Paper Review Title: Protocol-Aware Recovery for Consensus-Based Storage
Reviewer: swayanshu shanti Pragnya

1. Summarize the (at most) 3 key main ideas.

1. In this paper, they have introduced a protocol aware recovery(PAR) which will exploit protocol specified knowledge from distributed systems in a correct manner.

2. They demonstrated the efficiency and accuracy of PAR by implementing Corruption tolerant replication.

3. They have created another system named as Replicated state machine which will replicate PAR mechanism.

4. They tried to show the comparison and the efficient recovery method between 2 CTRL versions named as LogCabin and ZooKeeper.

2. State the main contribution of the paper

The primary contributions are-

1. **Creating PAR which will extract protocol specified knowledge in a correct way from any distributed system.**
2. **While comparing Zookeeper they have also proposed Leader initiated snapshotting, global commitment determination and crash corruption disentanglement.**
3. **Storage fault handling by using different techniques like mark non-voting, truncate and reconfigure.**
4. **Log and Snapshot recovery where they recovered the notes from chunk level in a faulty data.**
5. **They have implemented a paradigm named as RSM which will run on many servers with same initial state, same sequence of inputs to produce same inputs.**

3. Critique the main contribution

3.a. Rate the significance of the paper on a scale of 5 (breakthrough), 4 (significant contribution), 3 (modest contribution), 2 (incremental contribution), 1 (no contribution or negative contribution). Explain your rating in a sentence or two.

It is a significant contribution so I would rate as 5 because the methodology includes the following-

1. **As the paper talks about the most crucial topic of recovery and implementing a complete new approach must be very challenging but throughout the paper, they explained tit by bit.**
2. **Recovery approaches like protocol oblivious and protocol aware the low complexity, performance, no need of extra node and fast computation factors showed the model competence which they achieved.**

3. **Loss of any kind of data or file or log is very petrifying so the approaches create an overall impact of consolation.**

1. ***b. Rate how convincing is the methodology: do the claims and conclusions follow from the experiments? Are the assumptions realistic? Are the experiments well designed? Are there different experiments that would be more convincing? Are there other alternatives the authors should have considered? (And, of course, is the paper free of methodological errors.)***

All the methodologies are convincing which includes the following majors,

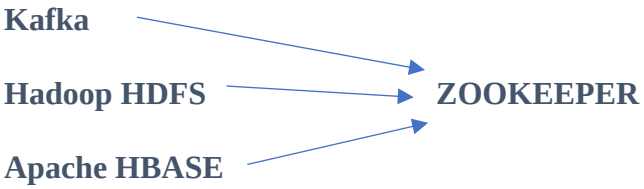
1. Redundancy enabled reliability which helps distributed storage systems from system crash or network failure.

2. For *recovering faulty data* they fixed it by using *Zookeeper* which clear all the states in Zookeeper and restart in the kernel.

3. For ignoring *global data loss protocol oblivious* is another method which is again based on the concept of delete and rebuild approach with *replicated data*.

4. Protocol aware recovery which will focus on replicated state machines

Example-



5. For stopping violation on safety they have introduced Corruption Tolerant RepLication(CTRL) which exploits properties of RMS protocol to avoid inconvenience.

3. c. What is the most important limitation of the approach?

Though it's a brilliant paper but still have few limitations.

1.Recovering from storage faults correctly from a distributed system can be time taking.

2. Most of their recovery approach includes *protocol-Oblivious* which may cause unsafety and low availability.

3. For correctness and quick recovery the protocol does not answer "*protocol- aware method*"

4. Rate the writing in the paper on a scale of 5 (great) to 1 (muddled), and justify your ranking. Did you have to re-read sections? Were algorithms clearly explained? Did the paper have a logical flow?

Writing-5

The paper had a structural flow and proper explanation of methodology like starting from the paper objective, recovery methodology, limitations, future work for improvising the performance and correctness are well explained.

As the authors clearly explained the design, implementation and method behind each protocol, it's easy to understand the logic.

The usage of Protocols like: ***RSM, CTRL, RAID, LogCabin, Zookeeper are well explained.***

5. Answer one of the following three questions (whichever is most relevant for this paper):

1. What lessons should system researchers and builders take away from this work?2. What is the lasting impact of this work? 3. What (if any) questions does this work leave open?

By answering the question 1 these are the following points which researchers can work further,

1. Adding protocol-aware to recover quickly

2. Improvising redundant copies which can improvise the bad data recovery.

3. Upgrading PAR to deal with quoram based vulnerable systems.

4. Focusing on reliability for distributed systems which will be robust for storage faults.