

Calcolo per CDF-2

Revisione annuale 2003

- Evoluzione piano di calcolo CDF (10min)
- Stato e Piano CDF-italy (10min)
- Richieste (referees)

Calcolo 2003/4: bottom lines

● 2003

- abbiamo analizzato i dati!
 - ☞ risorse riservate a FNAL ci hanno salvati
- MC: fatto tanto, ma ne dovremo fare molto di piu'
- CNAF: espansione finanziata a maggio non c'e ancora
 - ☞ ma e' stato importante per lavoro $W \rightarrow$ plug, dati e MC

● 2004

- b-physics diventata item centrale, bisogni "capiti"
- i dati aumentano, i bisogni aumentano (molto + di L)
- il budget di FNAL continua ad essere insufficiente, malgrado $O(1M\$/year)$ per analisi
- MC "scacciato" da FNAL \rightarrow Canada + ? + ognuno per se'
- **L'anno del CNAF** (ahead of schedule, no resource replica)

Piu' fisica Piu' **B**ella, grazie SVT !

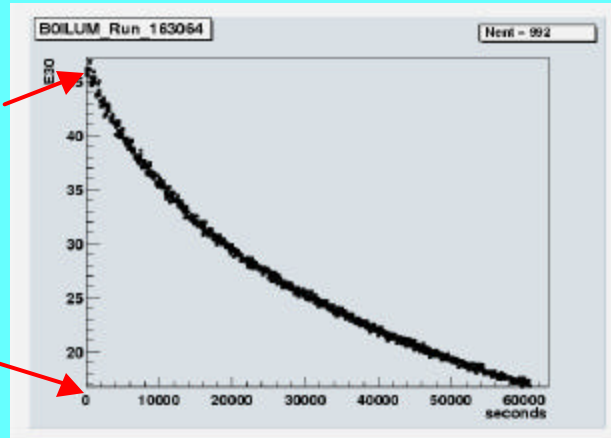
- **Fisica del B inserita alla pari nel piano di calcolo**
 - tre anni fa dicevamo: pianifichiamo per High-Pt, il B ? 100TB di PADs sono troppi, qualche Santo ci salvera'
- **Tutti i piani erano basati su estrapolazioni Run1**
 - Non avevamo idea di come affrontare la fisica del beauty
 - Non pensavamo nemmeno di poter fare fisica del charm
- **Abbiamo rivisto il piano di calcolo alla luce dell'esperienza**
 - Come previsto, il B e' difficile (tantissimi dati)
 - Servono risorse, intelligenza ed organizzazione
 - ☞ raccogliere un grandissimo campione e finalizzare offline le selezioni canale per canale, ognuno efficiente al 1~3%
 - ☞ appena possibile aggiungere singoli canali al trigger con tagli ad hoc per aumentare efficienza
 - ☞ coordinare il lavoro di analisi sul Mega-sample inclusivo

Perche' i dati da analizzare dipendono poco da L (DPS: Dynamic PreScaling)

1. Luminosity changes by a factor 3 in a 16 hour run

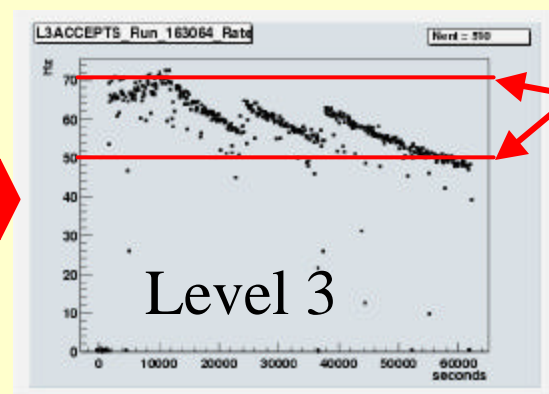
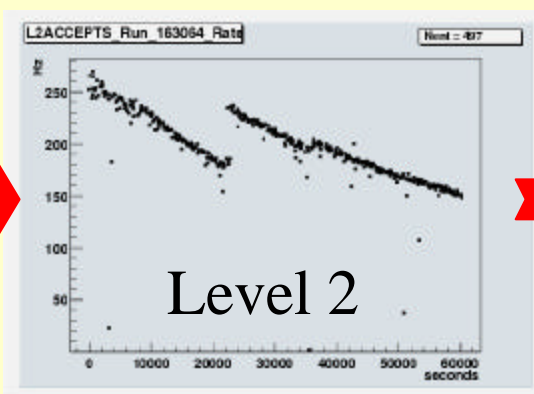
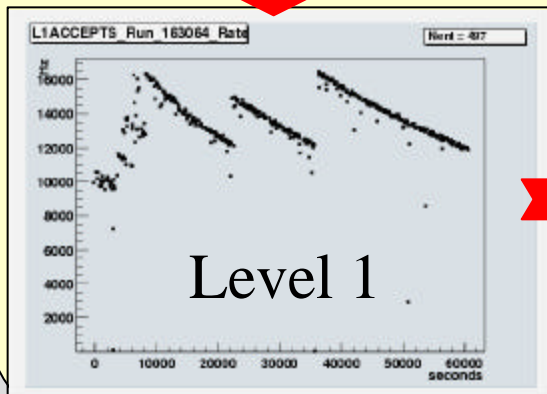
4.5 E31

