Syllabus

- Introduction to fault-tolerant distributed systems
- Models of distributed systems and related faults
- Data replication
- Distributed consensus
- State machine replication
- **Database replication**

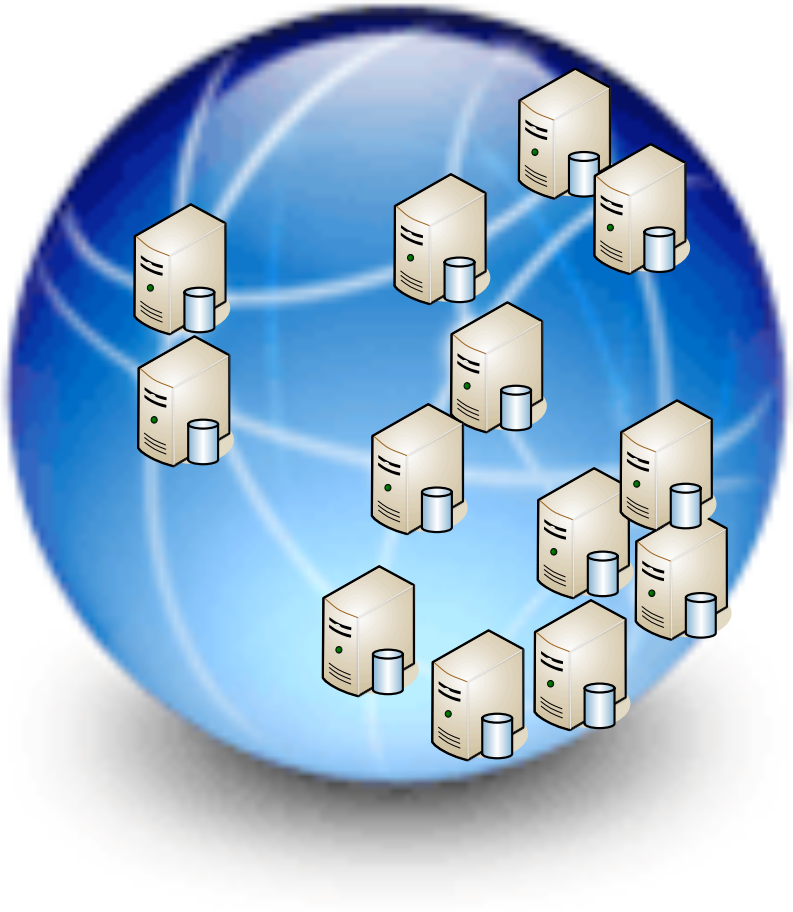
Database replication

- Basic Approaches
- Specific Database Replication Protocols

Database replication

Transactional database management systems

- Transactional requests
- Multi-interaction requests
- Highly concurrent servers
- Replication goals
  - Availability
  - Performance
- Replication challenges
  - Scalability under generic workloads
  - Preserved Consistency Criteria



Database replication

Basic approaches

- The replication of a DBMS is a particular case of state machine replication (not data!)
- DBMS typically process transactions which are special beasts with particular “life styles and social etiquette”
- DBMS tend to be highly concurrent systems and may adopt several concurrency control mechanisms. Concurrency is a source of non determinism!
- The replication of a DBMS tend to be based on the primary/backup technique, ie. passive replication, and is usually distinguished as being either:
  - Synchronous or Asynchronous
  - Single or Multi-master

