CASE STUDY: Enterprise High-Availability (HA) Cluster Engineering

Project Scope: 3-Node Hyperconverged Infrastructure (HCI) with ZFS Replication

Architect: David Culp, Principal Solutions Architect

Key Value: Verified <120 second automated failover for mission-critical medical workloads using zero-shared-storage architecture.

---

1. EXECUTIVE SUMMARY

To eliminate hardware single points of failure in a resource-constrained environment, I engineered a **3-node High-Availability cluster** using **Proxmox VE**. Unlike traditional clusters that require an expensive, centralized SAN (Storage Area Network), this architecture utilizes **ZFS Replication** to maintain data synchronization across local NVMe storage. The result is an enterprise-grade "Self-Healing" private cloud built on high-performance, cost-effective hardware.

---

2. THE CHALLENGE

- **Hardware Fragility:** In a standard single-server setup, a motherboard or power supply failure results in a total blackout of clinical services.
- **Storage Bottlenecks:** Centralized SANs are often a single point of failure themselves and introduce significant latency and cost.
- **The "Split-Brain" Risk:** In cluster engineering, if nodes lose communication but stay powered on, they may attempt to write to the same data simultaneously, causing catastrophic corruption.

---

3. THE ARCHITECTURAL SOLUTION

I implemented a robust HA logic stack that prioritizes data integrity and automated recovery.

A. Storage: ZFS Replication (The "No-SAN" Strategy)

- **Frequency:** Configured ZFS-to-ZFS replication jobs to run every **60 seconds**.
- **Benefit:** Ensures that a near-identical copy of the virtual machine's disk exists on all nodes at all times, allowing for rapid recovery without waiting for multi-terabyte data transfers.

B. The Logic: Quorum & Fencing

- **Quorum:** Established a 3-node minimum to ensure a mathematical majority is required before the cluster can make "life or death" decisions about a VM.
- **Fencing (Watchdog):** Configured hardware-level fencing. If a node becomes unresponsive, the cluster "fences" the node to prevent data corruption before restarting services on a healthy node.

---

4. THE "PULL-THE-PLUG" VALIDATION

To prove the resilience of this architecture, I conducted a physical stress test:

1. **Baseline:** Mission-critical VM (EHR Database) running on `pve-01`.
2. **The Event:** Manually disconnected the power supply from `pve-01` during a simulated high-load period.
3. **The Recovery:**
    - **0-60s:** Cluster detects "Node Down" state; Quorum remains maintained by `pve-02` and `pve-03`.
    - **60-120s:** HA Manager marks the node as dead and initiates the boot sequence on `pve-02`.
    - **Result:** Services restored and reachable via network in **under 2 minutes** with zero manual intervention.

---

5. TECH STACK

- **Hypervisor:** Proxmox VE 9.x (Debian-based)
- **Filesystem:** ZFS (RAID-Z1 for local redundancy)
- **Management:** Proxmox Datacenter Manager (PDM) for multi-cluster visibility.
- **Backup:** Proxmox Backup Server (PBS) with deduplication and encryption.