Update on MSG Messaging System for the Grid

Daniel Rodrigues

- Flashback

- Reporting on MSG
  - Stability
  - WAN
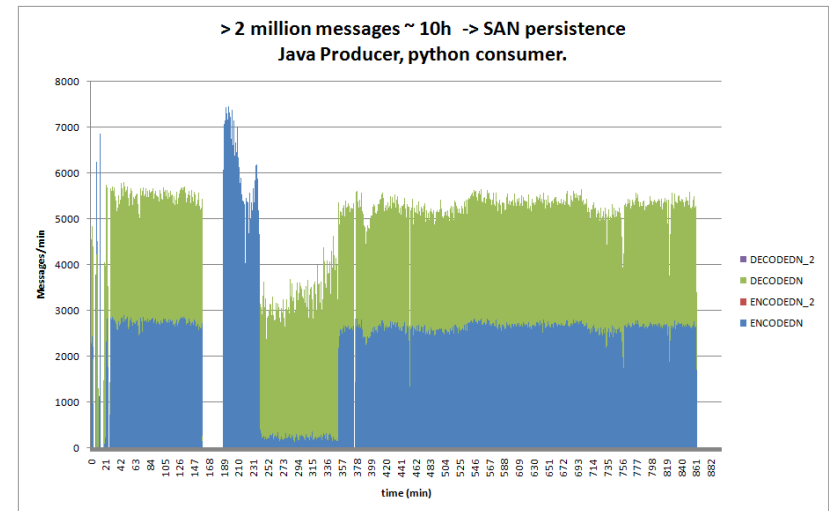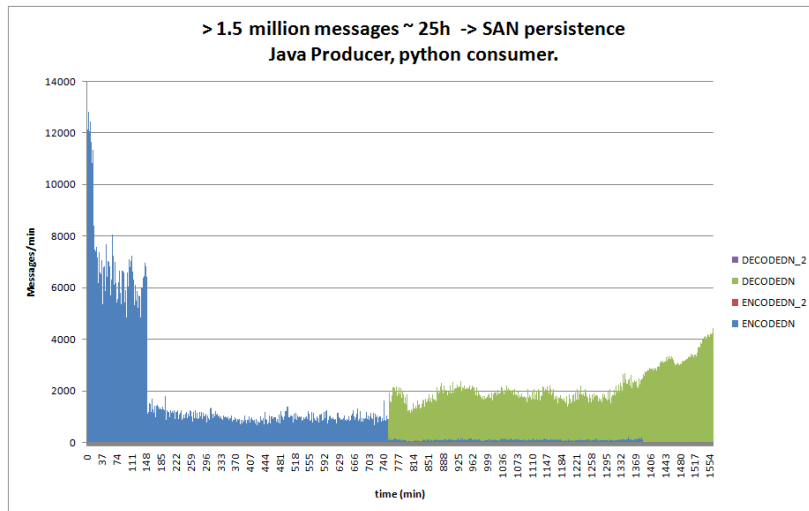 - SSL

