STAR Computing overview ...

Jérôme LAURET
jlauret@bnl.gov
STAR Regional Meeting – Dubna (Russia)
November 2003

Overview

- STAR Computing overview ... (What we stand for, Organization)
- How to get an account, help, etc ...
- Resource, usage and Monitoring

- Coming run data dilemma

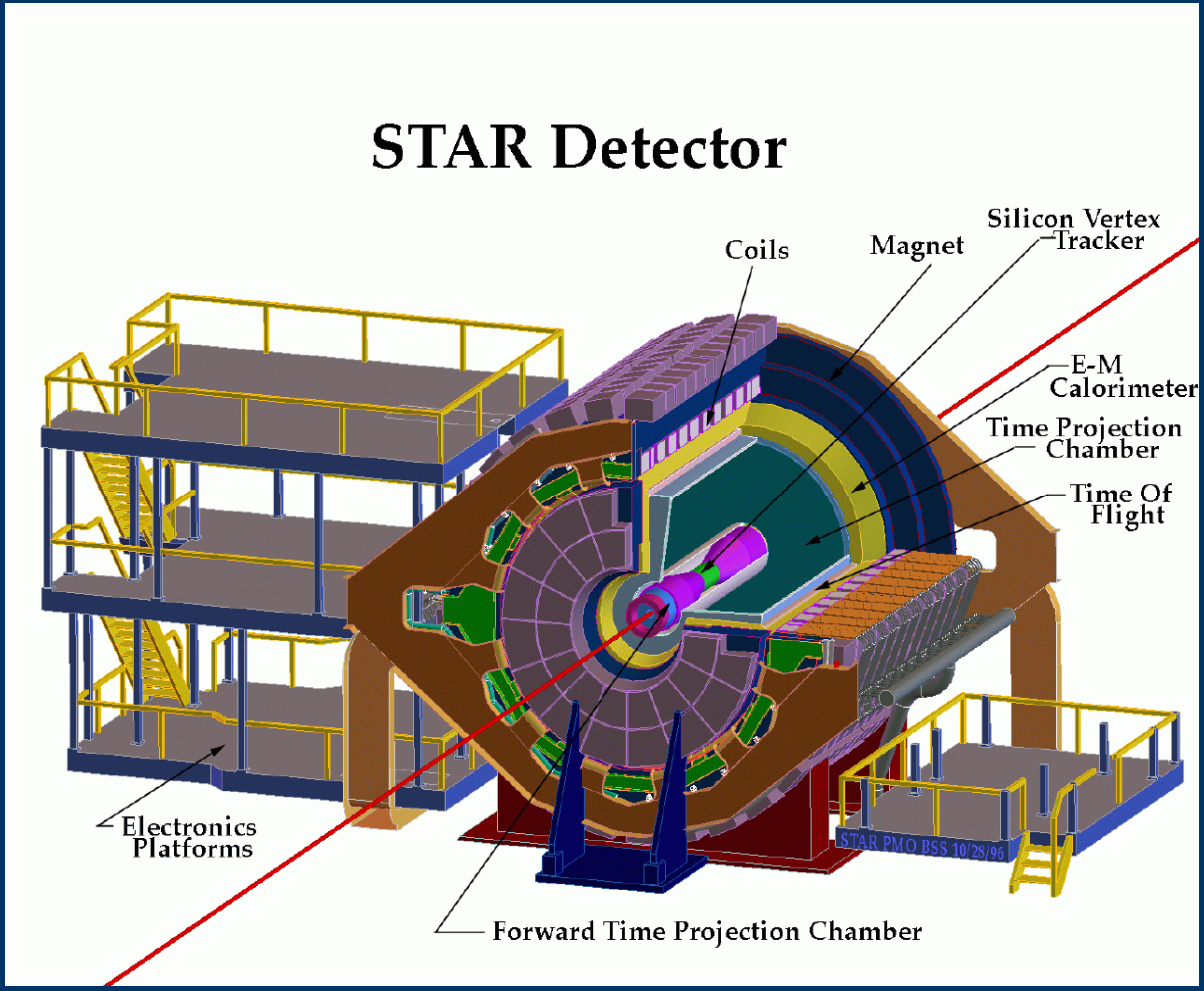
- Distributed Computing
- What is Grid about ?? (and what STAR is doing)
- How does it makes better science ??

Intro ...

137,000 pads ~ 70 M pixel
If Zero suppressed
~ 10 M pixels

Event size comparable to an image taken by a Digital Camera ...

