

BAR Fault Tolerance for Cooperative Services

Aiyer et al.
SOSP '05

Presented by John Otto

2008 Winter – EECS 345 – Distributed Systems
Northwestern University

Target and Motivation

- Cooperative services where authority admits nodes (provides bound on Byzantine nodes)
- Models exist for Byzantine or Rational nodes, but not together

Model

- Protocols provide guarantees under two classes
 - Incentive-Compatible Byzantine Fault Tolerant
 - Byzantine Altruistic Rational Tolerant
- For co-worker, dormitory, or non-profit group backup, with authority to distribute keys and admit nodes

Architecture

Architecture	Prototype		
Level 3: Application	BAR-B Backup		
Level 2: Work Assignment	Guaranteed Response	Periodic Work	Authoritative Time
Level 1: Primitives	Replicated State Machine		
	Message Queue		

Figure 1: System architecture

State Machine

- TRB
- Message queue, bubbles
- Balanced Messages
- Penance
- Timeouts and Garbage Collection
- Global Punishment

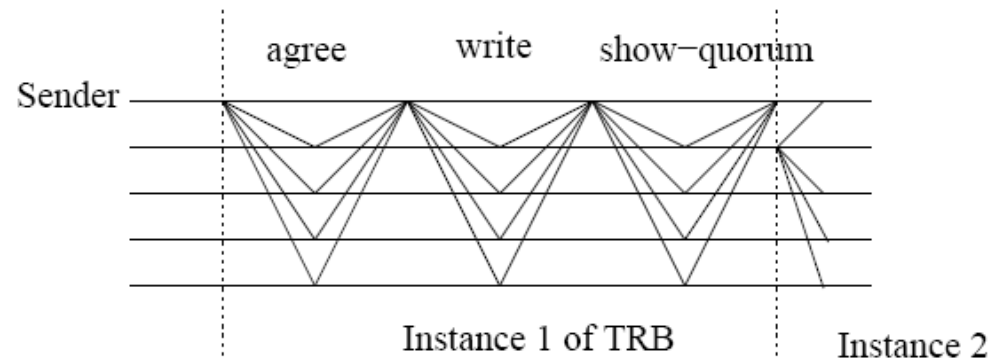


Figure 2: Terminating Reliable Broadcast (TRB) phases.

