DAQ 100 effects on Software and Computing.

- DAQ 100 overview
- The easy part ...
- The less easy part ...
- The definitly less easy part ...
- Now, thinking about it ...
- Solutions and ...
- (surprise title)
- Side effects, thoughts
- Final note

DAQ 100 overview and ...

- **Best described by the project's subtitle ...**

  “ How to store 100 central events/scdes ...”

- **Two main changes**

  ~ Double the data rate (30 MB/sec to 50 MB/sec)

  Compressed Data Stream (online clustering)

- **Assumed from now on ...**

  number of files x2, 5x more events per file,
  same size file, same RHIC duty factor ...

Leads to 10 times more events
Twice the rate to HPSS ...

The easy part ...

- **10 times more events ...**

  Database event level records become large ...

  Some effort but technical solutions exists ... (Jeff Porter)

- **Can we actually save this on HPSS ??**

  50 MB/sec requires more drives, twice the storage

  Minimum of 6 drives for sinking (4/2)

  Have 1100 tapes free ~ 65 TB of space

  24 M events (daq)

  120 M events (daq100)

  13 M events this year, 130 M planned

  Money solves this part ... Cost 200 K$

The less easy part ...

- 10 times more events ~45% gain in processing time ?
  - How long does it take to process ...
  - The answer is ... 2 weeks
- But seriously !!!
  - With an efficiency factor of ~ 60%

This year	1.5 M minbias events	0.5 M central
Next year	2.2 M minbias	0.7 M central

Just with the speed gain ...

Sounds good so far but ...

The definitly less easy part ...

- **Time for the complete set**

This year **4.7 M minbias in 6 weeks (1.5 month)**
3.6 M central in 14 weeks (3.5 month)

Next year **47 M minbias in 60 weeks (1 year and 3 month)**
36 M central in 140 weeks (~ 3 years)

That's ~ 4 years to go through the entire set ...

- **Solution : Put more money in ...**

100 nodes with (hopefully) twice the CPU speed, would cost ~ 400 K

Oups !! We exhaust our entire Computing budget !!!

Now, thinking about it ...

- All of this is fine, but it assumes we can actually process the data ...

  Haunting us for a while ...

  Memory leaks ... 550 kB/minbias event
  700 kB/central event

  That's 260 to 340 GB of memory after 500 events ...

- How does the user process all of this data ?

  i.e. With no increase in computing power ...

- Offsite data transfer ??

  With the StEvent files now being of equal size comparing to the daq file and the data volume scaling by 10 ...

Solutions and ...

- **Farm Uptime and load**
  More tools to have the queues always filled up and automated recovery of crashed ones ... Realistic ?
- **Code re- work ?**
  After profiling, we may get a better CPUtime/event. 12% means ½ year gain ... Worth a look ...
- **Speeding the IO** Victor Perevoztchikov
  Working on StEvent IO to speed up reading/writing
  Would benefit only some users ... (factor of 2)
- **MORE MONEY !!!**

Coup de Grace ...

- **By the way, where do we store all of this ??**
  Did I mentioned that the computing budget is already consumed ???
- **To think about ... (I mean : seriously think about !!)**
  Output production in MicroDst format
  Can work ONLY if PWG agree on a common format
  Make sens only if size is significantly smaller than StEvent
- **MORE MONEY !!!**

Side effects ... thoughts

- **More coordination in data transfer**

  More tools a-la DataCarousel will be required to manage vast amount of data.
  Distributed Disk project Claude Pruneau ?
  HRM project Pdsf, Arie/Alex

- **Will need a strong FileCatalog** Adam Kisiel/myself

- **MORE MONEY !!!**

Final note

- **YES !! We might make it ...**
  No, it won't be easy !!!

Lots of efforts, but enough ideas/solutions

Will require common effort (PWG, user education ...)

More planning and coordination ...

48 hours a day of work ...

Hum ... And maybe ... **MORE MONEY !!!**