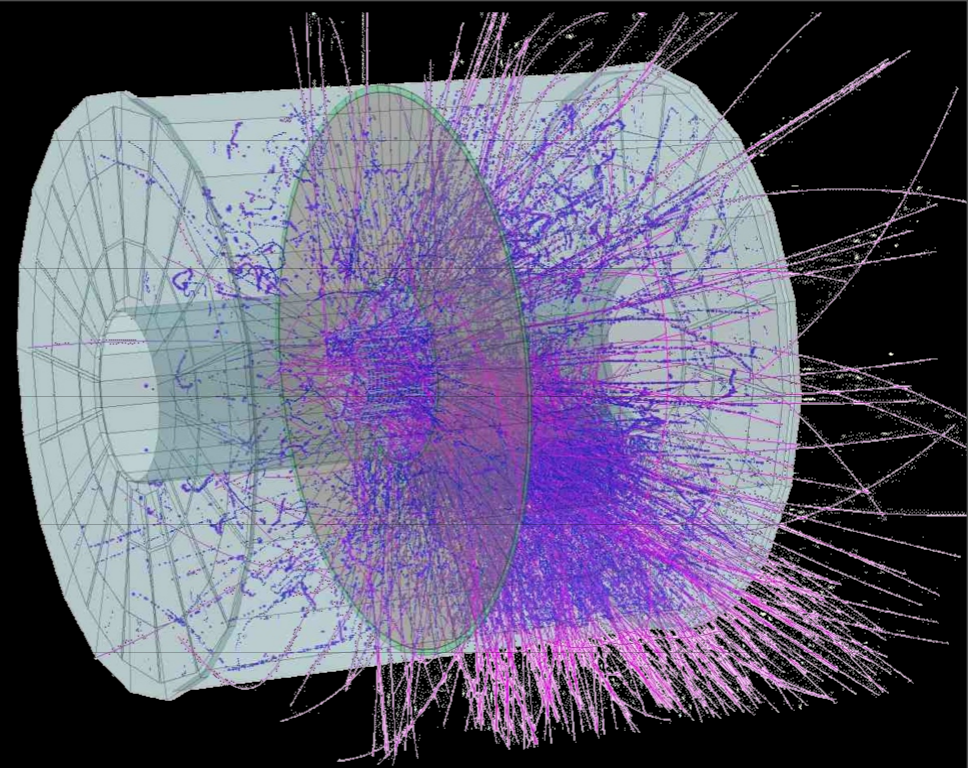


The ALICE Computing Status and Readiness

Federico Carminati
LHCC, July 1, 2008

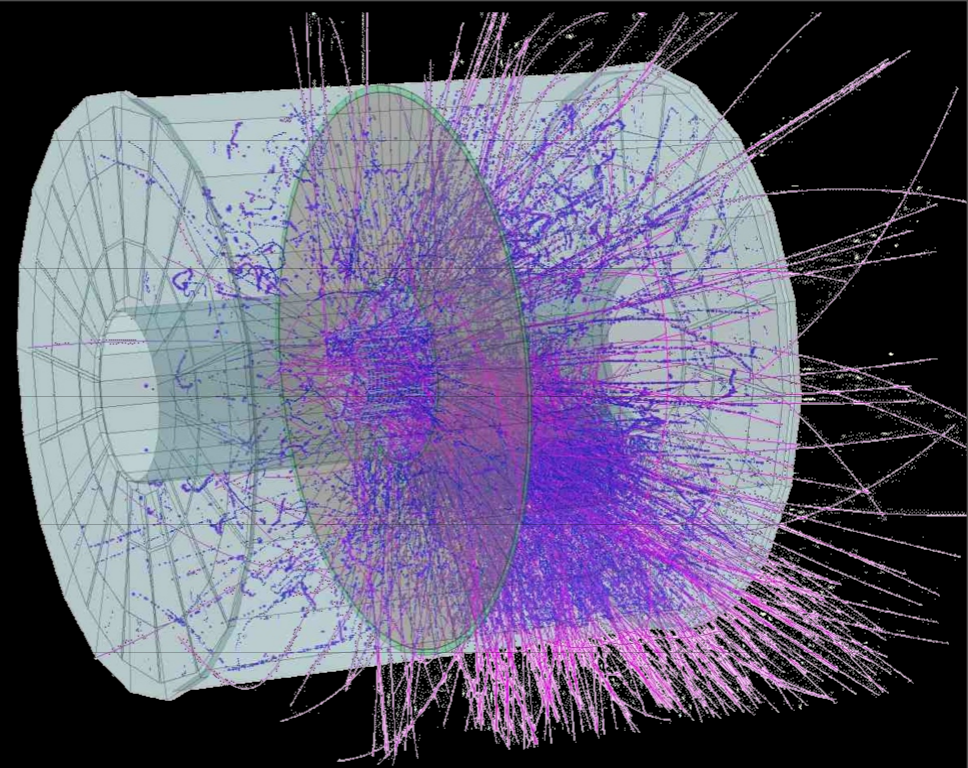


- Registration of data in CASTOR2 (T0) and on the GRID
- Replication T0->T1
- Conditions data gathering and publication on the GRID
- Reconstruction: quasi-online, pass 1 at T0 & pass 2 at T1s
- Replication of ESDs to CAF/T2s
- RAW data quality control
- MC production
- User analysis @CAF & @T2s

ALICE @ CCRC

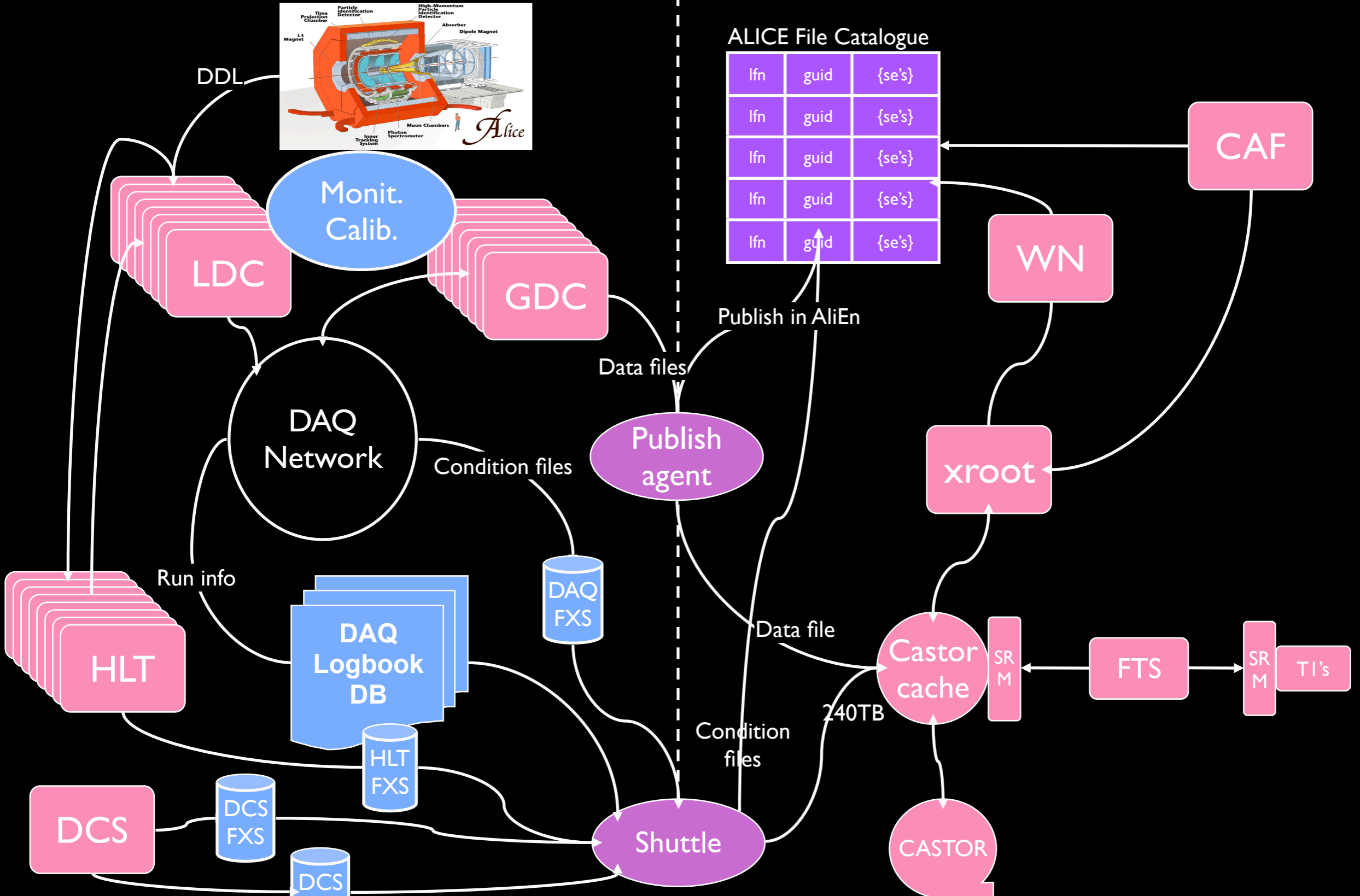
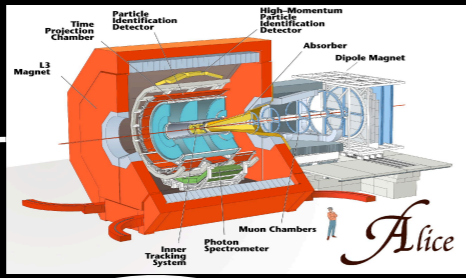
task list

- Registration of data in CASTOR2 (T0) and on the GRID
- Replication T0->T1
- Conditions data gathering and publication on the GRID
- Reconstruction: quasi-online, pass 1 at T0 & pass 2 at T1s
- Replication of ESDs to CAF/T2s
- RAW data quality control
- MC production
- User analysis @CAF & @T2s



