# CDFII Computing Status

OUTLINE:

- New CDF-Italy computing group organization
- Usage status at FNAL and CNAF
- Towards GRID: where we are
- Plans and requests

# CDFII Computing group changes

Presented by Stefano Belforte @CN1 for CMS

- S. Belforte keeps for the moment 20% (1 day/week) on CDF: necessary to guarantee the transition toward a new "responsabile nazionale"
- The CDF-Italia projects for software and data handling are mature and well under control :
  - transition from "our machines at FNAL" to "our farm at CNAF" completed with success
  - our CNAF farm merging into the Tier1 common pool almost completed
  - new software infrastructure for CAF re-written by Igor Sfiligoi co-head "Computing and Grid" for CDF at FNAL

# CDFII Computing group changes

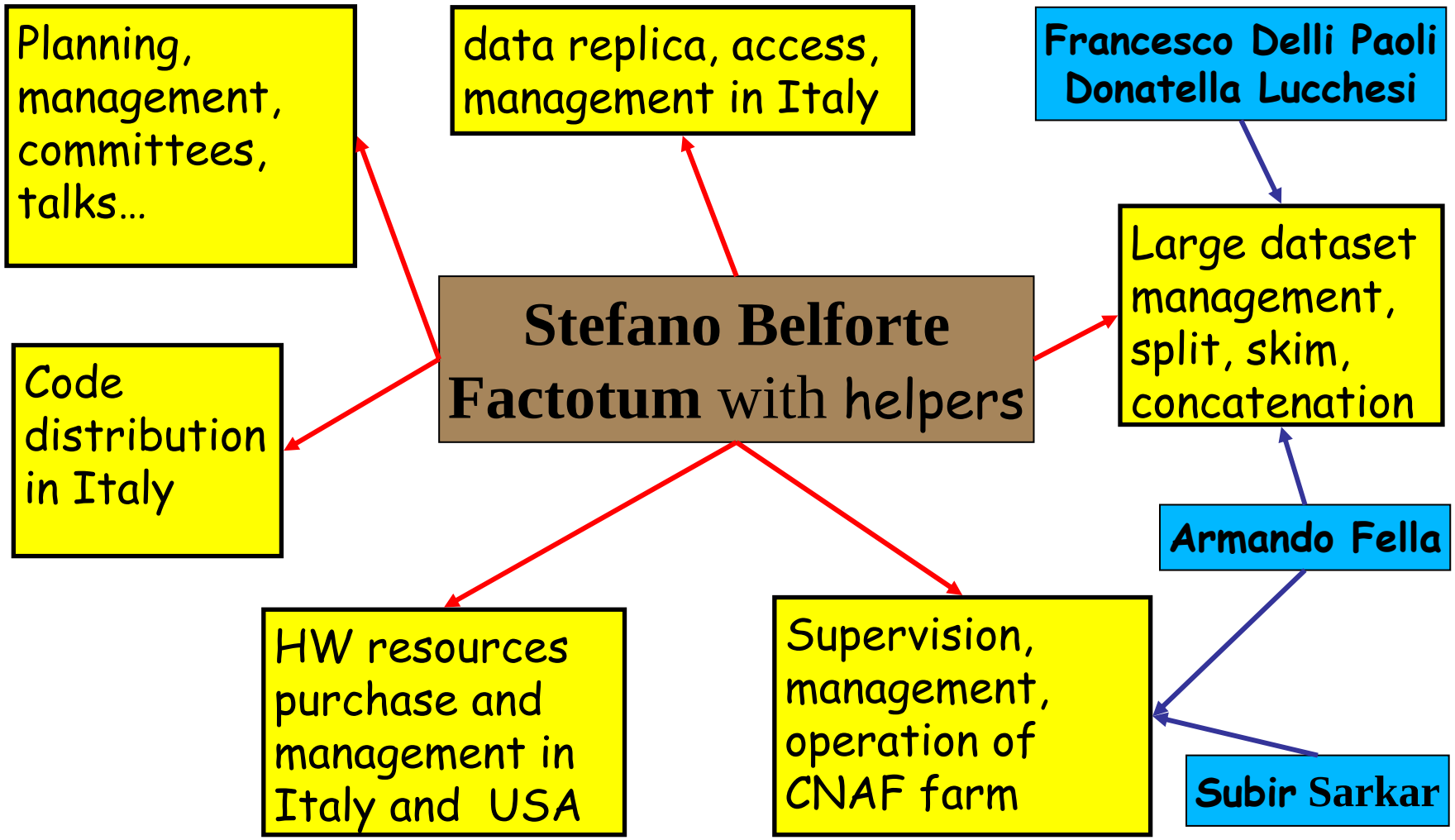
Presented by Stefano Belforte @CN1 for CMS

