Computing issues for the future ...

Jérôme Lauret

Computing issues for the future ...

Jérôme Lauret

- **Planning / funding**
- **What can we do**
- **Sociology**
- **Aging detector, upgrades**

Issues

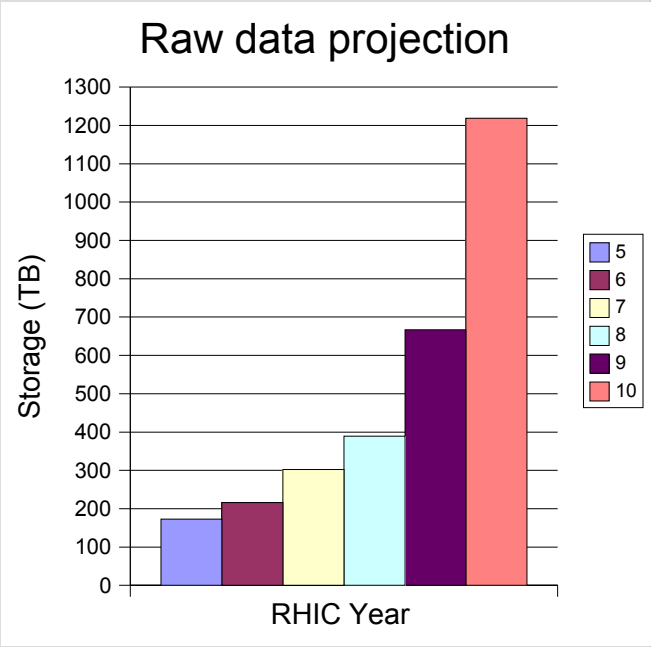
- Planning / funding
- What can we do
- Sociology
- Aging detector, upgrades

Planning, funding

- Grid activities are 2/3 years planned in advance (see proposals, timetable, roll-out plans)
- 2004, NSAC sub-committee were presented a “computing needs for the next 10 years”
- Currently under work, the network needs for the next 10 years
- Usually all of this done (due) before Physics plans are known
- **Missing (agreed) - “Tuning”**
  - **(re)-tie to STAR Physics Program & known funding profile**

Resource evolution ...

- "In"
  - As good as it could get ... not far off BUR
    - Long Au+Au at the end of the decade
    - Possible daq1000 effect: 2008 15%, 25%, 40% of DAQ1000 full "capacity"
  - Storage/CPU scales as raw data modulo adjustment
  - **Forced** to use **1.2 production pass**



