SAN design

Dr Peadar Grant

April 10, 2024

Contents

1 Purpose

Online host-specific block storage essentially replacing directly-attached block devices on the hosts.

Online clustered block storage allowing a number of hosts to share a block device that is formatted with a clustered/shared file system.

Offline block storage replacing removeable drives on hosts. This may be for archival purposes.

Backup where the SAN facilitates access to block devices for backup.

- May be disk-backed LUNs on storage appliance.
- May also include tape-based backup systems on SAN-attached tape equipment.

1.1 Host-specific considerations

- Will hosts be booting from the SAN?
  - There are a number of ways to provision this.

2 SAN type

Motivations for IP SAN:

1. Low/no-cost entry
2. Leverage existing IP knowledge
3. Re-use of infrastructure (incl. WAN)

Motivations for FC SAN:

1. Performance
2. Reliability
3. Inherent SAN/LAN demarcation.

2.1 Additional considerations

- Problem-solving technologies available: FCoE, FCoIP, iSCSI-FC bridging.
  - Unwise to base network around these technologies.
- Depending on organisation structure and culture, there may be different teams responsible for storage vs general-purpose networking.

3 High Availability

SANs hold high-value information critical to continuity of service. A naively designed SAN may exhibit worse availability characteristics than simple direct-attached storage on individual hosts.

3.1 Storage layer availability

Storage appliances normally include a number of features to ensure high availability, Figure 1.

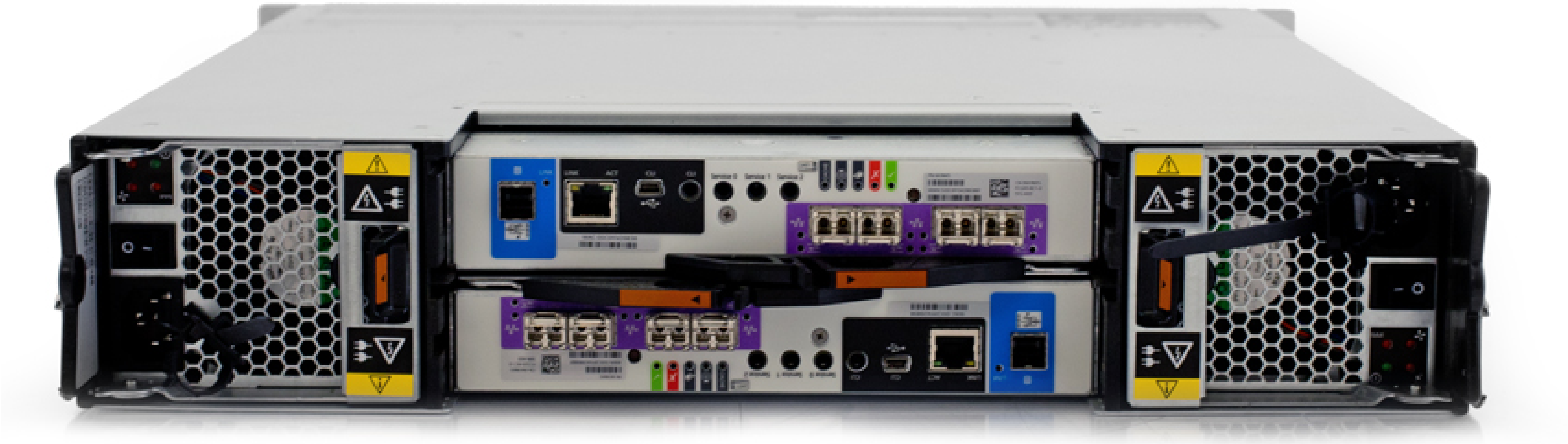

Figure 1: Rear of Dell PowerVault storage appliance

Disk redundancy where volumes / LUNs are provisioned such that the array is redundant against disk failures. May include RAID, hot spares etc.

Dual power-supply where each storage applicance has two Power Supply Units (PSU).

- Each PSU independently capable of supporting the load.
- Data centre environment should have dual power supply to take full advantage of dual PSU.

Redundant cooling components permitting the storage appliance to continue operating despite the failure of 1 (or sometimes more) fans.

- Should be easily replaced.
- Often hot swappable.

Dual controller where the storage array has two controller modules (i.e. high-performance embedded computer).