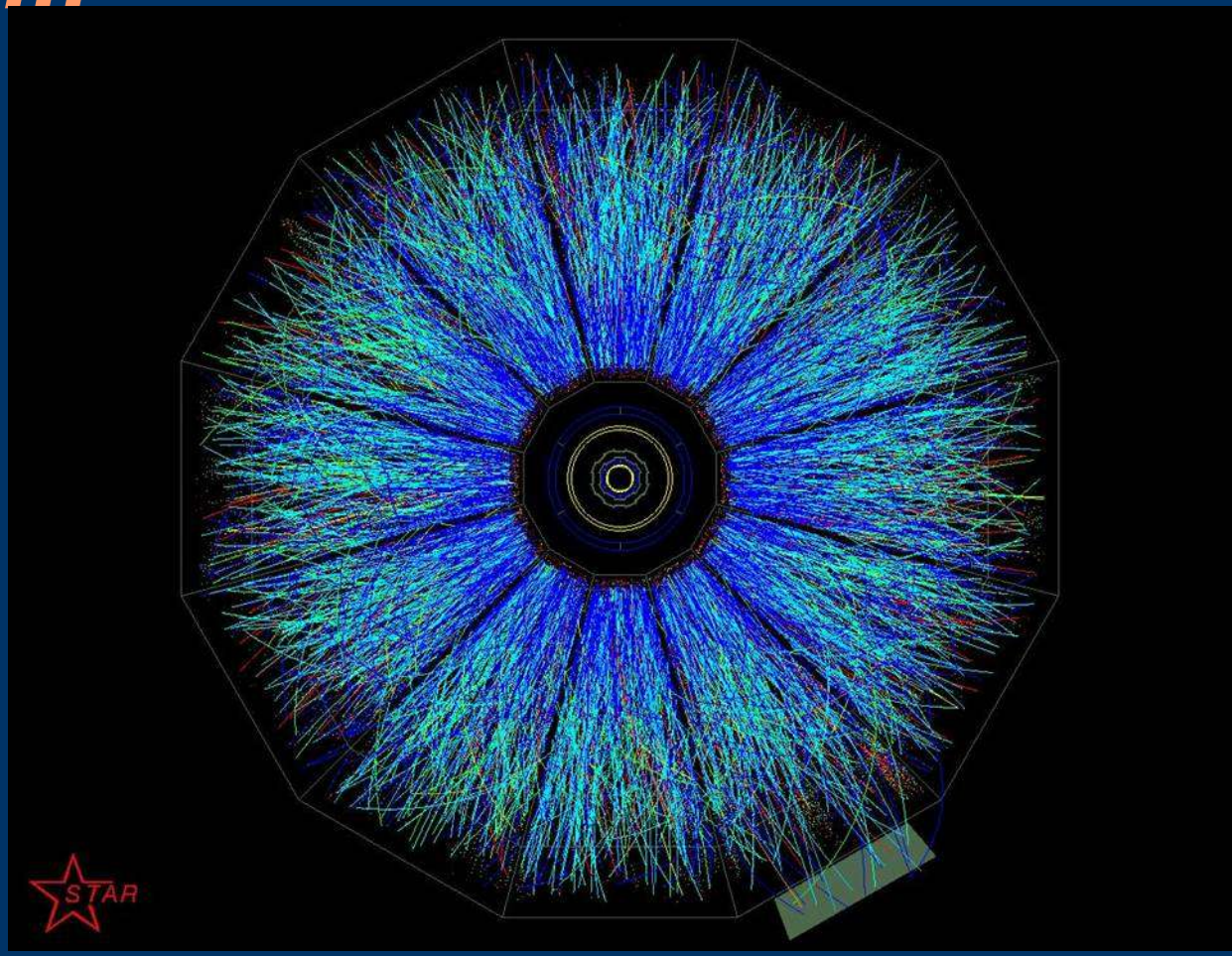
We have to deal with M of them and reduce the information to usable quantities



What are we ...

Software & Computing

Software and design
provide the tools,
Computing power for
data reduction and
analysis
Organization, planning,
priorities, plan



The lost child of STAR is actually trying to help getting Physics out of the above mess

General Organization

http://www.star.bnl.gov/STAR/comp/org/

S&C Leader (that guy giving this talk ... ☺

4 main areas with a head person (Leader)

Simulation
Maxim Potekhin

Reconstruction
Yuri Fisyak &
Manuel Calderon

Infrastructure

Database & Online Micheal DePhillips

Geant & Ev. Generator
Ronald Longacre

Embedding
Eric Hjort
Olga Barannikova

Event Display Root Valeri Fine
Victor Perevotchikov

Grid
Gabriele Carcassi
and many others

Global Tracking Project
Claude Pruneau (ITTF)

Computer support
Wayne Betts

QA
Lanny Ray

Production
Lidia Didenko (BNL) / Eric Hjort (PDSF)

Calibration
Gene Van Buren

13 Software sub-system coordinators
(TPC, FTPC, SVT, B-EMC, E-EMC, TOF, FPD/BBC ..) + you (??!)

How to get an account, help, etc ...

