Faking a Failover

Over the Top
With
Samba Clusters

Christopher R. Hertel
Samba Team
May 2017

Introductions

Introductorationaryesquenessesism

Me:

- Samba Team Elder
- SMB Wizard with:

WEKA.IO

The opinions expressed are my own and not necessarily those of my employer, my spouse, my spirit familiar, the Internet, or the monster in the closet.

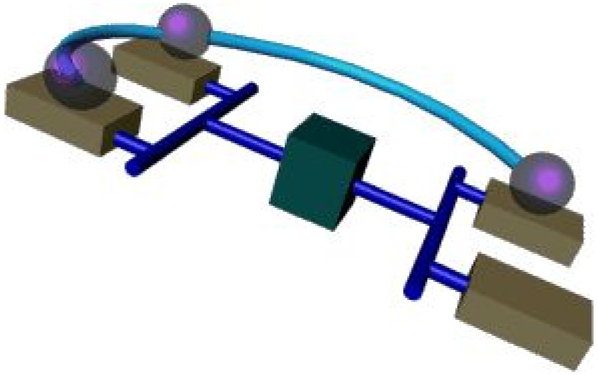
Introductorationaryesquenessesism

Mantra

In theory,
theory and practice are the same.

In practice, they're not.

Quick Review

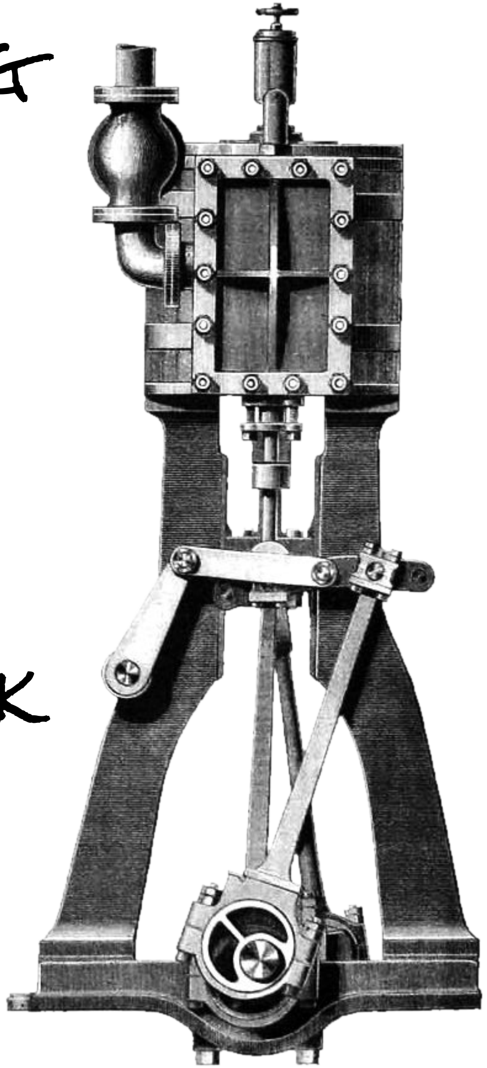


Samba, CTDB, and Clusters

Review: Samba

A Big Giant Semantics Engine

- Locking and Sharing
- Access Controls
- File Attributes
- File Names
- Weird Behaviors



Samba has to keep track of a lot of STATE.

Review: Samba TDBs

Using TDBs:

- Shared Access to State Information
- Atomicity / Consistency
- Resilience (survives reboot)

TDBs, generally, form Samba's state machine.

Review: Samba & CTDB

Using CTDB:

- Volatile State
  - Changes rapidly
  - May be safely lost when a server node is lost
- Persistent State
  - Less dynamic
  - Must be consistent

Provides a distributed state machine.