Partitioning Work

- Guaranteed Response
- Witness Node
- State Limiting

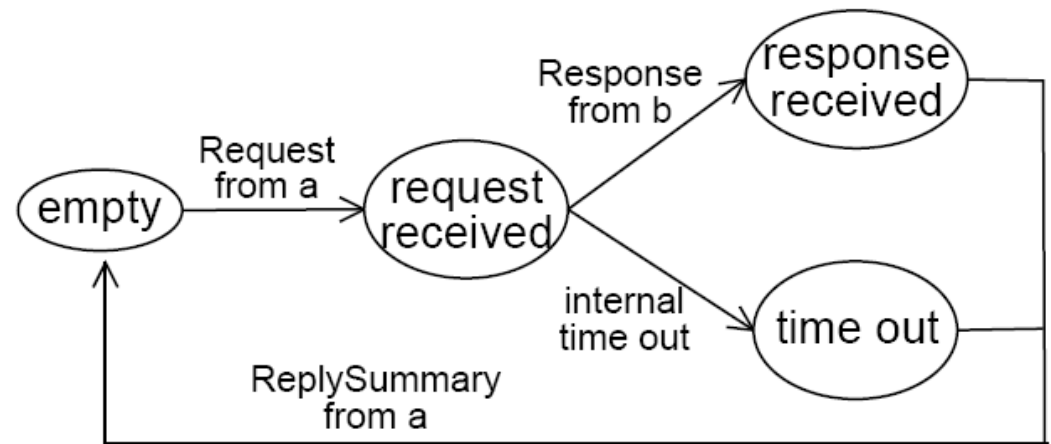


Figure 3: Basic Guaranteed Response protocol

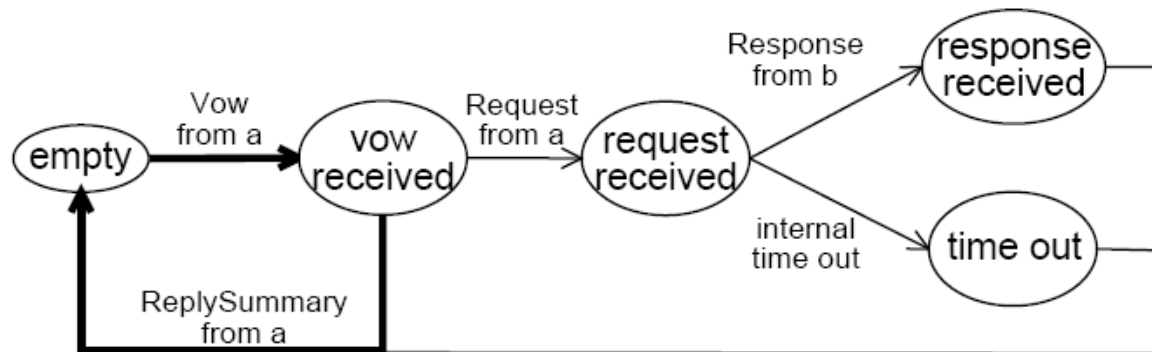


Figure 4: Guaranteed Response protocol with fast path

Backup Application

- Arithmetic Coding
- Commands
 - Store
 - Retrieve
 - Audit
- Sanctions
- Recovery Mode