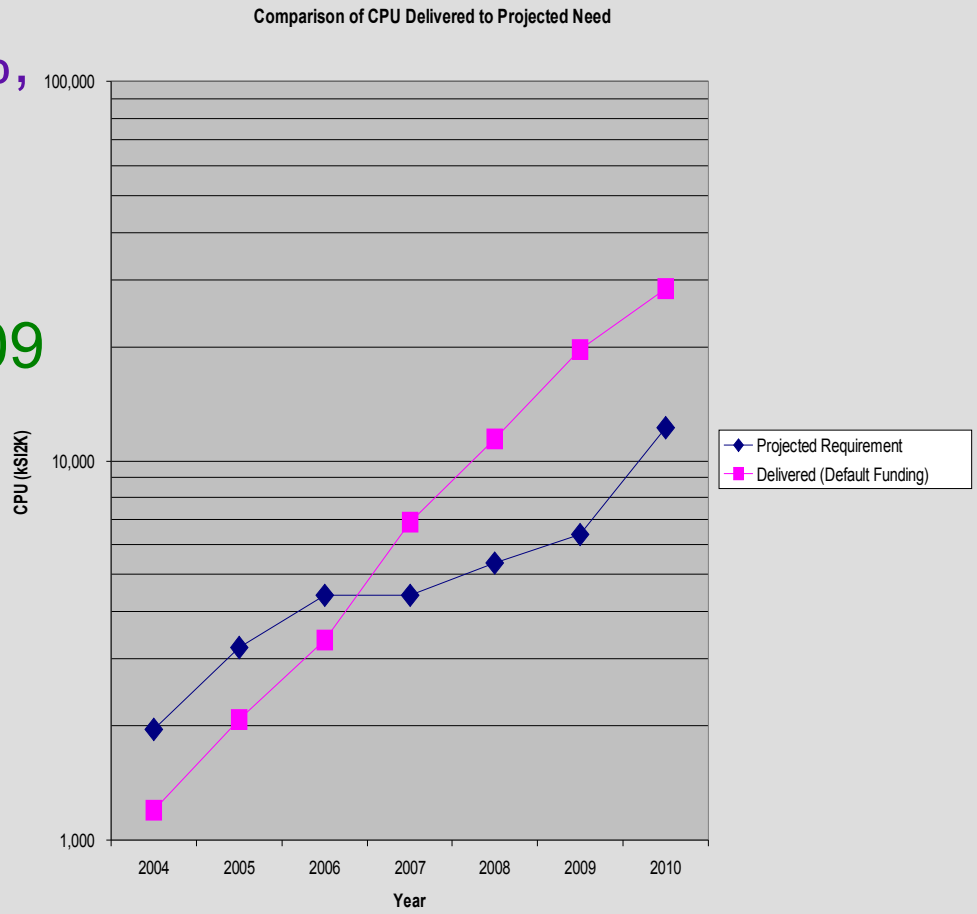
CPU in kSI2k

	2004	2005	2006	2007	2008	2009	2010
Projected Requirement	1,958	3,212	4,413	4,413	5,352	6,366	12,237
Delivered (Default Funding)	1,200	2,075	3,364	6,903	11,393	19,648	28,418
Delivered (Augmented Funding)	1,200	3,212	4,413	8,131	12,498	20,642	29,313
Adjusted requirement	4,814	7,897	10,849	10,849	13,158	15,650	30,083

Adjusted for STAR only

Resources ...

- "Out"
  - Shortfall is Y4 +75%, Y5 +10%, Y6 ~+0%
- Within those assumptions
  - 40%, 60% DAQ1000 in 2008/09 blows RCF budget by ~ 2 M$

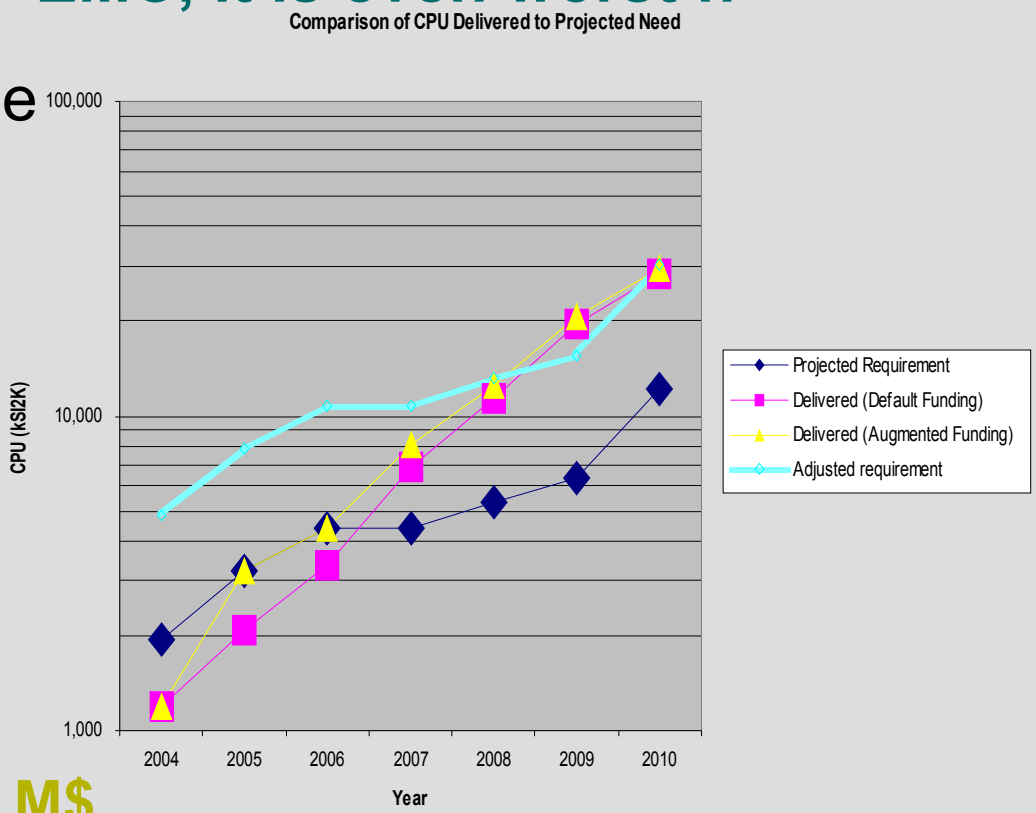


Reality ...