Review: Samba Clusters

Samba Clusters:

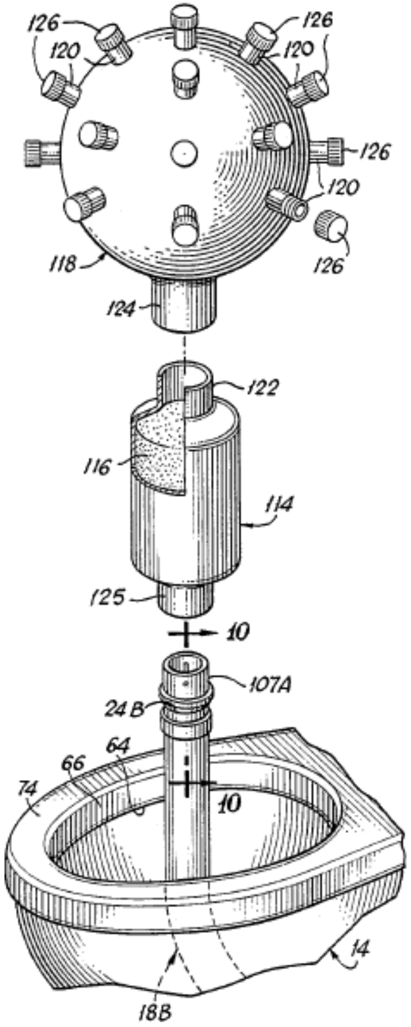
- Provide Windows Semantics
- Coordinate cluster-wide state
- Compensate for missing features in the underlying FS
- Tools for cluster management
- Hard Failover

When a server node fails, clients can reconnect to any cluster node.

Quick Review

Durable, Resilient, and Persistent Handles

Fig-9



Making Handles Sticky

Durable Handles

- SMB2.0 / Windows Vista
- Designed with WiFi in mind
- Requires an OpLock
- Limited state exposed

Samba provides limited support for Durable Handles in a Single-Server Configuration. (Not intended for Use in Clusters.)

Making Handles Sticky

Resilient Handles

- SMB2.1 / Windows 7
- Stronger guarantees
- Doesn't need an OpLock
- Tracks byte-range locks
- Separate IOCTL call required

Samba doesn't support this, but it could be implemented in the VFS layer by catching the IOCTL call. Still, not intended for clusters.

Making Handles Sticky

Persistent Handles

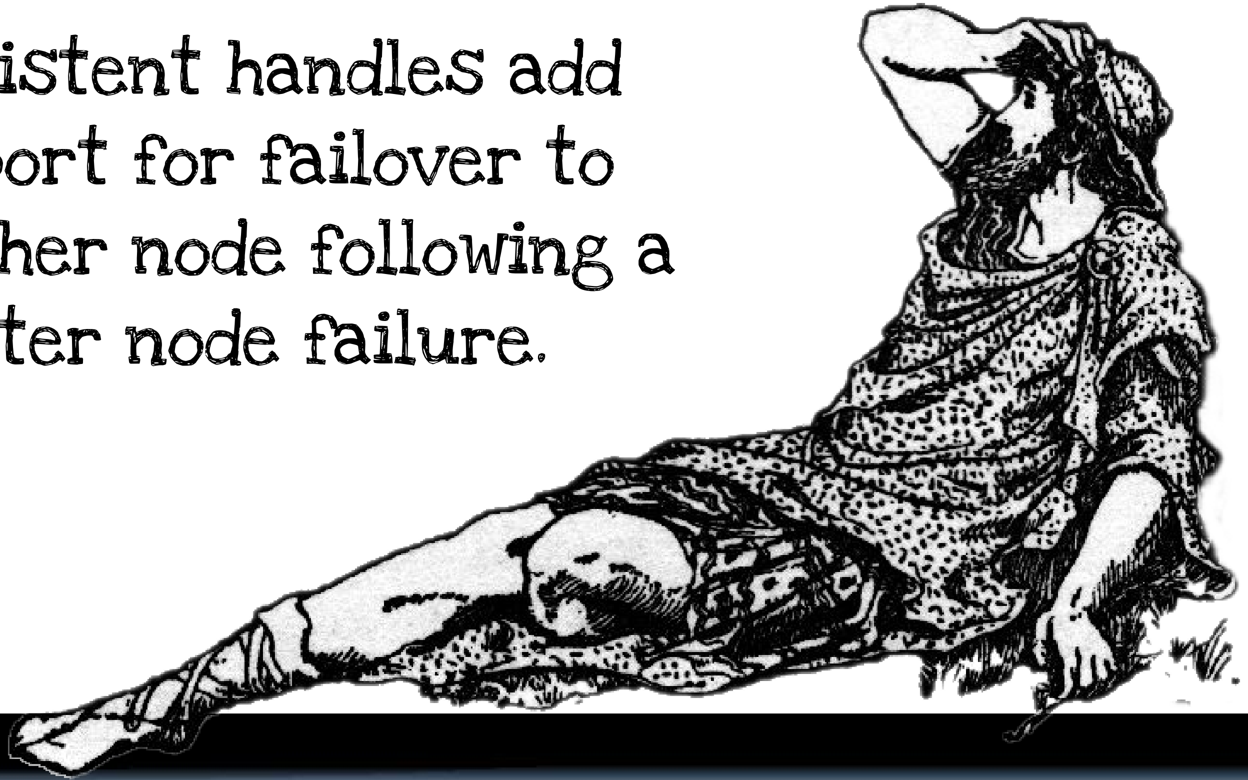
- SMB2.2 (3) / Windows 8
- Real cluster failover
- Automatically requested

