By Jeroen Klaver & Roel van der Jagt

Distributed file storage

Content

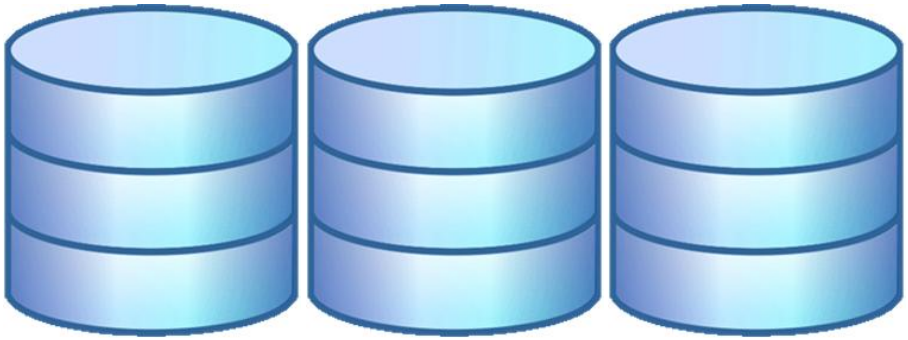
- Background
- Research question
- Test approach
- SURFnet design
- Conclusion
- Future work

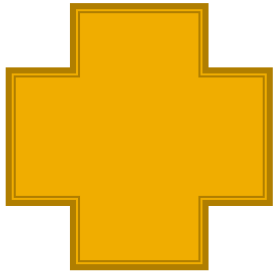
Background

- SURFnet is the Dutch NREN
- Storage explosion

Background

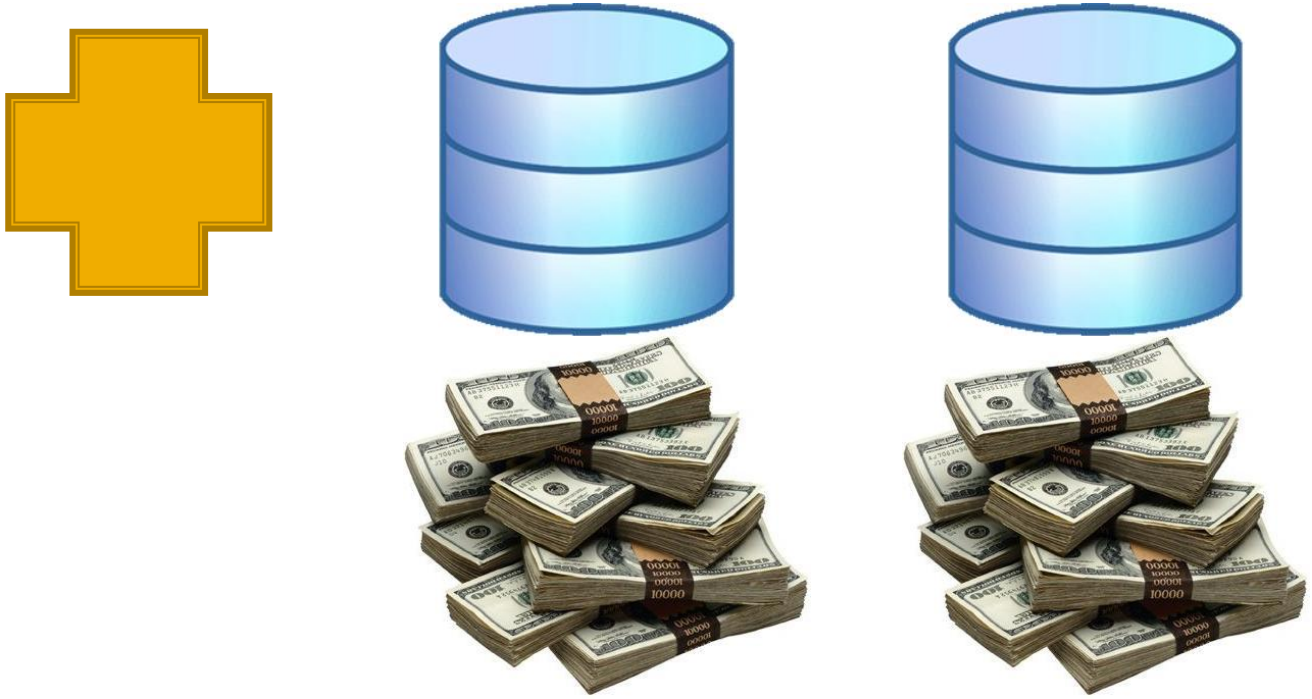
- Add more storage





Background

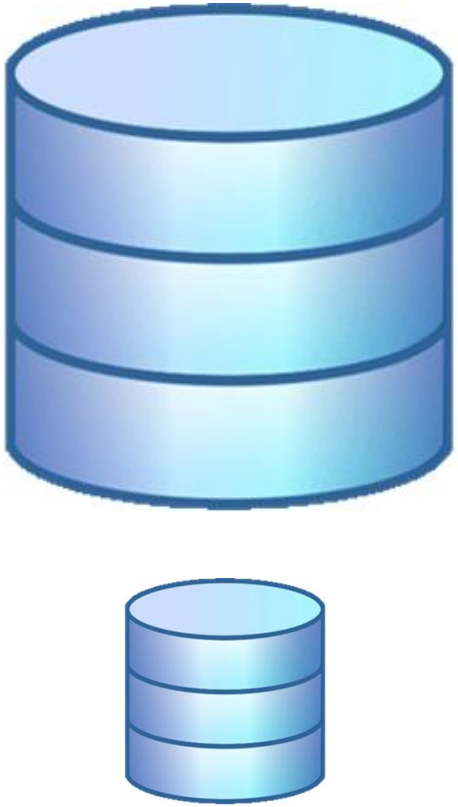
- And more…



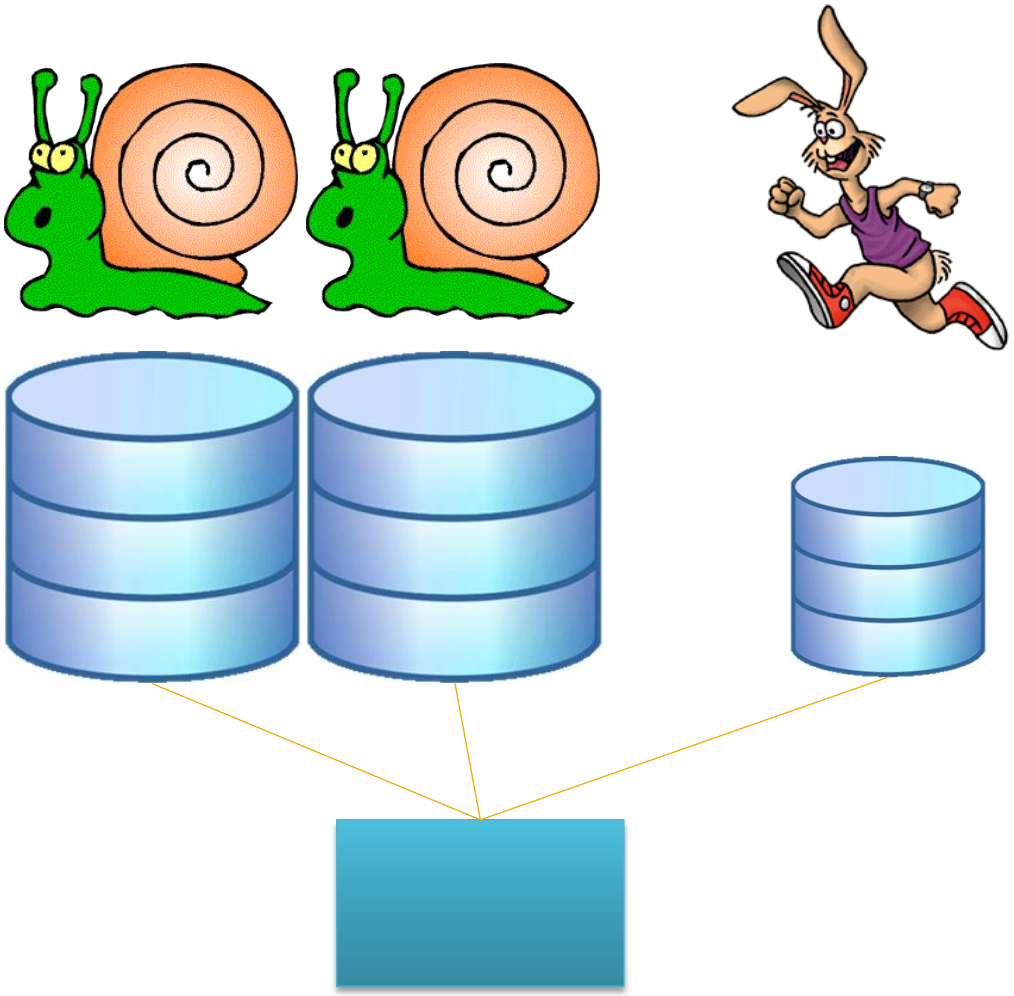
- But adding includes high costs and new bottlenecks

