Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Hyrise-R: Scale-out and Hot-Standby through Lazy Master Replication for Enterprise Applications

David Schwalb, Jan Kossmann, Martin Faust, Stefan Klauck, Matthias Uflacker, Hasso Plattner

Hasso Plattner Institute, University of Potsdam, Germany

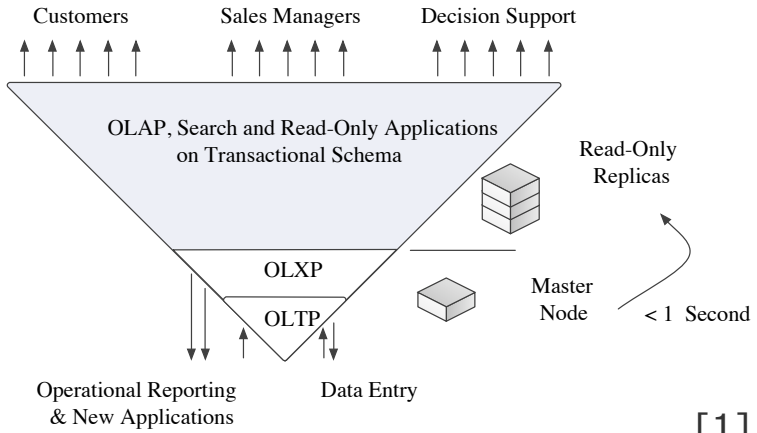
IMDM 2015

Motivation

New enterprise applications ..

- Growing number of users
- Increasingly complex queries
- Interactive data exploration

.. require scalability



[1]

Scale-up vs. scale-out

(+ availability)

Related Work

Theoretical replication models and comparison [2]

- Eager vs. lazy
- Group vs. master

Implementations

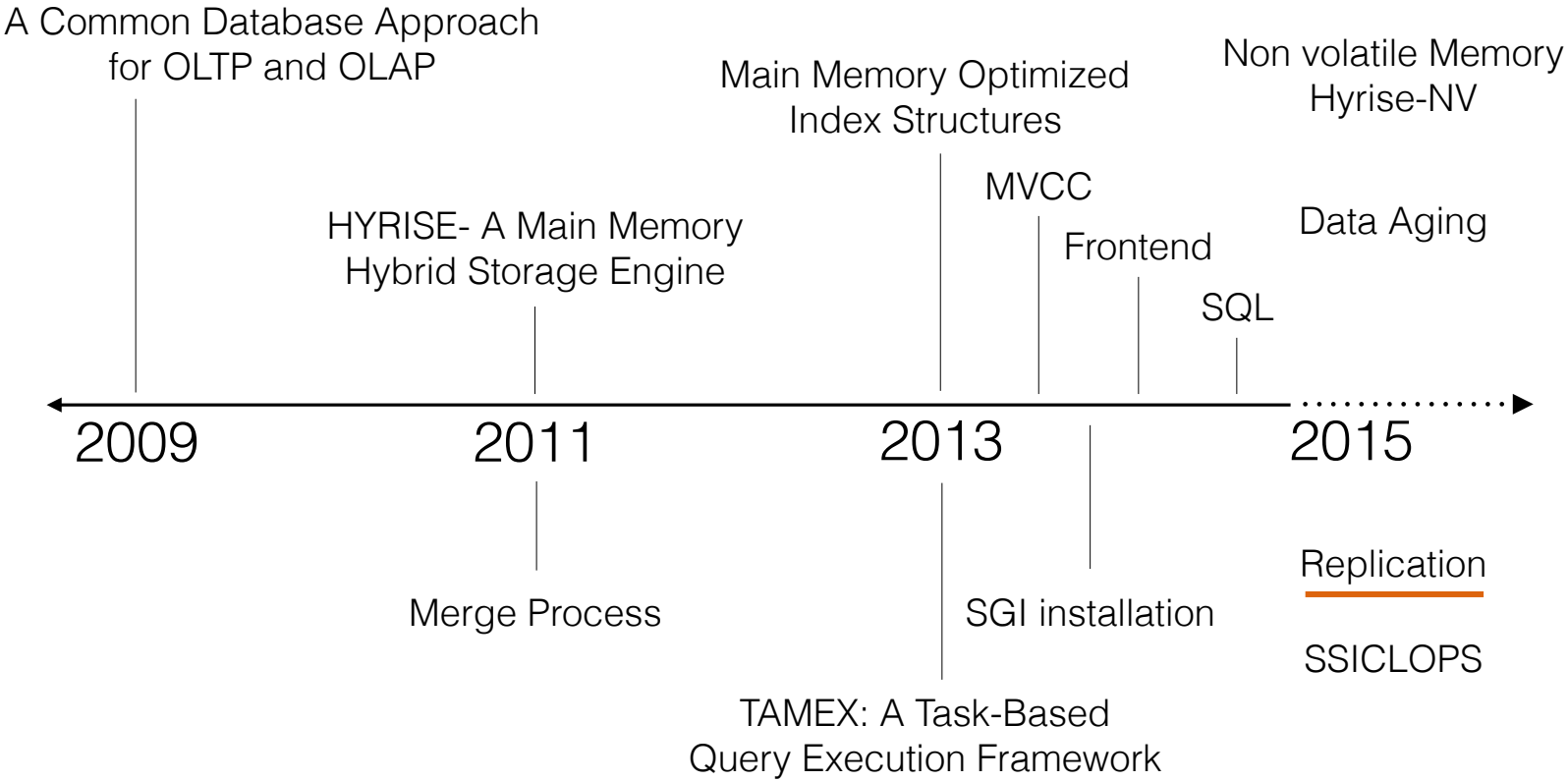
- Postgres-R – Eager group replication based on shadow copies [3]
- ScyPer – Lazy master replication with row layout for master node [4]
- ..