Persistent handles were added specifically to support Continuous Availability (CA).

Making Handles Sticky

Crash Recovery

- Durable/Resilient handles provide file-handle recovery following a brief network outage.
- Persistent handles add support for failover to another node following a cluster node failure.



Why Fake a Failover?

Fake Failover

What the Heck?

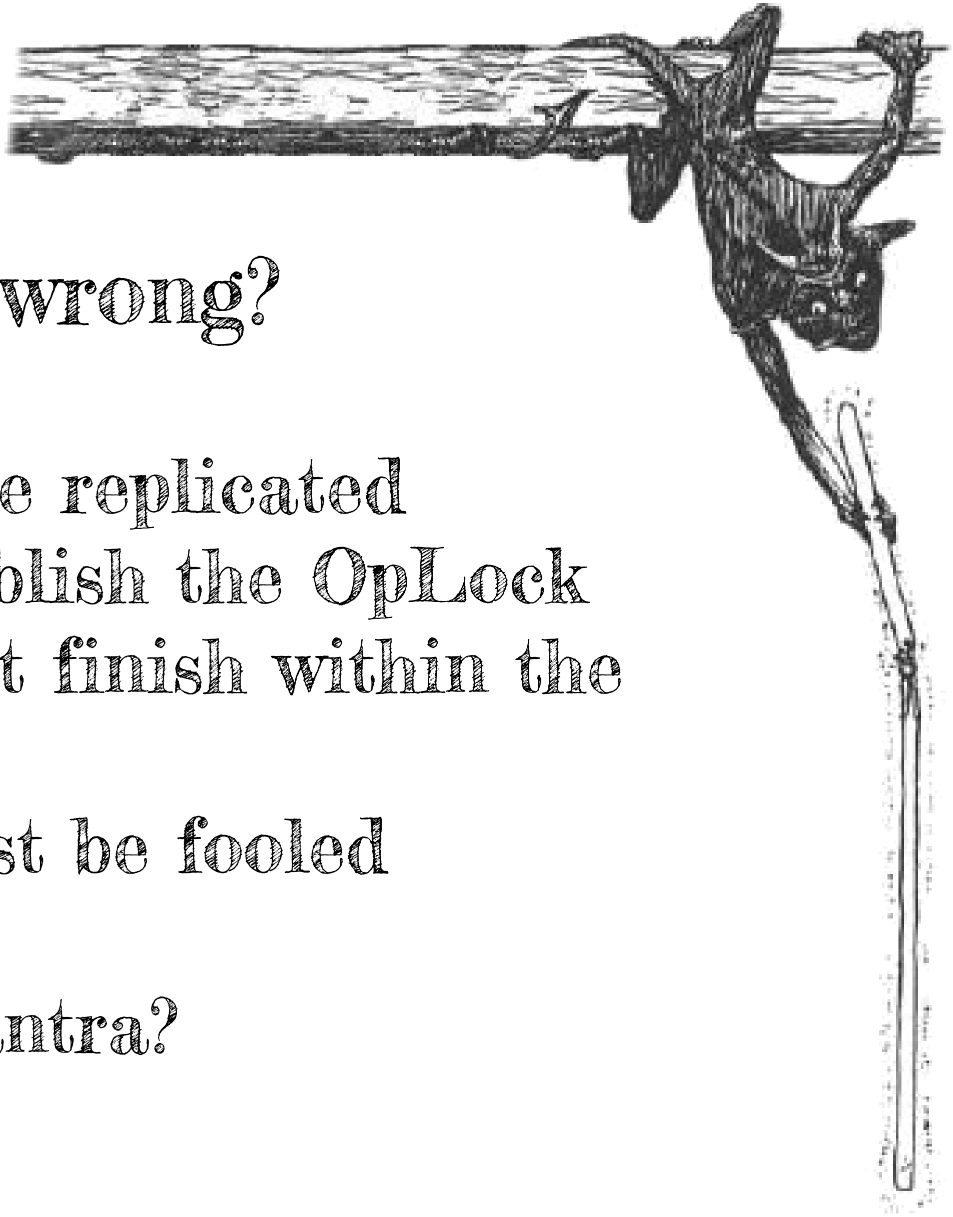
What do you mean by Fake Failover?