- Normally works so that either controller can independently provide all services.
- May be hot swappable, Figure 2.

Remote management capabilities that allow the storage appliance to alert administrators should any failures occur. Essential that all alerts are enabled and acted upon to avoid storage appliance remaining in a degraded state (e.g. failed PSU) as redundancy will hide this situation.

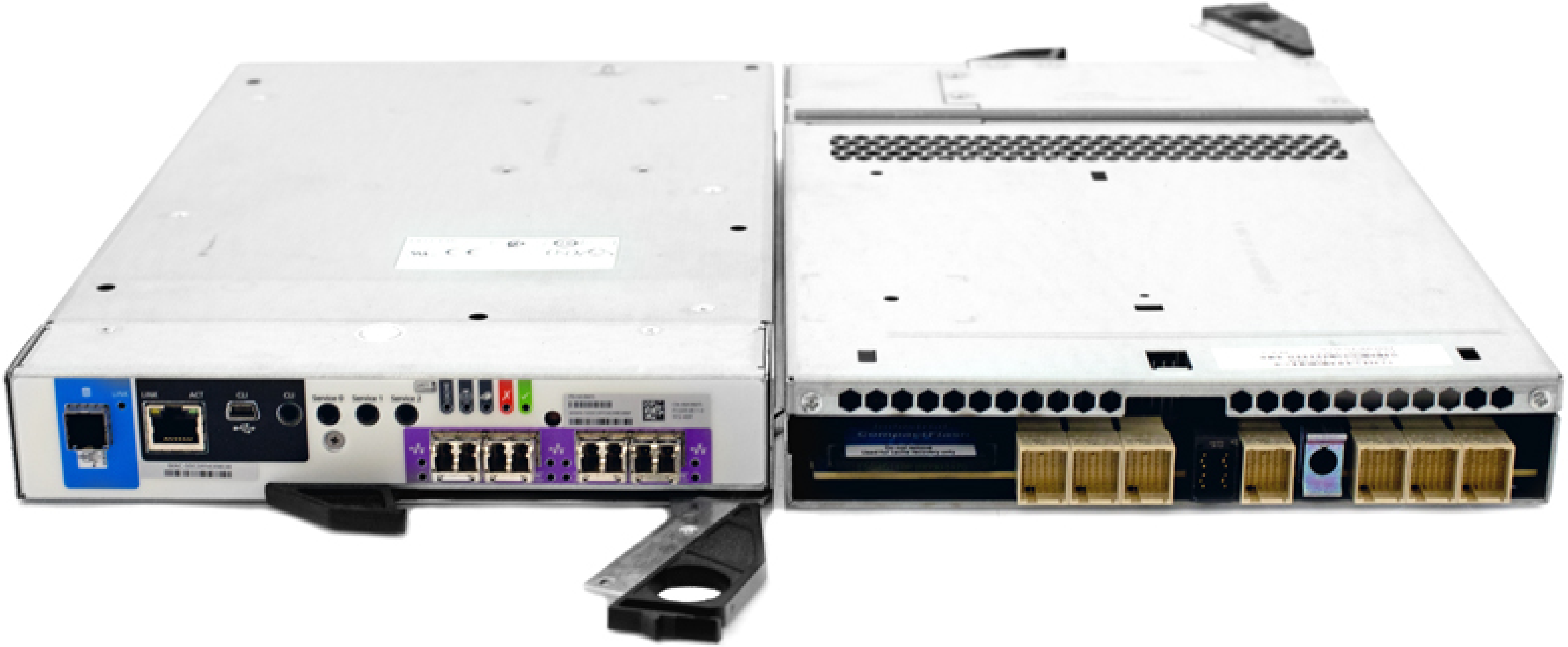

Figure 2: Hot-Swap Controller

3.2 Fabric layer availability

The fabric (regardless of type) can represent a significant single point of failure. Consider the simple SAN layout in **??**.

