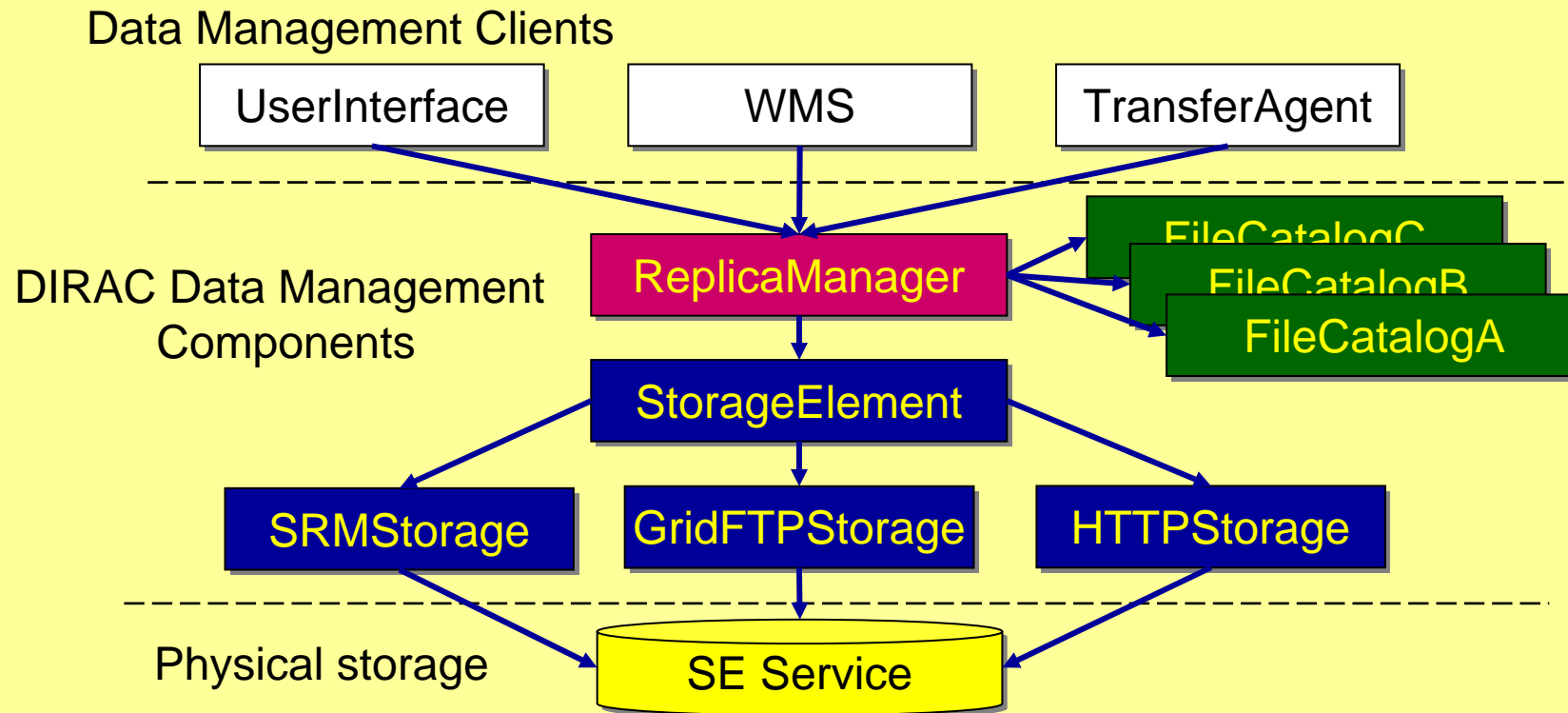


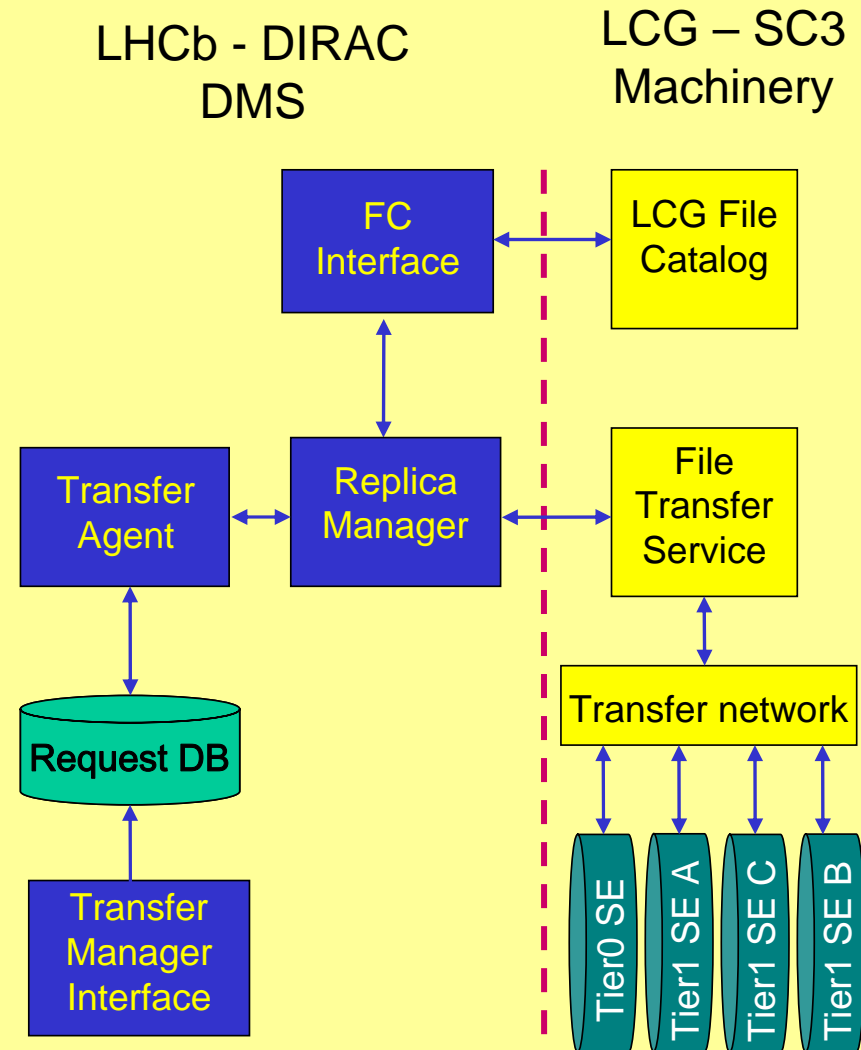
Castor Review – LHCb Experiences

- Use of Castor2
 - ▶ Introduction during SC3
 - ▶ ...till current LHCb instance
- Experience during staging operations
- LHCb Migration to Castor2
- Users' experience

- Data Management handled by DIRAC
 - ▶ LHCb specific grid software



- Central Data Movement model based at CERN.
- Replication of 8TB of data to 6 Tier1 sites
- DIRAC Transfer Agent submit/monitor FTS jobs
 - ▶ Removed from actual interaction with underlying storages
- This lead to a problem...



- Data to be used for SC3 was stored on tape
- FTS issues 'SRM Get' triggering stage
 - ▶ Operation longer than timeout of FTS Agent
 - ▶ 'Failed on SRM get: SRM getRequestStatus timed out on get'
- DIRAC retry policy can cope but:
 - ▶ Fills transfer slots
 - ▶ Dramatically reduces effective bandwidth
- Decide to pre-stage to make progress
 - ▶ ~ 50,000 files in dedicated Castor2 pools
 - ▶ ~ 10TB used for CERN-T1 transfers

