100 Million events, what does this mean ??

STAR Grid Program overview

Current and projected

- Always nice to plan for lots of events, opens new physics topics, rare particle detailed studies (flow of multi-strange particles etc …)
- Lots of numbers to look at in the next slides …

Au+Au 200 GeV projections 1

Au+Au 200 Production Central	**Min**	**Max**	**Expected Average**	**DAQ 100 (25%)**
1 event	85sec	115sec	100 sec	75 sec
1 M event	24k CPU hour	32 k CPU hours	17 k CPU hours	13 k CPU hours
Full RCF farm (150 nodes, 2 slots)	80 CPU hours	100 CPU hour	90 CPU hour	67 CPU hours
	3.3 days	4.4 days	3.75 days	2.8 days
Extrapolation to 100 Million	327 days	444 days	375 days	281 days
80% efficiency		555 days	470 days	352 days
1.2 passes			564 days	422 days

Pause

- That’s right, a year IS 365 days !!!
- We are now speaking of moving to a year-based production regime …

- Gotta be better for minimum bias right ??

Au+Au 200 GeV projections 2

Au+Au 200 Minimum Bias	**Min**	**Max**	**Expected Average**	**DAQ 100 (25%)**
1 event	32 sec	50 sec	45 sec	34 days
1 M event	9 k CPU hour	14 k CPU hour	12 k CPU hour	9.5 k CPU hour
Full RCF farm (150 nodes, 2 slots)	30 CPU hour	46 CPU hour	42 CPU hour	32 CPU hours
	1.3 days	2 days	1.8 days	1.4 days
Extrapolation to 100 Million	124 days	193 days	174 days	131 days
80% efficiency	155 days	242 days	217 days	164 days
1.2 passes			261 days	200 days

Useful exercise

	50 M central	175 M minbias	Total if requested
No DAQ100 processing time	282 days ~ 9 ½ month	456 days = 1 year and 3 months	738 days ~ 2 years
DAQ100 processing time	211 days ~ 7 month	350 days ~ 1 year	561 days ~ 1 ½ year
Total storage Size of event.root	105 TB	129 TB	234 TB
Total storage Size of MuDst.root	~ 18 TB	~ 18 TB (factor 7 used)	~ 36 TB
Number of files estimated (MuDst)	½ a million files	1 Million	1 ½ Million
Total estimation (data management)	2 ½ Million files	5 Million files	7 ½ Million

Number of files estimated using the current number of events / file. DAQ10 implies a reduction by ~ 5

Immediate remarks

- 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 …

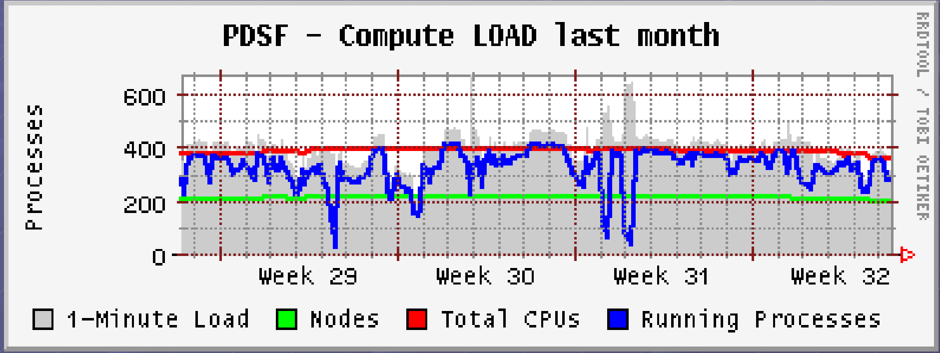
- Processing time and data sample very large
  - Need to off load user analysis (running where we can). Data production is not ready for multi-site …
  - Code consolidation is necessary (yet another reason for cleaning)
  - MuDst transfer alone from BNL to PDSF (at 3 MB/sec) would take 145 days …

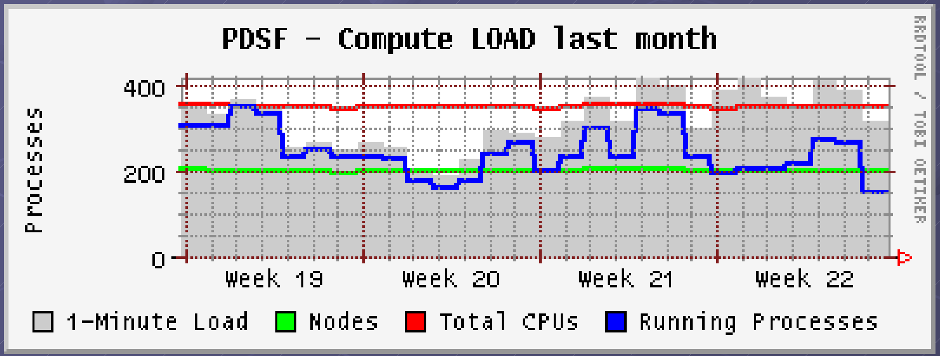
What can we do ??