1.5 E31



2. Triggers at Level 1 are automatically prescaled

3. Rate to tape stays in [50,70] Hz at all times



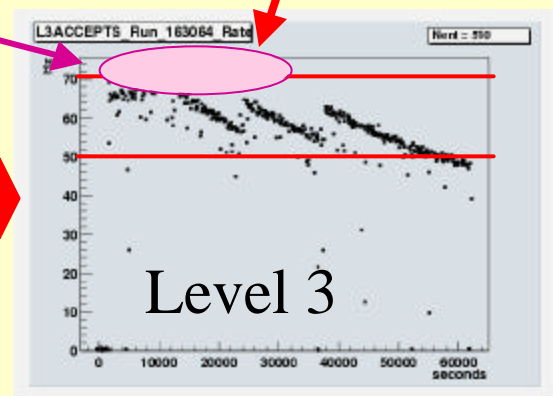
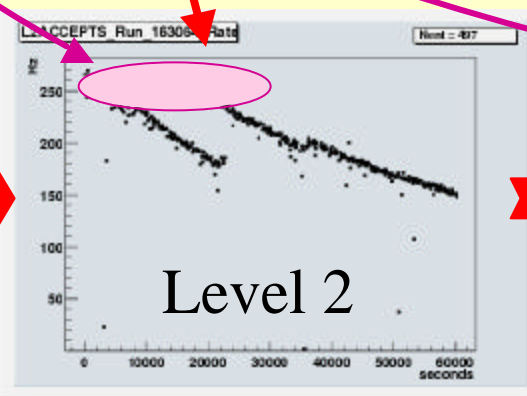
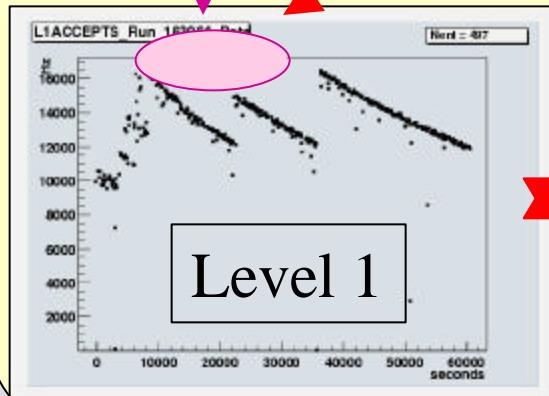
Anzi: eventi da analizzare cresceranno ancora piu' velocemente del tempo di run !

Sono triggers "buoni" quando si prescala si perde fisica !
Il Bs "sta qua dentro"

Bisogna rimediare !

Trigger upgrades
(talk di Paola)

Tape logger upgrade
Implicazioni per calcolo



DAQ upgrade: piu' dati su nastro 80 → 360 Hz in 3 anni

- Invece di cavalcare stancamente un raddoppio di statistica annuale regalato dal tevatron (sfumato), **incremento della efficienza di raccolta dati**
- Vitale per fisica del B, ma anche per la statistica dei campioni di calibrazione
 - Nel Run1 sia M_W che M_{top} sono state **limitate dallo errore sistematico**, non dalla statistica
 - ☞ Jet energy scale e risoluzione: low pt jets, $Z \rightarrow b\bar{b}$
 - ☞ B-tagging systematics: inclusive heavy flavour jets
 - ☞ EM cal fine tuning: low pt leptons and gamma's
- 2003: Compressione dati: 80→120Hz
- 2005: tape logger upgrade :120 → 240 Hz
- 2006: tape logger upgrade: 240 → 360 Hz

Evoluzione piano di calcolo di CDF

Gli esami non finiscono mai

- 1. 1997:** stime basate su estrapolazione Run1 (big SMP's)
 - Uccise da OO e SVT (10M\$, spesi tutti entro il 2001)
- 2. 2001:** 1 autore – 14 pagine
 - parte di old-CAF review, nessuna stima di costo
 - needs assessment basato su high Pt datasets
 - $O(1000)$ CPU + $O(100)$ TB per analisi $2fb^{-1}$
- 3. 2002:** 10 autori – 27 pagine
 - richiesti 2M\$/anno a Fnal
 - MC ancora un'incognita
 - basato su "dati scalano con L" (vero solo per high Pt)
- 4. 2003:** 24+ autori – 67 pagine
 - richiesti 3M\$/anno a Fnal

Viva il nuovo piano, lunga vita al nuovo piano

http://www.ts.infn.it/~belforte/offline/2004/cdf6640_computing_plan.ps.gz

■ Il nuovo documento:

CDF Plan and Budget for Computing in Run 2:
Second Annual Edition

- **Incorpora DAQ upgrades**
- **Incorpora una componente dei dati che scala con il tempo di run, non con la luminosita'**

Contributions from

Stefano Belforte, Lucio Cerrito, Glenn Cooper, Matthew Herndon, Jason Harrington, Chris Hays, Beate Heinemann, Robert Wagner, Elizabeth Sexton Kennedy, Robert Kennedy, Arthur Kreymer, Ashutosh Kotwal, Dmitri Litvintsev, Victoria Martin, Mario Martinez-Perez, Mark Neubauer, Manfred Paulini, Rick Snider, Rick St. Denis, David Tang, Al Thomas, Andreas Warburton, David Waters, Steve Wolbers, and the CDF Institutional Computing Representatives Board.

- **Incorpora esperienza Conferenze invernali 2003**

necessary to meet the physics goals over the next 3 fiscal years.

- **Ha una sezione su computing remoto e sulla via di CDF alla GRID (INFN-Grid non partecipa, peccato)**

| | | |
|----------|---|-----------|
| 9 | Computing Facilities Outside Fermilab | 47 |
| 9.1 | History and Perspective | 47 |
| 9.2 | Status of available off site resource | 49 |
| 9.3 | MC Production | 51 |
| 9.4 | Incorporating off site resources: toward a CDF Grid | 51 |
| 9.4.1 | The Vision | 51 |
| 9.4.2 | The Tools | 52 |
| 9.4.3 | The Financial Side | 53 |
| 9.5 | Offsite Productivity Issues | 53 |

CDF GRID: do it, see if it works, see if people use it, see how effective is it, decide on \$

- **Unico "impegno" per ora dal CANADA: 1 milione di eventi MC al giorno, e responsabilita' coordinamento MC production**

❖ **Lavoro per IFC**



Conclusions



- λ Computing requirements will scale with the size of the run 2 dataset.
 - └ Increased data over next 3 years will require ~10 times more computing.
 - └ Moore's law should prevent the cost from exploding.

- λ Computing procurements required to meet CDF needs
 - └ \$2M in FY04: driven by increased CPU to analyze extra events logged.
 - └ \$3M in FY05: from an additional doubling in data logging (CSL upgrade).
 - └ \$3M in FY06: from 50% increase in data logging (DAQ upgrade).
 - └ Additional operating expenses of roughly \$0.4 M per FY.

