



LCG Storage Management Workshop, CERN, 7th April 2005

LFC and Fireman Performance Measurements

Caitriana Nicholson / IT-GD, CERN / Glasgow University
Craig Munro / IT-GD, CERN / Brunel University



EGEE
Enabling Grids for
E-science in Europe



- Aims of Testing
- Test Methodology & Setup
- LFC Performance Results
- FiReMan Performance Results
- Conclusions



- Data Challenges of 2004 exposed limitations in LCG Data Management tools
- LCG File Catalog developed to address problems with the RLS
- Suite of tests developed to check the functionality and performance of the LFC
- Comparison required of performance of
 - EDG RLS
 - Globus RLS
 - LFC
 - FiReMan

- Multi-threaded C client program written to test each type of operation (insert, query, delete etc)

```
./create_files -d /grid/dteam/caitriana/insert/  
-f $num_files -t $num_threads
```

- C programs wrapped by Perl scripts
- Typically, each operation performed several thousand times in the client program (`$num_files=3000`) and mean result returned
- Client program called several times from Perl script and mean result taken
- Any entries removed before next test run

- Oracle DB on Xeon 2.4GHz
- PIII 1GHz, 512MB server running 20 threads
- PIII 853MHz, 128MB client with configurable number of threads (single client tests)
- 10 x PIII 1GHz, 512MB clients with configurable number of threads (multiple clients)
- 100 Mb/s LAN
- Insecure LFC

Comparison against published results for insecure catalogues

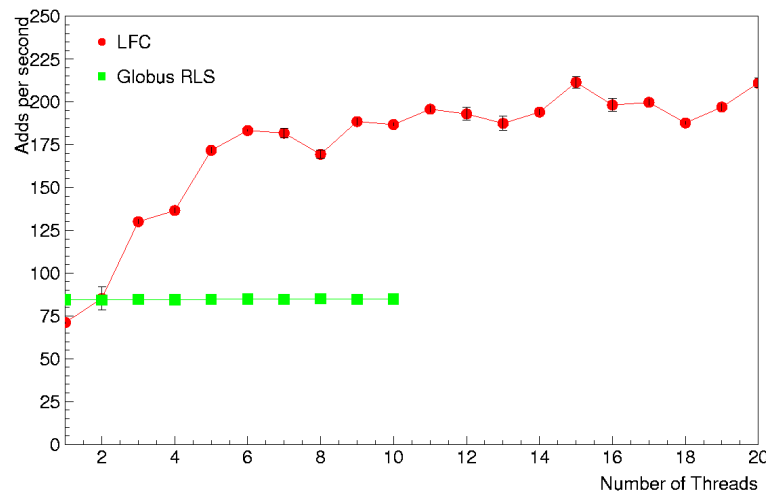
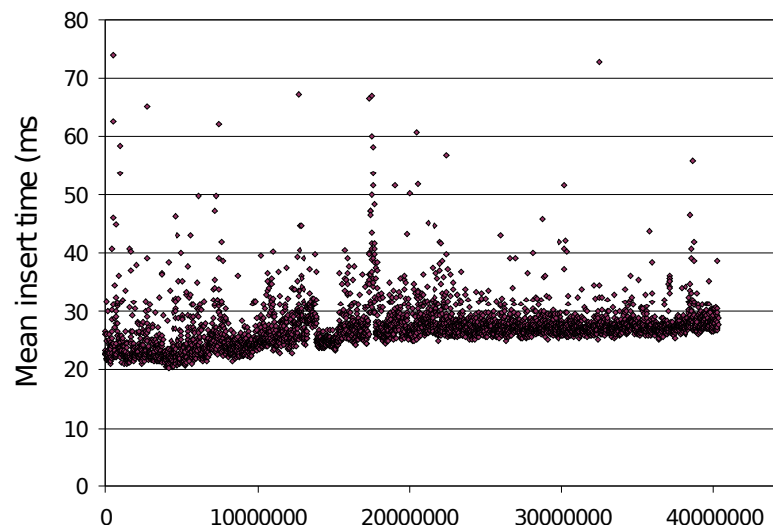
Security overheads now being tested

- Quality of machines used should be noted when comparing LFC and RLS results!

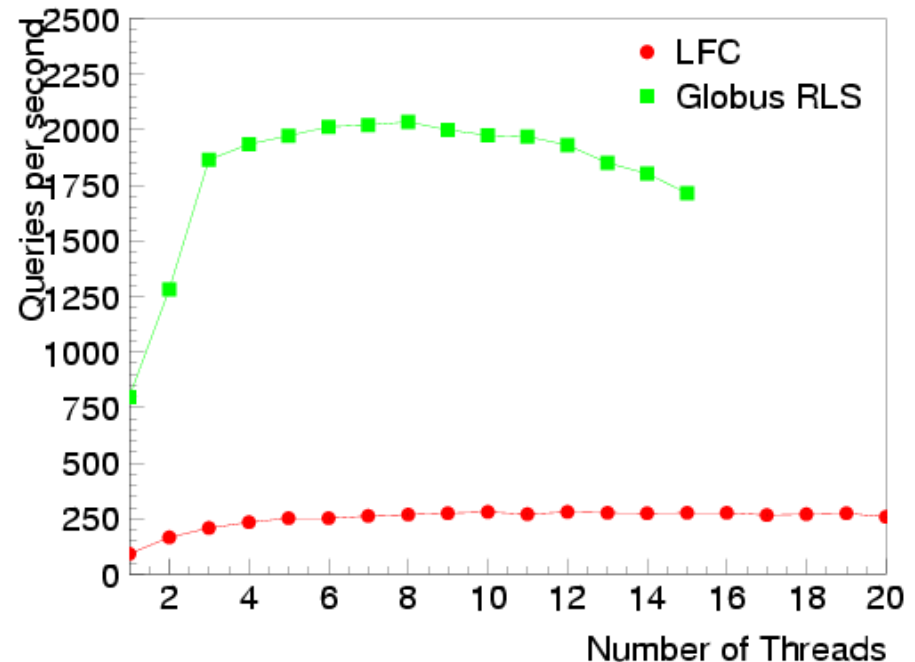
SPEC CINT2000 values:

	LFC	Globus RLS	EDG RLS
Server	420	810	420
Single Client	420	220	420
Multiple Clients	400	220	420

