Computing Overview

From Requirements, to STAR-Trek (or is it ?), to the present ...

Jérôme Lauret, Collaboration Meeting, July 2002

Computing Requirements

- Done every year, resources are assessed. Assumed :
  25 weeks of running
  60% machine Duty Factor
  ½ DAQ 100 @ 20 Hz i.e. a total number of events ~ 172 M (average pp/AuAu made at the time extrapolated from Year I)

  Came up with some alarming conclusions

  Required resource well beyond projected requirements (x2.5) while funding remains the same ...

Requirements consequences ...

Tape storage within this scenario comes to 650 k$
... not a big deal ...

Priority made to reconstruction time (within a year)

Required disk storage fail short by a factor of 10

BTW : How do we achieve DAQ100 (or whatever rate it will be) with last year 40 MB/sec data sinking ??

Parenthesis (DAQ100) ...

- **2001**
  - Used only 1 Gbit line last year
  - Data rates
    ~ 40 MB/sec with local buffer
    Large variation on short time scale
    HPSS caching not great (migration slow)

- **2002/2003**
  - 2 lines available
  - Can we do better ?

Parenthesis (DAQ100) ...

Called for an HPSS upgrade plan 2 month ago ...

- New drives 9940B coming our way 60 -> 180 Mb/sec
- New servers as mover machine 80 MB/sec per controller possible (shooting for 200 MB/sec agregate)
- We can use (theoretically) performant RAID for Caching ...
- Requires yet another HPSS software upgrade

▶ Starting around mid-August/September ...

You will probably hear me scream & curse ...

Infrastructure (now)

- ~ 20 TB of disk space
  Just enough for our Common MuDst with one production pass only...

- 124 nodes for reconstruction

- 80 nodes for users (49 interractive)
  And users are suffering ...

Infrastructure (easy part)

- Scaling the CAS & CRS nodes (x 2)

  Some work to do on the CRS side

  Current HPSS resources short by 50%

  Doubling it = tripling what we have (200 K to find or we have to leave with what we have)

  CRS Software overloaded

  More Servers to manage transaction (Tony Chan/RCF)

  More I/O to the centralized disk ...

  Do we believe it will ever become stable ??

  (the answer is in next page's background)

Infrastructure (the shock)

- Cannot scale the disk space indefinitly ...

  Costly servers (don't have ... the money), requires more backbone support (currently overloaded).

Only asked for 10 more TB (50 % increase)
So what ?? Shall we start to panic !!

The (not quite) final frontier ...

- Answer may rely on distributed disk approach. Huu !!??

  ... jobs goes where the files are ...

  Several pieces (almost) in place

  Scheduler
  DRM
  FileCatalog

  ... budget shaped to reach this goal

The GRID, going down the rabbit hole ?

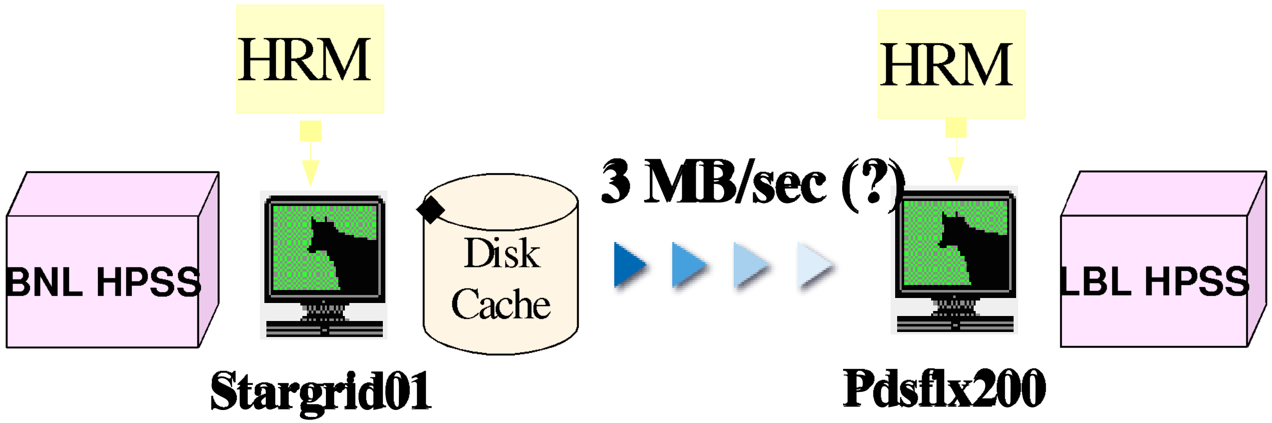
- Several tools exists already designed for computing intense tasks and large scale experiments with spread resources ...

