Principles of Distributed Database Systems

M. Tamer Özsu
Patrick Valduriez

Outline

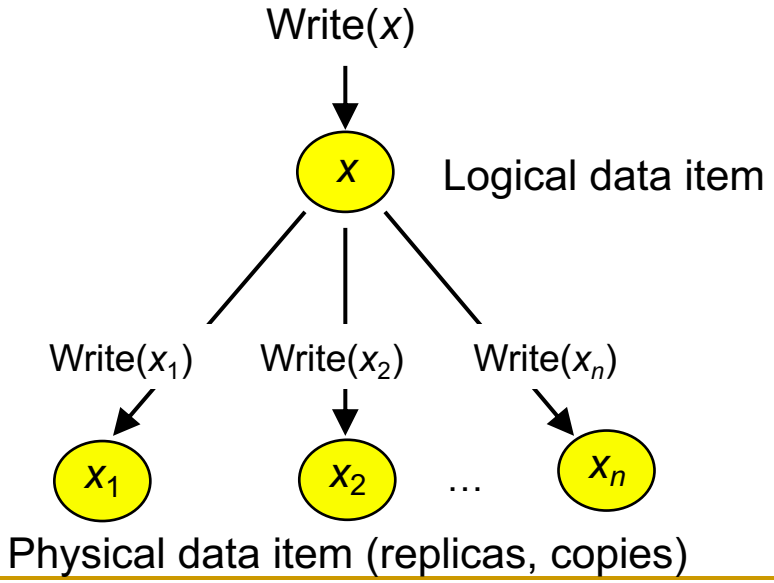
- Data Replication
  - Consistency criteria
  - Update Management Strategies
  - Replication Protocols
  - Replication and Failure Management

Replication

- Why replicate?
  - System availability
    - Avoid single points of failure
  - Performance
    - Localization
  - Scalability
    - Scalability in numbers and geographic area
  - Application requirements
- Why not replicate?
  - Replication transparency
  - Consistency issues
    - Updates are costly
    - Availability may suffer if not careful

Execution Model

- There are physical copies of logical objects in the system.
- Operations are specified on logical objects, but translated to operate on physical objects.
- One-copy equivalence
  - Transaction effects on replicated objects should be the same as if they had been performed on a single set of objects.



Replication Issues

- Consistency models - how do we reason about the consistency of the "global execution state"?
  - Mutual consistency
  - Transactional consistency
- Where are updates allowed?
  - Centralized
  - Distributed
- Update propagation techniques – how do we propagate updates to one copy to the other copies?
  - Eager
  - Lazy

Outline

- Data Replication
  - Consistency criteria
  - Update Management Strategies
  - Replication Protocols
  - Replication and Failure Management

Consistency

- Mutual Consistency
  - How do we keep the values of physical copies of a logical data item synchronized?
  - Strong consistency
    - All copies are updated within the context of the update transaction
    - When the update transaction completes, all copies have the same value
    - Typically achieved through 2PC
  - Weak consistency
    - Eventual consistency: the copies are not identical when update transaction completes, but they eventually converge to the same value
    - Many versions possible:
      - Time-bounds
      - Value-bounds
      - Drifts