- λ Budget weighted towards analysis CPU
 - └ ~ 40% analysis CPU
 - └ ~ 20% tape drives
 - └ ~ 15% disk
 - └ ~ 10% networking
 - └ ~ 10% reconstruction farm CPU
 - └ ~ 5% miscellaneous

Reviewers: "Good plan"
APPROVED

Robert M. Harris, Fermilab CD/CDF

17

FNAL Plan, 2003 edition

non sono rose !!!

- **Fare a Fermilab solo il minimo :**

- FCC solo per produzione ed analisi. **No MC**
- Reprocessing: 20% degli eventi (30% in Run1)
- **Analisi (>50% del budget)**
 - ☞ high-Pt "come prima" (5nb=1TB/1day x 200 user)
 - ☞ B "per quello che si puo'": 15 user (su 60) / 25 day

- Cio' nonostante la spesa rimane > 2M\$/y "guidance"

| Anno fiscale | FY01 | FY02 | FY03 | FY04 | FY05 | FY06 |
|--------------|---------|------|------|------|------|------|
| M\$ chiesti | 2 | 2 | 2.1 | 1.9 | 3.0 | 2.9 |
| M\$ avuti | 1 (+1→) | 2+1 | 1.4 | | | |

- FNAL fa la sua parte ! Ma non basta

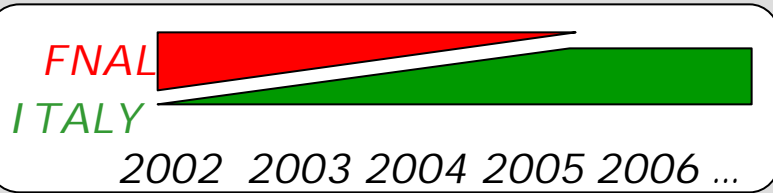
- MC "organizzato" → Canada + ?
- MC "single user" → ognun per se'
- Tagli al budget ? More reprocessing ? → Tagli alla analisi

Il nostro (Italia) piano non e' cambiato !

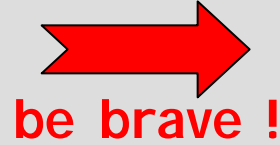
L'analisi si regge su **3 gambe**

- 1) Produzione campioni skimmati/ntuple a partire dai DST O(10-100TB): **FNAL** (tier 0+1 ?)
- 2) MC ed analisi ripetute di campioni ridotti O(1-10TB): **CNAF** (tier 2+3 ?)
- 3) Sviluppo codice (edit, compile, link, run, debug) ed analisi interattiva Paw/Root O(100GB) : **desktop e/o clusters di sezione** (o desktop "cicciuto" a FNAL) (tier 4 ?)
 - desktops da dotazioni rinnovati ogni 5 anni, in media hanno 2.5 anni di eta' (3 anni = vecchio per analisi); cpu/disco su desktop inefficiente per gruppi grossi con rapido turnover
 - piccoli clusters fondamentali (per ora)

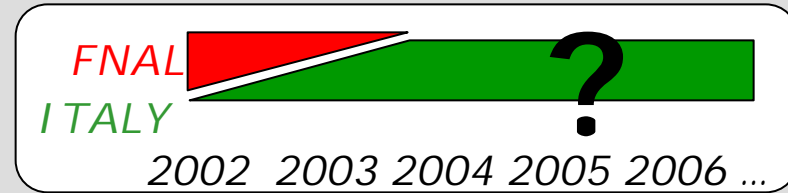
Le farms: the Big Plan



24 Giugno 2002



be brave !



24 Settembre 2003

● 2004: l'anno del CNAF (> 1 anno prima !)

➤ Non compriamo altro a FNAL

☞ nel 2005 FNAL spegnerà le ns. macchine ?

➤ Rendiamo CNAF ~ quota I NFN a FNAL

➤ Vediamo come va

➤ **Scelta coraggiosa**: farm CNAF a oggi : 10 cpu, 1TB, 3 utenti, ruolo marginale nell'analisi

➤ **Scelta sicura per I NFN**: cpu non usata → altri

CNAF per tutti

- **La nostra promessa: la CPU che non useremo (if any) e' disponibile per altri, per questo stiamo cambiando il software !**
- Cpu non-usata → per CDF MC (da "ora") = contributo a CDF che possiamo monetizzare a differenza di CPU non-usate a FNAL che "va persa"
- **Fbsng → Condor** (Igor Sfiligoi e Bologna's Condor team)
 - ☞ possibile **accesso via Condor/Grid a tutto INFN**
 - ☞ **CPU non usate → a chi ne ha bisogno**
- Macchine non usate: CDF ← → LHC in 1 giorno
- **Investimento al CNAF sono soldi "sicuri" per INFN**
- Con un po' di fatica anche disco puo' essere reso "dinamico"

Richieste farm CNAF 2004

- **Richieste nei moduli:**

- **23 TB**

- **70 duals**

- **Rationals – DISK**

- **Arrivare ad almeno 30 TB** (dettagli nelle slides di backup)

- 10 TB high pt

- 15 TB tertiary data set B

- Speriamo di gestirne una buona parte come cache dinamica

- 5TB User's space (200GB x 25 users)

- N.b. 100GB/user = standard FNAL

- **Rationals – CPU**

- A fine 2003 48 duals x 2.2GHz = 211GHz

- Equivalenti a 30 duals @ 3.5GHz

- 2004 ➔ **arrivare a 100 duals @ 3.5 GHz = 700GHz**

- 300 per analisi + 400 per MC (dettagli nelle slides di backup)