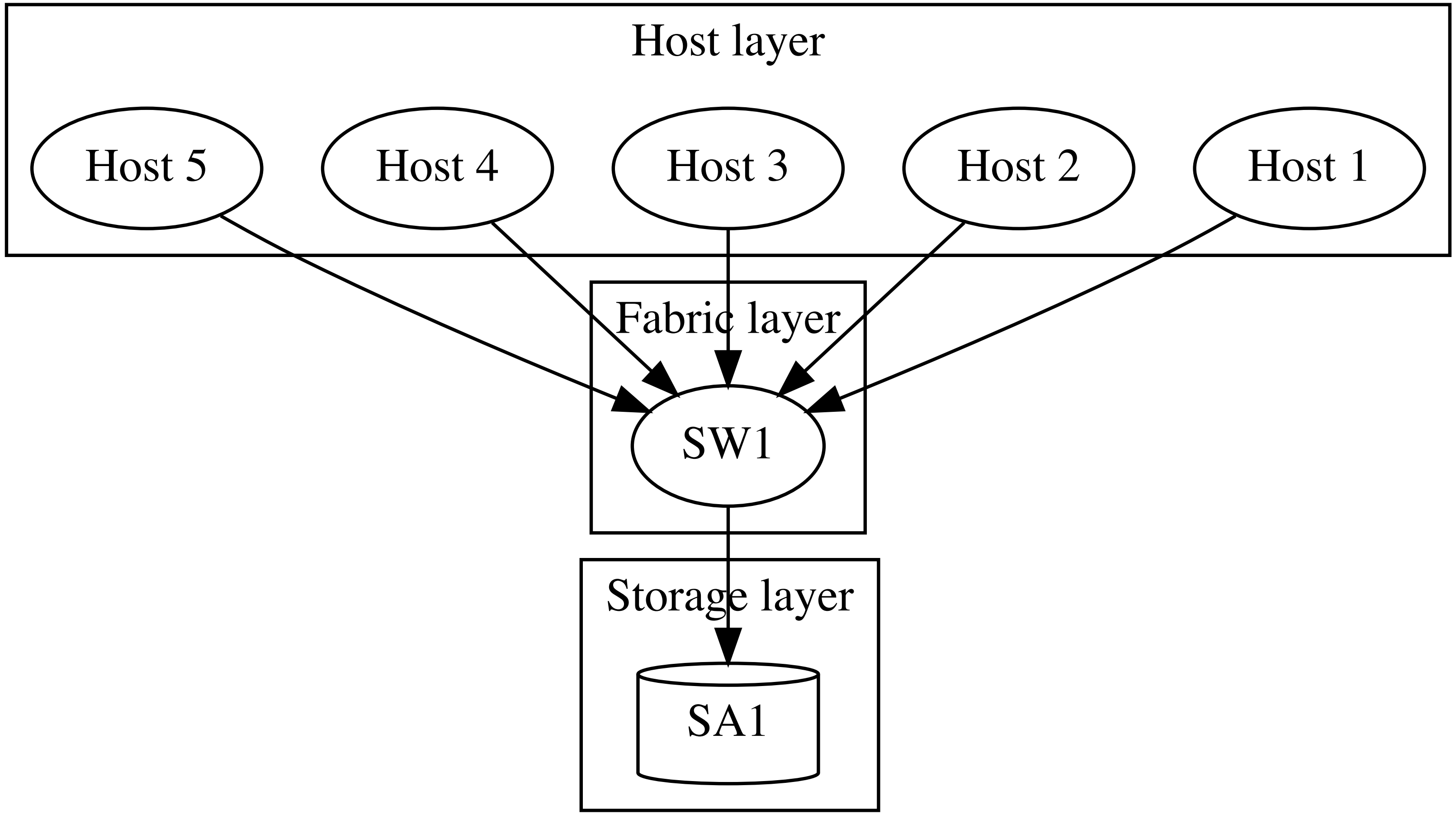

Figure 3: Fabric with no redundancy

This is vulnerable to a failure of the switch SW1. Duplicating the SAN fabric averts this, Figure 4.