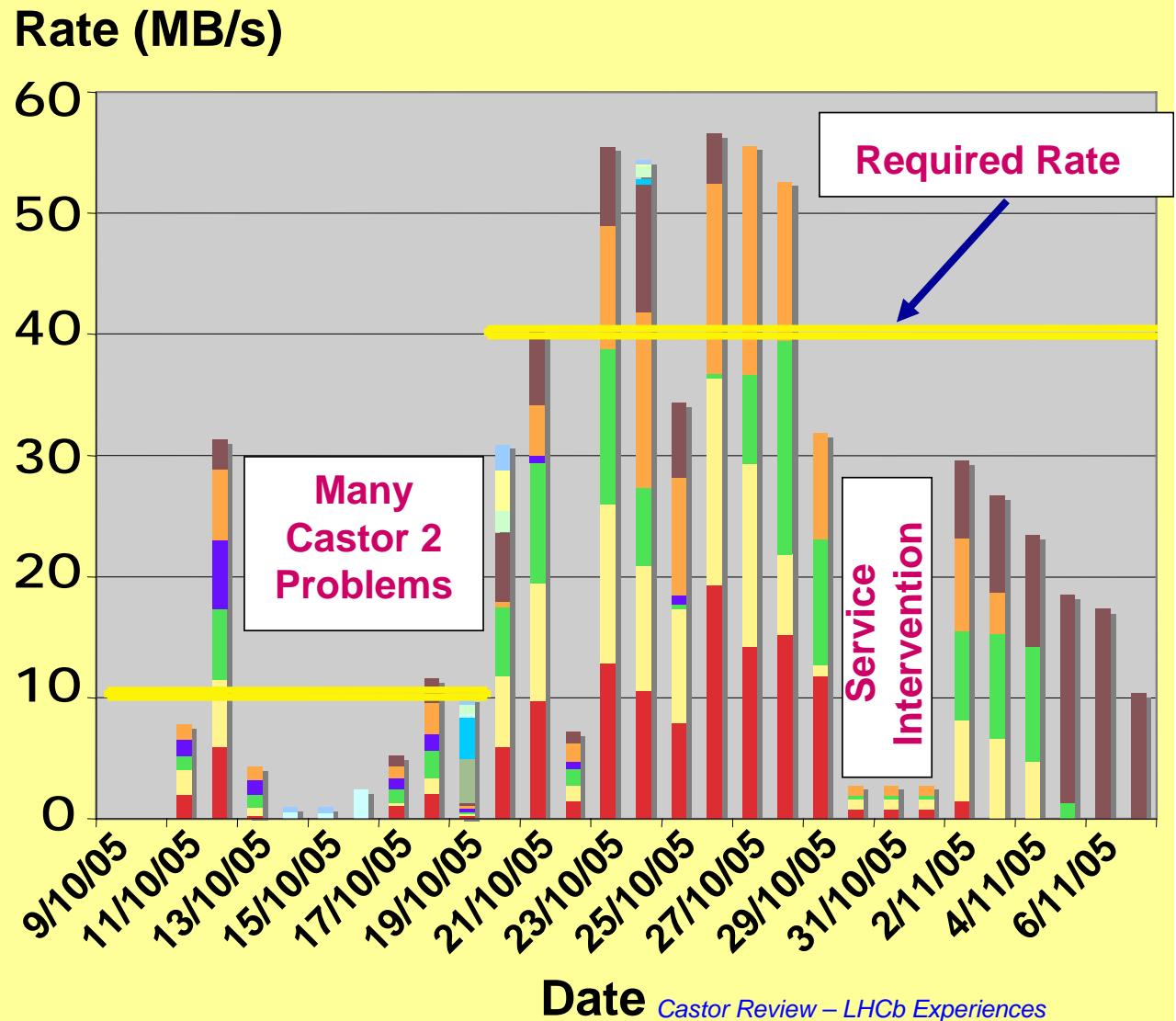
- Castor2 pools set up ~12/10/05
 - ▶ Began staging data into pools
 - ▶ 13-14/10
 - ▶ CASTOR2 Oracle instance problem blocking scheduling
 - ▶ 15-16/10
 - ▶ LSF hang -> DB load too high -> 'NO_Contact' alarm
 - ▶ 17/10
 - ▶ Stager dead
 - ▶ 19/10
 - ▶ Intervention

■ Pre-intervention:

- ▶ Many problems
- ▶ System felt unstable

■ Post-intervention:

- ▶ Stable staging and access
- ▶ Peak rates out of Castor at double SC3 target rate.