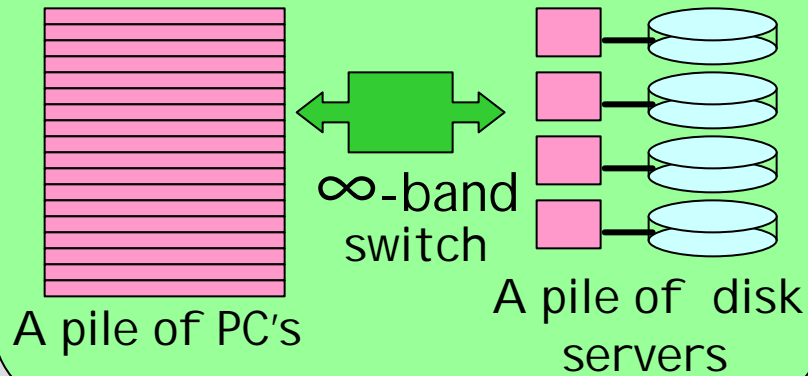
Piano Temporale

- Fine 2003
 - In opera quanto acquistato quest'anno
 - La prima vera farm usabile per le conferenze invernali
- Gennaio/Febbraio 2004
 - Ordine per 50% della richiesta 2004
 - Bisogna coprire i bisogni 2004, non 2005 !
 - In funzione a Maggio, per le conferenze estive
- Giugno 2004
 - Ordine per il restante 50%
 - Sfruttare il miglioramento prezzo/performance
 - In funzione a Novembre per le conferenze invernali
-

Don't starve the farm !

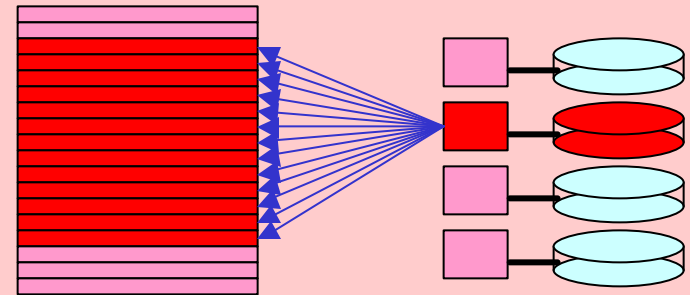
Una farm di analisi non e' una pila di dischi e PC

the plan, easy and clean



the trouble: Cpu's can't get data

Lo user seleziona 1TB di dati, 100 processi cercano di accederli



I RIMEDI

- fast access: no NFS
- data striping and replica
- access throttling
- Plan for (soft) failure
 - more time OK (cpu fails)
 - no result BAD (data lost)
- fault tolerance (**every piece will fail**, e.g. Raid controller)

LA SOLUZIONE FINALE

- Enstore (tapes)
- dCache (enstore cache, disk access, load shaping)
- SAM (data catalog, requests reordering, minimize tape ops)
- massimo disco (2.1kEu/TB)
- A server fail ? Use another

Accesso dati al CNAF

(il maggior ? per una farm di analisi)

- Un chiaro sforzo per avere hardware fault tolerant
 - per ora performance non un parametro della architettura
 - potrebbe richiedere un diverso rapporto disco/server
 - potrebbe spingere per soluzioni di tipo/costo diverso
- Data access per i workers: do it yourself. Will GRID ever help ?
- Il nostro piano (will prevent farm starving for data)
 - SAM + hand hack to copy data to local disk + fcp (fnal requests queueing tool), not elegant, not optimal, but will work
 - No clear solution for fault tolerance (yet)
 - ☞ SAM could restore data from Fnal, slow ... SAM + Castor ?
- Il secondo ? : accesso al database (latency, load on fnal server)
 - Wait and see. SQL replica of Oracle getting closer
- **CONCLUSIONE:** motivi di ottimismo, ma potrebbe non essere così efficiente come la nostra quota nella CAF@FNAL

L'interattivo : fatte le ntuple sulle farms, I plots per Phys.Rev.

- Le sezioni
- I desktop a FNAL (eta' media 2.5 anni)
- Inevitabilmente...
 - tante piccole richieste ($\Sigma \approx 60\text{kEu}$)
 - la gioia dei referees
- ❖ Come evitare il MicroManagement ?
- ❖ "Tutti" devono fare istogrammi con Root, unificare ?

Il software

- Ci diamo tanto da fare per
 - far funzionare le farms
 - aumentarne l'efficienza
 - avere voce in capitolo in cosa fa FNAL
 - entrare in "Grid"

Attività software "farms" 2003

- User's monitor (Lazzizzera/Amerio - Trento + Pinazza- Bologna + Casarsa/Belforte - Trieste)
 - Monitoraggio operazione batch system
- I CAF e CAF-mailer (Sfiligoi - LNF)
 - Gestione disco locale e report jobs sulla CAF
- Condor (Sfiligoi - LNF + Mazzanti/Semeria - Bologna)
 - Transizione batch system FBSNG → Condor
- SAM development (Roberto Rossin - Padova)
 - Agevolazione transizione DFC→SAM
- CAF at CNAF (Sidoti - Pisa + Belforte - Trieste)
 - Installazione, gestione, supporto utenti
- SAM test and tune (Belforte/Casarsa - Trieste)
 - Installazione in Italia, configurazione, test, documentazione
- I/O monitor (Armando Fella - Pisa)
 - Monitor I/O nella CAF: wait for tape/cache, MC o Ana etc.