Online

Offline

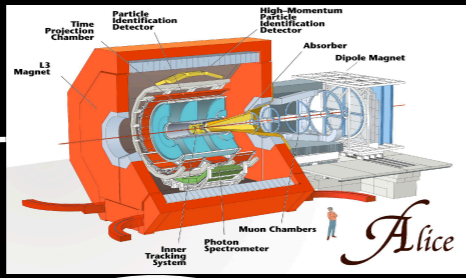


ALICE File Catalogue

lfn	guid	{se's}
lfn	guid	{se's}
lfn	guid	{se's}
lfn	guid	{se's}
lfn	guid	{se's}

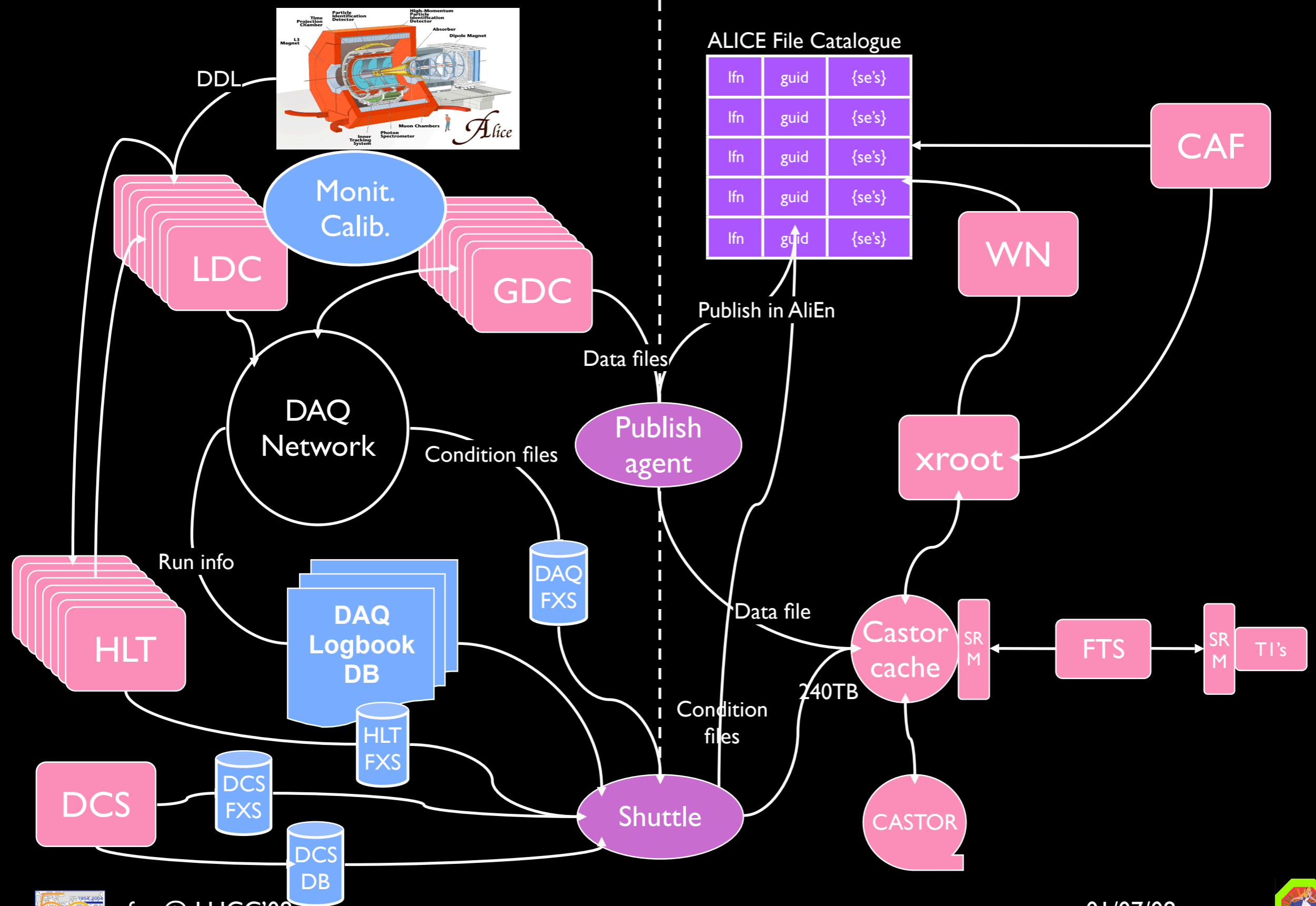


High-level dataflow



Online

Offline



- Detector commissioning
- Offline upgrades
 - New VO-Box (SLC4/gLite 3.1 - 64 bits)
 - New AliEn
 - Tuning of reconstruction
 - Fast-lane calibration & alignment



Activities in May - June

- Detector commissioning
- Offline upgrades
 - New VO-Box (SLC4/gLite 3.1 - 64 bits)
 - New AliEn
 - Tuning of reconstruction
 - Fast-lane calibration & alignment



- Despite Grid upgrades, steady RAW registration & replication
- Many thanks to CERN IT/FIO for the efficient CASTOR2 upgrade (coinciding with ALICE DAQ and AliEn central services upgrade)
- 88% of the expected volume (96TB)
- Short runs -> small files, as expected

2543 runs		113644 files	84.25 TB
-----------	--	--------------	----------



Data volumes

- Despite Grid upgrades, steady RAW registration & replication
- Many thanks to CERN IT/FIO for the efficient CASTOR2 upgrade (coinciding with ALICE DAQ and AliEn central services upgrade)
- 88% of the expected volume (96TB)
- Short runs -> small files, as expected

2543 runs		113644 files	84.25 TB
-----------	--	--------------	----------



```
/alice/data/2008/LHC08b/000037956/raw/ > ls -al
-rwxr-xr-x  alidaq  alidaq          530440 Jun 12 20:34 08000037956000.0.tag.root
-rwxr-xr-x  alidaq  alidaq       9776226240 Jun 12 20:55 08000037956000.10.root
-rwxr-xr-x  alidaq  alidaq       9769266079 Jun 12 20:55 08000037956000.20.root
-rwxr-xr-x  alidaq  alidaq       9770767560 Jun 12 20:55 08000037956000.30.root
-rwxr-xr-x  alidaq  alidaq       9769218809 Jun 12 20:55 08000037956000.40.root
-rwxr-xr-x  alidaq  alidaq       9768620353 Jun 12 20:55 08000037956000.50.root
-rwxr-xr-x  alidaq  alidaq       9777456300 Jun 12 20:37 08000037956000.60.root
-rwxr-xr-x  alidaq  alidaq       3391443105 Jun 12 20:35 08000037956000.70.root
```

- 10GB RAW will enter production on 3/7/08
 - Testing completed: no writing (rfio) or (xootd) degradation
- MSS requirements for large files and sparse access to tapes
- Fewer entries in the Grid file catalogue
- Larger size of secondary and tertiary files
 - 10GB RAW -> ~1GB ESDs -> ~100MB AOD



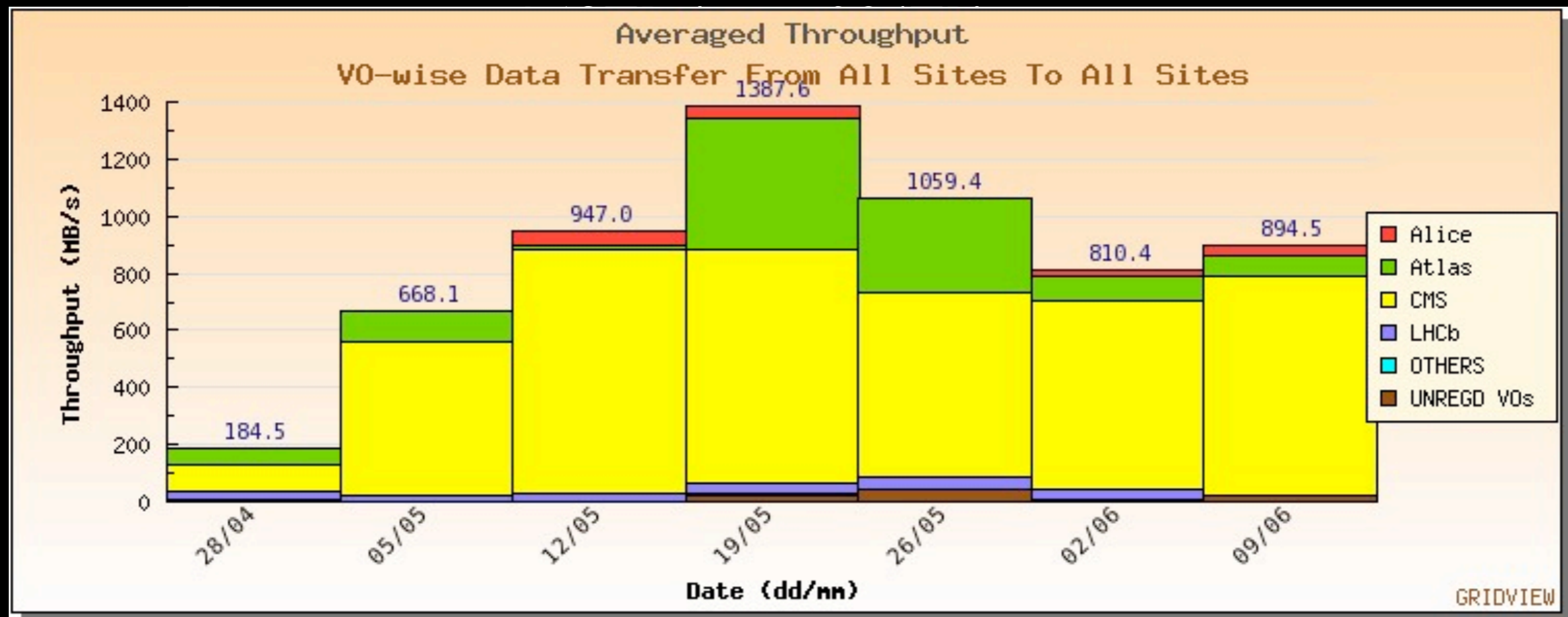
File sizes

```
/alice/data/2008/LHC08b/000037956/raw/ > ls -al
-rwxr-xr-x  alidaq  alidaq          530440 Jun 12 20:34 08000037956000.0.tag.root
-rwxr-xr-x  alidaq  alidaq       9776226240 Jun 12 20:55 08000037956000.10.root
-rwxr-xr-x  alidaq  alidaq       9769266079 Jun 12 20:55 08000037956000.20.root
-rwxr-xr-x  alidaq  alidaq       9770767560 Jun 12 20:55 08000037956000.30.root
-rwxr-xr-x  alidaq  alidaq       9769218809 Jun 12 20:55 08000037956000.40.root
-rwxr-xr-x  alidaq  alidaq       9768620353 Jun 12 20:55 08000037956000.50.root
-rwxr-xr-x  alidaq  alidaq       9777456300 Jun 12 20:37 08000037956000.60.root
-rwxr-xr-x  alidaq  alidaq       3391443105 Jun 12 20:35 08000037956000.70.root
```

- 10GB RAW will enter production on 3/7/08
 - Testing completed: no writing (rfio) or (xootd) degradation
- MSS requirements for large files and sparse access to tapes
- Fewer entries in the Grid file catalogue
- Larger size of secondary and tertiary files
 - 10GB RAW -> ~1GB ESDs -> ~100MB AOD

