T0/1/2 to T0/1/2 relationship

F.Carminati

June 12, 2006

T2 workshop@CERN

ALICE computing model

- For pp similar to the other experiments
  - Quasi-online data distribution and first reconstruction at T0
  - Further reconstructions at T1's
- For AA different model
  - Calibration, alignment, pilot reconstructions and partial data export during data taking
  - Data distribution and first reconstruction at T0 in the four months after AA run (shutdown)
  - Further reconstructions at T1's
- T0: First pass reconstruction, storage of RAW, calibration data and first-pass ESD's
- T1: Subsequent reconstructions and scheduled analysis, storage of a collective copy of RAW and one copy of data to be safely kept, disk replicas of ESD's and AOD's
- T2: Simulation and end-user analysis, disk replicas of ESD's and AOD's

Pledged by external sites versus required MoU

		2007		2008		2009		2010	
		T1	T2	T1	T2	T1	T2	T1	T2
CPU	TDR requirement (MSI2K)	4.9	5.8	12.3	14.4	16.0	18.7	20.9	24.3
	Missing %	-44%	-40%	-43%	-56%	-29%	-41%	-24%	-53%
Disk	TDR requirement (PB)	3.1	1.5	7.9	3.7	10.2	4.8	13.3	6.2
	Missing %	-61%	-35%	-61%	-48%	-51%	-28%	-46%	-32%
MS	TDR requirement (TB)	2779	-	6947	-	9031	-	11740	-
	Missing %	-45%	-	-45%	-	-15%	-	-9%	-

Relations T1 – T2

- In the ALICE Computing TDR there are no particular relations among Tier1 and Tier2 sites
- All sites of each category share their tasks
  - Reconstruction, scheduled analysis, and unscheduled analysis and Monte-Carlo production
- Relations T2 and T1 are in terms of data storage
  - MC data and AOD from unscheduled analysis produced at T2 are shipped to the “closest” T1 for custodial storage
- In countries with a T1, T2’s in the country refer to it
  - This is the case in France, Germany and Italy
- In other countries, the T1 should ideally be the one with the best bandwidth

Relations T1 – T2

- The main question is the impact in terms of storage @ T1's and network resources
- We have estimated it using only MC data, which provide the bulk of data at Tier2
- Available resources, pledged so far to ALICE, only allow producing about 50% of the MC data required by our Computing Model
- Storage and bandwidth, assuming that all T2's absorb proportionally the 50% deficit

Disclaimer

- During the Rome GDB it was asked to experiments to provide the T2-T1 relationships
- ALICE said that it would have preferred LCG to handle the first version of the table
  - But this task was pushed back on experiments
- We have now a tentative table, however
  - It does not follow from our computing model
  - It allows however to solidify the relations T2 – hosting T1

Analysis of first data in 2007

Resources required to analyze the first pp data in 2007

Working hypothesis	# real events		50,000.0
	# reconstruction passes		10.0
	# analysis passes		100.0
	# MC events		######
Processing	Real reconstruction	kSI2K s	3.5E+06
	Real analysis	kSI2K s	4.6E+05
	Real total	kSI2K s	4.0E+06
	MC generation	kSI2K s	4.5E+06
	MC reconstruction	kSI2K s	7.0E+05
	MC analysis	kSI2K s	9.2E+06
	MC total	kSI2K s	1.4E+07
	Real + MC total	kSI2K s	1.8E+07
Storage	Real	MB	8.3E+05
	MC	MB	9.9E+06
	Total	MB	1.1E+07
Avalilable resources in 2008	CPU @ T1[1]	kSI2K	9.5E+00
	CPU @ T2[2]	kSI2K	8.3E+00
	CPU total	kSI2K	1.8E+01
	Disk @ T1	MB	4.0E+06
	Disk @ T2	MB	2.3E+06
	Disk total	MB	6.3E+06
Processing time		h	2.9E+02
Disk occupation		%	1.7E+02
Processing units for 1 week	MC	kSI2K/week	2.4E+01
		kSI2K/day	1.7E+02
	Real	kSI2K/week	6.5E+00
		kSI2K/day	4.6E+01
			5.8E+00

PDC'06 progress and plans

- Production of p+p and Pb+Pb events
  - Conditions and samples agreed with PWGs
  - Data migrated to CERN
  - Duration – 30-45 days (ongoing)
- Push-out to T1s (through FTS) - July
- Registration of RAW data -> first pass reconstruction at CERN -> secon pass reconstruction at T1s – August/September
- User analysis on the GRID and at CAF – September/October

Grid software deployment and running

- LCG sites are operated through the VO-box framework (see talk of Patricia)
  - All ALICE sites should provided one
  - Relatively extended deployment cycle, a lot of configuration and versioning issues had to be solved
  - Situation is quite routine now
- Data management
  - This year – xrootd as disk pool manager on all site SEs
  - The installation/configuration procedures have just been released
  - xrootd integrated in other storage management solutions (CASTOR, DPM, dCache) – under development
- Data replication (FTS)
  - Lower-level service for p-2-p data transfers
  - We use it for scheduled replication of data between the computing centres (RAW from T0->T1, MC production T2->T1, etc...)
  - Fully incorporated in the AliEn FTD, to be extensively tested in July

VO box support and operation

- In additional to the standard LCG components, the VO-box runs ALICE-specific software components
- The installation and maintenance of these is entirely our responsibility:
  - Regional principle, few named experts handling the installation and support (alicesgm), emphasis on overlap and redundancy:
    - CERN – Pablo Saiz (also AliEn expertise), Patricia Mendez Lorenzo, Latchezar Betev
    - Italy – Stefano Bagnasco
    - France – Artem Trunov, Jean-Michel Barbet (Subatech Nantes)
    - Germany – Kilian Schwarz, Jan-Fiete Grosse Oetringhaus (Muenster)
    - Russia – Mikalai Kutouski, Eygene Ryabinkin
    - Romania – Claudiu Shiaua
    - All others (NIKHEF, SARA, RAL, US) - Patricia Mendez Lorenzo, Latchezar Betev
    - We wish to enlarge this list with more experts (if they become available)
    - Installation, maintenance and operation procedures are quite well documented (major effort by the AliEn team and Stefano)
- Site related problems are handled by the site admins
- LCG services problems are reported to GGUS

Conclusions and outlook

- The last exercise before data taking (PDC'06) has started as planned
  - T2s are playing an important role in its operation – both in terms of resources and services testing
  - It is a test of all Grid tools/services we will use in 2007
    - If not in PDC'06, good chance is that they will not be ready
  - It is also a large-scale test the computing infrastructure – computing, storage and network performance

Conclusions and outlook (2)

- We have (some untested) all pieces needed to run production on the Grid
- The exercise has started 1 ½ months ago and will continue until the end of the year
- At the moment, we are optimizing the use of resources, improving the stability of services at all sites
- Next phase of the plan is a test of the file transfer utilities of LCG (FTS) and integration with AliEn FTD
- Finally we will do the most difficult task – analysis on the GRID


"That's
all
folks!"