Transactional Consistency

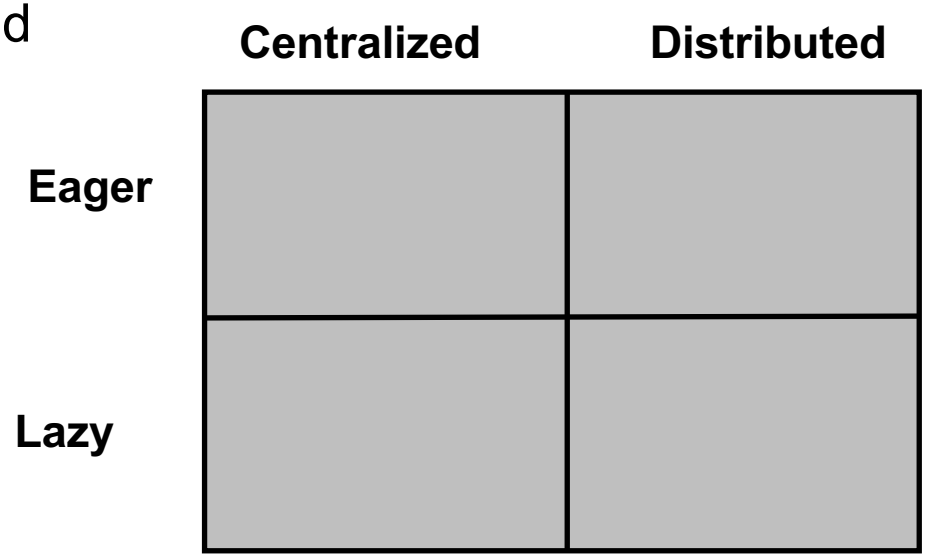
- How can we guarantee that the global execution history over replicated data is serializable?
- One-copy serializability (1SR)
  - The effect of transactions performed by clients on replicated objects should be the same as if they had been performed *one at-a-time* on a single set of objects.
- Weaker forms are possible
  - Snapshot isolation
  - RC-serializability

Outline

- Data Replication
  - Consistency criteria
  - Update Management Strategies
  - Replication Protocols
  - Replication and Failure Management

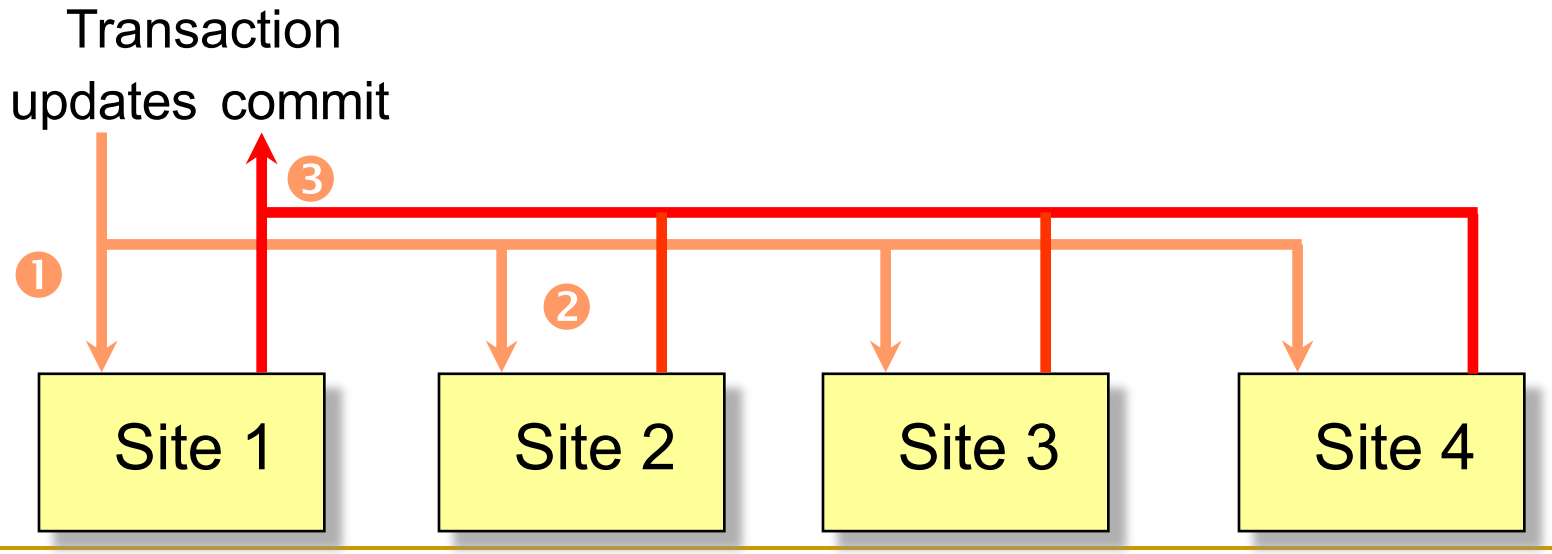
Update Management Strategies

- Depending on when the updates are propagated
  - Eager
  - Lazy
- Depending on where the updates can take place
  - Centralized
  - Distributed