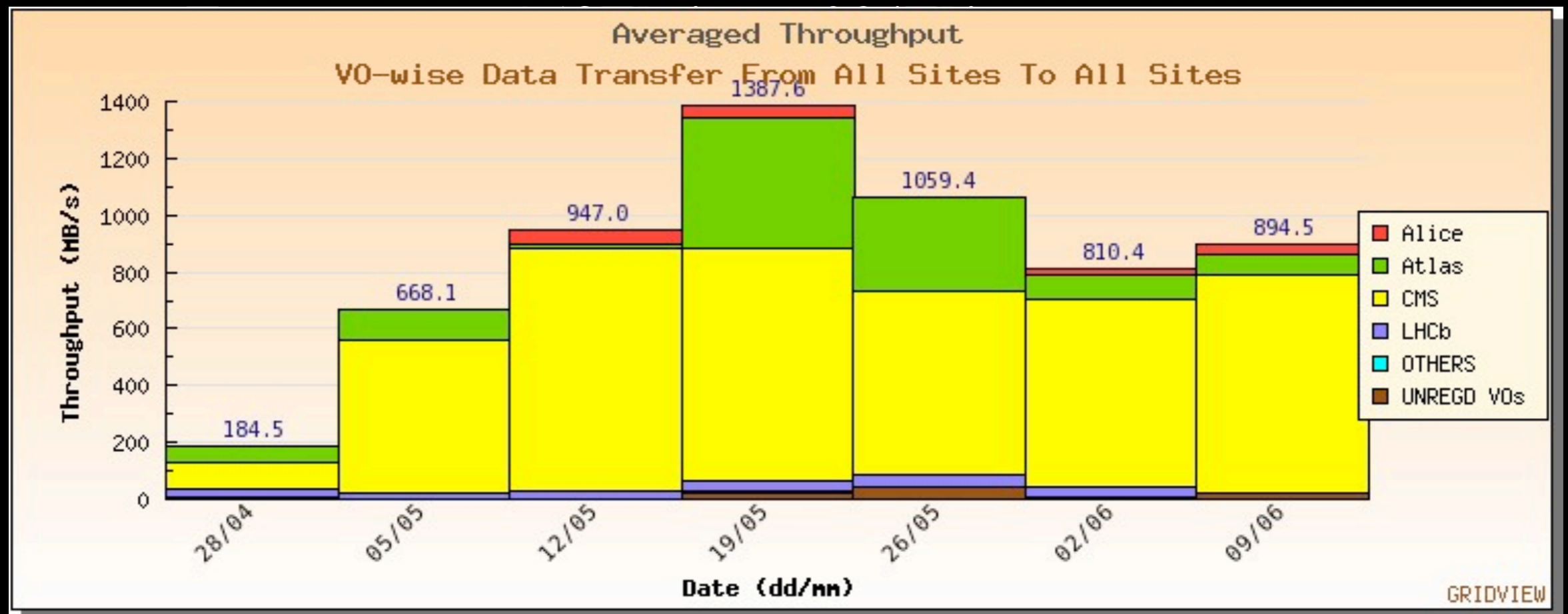


- Only selected (calib + align) RAW replication
 - Started immediately after registration@T0 from data on disk
 - According to TI capacity share
- 60MB/s (when data present) => SDTY pp data taking scenario



Replication

- Only selected (calib + align) RAW replication
- Started immediately after registration@T0 from data on disk
- According to TI capacity share
- 60MB/s (when data present) => SDTY pp data taking scenario



- Collection of data from experiment RAW and DAQ/DCS/HLT databases
 - Calibration software running in DAQ framework
 - External conditions sources – LHC machine parameters, magnet status, etc...
 - Publication of data through Shuttle framework in OCDB
 - Special portion of the AliEn FC for fast MD search
 - Data itself are plain root files
 - Data replicated to TIs & accessed by reco & some analysis tasks

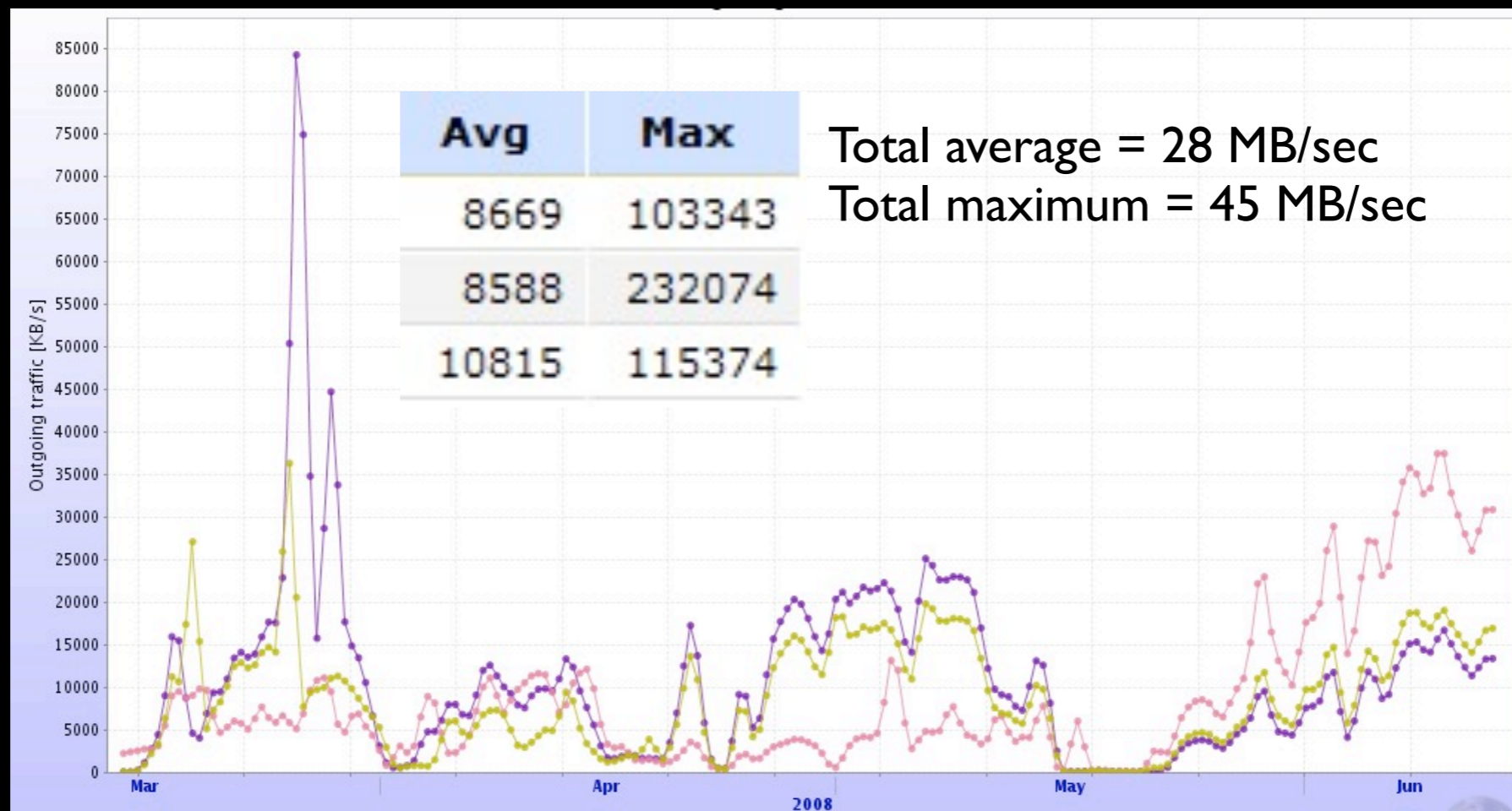


Conditions data

- Collection of data from experiment RAW and DAQ/DCS/HLT databases
- Calibration software running in DAQ framework
- External conditions sources – LHC machine parameters, magnet status, etc...
- Publication of data through Shuttle framework in OCDB
 - Special portion of the AliEn FC for fast MD search
 - Data itself are plain root files
 - Data replicated to TIs & accessed by reco & some analysis tasks

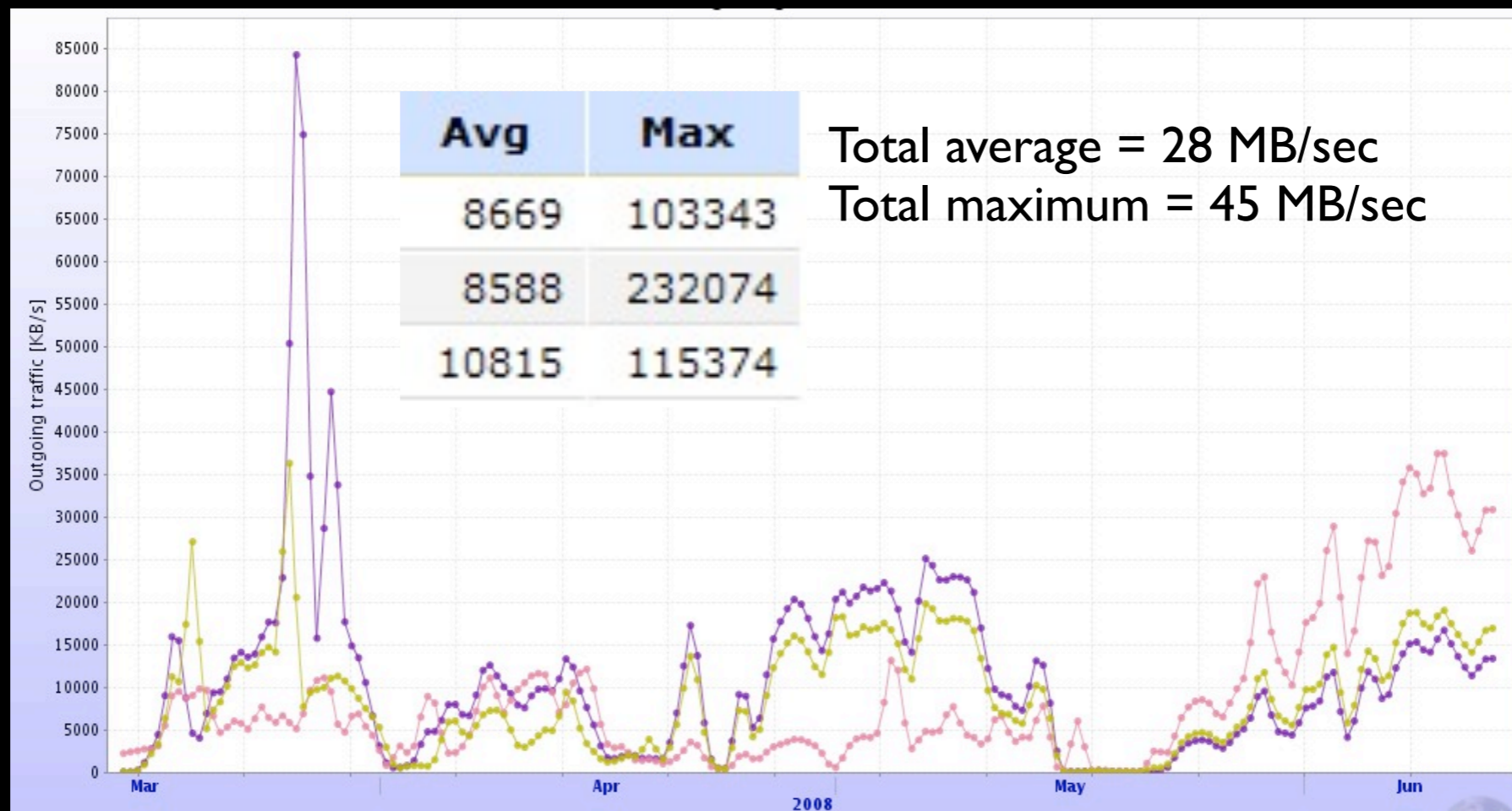


- OCDB in operation since mid-2007
- Stress- tested with Grid jobs
- Including failover for non-accessible replicas



Conditions data (2)

- OCDB in operation since mid-2007
- Stress- tested with Grid jobs
- Including failover for non-accessible replicas



Jobs efficiency (cpu time / wall time)

Farm	Last value	Min	Avg	Max
CERN-L	67.98	0	71.42	101.3
CERN_gLite	63.75	0	69.39	100
Total	65.86		70.4	

Running Jobs



Data processing

Jobs efficiency (cpu time / wall time)

Farm	Last value	Min	Avg	Max
CERN-L	67.98	0	71.42	101.3
CERN_gLite	63.75	0	69.39	100
Total	65.86		70.4	

Running Jobs



- ‘first few days of LHC running’ events / cycle
- High intensity production – 10 days for all & fast data re-processing
- Analysis ongoing (Grid and CAF)

