

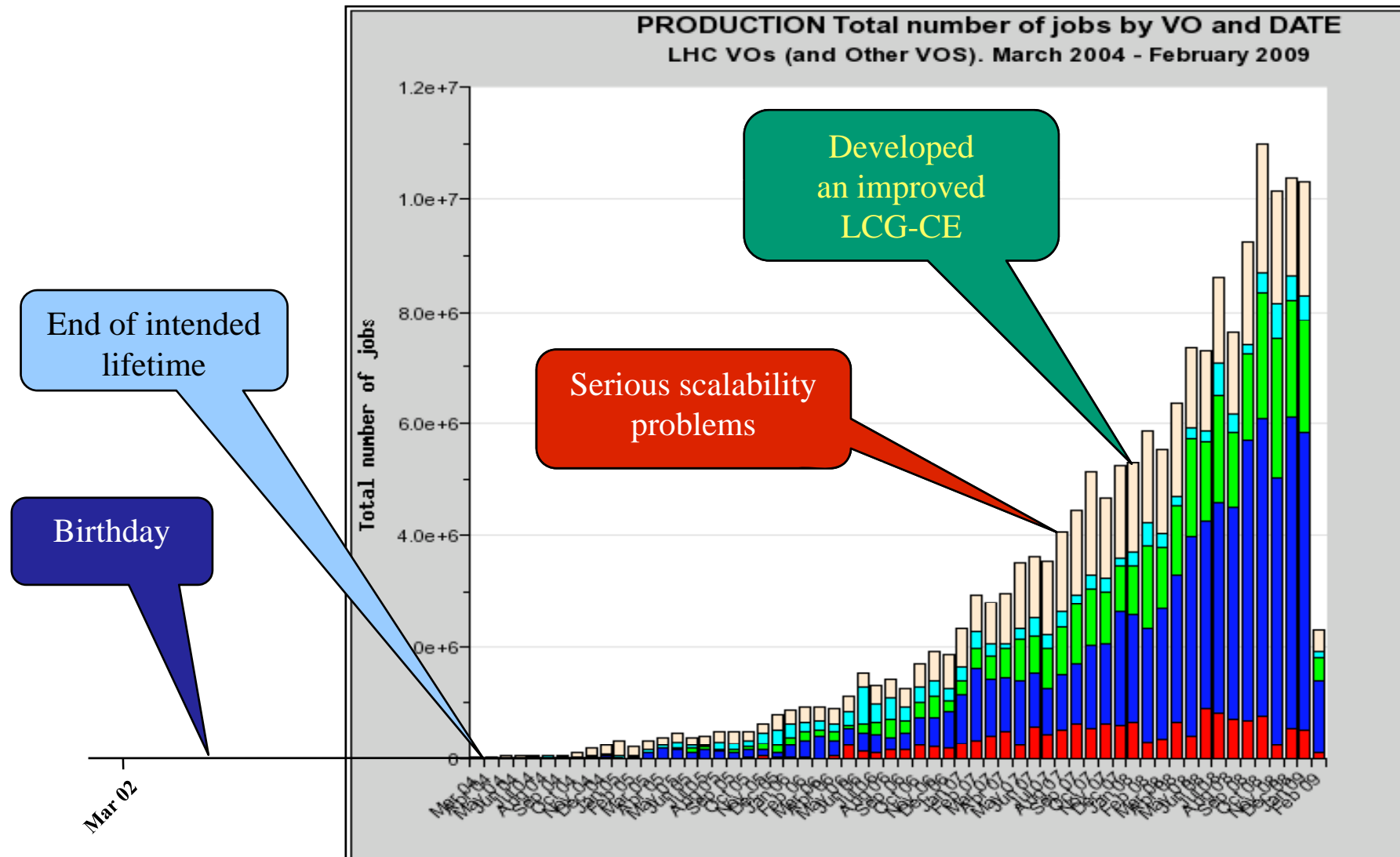
Current status of LCG-CE

Andrey Kiryanov, PNPI

History of LCG-CE

- Introduced at the end of the European Data Grid Project
 - Before 2003 – the oldest piece of distribution which is still there
 - Focus on improved stability
 - Limited effort
- Intended as a stop gap solution for the CE soon to come
 - Proper architecture
 - Stability by design
 - Scalability
 - Push & Pull
 - etc.