What we truly need is
2.5 minimal / 3.5 (reasonable) passes ....
Anyone who have worked on a TPC based detector knows that 1 pass does not make it + EMC, it is even worst !!

The difference between the blue curves is what we EXACTELY need in terms of EXTERNAL (non existing) resources.

Since NSAC presentation said, "*all is OK, don't worry*" (a strategy ?), passing this deficit as a need is difficult



$$ for the yellow curve = + 1.6 M$...

Reality (#2) ...

- But I think it all went in smoke (as well as any chances of the 1.6 M$ increase / recovery) with the repair needed here ...

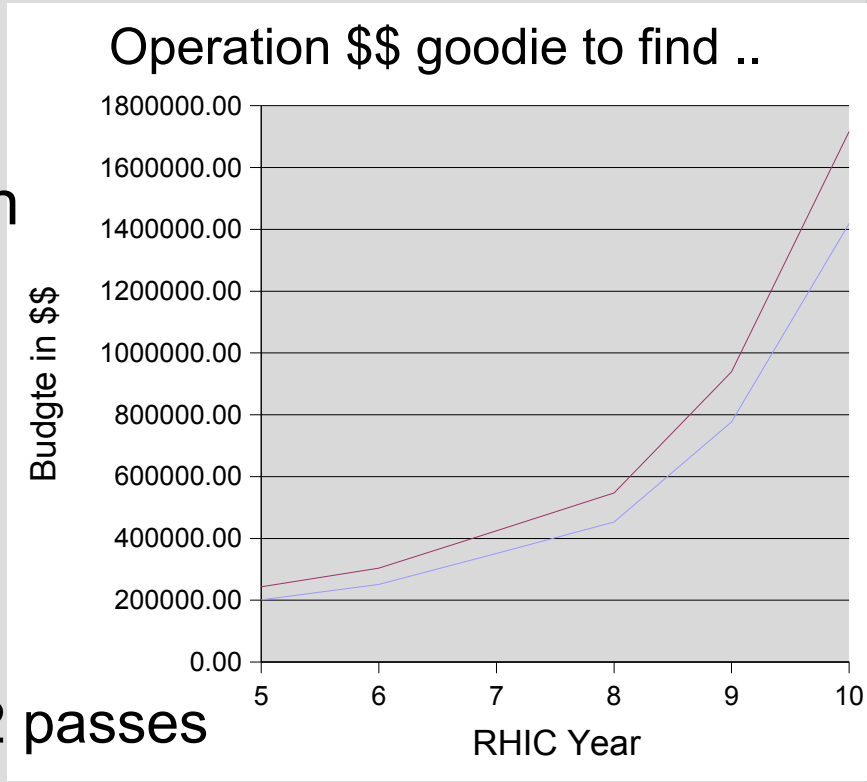
The bearing was damaged

We need it but, can we even afford > 1 pass production at this stage ??

Operation $$

- Projections do not include OPS
  - tape price is unknown
  - # pass insanely set by what we can buy as storage ...
  - NB: each production pass is big $$ in storage

- Is this a limiting factor ?
  - 200 k$ ~ 380 TB ↔ 170 M events,1.2 passes
  - 400 k$ ~ 660 TB ↔ 300 M events, 1.2 passes
  - etc ... minor details on species



Tell me the number of events and I will tell you the price keep in mind that 1 production pass is ~ ½ the init. price

What does this mean ??

- We can do more planning as per the current resources (PDSF/RCF)
- We can look in detail what we can do, cannot do, "*plan for the best we can with what we have*"
  AND
  Find solution to cover for missing resources by proof of principle + reaching far/remote resources (Grid)

What I think of solution #1

"*Much work Remains to be done before we can announce our total failure to make any progress ...*"

- Does not allow us to deploy the past years strategy for distributed computing, no chance to **get access** to **remote resources and / or spread funding (**STAR != US alone**)**

Truth about Computing funds

- **FY05 Computing budget**
  - OSCAR -10%
    - HEP -12% cut / **NP +5%** **(**actually STAR since Jlab/LQCD separate**)** + 2 SBIR in FY05
- **FY06 Computing budget**
  - OSCAR -11% . Details HEP / NP not looked into but
    - HEP ~ – 15% (NP ?)
    - Have been asked a few times if "*the STAR Collaboration really want Grid*" or for "*the opinion of the rest of the STAR management*"
  - In a fierce funding, our part of the pie should be defended as fiercely (and it is to some extend, not giving up ...)
    - Predict that lack of statement from "STAR" is badly hurting
    - Any "plan" is useless without more support and understanding (appearance of a single-handed effort)