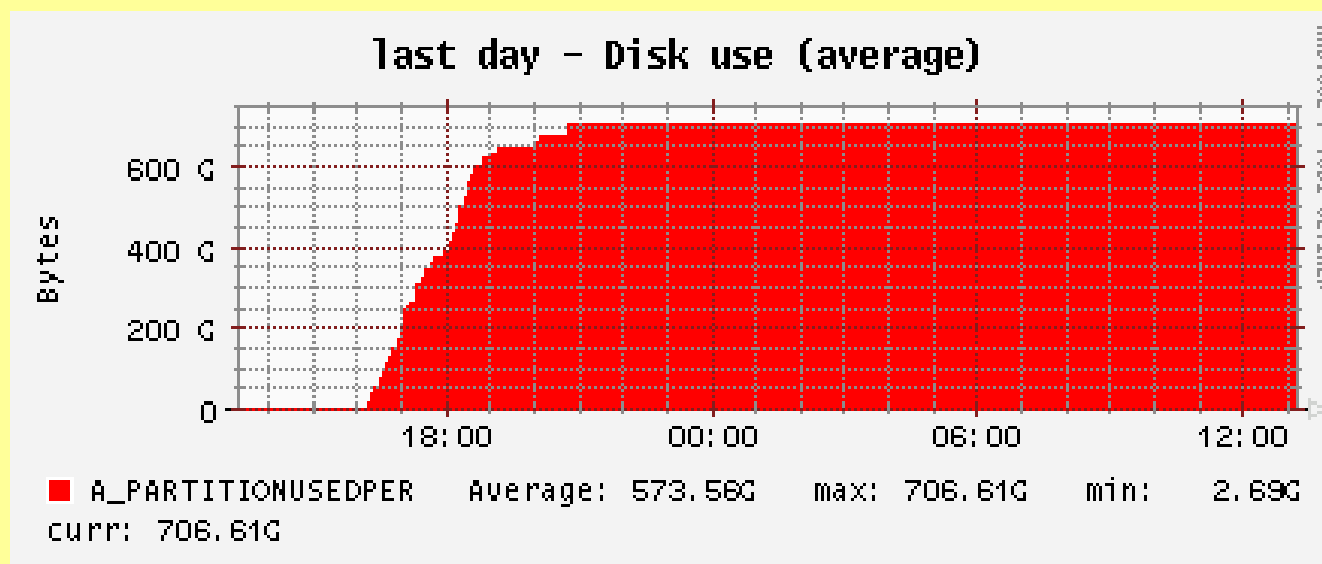


- Several large stage requests during SC3
 - ▶ Observed some requests being returned:
 - ▶ 'Error Internal error'
 - ◆ Request effectively lost
 - ▶ Receive confirmation of receipt of a portion of files
 - ◆ i.e. submit stage request for 100 files
 - ◆ Stager returns receipt of a fraction:
 - 'Received 58 responses'
 - ▶ Conducted subsequent stages to trace behaviour

- 1138 analysis files staged to lhcb Castor2 instance, *lhcbdata* pool (4/4/06)
 - ▶ stager_get with 50 files
 - ▶ server response -> issue next request
 - ▶ ~15 mins for submission of all files
- Only 2 files encountered problems
 - ▶ 'unexpected RFIO error...not retried and the request failed'
 - ▶ Retry logic in place as result
 - ▶ 'tape recall on way to exit and too late for new file'
 - ▶ Exit handling reviewed

- The lemon 'disk use' plot below shows progress of stage
 - ▶ First 600G staged at constant rate
 - ▶ ~2.5hours
 - ▶ ~240G/hr or ~400files/hr

- After this rate decreases
- Queues not saturated?



- 10,000 RAW files staged to lhcb Castor2 instance, *wan* pool (26/5/06)
 - ▶ Same method as before
- Mid-stage 'lxfsrc6201' went down
 - ▶ Many files disappeared from stager
 - ▶ When machine returned all requests returned
 - ▶ All (bar two) files staged successfully
- Impressed by resilience and ability to recover

- Contact with the castor group by email
 - ▶ castor-deployment@cern.ch
- Found this to be responsive and helpful
- Problems/bugs taken onboard/fixed quickly

- For new data (i.e. using latest software)
 - ▶ Initial problem as ROOT was not Castor2-aware: fixed since ROOT 5.10.00c (13/4/06)
 - ▶ Gaudi forcing Castor 2 usage since 17/5/06 (next Gaudi release)
- Legacy data (using ROOT 3...)
 - ▶ Private version of ROOT 3 built for Castor2 on 17/5/06
 - ▶ Castor2 usage forced since 30/5/06
- Default mapping of all LHCb users to Castor2 on 8/6/06 (today!)

- Very little experience so far (late migration)
- Users mapped by default to *default* pool
 - ▶ Files disk resident on *lhcbdata*
 - ▶ Small penalty (in time negligible), but bad usage of disk space, due to the copy between *lhcbdata* and *default*
 - ▶ No plan to let users define their mapping manually
 - ▶ Expect that moving to using SRM and related mapping will fix this caveat
- Welcome the distribution of `libshift.so` through LCG (allows consistency in applications)

- Began using Castor2 as part of SC3
 - ▶ Initially many problems
 - ▶ After intervention increased stability
 - ▶ Subsequent stages (mostly) without problems
- Castor team responsive when contacted
- LHCb migration delayed by POOL problem
 - ▶ Users mapped to Castor2 as of today
 - ▶ Their experience will be reported

Questions...?