The infrastructure and usage growth (here since 2004)



Scalability problems

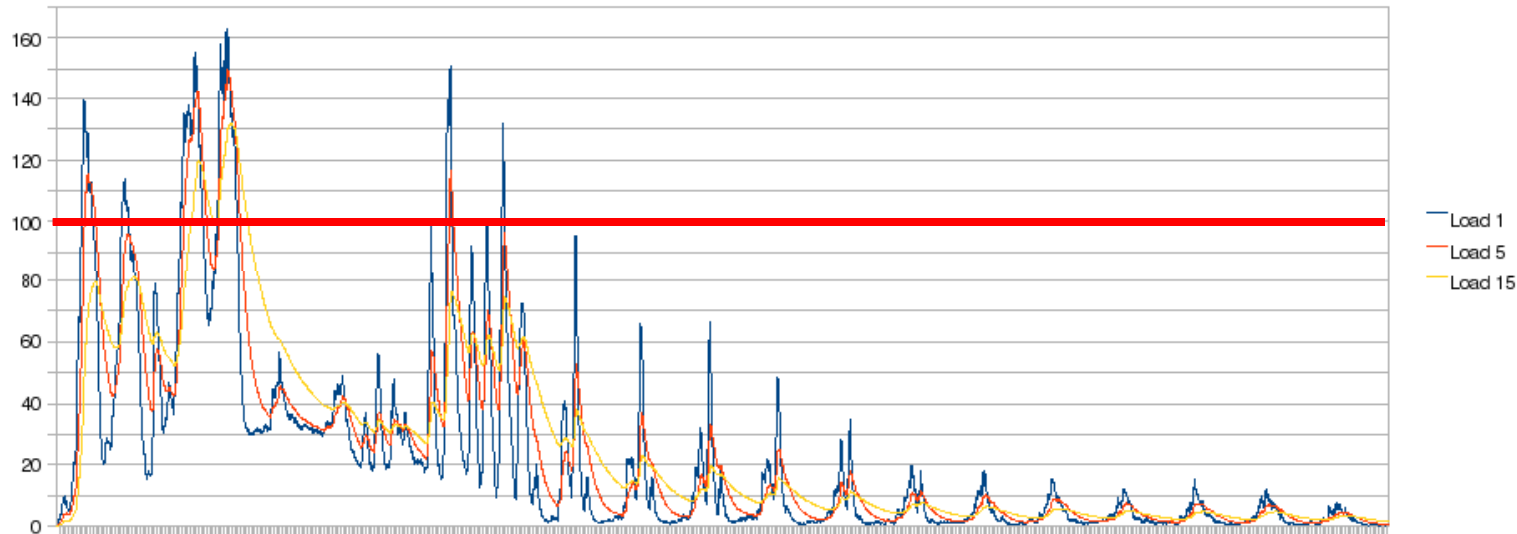
- Globus (and consequently LCG-CE) was not initially designed to deal with so many users and jobs at once
 - It's a problem of the **Globus-2/GRAM** architecture
 - Situation became even worse with introduction of pool accounts and VOMS roles
- “**One process per job**” or “**one process per user**” approach
 - Globus tries to serve all requests in parallel, no queues or other machinery to control this
- As a consequence more than **10 users** working in parallel can render a LCG-CE completely irresponsive (**load average over 100**)
 - Linux kernel is still far from perfect in handling such a heavy load

Latest improvements

- Improved Jobmanagers (globus-***-marshal patches)
 - A memory-persistent daemon was developed to avoid Perl code recompilation and to provide a queue-based control over parallel requests from binary jobmanagers
 - Perl script was replaced by a tiny client written in C, that communicates with daemon via domain socket
- Grid Manager Monitor Agents (globus-gma patch)
 - An existing machinery inside “fork” jobmanager was used to detect incoming agents and replace them with modified “light” ones
 - A monitoring daemon was developed to supply job state information to “light” monitoring agents

Test results (1)

