CDF needs at Tier 1

- Many details in slides for (your) future reference
- Will move faster with talk
- Not really clear what I have to say today (newcomer)
- Try to address both motivations (Comitato di Gestione) and hardware specifics (Comitato Tecnico)
- More info in my web page and links therein
  - http://www.ts.infn.it/~belforte/index_offline.html

A bit of History

- We thought about a central analysis center in Italy in '99 and rejected it
  - No INFN site volunteered
  - Worried about manpower needs
  - Worried about guaranteeing consistent software environment, data base export, network speed
  - Leave hardware and sysman to professionals, focus on data analysis

- CSN1 approved a plan to
  - Analyze 2ndary data sets at FNAL
  - Copy Ntuple to INFN sites in Italy
  - Spend up to 1M$ in computers at FNAL
    - ☞ ~150K$ assigned and spent so far

What's new

- Tier1 is coming along
- Software distribution proved to work very well
- Distributed DB access is a global CDF need and will have to be solved anyhow
- Networks grew much more then expected
- GRIDs are coming and many CDF collaborators are pursuing them
- FNAL facility will not grow to cover full needs (esp. interactive)
- Mini-farms in INFN sites will not grow beyond <10 nodes

CDF needs from Tier1

- Want to move computing of CDF Italian group from Fermilab to Italy
  - Do not bring in Italy bulk data reconstruction
  - Not a fixed share of overall CDF needs
- Common trend of other CDF collaborators
- Look ahead at next 8 years of CDF Run2 data analysis
  - Exploit fast EU network vs. slow TransAtlantic link
  - CDF moving toward GRID architecture of Linux clusters
  - CDF Italy participates in DataTAG test
- Looking forward to use GRID tools for
  - Job submission etc. now
  - Data management later

Time frame

- Tevatron Run 2a: 2001-2004
- Tevatron Run 2b: 2005-2007

- Data Analyis = continous process, from this year till after LHC starts, 2-year overlap likely
- ➔ CDF data analysis: until 2010

- 2002: Data at FNAL, copy NTUPLE to INFN sites
- 2003: transition
- 2004+: Data in Italy, copy NTUPLE to INFN sites

- Monte Carlo (largest uncertainety)
  - Large scale productions (1K cpu-days) @ FNAL
  - Small/private/fast (< 100 cpu-days) @ Tier1

Medium term (Run2a) needs for Tier1

- Limited to Run2a: 2005
- Goals:
  - Eliminate all italian hw at FNAL but "X-term" desktop
    - replicate 2ndary data sets
    - create/store 3rtiary data sets
    - create/store ntuples
    - allow interactive analysis of 3rtiary and ntuples
  - Eliminate all non-desktop from INFN sites
    - no "CDF" environment outside Tier1
    - limited PAW/ROOT only at local sites on small final samples for "publication refinement"
  - Share resources with other CDF groups, esp. Europe
    - collaboration started with UK + Spain, Germany to join later this year

Data (Orders of Magnitude) for Run2a

- 1 PB (1mary=everything) a FNAL (care of Fermilab)
- 100 TB (2ndary(part)) a FNAL (care of Fermilab)
- 10 TB (2ndary(part)) a FNAL (INFN founded)
- 10 TB (2ndary(copy)) al Tier1 (growing to replace previous)
- 1 TB (3rtiary) al Tier1
- 1 TB (Ntuple) al Tier1

- By Physics Data Set (i.e. selected process, e.g. W+jets):
  cross section = 5nb ➔ 10^7 events
  - 2nd = 1TB (100KB/event) ➔ a few DS at Tier1
  - 3rd = 100GB (select 1/10) ➔ all "interesting" DS at Tier1
  - Ntuple = 1GB (1/10 selection * 10KB/event or
    all events*1KB/event or ...) x "a few"
    x N_users ➔ all at Tier1

Overall Hardware Needs

- Data Storage (2003 + …)
  - 10TB + 10TB/year for 2ndary/3rtiary
  - 3TB + 1TB/year for interactive
- Analysis CPU
  - 10 "1GHz" CPU / TB of data (from 2001 benchmark)
- Interactive CPU/Disk
  - 2 "up-to-date" CPU / user x 40 users
  - 300GB / user x 40 users (growing with technology)
    - ☞ size this from comparison to resources available to US students at Fermilab (typical University owned PC's in offices: 5~7K$ per desk every 3 years)
- MonteCarlo CPU (Gen+Sim)
  - ~40 "up-to-date" CPU's (possible underestimate)

How/When to get started

- Work is resource-limited now (in spite of low Luminosity)
  - Need (order of): 2TB disk, 10 analysis CPU, 10 interactive CPU as of "yesterday"
- Expect some relief from new CAF at FNAL by summer
  - Slow turn on curve for effective usage
  - Saturation by starved users
  - Summer conference rush
- Better start moving some activity already
  - Learning process
  - Feed in experience from FNAL
  - Test/Learn/Develop new tools
- Release early – Release often – Listen to users
- Start with important but not critical activity

Short Term Tasks (next 12 months)

- Monte Carlo (H → tau-mu, top → 6jet)
  - Using GRID framework
  - 1 week x 20 PC every 1~2 months
  - Output = Ntuple = O(10GB) saved "at home"
    - temporary disk store, tape store welcome
- Secondary data sets (no backup, maybe no NFS (rootd))
  - Non-Grid copy to begin with
  - Goal: test tools on small scale (1/10) real exercise
    - hadronic multijet b-tagged data set for t → 6j and H → bbar
    - 200GB+50GB/month starting "now", 4-10 CPU's
  - Add larger data once "it works"
    - B→ hadrons tertiary data sets
    - 2~3 times larger

Another Short Term Task (2002)

- Interactive facility (non absolutely needed, but much welcomed), sized for 20 users :
  - Single place for code development/debug
  - Share 3rtiary data sets and ntuple
    - in general all sorts of small, unstructured data sets
  - Avoid replication, avoid resource waste
  - Resource sharing allow faster startup for new persons

- Already a significant impact with:
  - 500GB + 100GB/month user's data (RAID)
  - 4CPU + 2CPU/month
  - Interactive access (non-GRID)
  - Common home & code (10GB/user x 20 + 5GB) backed up
    - disk quotas, NFS mounts

Summary of short term desiderata

- Simulation
  - 20 nodes x 10 days every 1~2 months
  - Starting May~June (try DataTAG testbed in April)
- Data analysis
  - 200GB growing to 2TB by Christmas
  - 1 processor / 100 GB
  - Starting June~July
- Interactive
  - 500GB + 10 processors
  - Doubling by Christmas
  - Starting "now"

Would like to know user configuration

- Batch ? Which ? How ?
  - Queues, priorities, provision for privileged users
  - Automated job replica ? e.g. now at Fnal submit one exe on 100 files, automatically splitted in 100 jobs potentially running in parallel
  - Have your own to test ? Import ours from Fnal ?
- Tape storage
  - Which DB ? HSM ?
- Interactive ?
  - May like to copy from here to Fnal
  - How to deal with many competing users

- Should/could we join a "user committee/community" ?