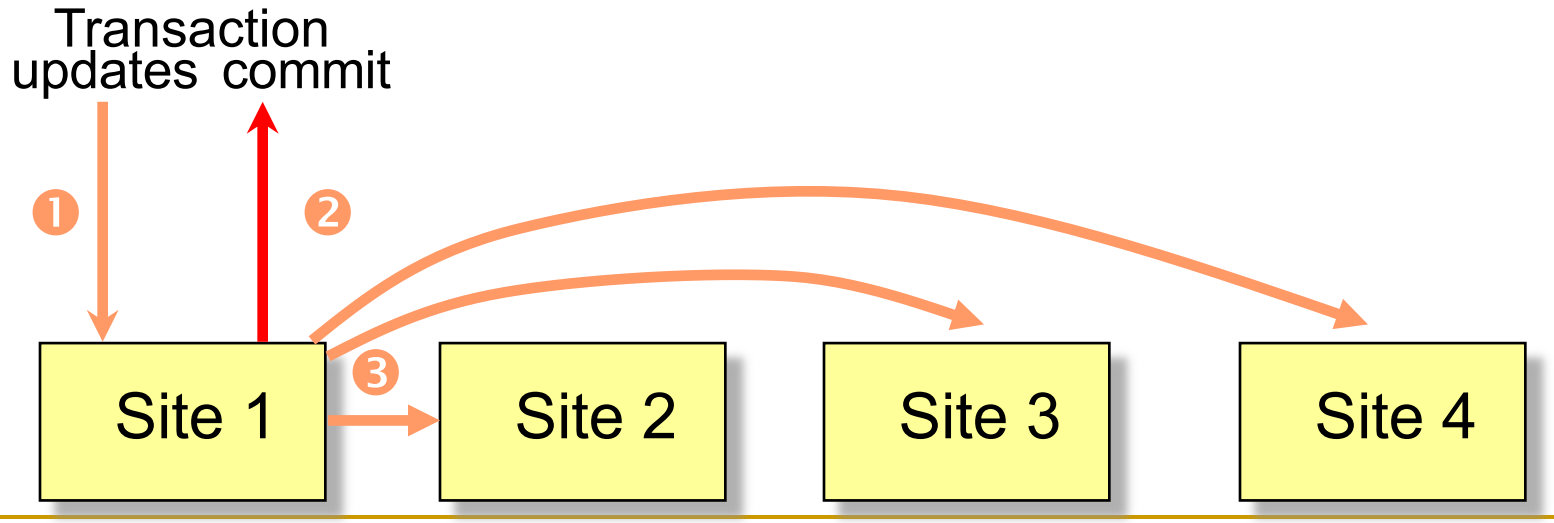
Eager Replication

- Changes are propagated within the scope of the transaction making the changes. The ACID properties apply to all copy updates.
  - Synchronous
  - Deferred
- ROWA protocol: Read-one/Write-all



