Reqs from CDF Run II on INFN computing infrastructures

- **Usual things:**
  - system management (-DEC +Sun&SGI ++Linux)
  - ROOT
  - printers, news, mail, web… standard stuff (AFS maybe)
  - gnu software
  - maybe KAI compiler for C++
- **New stuff:**
  - network
  - network
  - network

Network usage for Italian CDF group

- Usual caveat: strategy for analysis outside Fermilab not defined yet (workshop February 12 at Fnal). Will use my guess (as usual). Everything may still change (as usual).

- Data Copy
- Remote Analysis
- Remote Monitoring
- Remote Shift
- Collaboration at distance
  - as addition to “old” codec-based videoconf (that we keep!)

CDF Data Size in Run II

- Data logging: 75 Hz for 2 years (2 inverse fb)
- Raw event size: 250 KB
- Add analysis data and data overlap: 520 KB
- Reduced Physics Analysis Data set (PAD) = 60 KB/event
    - ☞hope to get ≈ 30KB with ROOT compression
- **Overall data set for Run II: ~1 PB** = 1000 TB
- **Overall PAD size: ~200 TB**
- FNAL data storage (all data are at FNAL, at least!)
    - 30 TB disk (PADs + robot cache)
    - 1PB Emass tape robot **(~100GB/tape**)
- 200 physicists doing analysis at FNAL, about 20 at Pisa, but simple scaling don't work (can't have 3 TB disk + 100 TB robot in Pisa, data can not be parted among physicists).

Network needs for Analysis

- 3 SCENARIOs (extremes and middle way):
  - copy all PADs in Italy, need almost no net to US
  - leave all PADs & "ntuple" in US, use Xterm from Italy
  - copy some PADs here, keep most ntuples here (copy/create)
- Difficult to estimate requirements. Better the other way around. Given 4 Mbit/sec dedicated to CDF, what can we do ?
  - 4 Mb/sec = 200GByte/week = 2 tapes/week, can't beat DHL !
    - 1 tape a day = 100GByte/day = 10 Mbit/sec
    - **PADs don't travel on the net**
  - 4Mb/sec / 10 users = 1GByte/5hr/person for copying ntuples
    - one analysis = 1/10th data $\Rightarrow$ PAD=20TB, Ntuple=20GB ? refreshing Ntuple takes 4 days min ! More data, more users..
    - Converging argument: 10GB ntuple/physicist = minimum !
    - **Can't make Ntuple offsite and copy locally on the net**

Network solutions for Analysis

- Working remotely (i.e. use FNAL CPU) is not so bad:
  - 4Mb/sec / 64 users = 64kb/sec/person = run PAW at FNAL from Xterm in Italy, more then needed.
  - still need internet to copy docs, printouts etc.. Note: presently WWW to Fnal downloads at ~20 kbit/sec and is just fine.
- Will **mostly run at FNAL** and **ship** DHL **large ntuples** and small or often needed PADs (20TB = 200tapes, already handled in Run I that many Exabyte). **Copy short ntuples** (≤100Mbit) on the net (keep transfers < 1hour)
- This requires little more then present bandwidth, e.g. ~2 Mbit/sec for CDF probably OK.
- **BUT**: need to run analysis jobs at FNAL, **desperately needs better interactive response**: now 300msec per key (ping from TS, BO, PS), unacceptable ! Interactive must be priviliged. Low throughput is OK, but **response time must be shorter**.

Remote monitoring

- Offline ~ok: need guaranteed interactive work, probably OK even if heavy graphics is slow. Transfer large data sets via tapes.
- **Online: a new front**.
  - **1. Silicon Vertex Tracker**: crucial part of Level 2 trigger, a lot of online monitor, care, debugging: experts in Italy (only!)
  - **2. Internal Silicon Layers**: our largest contribution to new detector, want to keep ownership after installation, needs on-line assistance from Pisa as well.
  - **3. SVXII, MUON, CALORIMETER TOF…** many more crucial detectors with large/predominant competence in Italy
  - **4. Remote Control Room**: do shifts in Italy, save on travel
  - 4 is about saving money, 1, 2 and 3 about getting proper credit and physics rewards from present large efforts: keep ownership of our detectors !
  - **NEED: few X-terminals that work as extensions of FNAL LAN**