PDC 08/LHC08b6	First physics pp, Phojet, No field, 10 TeV	Completed
PDC 08/LHC08b5	First physics pp, Pythia6, No field, 900GeV	Completed
PDC 08/LHC08b4	First physics pp, Pythia6, No field, 10TeV	Completed
PDC 08/LHC08b3	First physics pp, Phojet, 5kG, 10TeV	Completed
PDC 08/LHC08b2	First physics pp, Pythia6, 5kG, 900GeV	Completed
PDC 08/LHC08b1	First physics pp, Pythia6, 5kG, 10TeV	Completed



MC for first physics

- 'first few days of LHC running' events / cycle
- High intensity production – 10 days for all & fast data re-processing
- Analysis ongoing (Grid and CAF)

PDC 08/LHC08b6	First physics pp, Phojet, No field, 10 TeV	Completed
PDC 08/LHC08b5	First physics pp, Pythia6, No field, 900GeV	Completed
PDC 08/LHC08b4	First physics pp, Pythia6, No field, 10TeV	Completed
PDC 08/LHC08b3	First physics pp, Phojet, 5kG, 10TeV	Completed
PDC 08/LHC08b2	First physics pp, Pythia6, 5kG, 900GeV	Completed
PDC 08/LHC08b1	First physics pp, Pythia6, 5kG, 10TeV	Completed



RAW Production Cycles

Production	Description	Status	Run Range	Recorded chunks	Processed chunks	Comments
PDC 08/LHC08b	LHC08b global partition	Running	34784 - 42365	26,085	25,452	RUN III cosmics
PDC 08/LHC08a	LHC08a global partition - MUON reco	Running	21392 - 26024	8,336	166	RUN II cosmics (Muon)
PDC 08/LHC08a	LHC08a global partition	Completed	18047 - 26042	37,064	33,847	Feb 08 cosmics

LPM		Only LPM jobs	AliEn	Detector	Software versions	
Run#	Chunks	Processed	Date/PID	Partition	ROOT	ALIROOT
				- All -	- All -	- All -

37040	140	136	97.1%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
37028	347	336	96.8%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
36148	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
35366	10	9	90%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
35358	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
34784	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
TOTAL	26,085	25,452	97.6%	67 jobs			

- Run lists/priority from detector groups
- Quasi-online (pass 1 @ T0) production is ongoing
- Target: 48 hours after the data collection

RAW data reconstruction

RAW Production Cycles							
Production	Description	Status	Run Range	Recorded chunks	Processed chunks	Comments	
PDC 08/LHC08b	LHC08b global partition	Running	34784 - 42365	26,085	25,452	RUN III cosmics	
PDC 08/LHC08a	LHC08a global partition - MUON reco	Running	21392 - 26024	8,336	166	RUN II cosmics (Muon)	
PDC 08/LHC08a	LHC08a global partition	Completed	18047 - 26042	37,064	33,847	Feb 08 cosmics	

LPM		Only LPM jobs	AliEn	Detector	Software versions		
Run#	Chunks	Processed	Date/PID	Partition	ROOT	ALIROOT	
				- All -	- All -	- All -	
37040	140	136	97.1%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
37028	347	336	96.8%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
36148	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
35366	10	9	90%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
35358	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
34784	2	2	100%	24 Jun 2008	LHC08b	v5-19-04	v4-13-Rev-03
TOTAL	26,085	25,452	97.6%	67 jobs			

- Run lists/priority from detector groups
- Quasi-online (pass 1 @ T0) production is ongoing
- Target: 48 hours after the data collection

- Critical data immediately replicated to CAF@CERN & T2s with large detector communities
- Fast lane analysis on CAF
- User Grid analysis is ongoing in parallel
- ~47 detector experts analysing data on the Grid



Analysis

- Critical data immediately replicated to CAF@CERN & T2s with large detector communities
- Fast lane analysis on CAF
- User Grid analysis is ongoing in parallel
- ~47 detector experts analysing data on the Grid



- <http://twiki.cern.ch/twiki/bin/view/ALICE/CCRC08>

Rank	Definition	Max downtime (hrs)
10	Critical	2
7	Serious disruption	8
5	Reduced efficiency	12

Excellent support from site experts

Service	Rank	Comment
Site VO boxes	10	If restored within 2 hours - no loss of jobs
CASTOR2+xrootd@T0	10	Same as above
MSS@T1s	5	
FTS T0->T1	7	Will affect primarily second pass reconstruction at T1s
gLite WMS or RB	5	Failover mechanism 1/3 used, efficiency will be affected if all 3 fail
PROOF@CAF	5	Especially relevant during daytime

- Critical services – fully identified and tested, SLA@T1s OK, T2s – on a site by site basis (generally OK)

The Critical Services

- <http://twiki.cern.ch/twiki/bin/view/ALICE/CCRC08>

Rank	Definition	Max downtime (hrs)
10	Critical	2
7	Serious disruption	8
5	Reduced efficiency	12

Excellent support from site experts

Service	Rank	Comment
Site VO boxes	10	If restored within 2 hours - no loss of jobs
CASTOR2+xrootd@T0	10	Same as above
MSS@T1s	5	
FTS T0->T1	7	Will affect primarily second pass reconstruction at T1s
gLite WMS or RB	5	Failover mechanism 1/3 used, efficiency will be affected if all 3 fail
PROOF@CAF	5	Especially relevant during daytime

- Critical services – fully identified and tested, SLA@T1s OK, T2s – on a site by site basis (generally OK)

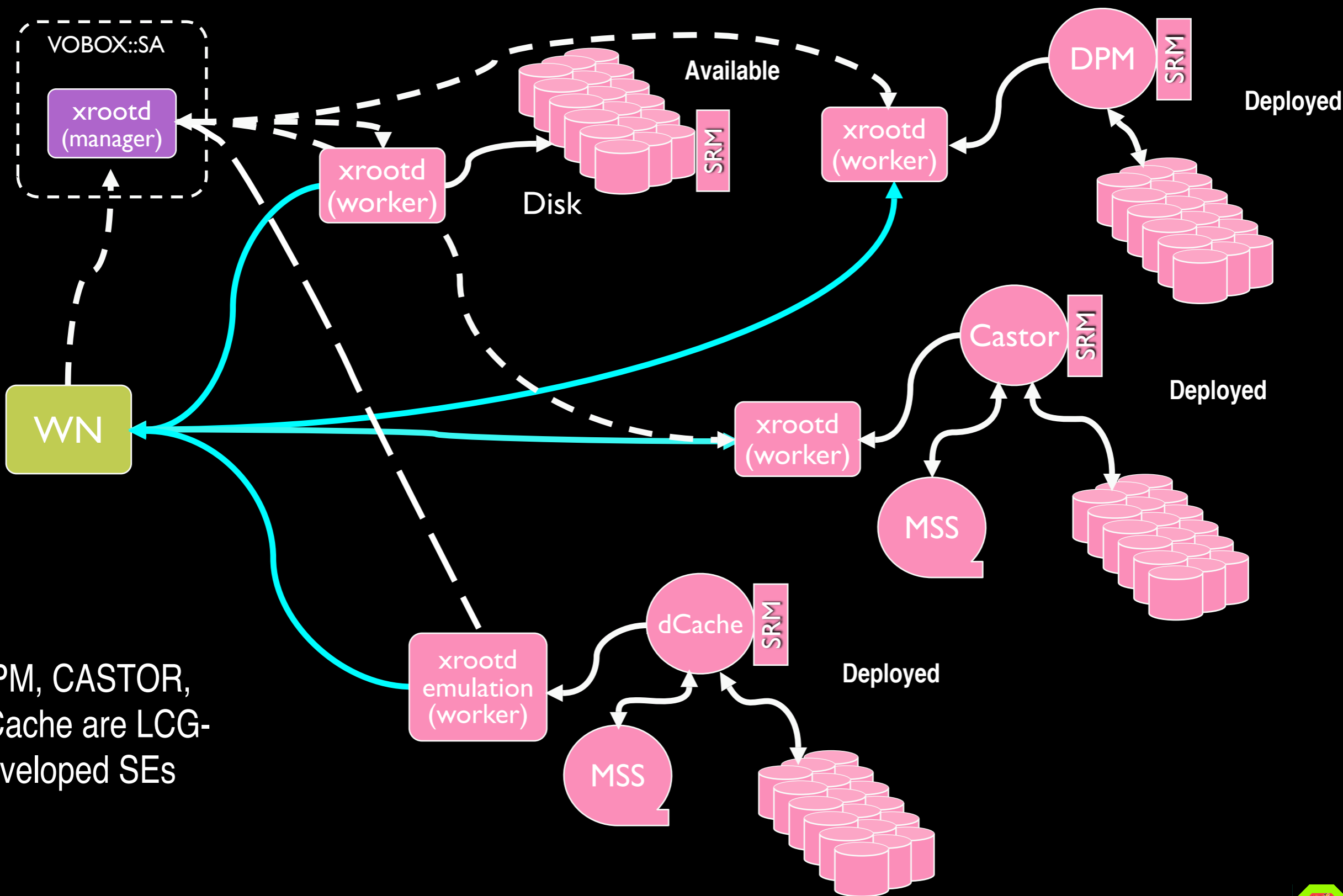
- **Strict weekly release schedule for AliRoot**
 - Coordinated with detector groups
 - Essential for integration and validation of detector algorithms in DAQ/HLT systems (shared code)
 - Some effort to explain to the community that “a release is a release”
- **Simultaneous deployment on the Grid for RAW production and ESD analysis**
- **Titanic effort to remain below 2GB is paying off**



Application software

- Strict weekly release schedule for AliRoot
 - Coordinated with detector groups
 - Essential for integration and validation of detector algorithms in DAQ/HLT systems (shared code)
 - Some effort to explain to the community that “a release is a release”
- Simultaneous deployment on the Grid for RAW production and ESD analysis
- Titanic effort to remain below 2GB is paying off