Background

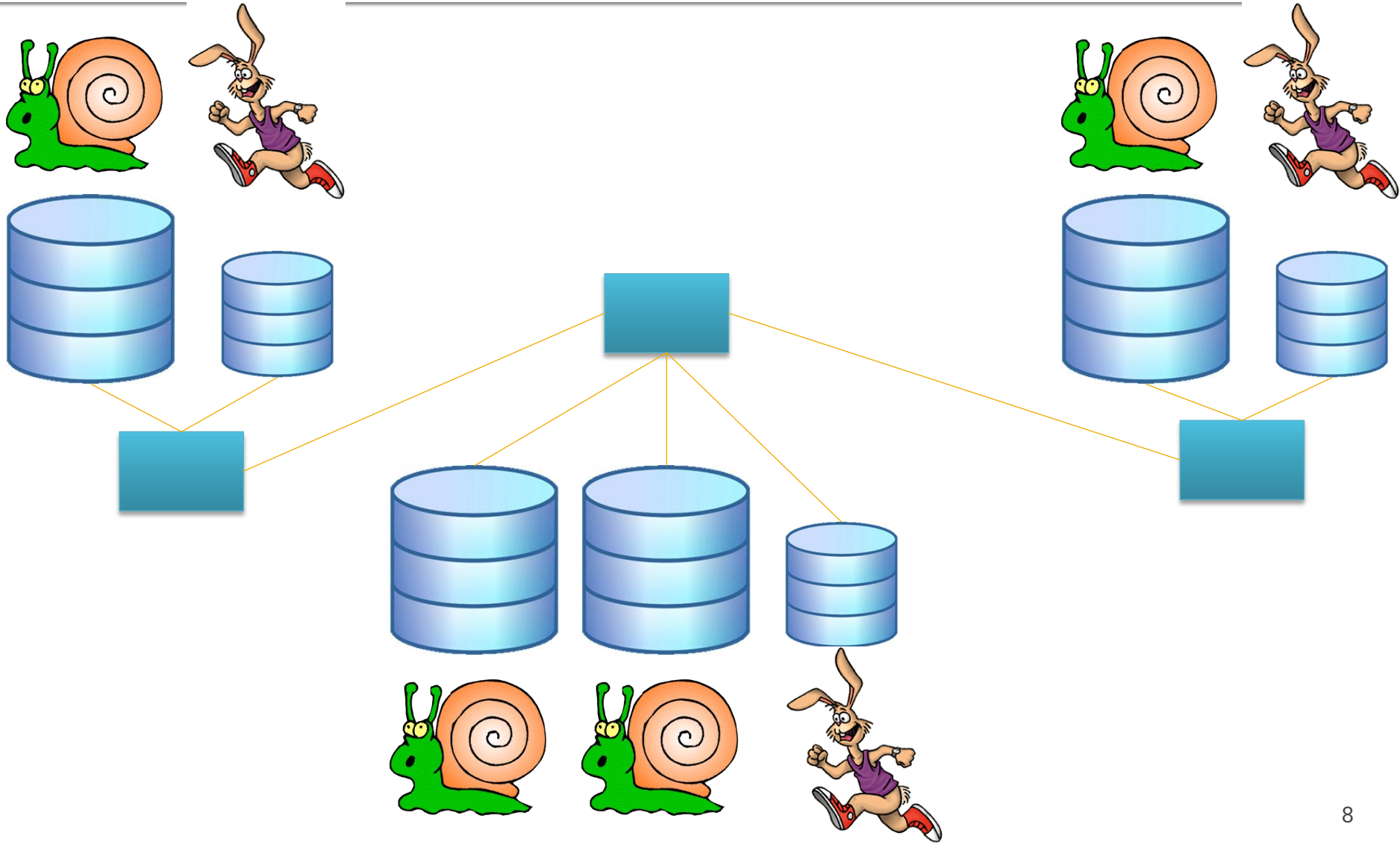
- SURFnet thinks there is a smarter way



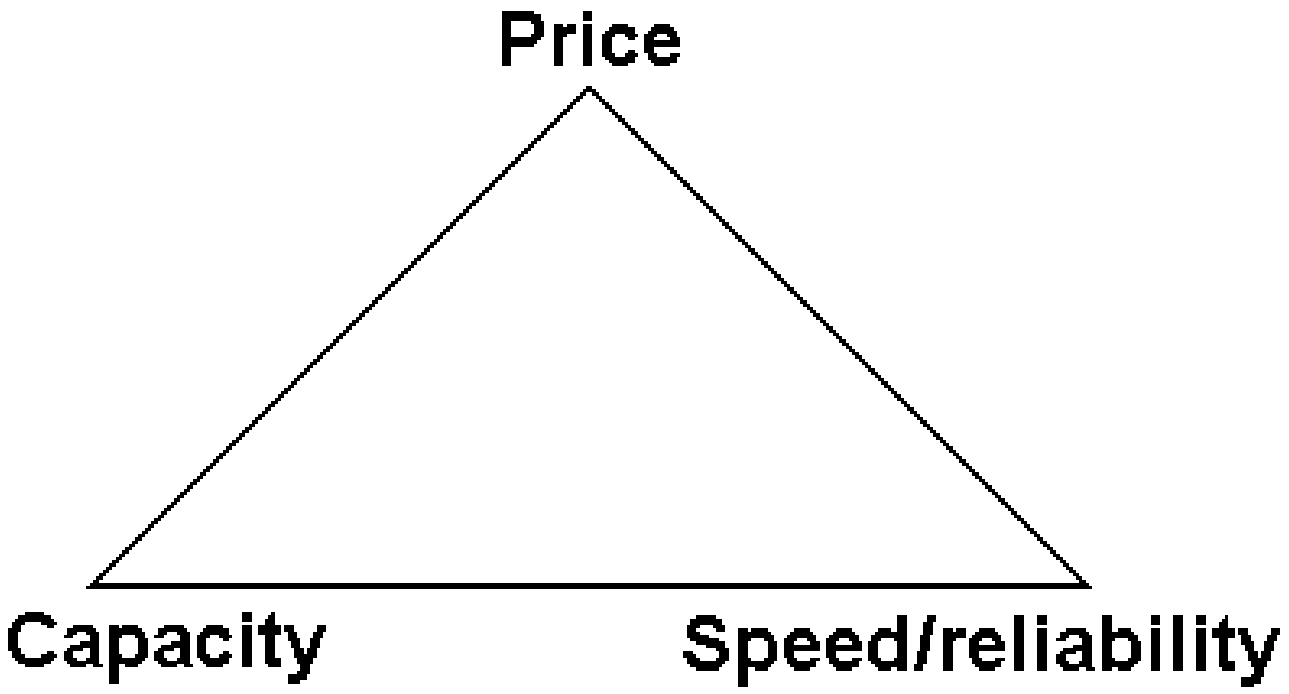
Background



Background



Background



Main question

- What infrastructure and open source tools provide SURFnet or the participants a scalable and distributed any-kind-storage solution?