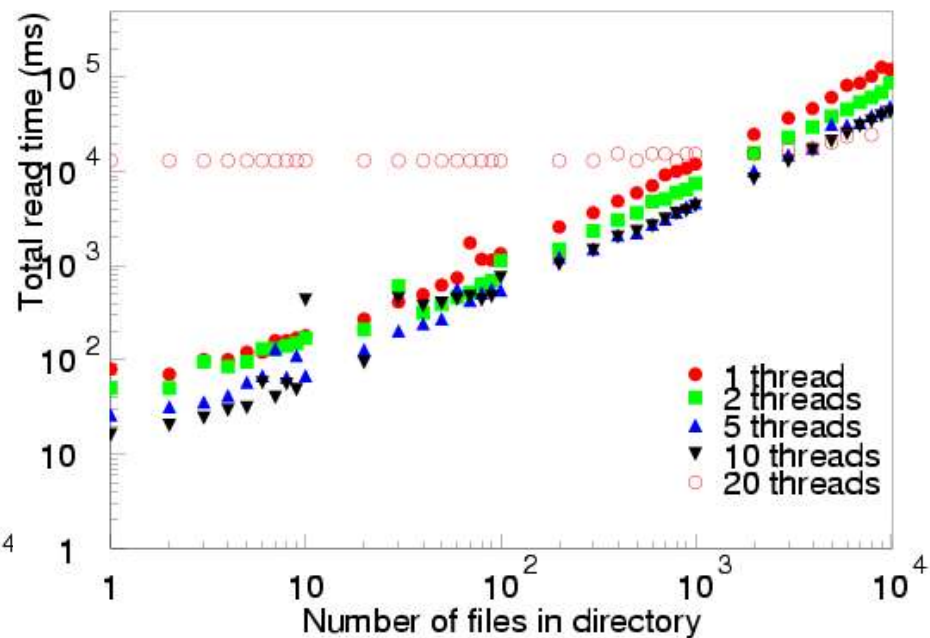
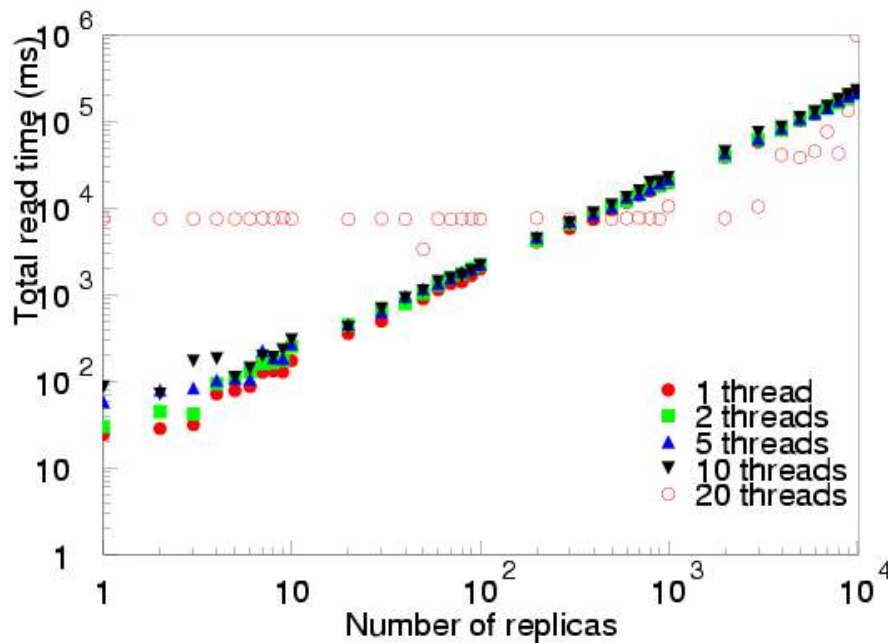
- Mean insert time as number of entries increased up to 40M remains below 30 ms
- EDG mean insert time was ~40 ms with 500,000 entries
- Insert rate, with increasing number of client threads, for ~1M entries
- Increases up to ~200 adds/sec up to server thread limit
- Globus RLS gave ~84 adds/sec when run with consistency



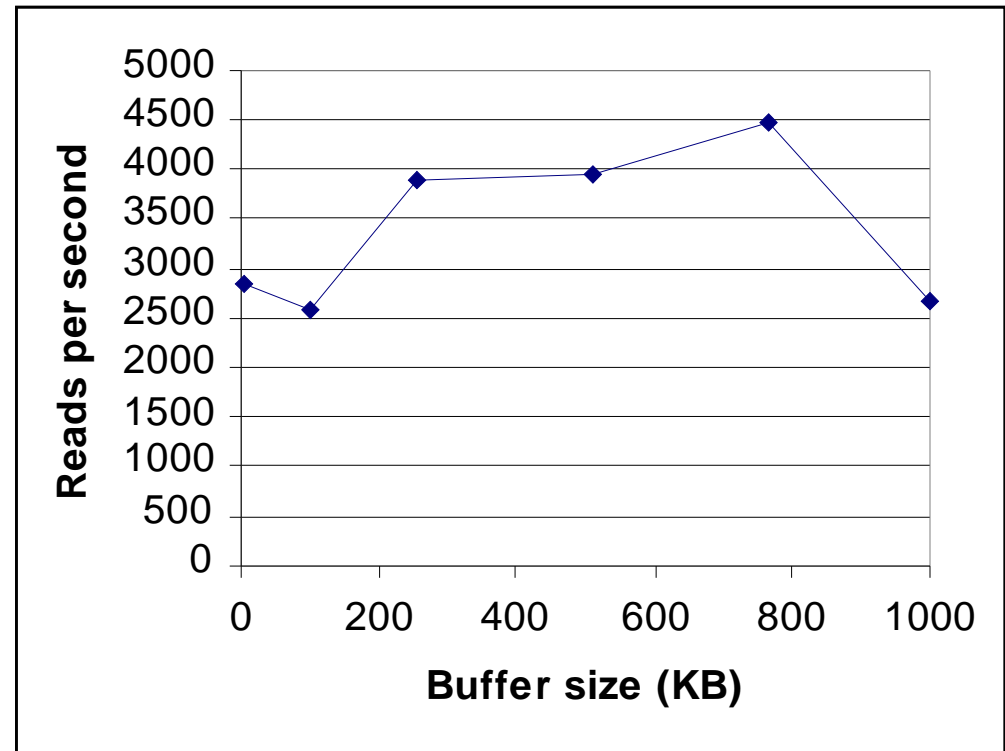
- Rate of querying for a single LFN, increasing number of client threads, ~1M entries
- Not really comparable with Globus results
 - RLS does 1-to-1 lookup
 - LFC stat() returns system metadata, checks permissions...
- EDG RLS rate ~63 queries/sec for 1 thread; LFC with 1 thread gives ~90 queries/sec