- Ultimate Goal ??
  User submits jobs
  Results come back ... (so what ??)

  ... except that you were running in China, or maybe somewhere else (who knows) but it apparently worked.
  Basic requirement : have the data, the code, the db 'there'

Other non SCI-FI items ...

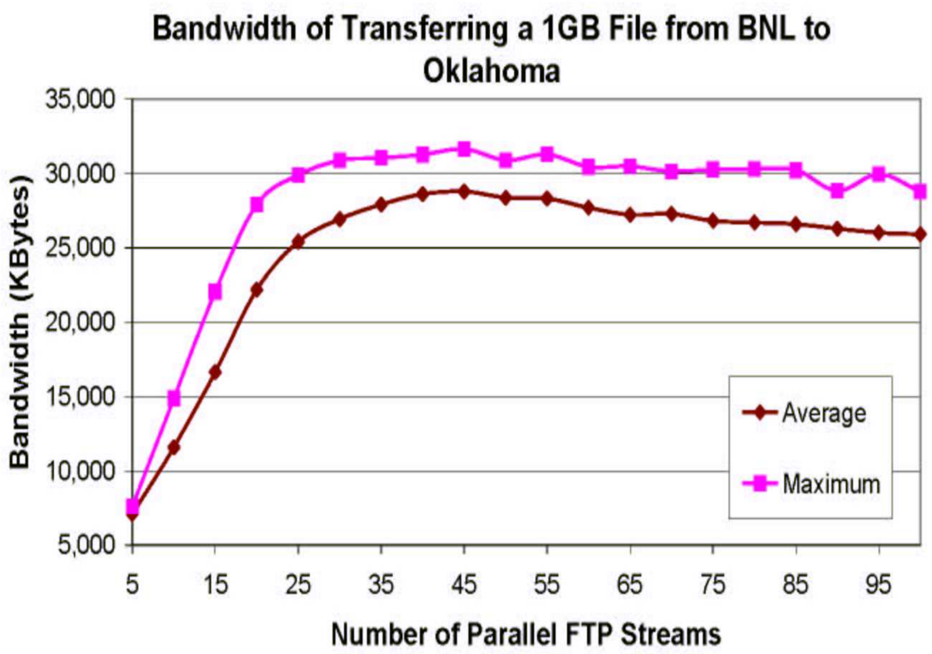
- Transfer of files using GRID tools actually happening
  - HRM running between BNL and LBL
    - Tool installed and tried out in January
    - Data transfer to PDSF is (almost) automated
      - Eric Hjort, Alex Sim, Arie Shoshani, Dantong Yu, ...



Other non SCI-FI items ...