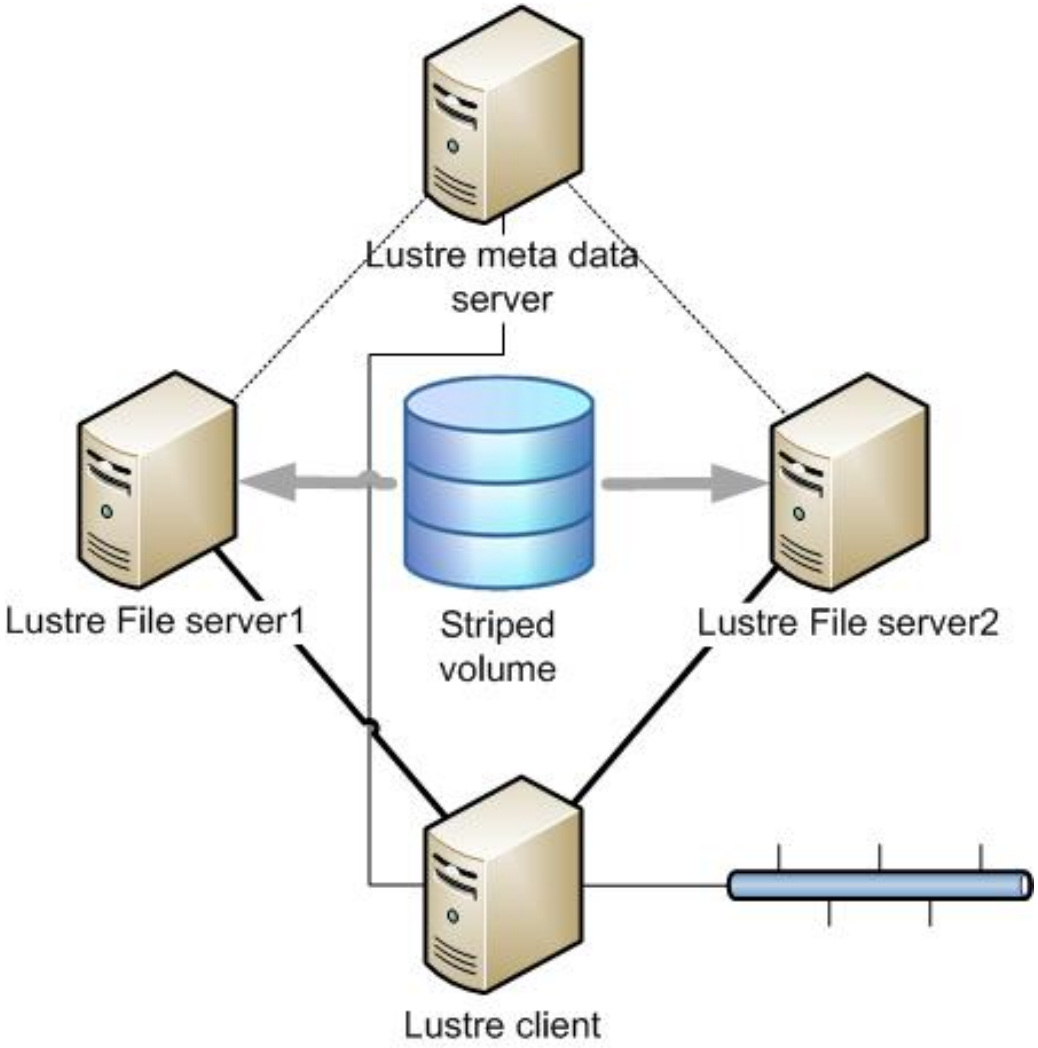
Product inventory

- Coda
- Lustre
- GlusterFS
- XtreemFS
- Ceph
- PVFS
- MooseFS
- Hadoop

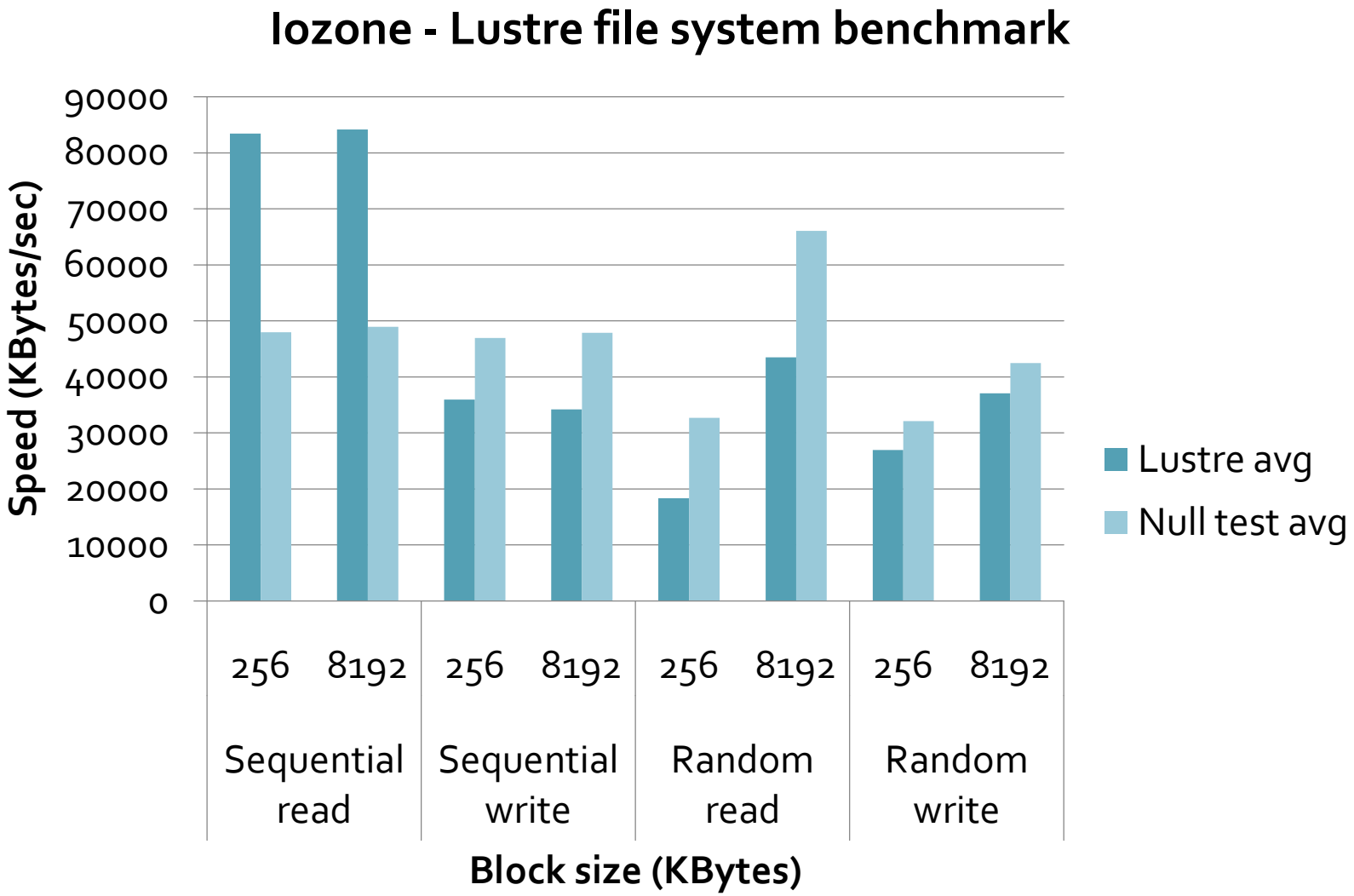
Test approach

- First stage
  - Null test
- Second stage
  - Architectures in the lab environment
- Third stage
  - Architectures in practice