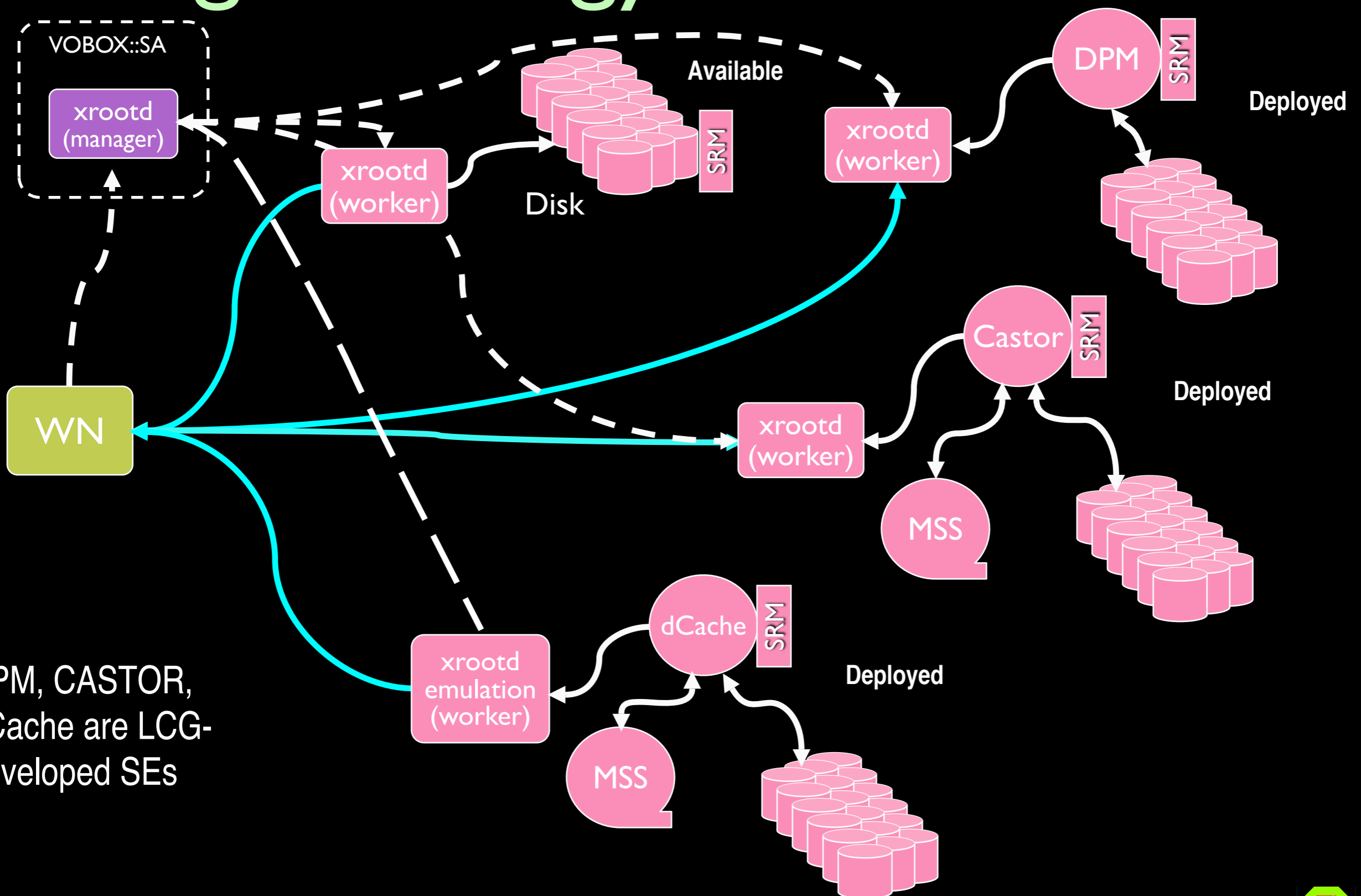




DPM, CASTOR, dCache are LCG-developed SEs



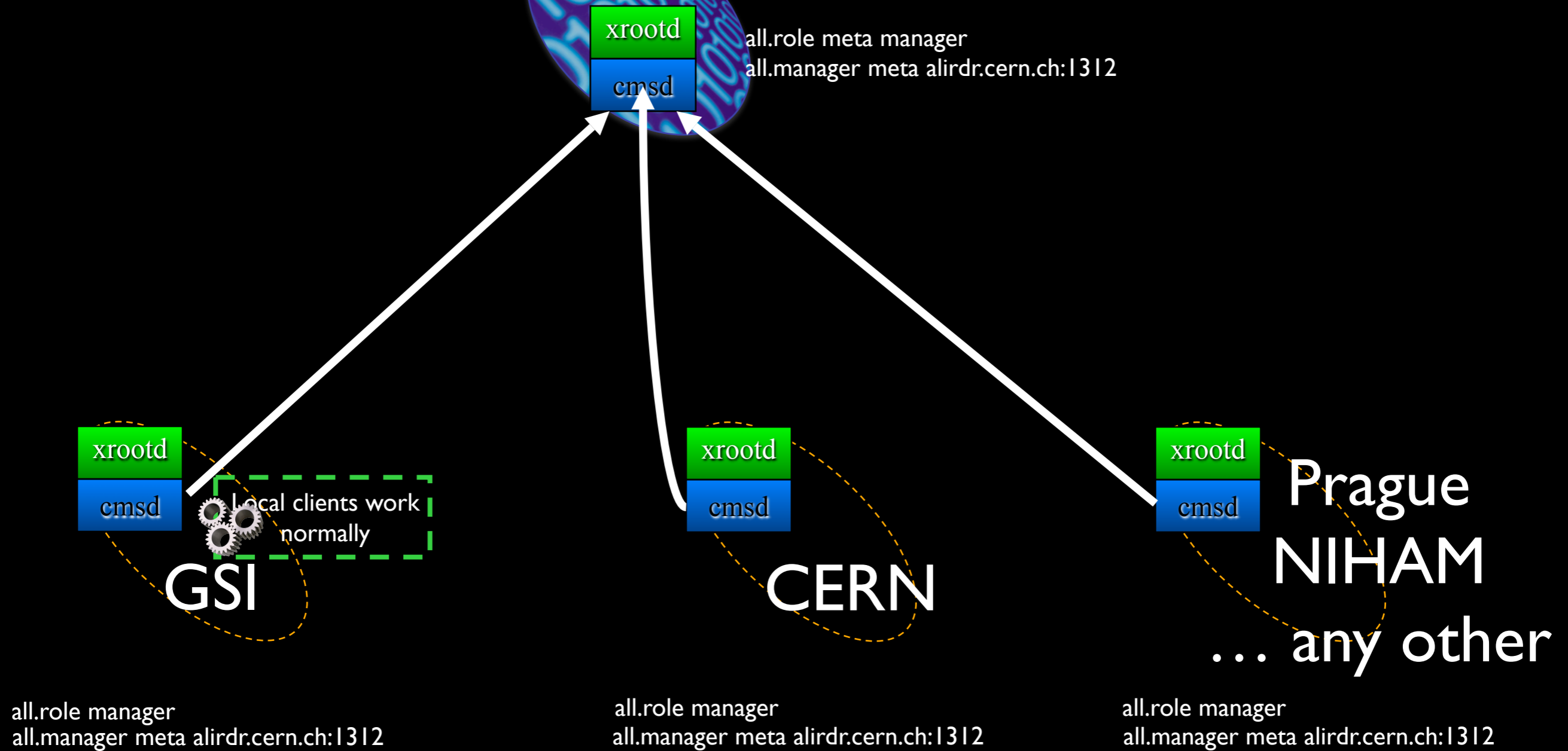
Storage strategy



DPM, CASTOR, dCache are LCG-developed SEs

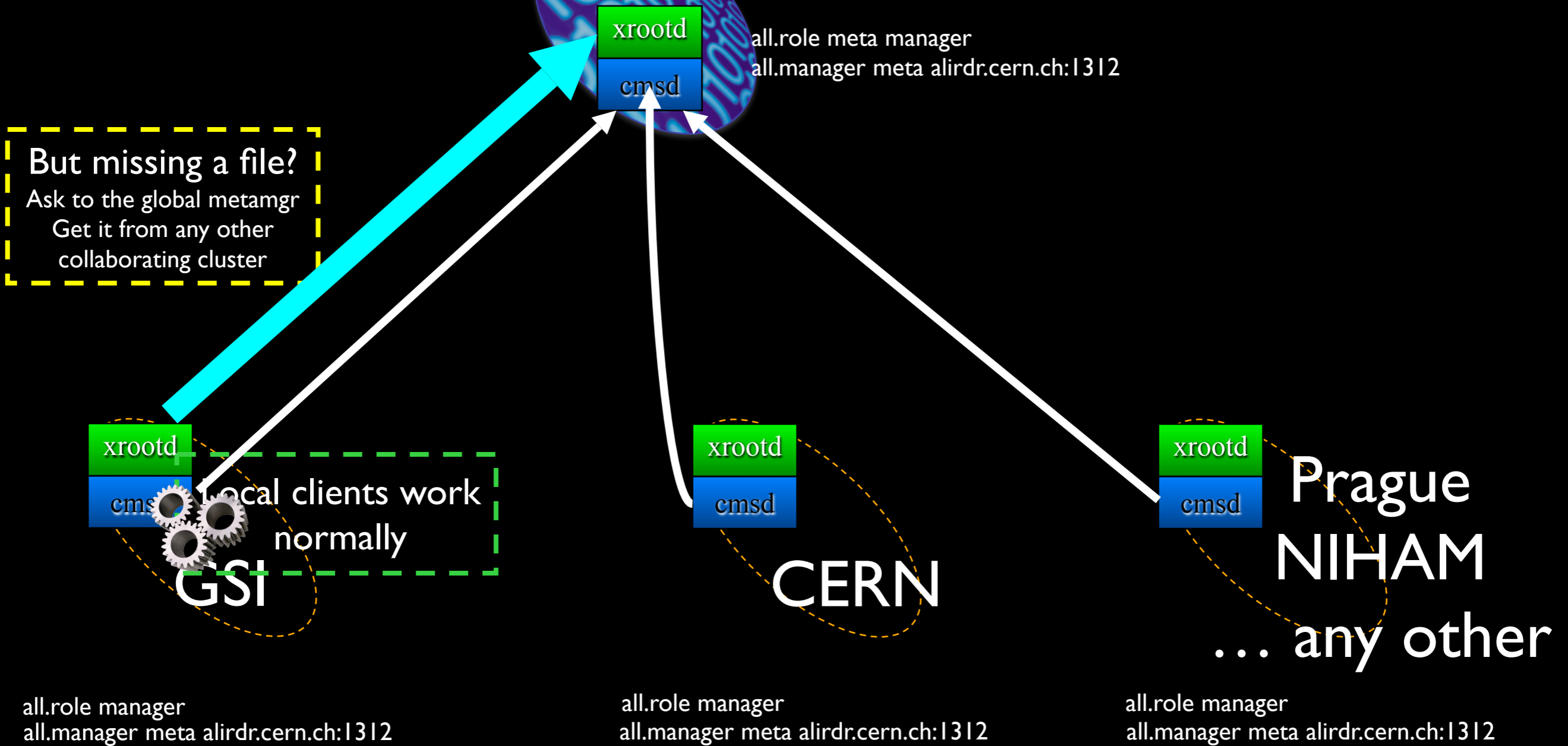