- Transition toward GRID started and very promising
- Our budget ~0 Euro since two years

- New people arrived in the latest 12 months and new very important role of Igor Sfiligoi:
  - Donatella Lucchesi (art. 23 – Padova)
  - Francesco Delli Paoli (borsa tecn. INFN – Padova)
  - Armando Fella (borsa tecn. INFN – Pisa)
  - Subir Sarkar (A.R. tecn. Tier1 – CNAF)
  - Daniel Jeans (A.R. tecn. INFNGRID – Roma)
- New group organization

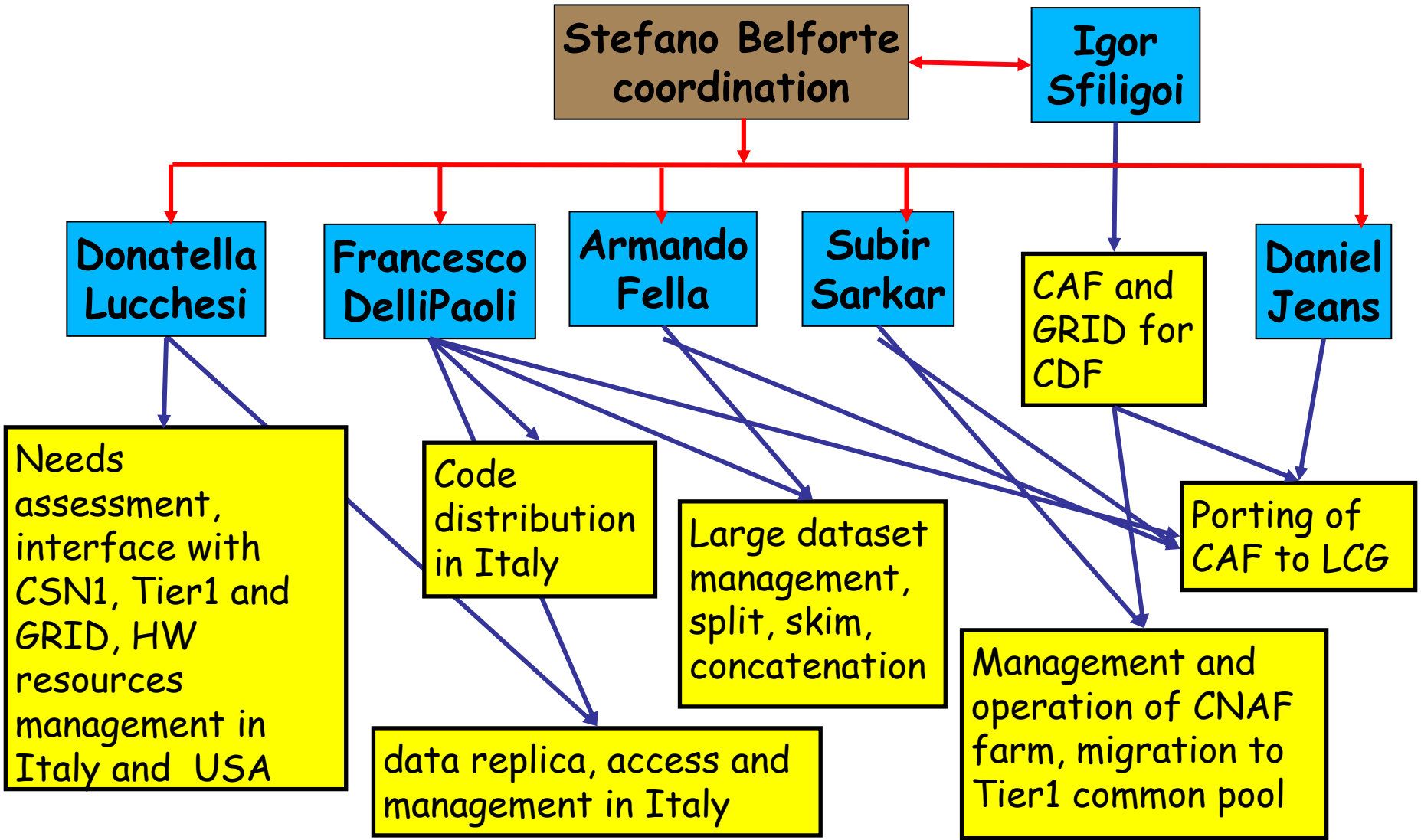
# CDFII-Italy Computing before

Presented by Stefano Belforte @CN1 for CMS