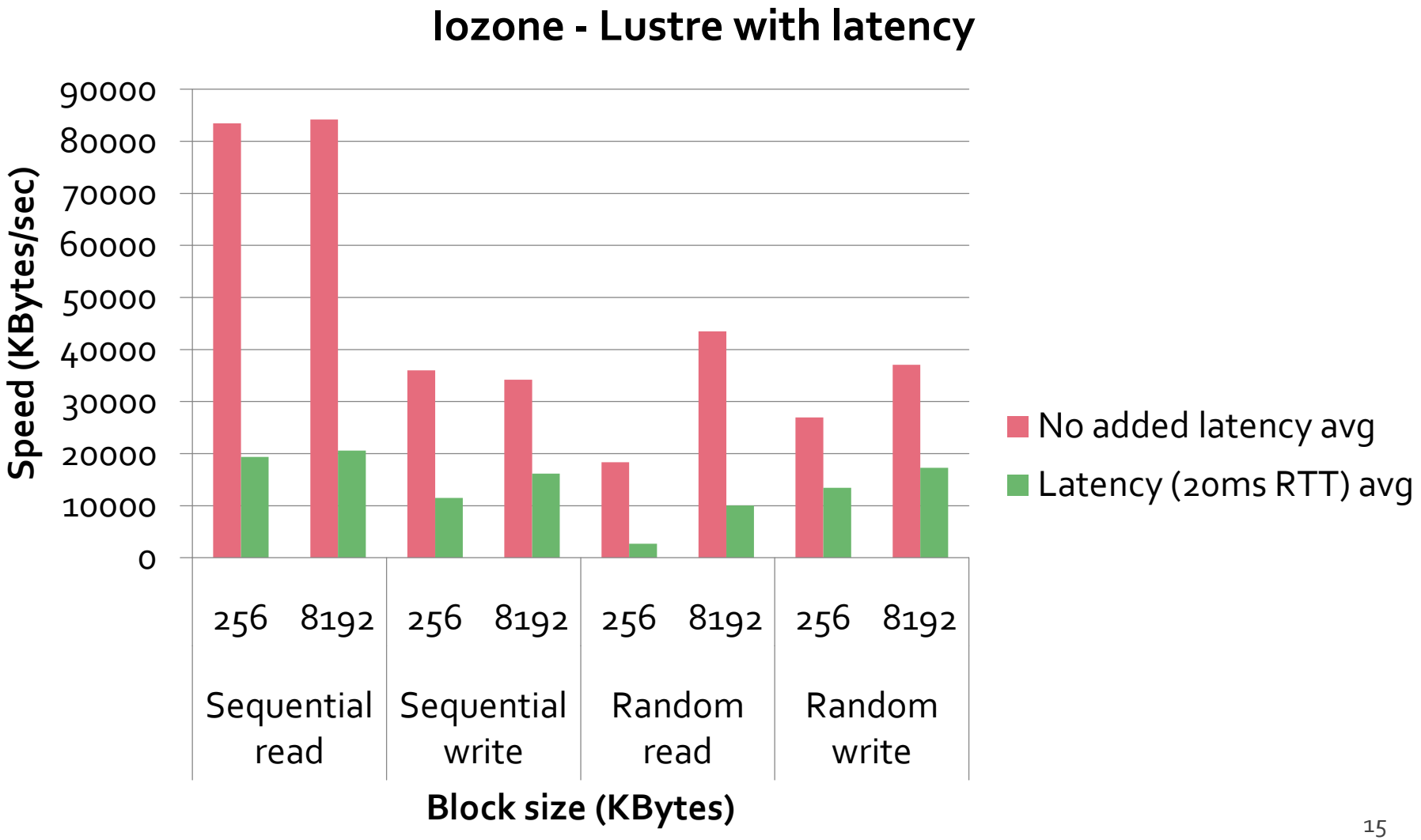
Lustre Test Approach



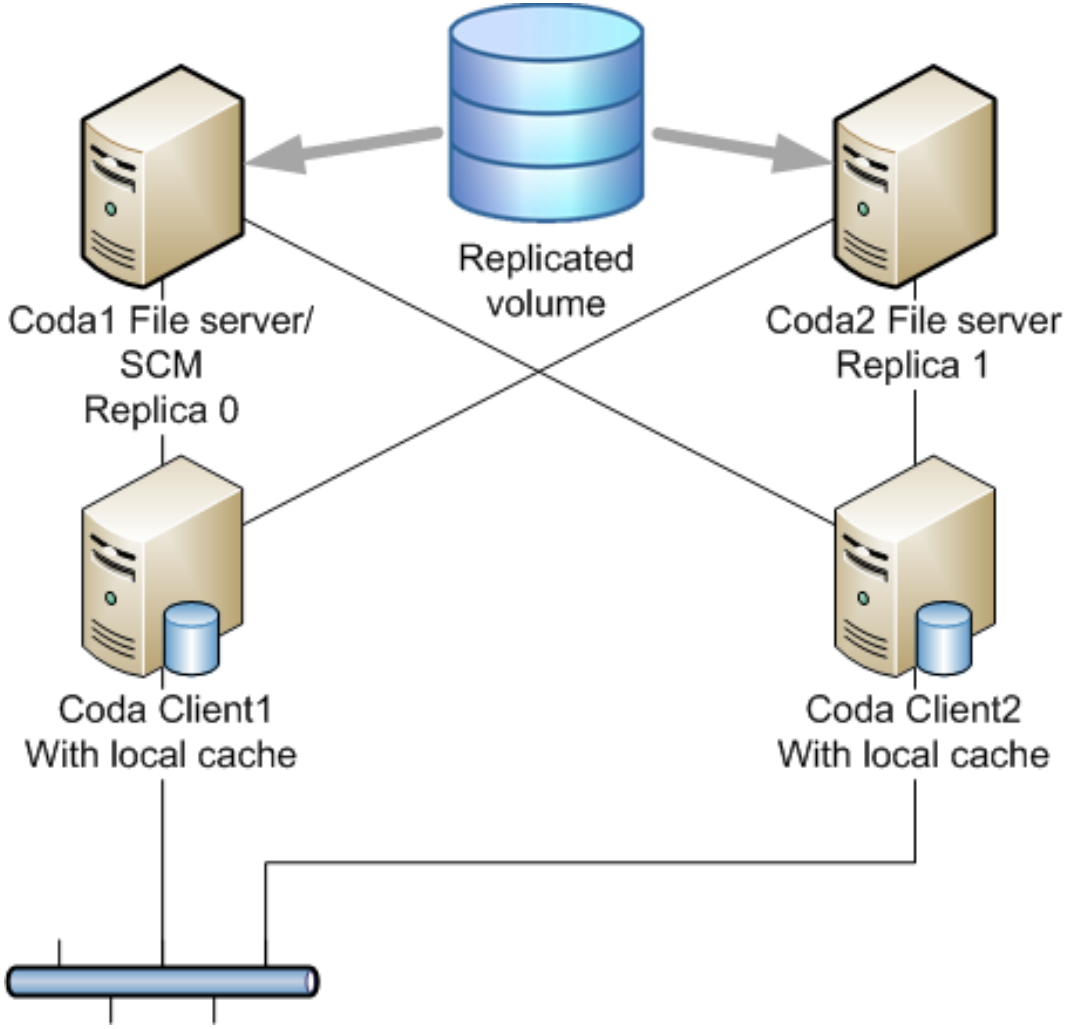
Lustre Performance



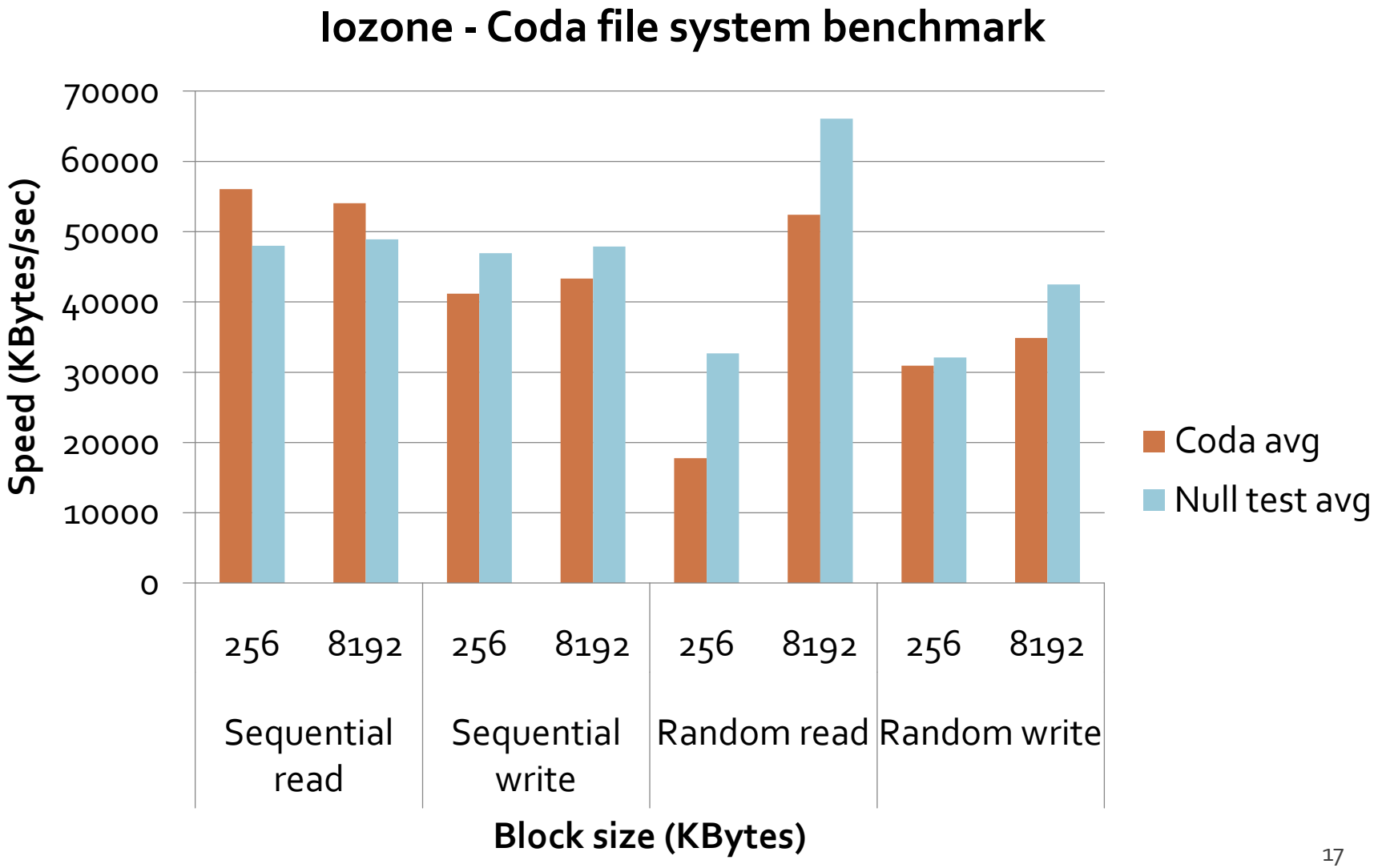
Lustre Performance – Latency impact



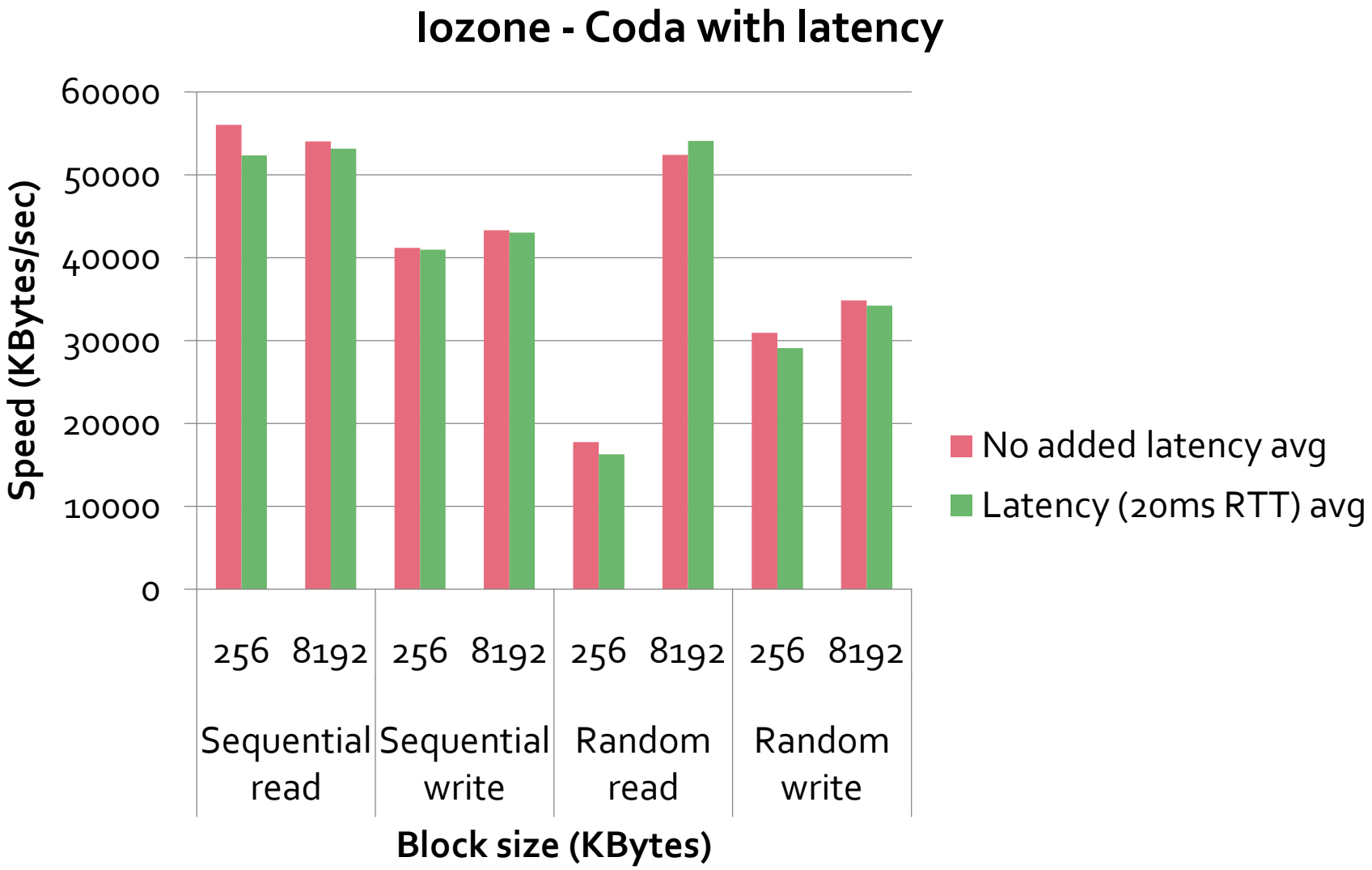
Coda Test Approach



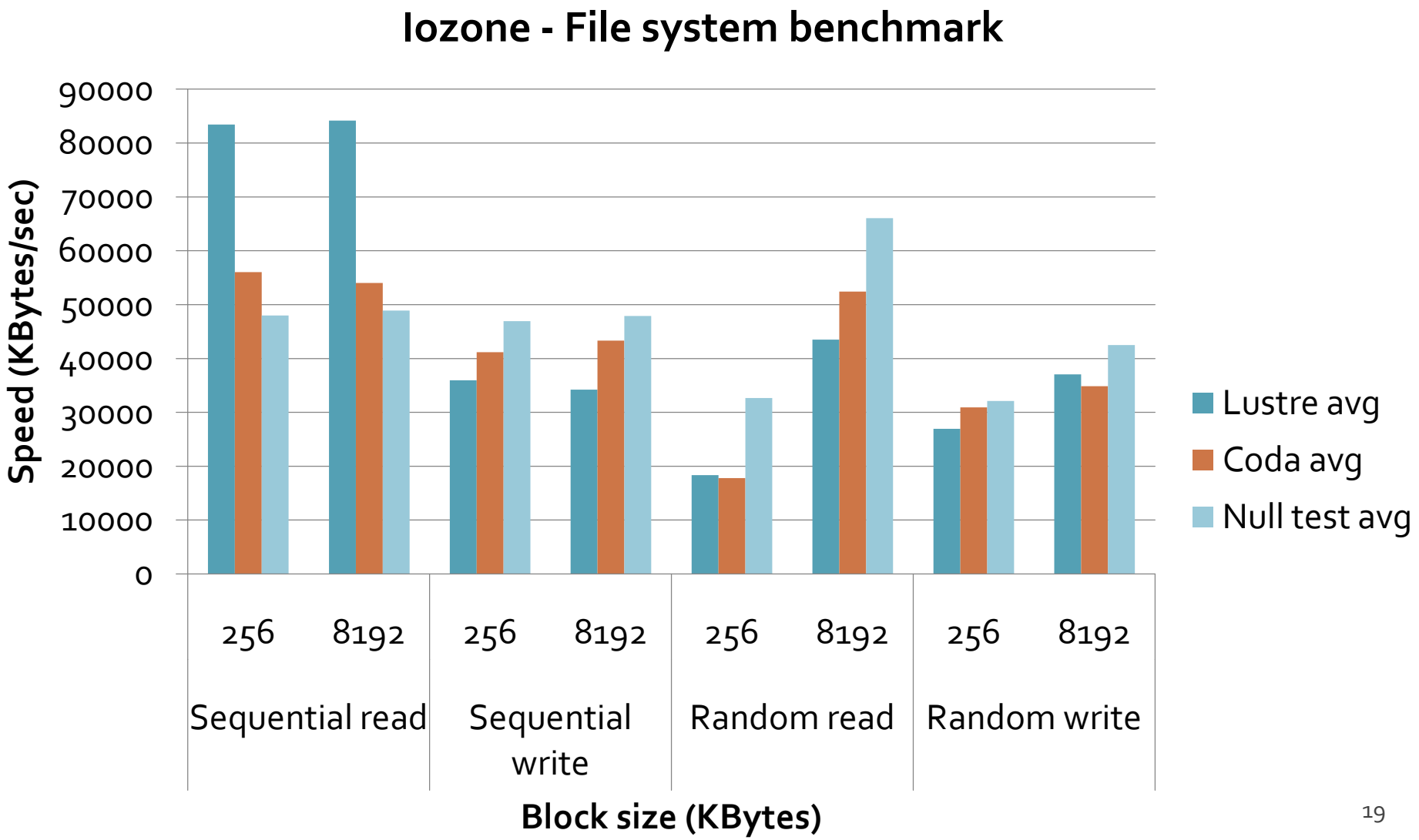
Coda Performance



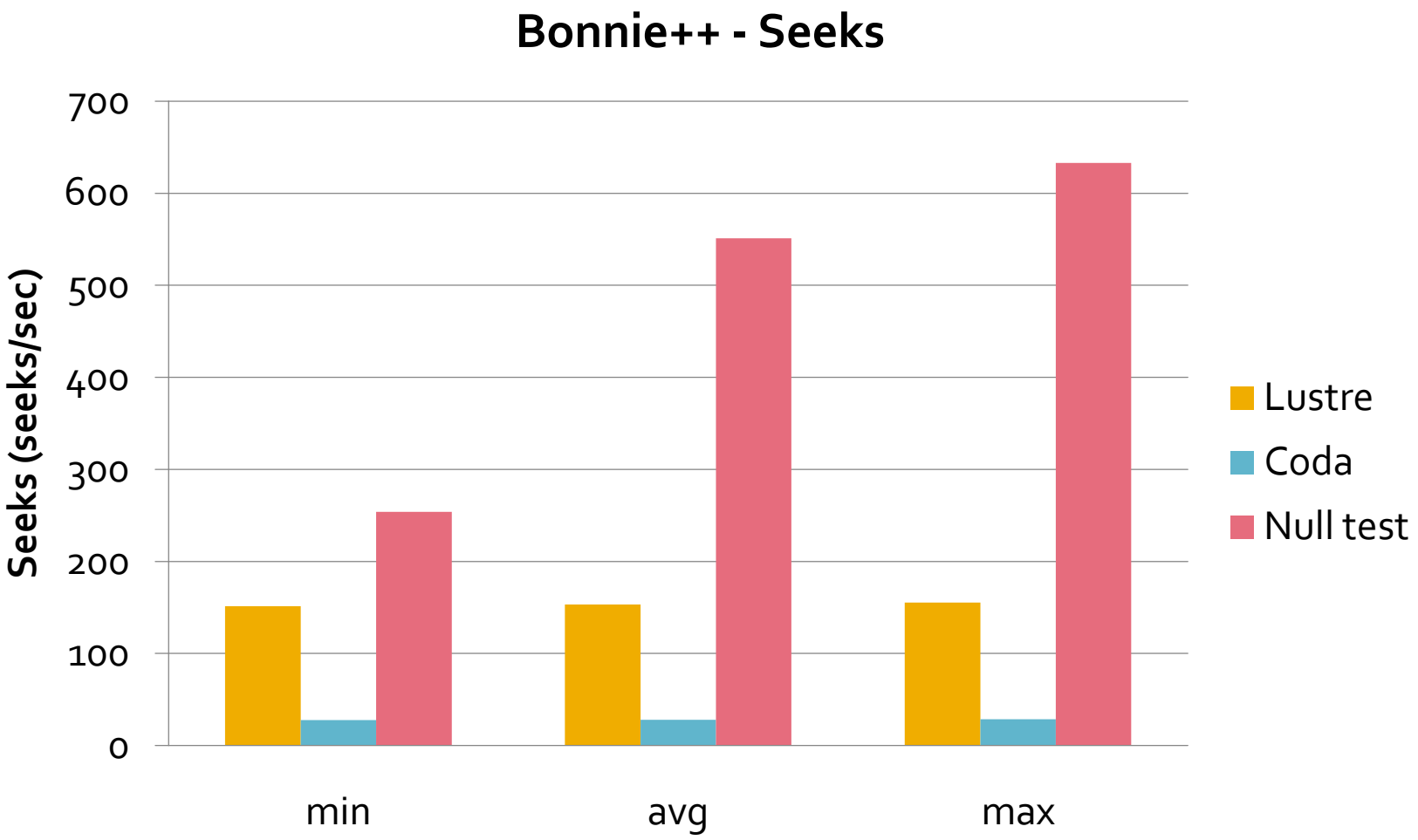
Coda Performance – Latency impact



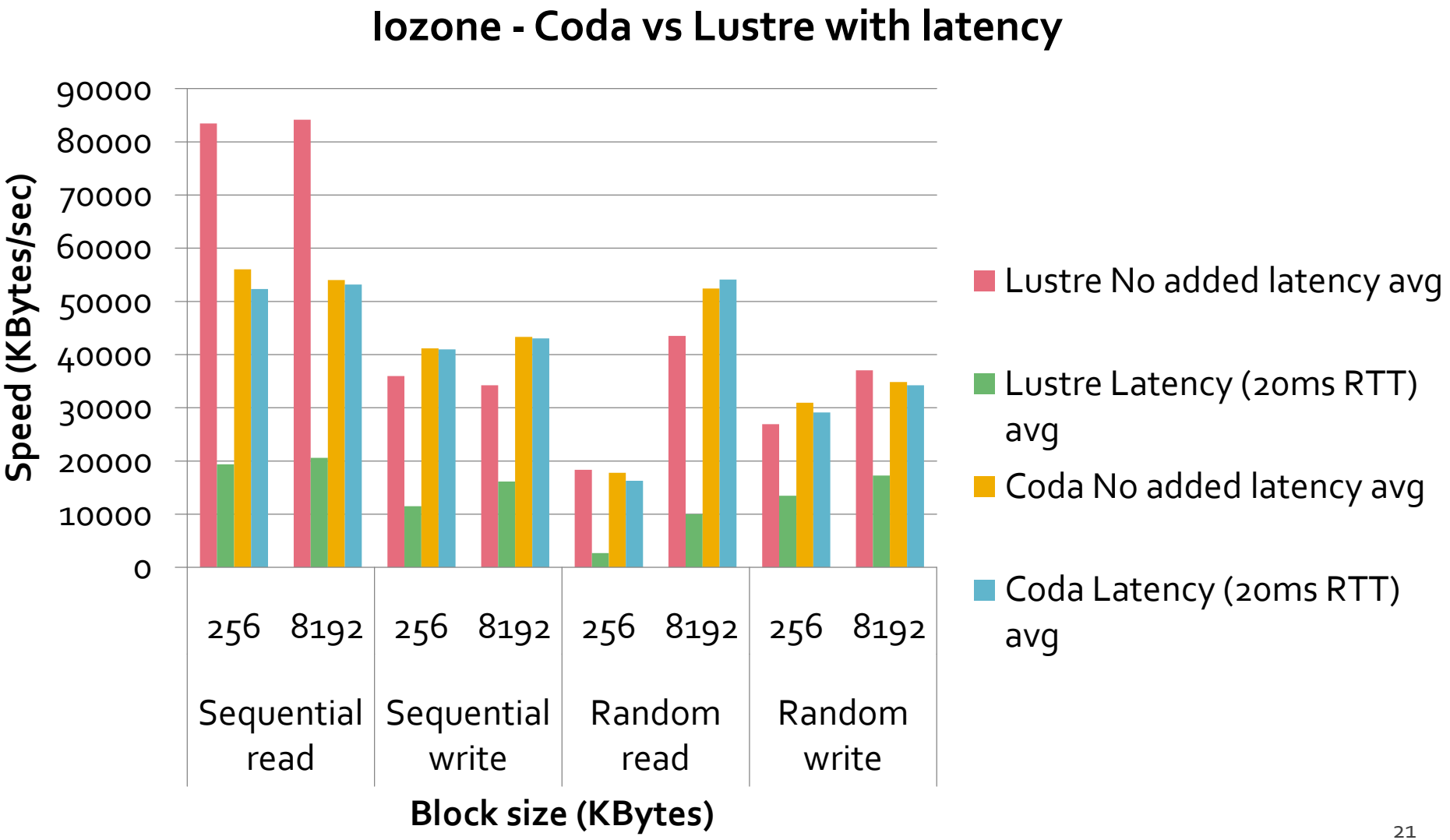
