



FAULT TOLERANCE IN LARGE-SCALE DISTRIBUTED SYSTEM

Advanced Data Management's Project

A.Y. 2022/2023, UNIGE

Edoardo Pastorino, 5169595



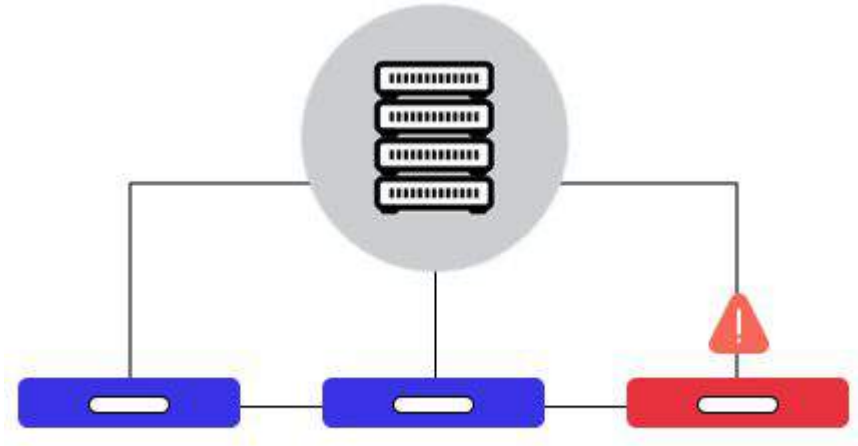


INTRODUCTION

WHAT IS FAULT TOLERANCE?

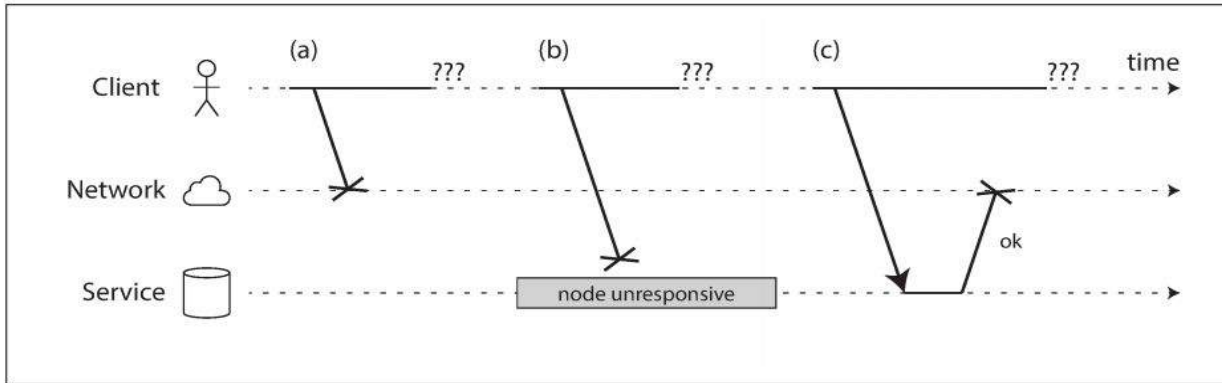
Fault Tolerance, refers to the capability of a certain system (computer system, cloud cluster, network...) to deal with system failures to continue operating even if one or more of its components fail.

A fault is the failure of a component of the system and can be associate, for instance, to the crash of a node, or to the happening of an error during execution of some operation.





MAIN FAULT TOLERANT ISSUES

Network fault, that occurs when a part of the considered network is cut off from the rest.



Byzantine failure, byzantine fault-tolerant system if it continues to operate even if some nodes are malicious.

Main fault tolerance's process' issues are failure detection and system recovery.



