Implementation of Viewstamped Replication Protocol

Indian Institute of Technology Guwahati

November 4, 2024

Overview

Problem Statement

- Build a highly available, replicated DNS server that ensures consistency of updates
- Our emphasis is on the availability of the DNS server. Thus, replication protocols are more suitable and consensus protocols.

What is Viewstamped Replication?

- A method for achieving fault tolerance in distributed systems.
- Utilizes multiple replicas of a data store to ensure high availability.
- Operates under the assumption that some replicas may fail or become unreachable.

Key Concepts

- **Primary Replica:** Handles all read and write operations and coordinates with other replicas.
- **Backup Replicas:** Store copies of the data and respond to write modifications made.
- **Client Requests:** Clients interact primarily with the primary replica.

Normal Operations Overview

- Consists of read and write operations executed under normal conditions.
- Primary replica processes write requests and disseminates updates to backups.
- Primary replica processes read requests and sends the response to the client. No need to inform or consult the backup replicas
- Backup replicas respond to the prepare requests sent by the Primary replica for a write request.

Read Operation Flow

Process

- Client sends a read request to the primary.
- Primary responds directly because it has the latest data.

Write Operation Flow

Process

- Client sends a write request to the primary.
- Primary updates its data store and logs the operation.
- Primary notifies all backup replicas to update their stores.
- Acknowledgment sent back to the client once at least f replicas are updated where f is the most number of failures.

Handling Failures

- If the primary fails, a backup must take over.
- This backup is the replica that has the lowest ID among the IDs that are greater than the previous primary replica and is active.
- Client operations must be rerouted to the new primary.

View Change Overview

- Essential for maintaining consistency when a primary replica fails.
- Involves a change in the primary replica and potential new views.
- Requires communication among replicas to agree on the new primary.

View Change Process

Steps

- Detection of primary failure.
- Backup replicas communicate their logs and states.
- Selection of the most up-to-date logs at the new primary.
- The new primary informs all replicas of the new view.

Architecture Overview

- The system architecture includes:
  - Client nodes
  - Primary replica
  - Backup replicas
- Communication through gRPC (Remote Procedure Calls).

Technologies Used

- Programming Language: Go, Protobuf
- Communication Protocol: gRPC
- The code is only a simulation. It creates the clients and replicas. The communication between them takes place with the help of the declared gRPCs.

Testing the Protocol

- No failures
- Failure of the primary
- Failure of 2 replicas (1 primary, 1 backup)
- Failure of 2 replicas (1 primary, 1 to-be-primary)

Learnings & Takeways

- Writing gRPCs and using them as a way to communicate
- Utilizing Go routines
- Learning a new algorithm: View Stamped Replication

The End