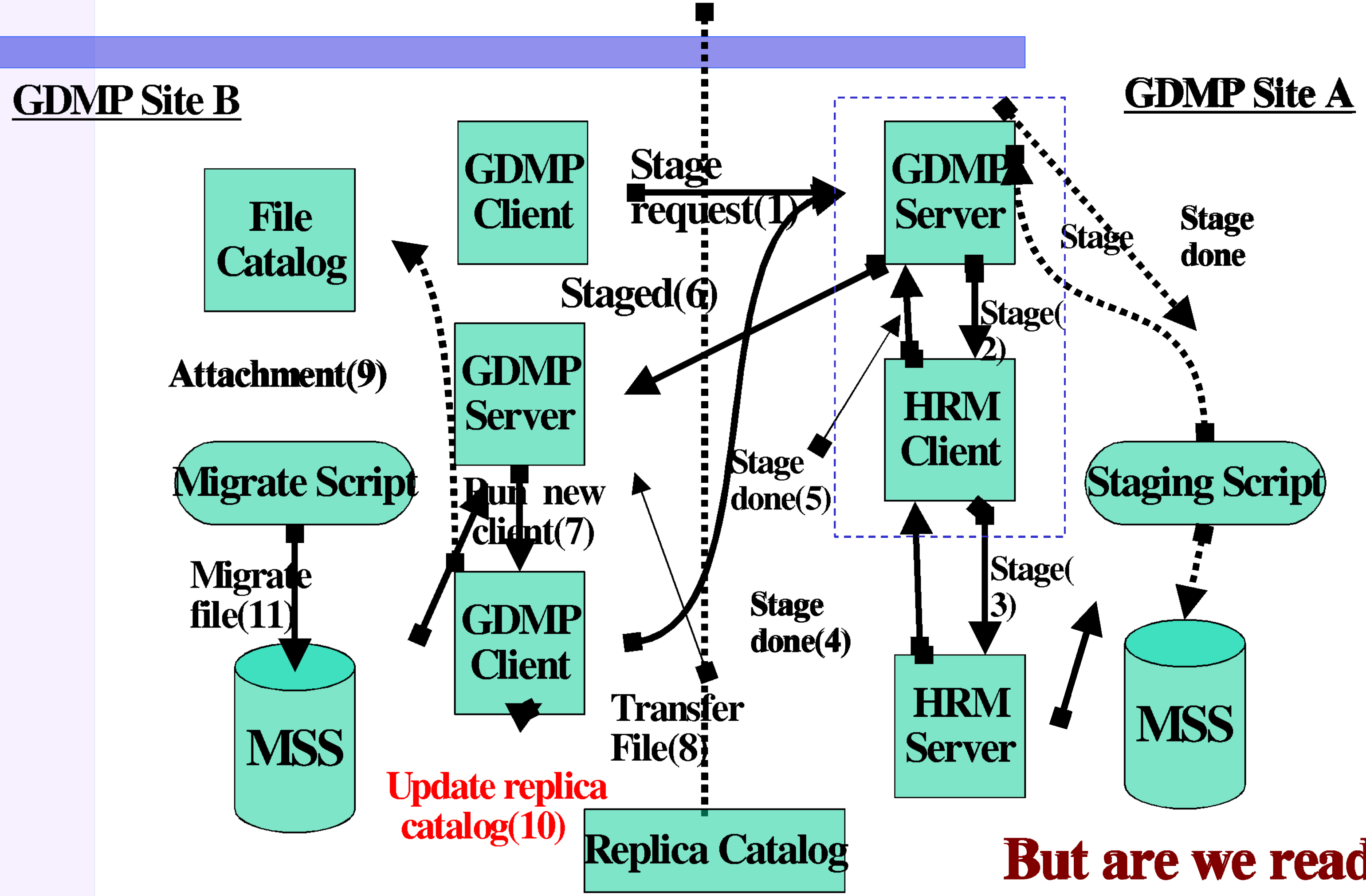
- Network allows it now : OC12 (70 MB/sec)

We are currently limited by disk I/O



Test performed by Dantong Yu/RCF

The final frontier ...



But are we ready ???

Grid & MySQL

- All of our reconstruction software uses DDB
- ... and we currently have 6 MySQL mirror servers
  3 @ BNL, PDSF, Sao Paulo, WSU
- Number to double in 2003

Needed a tool to monitor all mirrors (admin)

