A System Recovery Benchmark for Clusters

Ira Pramanick, James Mauro, Ji Zhu

Sun Microsystems, Inc.

ira.pramanick@sun.com

Sun microsystems
We make the net work.

Outline

- Introduction
- System Recovery Benchmark (SRB) Requirements
- SRB-X & Its Properties
- Variables Impacting SRB-X
- Running SRB-X
  - Results Metrics
  - Examples
- Summary

Introduction

- High availability clusters increasingly common in IT environments today.
- No systematic and consistent methodology to perform availability assessment of clusters.
- Availability benchmarking is relatively new, even for single systems.

Availability Benchmarks

- Benchmarks need to be:
  - Repeatable
  - Portable
  - Practically useful
- Benchmarks need to capture results via simple metrics.

Background

- System Recovery Benchmarking (SRB)
  - Framework for benchmarking automatic recovery of systems.
  - SRB focuses on the automatic recovery capabilities of a system.
  - **Basic premise**: computer systems implement quick recovery designs.
  - Has been previously used to develop a system recovery benchmark for single systems.

SRB-X

- An **SRB** for **HA clusters**.
- SRB-X is
  - Repeatable
  - Portable
  - Characterized by a single metric
  - Measures a common outage mode
- **Goal**: drive industry-wide availability benchmark adoption for HA clusters.

SRB Requirements

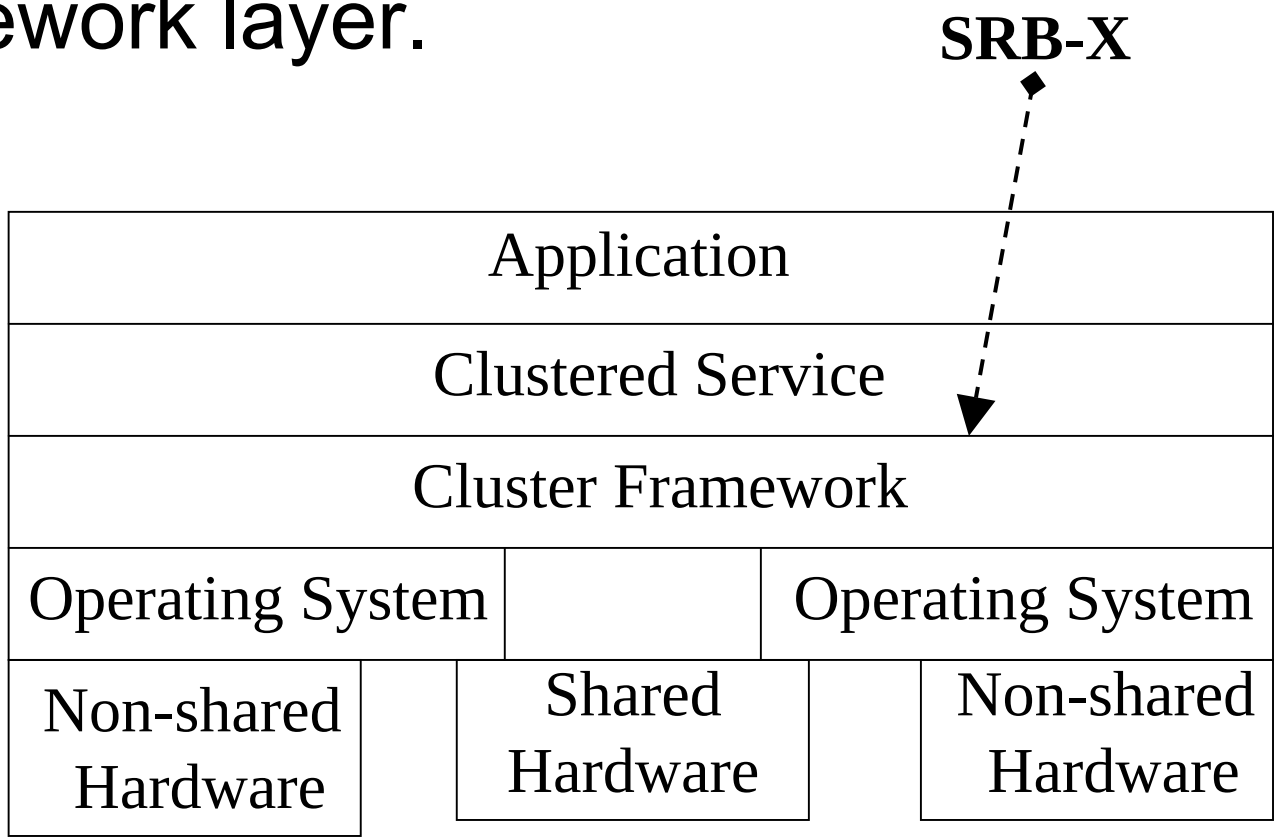
- SRB-Requirement-A
  - SRB trigger -> system being unavailable.
- SRB-Requirement-B
  - SRB trigger -> immediate system outage.
- SRB-Requirement-C
  - Return to service should be automatic.
- SRB-Requirement-D
  - Return to service -> working, usable state.
- SRB-Requirement-E
  - Tuneable parameters should be set to defaults.