Issues

- Planning / funding
- What can we do
- Sociology
- Aging detector, upgrades

HPSS/ drive / LAN

- Capacity – Not random luck (also fought for)
  - Each drive does 12 MB/sec x 10 for STAR alone – While we cannot produce the data, we can "take" it

	FY '04	FY '05	FY '06	FY '07	FY '08	FY '09	FY '10
CPU (kSPECint2000)	1,958	3,212	4,413	4,413	6,407	7,500	12,745
Disk (TBytes)	333	677	1,053	1,053	1,322	1,322	2,088
Disk (GBytes/sec)	3.3	6.8	10.5	10.5	13.2	13.2	20.9
Tape (PBytes)	2.2	4.5	7.8	9.7	13.8	17.1	24.2
Tape (MBytes/sec)	560	854	1,247	1,247	1,491	1,491	2,602

DAQ1000 can start gently in 2006

- Data samples
  - ~ fits within the 2 Gb links in place (125 MB/s x 2)

- Computing model / network needs

Scenario with 3.5 estimates	FY05	FY06	FY07	FY08	FY09	FY10	
Transfer rate needed for months=6 (MB/sec)	34.88	43.39	66.43	104.39	168.72	319.39	Not viable
Transfer rate needed for months=3 (MB/sec)	69.75	86.78	132.86	208.77	337.44	638.79	Viable
Transfer rate needed for months=2 (MB/sec)	104.63	130.18	199.29	313.16	506.16	958.18	*More than Comfortable*
Known previous requirements Mb/sec	1350	2025	3050	4600	4600	4600	
	OC48	OC192	OC768	OC786			

Storage & Cataloging

- Large data samples implies issues with
  - Keeping track of where the files are
    - Current catalog will "survive" up to 2010, re-writte in C/C++ to ensure longevity/availability (*Maciek Jedynak* - Varsaw)
  - Maximizing file location / resource consumption
    - IO congestion happen every day ...
    - Need scalable solutions

- Existing system?
  - Scalable persistent storage solutions. Several ... (**PANFS**)
  - Even more Grid inspired (or necessity) driven tools (SRMs/DRM, dCache, **XROOTD**, SANFS, GUPFS, ...)
  - **XROOTD**: (*Pavel Jakl* - Prague)
    - Provide load balancing, dramatic failure recovery, fail-over
    - Manages space (restore files from HPSS if not present). Later will be able to bring files "*from remote*" (STAR/SDM/SLAC/IN2P3)
    - Name space conservation for catalog (redirector)

CPU future, how to recover ??

- **Analysis are challenged in event-rich experiments**
  - Analysis on the Grid
    - *GridCollector* (*SciDac* funded) allows accessing a sub-set of events, REGARDLESS of the location of the file
      (*Kensheng Wu* - SDM/LBNL)
  - SUMS, SBIR & Grid Submission Service
    - "Scheduler" like. *SciDac/PPDG, SBIR Phase I* with *Tech-X, SBIR Phase II with* SH
- **Status**
  - Running simulation is fine (done it)
  - Hard time to pull through reconstruction
  - Issues span from Network, knowledge/manpower, ...

Although requires more work, we COULD be fully Grid enabled for any production challenges NOW ... if we really want it.
Gain ~ + 0.4 passes with Terra-Grid resources (Mort Kaplan)

Reco / Simu

- **While ITTF is coming ... and coming ... and coming ...**
  - Really moving to next phase ... Trying to support the long term
  - While allowing rapidly evolving software/hardware landscape
  - And provide stability to users, R&D etc ...
- **Especially**
  - Evaluation of VMC as “an” approach allowing G3 to G4 smooth transition + reco/simu geometry reconciliation
  - Would like to see “a” geometry description independent on simulation package (GDML, AGDD, ...) i.e. single source (*Sunil Manohar Dogra*)
- **Other**
  - High eta/forward tracking (FTPC/PMD interest) (*Pawan Netrakanti*)
  - Pile-up rejection vertex finder (delayed?)
  - Track tagging and multiple-event
    - If several vertices are found, corresponding to several real event, what is the definition of a primary ??
  - HFT “frames” makes no sens to me so far as per reco approach (but there is time to speak about it)

Issues

- Planning / funding
- What can we do
- Sociology
- Aging detector, upgrades

Sociology, NP culture

- **Generically**
  - Changing collaboration mode of thinking as per resource usage is damned hard ...
  - Tools like the `DataCarousel` or SUMS took years to be adopted

- **Difficult times means difficult measures but also ...**
  - To survive, enforcement (therefore frustration) will be needed. No other method have proved itself to work.
  - Get accustom to Space recycling, data re-location, etc ...
    - Quota & file deletion, disk cleaning, backup, compression

  Any other mean seem to lead to user striving to cheat the system. We all gain by playing the rules (resource sharing not take-over). If you have a good idea, let's talk.