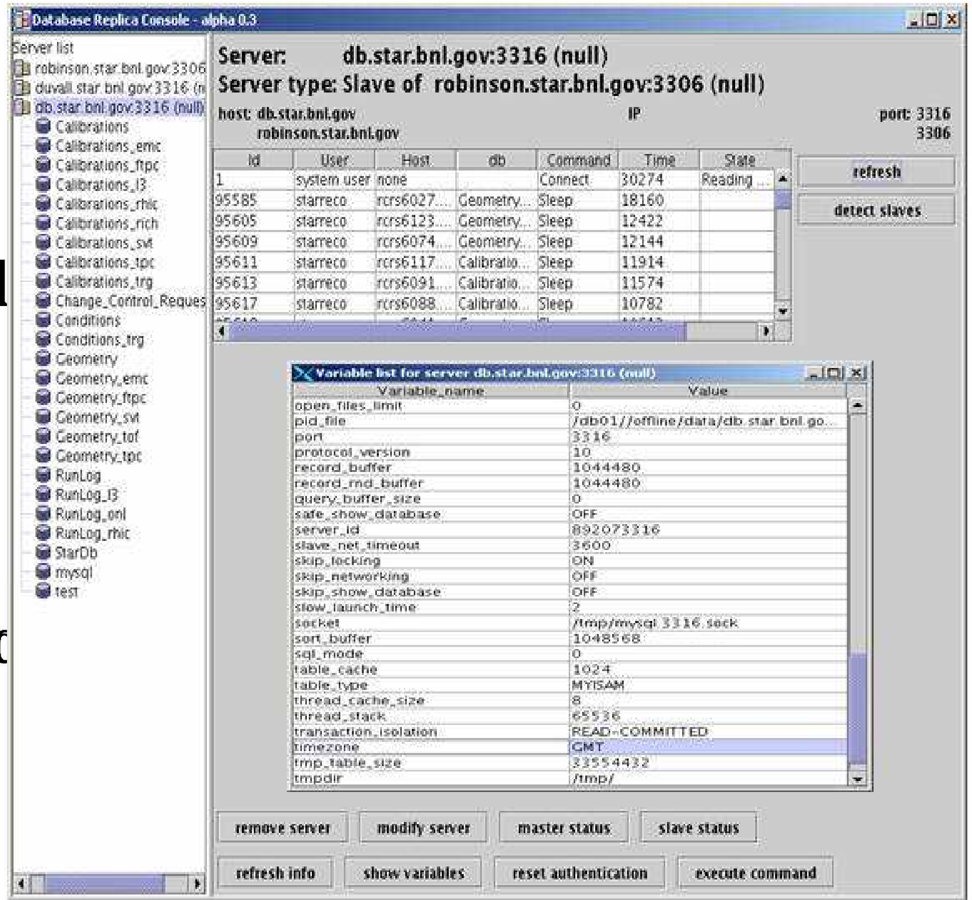
Increases data availability

In order to successfully execute jobs in a GRID environment, we need database replication

Grid & MySQL

- **Tool exists**
  Next generation will allow to install a mirror using GRID technology.

- **Why GRID ?**
  Single authenticatio
  Security



Any other ways ?

- Resource sharing will greatly help
  Cope with the sad reality increasing <-> same money
  Integretting GRID technologies (easiest with homegeneous environment)
  3 of the RHIC experiments + Atlas joined . Yeahhh !!!

- Investigating other Compilers (reconstruction)
  ICC compile code reported to be 40% faster than gcc/egcs
  CAN'T IGNORE THAT !!

- Code improvement : ITTF (next slides)

Back to earth:what's done, what's new?

- Production Status (*Lidia Didenko*)
  Lot's of problems Farm-wise (real heavy load stressed every components to limits but we managed somehow)
  Problems in analysis ? Means we need to re-think our QA strategy ...

- Embedding QA Coordinator on board (Olga Barannikova)

- BNL Computer Support person (Wayne Betts)
  Will be involved in online cluster setup,...

Software Progress ...

- Lots of progress with several sub-systems
  EMC
  (*Suaide*)
  FTPC
  (*Putschke*)
  SVT
  (*Takahashi*)

  Required a few worthwhile side production passes ...

Software ? Remember the shroom'

- Common MuDst
  Seems like the MuDst file came right on time ...
  We would have been in trouble without it !!
  But we need to push it further, and shrink it down (time for cleaning) (*Frank Laue*)
  Soon (this coming year that is), the 2 weeks promised event.root lifetime will become a hard reallity ...

IO SEC. 100 METERS

- Can we do the same with Geant ??
  StMiniMcEvent (*Mr de la ...*)

Online, what's on its way ...

- L0, L2 & L3 triggers usage, accounting & cataloguing needed.

- Calibration & Monitoring
  - TPC Pad level gains via Pulsers for DAQ100
  - Automated calibration pass ?? (drift velocity, gas-gain)
  - ...

- A new Electronic log book is being developped (BNL Java course). This means, we may need one more shift : The paper logbook burning patrol

Reco, what's coming ?

- **DAQ100 (or is it DAQ60 ?) i.e. Online cluster finder**
  Has been tested offline (Several passes made at S&C meetings) **(*Dan Magestro*)**

Need feedback and final word from PWG

Reco, what's coming ?

- ITTF is coming to maturity (see *Claude & Claude*'s talk)
  Called for the project review second half of September
  Lots of activities from within and outside ITTF
  - V0/Xi finder (*Julien Faivre*)
  - Kink (Lee Barndby)
  - Vertex Finder (*David Hardtke*)

Reco, what's coming ?

- **My 2 cents ITTF success condition (grossly un-subliminal message)**

We need PARTICIPATION & FEEDBACK FROM PWG

Same story than for DAQ100 ...
Either we are serious about it, or we aren't !!...

Beta-Testers exists, always room for more ...
Lacking representative from several PWG

Conclusions ...

... Only two

There is a BIG challenge coming ahead ...

There are several drastic code changes coming our way
We need PARTICIPATION & FEEDBACK FROM PWG