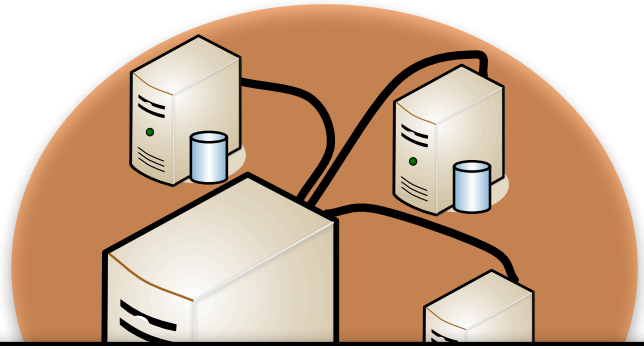
Database replication

Consistent replication requirements

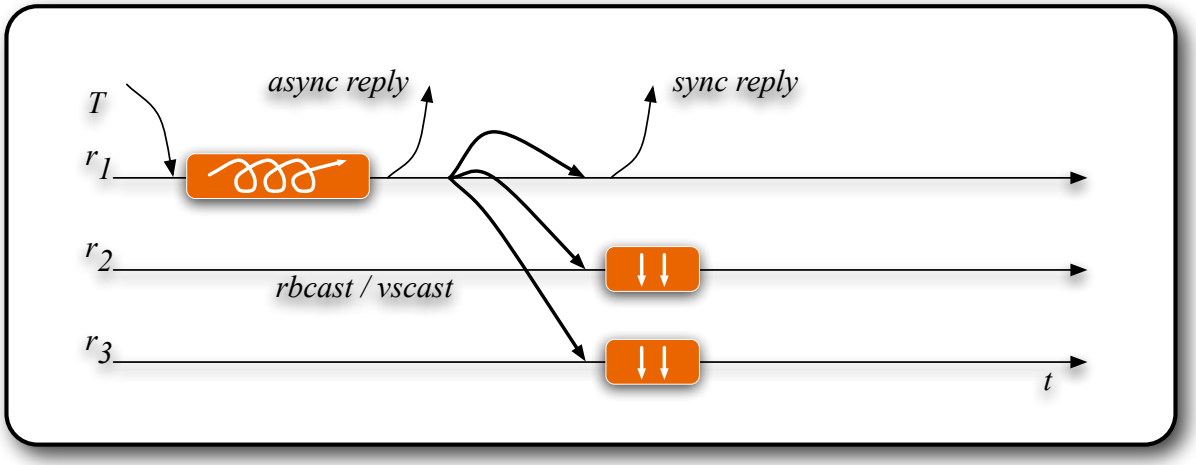
- Strictly consistent database replication to provide transparent distribution and fault tolerance
  - Strictly consistent replicas on transaction boundaries
  - No reconciliation
  - Automatic fail-over without loss of updates
- Focus on group communication based techniques instead of traditional distributed locking and atomic commitment protocols
- Multi-master, update-anywhere protocols

Database replication

Primary-backup

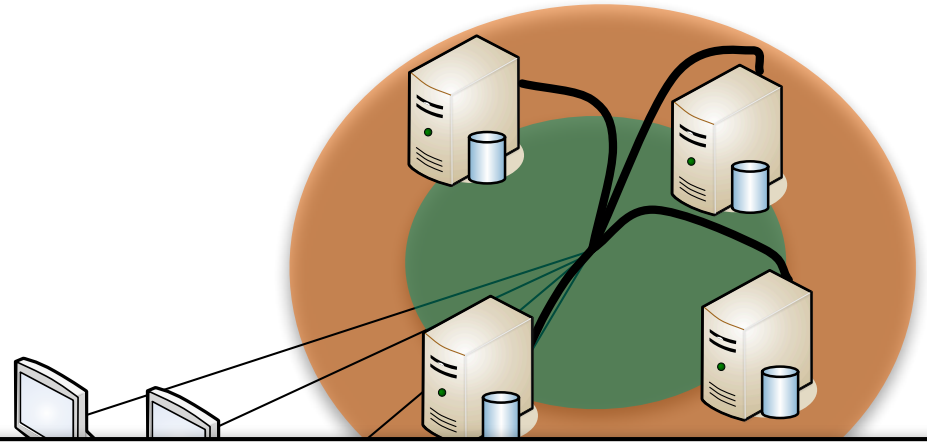


- Asynch PB is standard replication in most DBMS
- Asynch vs. synch update propagation
- Adequate to handle non-deterministic servers
- Extensible to multi-master (partitioning, reconciliation)