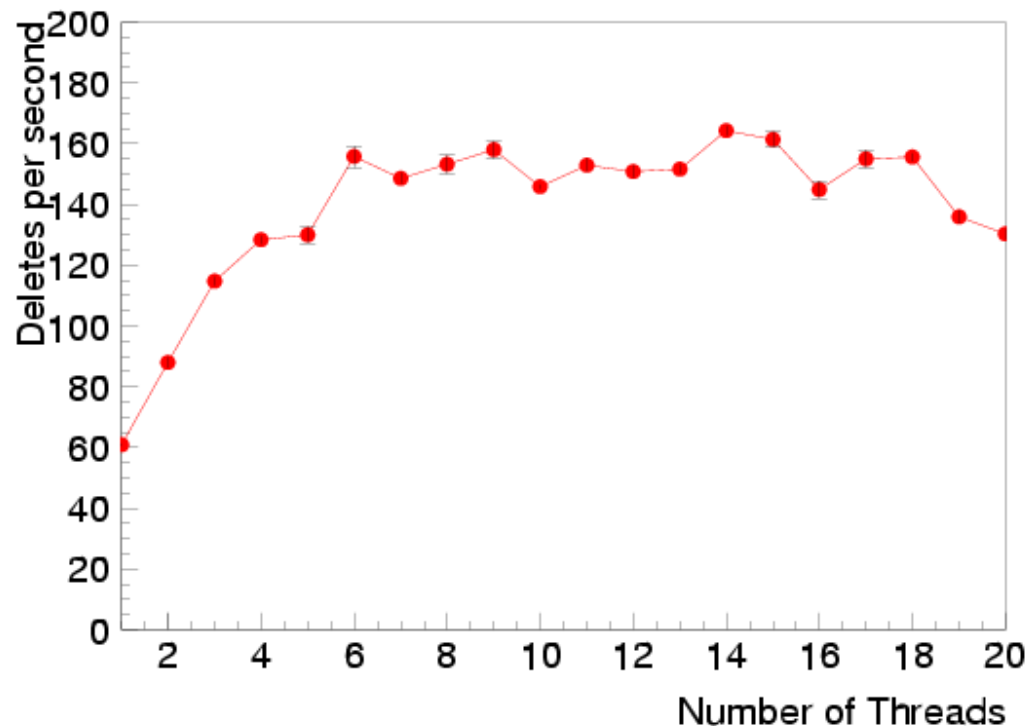
- Time to list and stat all replicas of a file proportional to number of replicas
- Time to read a directory is directly proportional to directory



- Default buffer size in LFC is small (4 KB)
- Tuning the buffer size leads to improved performance for readdir()
 - Time to read directory of 100000 entries measured with varying buffer sizes
- If bulk queries implemented, they should show similar behaviour



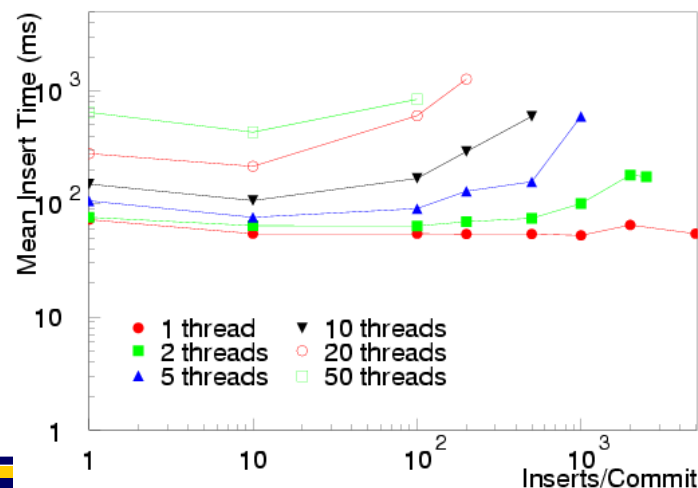
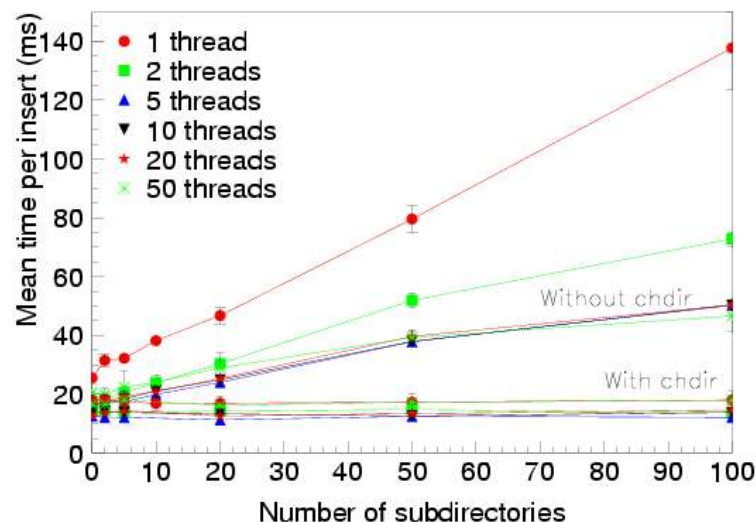
- Delete rate, with increasing number of client threads, for ~1M entries
- With 1 thread, time per delete ~16 ms
- EDG RLS with 1 thread gave ~30 ms per delete
- No comparable results for Globus RLS



- Performing a `chdir()` before many simultaneous operations in the same directory improves performance significantly
- Using transactions gave loss of performance with single client