SRB-X Property 1

- SRB-X measures recovery for data center clusters, where HA data and data integrity are requirements for the cluster.

SRB-X Property 2

- SRB-X measures recovery at the cluster framework layer.



SRB-X Property 3

- SRB-X measures outage when one of the nodes goes down.

Cluster reconfiguration outage

- Applies to both *failover* and *scalable* services.

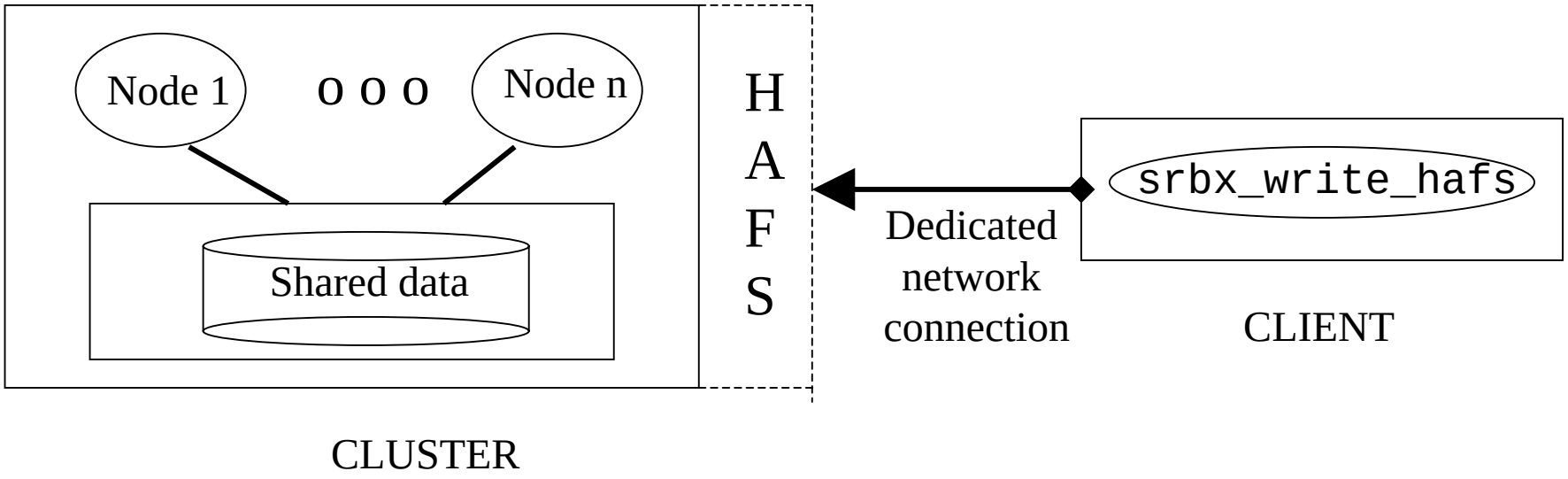
HAFS

- Minimal HA file system.
- Used to measure framework recovery.
- Implies shared data.
- Does not specify implementation of this service.

SRB-X Property 4

- The cluster is recovered if a write command to shared data in HAFS completes correctly.

Write command: `srbx_write_hafs`



SRB-X Property 5

- SRB-X recovery calculation includes the outage trigger detection time.
  - Detection time is an integral part of a cluster's recovery.
  - Default setting for this detection time is required.

