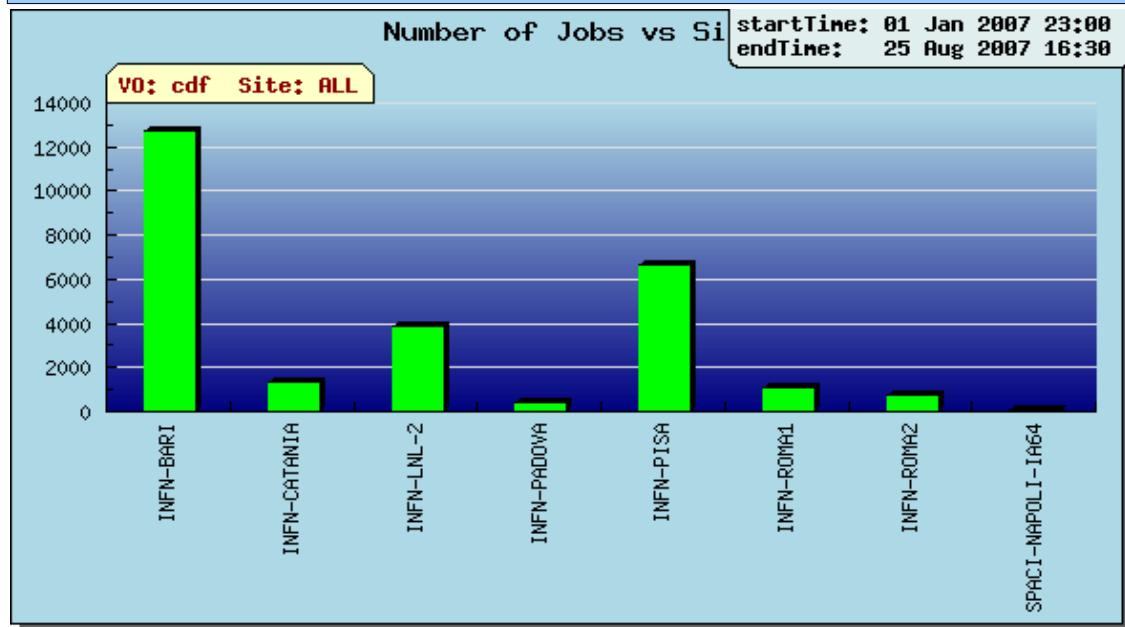
- ✗ Data hosted on GPFS ~ 50 TB used, ~20 TB free, new data are coming
- ✗ Working area ~ 5 TB mainly for italians.
- ✗ New Storage Element to temporary storage Monte Carlo

# LcgCAF configuration & usage

## List of sites accessed

INFN-T1	Italy
INFN-Padova	Italy
INFN-Catania	Italy
INFN-Bari	Italy
INFN-Legnaro	Italy
INFN-Roma1	Italy
INFN-Roma2	Italy
INFN-Pisa	Italy
FZK-LCG2	Germany
IEPSAS	Slovakia
IFAE	Spain
PIC	Spain
IN2P3-CC	France
UKI-LT2-UCL-HE	UK
Liverpool	UK