Extra time being spent in DB

Requires further investigation



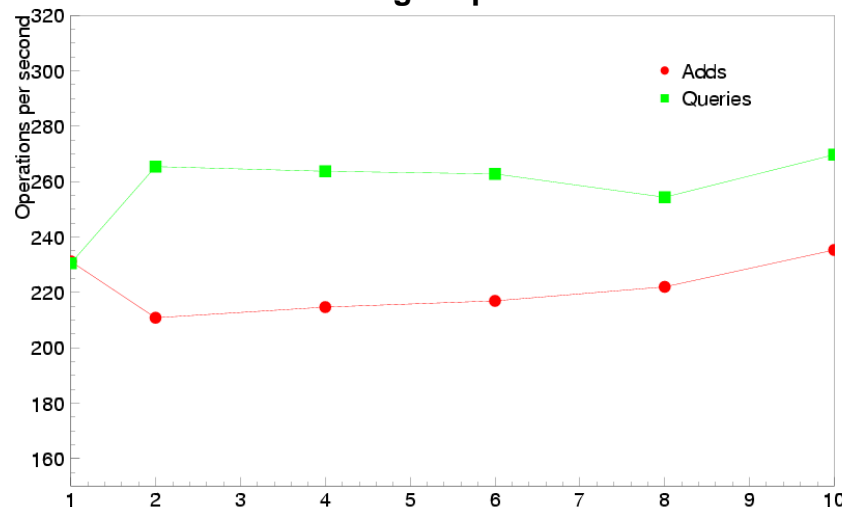
- Tests run simultaneously on varying number of client machines
- Clients all running with 10 threads
- Insert and query rates measured with:

Single operations

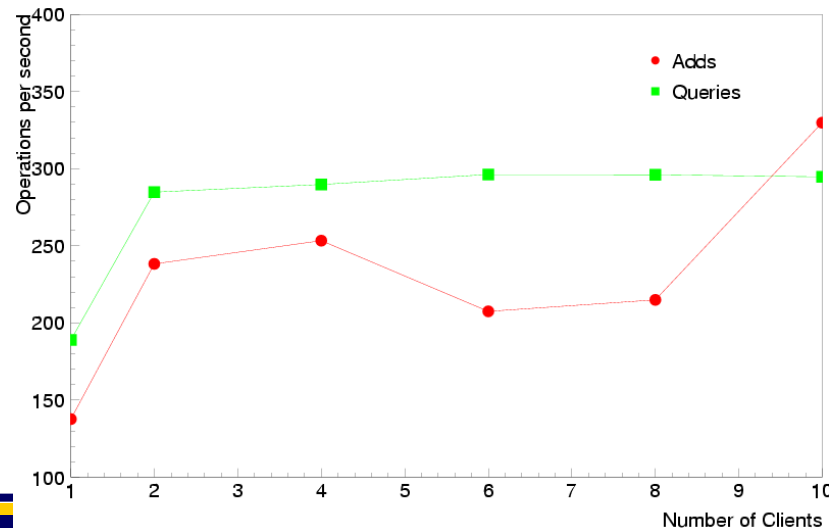
Transactions (100 operations per transaction)

- No reduction in operation rate as number of clients increases

Single Operations



Transactions





- LFC has been tested and shown to be scalable to at least:
 - 40 million entries
 - 100 client threads
- Performance improved with comparison to RLSs
- Stable :
 - Continuous running at high load for extended periods of time with no crashes
 - Based on code which has been in production for > 4 years
- Tuning required to improve bulk performance





- LFC and FiReMan tests performed on identical hardware
- Catalogue Server and Oracle DB running on same machine
 - Dual Xeon 2.4Ghz with 2048 MB RAM
- PIII 800 Mhz, 512 MB RAM Client with configurable number of threads
- 100 Mb/s LAN
- Insecure FiReMan and LFC

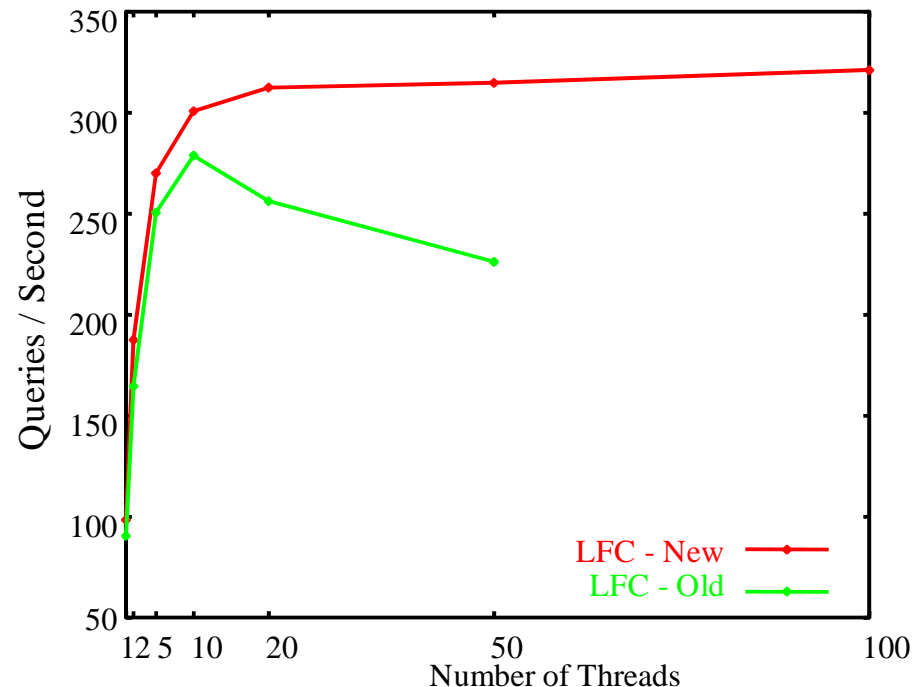
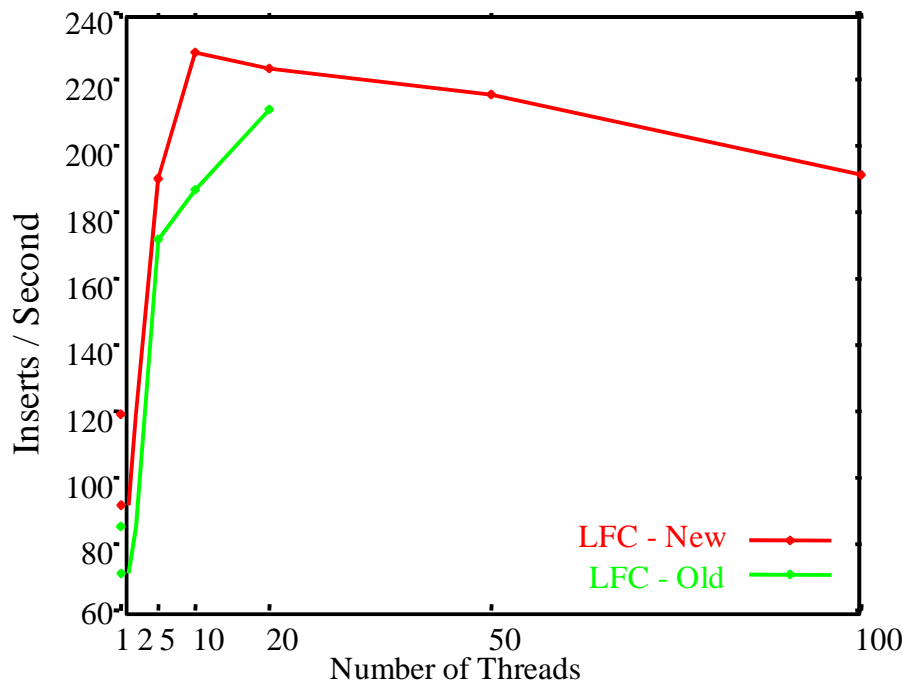