Software plans

- Keep up the good work (slide precedente)
- Interactive Grid (Condor+SAM+CAF+Proof) :
UCSD(CAF designer) + MIT(Proof developer) + INFN (I. Sfiligoi)
 - Analisi di O(100GB) in “minuti”
 - Non siamo gli unici a volerlo fare, ma tra i pochi ad avere esperienza di una farm di analisi >>100 nodi con >>100 users
 - Progetto ambizioso, grandi speranze, grande visibilita' e ricaduta. Prototipo fine 2003, se va se ne riparla
- CNAF/I talia: come Maggio + bbftp (24Mbit/sec Fnal→Ts)
 - Contiamo ancora di farcela con:
 - ☞ system management da personale Tier1
 - ☞ una posizione al CNAF dedicata al supporto sw (dal 2004)
 - ☞ sviluppo, setup up, cose complicate: CDF (6 nomi)

Conclusioni

- Abbiamo finalmente capito perche' i bisogni di calcolo non sono calati malgrado la luminosita' sia bassa
- C'e' un nuovo piano (sempre il migliore)
- Abbiamo capito dove faremo il MonteCarlo (fuori FNAL)
 - ☞ dobbiamo ancora definire esattamente quanto
- Potranno esserci difficolta' per la analisi a FNAL
 - ☞ confidiamo nel CNAF
- Full speed su farm analisi e MC a CNAF. Attenti a:
 - Performance ed architettura disco→cpu
 - Supporto sistemistico: 24x7 ← → 7:12x5
 - ☞ hang/crash di nodi e gateway... hw ? Kernel ?
 - Nostro software ancora in evoluzione e da installare
- Costi ben inferiori a quanto preventivato 2 anni fa
 - accordo verbale a C.Gandolfo: 1M\$ a FNAL per Run2a (3 anni)
- Spese spostate dagli USA all'Italia

The Plan (as presented 24 Jun 02 in CSN1) evolving... and evolving...

| year | Luminosity | | ANALYSIS FARM | | | contingency 40% (Keuro) | Requested per year (Keuro) | Assign |
|---|----------------------|-------------------|---------------|-------------|----------------|-------------------------|----------------------------|-------------|
| | Planned (Church) | Target (adjusted) | disk (TB) | CPU (duals) | cost/y (Keuro) | | | |
| 2001 | commissioning | | 0.6 | 0 | | | 43 | 43 |
| 2002 | 0.3 | 1.0 | 20 | 80 | 336 | 0 | 336 | 269 |
| 2003 | 1.2 | 2.0 | 40 | 140 | 266 | 106 | 372 | 264 |
| 2004 | 2.5 | 3.5 | 70 | 200 | 285 | 114 | 399 | 290 ? |
| TOTAL cost Analysis Farm at FNAL + 40% conting. for Run2a expanded to 3.5 fb-1 | | | | | | | | 1150 |
| 2005 | 4.1 | 6.0 | 110 | 280 | 331 | 132 | 463 | |
| 2006 | 7.6 | 9.5 | 180 | 350 | 298 | 119 | 417 | |
| 2007 | 11.3 | 13.5 | 250 | 430 | 331 | 132 | 463 | |
| 2008 | 15 | 18.0 | 330 | 500 | 288 | 115 | 403 | |
| TOTAL cost for Analysis Farm at CNAF + 40% conting. for Run2b (15 fb-1) | | | | | | | | 1746 |
| TOTAL BUDGET CENTRALIZED COMPUTING FOR ANALYSIS 2001-2008 | | | | | | | | 2896 |



866 ?

❖ Only analysis farm. ~~No MC~~. No interactive.

- 2004: expand farm at CNAF but not FNAL
 - ▶ No resource duplication.
- Staying well within initial estimate and CSN1 "agreement":
 - ▶ 1Meuro until 2004
- Discuss 2005-... in 2004 (Jun or Sep)

SPARE SLIDES

Nulla di nuovo dopo questa

● Se la vita ti da limoni...

- Luminosita' ristagna
- Upgrade silicio non si fa
- Higgs sta sfumando
- Bs yields << aspettative (2fb-1 non bastano)
- Lum sale, DAQ dead time sale (gia' al 5%) → prescalare e/o alzare le soglie

● Fatti una limonata!

- No shutdown lunghi, prendiamo dati, abbassiamo le soglie, aumentiamo i rates, spremiamoci l'ingegno, rimbocchiamoci le maniche !!!
 - ☞ SVT: 4/5, faster timing, RoadWarrior, AM upgrade
 - ☞ 3-track trigger, SDPS, CSL upgrade, Pulsar (→ plug e/mu in SVT)
 - ☞ review trigger table: hadronic B effic. 1% → 2~3~? %

Piu' fisica Piu' **B**ella

- Fisica del B inserita a pieno titolo nel piano di calcolo
 - tre anni fa dicevamo: pianifichiamo per High-Pt. Per il B ? 100TB di PADs sono troppi, qualche Santo ci salvera'
- Tutti i piani erano basati su estrapolazioni Run1
 - Non avevamo idea di come affrontare la fisica del beauty
 - Non pensavamo nemmeno di poter fare fisica del charm
- Il successo di SVT e' stato travolgente
 - CDF e' un attore di primo piano nel panorama del B
 - I gruppi di analisi di B e c sono affollati, dinamici, esuberanti
 - Sull'onda della competenza sul trigger abbiamo una **posizione di leadership anche sulla analisi che vogliamo mantenere**
- E non dimentichiamo l'alto Pt:
 - $Z \rightarrow b\bar{b}$ $t \rightarrow 6\text{jet}$ $W \rightarrow \text{plug}$: tutti canali "nostri"

Sociologia: CDF = collaborazione competitiva

- A CDF i talk a conferenza e le posizioni di responsabilita' non si assegnano col manuale Cencelli:
 - Chiunque lavora su quello che vuole, come puo' e gli pare
 - **Ognun per se' e vinca il migliore**
 - Le presentazioni ai meeting di analisi sono libere e le fa il lavoratore, non il capo
 - Chi fa bene, e' premiato con lodi, visibilita' ed altro lavoro da fare
- Fermilab e' lontano, ha un clima orrendo, si lavora e basta
 - **I ns. Giovani fanno tanta fatica**
 - Non guadagnano soldi o posti
- **Devono almeno avere la soddisfazione di fare fisica**
 - Che e' il motivo per il quale fanno CDF in primo luogo
 - **Chiediamo alla CSN di metterli in condizione di competere e vincere le difficili sfide dell'analisi**