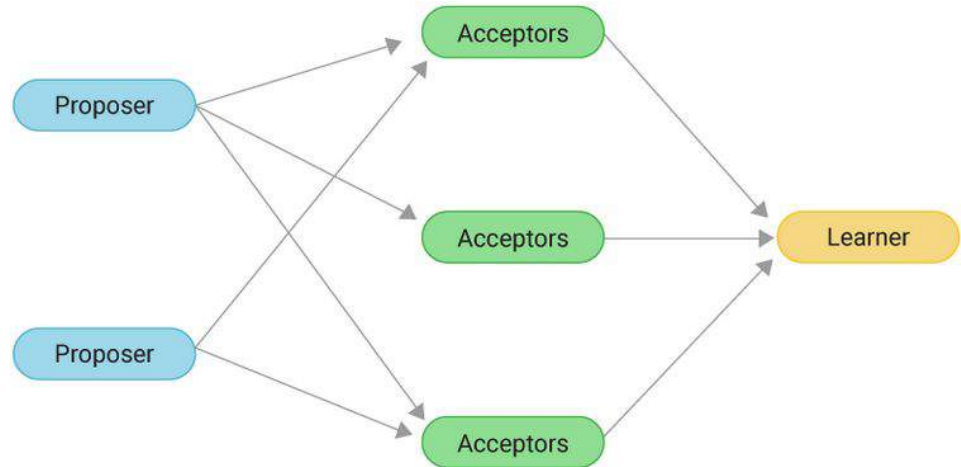
PAXOS CONSENSUS PROTOCOL

WHAT IS PAXOS?

Paxos is a protocols for solving consensus in a network of unreliable or fallible processors.

The main goal is try to ensure that a single value, among all the values proposed by a collection of processes, is chosen.

It is divided in the selecting phase, in which a value is selected among proposal, and the learning phase, in which we inform all the other processes about the selected value.



PAXOS' FAULT TOLERANCE

The system requires $2m+1$ servers to tolerate the failure of m servers.

The Paxos algorithm can be used in practice, in a fault tolerant system, to elect a leader process when the previous leader fails.

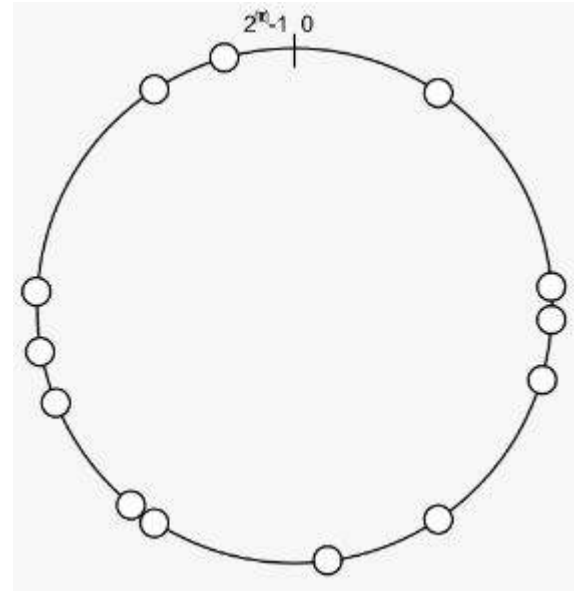


CHORD ALGORITHM

WHAT IS CHORD?

The Chord protocol is a lookup protocol that try to support the efficient location of data items inside the nodes in a peer to peer distributed system, based on consistent hashing for assigning the data items into a ring space address.

It allows to obtain good properties, such as high scalability, availability and a proper load balance.



CHORD'S FAULT TOLERANCE

To increase robustness against faults, each node n in the ring space address maintains a list of r successors. In this way even if some nodes that are successors of n fail, n is able to direct a right lookup result. Only when all the successors node in the r list of n fail we have problems, but this is very less probable.

This algorithm is also able to handle a situation in which a node n leaves the Chord ring voluntarily, as if it were a failure of that node.



PAXOS VS. CHORD

DIFFERENCE BETWEEN PAXOS AND CHORD

Aspect	Paxos	Chord
Approach	Message passing	Consistent hashing
Reference Architecture	Master-slave	Peer to Peer
Complexity	N	$\log(N)$
Fault Tolerance Guarantee	More than 50% of node can crash	when 50% of nodes fail, only 1.2% of lookups fail.



FAULT TOLERANCE IN NEO4J

- Raft protocol, divided in leader election phase, when a leader fails a new one is elected, and log replication phase, the leader can accept the client requests and propagates them to the follower nodes.
- Multi-Paxos protocol, a way for executing different instance of basic Paxos. Reducing the failure-free message delay and the overhead in communication





FAULT TOLERANCE IN MONGODB

Uses a raft algorithm with some differences:

- A pull based approach instead of a push based approach.
- Introduces an optimization that applies modification when it is added and after commits the log.
- The user can select a personalized timeout.
- We can also assign priorities to the nodes.
- MongoDB doesn't use snapshots approach when a new server is added to the system.
- In the catching phase, the leader cannot accept immediately new write requests for not losing data.
- More roles for a node (arbiter and non-voting member).