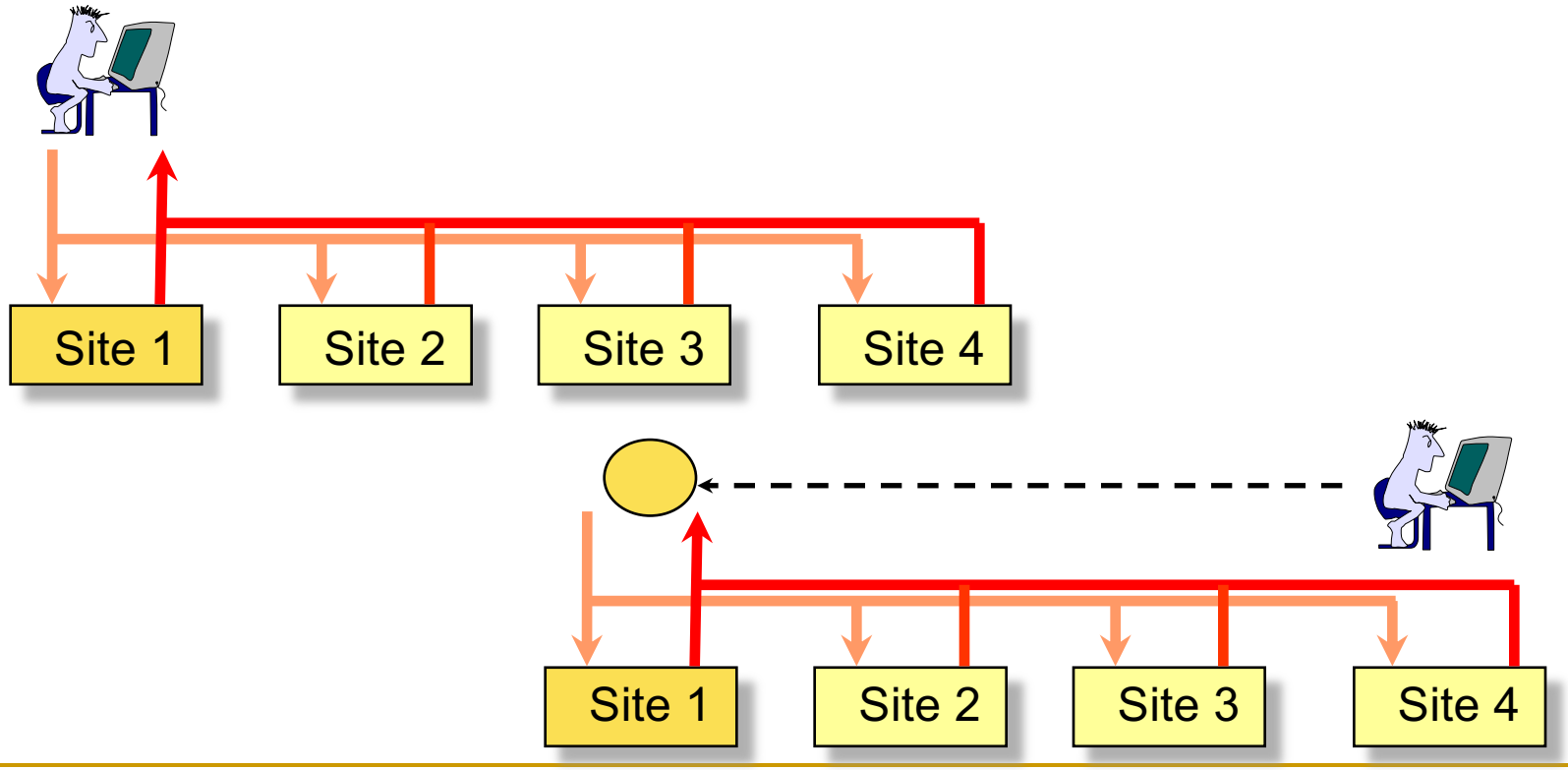
Lazy Replication

- Lazy replication first executes the updating transaction on one copy. After the transaction commits, the changes are propagated to all other copies (refresh transactions)
- While the propagation takes place, the copies are mutually inconsistent.
- The time the copies are mutually inconsistent is an adjustable parameter which is application dependent.