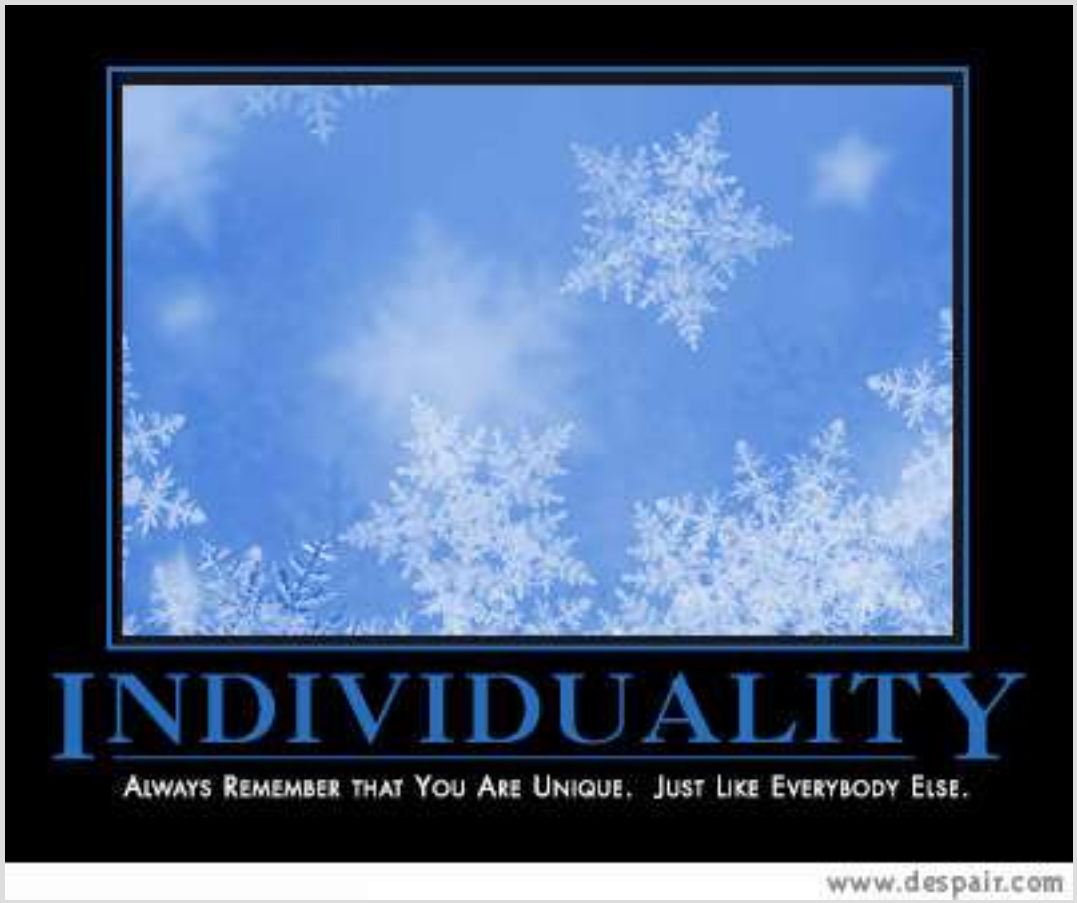
  There will be NO magic money, no magic resources, **we are down by ~ 40%** (FY05/FY06) comparing to 1.2 baseline

Sociology, Physicists supremacy

- Difficult times requires increased wisdom and self-assurance
  - Computing (still) considered (to advertised) as an under-worth activity
    - Not unique to Computing but ...
    - What it does is discourage the youngest by diminishing the worth of their work, in a time when we need more talents and more knowledge, facing bigger problems
  - There is a LONG way to go for this change of mentalities (may not even be possible as counter-human being nature)

    While waiting ... one perspective ...