# CDFII-Italy Computing group

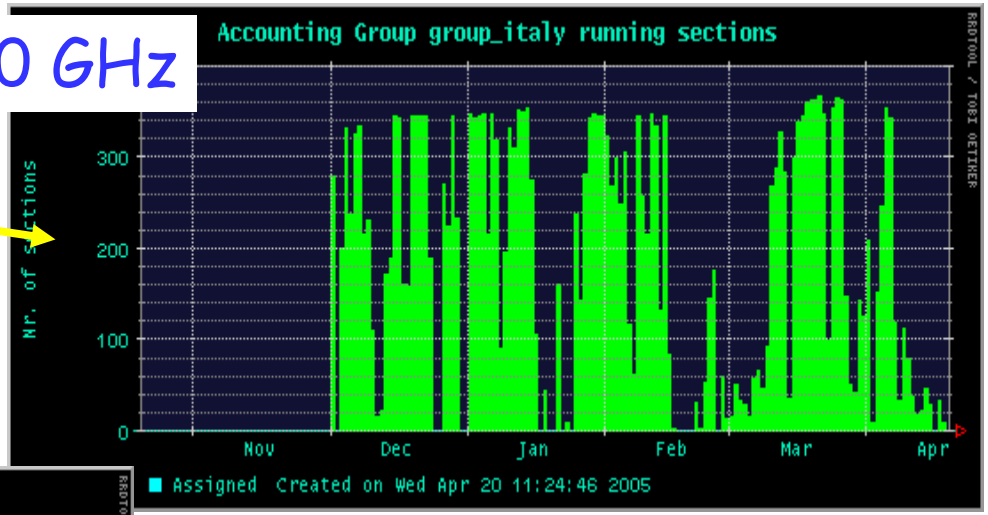
Presented by Stefano Belforte @CN1

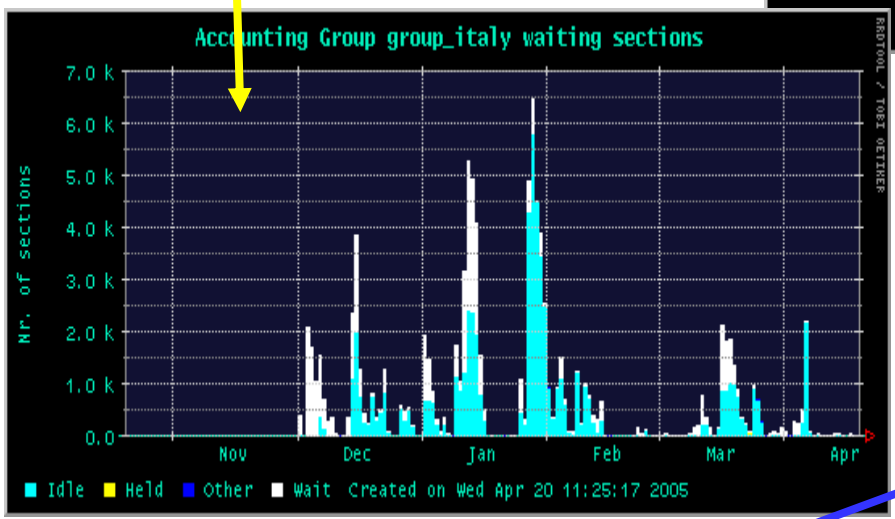


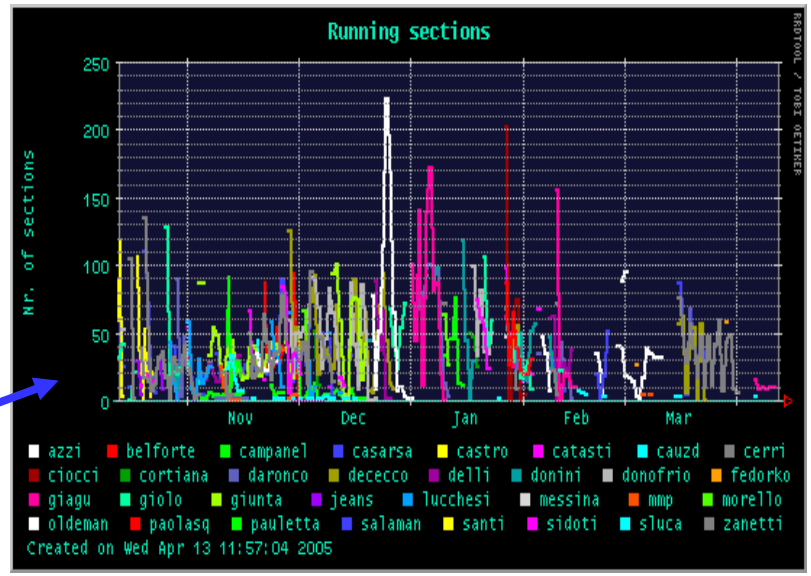
# CAF status at Fnal (latest 6 months)