- Main Facilities
  PSDF (Parallel Distributed Systems Facility) and **RCF** (Rhic Computing Facility)
- Account : Where to get one & procedure
  Always start asking your council representative for adding you in the list of STAR collaborators (Liz Mogavero)
  Request an account
  `http://www.rhic.bnl.gov/RCF/UserInfo/GettingStarted/NewUser/`
  `http://pdsf.nersc.gov/starting/newuser_form.html`
- Login in on the RCF
  % ssh rssh.rhic.bnl.gov
  ... use Kerberos password (was mailed to users)
  % rterm
- PDSF: `http://pdsf.nersc.gov/index.html`

How to get an account, etc ...

- Mailing lists (Hypernews & `lists.bnl.gov`)

  - Don't try to register to Hypernews before getting an RCF account (it will disappear within 24 hours)
  - Follow link from the computing page ...
  - Always login before using the Web interface (pleazzzzzzz!!)
  - An extra password is required for protected/ lists ...

  starsoft-hn@www.star.bnl.gov STAR Software of General interest
  starsofi-hn@www.star.bnl.gov STAR Infrastructure
  offsites-hn@www.star.bnl.gov Off site facilities support
  ... find more on `http://www.star.bnl.gov/STAR/html/mail.html`

- Email account at RCF

  Imap: `{rcf|rmail}.rhic.bnl.gov` : has NOTHING to do with the RCF account (not the same password)

Where do I get help in STAR ??

- Getting help - Email
  `starsofi-hn` for infrastructure issues
  `starsoft-hn` for software of general interest
  specialized lists (EMC, Scheduler, Catalog, MuDst, ...

- BugTracking system (guest)
  - for STAR specific issues
  - `http://www.star.bnl.gov/rt2/`
  - Please, describe the problem as accurately as possible (with a "how to reproduce", macro, input you used etc ... We cannot guess what "*my code does not work, please help !!*" means.
  - "*my code was working yesterday*" as been proved to be an unreliable statement (dev, worked miraculously before, different farm, ...)

Resource, usage and Monitoring

- Disk space layout

`/star/data01/pwg` OR `/star/data02/pwg` PWG reserved

`/star/data05/scratch` scratch space ONLY !!!
beware, files will be deleted there after a while ...

`/star/data03/daq` DAQ files only

`/star/data04/sim` Simulation output

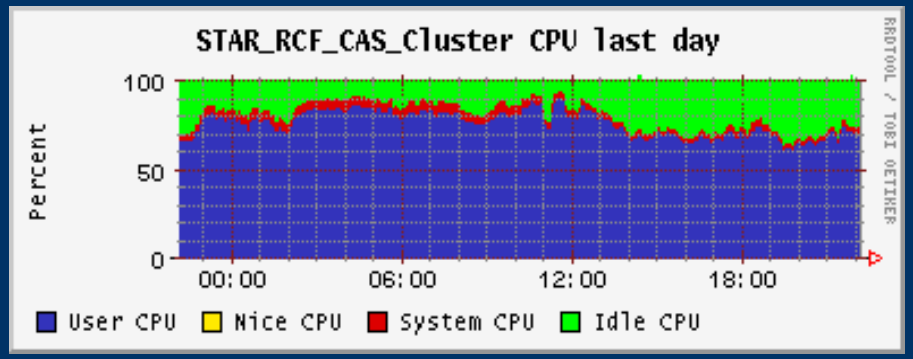
} Specific usage

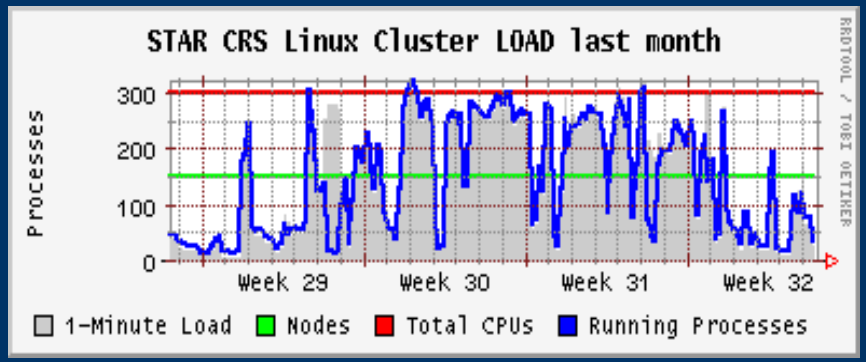
the rest (~ 30 TB) is reserved for production and is barely enough.