Sociology, perspective



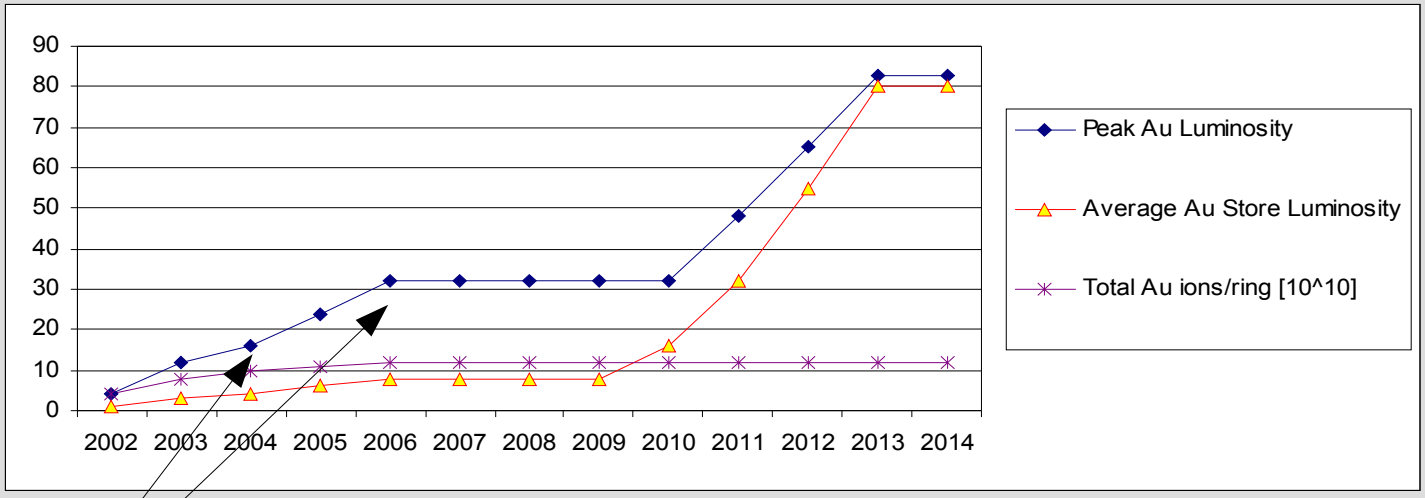
Issues

- Planning / funding
- What can we do
- Sociology
- Aging detector, upgrades

Calibration / Detectors ??

- We heard a status from Gene GridLeak on Wednesday
  - Great work as usual, even greater Jim/Gene tandem ...
  - But worried about
    - **TPC performance and Specs not well understood, we don't have to "use" luminosity but we like push-ing too ...**

Does not matter if you look at peak or average, it is still x2



There is a factor of x2 between 2004 and 2006 in Roser's projections

What's going to happen past 2010 ???

Calibration / Detectors

- **Equally worried about**
  - Possibly aging TPC (and/or TPC expertise and personnel, no dis-respect meant) on a 10 year scale
- **To avoid "patch-work" and a tendency for last minute re-action**
  - TPC should be reviewed as ANY other sub-system (a good time to think about performance and manpower)
  - If a long term detector, expertise should be secured
- **Why would this be a Computing issue (??)**
  - Because it currently
    - Often unexpectedly divert manpower in sub-system problems
    - Delay production schedules already hard
    - Work mode not compatible with 1.2 passes (or whatever we believe viable but minimal) AND preparation for the future

Summary 1

- Computing funds
  - Need to find away to revert NSAC either smart or miss-guided message. In difficulty until 2008/2009 unless we can revert 1.2 passes to 2.5 (or better 3.5). In the mean time: careful about DAQ1000 (not before 2008, within max-capability %tage)
  - Funding as-is is so "*not good*" it may (will) prevent multiple passes by simple storage cost. One mistake and we lose !!
  - OSCAR/SciDAC questioning single handed effort
    - Bad for our Grid efforts, bad for OSCAR funding in general
    - Will go from bad to very bad if un-addressed
    - Not going to Grid is missing remote resource opportunities
- What we can do – A lot.
  - Ideas and solutions for immediate needs are endless
  - Often medium-term critical projects turned into service tasks
  - Manpower is short: often preventive to exploiting resources at hand

Summary 2

- Sociology
  - We HAVE TO change our ways of work to be able to continue to work. “It is ALL mine” time are over, sharing has come (I know, it is as hard as in kinder-garden)
  - A lot to do to change mentality toward computing and pursue the building of needed strength
- TPC specific
  - Physics relying on TPC tracking with no consistent planning for TPC is suicidal
  - Need a thoughtful review to evaluate performance, especially RHIC-II era
  - If problem, time is short (very short to prohibitive)
  - Current mode of operation prevent from attending other issues