SRB-X Property 6

- The number of nodes should be fixed across an instance of measurements and comparison in SRB-X.
  - Reconfiguration time is typically proportional to the number of nodes.
  - Results from clusters of different sizes should not be directly compared.

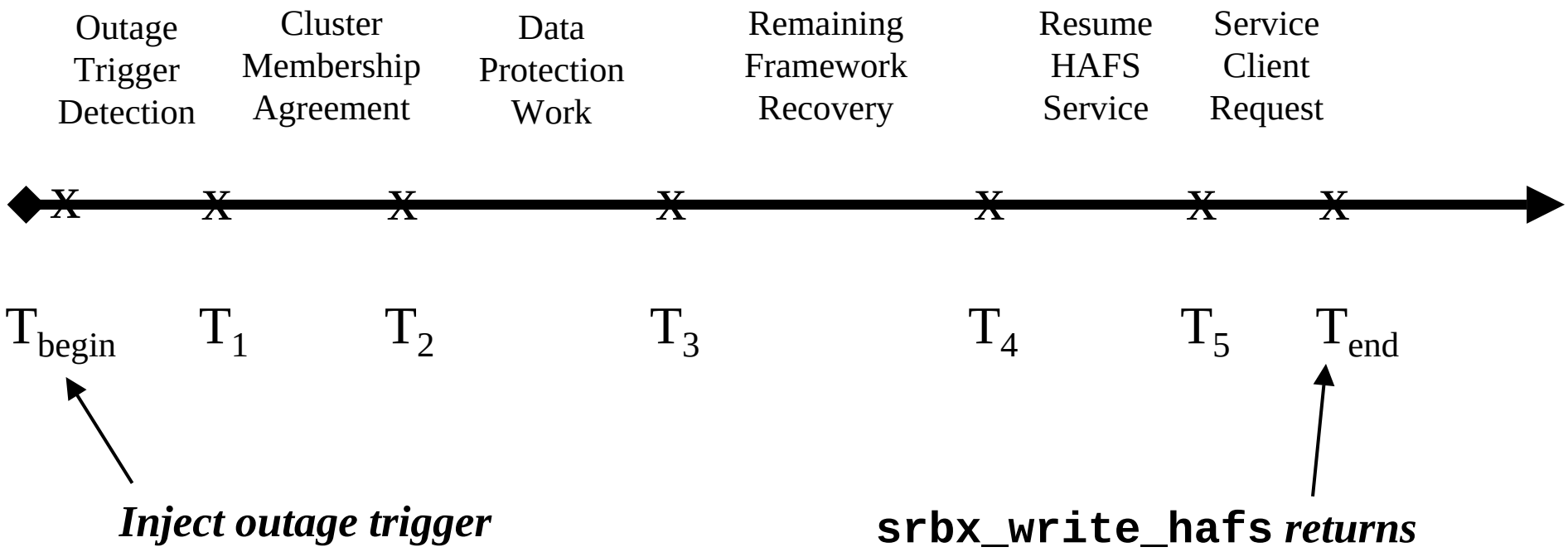
SRB-X Property 7

- Cluster nodes should be homogenous for SRB-X measurements.
  - Lowest common denominator across all HA clusters.
  - Most deployed clusters have homogenous nodes, so this represents the common case.

SRB-X Property 8

- The configuration should be the minimal required to provide HA data access, as documented by the vendor in question.

SRB-X Recovery Components



Variables Impacting SRB-X Recovery

- Outage trigger independence – a result of SRB-Requirement-B.
- Load dependency
  - No load measurements
  - Synthetic load measurements

SRB-X Metric

$$SRB\text{-}X = 3600/(t_{no\text{-}load} + t_{syn\text{-}load})$$

- Single numeric value.
- Larger value -> better result.
- 3600 in numerator: value of an hour in seconds
  - Yields SRB-X value greater than 1.

Examples

	Cluster A	Cluster B
$t_{no\text{-}load}$	65s	90s
$t_{syn\text{-}load}$	90s	140s
SRB-X	**23.23**	**15.65**

Summary

- SRB-X: First proposal for an industry wide availability benchmark for HA clusters.
- SRB-X: repeatable, portable, has a simple results metric, and represents a common cluster outage scenario.
- Currently working on the SRB-X specification.
- Investigating issues in benchmarking recovery at the upper layers in an HA cluster.