- Future role on OA

- Conclusion

- We had already the opportunity to present you what is MSG (again):
  - A novel approach to exchanging information
  - Asynchronous
  - Reliable
  - Scalable
- Suitable for the monitoring environment
- First tests looking promising
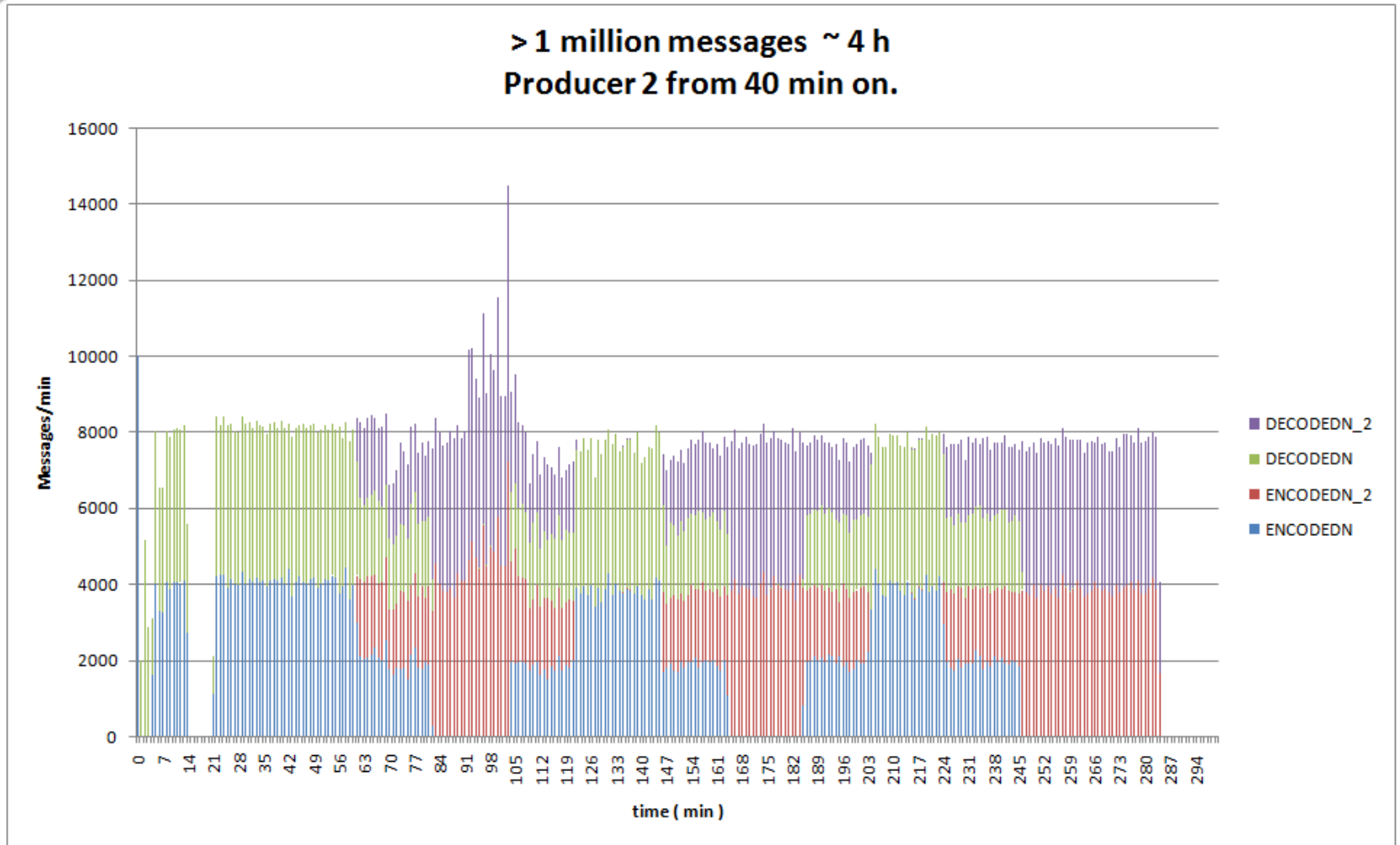- Some prototypes in place

- Since 1 year:
  - Prototype with different options for the clients
  - Different versions and setups of ActiveMQ
  - Working with people with use cases
  - Putting everything in place
  - Running the system and troubleshooting
- Currently used in different tools:
  - SAM, Gridview, JobMonitoring
  - Monitoring and improvements on place
  - Going through the pain of releasing properly