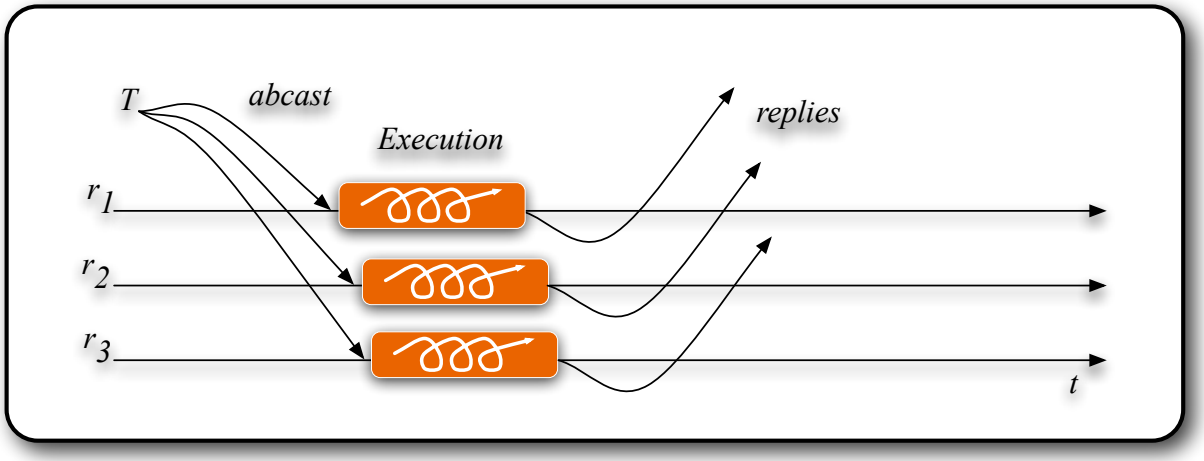


Database replication

Replicated state-machine



- Simplicity and failure transparency
- Requires deterministic processing (SQL, scheduling)



Database replication

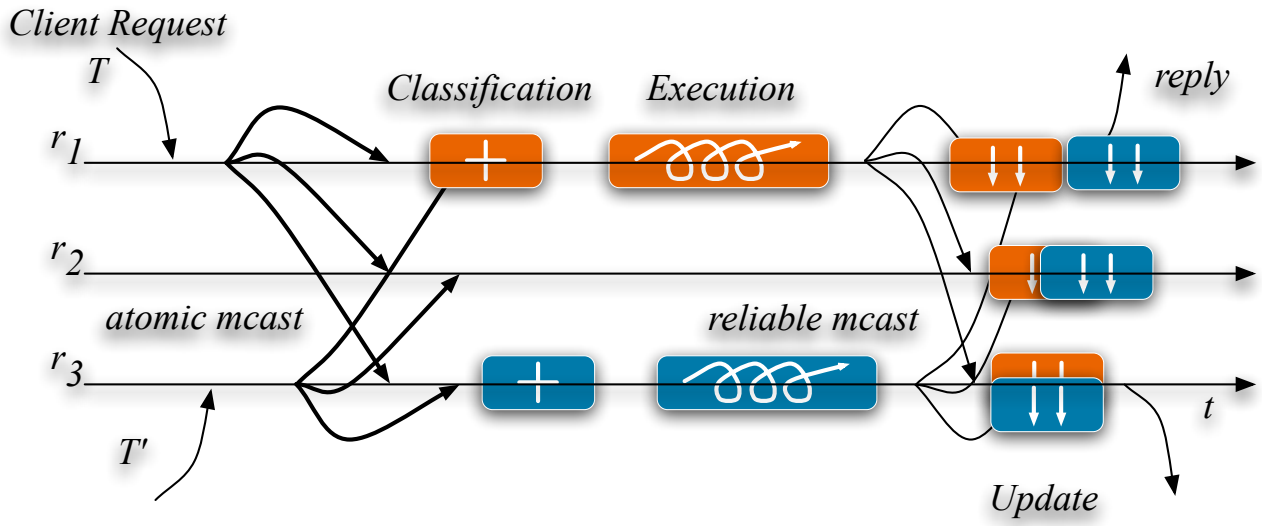
Specific Database Replication Protocols

- Build on the classical Replicated State-Machine
- Adopt the passive replication paradigm
- To globally serialise conflicting concurrent transactions share the use of an atomic multicast primitive
  - Message delivery is Atomic and Totally Ordered
- Ensure 1-copy-equivalence as consistency criterion

Database replication

Conservative execution

- Transactions are classified in conflict classes
- Transactions are ordered before their execution
- Conflicting transactions are sequentially executed