Coda vs Lustre - Performance



Coda vs Lustre - Seeks

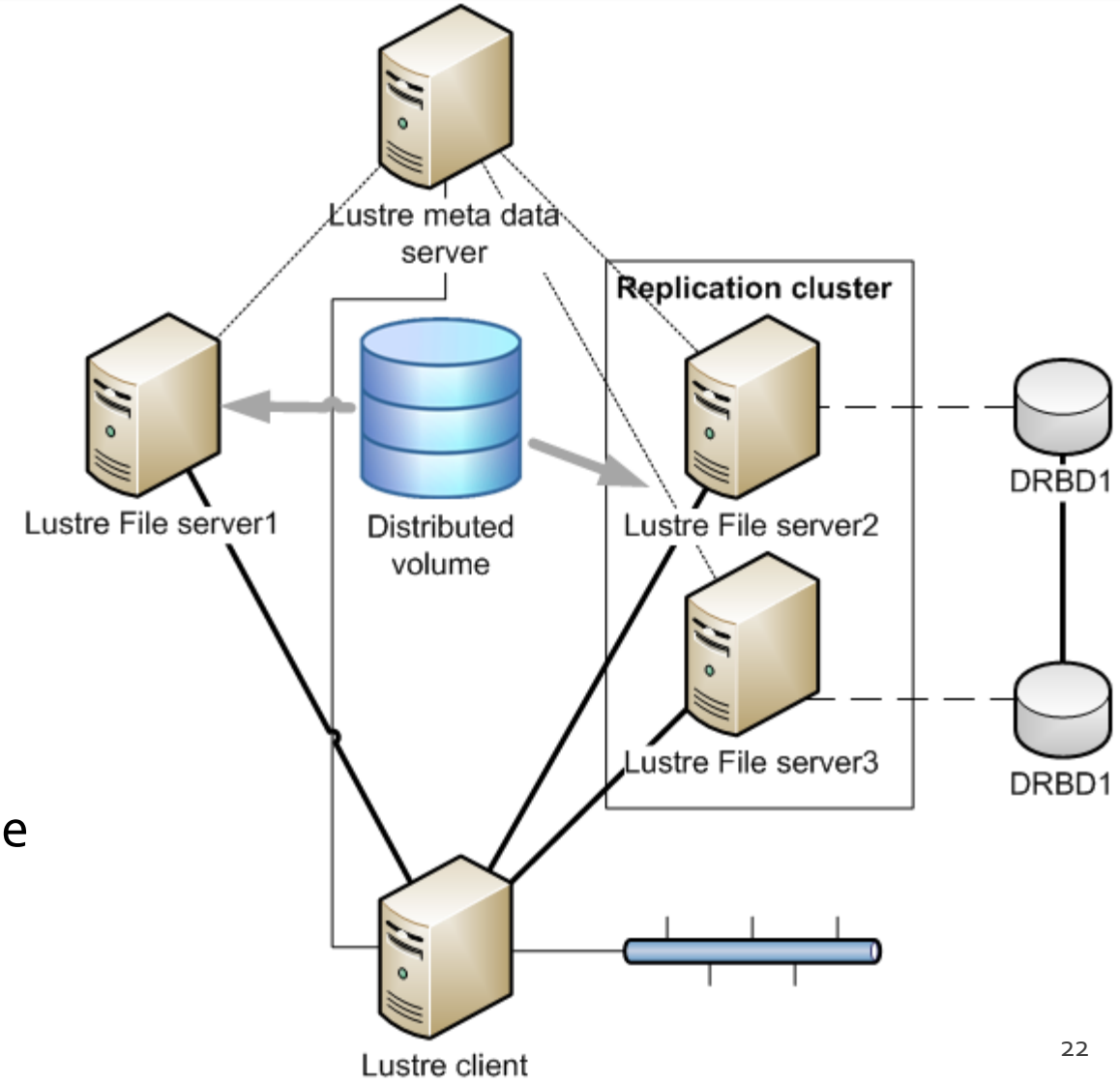


Coda vs Lustre – Latency Performance



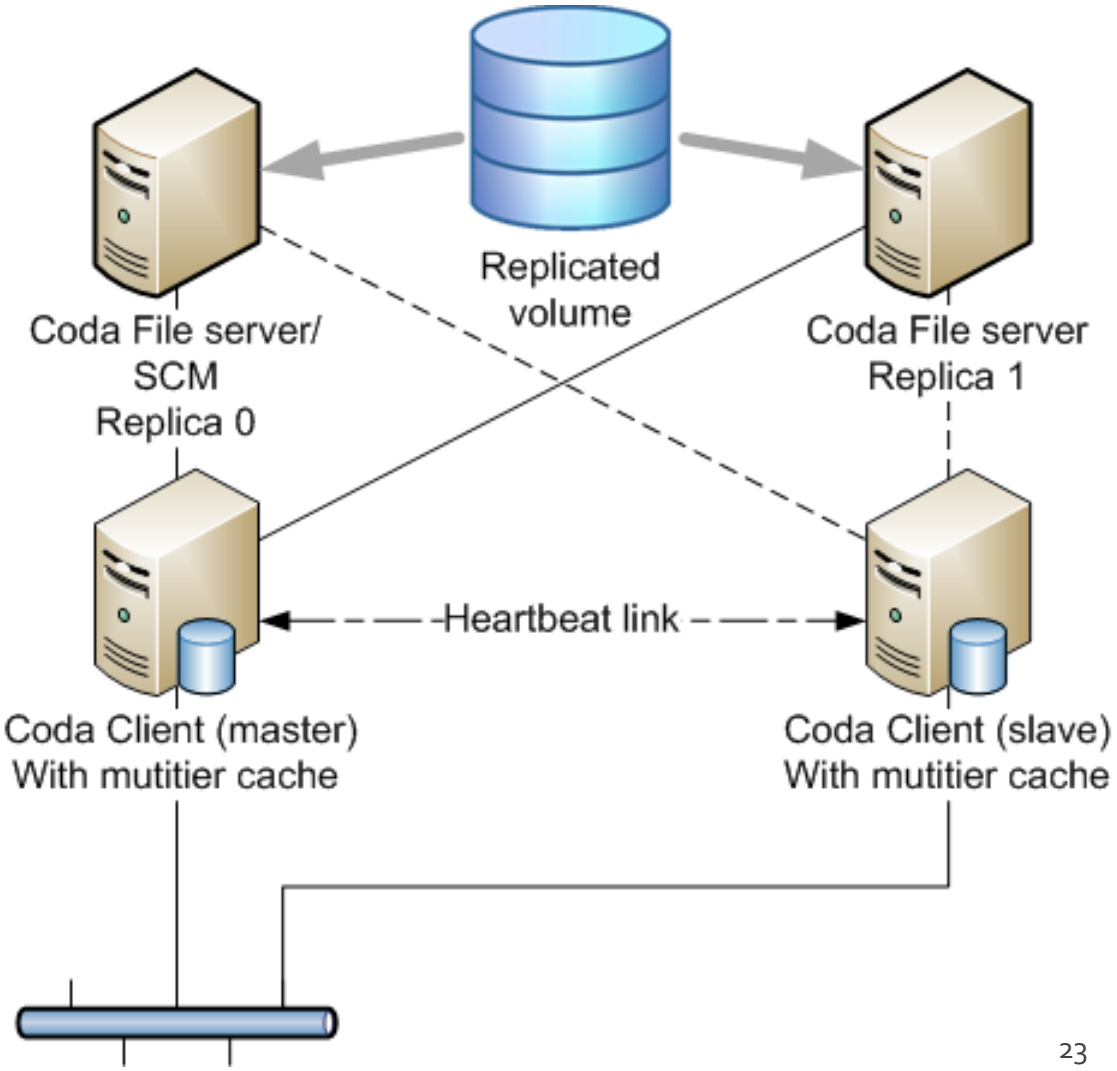
Lustre SURFnet Design

- Requirements
  - Scalable
  - Available
  - Durable
  - Performance
  - Dynamic
  - Cost effective
  - Generic interfaces
  - Open protocols
  - Geographic dispersion
- Deployment \ Maintenance
- Security features



Coda SURFnet Design

- Requirements
  - Scalable
  - Available
  - Durable
  - Performance
  - Dynamic
  - Cost effective
  - Generic interfaces
  - Open protocols
  - Geographic dispersion
- Deployment \ Maintenance
- Security features



Conclusion

- Coda:
  - Hard to configure
  - Can not handle large files
  - Promising architecture
  - Low latency impact
- Lustre:
  - DRBD is needed for replication / latency reduction
  - Latency has impact
  - No multi-tier
  - Promising future