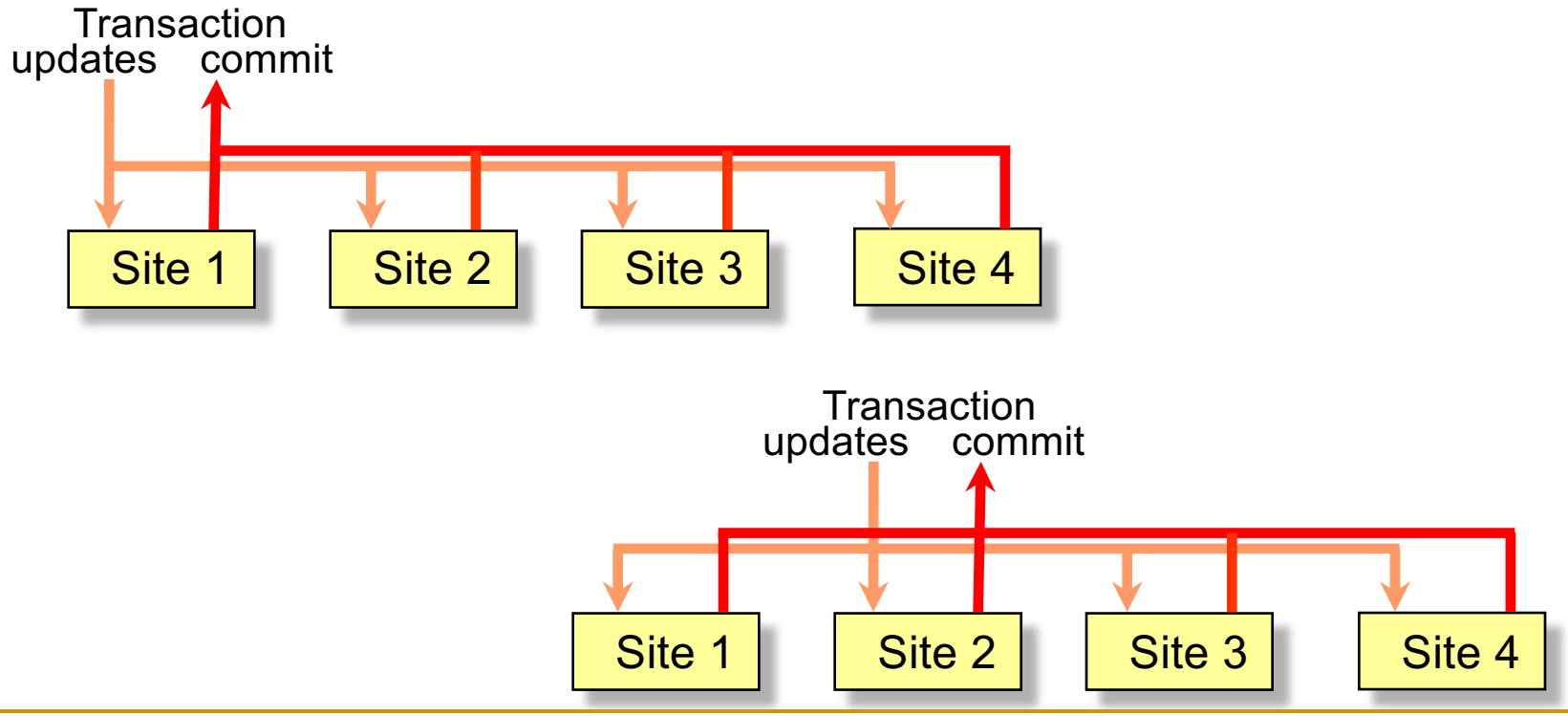
Centralized

- There is only one copy which can be updated (the master), all others (slave copies) are updated reflecting the changes to the master.



Distributed

- Changes can be initiated at any of the copies. That is, any of the sites which owns a copy can update the value of the data item.