Hyrise

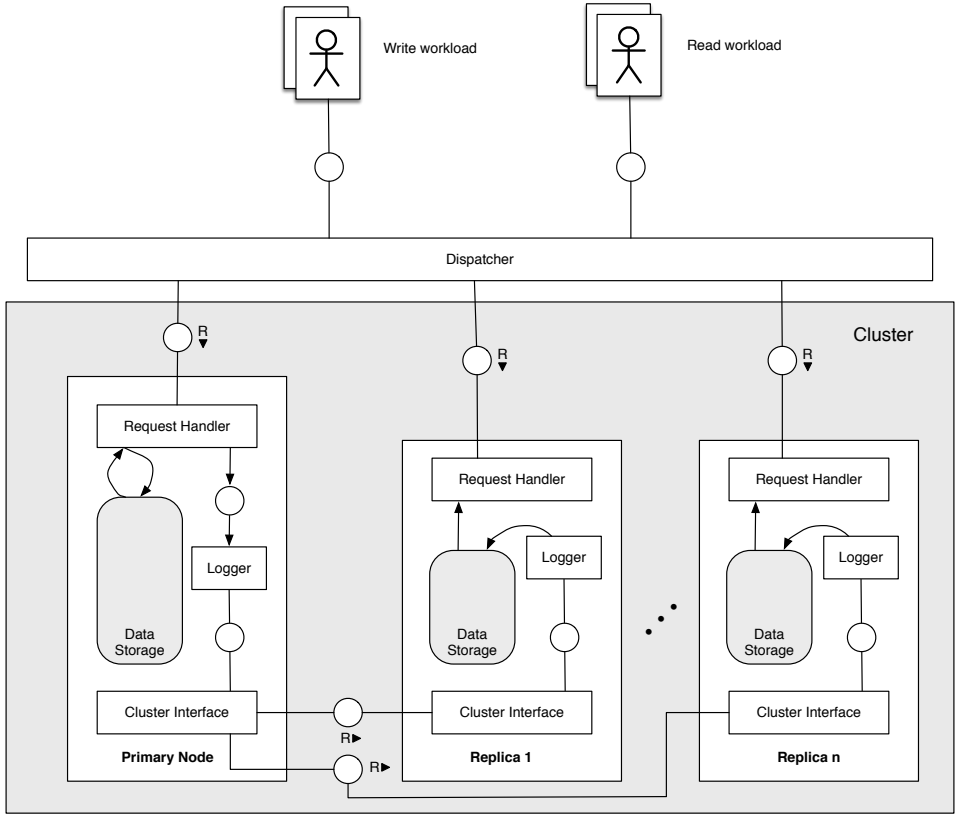
Storage engine developed at **HPI for research** and prototyping, initially focused on main memory processing and **hybrid storage layouts** of tables