## CAF reserved queue ~300 GHz

All slots used, relative high demand before analysis approvals for conferences



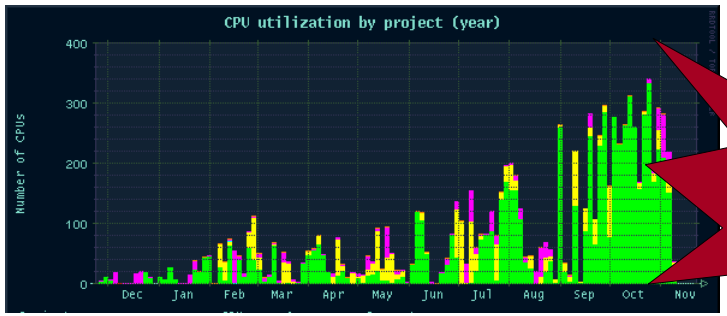




## Italian Users on CondorCaf

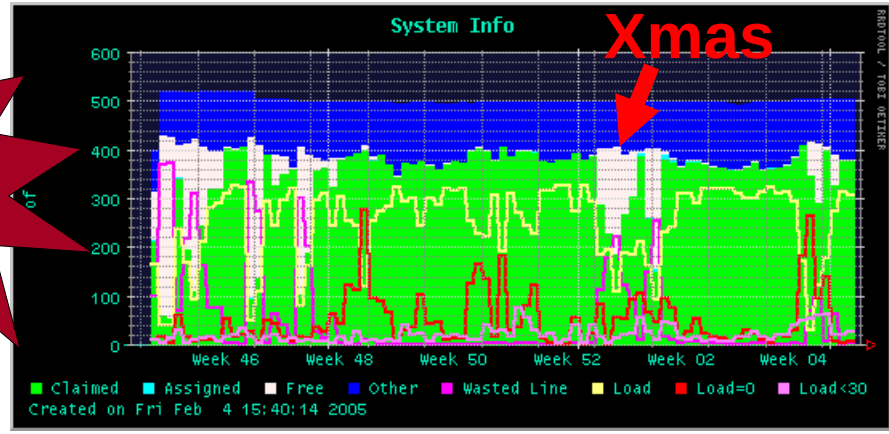
# CDF al CNAF 2004:

http://cdfmon.cr.cnaf.infn.it:8081

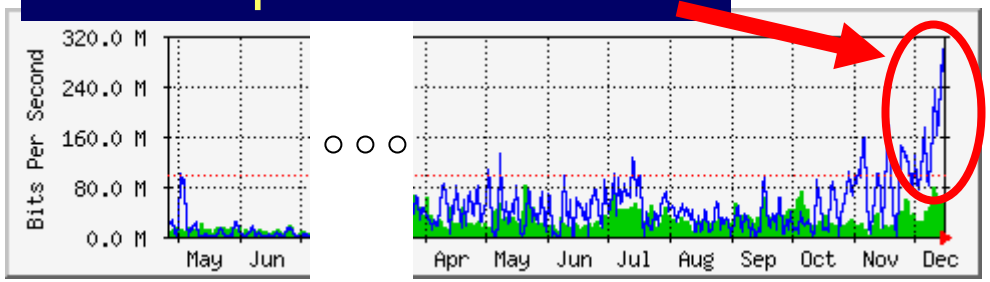
From 8 to 80 boxes. From FBSNG to Condor



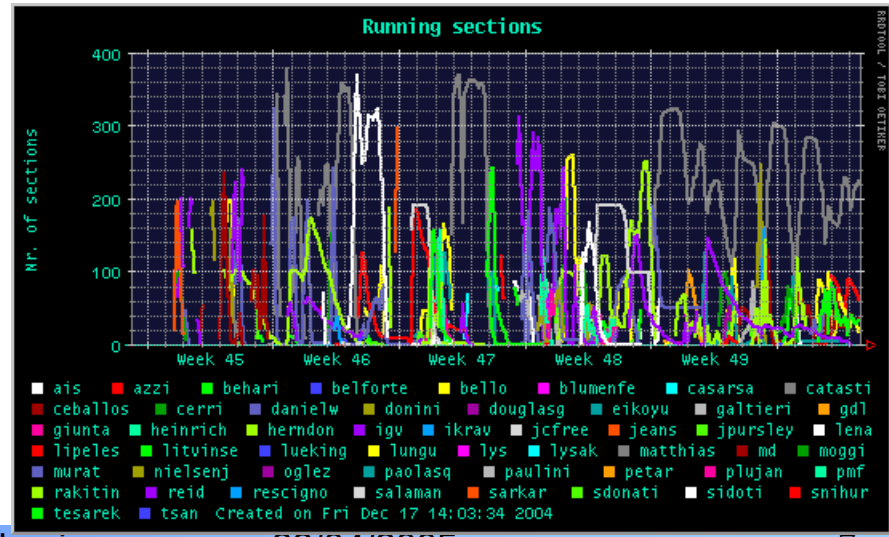

CPU Satura !