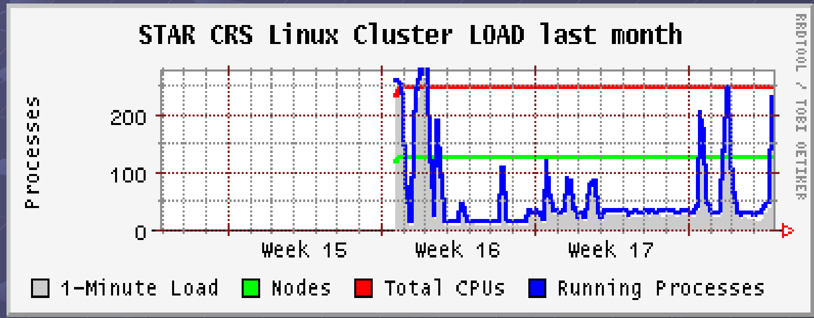
- Several ways to reduce CPU cycles, the usual suspects
  - Code optimization (has its limits / hot spots)
  - Try ICC ??
  - Better use of resources
  - Offload user analysis (expands farm for production)
  [smells like grid already]
  - Bring more resources / facilities
  - Any other ideas ??

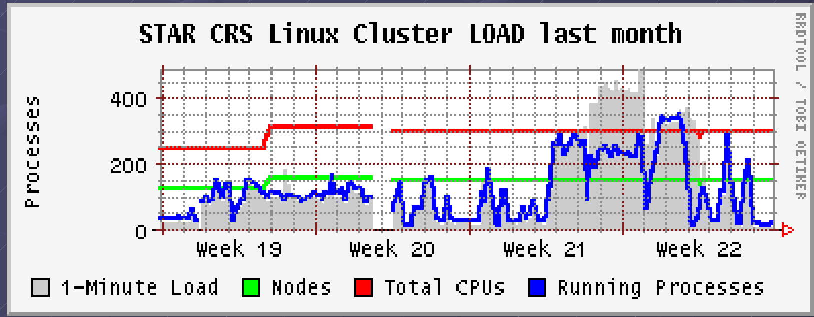
- Data taking & Analysis side
  - Reduce the number of events : Trigger
  - Smart selection (Selected stream - Thorsten)

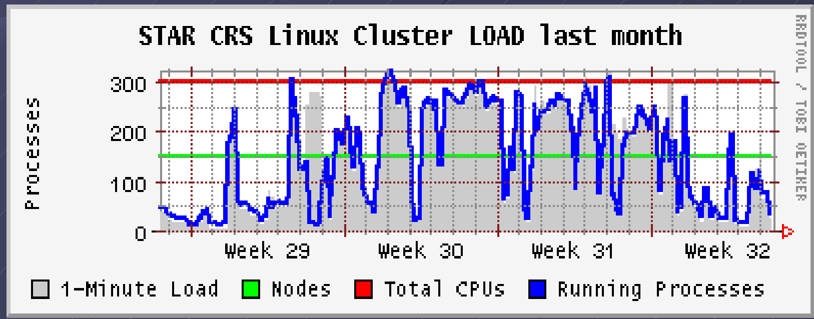
Better use of existing facilities ??











PDSF resources seems saturated
CRS/CAS load balancing is not …

More external facilities ??

- Investigation of resources at PSC
  - Processors there are 20% faster than a Pentium IV 2.8 GHz
  - Except that there are 700x4 of them ALREADY there and eager to have `root4star` running on them
  - AND if we build a good case, we can get up to 15% of that total (NSF grant) = that's 50% more CPU power comparing to 100% of CRS+CAS+PDSF
- Network ? 30 MB/sec (TBC) and part of the TeraGrid
- From "worth a closer look" in February, I say "GOTTA TRY".

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)
  - PPDG (Particle Physics Data Grid)
- PP what ??

  STAR is part of PPDG since Year1 (2 years ago)
  CS & Experiments working together

  We collaborate with : SDM (SRM), U-Wisconsin (Condor),
  J-Lab and even possibly Phenix …

What do we 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

Catalogue

- FileCatalog (MetaData / **Replica Catalog**) development *myself*
- Site-site file transfer & Catalog registration work *myself & Alex Sim*

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 *<u>Kensheng (John) Wu</u>*

What do we (still) Grid about ??

Monitoring

- Ganglia & MDS publishing *Efstratios Efstathiadis*

Database

- MySQL Grid-ification

Richard Casella & Michael DePhillips

Projects *:*

- Condor / Condor-G Miron Levny
- JDL, WebService project with J-Lab (next generation of grid architecture) Chip Watson

Much more to do … See `/STAR/comp/ofl/reqmts2003/`

If you are interested, will take you …

How does it change my life ??

- Remote facilities (big or small)
  - file transfer and registration work allows moving data-sets with error recovery (no need to “pet” the transfer)
  - 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
  - May not like it but … 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 …

Conclusion

- Hard to get back to slide one but …

  - Be ready for YEAR long production, we are at the one order of magnitude off level …

  - With such programs, we MUST integrate other resources and help others to expand mini-farms

- Grid

  - Tools already exists for data management

  - Must take advantage of them

  - More work to do for a production Grid … but coming (first attempt planned for the coming year)