- Dictionary and bit-vector compression
- Main/delta architecture with merge process
- Hybrid row and column layouts of tables
- Supports vertical and horizontal partitioning

Hyrise – Research History



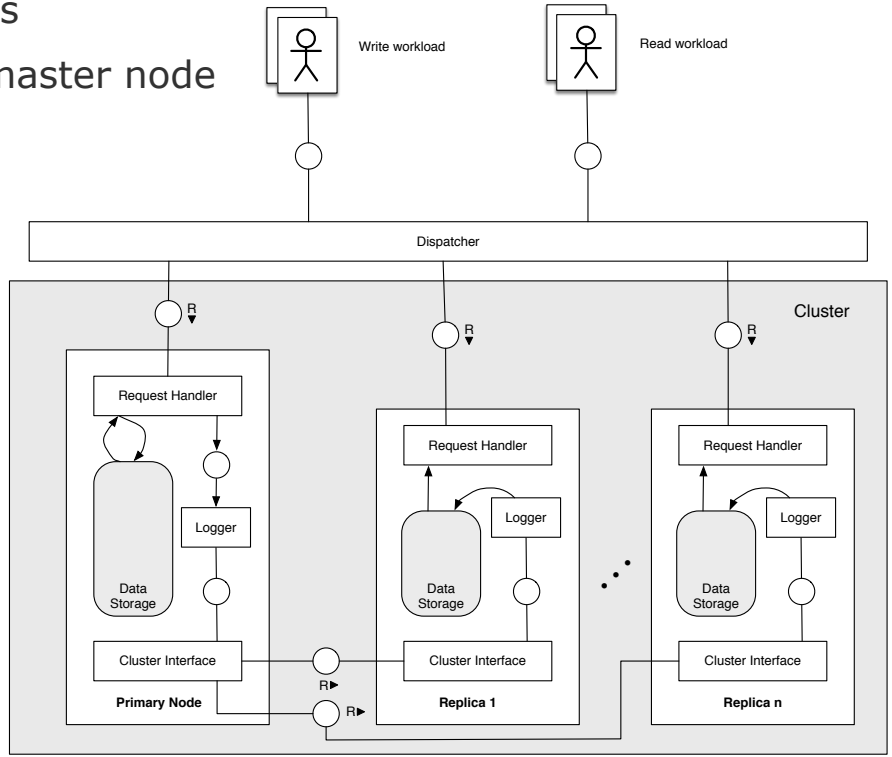
Hyrise-R


Write workload
Read workload
Dispatcher
Cluster
R
R
R
Request Handler
Logger
Data Storage
Cluster Interface
Primary Node
Request Handler
Logger
Data Storage
Cluster Interface
Replica 1
Request Handler
Logger
Data Storage
Cluster Interface
Replica n
R
R