Original LCG CE

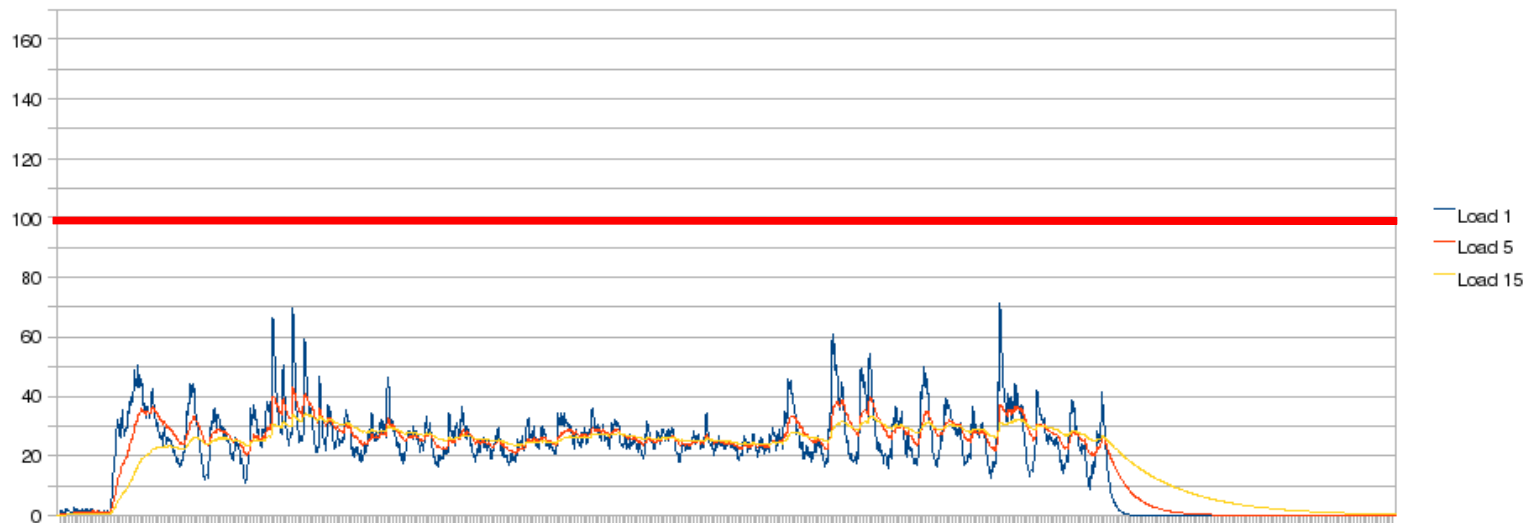


Original
version:

10 users

1000 jobs

New Modified LCG CE, 30 users

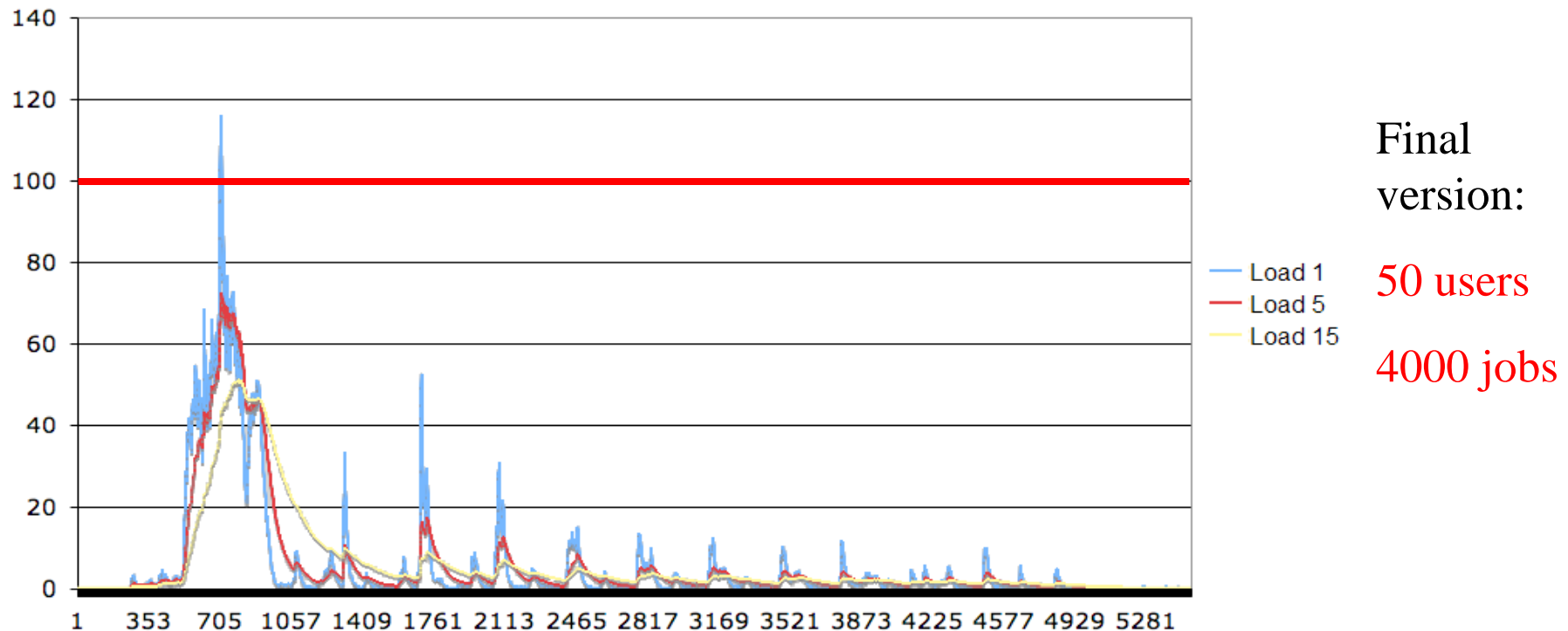


Modified
version (no
globus-gma
yet):