ALICE global redirector



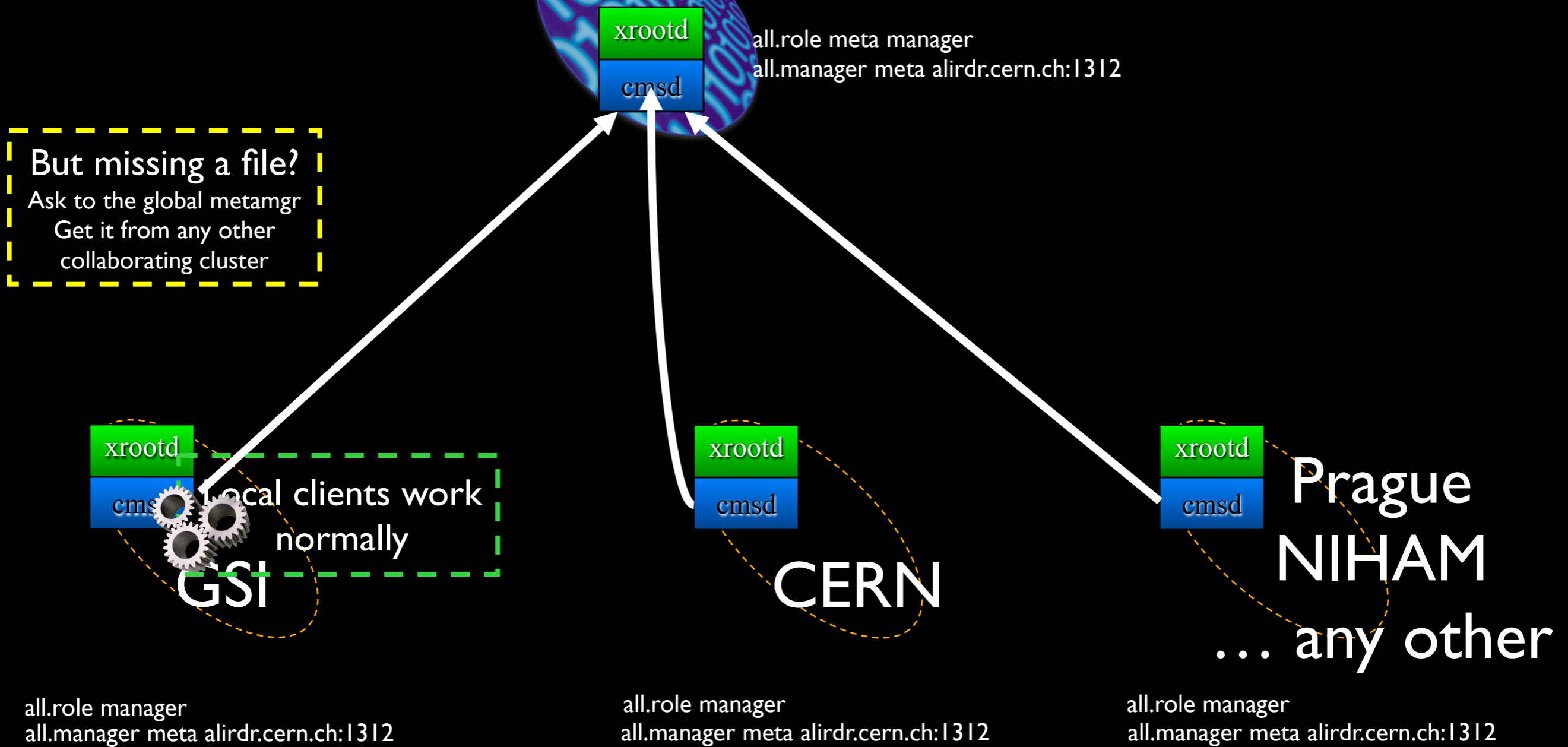
Storage tech: new trends

ALICE global redirector

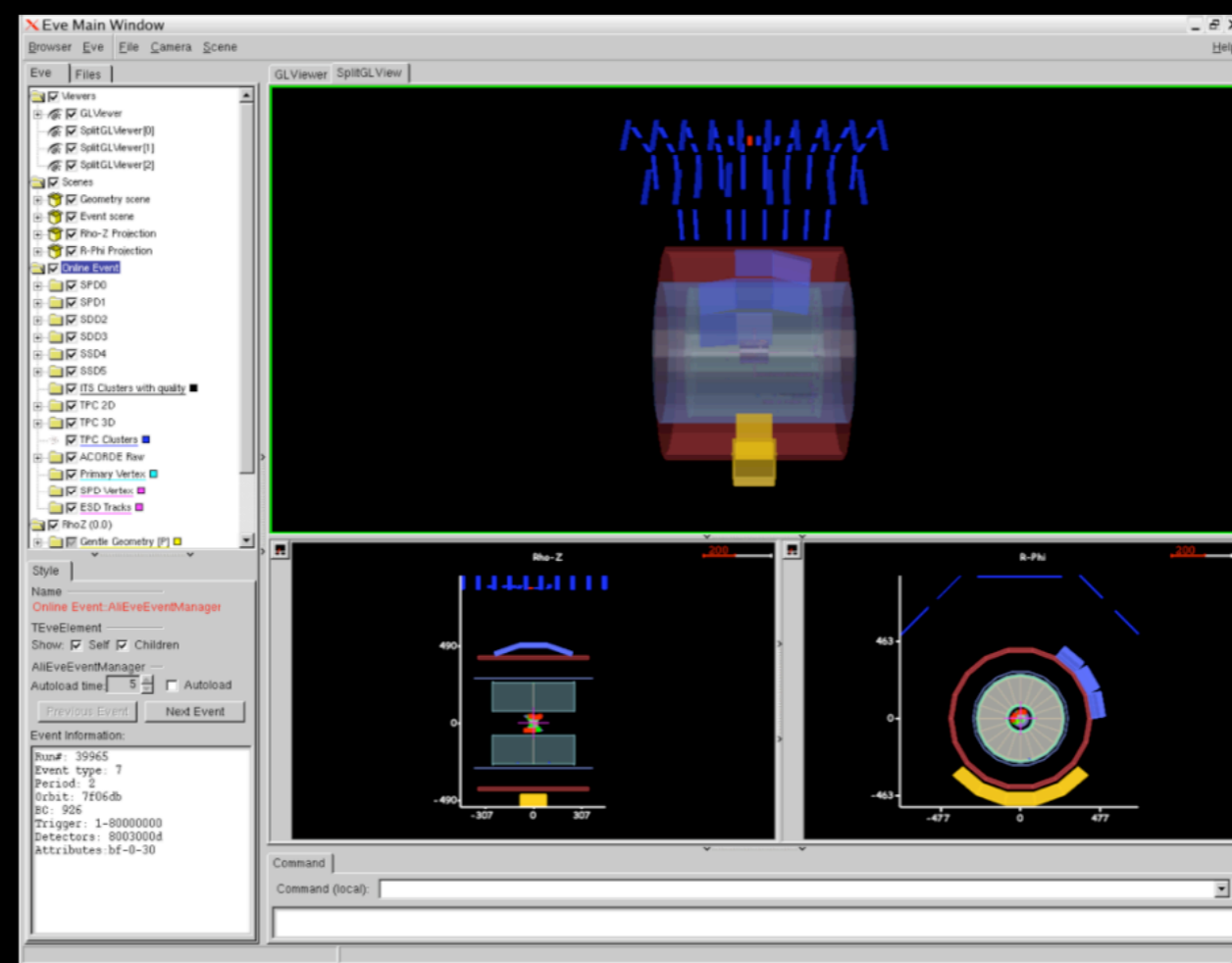
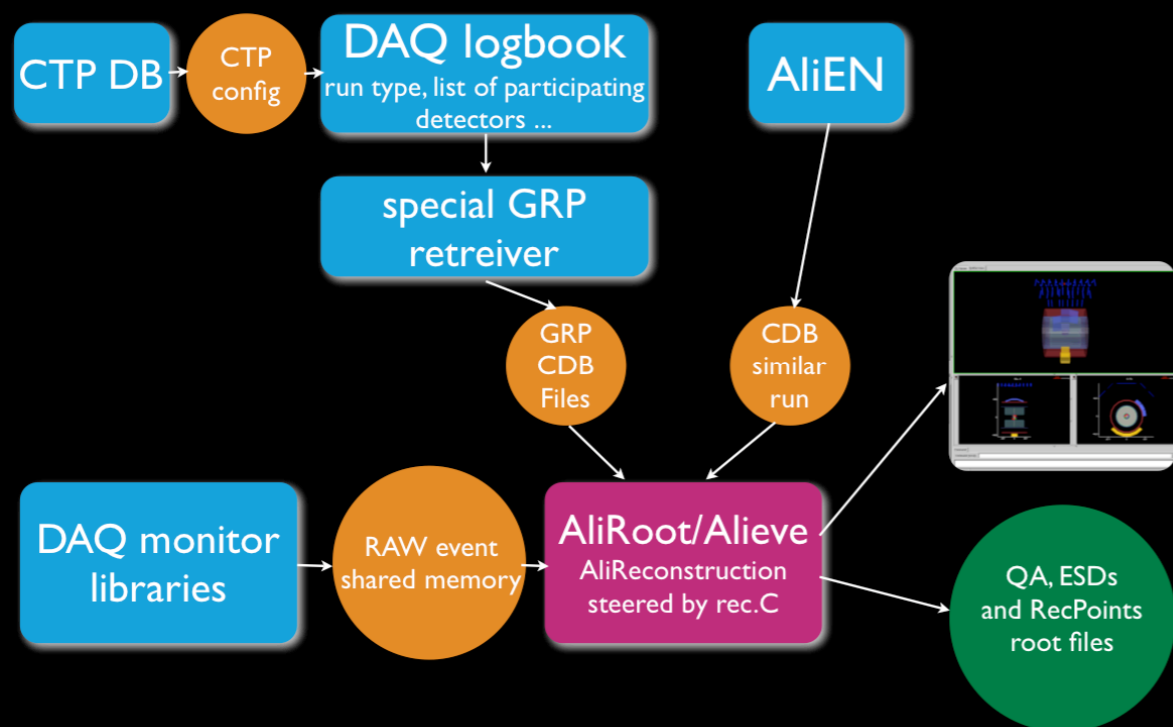
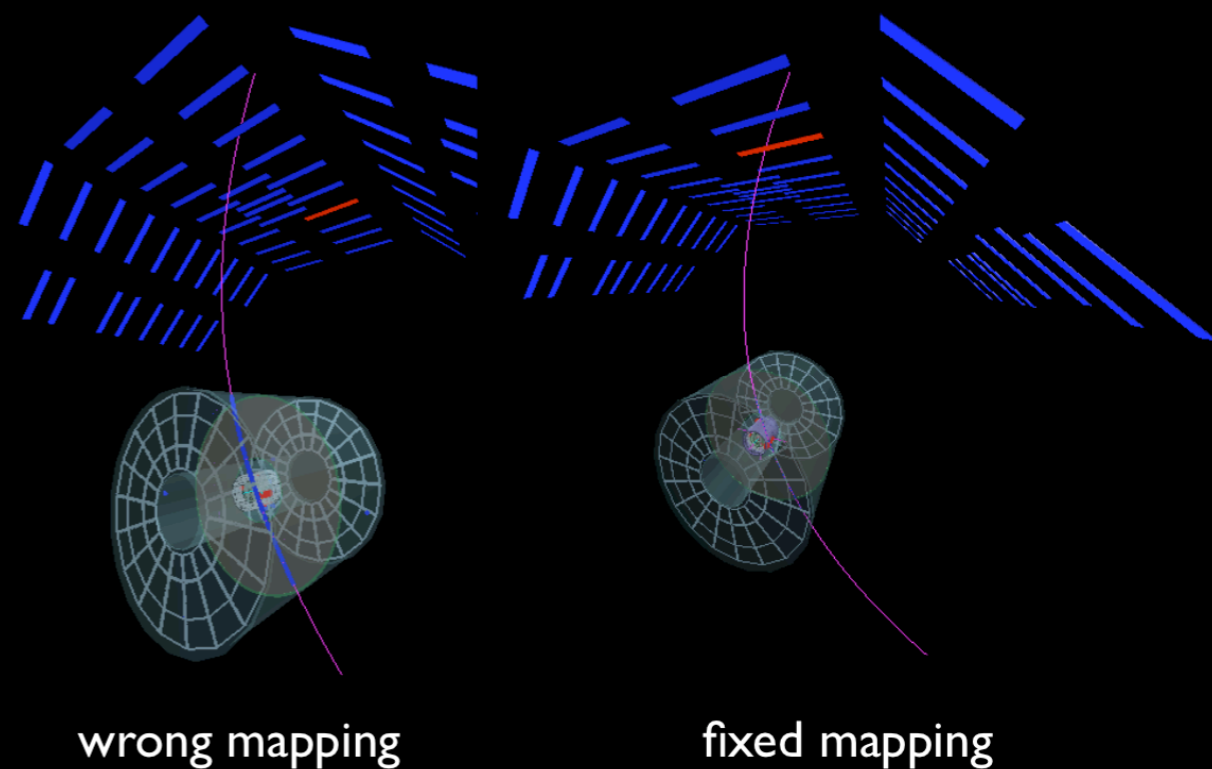