Dispatcher

Redirect queries to cluster nodes

- Transactional workload -> master node
- Reads -> all cluster nodes



Replication Mechanism

Logs are written to file system + send to cluster interface

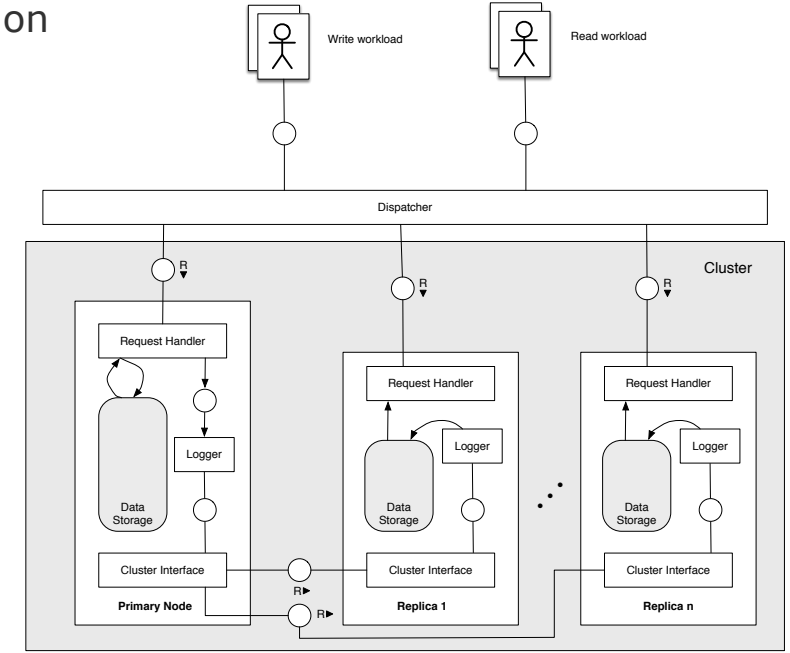
Cluster interface sends (dictionary encoded) log information to replicas

Frequency is configurable and based on

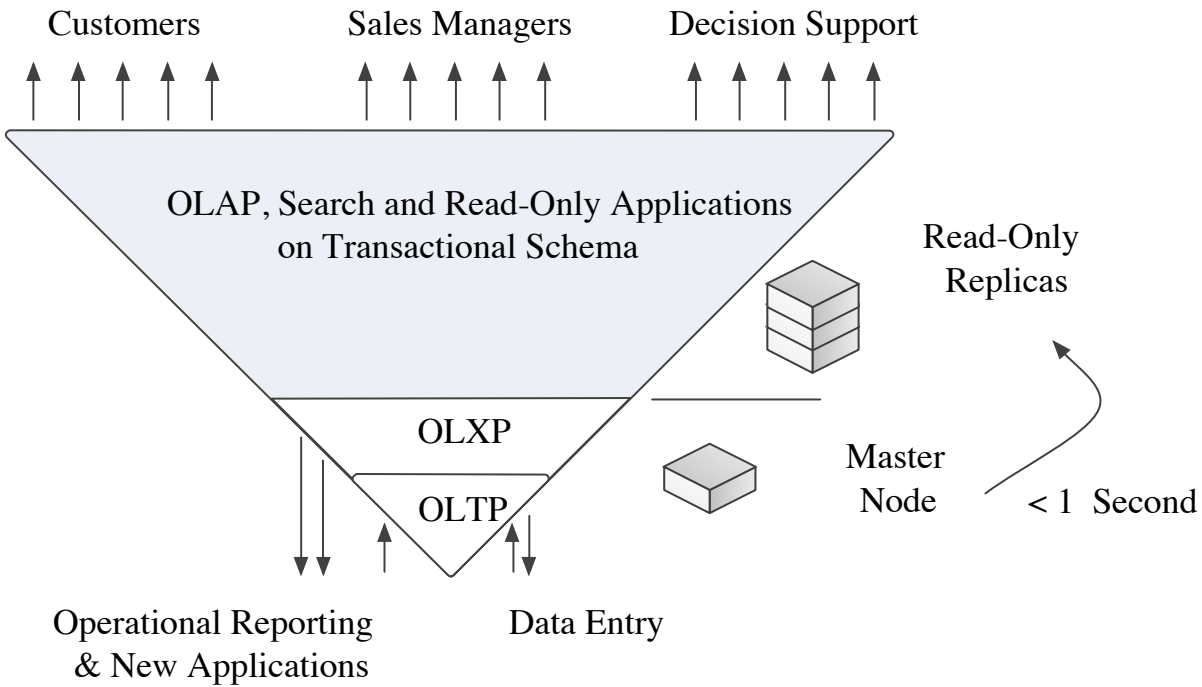
- Number of calls
- Exceeding buffer size
- Time since last transmission