Reporting on MSG

- Stability:
  - How has MSG been running (production)?
  - Load still far from testing limits:
    - Sustained ~50 records/s for the past few months
    - With peaks of 600/s every night (jobDetails)
    - Each record varies, but from a few bytes to kB
  - Recovered whenever a broker would go down (usually memory): As expected!
    - However, time of failover dependent on message store size
    - Increased committed memory to JVM from 512 kB to 4 Gb
    - Broker hasn't go down for a month! ☺
  - However, one month ago…
    - Brokers are using SAN message store over AFS (lxmrrb3705)
    - Kernel panic occurred with a tricky problem...
    - File lock not released, even though machine was down
    - There's the need for human intervention (machine was in maintenance)
    - A question to Openlab folks:
      - The file is locked (/proc/locks)
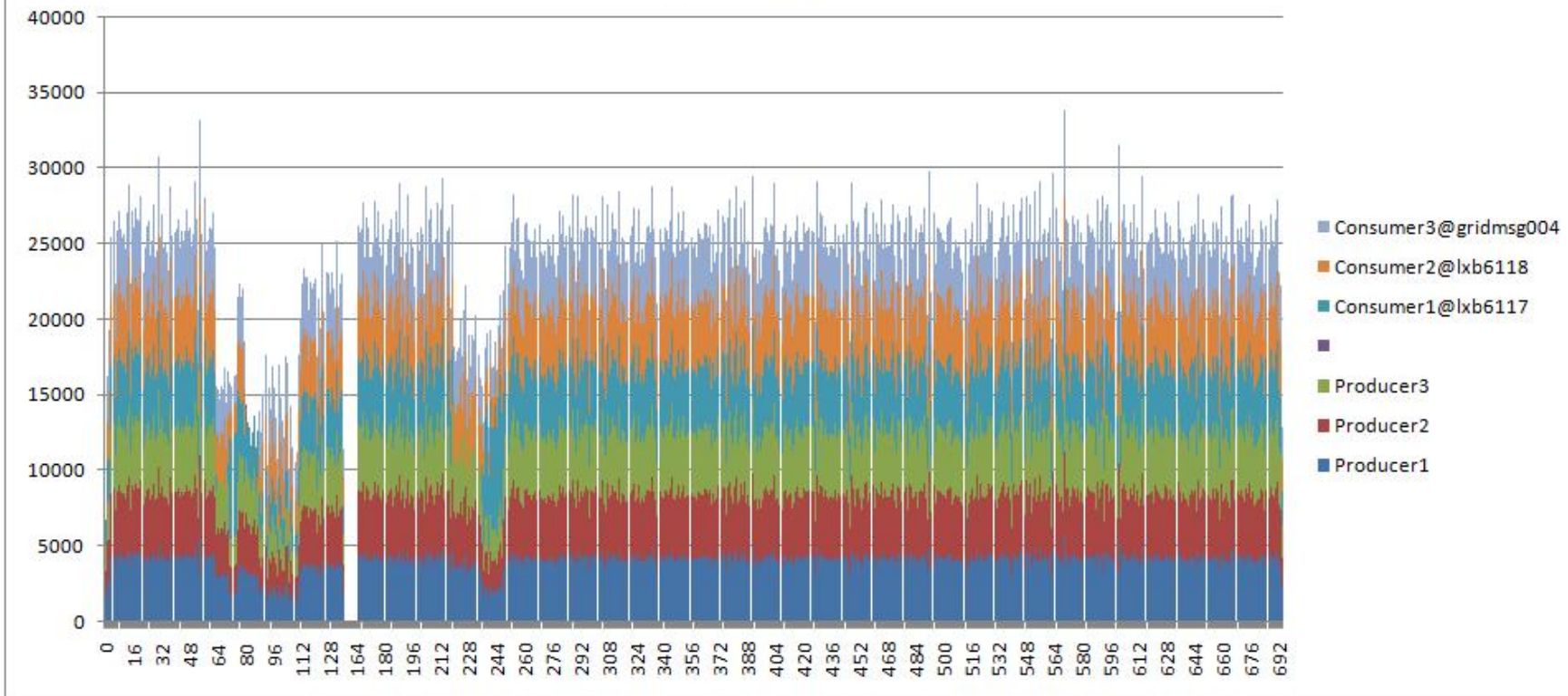   - Pid is not of a process in the machine
      - Implications!

```
root@gridmsg002:~
[root@lxmrrb3705 proc]# cat /proc/locks
1: POSIX  ADVISORY  WRITE 24296 08:06:98328 0 EOF
2: POSIX  ADVISORY  WRITE 24287 08:06:98327 0 EOF
3: POSIX  ADVISORY  WRITE 14759 08:06:6144015 0 9223372036854775806
4: POSIX  ADVISORY  WRITE 14759 08:06:6144092 0 9223372036854775806
5: POSIX  ADVISORY  WRITE 14759 08:06:6144064 0 9223372036854775806
6: POSIX  ADVISORY  WRITE 14759 08:06:22970372 0 9223372036854775806
7: FLOCK  ADVISORY  WRITE 22406 08:06:98331 0 EOF
8: POSIX  ADVISORY  WRITE 21957 08:06:98332 0 EOF
[root@lxmrrb3705 proc]#
```

  - Lost messages on JobDetails?
    - No, not a Messaging System problem!
    - But we need to improve feedback (DeadLetters channels)
  - So, we have a stable system, but we need a minimum degree of maintenance.

- # WAN

  - ## Will have a major role in Operations Automation.
    - (we'll see later)

 - ## Network of brokers will have a central role.

  - ## So, testing!