Storage tech: new trends

ALICE global redirector

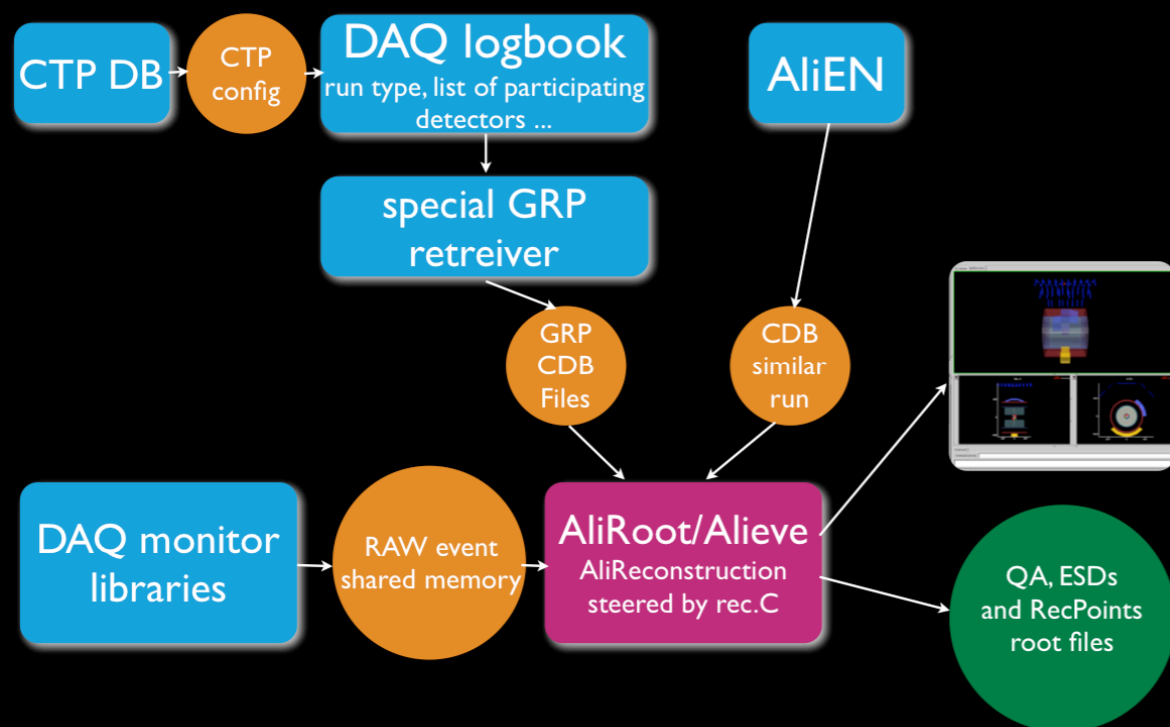
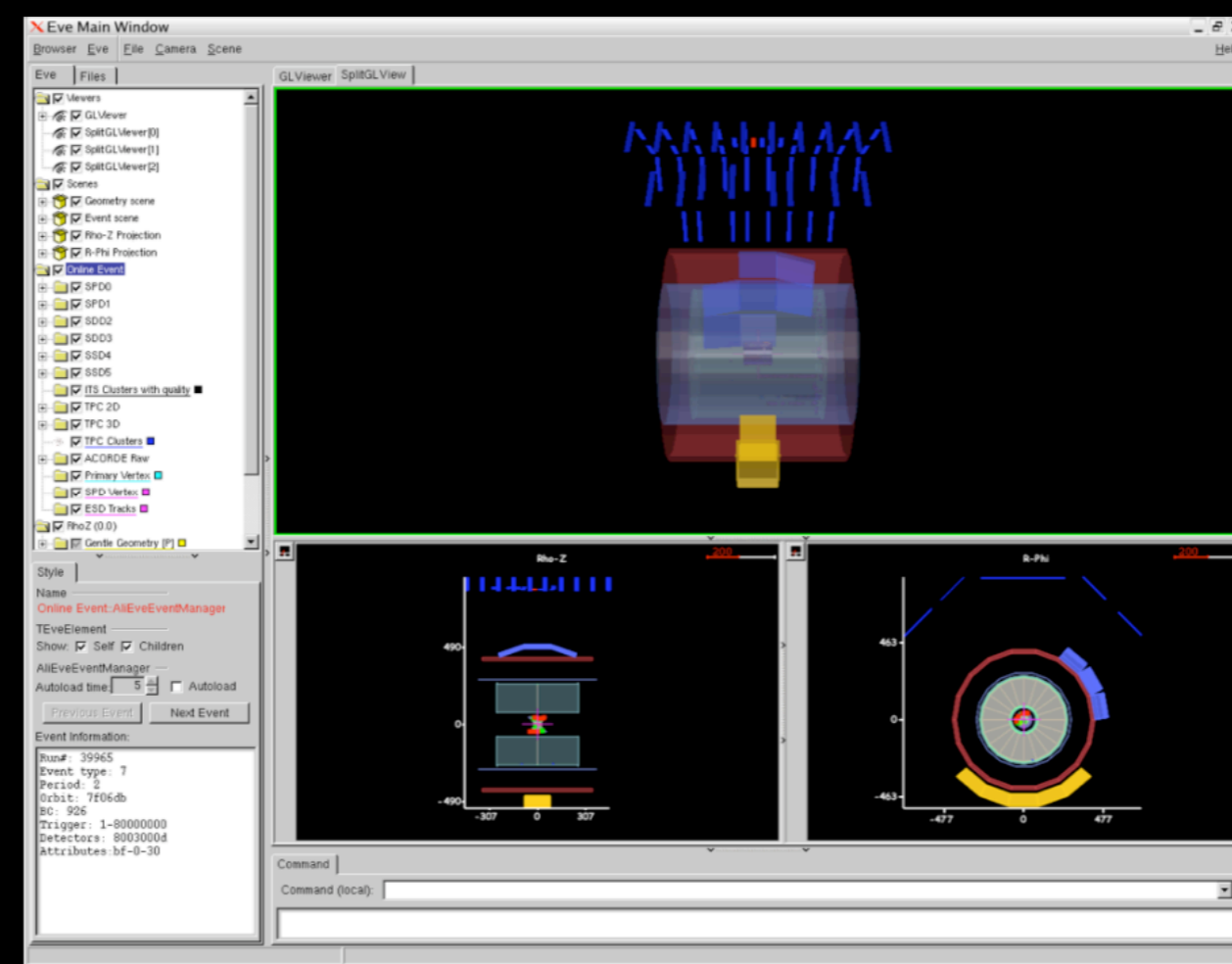
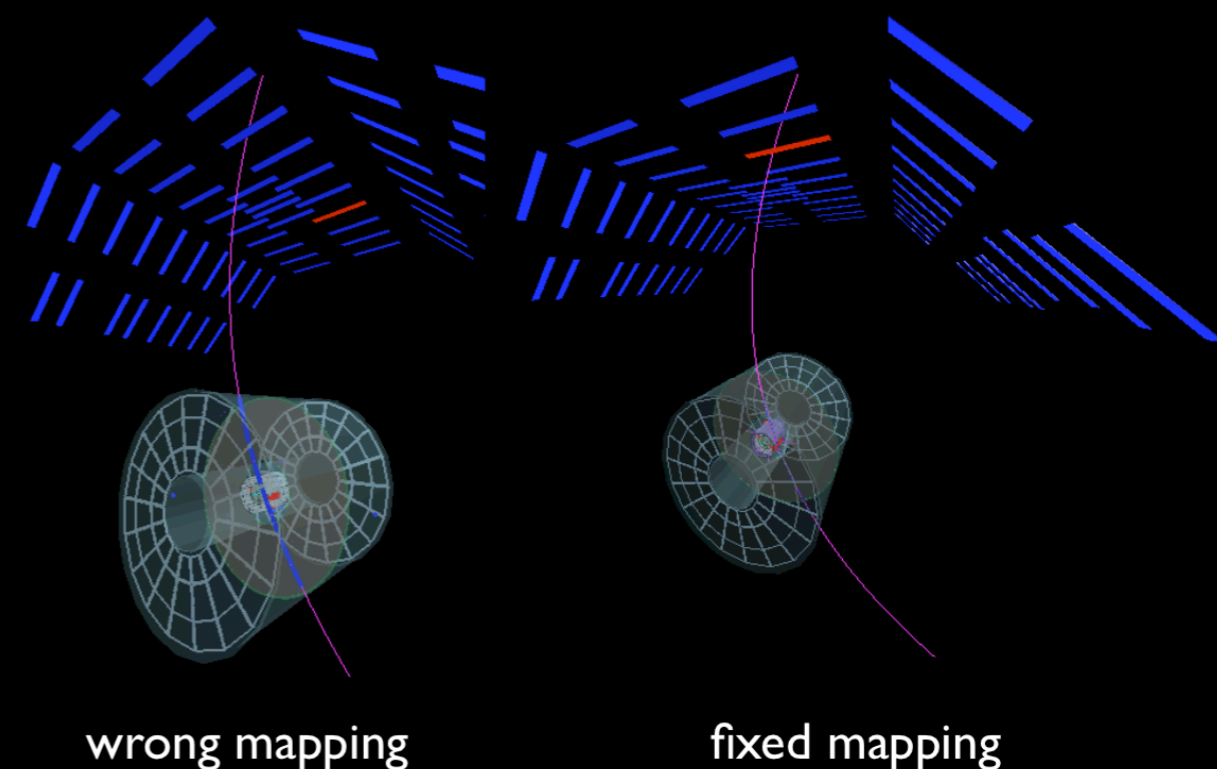


- Recent development
- Very useful for high-level QA and debugging
- Integrated in the AliEVE event display
- Full Offline code sampling events directly from DAQ memory