TCP with nanomsg

- Survey pattern allows replicas to acknowledge reception
- Heartbeat protocol for failover



Why Hyrise-R is a good fit for Enterprise Applications

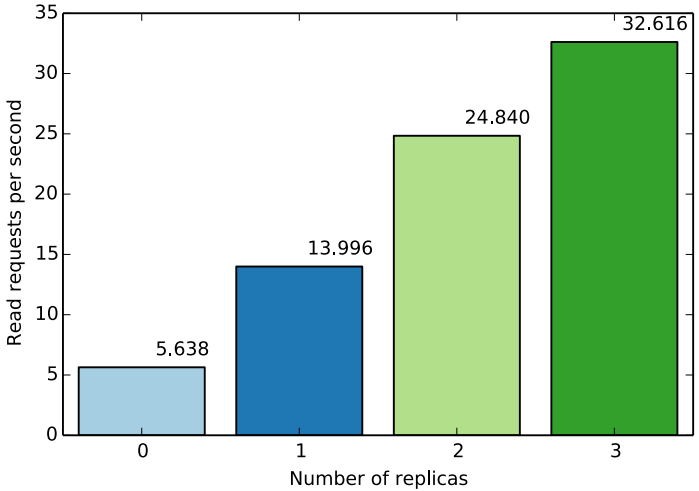


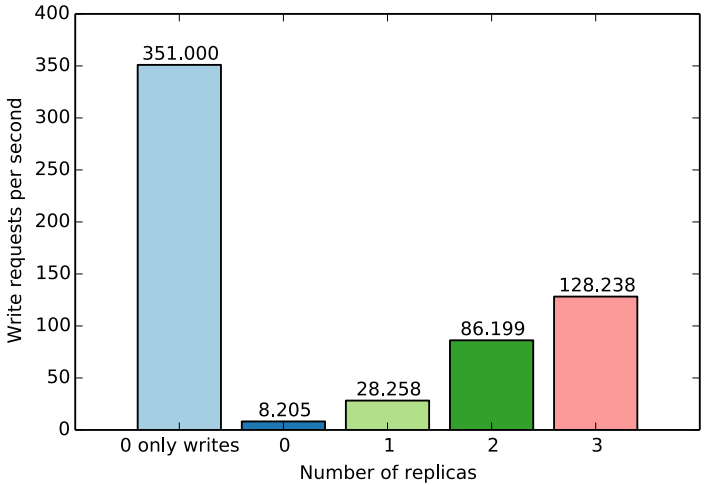
[1]

Evaluation
on Amazon EC2 cluster

5 machines with

- Intel Xeon E5-2666 v3 (36cCPUs; 10 cores @ 2.9GHz)
- 60 GiB main memory





Conclusion

Hyrise-R – a system to cluster Hyrise instances using lazy master replication

- Dictionary compressed logs for updating replicas
- Heartbeat protocol for failover
- Benchmarks on Amazon EC2 cluster

Future Work

- Extend query dispatching and distribution
- Extend mixed workload measurements (ch-beCHmark)

References

[1] H. Plattner. The Impact of Columnar In-Memory Databases on Enterprise Systems. 2014

[2] J. Gray, P. Helland, P. O'Neil, and D. Shasha. The dangers of replication and a solution. 1996

[3] B. Kemme and G. Alonso. Don't be lazy, be consistent: Postgres-r, a new way to implement database replication. 2000

[4] T. Mühlbauer, W. Rödiger, A. Reiser, A. Kemper, and T. Neumann. Scyper: Elastic olap throughput on transactional data. 2013

Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Thanks

Stefan Klauck
stefan.klauck@hpi.de
http://epic.hpi.de