> 1.5 million messages ~ 25h -> SAN persistence
Java Producer, python consumer.

> 2 million messages ~ 10h -> SAN persistence
Java Producer, python consumer.

> 1 million messages ~ 4 h
Producer 2 from 40 min on.

Long Run, >7 million messages, 3 different channels ~9h
Broker gridmsg003 (4Gb ram)

- ## Finally, what about WAN?

WAN sending records/minute Croatia broker Java-Publisher - msgConsume2Oracle
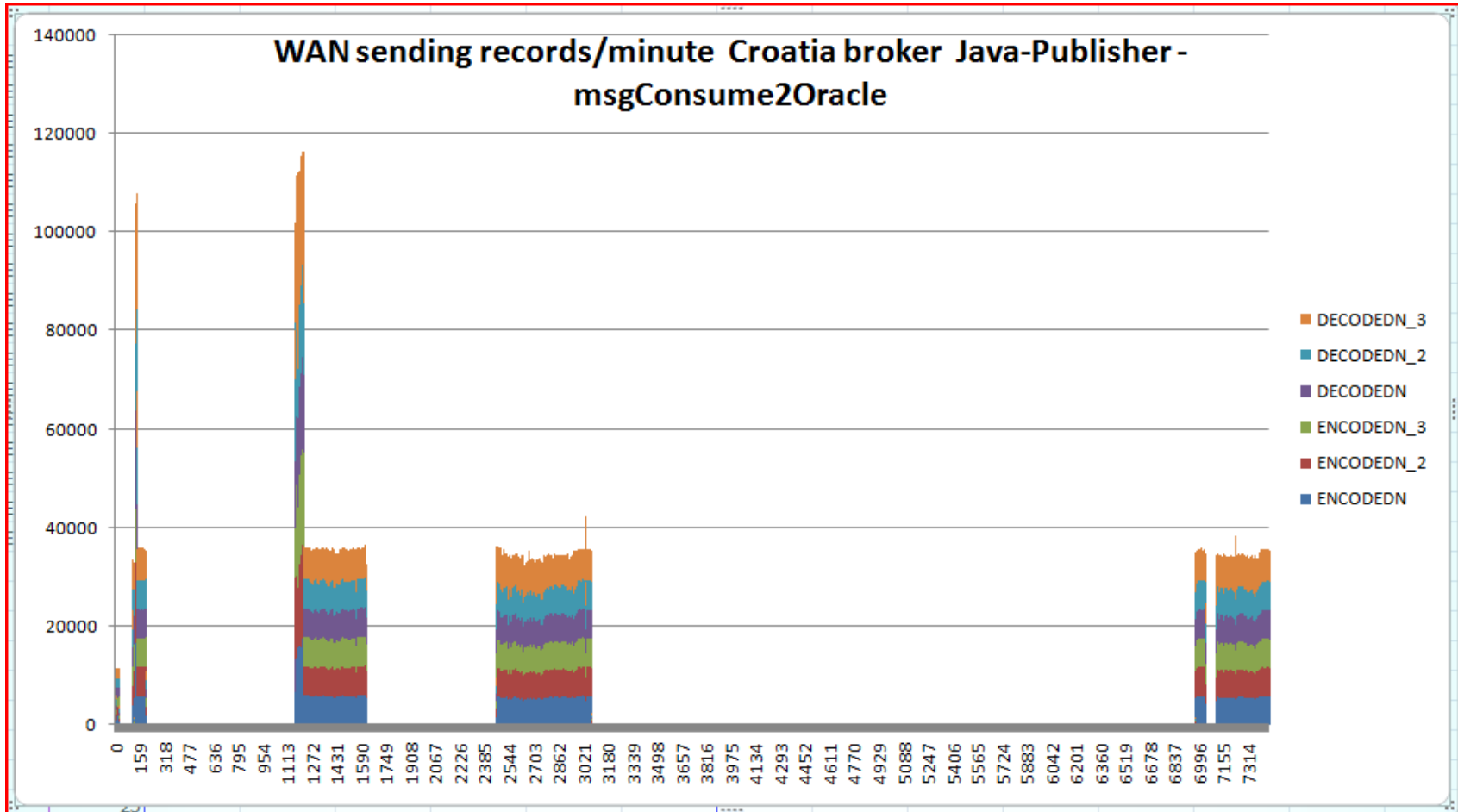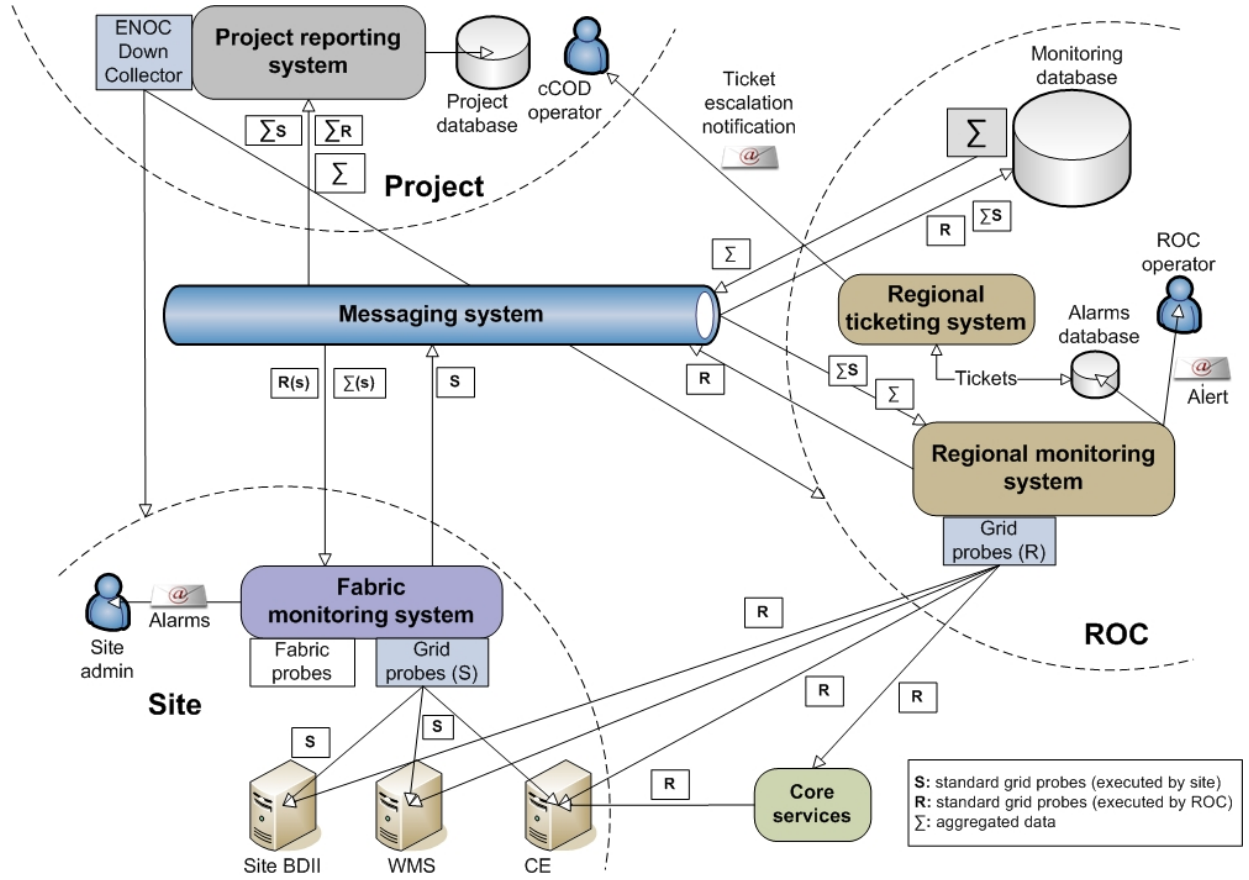
- Lag is not so important…
  - But, as reference, it was for WAN ~ 63 ms
  - And ~ 0 ms at cern.

- **Operations Automation**
  - Undergoing major revision
  - MSG central to the new architecture

- **Messaging System performing well**

- **Need minimal maintenance**
  - Be aware there might be unknown problems
  - Good performance under predicted conditions (failures on the clients, messages properly persisted)
  - Uptime satisfatory, not 100% due to maintenance.

- **Moving now to Network of Brokers in WAN**
  - With Zagreb University Computing center
  - Same throughput (dedicated machines on the other end)

- **Raising Security priority!**
  - SSL on the working canvas.

- **https://twiki.cern.ch/twiki/bin/view/LCG/MessagingSystemforGrid**

Thank you for your attention.