

# Paper 1: Li et al.

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

**Title** Making Geo-Replicated Systems Fast as  
Possible, Consistent when necessary  
*10th USENIX Symposium on Operating Systems Design and  
Implementation*

**Authors**

**Date**

# Motivation:

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- 1 To improve user-experience, services replicate system state across geographical diverse sites.
- 2 Performance vs Consistency
  - Amazon's Dynamo - eventual consistency where state temporarily diverge.
  - Yahoo PNUTS - avoids state divergence by requiring all operations that update the service state to be funneled through a primary site and thus incurring increased latency.

# Overview:

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- 1 RedBlue Consistency - Blue operations execute locally and are lazily replicated. Red operations are serialized with respect to each other and are immediately cross-site coordinated.
- 2 Conditions under which operations must be colored red or blue.
- 3 Decomposing operations into two components - a generator operation and a shadow operation.

# Properties of Geo-Replicated Systems

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- 1 Low latency - Operations should proceed after contacting a small number of users.
- 2 Causality - Monotonicity of user request within session and also preserving causality across clients
- 3 State Convergence - All replicas have executed the same set of operations
- 4 All operations should return a single value.
- 5 The system should provide a set of stable histories and support for general operations.
- 6 The system should preserve a set of invariants.
- 7 Eventual Propagation

# Related Work: Target end-to-end Properties

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

Consistency level	Example systems	Immediate response	State convergence	Single value	General operations	Stable histories	Classification strategy
Strong	RSM [20, 31]	no	yes	yes	yes	yes	N/A
Timeline/snapshot	PNUTS [8], Megastore [3]	reads only	yes	yes	yes	yes	N/A
Fork	SUNDR [24]	all ops	no	yes	yes	yes	N/A
Eventual	Bayou [38], Depot [26]	all ops	yes	no	yes	yes	N/A
	Spore [12], CRDT [33]	all ops	yes	yes	no	yes	N/A
	Zeno [34], COPS [25]	weak/all ops	yes	yes	yes	no	no / N/A
Multi	PSI [35]	cset	yes	yes	partial	yes	no
	lazy repl. [19], Horus [39]	immed./causal ops	yes	yes	yes	yes	no
RedBlue	Gemini	Blue ops	yes	yes	yes	yes	yes

Table 1: Tradeoffs in geo-replicated systems and various consistency levels.

# Related Work: Strong vs Weak Consistency

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# Related Work: Other

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# System Model

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1



# RedBlue Consistency

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# RedBlue Consistency - Definition

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# State Convergence and RedBlue Bank

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# Replicating side effects - Defining shadow operations

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

# Revisiting RedBlue consistency

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

1

## Paper 2: Lloyd et al.

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

### **Stronger Semantics for Low-Latency Geo-Replicated Storage**

*Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI13)*

Wyatt Lloyd, Michael J. Freedman, Michael Kaminsky, and  
David G. Andersen

April 2013

# Main Idea

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Main Idea

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Take slight hit in throughput to get stronger version of consistency



# Main Idea

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Take slight hit in throughput to get stronger version of consistency
- Causal Consistency Instead of Eventual Consistency (causal is stronger)

# Main Idea

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Take slight hit in throughput to get stronger version of consistency
- Causal Consistency Instead of Eventual Consistency (causal is stronger)
- We require low latency

# Main Idea

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Take slight hit in throughput to get stronger version of consistency
- Causal Consistency Instead of Eventual Consistency (causal is stronger)
- We require low latency
- Extend previous systems: Cassandra and COPS

# Contributions

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Contributions

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Eiger
  - Low Latency
  - High throughput (slightly lower than Cassandra)
  - Causal Consistency (rather than eventual as in Cassandra)

# Contributions

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Eiger
  - Low Latency
  - High throughput (slightly lower than Cassandra)
  - Causal Consistency (rather than eventual as in Cassandra)
- Read Only Algorithm

# Contributions

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Eiger
  - Low Latency
  - High throughput (slightly lower than Cassandra)
  - Causal Consistency (rather than eventual as in Cassandra)
- Read Only Algorithm
- Write Only Algorithm

# Background

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography



# Background

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Cassandra
  - Eventual Consistency

# Background

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Cassandra
  - Eventual Consistency
- COPS

# Consistency - Causal versus Eventual

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Consistency - Causal versus Eventual

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

# Column Family Data Model

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Column Family Data Model

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

# Column Family Data Model

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- p1
- p2

# Column Family Data Model

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- p1
- p2
- p3



# Eiger

## Consistency in the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Eiger

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

# Eiger

## Consistency in the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

■ p2

# Eiger

## Consistency in the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- p1
- p2
- p3

# Evaluation

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Evaluation

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Versus Cassandra
  - Within 7% of throughput Using Facebook-like data
    - Ops/sec
    - Keys/sec
    - Columns/sec

# Evaluation

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- Versus Cassandra
  - Within 7% of throughput Using Facebook-like data
    - Ops/sec
    - Keys/sec
    - Columns/sec
- Versus COPS

# Follow Up Research

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography



# Follow Up Research

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

# Ideas for Future Research

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

# Ideas for Future Research

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

■ p1

# Bibliography

Consistency in  
the Cloud II

Satabdi  
Aditya and  
Shannon  
Harwick

Li et al.

Lloyd et al.

Bibliography

- **Stronger Semantics for Low-Latency Geo-Replicated Storage**, *Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI13)*, Wyatt Lloyd, Michael J. Freedman, Michael Kaminsky, and David G. Andersen, April 2013