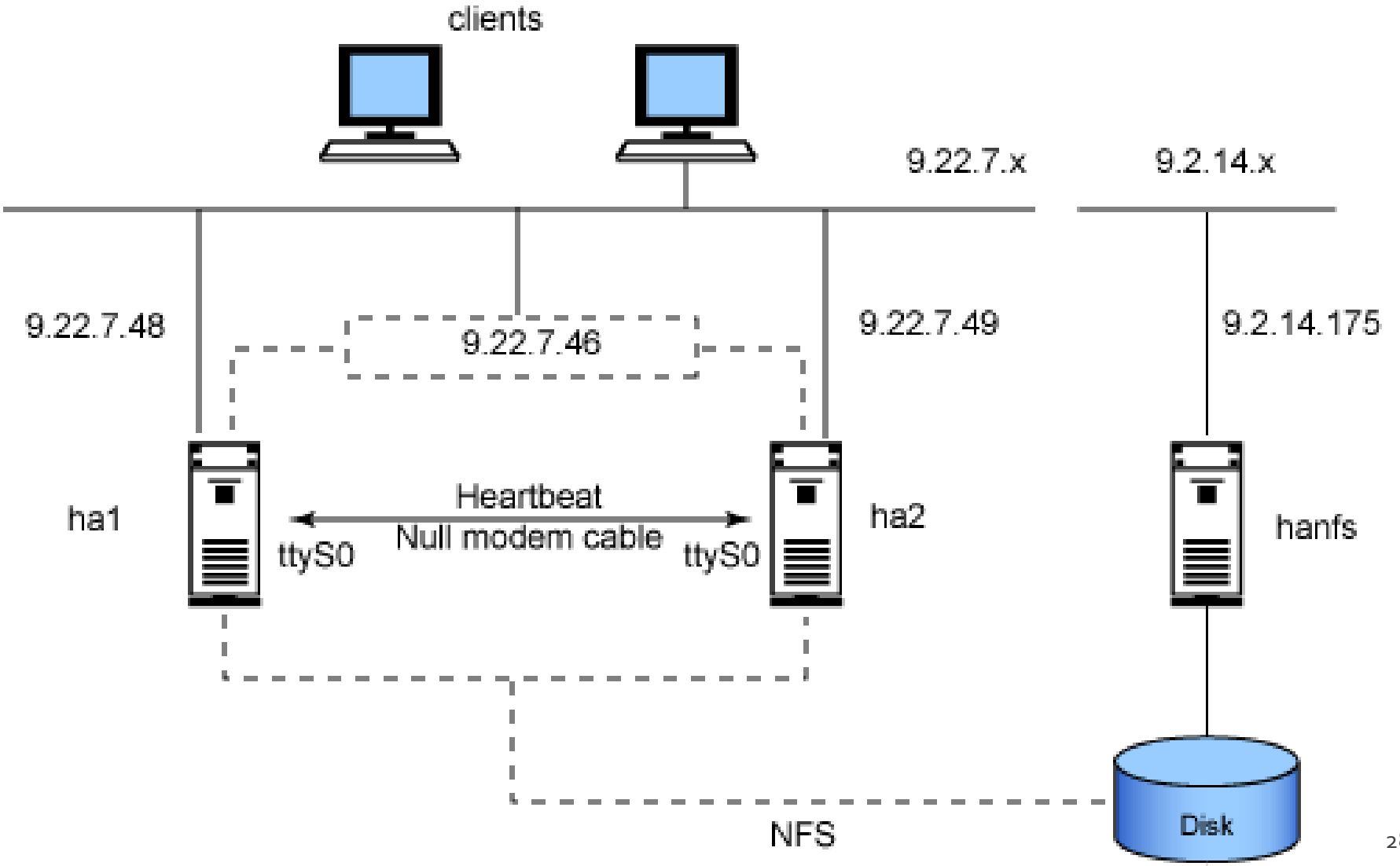
Future work

- Review of GlusterFS
- Test environment without VMWare ESXi
- Further tuning and configuring of Coda
- Coda and ZFS tests
- Lustre and DRBD tests

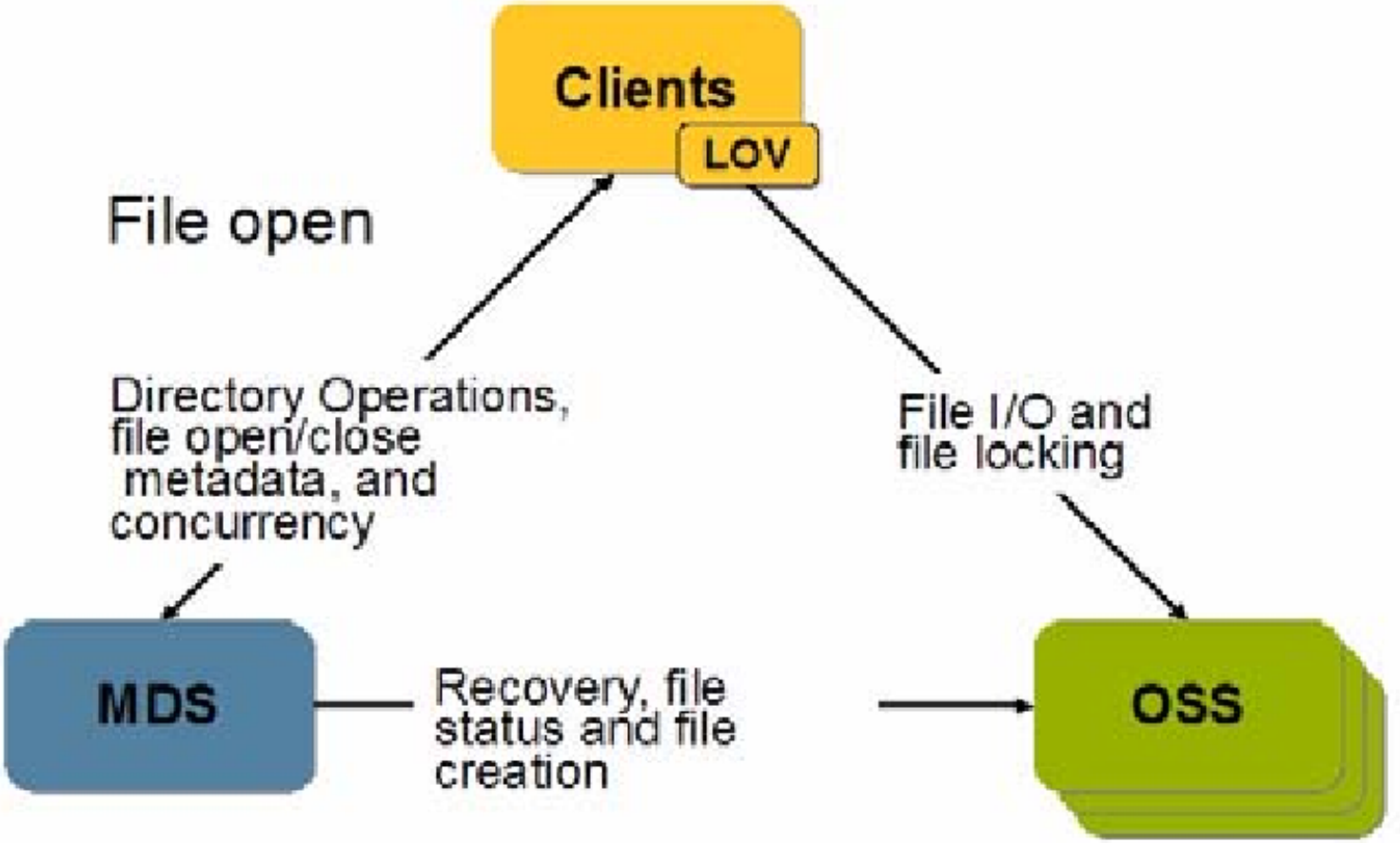
Questions & Feedback

(Source: mastergoogle.com)