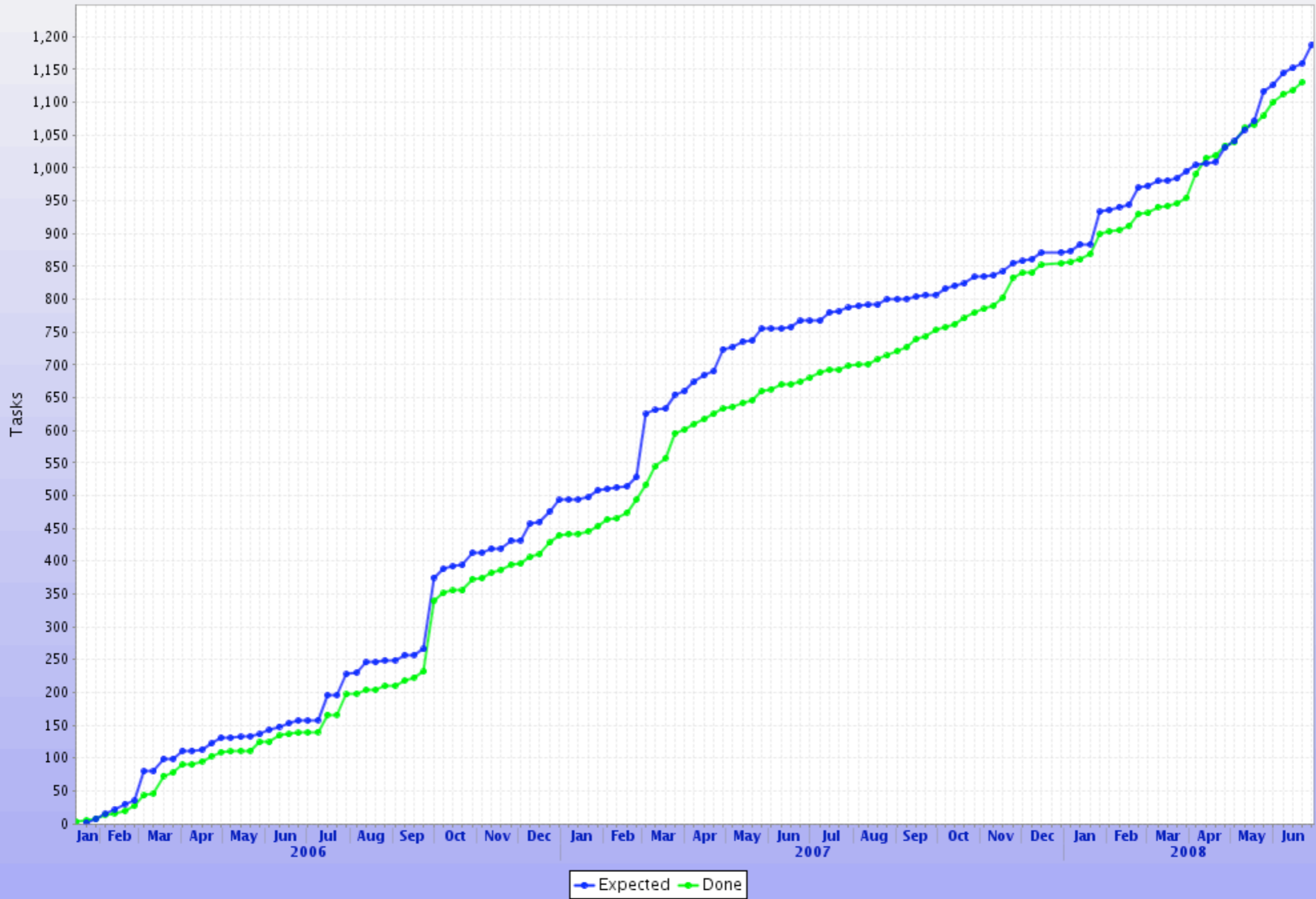


Quasi-online reco

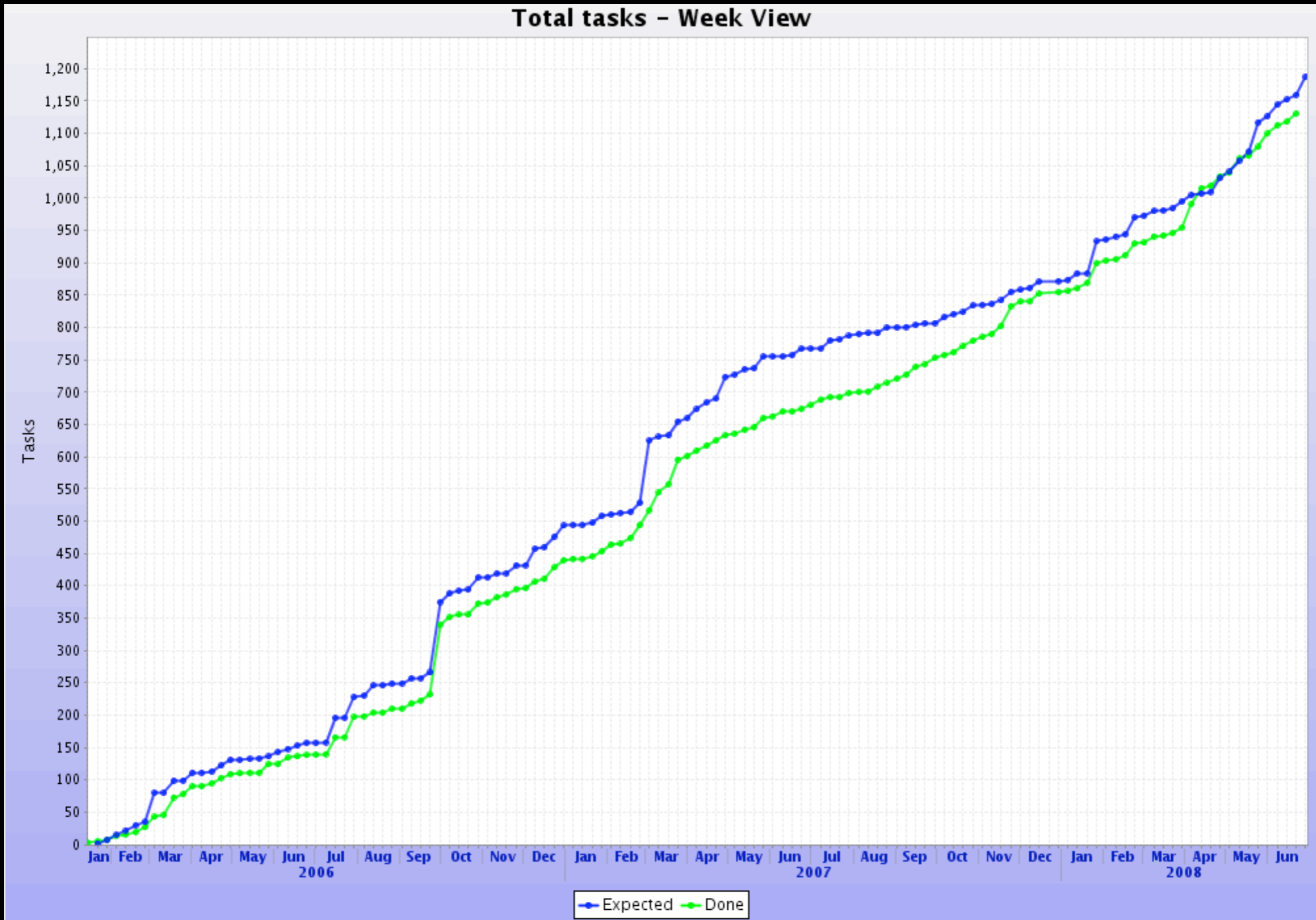
- Recent development
- Very useful for high-level QA and debugging
- Integrated in the AliEVE event display
- Full Offline code sampling events directly from DAQ memory



Total tasks - Week View



Planning



Parameter	Now	CTDR	Ratio
pp RAW	1.0MB	0.2MB	5
Pb RAW	35MB	13.8MB	2.5
ESD pp	0.04MB	0.04MB	1.0
ESD Pb	6.3MB	3.0MB	2.1
AOD pp	5kB	16kB	0.3
AOD Pb	1.3MB	0.34MB	3.8
Reco pp	6.8s	6.5s	1.0
Reco Pb	800s	810s	1.0

Pledged by external sites versus required (new LHC schedule) all											
		2008		2009		2010		2011		2012	
		T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
CPU	Requirement (MSI2K)	10.1	12.5	19.9	14.3	23.5	25.0	30.5	32.5	39.7	42.2
	Missing %	-33%	-28%	-37%	-7%	-32%	-27%	-48%	-44%	-60%	-57%
Disk	Requirement (PB)	3.9	1.7	6.8	4.0	12.0	4.3	16.8	5.6	22.7	7.3
	Missing %	-31%	6%	-30%	-10%	-35%	9%	-53%	-17%	-65%	-36%
MS	Requirement (PB)	5.7	-	12.4	-	19.8	-	27.3	-	34.1	-
	Missing %	-40%	-	-36%	-	-34%	-	-52%	-	-62%	-



Resource overview

Parameter	Now	CTDR	Ratio
pp RAW	1.0MB	0.2MB	5
Pb RAW	35MB	13.8MB	2.5
ESD pp	0.04MB	0.04MB	1.0
ESD Pb	6.3MB	3.0MB	2.1
AOD pp	5kB	16kB	0.3
AOD Pb	1.3MB	0.34MB	3.8
Reco pp	6.8s	6.5s	1.0
Reco Pb	800s	810s	1.0

Pledged by external sites versus required (new LHC schedule) all											
		2008		2009		2010		2011		2012	
		T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
CPU	Requirement (MSI2K)	10.1	12.5	19.9	14.3	23.5	25.0	30.5	32.5	39.7	42.2
	Missing %	-33%	-28%	-37%	-7%	-32%	-27%	-48%	-44%	-60%	-57%
Disk	Requirement (PB)	3.9	1.7	6.8	4.0	12.0	4.3	16.8	5.6	22.7	7.3
	Missing %	-31%	6%	-30%	-10%	-35%	9%	-53%	-17%	-65%	-36%
MS	Requirement (PB)	5.7	-	12.4	-	19.8	-	27.3	-	34.1	-
	Missing %	-40%	-	-36%	-	-34%	-	-52%	-	-62%	-



- Most of the people in the CORE Offline team are PJAS and Fellows
 - Some of the key PJAS will have to leave end 2008
 - Fellows are now deviating from computing to physics
- The staffing situation of core Offline remains a risk factor
- INFN in-kind contributions are highly appreciated, but it is the only FA to do so
- Strong dependence on some dedicated IT / LCG / EGEE support
 - Already slashed by more than half in the last year!



Staffing situation

- Most of the people in the CORE Offline team are PJAS and Fellows
 - Some of the key PJAS will have to leave end 2008
 - Fellows are now deviating from computing to physics
- The staffing situation of core Offline remains a risk factor
- INFN in-kind contributions are highly appreciated, but it is the only FA to do so
- Strong dependence on some dedicated IT / LCG / EGEE support
 - Already slashed by more than half in the last year!



- AliEn, xrootd and proof support
 - Excellent, but out of our control (IT, LCG, EGEE and SFT)
 - It has to continue at least at this level (NO plan B!)
- Staffing of Core offline
- Computing resources



Risk factors

- AliEn, xrootd and proof support
 - Excellent, but out of our control (IT, LCG, EGEE and SFT)
 - It has to continue at least at this level (NO plan B!)
- Staffing of Core offline
- Computing resources

- Past months activity centred on
 - Detector commissioning and upgrades of the ALICE Grid installations
 - Collection, replication and processing of calibration and alignment data from the experiment
 - Testing and tuning of Grid services ongoing – generally smooth operation
 - Deployment of xrootd-enabled storage at T2s and extension of capacity at T0/T1s
- June and beyond period
 - Full experiment data taking and reconstruction at T0/T1s
 - MC production and analysis at T2s
- We are (reasonably) ready for the initial LHC data



Summary

- Past months activity centred on
 - Detector commissioning and upgrades of the ALICE Grid installations
 - Collection, replication and processing of calibration and alignment data from the experiment
 - Testing and tuning of Grid services ongoing – generally smooth operation
 - Deployment of xrootd-enabled storage at T2s and extension of capacity at T0/T1s
- June and beyond period
 - Full experiment data taking and reconstruction at T0/T1s
 - MC production and analysis at T2s
- We are (reasonably) ready for the initial LHC data