December: + 22TB disk data import ~200Mb/s




- NFS stinks, GPFS ok
- RFIO ok
- Condor great
- Glide-In in progress



# Towards GRID: Glide-in

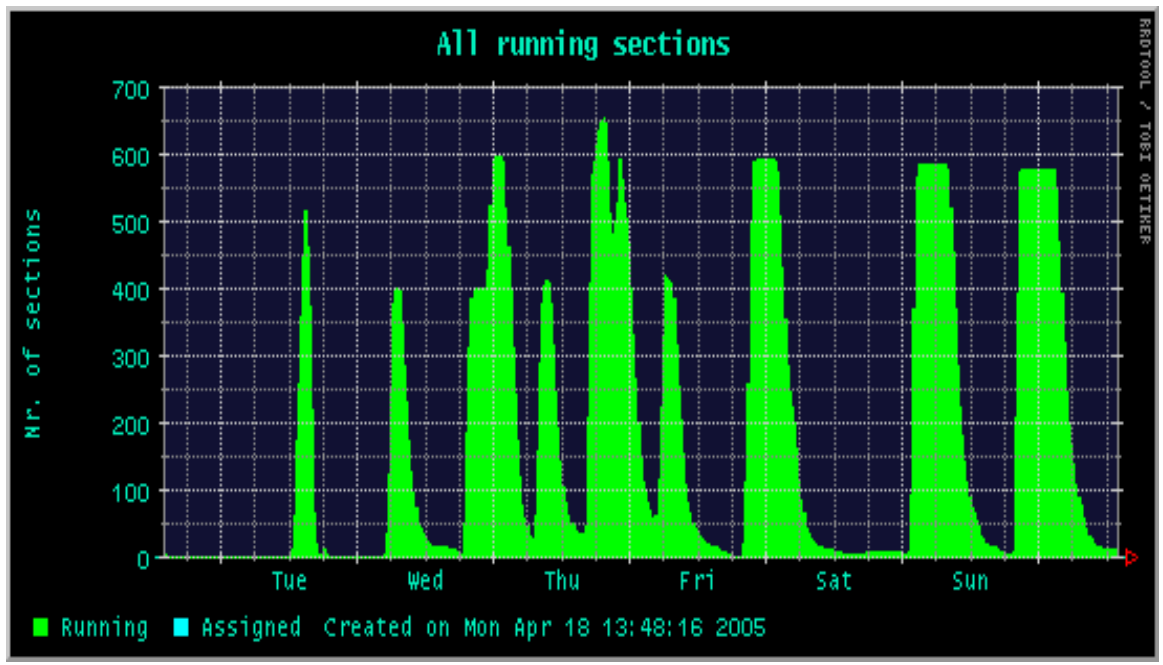
Presented by Igor Sfiligoi @IFC, April 15

- Just a small change over the production Condor based CAF
- Use glide-ins to extend the Condor pool
  - A dedicated head node must be installed and maintained near the Computing Element
  - CDF user policy in our hands
  - Use dedicated storage
- In beta testing stage now at CNAF
  - Test users have access to 500KSI2K of CPU

Work done by Subir Sarkar and Igor Sfiligoi

# Towards GRID: Glide-in

- CDF users are submitting jobs via GlideCaf to common resource @CNAF (farm LSF)
- We have access to 1.1 THz in principle
- Need to define CDF dedicated resources when LHC Data Challenge will be in



# Bottom Lines

- Everything works:
  - CAF @ FNAL and CAF @ CNAF **OK**
  - Resources @ CNAF **OK**, usage ~80%
  - CAF extension @ CNAF to "pool comune LSF" **OK**
  - LCG portal **OK**
- Users happy, analysis not limited by resources
- Impossible to define now if and when we will have crisis
- Compare CDF 2008/9 with "Tier1"=resources pledged to INFN for LCG @CNAF as RRB of LCG of April

http://lcg.web.cern.ch/LCG/C-RRB/2005-04/RRB7_Report050405.pdf

- CDF cost is low
  - 2005/2006 covered by what already approved
  - future ➔ next slides