Remote cooperation

- Not only X-Term. Need to replace physical contact
- Want to be able to collaborate (sit here next to me and help me with your program) with a collegue at Fnal. Avoid moving the full group back and forth. Keep presence on-site tied to work at home.
  - one-to-one videoconference (not what we used so far)
  - PC-style
  - very little demand from video
  - whiteboard and/or window sharing a must
- also useful for remote shifts
- need support from local system managers, central guidelines on equipment/software etc. then… like before:
  - **need X-terminal that works as extension of FNAL LAN**

Extending FNAL LAN

- Needed for online monitoring and debugging and remote control room
- Do not need full 1Mb/sec Ethernet, need more tests but probably 64 Kbit/sec is good enough (128Kb/sec is fine)
- Do not need it at all times (problem in FNAL, phone call to expert in PISA, get the line for a few hours)
- It has to be a guaranteed connection, something to rely on, if it fails a few times, the collaboration will request expert on site
- Can "you" do it ? Can I have a dedicated slice of the net kept aside for just my terminal ?
- Do we have to buy ISDN by the minute as for videoconference ?
- Need agreement on direction to move.

Is PTT ISDN the way ?

- **ISDN works**
  - physicists in Batavia work at home happily with FNAL calling back on commercial ISDN, 64k upgraded to 128k automatically when usage > 48Kb/sec (seldom unless heavy web graphics). The lab also provides the networking equipment at home.
- **ISDN exists**
  - just on Italian Telecom Phone Book, can have it everywhere
- **ISDN is end-to-end**
  - gurantees bandwidth all the way from one room in Italy to CDF control room LAN, no matter what.
- **ISDN is (relatively) cheap**
  - 64kb/sec = 55 Klit/hr (27 Euro)
  - 5hr of remote work cost like 1 day in Fermilab
    + save the flight money and time (and jet-lag)

ISDN implications

- Need work, money, and infrastructure
- **Work**: find best configuration
  - ISDN card on PC + call to US ? Easy
  - Router + Call-Back from US ? Cheaper. More workplaces...
  - Is one line/site enough ? 1 in Trieste, 3 in Pisa ? 10 in Italy ? Get more “air time” for the same cost. Time sharing...
- **Money**: 20 users * 1hr/day average = 1MLit/day ~ 400MLit/year
  - probably OK if taken as saving on travel
  - still a significant cost. What is the INFNET cost ?
  - needed connection time could easily be 2 or 3 times more
  - needed connection number could be less if can share one line.
- **Infrastructure**: dedicated room(s) and PC(s) (unix WS?). Maybe router. Equipment at Fermilab as well ? A few tens of Mlit plus some effort and a lot of manpower.

The “normal” way

- Can we get ISDN-like quality at lower cost using QoS on INFNET ?
- **When** ? How ? Who ?
- **Simpler** (for the user !), even if same cost as PPT ISDN.
- Do we have to **pay** for it, like with Telecom ?
  - That’s allright: easier to avoid waste and to control usage
- Not just bandwidth for CDF, nor for INFN site, need end-to-end service: **bandwidth for user** !
- Must be **available ~immediately upon request** to be usefull to keep expert in Italy for needs in CDF Control Room
- **Our “schedule”**
  - Already would help (also across Italy !!)
  - Very helpful after summer 1999 (SVT installation)
  - Definitely needed starting in 2000
  - usage and need will grow with time (and data)

Conclusions

- **CDF RunII** means **Unix**, mostly on **PC**, then **Sun** and/or **SGI**, DEC may be dropped on the way.
- Mostly no special software needed but ROOT
- **will use more m-bone like one-to-one videoconferencing**
- **Network to FNAL must be improved**
  - **better interactive response**
  - higher bandwidth for data transfer (few Mbit/sec)
  - **good bandwidth on demand (QOS) for LAN-like connection**:
    - remote monitoring **-** remote shifts **-** remote collaboration
    - ☞10KBytes/sec ~ 64 kbit up to 128 kb/s
    - ☞commercial ISDN on PPT an option
    - ☞cutting it out of internet better, but soon (2000 is next year).
    - ☞Start ISDN then go on private line as usage increases in 2001 and later ? RunII lasts till 2005 at least ...