Forms of Replication

Eager

+ No inconsistencies (identical copies)
+ Reading the local copy yields the most up to date value
+ Changes are atomic
− A transaction has to update all sites
  − Longer execution time
  − Lower availability

Lazy

+ A transaction is always local (good response time)
− Data inconsistencies
− A local read does not always return the most up-to-date value
− Changes to all copies are not guaranteed
− Replication is not transparent

Centralized

+ No inter-site synchronization is necessary (it takes place at the master)
+ There is always one site which has all the updates
− The load at the master can be high
− Reading the local copy may not yield the most up-to-date value

Distributed

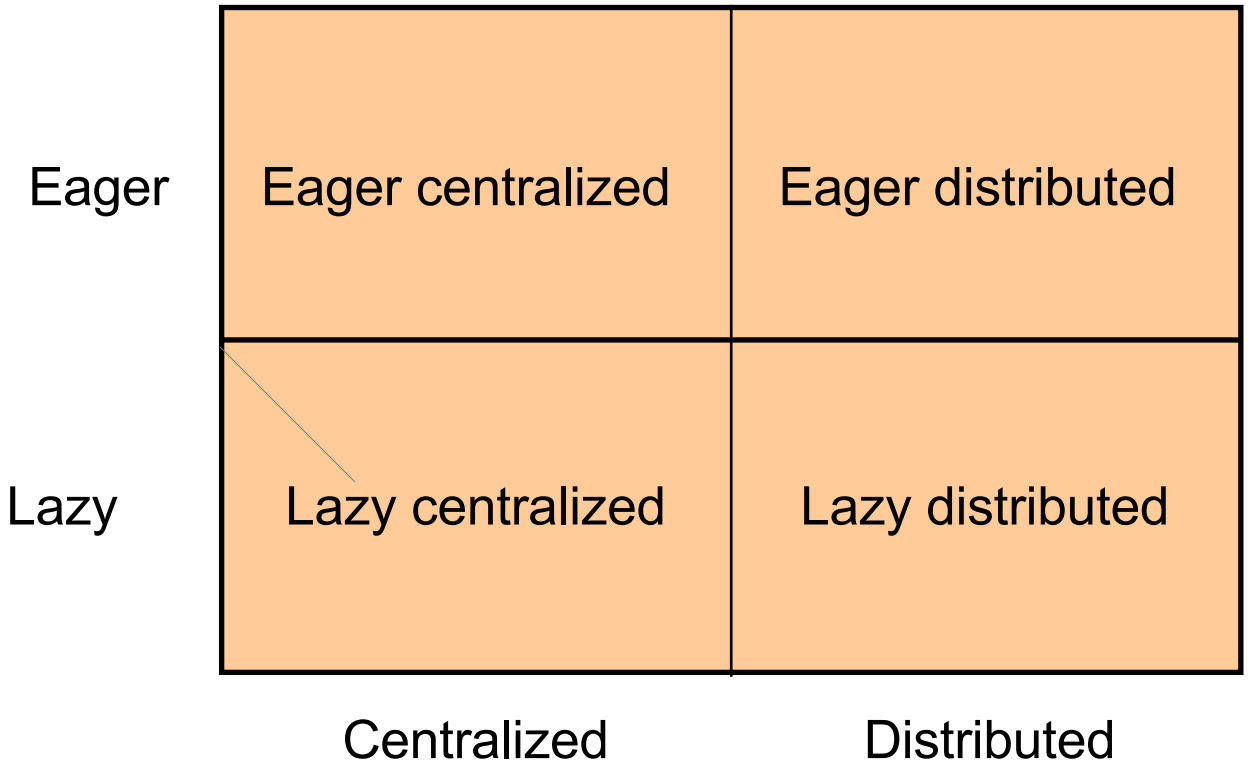
+ Any site can run a transaction
+ Load is evenly distributed
− Copies need to be synchronized

Outline

- Data Replication
  - Consistency criteria
  - Update Management Strategies
  - Replication Protocols
  - Replication and Failure Management

Replication Protocols

The previous ideas can be combined into 4 different replication protocols:



Outline

- Data Replication
  - Consistency criteria
  - Update Management Strategies
  - Replication Protocols
  - Replication and Failure Management

Failures

- So far we have considered replication protocols in the absence of failures
- How to keep replica consistency when failures occur
  - Site failures
    - Read One Write All Available (ROWAA)
  - Communication failures
    - Quorums
  - Network partitioning
    - Quorums