Sinopsi delle nostre CAF (Fnal e Bologna)

nero=certo **rosso=richieste da approvare**

| | FNAL | | | CNAF | | |
|-------------------|-------|-------|------|-------|-----|------|
| | Cpu | | Disk | Cpu | | disk |
| | duals | GHz | TB | Duals | GHz | TB |
| 2003 (FNAL owned) | 179 | 580 | 94 | | | |
| 2003 (INFN owned) | 62 | 232 | 14 | 48 | 211 | 7 |
| 2003 total | 290 | 930 | 164 | | | |
| 2004 (FNAL owned) | 338 | 1280 | 184 | | | |
| 2004 (INFN owned) | 102 | 470 | 24 | 118 | 700 | 30 |
| 2004 total | ~500 | ~3000 | >200 | | | |
| 2005 (FNAL owned) | 674 | 3700 | 288 | | | |
| 2005 (INFN owned) | 102 | 470 | 24 | ? | ? | ? |
| 2005 (total) | >800 | >4000 | >300 | | | |

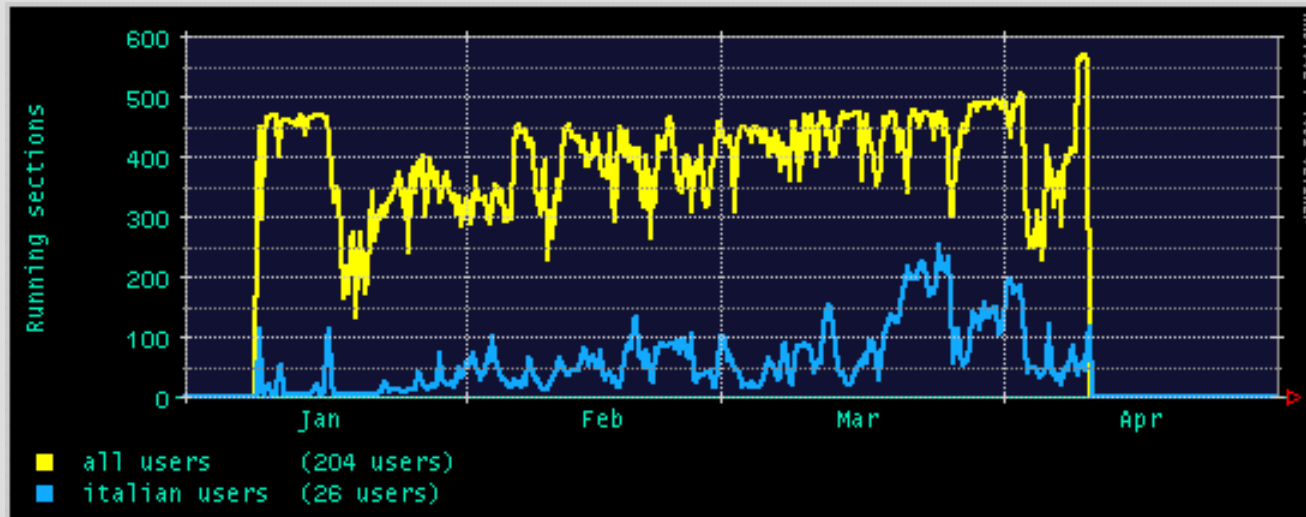
2003 at CNAF = attivo da autunno, 2004 = da ~estate

FNAL total = FNAL owned + INFN owned + UK+Japan+Germany+...

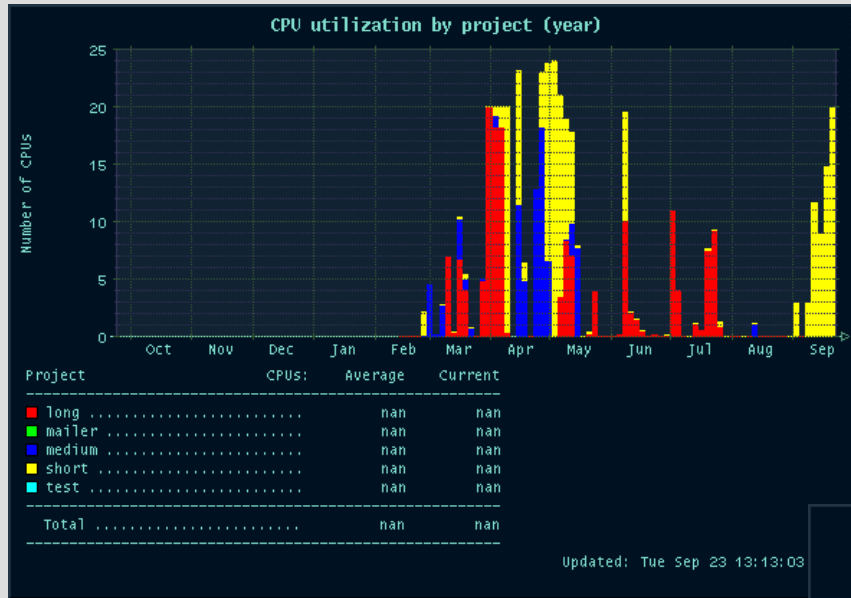
2004 al CNAF = se richiesta nei moduli (290KE) e' approvata interamente

CAF a Fermilab

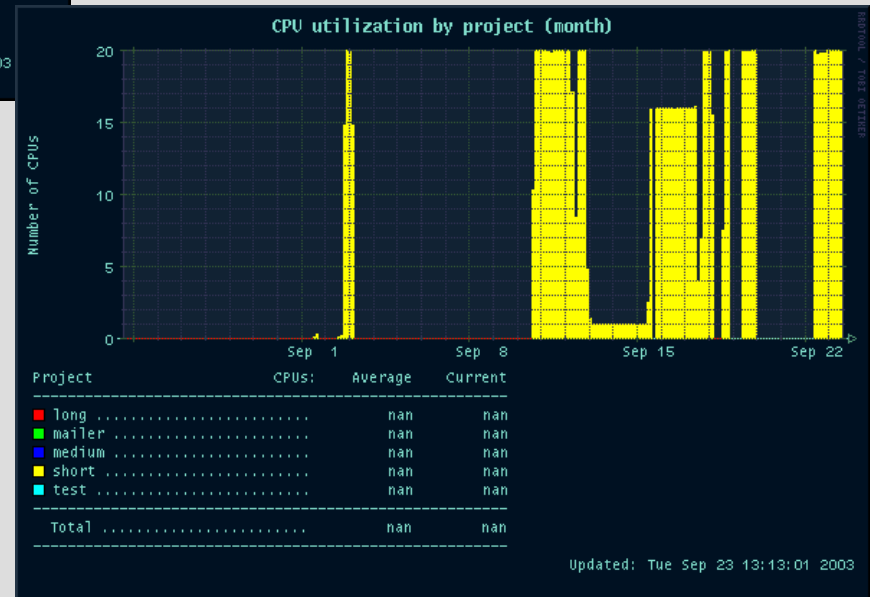
- File servers INFN pieni al 70%, aggiornamento quotidiano su <http://www.ts.infn.it/~belforte/offline/caf-disks.html>
- Uso delle CPU non di facile documentazione (avuti problemi col monitor), da una ispezione ~quotidiano (sb) la nostra quota e' occupata circa la meta' del tempo.
- L'importanza della coda riservata non e' stata tanto nella media (non e' una produzione MC "long distance") quanto nel permettere lo "spunto di picco" necessario per i blessings



