

Advanced Enterprise Computing - Lecturenotes SoSe2016

Julius Hülsmann

10. Mai 2016

Inhaltsverzeichnis

1	Replication and State Management	
	(25.04. - 09.05.)	2
1.1	Motivation and Background	2
1.1.1	Replication	2
1.2	Managing Replication	3
1.3	Implications of Replication	3
1.4	Paxos and CRDTs	3
2	Prototyping	3
3	Experiments	3
4	DevOps and Microservices	3
5	Reading Assignment	3
6	Lecturenotes	5
7	Begriffe und Abkürzungen	5

1 Replication and State Management (25.04. - 09.05.)

1.1 Motivation and Background

1.1.1 Replication

Definition - Replication Process of maintaining multiple Copies of an Entity (Data / Process / File ...)

Advantages of Replication in General

- *System Availability / Fault tolerance* in case
 - A Server fails
 - B Data is corrupted.
- *Performance / Scalability*
 - A Workloads are spread across distributed Replicas
 - B Geodistribution for processing demands in client's proximity

Disadvantages of Replication in General

- Consistency vs. Performance

Kinds of Replication In general there are the following kinds of „physical“ Replication. We do only consider (B).

!!!!!!!!!!!!!!ueberpruefen am Ende!!!!!!!!!!!!!!

Replication Strategies PAGE 17

Synchronous vs. Asynchronous

Synchronous / eager

Asynchronous / lazy

Primary Copy vs. Update Everywhere

Primary Copy / master

Update Everywhere / group

Name explanation Protokolltitle implementation Atomacity either perform Transaction entirely or roll-back atomic commitment protocol 2PC

ACID

Atomiticity:

Consistency: does not mean Data-Consistency but that the transaction produces consistent changes.

Isolation: Transactions are isolated from one another

Durability: Once the transaction is ready (commits) it remains.

Both the Atomiticity and the Isolation are managed by the **Transaction Manager** *It acquires locks on behalf of all transactions and tries to come up with a serializable execution, that is, make it look like the transactions were executed one after the other. If the transactions follow 2 Phase Locking, serializability is guaranteed. Thus, the scheduler only needs to enforce 2PL behavior.*

What happens to ACID in case of Replication? Atomicity can be guaranteed using 2PC (but expensive) Problem: Serialization order must be the same at all replicas.

Synchronous ACID properties apply to all copy updates

1.2 Managing Replication

1.3 Implications of Replication

1.4 Paxos and CRDTs

2 Prototyping

3 Experiments

4 DevOps and Microservices

5 Reading Assignment

Tabelle 1: My caption

	Procedure	Advantages / Disadvantages
Synchronous	1 propagate Data to everybody 2 Wait until everybody responded	- ACID (no Inconsistencies) - High response time (high execution time, response time) - Availability (in case one Copy fails)
Asynchronous	3 commit 1 Update local copy 2 commit 3 Propagate Data	- Response Time - Availability - Data inconsistency (local read does not always return the latest value) - No guarantee that the changes arrive at each copy - Replication is not guaranteed
Primary Copy	one Primary copy and several read-only copies	
Update Everywhere	Each site is able to initiate changes	

6 Lecturenotes

Lecture 05? start @ 81 Für Donnerstag paper mitbringen und Paxos anschauen.
2016-05-09

Paxos (Represent as State-machine) - P. 77

Proposer

Phase 1 - Proposer choses Number largr than any value chosen before by Proposer. - Broadcast the integer *prepare(n)*, e.g. prepare(50)

Acceptors a) Not respond at all b) *reject* Reject, in case a higher value has been accepted. 50 ; something b) *prommise(n)* in case 50 ; everything. Also Send everything that has already been accepted.

If prposer receives majority of prommise resposns, -; proceed to Phase 2 ELSE -; Phase 1

Phase 2 - Check whether any ;n, value; have been returned. - YES: take max n's value - accept (n, value)

Xtensions Paxos **Multi-paxos** Determine Leader once Stay in phase 2, attatch the leader identifier Leader is the one to accept values

Purpose: Optimize Speed (get rid of the first phase, Master-Slave setup)

Fast Paxos

Generalized Paxos - Assumption: The execution order does not matter.

CRDT Conflict free / Communitive replicated Datatypes

Some operations are commutative, others not.

State- Based vs. Operation based.

theoretically it is possible to converge them but ... practice

IDEA INTEGER - example: e.g. not store int values but operations (increment / decrement))

SET - example

State - based Set

7 Begriffe und Abkürzungen

Replication Strategy to maintain mutiple copies of an entity on multiple Servers.

Replica

CRDT *conflict-free replicated data*

Paxos

Commit In case a Transaction commits, it is ready.

Concurrency control protocol guarantees isolation of Transactions

2PL Two phase locking (one concurrency control protocol)

Snapshot Isolation other concurrency control protocol implementation

atomic commitment protocol guarantees atomicity

2PC Two phase Commit

Transaction Manager Middleware Component; Manages Atomicity and Isolation of Transactions

ACID Atomicity + Consistency + Isolation + Durability