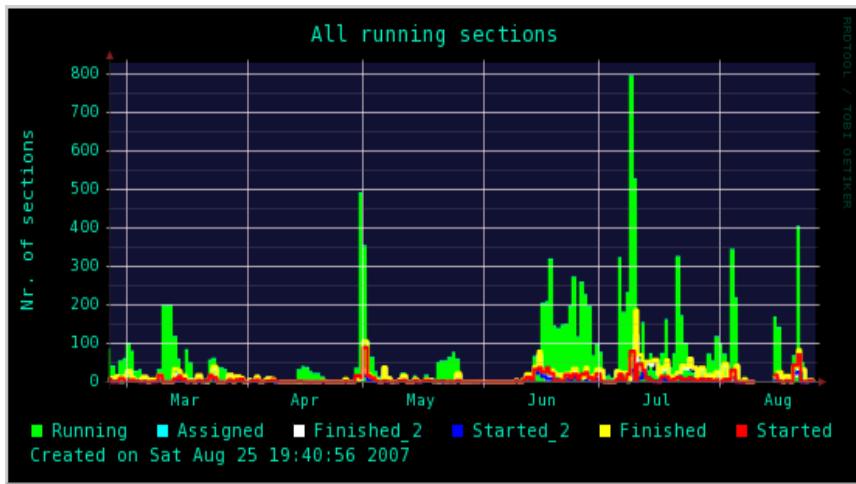
## Job distribution among Italian sites



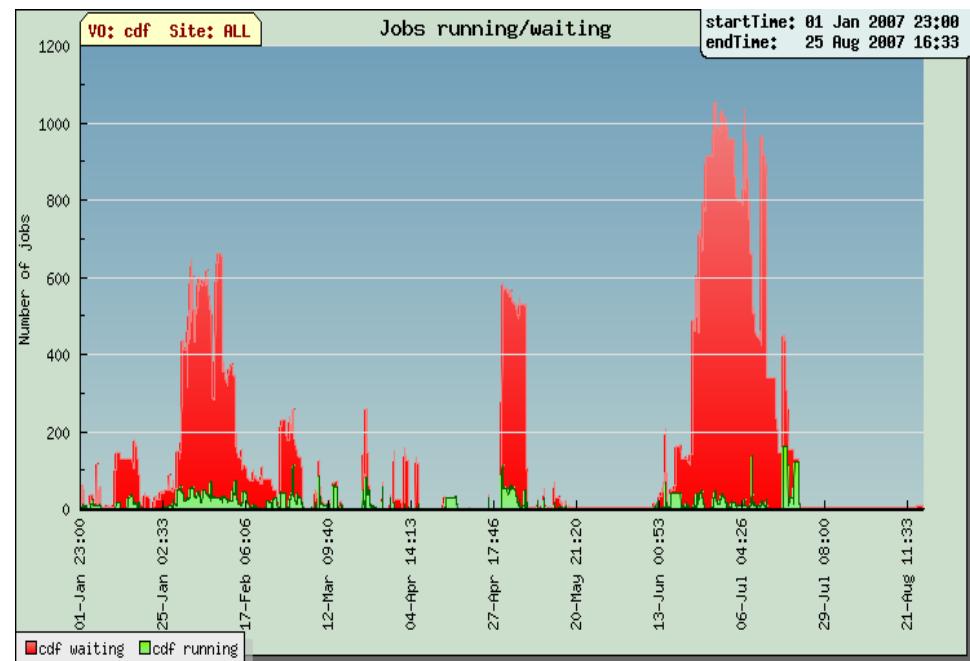
Tier 1 is not included because of monitor/site misconfiguration