CAF at CNAF



- a primavera $W \rightarrow \text{plug}$
- settembre inizio attività di simulazione Bs
- non e' così liscia



- ❑ dopo la prova si "parte" a meta' settembre
- ❑ doveva essere tutto giallo
- ❑ I gaps sono crash dei worker e soprattutto lo head node, soprattutto il weekend

CSL (i.e. tape logger) upgrade

20MByte/s (now) → 40 (2005) → 60 (2006)

- L'upgrade che ha un'impatto immediato su calcolo
- Eventi al secondo su nastro e' quello che conta per l'analisi
 - Ora 80Hz: 1/3 = low Pt stream: 80% puro, 50% da B
 - ☞ e tutti quei B persi per il prescalo..
 - ☞ e a CDF serve fattore 2~4 per competere con B-factories e fare Bs mixing
 - Autunno: 120Hz (compressione software raw data)
 - 2005: 240Hz
 - Studio dettagliato di come dividere la b/w e ridefinire trigger table in corso
- CDF Computing Plan (i.e. budget req. to FNAL) mostra che i bisogni non diminuiscono, >= 2M\$/anno
 - high-Pt comunque e' quella del "Run2a"
 - maggior appetito per campioni di calibrazione (Jets)

CDF Policy

- MC deve essere fatto fuori, 2 componenti
 - MC “organizzato”: MOU,MOF... (e.g. Canada 10^6 ev/day)
 - MC “user level”: ognun per se'
 - ☞ sviluppo e test su piccola scala di MC che poi si fanno centrali, toy MC's, MC di canali specifici, MC che aiutano (apparentemente) solo una analisi di interesse di poche persone
- Reprocessing, se serve piu' della frazione indicata: deve essere fatto fuori
 - MOU, MOF, GRI D... per ora solo pie speranze
- Sviluppare tools per rendere questo possibile
 - la cdf-grid
 - Presto (2005 ?) parte di “LA GRI D”

B group computing needs: MC

- Efficiency of hadronic triggers ~1%
 - Do some preselection, but cannot improve much (~x2)
- CPU time: 6s/ev for Pythia evts, 0.5s/ev Bgenerator on 2.5 GHz
- MC generated&used by Italians (on CAF+Glasgow):
 - B->hh 50M BGEN+fullsim, runbyrun (parametrize for physics results) (CAF)
 - LambdaB->ph 2.5M BGEN+fullsim (CAF)
 - bbar->X Pythia 10M (CAF/Glasgow)
 - Charm 3M Bgen 5M Pythia
 - B->D*X channels: 12M / 6M simulati
 - D*->D0pi 3 modes 100M/50M simulati
 - Bs-> mixing (3 modes) 6M bgen (segnale) = 63M (fondo+sistematica) simulati
 - B+->phiK+ 2M Bgen
 - J/psi K 6M
 - J/psi phi 2M
 - Bs->phi phi 2M
 - phiK* 4M
- Total 260MBgen+15M pythia = $2.2 \cdot 10^8 \text{s} = \underline{\underline{2500 \text{ CPU days}}}$
- Need at least 10x next year (more data, more precision)

B group computing needs: MC

- 2004: Data x3-x7 ->more quantitative -> go to physics papers ->more channel-specific MC
- Also, plan HUGE generic bb MC to understand backgrounds, other side tagging, validate Bgenerator
 - Efficiency <0.3%
 - Need 10^8 - 10^9 Pythia events-> 10^9 s= 10^5 CPU days (1K CPU x 3m)
 - Expect to do within CDF, but still need pilot productions !
 - ☞ and likley will need to add to "Toronto"
- CDF policy: move MC out of CAF, precious for data
 - We need it for our analyses anyway !
 - CAF designed for **15** B-users,spinning through every **25** days
- Will need CPU at home to make the MC we need, whether standalone, or in coordination with the CDF effort
- Will need CPU and disk at home to supplement limited data analysis capacity on FNAL's CAF

Dettagli 1: 15TB per B

- Dai primi 100pb⁻¹: 7TB di DST nel canale B inclusivo → SKIM!
- Esempio: una analisi su **Bs** (Donatella Lucchesi)
 - Bs → Ds pi → phi pi pi seleziona ~2% del totale
 - Bs → D0pi seleziona ~2%
 - B0 → Dpi → K pi pi seleziona ~9%
 - Totale circa 1TB da importare in Italia
- 100pb⁻¹ → 400, DAQ rate x2: ~**8TB** solo per questa analisi
- Almeno un altro canale di Bs verra' aggiunto
- Altro esempio: **B→3body** (Massimo Casarsa)
 - selezione accetta ~5% del totale
 - ☞ 1% solo per il canale esclusivo B→kkk
 - ~**3TB** aggiuntivi
- B→pipi (Diego Tonelli, Simone Donati)
 - Pochi dati, Tanta calibrazione (particle ID)
- Piu' il MC, le ntuple, un po' di campioni di calibrazione...