Evaluation

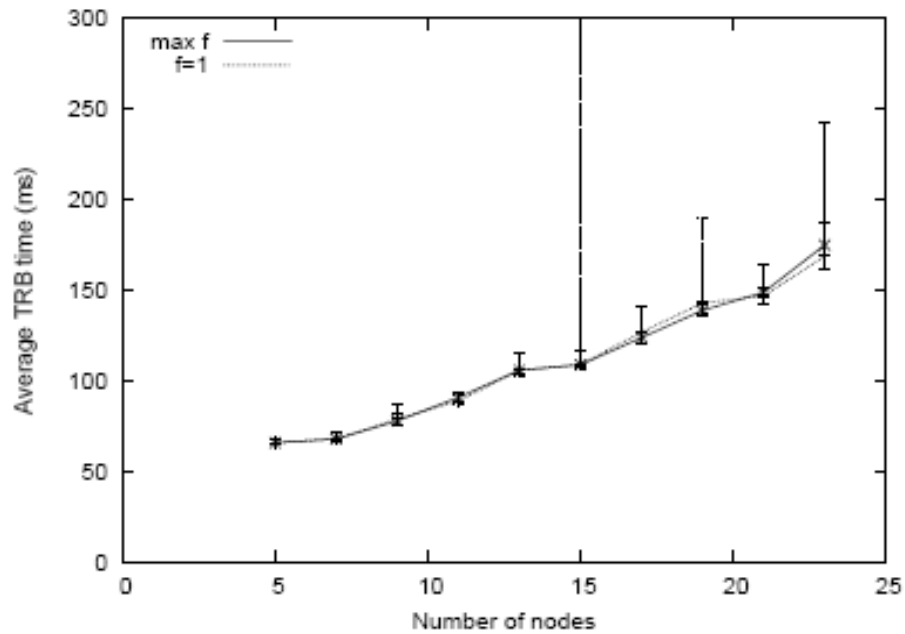


Figure 5: RSM performance as nodes are added

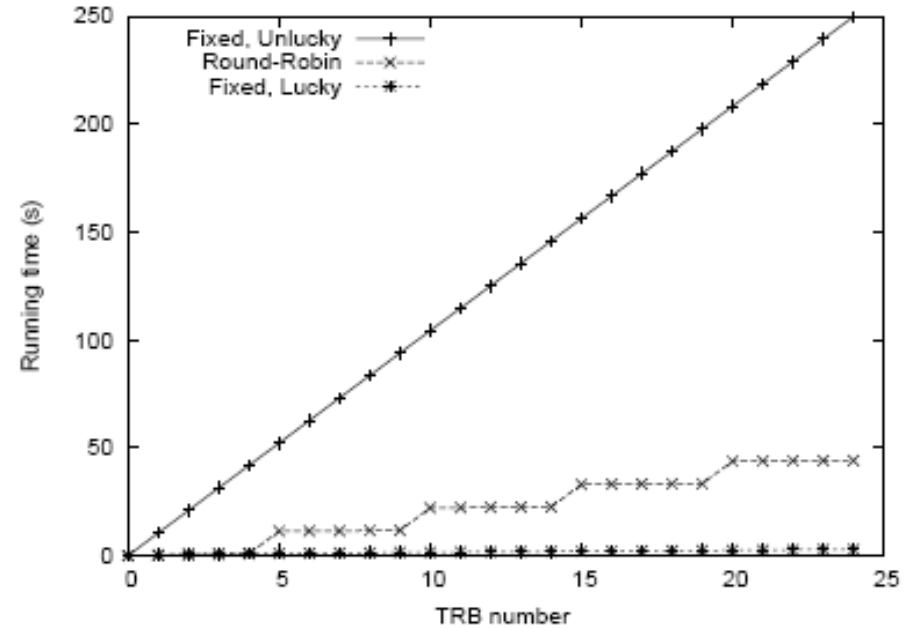


Figure 6: Impact of rotating leadership

Evaluation

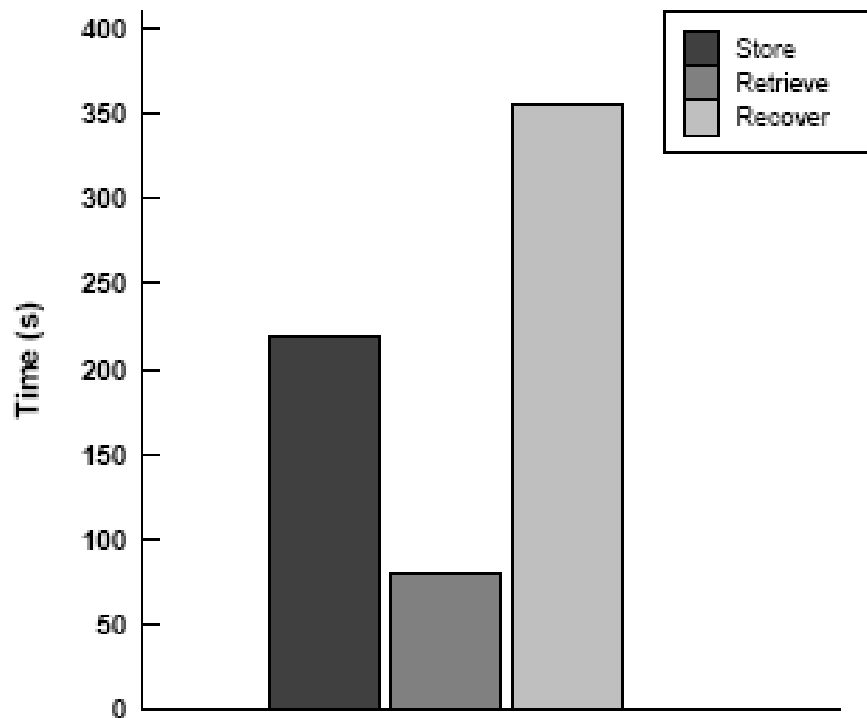


Figure 7: Operation time for 100 MB

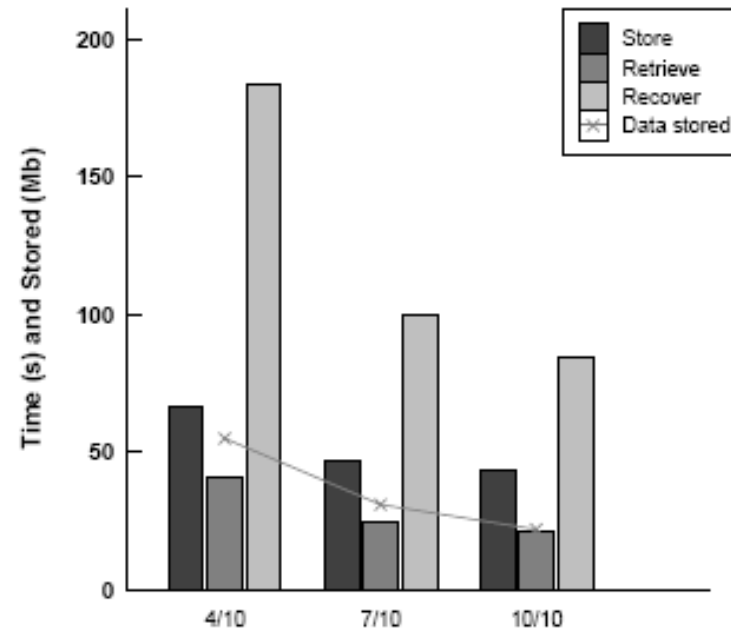


Figure 8: Operation time for 20MB at various encodings

Evaluation