# LcgCAF jobs distributions

Running jobs as seen by LcgCAF monitor



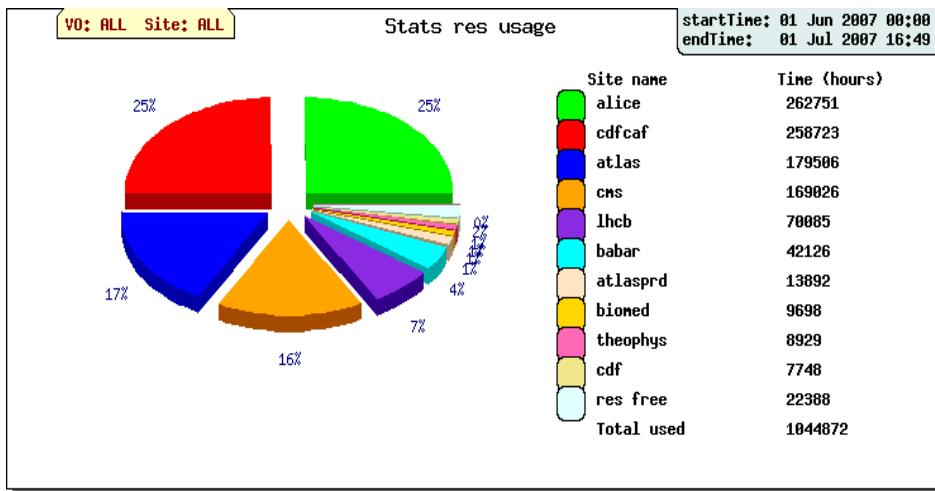
Running/waiting jobs as seen by GridIce



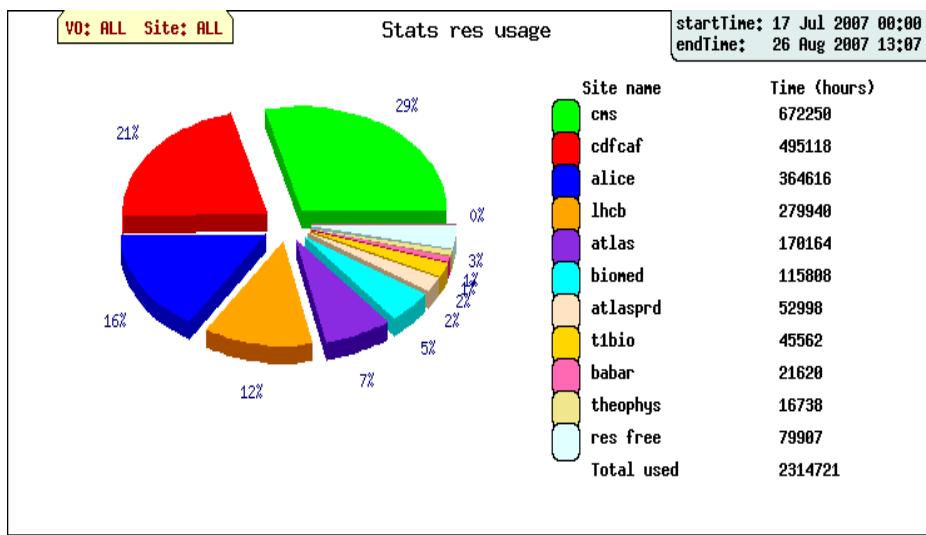
Recently LcgCAF being used more regularly

CDF jobs remain for long time on queues due to a not so efficient resources matching

# CDF Grid resources usage



CDF uses more than 20% of the available resources on Grid



# CDF Requests for 2008

Use F. Forti @CNS1 April 2nd

	2007-Q4			2008-Q4			2009-Q4			2010-Q4		
	CPU (kSi2k)	Disco (TB-N)	Nastro (TB)									
BaBar	680	200	0	1215	350	0	1215	350	0	1215	350	0
CDF	820	100	15	1161	170	15	1290	220	15	1420	270	15
Totale	1500	300	15	2376	520	15	2505	570	15	2635	620	15
<b>Acquistando in due anni</b>	<b>TOT €</b>			<b>540</b>			<b>80</b>			<b>62</b>		
Costo CPU				188			18			14		221
Costo Disco				352			62			48		462
Costo Nastro				0			0			0		0
<b>Acquistando nell'anno in corso</b>				<b>448</b>				<b>69</b>				<b>572</b>
Costo CPU				140			15			13		169
Costo Disco				308			54			41		403
Costo Nastro				0			0			0		0
	2007-Q4			2008-Q4			2009-Q4			2010-Q4		
	CPU (kSi2k)	Disco (TB-N)	Nastro (TB)									
LHCb TIER2	0	0	0	600	0	0	1200	350	0	1700	350	0
<b>Acquistando in due anni</b>				<b>129</b>				<b>84</b>				<b>268</b>
<b>Acquistando nell'anno in corso</b>				<b>96</b>				<b>72</b>				<b>218</b>

# CDF Requests for 2008

---

Requests made to CNS1:

$$\text{cpu: } \frac{(1161-820)}{(1161-820)+(1265-680)} * 188K\epsilon = 73K\epsilon$$

$$\text{Disk: } \frac{(170-100)}{(170-100)+(350-200)} * 352K\epsilon = 112K\epsilon$$