Dettagli 2: 10TB per high-Pt

- Dai primi 150pb^{-1} (tutto quello che c'è con SVX finora)
- $Z \rightarrow b\bar{b}$ (Tommaso Dorigo)
 - 3Mevents finora
 - Da una settimana, nuovo trigger: $20 \rightarrow 45\text{nb}$
 - Altri 250pb^{-1} con nuovo trigger = altri 10M events
 - DST+ntuple $\sim 300\text{KB/event} \rightarrow 4\text{TB}$
 - Finora 0.5TB di MC generato (1.8Mevents)
 - Serve almeno x3 MC
 - Totale $\sim 6\text{TB}$
- $\text{Top} \rightarrow 6\text{jet}$ (Ambra Gresele) $1.2\text{TB}/100\text{pb}^{-1} \rightarrow \sim 4\text{TB}$
- $W \rightarrow \text{plug}$ (Antonio Sidoti)
 - 0.8TB usati attualmente su CNAF
 - Totale per $400\text{pb}^{-1} \sim 2.5\text{TB}$
- Siamo già a 12.5, poi, jet calibrations, top mass..., $H \rightarrow b\bar{b}$..

Dettagli 3: 700GHz

- E' una speranza
- Se 3.5 troppo costoso (vogliamo comprare subito!) → si compra 3GHz e totale e' 631GHz invece di 700
- Analisi:
 - Good old rule: 1GHz/100GB (1 pass/day @ 10Hz)
 - ☞ 30TB → 300GHz
- Monte Carlo
 - Bgen: 25K CPU-days, Pythia 100K (?) CPU-days
 - ☞ 50K CPU-days @ 2.5GHz = 120K GHz-day
 - 400GHz x 1year (300 days) = 120K GHz-day
 - Siamo nel "right ball park"
 - N.B. 400GHz sono 16Hz di eventi con Pythia ~ 1M/day
- Efficienza non e' 100%, ci sono bisogni di picco, ... tutto sommato chiediamo quasi poco

La terza gamba: clusters nelle sezioni

- Richieste in linea con filosofia comune, globalmente di importo in linea con le aspettative, dettagli da discutere coi responsabili locali delle singole sezioni: **totale 62 KEuro**
- **BO: 0** OK per ora
- **LNF: 5 KE** progetto FBSNG→Condor(+Proof) (Igor Sfiligoi)
 - 2 server (mini-CAF) per sviluppo sw (scratch/reboot...)
- **PD: 7 KE** rafforzamento CPU per analisi locale
 - 2 server + rack + switch per collegare a macchina esistente
 - 4x700 Compaq del 2001 fornisce solo 2.8 GHz
- **PI: 27 KE** vedi slide ad hoc
- **RM: 20.5 KE** potenziamento per interattivo ed analisi
 - Un nuovo dual a FNAL per interattivo (3.5K)
 - 1 server + 2TB a Roma (4+13K)
- **UD: 1 KE** poco disco
- **TS: 1.5 KE** sostituzione 1 PC a FNAL
 - 2 acquistati nel 1999 (P2 400MHz), uno si e' rotto a Luglio

Interattivo a FNAL (i duals da tavolo)

- La situazione e' sempre la solita
- Pool interattivo usando nodi CAF, sperabilmente in beta test da Novembre/Dicembre, noi non siamo coinvolti. Se va
 - Spostiamo li' i dischi FC (2TB) da fcdfsigi2
 - Spostiamo li' ~10 nodi CAF piu' vecchi (2x2GHz)
- Long term solution: Interactive Grid basata su Proof (vedi dopo), ancora lontana
- Desktop a FNAL: ottimo accesso a dati (ntuple e.g.) su CAF servers e/o fcdfsigi2 via rootd, utili anche da Italia
- Dual CPU desktop = miglior GHz/\$ rispetto a due desktop
- Stima dell'anno scorso: 10 sistemi come tetto, ancora valida
- Al momento ce ne sono 4 installati
- Si chiede di portarli a 5 (1 richiesto da Roma)

Situazione desktops FNAL (svecchiamo ?)

- Vedi :
<http://www.ts.infn.it/~belforte/offline/pc-at-fnal.html>

| Anno di acquisto | 1999 | 2000 | 2001 | 2002 | 2003 | TOT | di cui duals |
|------------------|------|------|------|------|------|-----|--------------|
| Bologna | | | 1 | 1 | | 2 | |
| LNF | | | | 2 | | 2 | |
| Padova | | | 4 | 1 | 1 | 6 | 2 |
| Pisa | | 4 | 4 | 1 | 1 | 11 | 1 |
| Roma | | 1 | 4 | 2 | 1 | 8 | 1 |
| Udine | | | 1 | 1 | | 2 | |
| Trieste | 2 | | | | | 2 | |
| TOTALE | 2 | 5 | 14 | 8 | 3 | 33 | 4 |

- Eta' media nel 2004: 2.7 anni

Sviluppo Software per CAF

- FNAL:
 - Abbiamo investito molto in passato sul monitor CAF (creazione TS/BO), ma non siamo riusciti a rimanerne on top (carenza umana), ci riproviamo per la CAF a Bologna
 - Interactive Grid (Condor+SAM+CAF+Proof): Igor Sfiligoi
Ottimo progetto, grandi speranze, grande visibilita' e ricaduta. Prototipo fine 2003, se OK collaborazione con INFN-Grid ?
- CNAF/Italia: come Maggio + bbftp (24Mbit/sec fnal→ts)
 - Siamo indietro (lavorato su SVT). Serve (ce la faremo):
 - ☞ aggiornare CAF software
 - ☞ provare altre configurazioni di SAM (gridFtp, multiple stagers, 2 levels of local cache)
 - ☞ installare SAM al CNAF
 - ☞ spazio disco utenti al CNAF (icaf)



Setting the Scale

1/15 evts/sec for MC

1/3 evts/sec for reconstruction

~5 evts/sec for "typical" user analysis

~300 evts/sec for complex ntuple analysis

→ Organizational issues can have large impact on CPU power needed, or quality of service perceived by user.