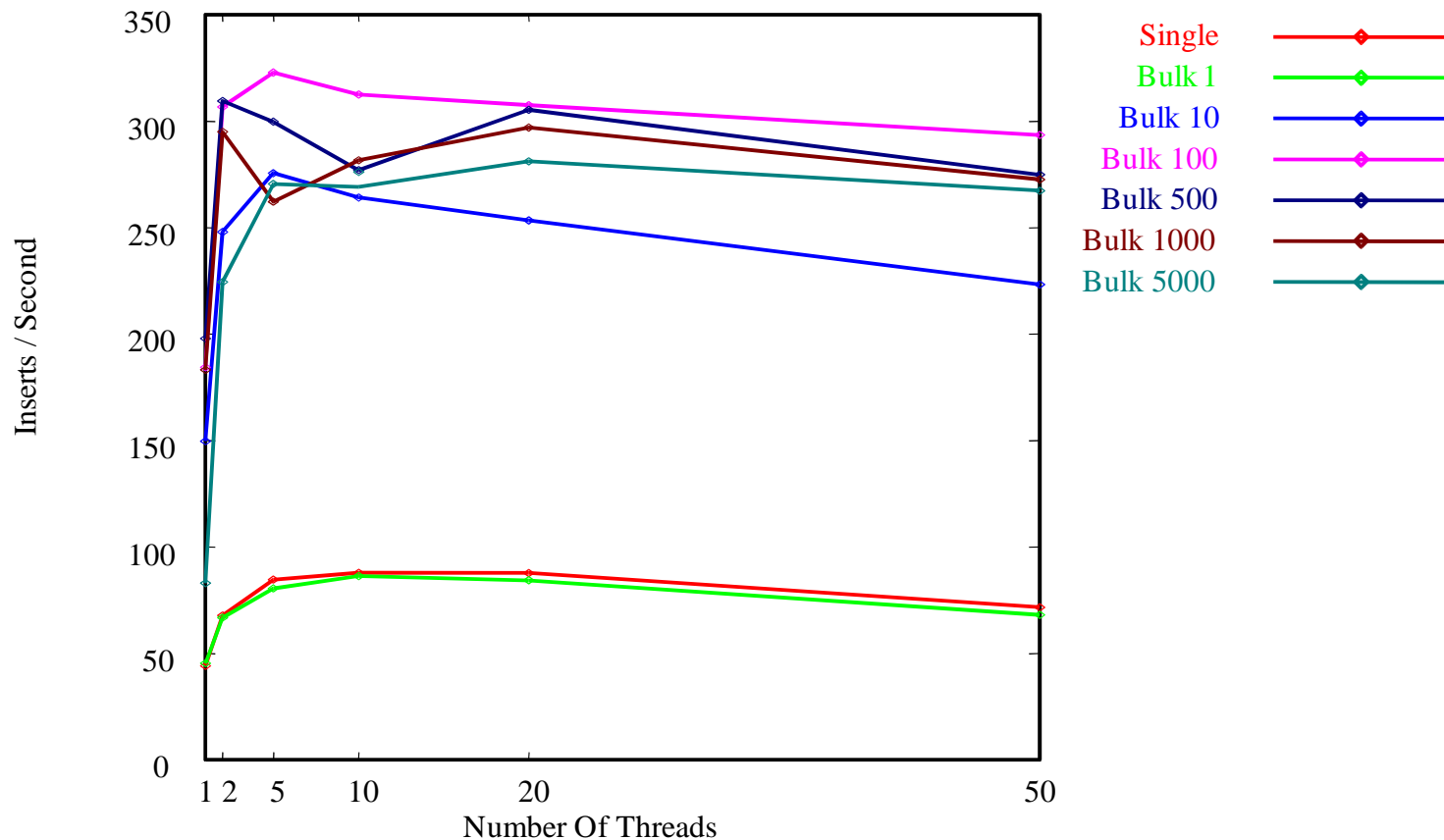
- Limited time -> limited scope of tests
- Only one multi-threaded client used to test server.
Difficult to explore all limits of server
- All tests performed over LAN, need to look at WAN
Influence of round trip time