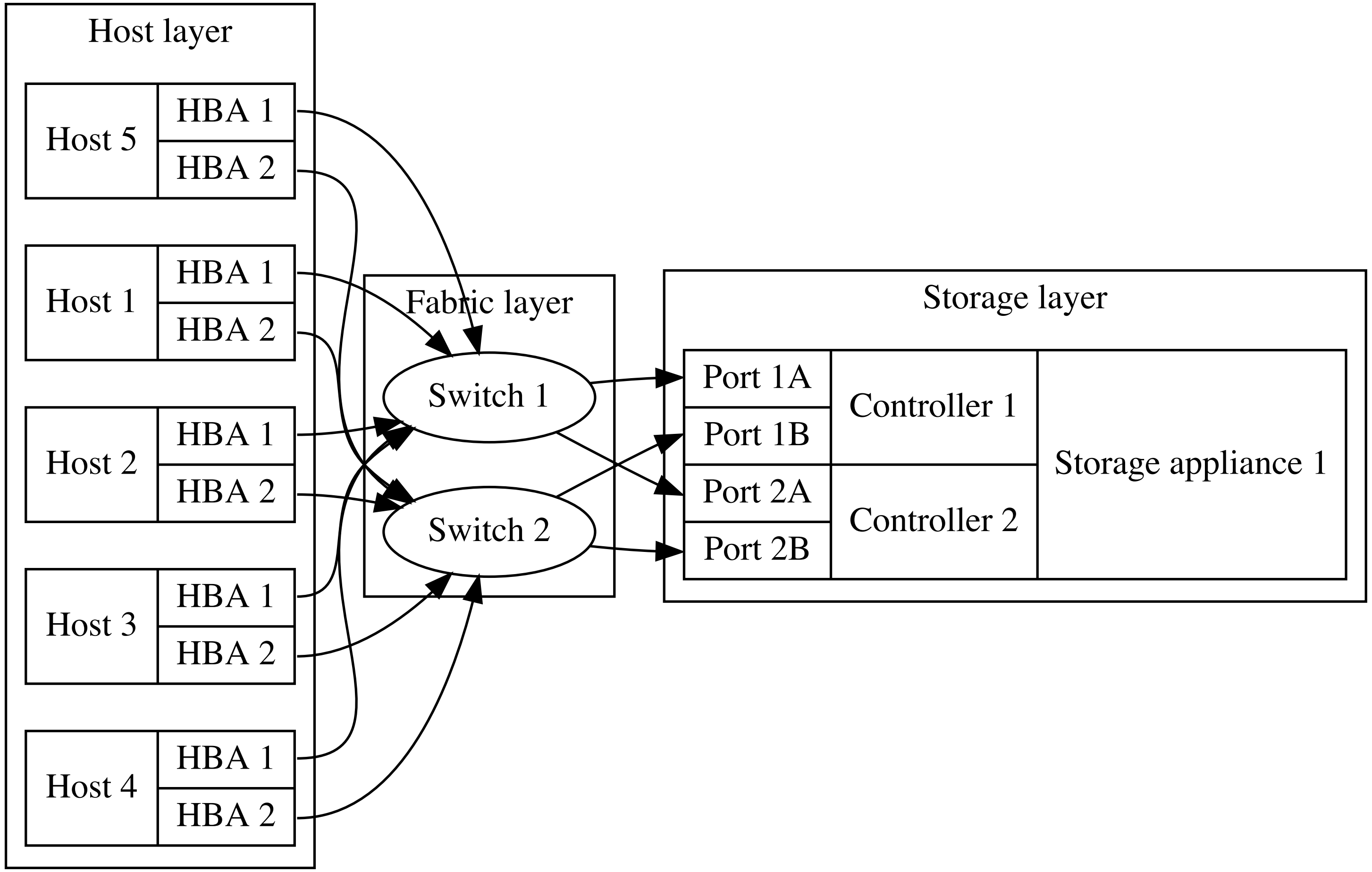

Figure 4: Fabric with redundancy

SAN fabric components (e.g. switches) should be dual-powered where possible. This may be via dual-PSU or via external transfer switches.

3.3 Host layer availability

Assuming that there is redundancy in the fabric, the host layer requires approriate redundancy to connect to the fabric.

To transparently utilise the redundant fabric we must ensure that multipathing is correctly configured.

3.4 Data centre environment

A SAN will normally be provisioned within a data centre environment. We should assume at a minimum:

Power supply is protected by UPS, possibly generator and has diverse A and B power paths.

Cooling is sufficient to meet the load and incorporates redundancy to maintain operation in the event of a failure.

Fire detection and suppression to protect data and ensure service availability.

Physical security appropriate to the data and service risk profile.

The provision of a SAN should dictate that the infrastructural services are sufficient.