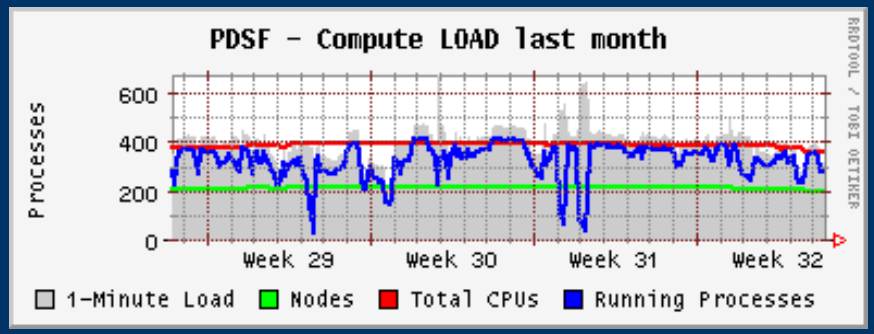
http://www.star.bnl.gov/webdatanfs/pub/overall.html

Resource, usage and Monitoring

- Lots of information from the **SOFI** page
  http://www.star.bnl.gov/STAR/comp/sofi/
  Disk resource and usage being monitored (you will find layout and usage there), CPU as well ...
  `http://www.star.bnl.gov/cgi-bin/nova/showMachines.pl`







Standards (a few examples)

- Many software problems can be avoided by following the Software rules (naming conventions)
  Why rules ?? Because 13 sub-systems software (and probably x10 different style) need to talk to each other ...

  `http://www.star.bnl.gov/STAR/comp/train/standards.html`
  `http://www.star.bnl.gov/STAR/comp/train/NamingAdvice.txt`

- New code follows a peer review process
  Ensures early compliance check
  May be regularly scheduled by sub-system (software decay)
  http://www.star.bnl.gov/STAR/comp/sofi/peer-review.html

Can I help ?? !!! YES YES !!!

- By working with one or the other software sub-system coordinators, you are definitely helping.
- By learning the STAR Software/ tools and its rule, or commenting on the documentation, you are helping ...
- By giving a tutorial today
- ...

- And if you really want to
  Service tasks
  `http://www.star.bnl.gov/STAR/comp/general/tasks.html`

Coming run data dilemma

- Always nice to plan for lots of events, opens new physics topics, etc …
- Production impacts
  Based on a gross assumption of ~ 70 M events

Au+Au 200 (minbias)	35 M central	35 M minbias	Total
No DAQ100 (1 pass)	329 days	152 days	481 days
No DAQ100 (2 passes)	658 days	304 days	962 days
Assuming DAQ100 (1 pass)	246 days	115 days	361 days
Assuming DAQ100 (2 passes)	493 days	230 days	723 days

will be challenging (3.5 passes average …)

More challenges ...

- Preceding numbers relies on DAQ100
  - Important for the data rate
  - To be done by December the 20th maximum
  - Some testing started ...

  Changes the DAQ Reading front hand software
  Possibly a good investment for uniformity online/offline

More challenges ...

- ITTF project (and impacts)
  - Equally important (breaks tracking time-linearity approach)
  - To be integrated on the first week(s) of January
  - First data will NOT be officially processed with old chain
  - Crisis management: new data through FastOffline (un-calibrated)
  - In April, stop support for the old tracking code

  Removal of most tables and possibly all FORtran code
  No need for PGI compiler anymore (medium term) ...
  Bye bye FORtran, all C++

Remains ... How are we going to handle all of this data ??

