Literature Review on Consistency Models and Conflict-Free Replicated Data Types in Distributed Systems

A Structured Study of Consistency Models and Conflict-Free Replication

Muhammad Minhal Manjee

Electrical Engineering and Computer Science

University of Wyoming

Laramie, WY, USA

mmanjee@uwyo.edu

1 Topic

This project conducts a literature review on the evolution, classification, and implementation of consistency models in distributed systems, focusing on both theoretical foundations and practical frameworks. Specifically, the study explores how the meaning of consistency has evolved from strong consistency and linearizability in early distributed systems to weaker, eventually consistent semantics that dominate modern, large-scale, fault-tolerant architectures. The review also examines Conflict-Free Replicated Data Types (CRDTs) as an approach for achieving strong eventual consistency in distributed environments, and the role of these models in non-transactional distributed storage systems.

2 Motivation

As distributed systems grow in complexity, particularly driven by large scale storage systems in cloud systems and global-scale databases, maintaining consistency among multiple replicas becomes a major challenge. Early systems favored strong consistency, but this came at the cost of availability and scalability, as described by the CAP theorem. Modern distributed systems often relax consistency guarantees to improve performance, giving rise to a rich spectrum of consistency models. Understanding these models is essential for designing efficient systems that balance correctness, performance, and fault tolerance. By systematically reviewing key literature, this project aims to clarify formal definitions and taxonomies of consistency models, explore CRDTs as a coordination-free approach for maintaining convergence and identify the major trade-offs and research challenges in non-transactional distributed storage systems.

3 Related Work and Resources

3.1 Primary Literature

- Viotti, P., & Vukolić, M. (2016). *Consistency in Non-Transactional Distributed Storage Systems.* ACM Computing Surveys, 49(1).
- Zalewski, M., et al. (2019). *Consistency Models in Distributed Systems: A Survey on Definitions, Disciplines, Challenges, and Applications.*
- Shapiro, M., et al. (2011). *Conflict-Free Replicated Data Types (CRDTs).* INRIA Research Report.

3.2 Supplementary References

- Lamport, L. (1978). *Time, Clocks, and the Ordering of Events in a Distributed System.* Communications of the ACM.
- Vogels, W. (2009). *Eventually Consistent.* Communications of the ACM.

3.3 Resources and Tools

- **Computing Resources:** Local workstation and optional ARCC cluster access.
- **Software Tools:** Zotero (citation management), Overleaf (LaTeX editing), GitHub (version control) and Draw.io (visualization).

4 Timeline and Deliverables

Week 1: Literature Setup & Initial Review

- Read all selected papers thoroughly to identify the overall research scope and key questions.
- **Deliverables:** Summary notes for each paper.

Week 2: In-Depth Reading (Consistency Models)

- Analyze the definitions, categorizations, and challenges of various consistency models. Focus on understanding the relationships among strong, causal, and eventual consistency.
- **Deliverables:** Summary table of consistency definitions and their hierarchy. Identification of common patterns, strengths, and limitations in existing models.

Week 3: In-Depth Reading (CRDTs)

- Study Conflict-Free Replicated Data Types (CRDTs), their mechanisms, convergence guarantees, and distinctions between state-based and operation-based models.
- **Deliverables:** Summary diagram of CRDT behavior. Tabulated overview of CRDT advantages, weaknesses, and applicability in distributed environments.

Week 4: Comparative and Thematic Analysis

- Synthesize insights across all three papers. Compare taxonomy, performance trade-offs, and practical implications. Identify overlaps and differences among consistency frameworks and replication strategies.
- **Deliverables:**
  - Comparative analysis matrix (models × features).
  - Summary table of distributed model trends, weaknesses, and representative works.
  - Preliminary gap analysis and future research challenges.

Week 5: Draft Literature Review Report

- Write draft sections: Introduction, Background, Consistency Models, CRDTs, Comparative Analysis, Discussion.
- **Deliverables:** First complete report draft with comparative and gap analysis.

Week 6: Revision, Visualization, and Synthesis

- Refine structure, add visuals, ensure accurate references.
- Integrate analysis of key challenges, research gaps, and future directions.
- **Deliverables:** Polished report draft with finalized analysis and visualizations.

Week 7: Finalization and Submission

- Refine report structure, incorporate figures/tables, and finalize sections on key challenges, research gaps, and future directions. Strengthen analytical insights and ensure citation accuracy.
- **Deliverables:** Final report and presentation slides.

5 Conclusion

This structured literature review plan provides a systematic framework for understanding the evolution of consistency models and replication mechanism in distributed systems. By examining CRDTs and a wide range of consistency semantics, the study aims to identify principles, limitations, and open research directions that will inform the design of future scalable and fault-tolerant data systems.