- Reconnect a *Durable* Handle...
- ...to a different node

Why do such a silly thing?

- Samba has Durable Handle Support
- Minimal state to keep
- More clients
- Prelude to Real Failover
- Why not?

Fake Failover



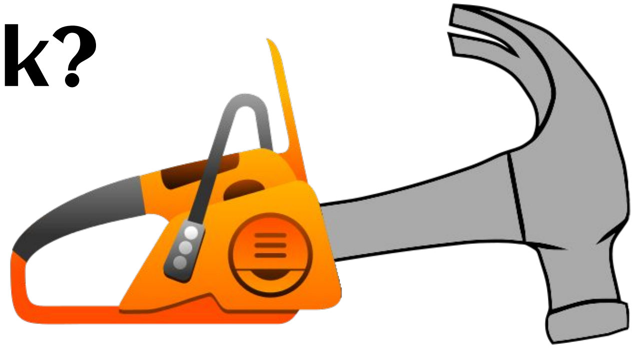
What could go wrong?

- State must be replicated
- Must re-establish the OpLock
- Failover must finish within the timeout
- Windows must be fooled

Remember our mantra?
This is theory.

Fake Failover

How would this work?

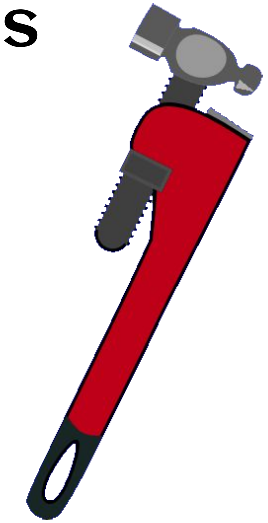


New semantics:

- **Reliable State**
  Handle ID, & any state that exists for the duration of the open

- **Ephemeral State**
  Uncommitted/Un-ACKed changes

Do not expect Durable Handles to survive a full cluster failure. (That's for Persistent Handles.)



Fake Failover

Implementation Options

- **MemCacheD**
  Distributed memory cache with client-driven replication
- **New CTDB modes**
  Uncommitted/Un-ACKed changes

New CTDB storage modes were presented earlier by Amitay/Martin.

EPILOGUE

...and then they showed up with their pitchforks and torches and questions...

Blank Slide