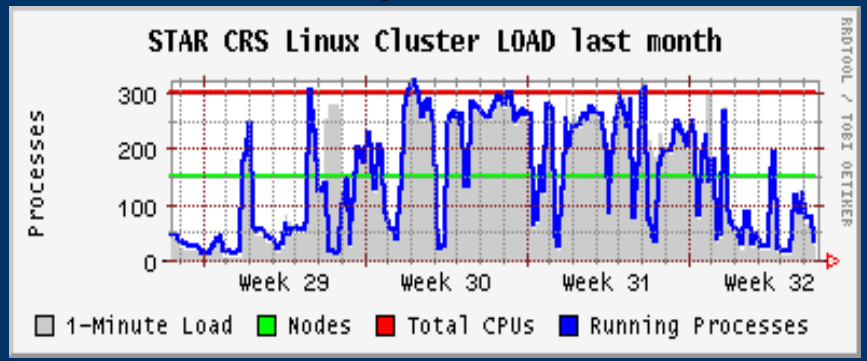
- ~ 7 Million files !!!?? Real Data Management problem
  - Resilient ROOT IO
  - Cannot proliferate more “kind” of files
  - Good luck with private formats …
  - Catalog better be scalable (and efficient)
  - Find a needle in a hay stack …

- Processing time and data sample very large
  - Need to off load user analysis (running where we can). Data production is not ready for multi-site …
  - Code consolidation is necessary (yet another reason for cleaning)
  - MuDst transfer alone from BNL to PDSF (at 3 MB/sec) would take 145 days …

A few things ...

- Several ways to reduce CPU cycles, the usual suspects
  - Code optimization (has its limits / hot spots)
  - Try ICC ?? (huum ... glibc issue found)
  - Better use of resources (Scheduler, farms, ...)
  - Offload user analysis (expands farm for production)

    [smells like grid already / get ready]
  - Bring more resources / facilities



Distributed Computing

- For large amount of data, intense data mining etc … distributed computing may be the key.

- In the U.S., three big Grid collaboration
  - iVDGL (International Virtual data Grid Laboratory)
  - GriPhyn (Grid Physics Network) : VDT
  - PPDG (Particle Physics Data Grid)

- STAR is NOT new to Grid activities
  part of PPDG since 2001, CS & Experiments working together

  We collaborate with : SDM (SRM), U-Wisconsin (Condor), J-Lab and Phenix … ITD/STAR = one team ...

What is Grid about ??

- **Data management**
  - **HRM** based file transfer — *Eric Hjort & SDM group*
  in production mode Since 2002, now in full production with 20% of our data transferred between BNL and NERSC. 2003 : **HRM** BNL to/from PDSF

- **Catalog**
  - FileCatalog (MetaData / **Replica Catalog**) development — *myself*
  - Site-site file transfer & Catalog registration work — *myself & Alex Sim*
  Final Replica Registration Service & defining necessary scheme to register files or datasets across sites

- **Analysis / Job management**
  - Resource Broker, batch (Scheduler) — *Gabriele Carcassi*
  - Interactive Analysis Framework solution — *Kensheng (John) Wu*

- **Monitoring**
  - *Ganglia & MDS publishing* — *Efstratios Efstathiadis*
  MonaLisa

What are we going to Grid about tomorrow ??

- Projects
  - Condor / Condor-G / VDT — Miron Levny
  - JDL, WebService project with J-Lab (next generation of grid architecture) — Chip Watson
- OpenScience Grid alignment
  Start federating Grid facilities in July ...
  Will possibly lead a unique effort in the US
- Grid is MUCH more ...
  It has deep consequence on the way we communicate and solve problems between groups (sometimes having conflicting scientific interests) as well as profound Geo-politics consequences ...

  STAR sharing CPU with Atlas, STAR/ITD, STAR/Phenix, ...

And so what ?? How does it makes better/faster science ??

- Remote facilities (big or small)
  - file transfer and registration work allows moving data-sets with error recovery (no need to “pet” the transfer, get rid of mindless tasks)
  - GridCollector does not require you to know where the files are, nor does the Scheduler (eliminate data placement task)
  - Grid enabled cluster bring ALL resources at reach
- Every day work
  - Mind set change : collection of data (will fit some analysis, some not)
  - Transparent interfaces and interchangeable components (long term)
  - Hopefully more robust systems (error recovery already there)
- Any other reasons ??
  - The Grid is coming, better get ready and understand it …
  - STAR is a running experiment with real data an practical needs (no sci-fi, plain real and harsh realities of the everyday life …)

Today's ...

- Please, beware of the Challenges ahead
- Grid is an opened invitation to YOU ...

- For today, focus will be on up-to-speed tutorials
  - More details on Infrastructure and STAR Software setup
  - Framework philosophy
  - StEvent IO model
  - GEANT overview and GEANT by example
  - Analysis and Simulation
  - MuDst
  - Batch (and a bit of everything else)