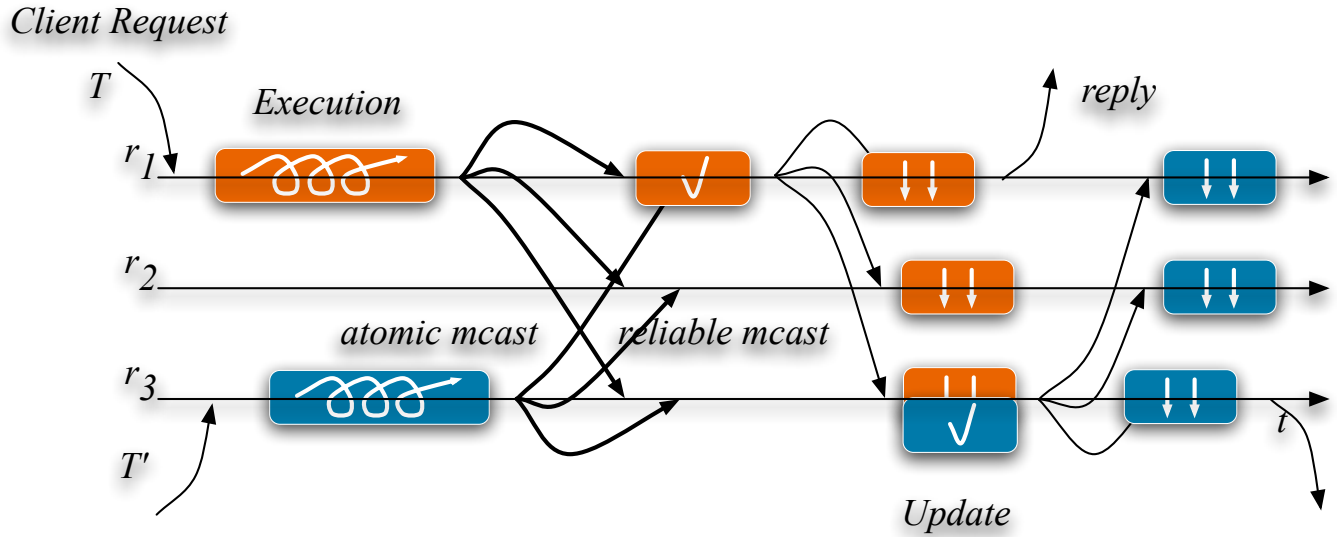


Database replication

Optimistic execution

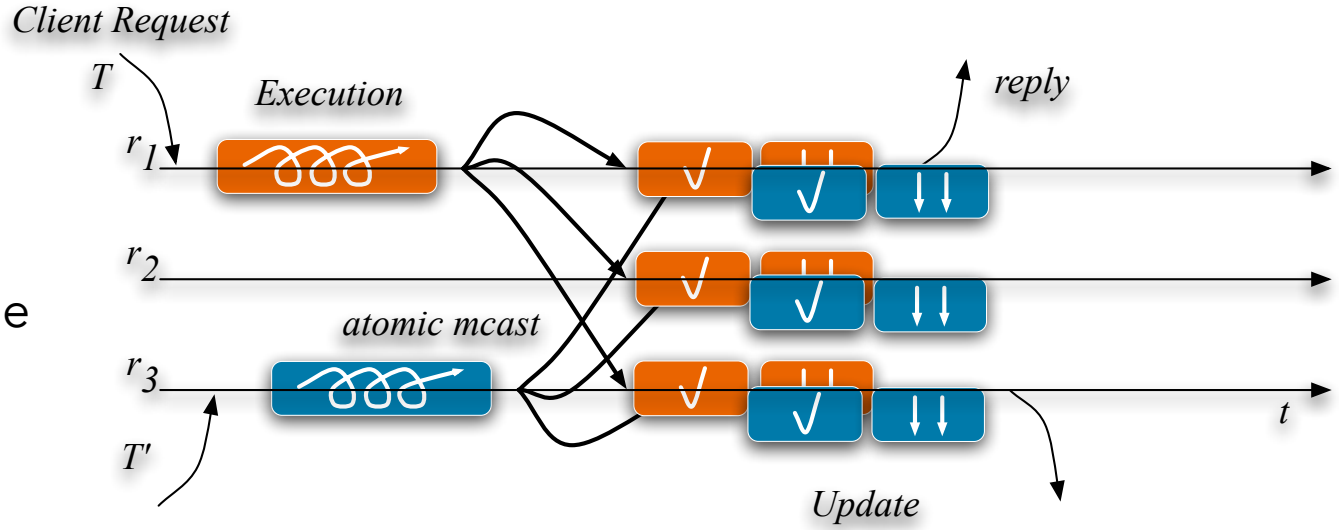
- Transactions are ordered after their execution
- Conflicting transactions may execute concurrently