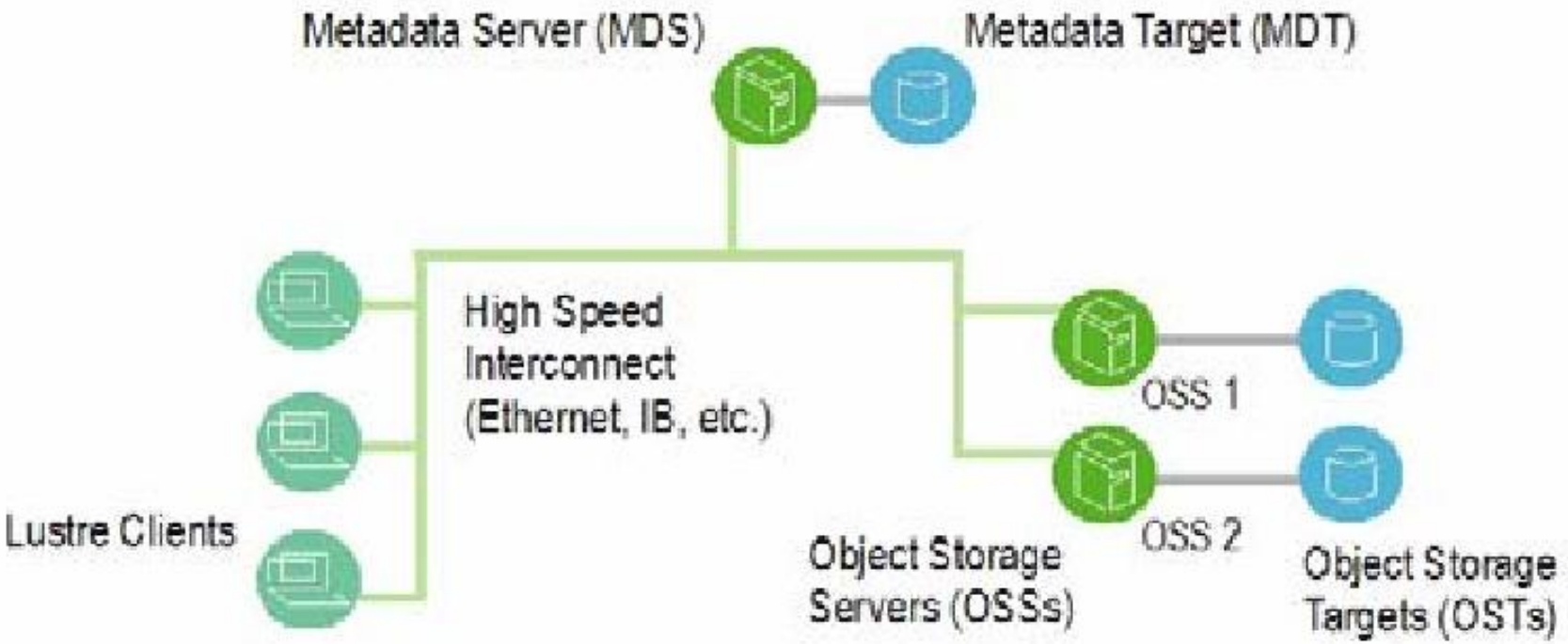
Appendix - Heartbeat



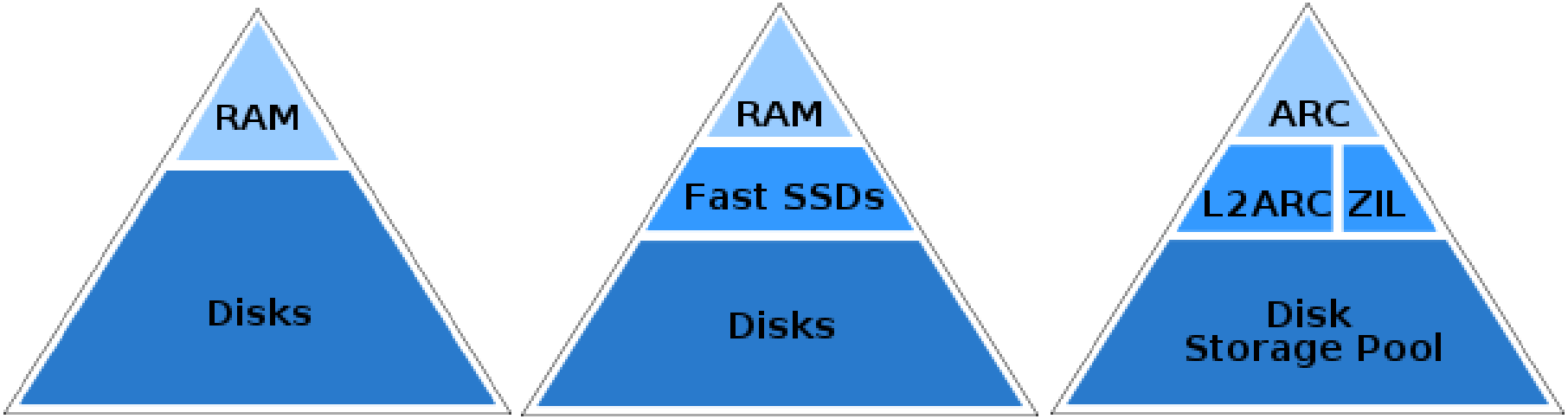
Appendix - Lustre



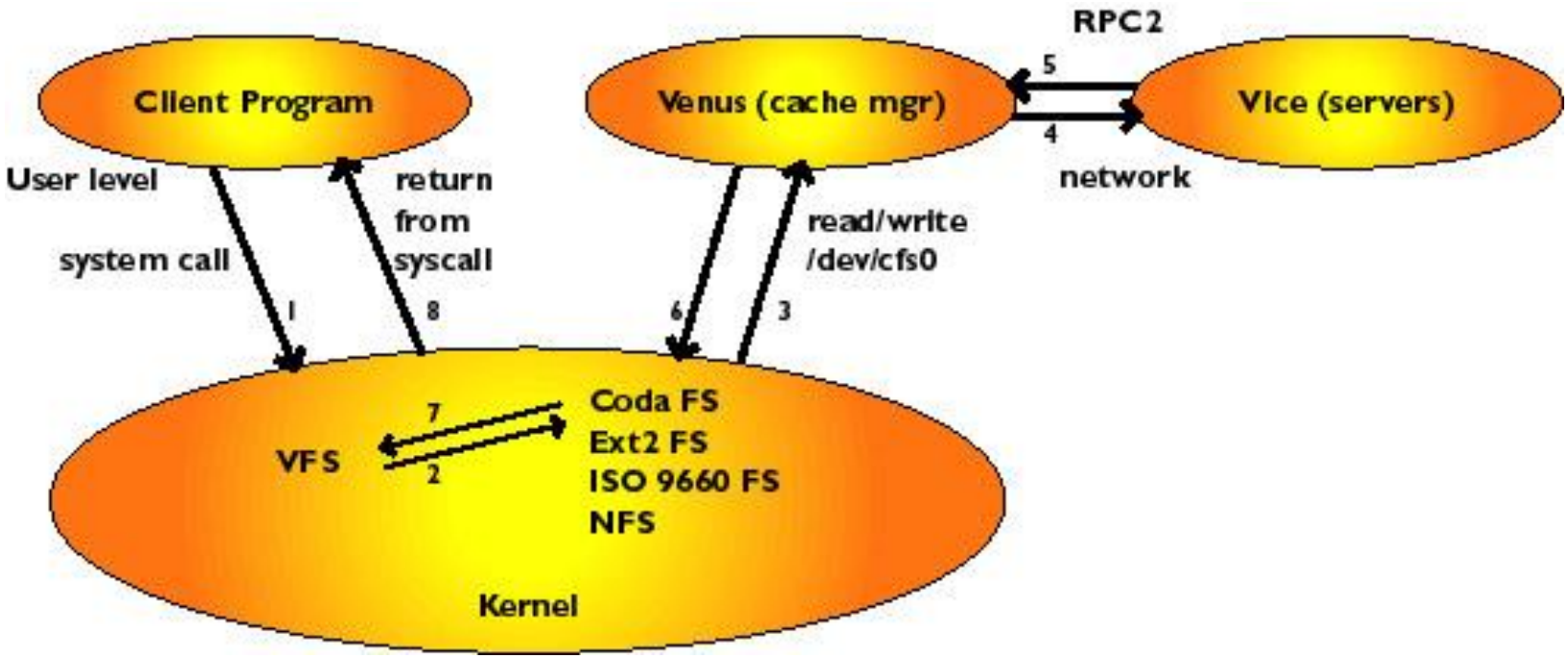
Appendix - Lustre



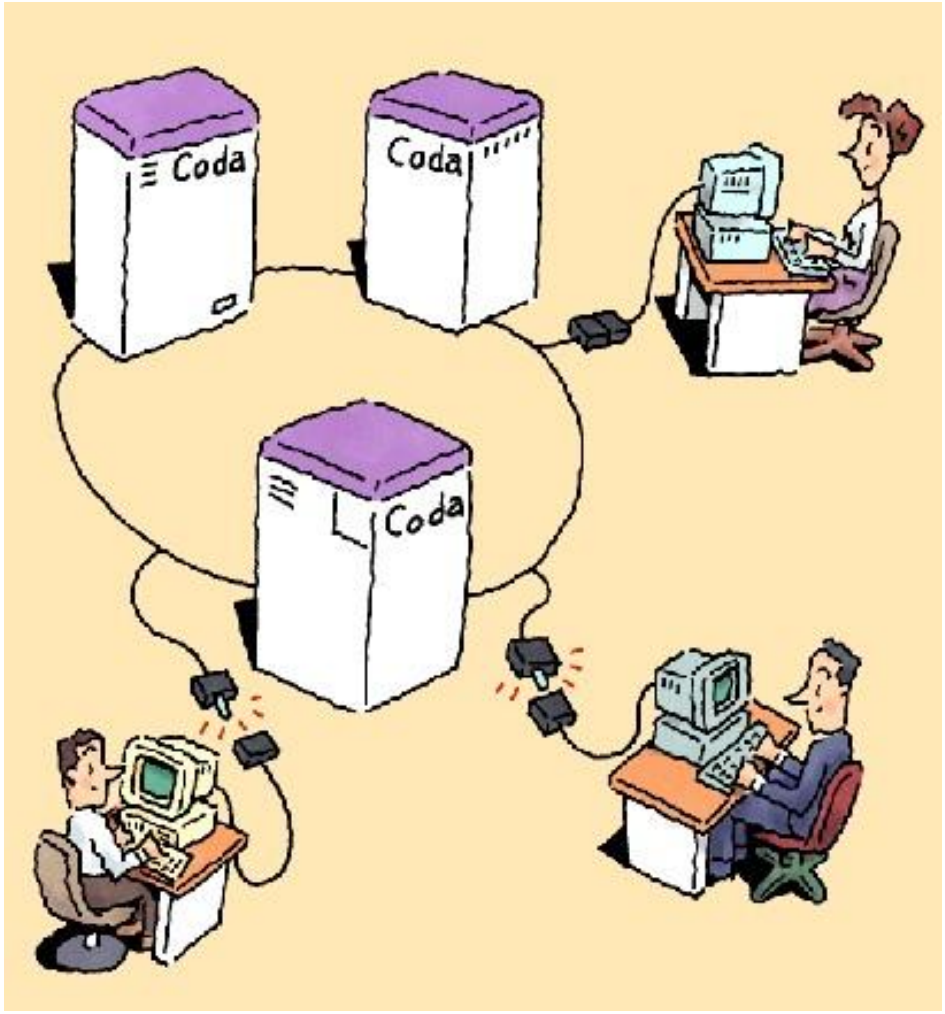
Appendix – ZFS L2ARC



Appendix - Coda



Appendix - Coda



Appendix – Seeks (seeks/sec)

- Coda

	min	avg	max
No added latency	27,6	28	28,4
Latency (20ms RTT)	9,5	9,8	10

- Lustre

	min	avg	max
No added latency	151	153	155
Latency (20ms RTT)	28,7	29	29,2