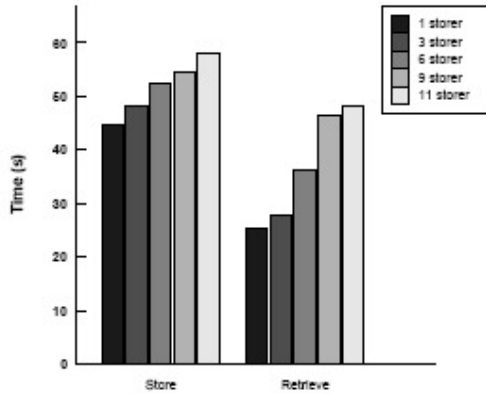


Figure 9: Concurrent operations

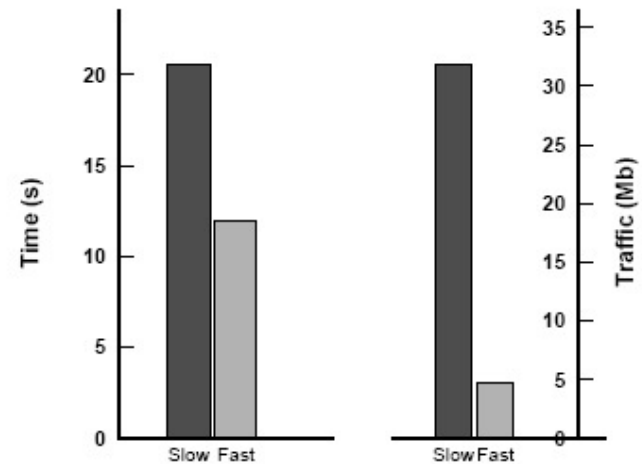


Figure 11: Impact of the fast path optimization

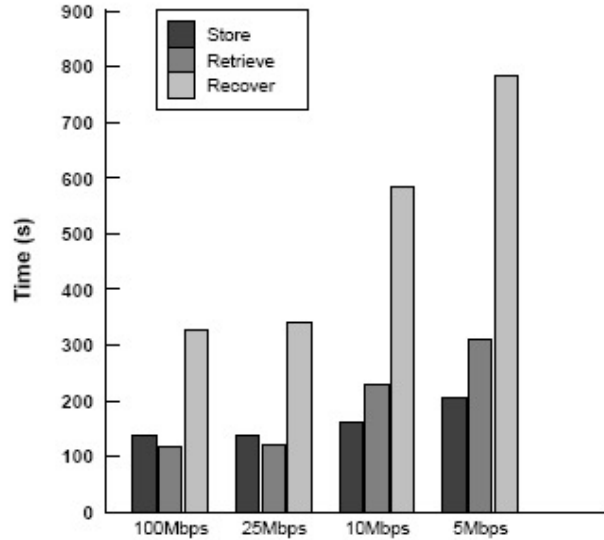


Figure 10: Operation under different network conditions

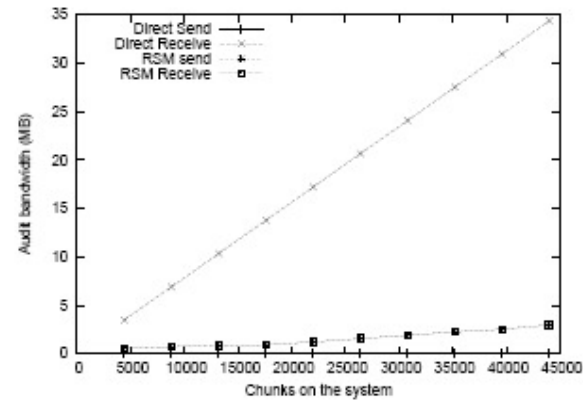


Figure 12: Cost of audit as capacity grows

during an audit increases as the duration of the audit increases be