

AFS Fiber Channel Network Test Plan

Dale Ghent 4/13/05
UMBC OIT/Syscore

Equipment:

- 3x Sun V20z with Qlogic FC HBAs
- 2x Apple Xserve RAID arrays, 4x 2.2TB
- 2x Qlogic SANbox 5200 FC fabric switches

Software:

- Solaris Volume Manager (SVM)
- Apple RAID Manager
- Qlogic SANbox Manager

Goals:

(listed in order of execution)

1. Test different RAID strategies (RAID levels 1, 10, 5) and formulate the standard RAID/LUN layout of the Apple Xserve RAID arrays. Aim for the best combination of fault-tolerance, useable disk space, and serviceability.
2. Test fault-tolerance of fiber channel fabric and note any unavoidable potential points of failure of the fabric, including physical faults (eg: cable breaks, switch failure) to accidental misconfiguration of fabric components.
3. Configure and test SVM Disk Sets on all hosts and determine host import/export requirements.

Goal 1 Summary and Methodology

This goal will involve testing of several different RAID configurations on both the Xserve RAID arrays and Solaris Volume Manager (SVM.) As stated, our goal will be to derive a configuration with allows us to maximize our available disk space but maintain performance and availability in the case of a catastrophic failure of either one of the Xserve RAID arrays or a portion of the fiber channel network.

Potential configurations all include use of SVM locally on each V20z server. This is required to maintain cross-array mirroring of volumes so that partial or complete loss of a single Xserve RAID array does not cease the availability of any one volume to the hosts. SVM also allows managed transfer of volumes from one host to another, enabling us to address the potential scenario of a host becoming unavailable due to catastrophic failure.

Potential Configurations include:

1. Two RAID-5 sets on each Xserve RAID and divided into n LUNs. Each V20z host would be assigned its own corresponding LUNs on each Xserve RAID and a SVM disk set would be configured to mirror these two LUNs between arrays
2. Multiple RAID-1 sets configured on each Xserve RAID array and as one LUN. LUN assignment and SVM configuration would be configured as prescribed in Configuration (1)
3. Two RAID-0 sets configured on each Xserve RAID array and divided into n LUNs. LUN assignment and SVM configuration would be configured as prescribed in Configuration (1). However, this configuration does not utilize any RAID redundancy features on the Xserve RAID arrays.

Goal 2 Summary and Methodology

Due to the comparatively physical size of the fiber channel fabric within Room 012C, there is more potential for failure of the network due to it not being contained in one or two side-by-side cabinets.

Possible fault causes may include:

1. Complete FC switch failure
2. Partial FC switch failure (dead switch port or SFP transceiver)
3. Loss of one or both of cross-room switch interconnects
4. Loss of one drive within the Xserve RAID sets.

Testing for fault tolerance:

1. Simulate single switch loss by cutting its power. Monitor hosts for anomalies, and reapply switch power to test host recovery and resulting data sync.
2. Simulate loss of one of hosts' and arrays' two paths to the fiber channel switches.
3. Simulate loss of one of the two interconnects between the two switches.
4. Pulling one random drive from a Xserve RAID.

While this testing is being executed, there will be constant monitoring of the hosts for any changes in performance. Any changes should be investigated and documented.

Goal 3 Summary and Methodology

SVM will serve as the cross-array mirroring mechanism. It will maintain data sync across a host's LUNs on each Xserve RAID array and mask loss of one of those LUNs from the rest of the system.

Basic SVM mirroring may be utilized to achieve this. However, we may take this one step further and use the Disk Set facility of SVM to enable us to transfer use of SVM volumes to another host. This creates a basic fail-over option in the event of the complete loss of a host. In this case, that host's SVM Disk Set would be transferred to another host and be served from there. Downtime would not be avoided, but would be minimized.

Potential Configuration of a Host Disk Set:

1. Create the host's Disk Set containing the two (or more) LUNs assigned to that host.
2. Within the Disk Set, configure a mirrored volume of those two LUNs
3. Assign a second host to that Disk Set configuration. This would be the fail-over or "Buddy Server" of that host.

Testing Disk Sets:

1. Testing of Disk Sets may be done in conjunction with Goals (1) and (2)
2. Test release (export) and take (import) of a Disk Set between two hosts
3. Note file system recovery issues that may present themselves.