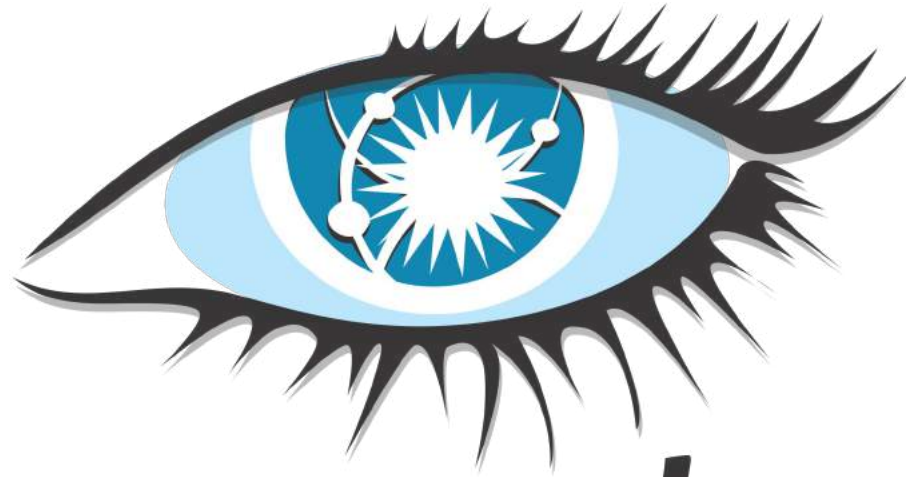




FAULT TOLERANCE IN CASSANDRA

Cassandra some of the best practice:

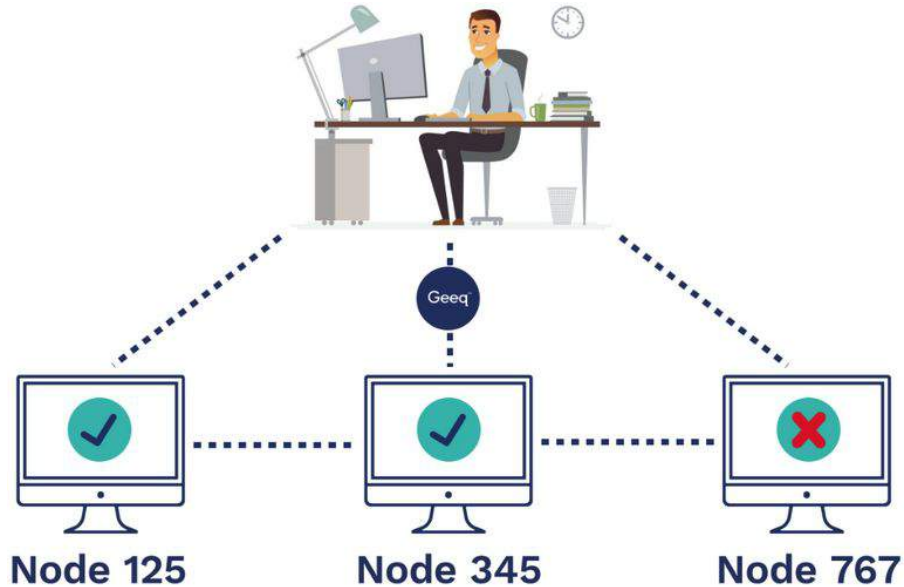
- Replication strategies
(NetworkTopologyStrategy,
SimpleStrategy)
- Probabilistic Quorum System and
Consistency Level.
- Gossip Protocol.
- Failure Detection (Phi Accrual).



cassandra



CONCLUSIONS



Fault tolerance is a needed features in a large scale distributed system, for dealing with data loss, network fault, leader crash and node crash in general, with procedure for finding the failure and with method for managing these failures, necessary for the data survival.



THE END

THANKS FOR YOUR ATTENTION