# CDF Requests for 2008

CNAF Plan March 2007																	
Experiment	%	2006			2007			2008			2009			2010			
		CPU KSI2K	DISK TB-N	TAPE TB													
ALICE	22%	154	16	77	286	110	143	1210	550	836	1870	880	1320	3520	1760	1870	
ATLAS	32%	224	40	112	416	160	208	1760	800	1216	2720	1280	1920	5120	2560	2720	
CMS	35%	245	86	123	455	175	228	1925	875	1330	2975	1400	2100	5600	2800	2975	
LHCb	11%	77	26	39	143	55	72	605	275	418	935	440	660	1760	880	935	
Total LHC TIER1		700	168	350	1300	500	650	5500	2500	3800	8500	4000	6000	16000	8000	8500	
BaBar		585	149	0	680	200	0	1215	350	0	1215	350	0	1215	350	0	
CDF		900	66	0	820	100	15	1161	170	15	1290	220	15	1420	270	15	
LHCb TIER2		0	0	0	150	0	0	600	0	0	1200	350	0	1700	350	0	
TOTALE GRUPPO I		1485	214	0	1650	300	15	2976	520	15	3705	920	15	4335	970	15	
AMS2		32	2	16	25	5	16	32	5	24	180	16	128	180	16	128	
ARGO		22	12	28	150	70	186	188	112	366	188	129	546	188	129	546	
GLAST					5	10	0	5	10	10	5	10	20	5	10	20	
MAGIC			1		20	5	4	20	4	8	20	4	12	20	4	12	
PAMELA			4		20	10	16	25	10	32	25	10	48	25	10	48	
Virgo		10	25	75	180	90	130	250	150	200	500	220	250	500	220	250	
TOTALE GRUPPO II		64	43	119	400	190	352	520	291	640	918	389	1004	918	389	1004	
All experiments		2249	426	469	3350	990	1017	8996	3311	4455	13123	5309	7019	21253	9359	9519	
All w/ overlap factor		1874	387	469	2792	900	1017	7497	3010	4455	10936	4827	7019	17711	8509	9519	
CNAF TOTAL (PLAN)		1874	387	469	3000	1000	1000	7497	3010	4455	10936	4827	7019	17711	8509	9519	
CNAF ACTUAL		1570	400	510													
Relative Contingency							0%										
Absolute contingency						0											
Zoccolo duro (TOTAL-CONTINGENCY)						3000	100										
INFN T1 P2P 2005		1800	850	850	2400	120					4100	11500	5800	6000			
INFN T1 P2P 2007		-	-	-	1300	50					3300	10000	5000	5000			

CDF requests:  
 KSI2K: 15% x Tier1  
 DISK : 6% x Tier1

# CDF Requests for 2008

---

- In the firsts 6 months of 2007 CDF at CNAF has used **444 KSP2K** if we project to the end of 2007 we expect to use **~900 KSP2K** that correspond to our request, but we have been reduced to **820**.
- In 2008 CDF can not go below **1100 KSP2K**.
- CDF survives thanks to the usage of Grid.it, part of Monte Carlo production has been moved there.
- Even with that CDF has always a lot of jobs in queues.

# Manpower

---

- D.L. and Rick Snider Co-head
- Fermilab:
  - ✓ SAM support
  - ✓ CAF+dCAF
  - ✓ CDF code and user support
- Italy:
  - ✓ CNAF: G. Compostella
  - ✓ LcgCAF: S. Pagan Griso
  - ✓ Manoj Kumar Jha postdoc Bologna computing
  - ✓ 2 Laureandi
  - ✓ AR at CNAF starting begin 2008

Manpower situation is getting better.

Aug. 28, 2007

# To conclude

---

Talk in 2007:

- ✗ G. Compostella "CDF Computing Experience", IFAE Napoli
- ✗ S. Pagan Griso "Experience of the running HEP experiments in using the EGEE/LCG infrastructure" EGEE User forum
- ✗ G. Compostella "Esperienza acquisita da CDF nell'utilizzo dell'infrastruttura GRID" WorkShop 2007 sul Calcolo e Reti dell'INFN
- ✗ S. Pagan Griso "CDF experience with Monte Carlo production using LCG Grid" CHEP 07
- ✗ Paper submitted to IEEE