# Disk Space

- CDF produces ~100 TB of data per year
- ~20% comes to CNAF ⇒ 20TB per year
- Same amount of disk space considered for Monte Carlo Ntuples and reduced datasets homemade
- End of 2005 CDF has 90TB (80 from CNAF)
- Small expansion: +40TB/year
- CDF is a small fraction of CNAF/Tier1 program
- Years of reference: 2008/9 CDF = 4~5% of Tier1

| Year | TB CDF | TB T1 | %CDF |
|---|---|---|---|
| 2004 | 10 | | |
| 2005 | 32 | 200 | 16% |
| 2006 | 90 | 850 | 11% |
| 2007 | 130 | 1500 | 9% |
| 2008 | 170 | 3500 | 5% |
| 2009 | 210 | 5600 | 4% |
| 2010 | 250 | 9000 | 3% |

# CPU: "Assisi" CDF requests

- Scenario presented in 2004 (3 times) based on CDF official expectations presented by Rick Snider and not changed as today
- CPU for CDF-Italy users:
  linear increase as 2004→2005
- CPU for CDF: 15% of total needs as presented
- In the reference year 2008/9 CDF ask for CPU as Atlas or CMS
- Not crazy, but difficult to make now the case
- Today proposal: we will ask so much when and if we need

# CPU: "Assisi" CDF requests

| Year | CDF (KSI2K) | T1 (kSI2K) | %CDF |
|---|---|---|---|
| 2004 | 86 | - | - |
| 2005 | 495 | 1300 | 38% |
| 2006 | 903 | 1800 | 50% |
| 2007 | 1505 | 2400 | 63% |
| 2008 | 2107 | 6300 | 33% |
| 2009 | 2709 | 10500 | 26% |
| 2010 | 3311 | 18000 | 18% |

# CPU: exercise "crescita zero"

- Assuming, as exercise, that starting from 2006 CDF does not buy any CPU we have 900 KSI2K
- In the reference year 2008/9 CDF is 12% of Tier 1 a bit less of ~1/3 of Atlas o CMS
- Not surprising, CDF will have ~10PB of data
- CDF will not be "insignificant" in any case

# CPU: exercise "crescita zero"

| Year | CDF (KSI2K) | T1 (kSI2K) | %CDF |
|---|---|---|---|
| 2004 | 86 | | |
| 2005 | 494.5 | 1300 | 38% |
| 2006 | 903 | 1800 | 50% |
| 2007 | 903 | 2400 | 38% |
| 2008 | 903 | 6300 | 14% |
| 2009 | 903 | 10500 | 9% |
| 2010 | 903 | 18000 | 5% |

# CPU: "in medio stat virtus"

- But, CDF does not disappear
- A small increase per year is needed:
  - luminosity will increase and data as well
  - to have something to count on in case GRID or FNAL are too loaded
  - give the possibility to redefine each year the increase (positive or negative)
- Assumption: increase of 300 KSI2K/year after 2006
- In the reference year 2008/9 CDF is 15% of Tier 1 to be compared with 12% if "a crescita zero"

# CPU: "in medio stat virtus"

| Year | CDF (KSI2K) | T1 (KSI2K) | %CDF |
|---|---|---|---|
| 2004 | 86 | | |
| 2005 | 494.5 | 1300 | 38% |
| 2006 | 903 | 1800 | 50% |
| 2007 | 1032 | 2400 | 43% |
| 2008 | 1161 | 6300 | 18% |
| 2009 | 1290 | 10500 | 12% |
| 2010 | 1419 | 18000 | 8% |

# Summary

- CDF-Italia has a group of people working on Computing and Data Handling which has been reorganized to cope with Stefano leave
- CAF@CNAF is highly used
- Several datasets stored @CNAF and used for analysis
- Big improvements towards GRID:
  - GlideCaf used to produce Monte Carlo
  - working hard to have LCG based CAF
- Requests made in Assisi still valid for disks and CPU
- Different scenario also presented to have possibility to discuss each year our needs based our and LCH usage