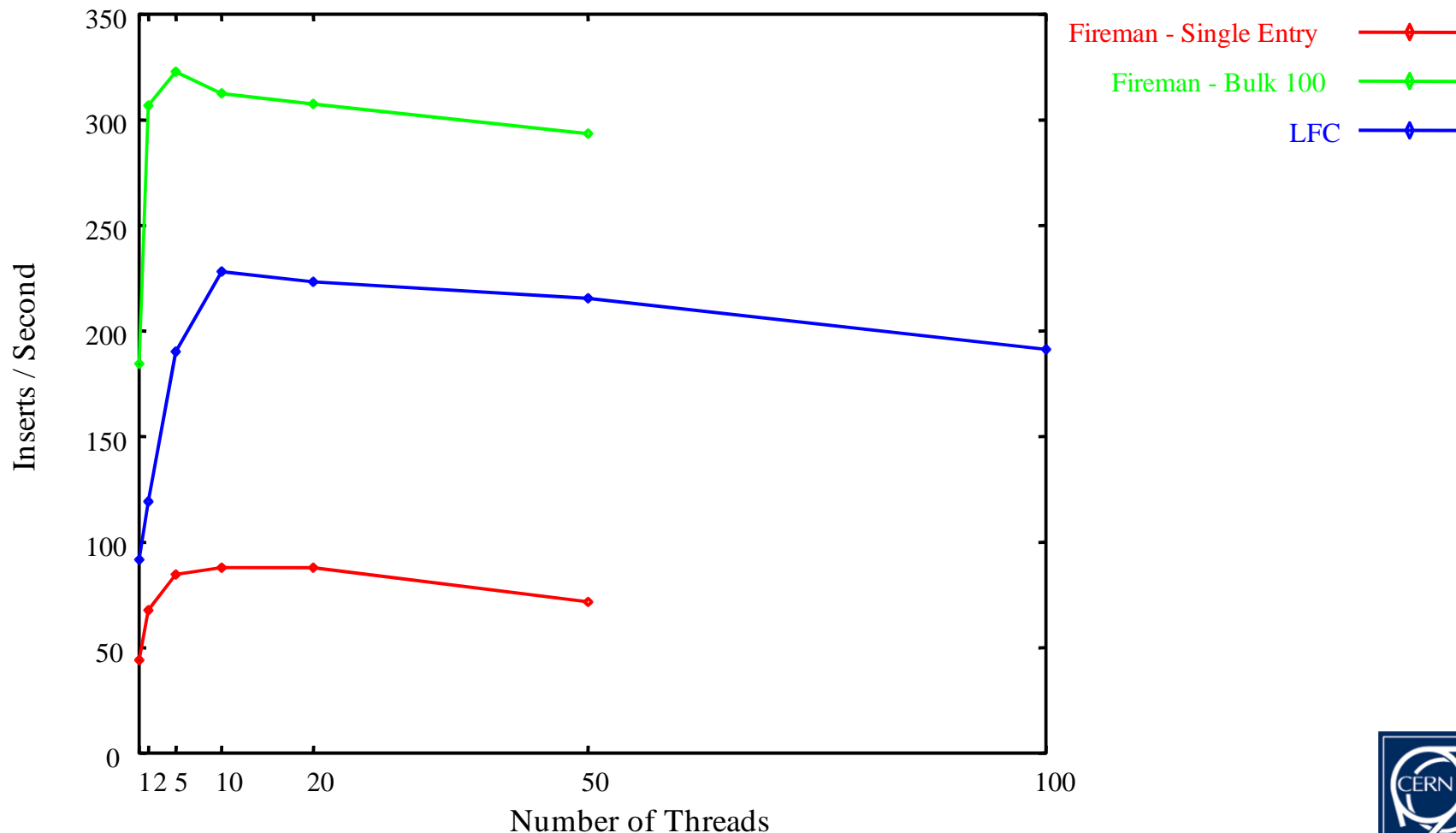
- LFC Tests repeated on identical hardware to give fair comparison
- LFC Tests performed with relative paths (chdir) and without transactions



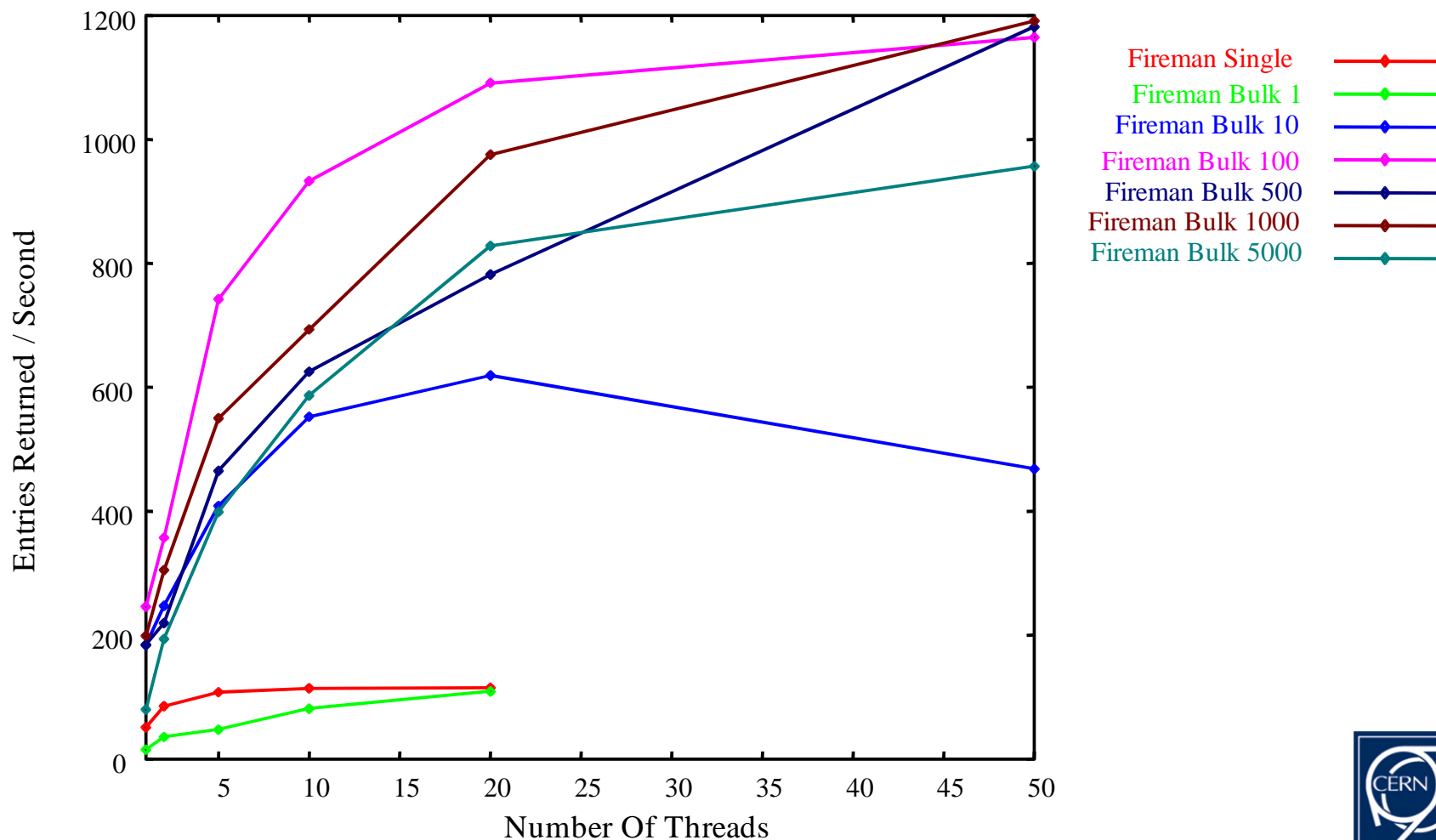
- Inserted ~1M entries in bulk with insert time ~5ms
- Insert Rate for different bulk sizes