Postgres-R



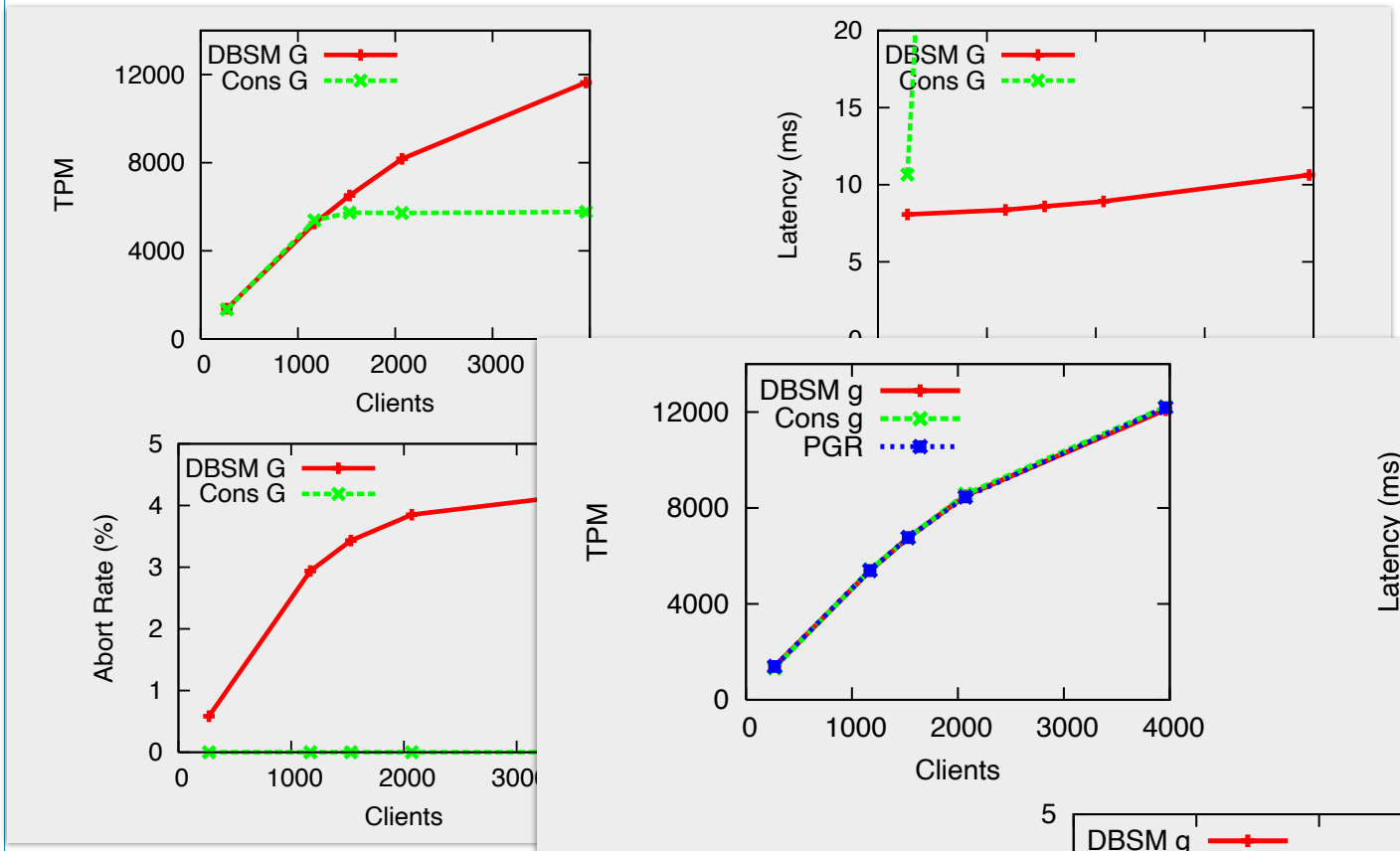
DBSM

DataBase State Machine

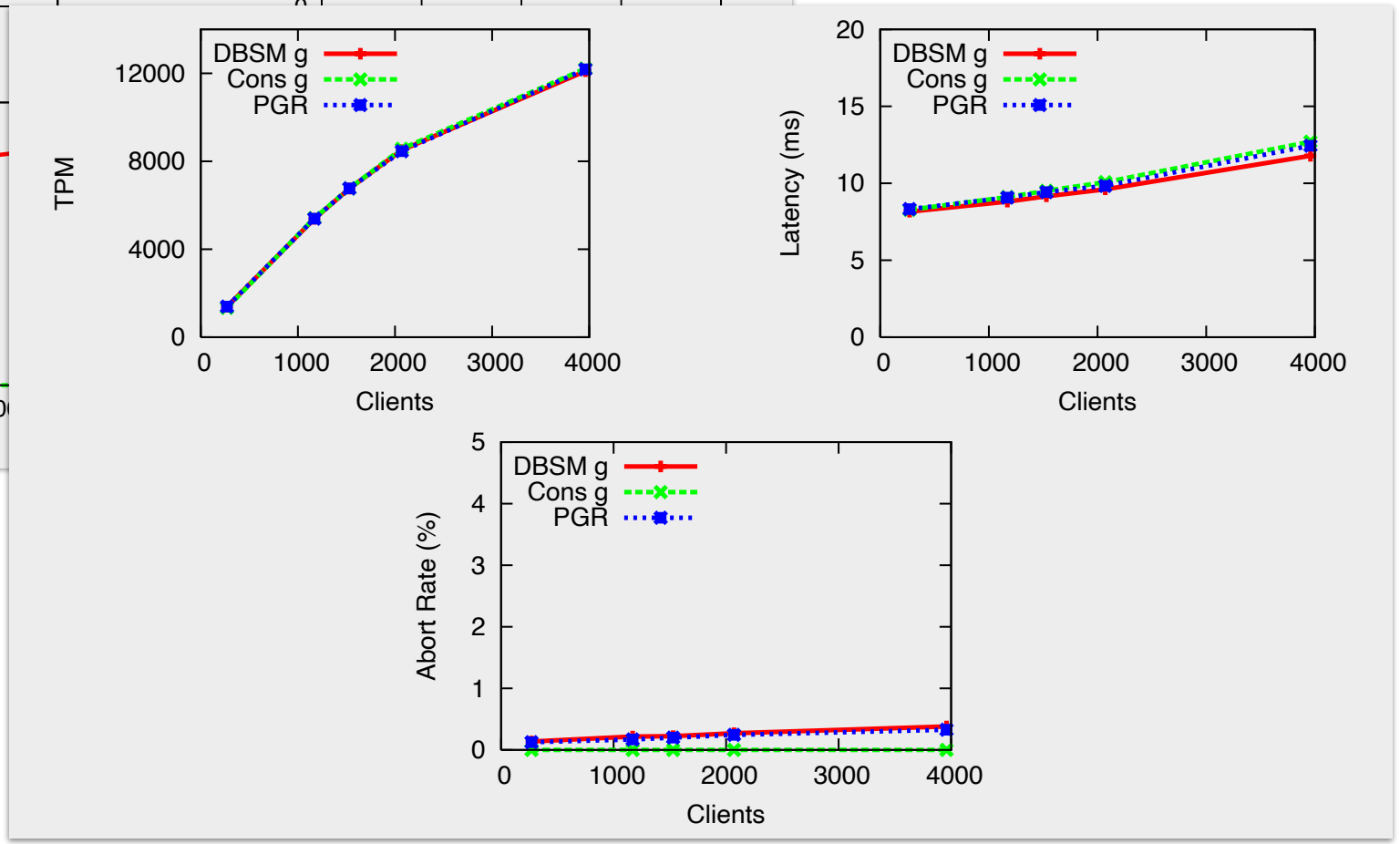


Database replication

Performance comparison



Coarse Grain CC



Fine Grain CC

Database replication

Reading material

- M. Wiesmann, F. Pedone, A. Schiper, B. Kemme, G. Alonso
  "Understanding Replication in Databases and Distributed Systems"
  INTERNATIONAL CONFERENCE ON DISTRIBUTED COMPUTING SYSTEMS (2000)

- F. Pedone, R. Guerraoui, A. Schiper
  "The Database State Machine Approach"
  DISTRIBUTED AND PARALLEL DATABASES (2003)

- Y. Lin, B. Kemme, M. Patiño-Martínez, R. Jiménez-Peris
  "Middleware based data replication providing snapshot isolation"
  ACM SIGMOD (2005)