# Backup

# Backup

## Total computing requirements

| | Assumed conditions | | | | Total requirements | | | | |
|---|---|---|---|---|---|---|---|---|---|
| FY | Int L. (fb^-1) | Evts (10^9) | Peak rate (MB/s) | (Hz) | Ana (THz) | Reco (THz) | Disk (PB) | Tape I/O (GB/s) | Tape Vol (PB) |
| 03A | 0.30 | 0.6 | 20 | 80 | 1.5 | 0.5 | 0.2 | 0.2 | 0.4 |
| 04A | 0.68 | 1.1 | 20 | 80 | 2.3 | 0.7 | 0.3 | 0.5 | 1.0 |
| 05E | 1.2 | 2.4 | 35 | 220 | 7.2 | 1.4 | 0.7 | 0.9 | 2.0 |
| 06E | 2.7 | 4.7 | 60 | 360 | 16 | 1.0 | 1.2 | 1.9 | 3.3 |
| 07E | 4.4 | 7.1 | 60 | 360 | 26 | 2.8 | 1.8 | 3.0 | 4.9 |

A = actual (FNAL o E = estimated

- Presented at Run 2 review last fall and Oct., 2004 IFC meeting
- Analysis CPU, disk needs scale ~ with number of events
- Changes in logging rate in FY2005 and FY2006 drive large increases in requirements

CDF IFC Meeting, April 15, 2005 Status of CDF computing plan and budget R. Snider 15

# dCAF @CDF

| Current Resources [*] | | | |
|---|---|---|---|
| Cluster Name and Home Page | Monitoring and Direct Information Links | CPU (GHz) | Disk space (TBytes) |
| Original FNAL CAF | queues, user history, analyze, ganglia, sam station, consumption | 1000 | 370 |
| FNAL CondorCAF (Fermilab) | queues, user history, analyze, ganglia, sam station, consumption | 2200 | (shared w/CAF) |
| CNAFCAF (Bologna, Italy) | queues, user history, analyze, resources, network, sam station, datasets, consumption | 480 | 32 |
| KORCAF (KNU, Korea) | queues, user history, ganglia, sam station, datasets, consumption | 178 | 5.1 |
| ASCAF (Academia Sinica, Taiwan) | queues, user history, ganglia, sam station, datasets, consumption | 134 | 3.0 |
| SDSC CondorCAF (San Diego) | queues, user history, analyze, ganglia, sam station, datasets, consumption | 380 | 4.0 |
| HEXCAF (Rutgers) | queues, cpu, sam station, datasets, consumption | 100 | 4.0 |
| TORCAF (Toronto CDF) | queues, user history, analyze, ganglia, disk status, sam station, datasets, consumption | 576 | 10 |
| JPCAF (Tsukuba, Japan) | queues, user history, ganglia, sam station, datasets, consumption | 152 | 10 |
| CANCAF (Cantabria, Spain) | queues, user history, ganglia, sam station | 50 | 1.5 |
| MIT (Boston, USA) (MC only) | queues, user history, analyze | 322 | 3.2 |
| | *Current Totals [*]:* | 5572 | 448 |

# Assisi requests

## CDF Analysis hardware plan

| | CDF ANALYSIS NEEDS | | | 15% | | |
|---|---|---|---|---|---|---|
| Year | GHz | TB | K$ | GHz | TB | K$ |
| 2004 | 3700 | 300 | 960 | 555 | 45 | 144 |
| 2005 | 9000 | 600 | 1800 | 1350 | 90 | 270 |
| 2006 | 16500 | 1100 | 1590 | 2475 | 165 | 239 |

## Roadmap for CNAF

| | CDF FARM AT CNAF | | | | | | |
|---|---|---|---|---|---|---|---|
| | for INFN physicists | | for CDF grid | | | CNAF tot GHz for CDF | Notes |
| Year | GHz | TB | 30% of our CPU | GHz to add | GRID GHz | | |
| 2004 | 950 | 38.5 | 285 | 200 | 485 | 1150 | "for INFN" already payed |
| 2005 | 1500 | 90 | 450 | 600 | 1050 | 2100 | discuss in Assisi |
| 2006 | 2000 | 150 | 600 | 1500 | 2100 | 3500 | discuss in 2005 |