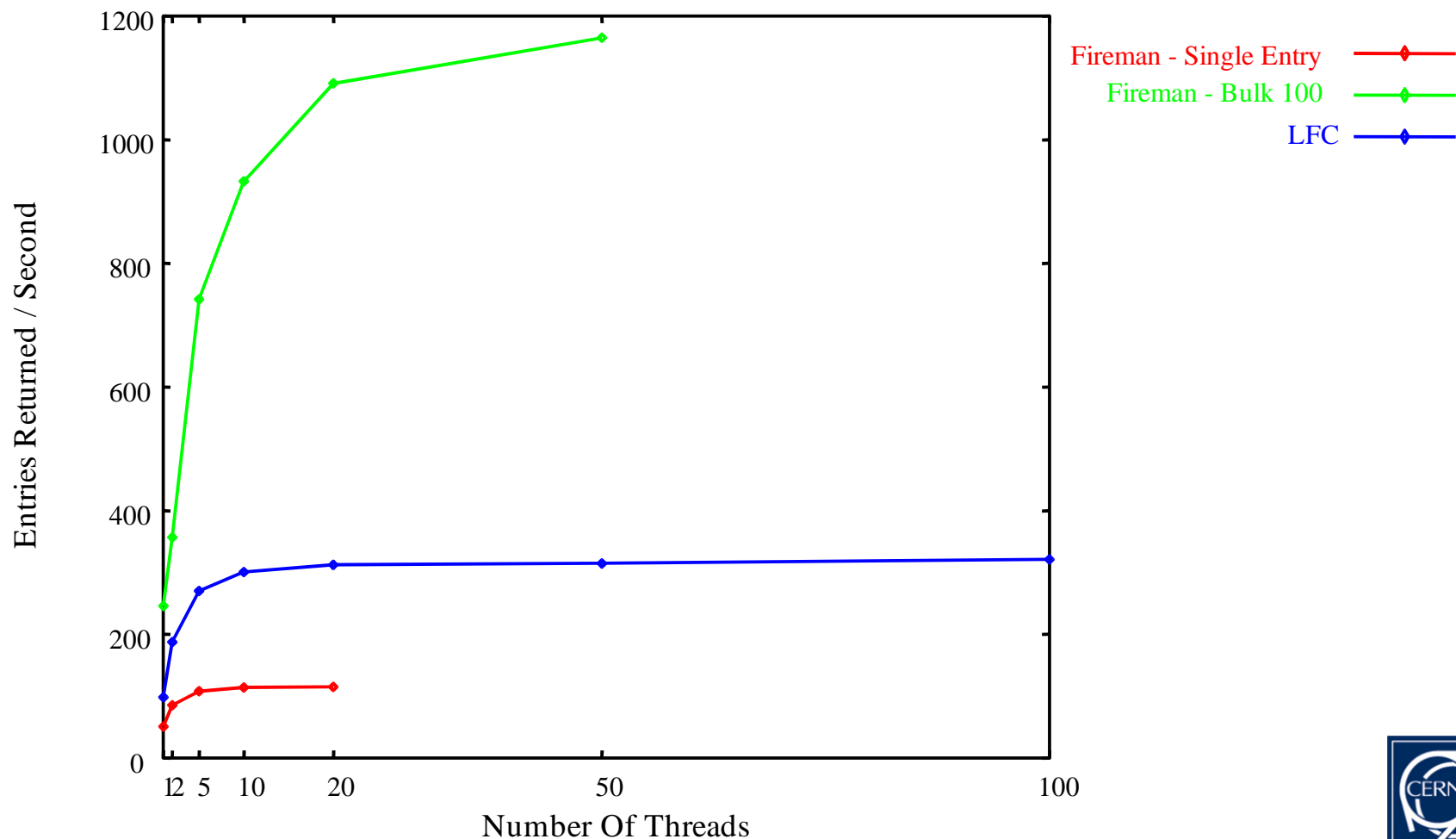
- Comparison with LFC:



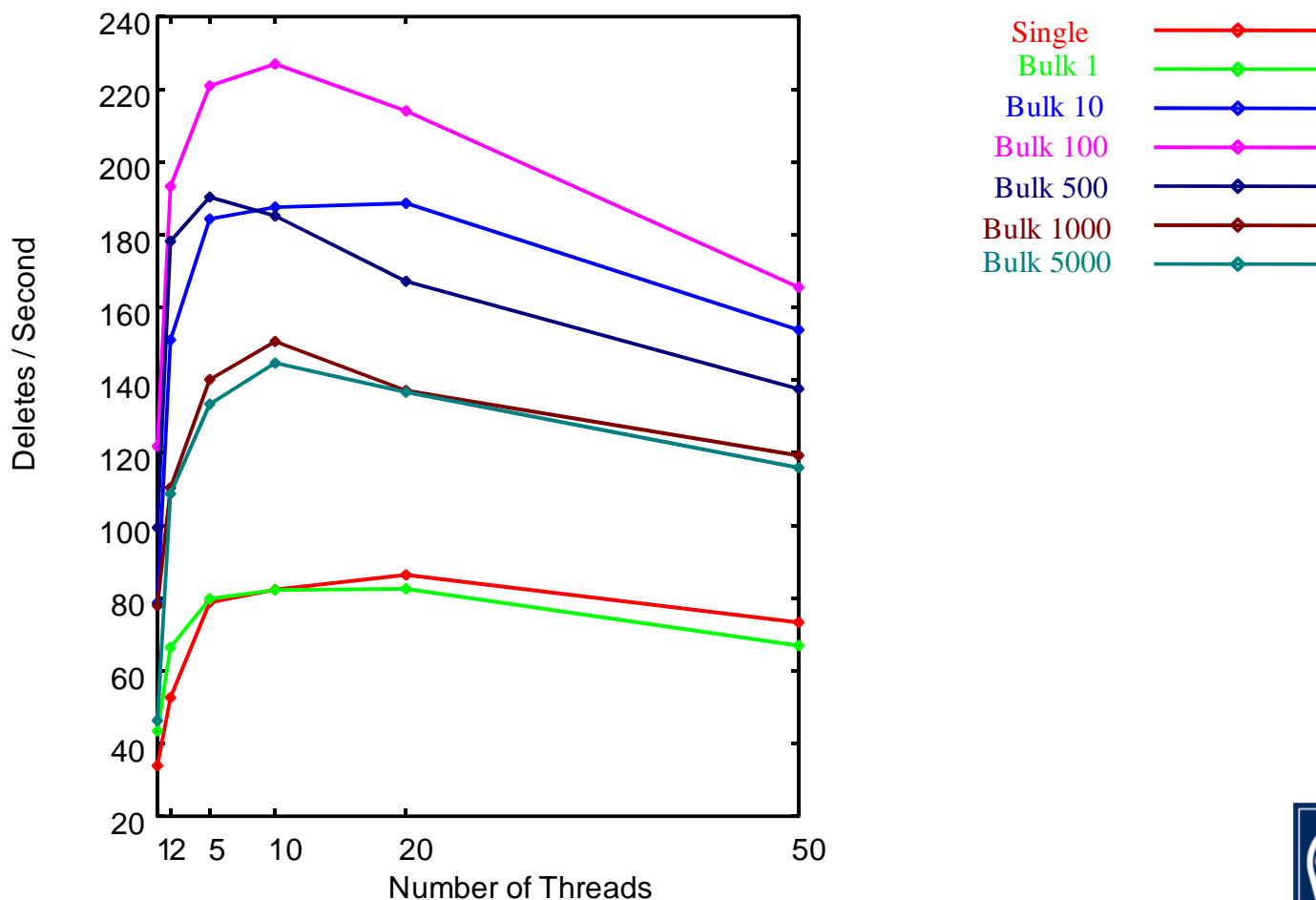
- Query Rate for an LFN



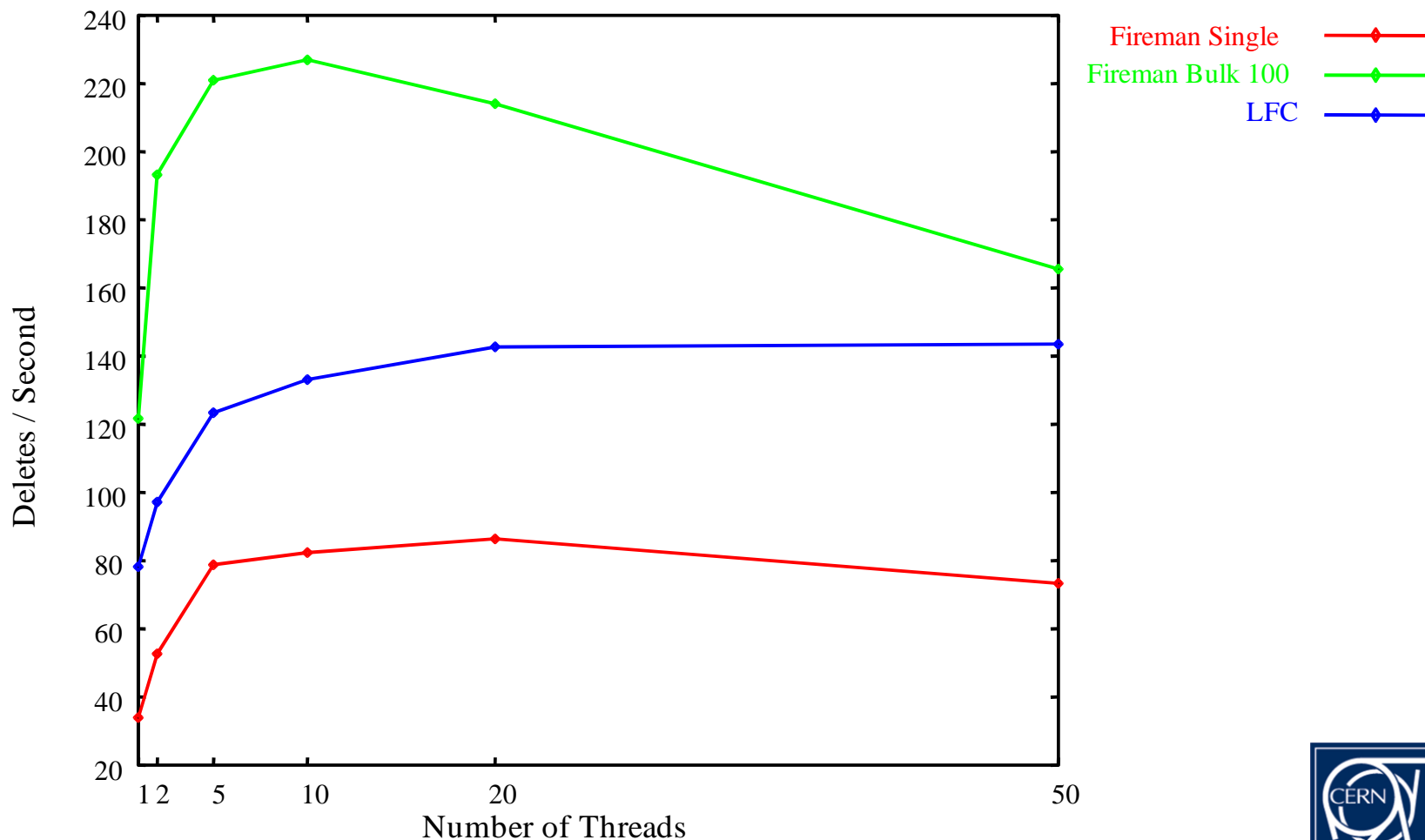
- Comparison with LFC:



- Rate LFNs can be deleted from catalogue



- Comparison with LFC:





- Both LFC and FiReMan offer large improvements over RLS
- Still some issues remaining:
 - Scalability of FiReMan
 - Bulk Entry for LFC
- More work needed to understand performance and bottlenecks
- Need to test some real Use Cases





Questions?



eGEE

Enabling Grids for
E-science in Europe