30 users

4000 jobs

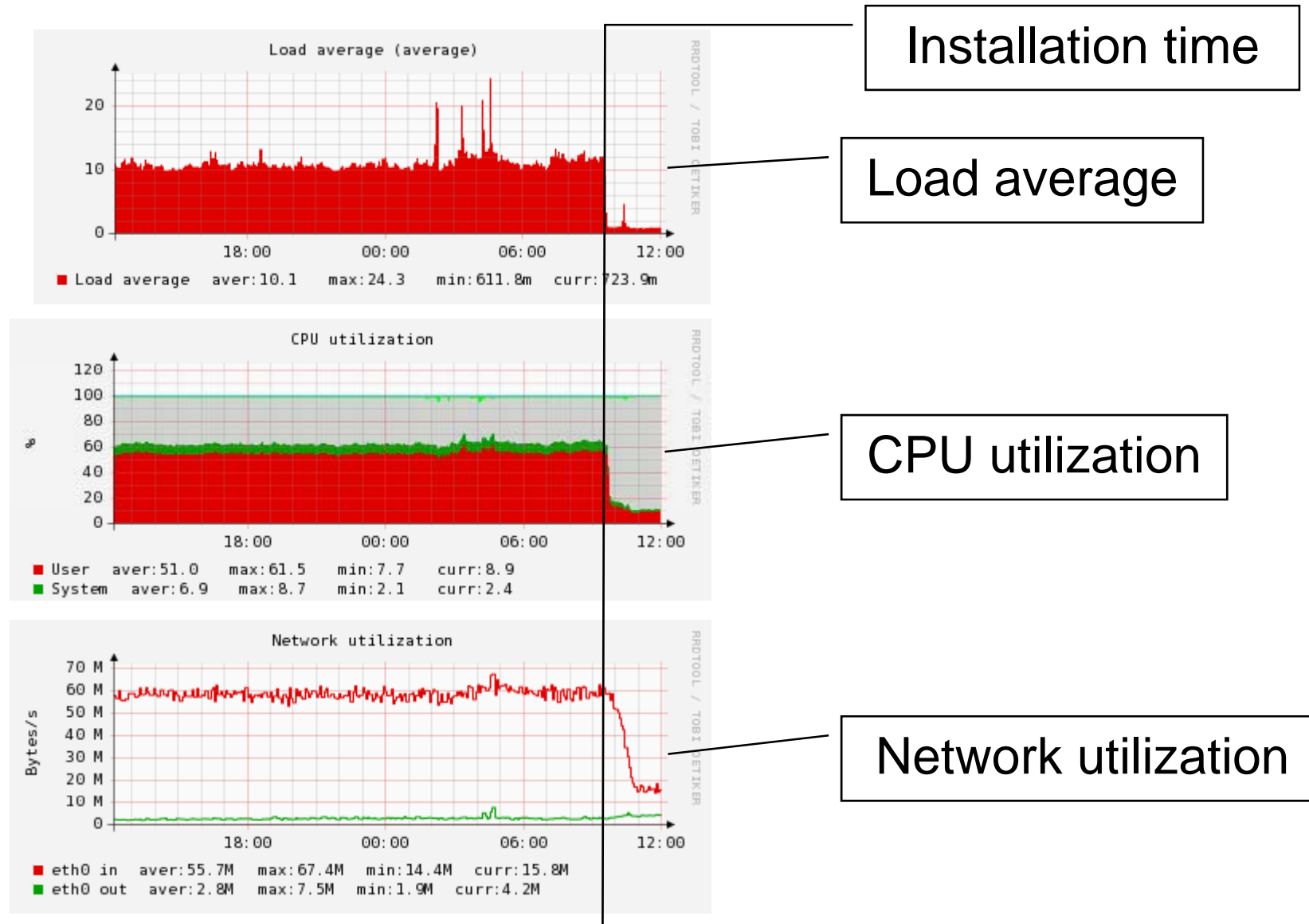
Test results (2)



Stress test of the final version with both globus-gma and globus-*
marshal patches installed

First load peak is caused by file transfers from WMS (gridftp)

Installation on CERN CE cluster



Conclusions (1)

- System load on CE is decreased by factor of 3 to 5
- Jobs start and finish faster
 - especially visible with lots of short jobs
- CE can handle significantly larger number of jobs **and** different users
- No interface changes, no need to modify existing software
- LCG-CE can be tuned for hardware with different CPU/disk performance
- Modified LCG-CEs are in LCG/gLite production since April 2008, already installed on most gLite 3.1 sites

Conclusions (2)

- How much time do we have gained?
- With 50 users and 4K jobs sites can handle the Production load with a few LCG-CEs
- For analysis we have to handle hundreds of users/roles with hundreds to thousands of jobs per CE
 - For this LCG-CE will probably not be good enough
 - CREAM-CE ?
 - Pilots ?

Thank you!