



# Best Practices

For a successful SAN deployment

## **Summary**

MetaSAN (as well as MetaSAN iSCSI, MetaLAN Server and MetaLAN) can be deployed in many ways to achieve a multitude of workflow topologies. In this document, Tiger Technology recommends specific guidelines that will ensure successful deployment of MetaSAN. Following these guidelines will maximize performances, ease of use, and reliability of MetaSAN in your facility.

## **What Is a Best Practice?**

Tiger Technology and its partners have developed significant operational experience with MetaSAN software and have documented a number of techniques that are most likely to result in a highly successful deployment. These techniques are called best practices, and this document describes them. Please note that these practices are recommendations, not requirements, as some recommendations may not apply to your circumstances. However, Tiger Technology believes everyone will benefit from thinking through these recommendations and making a conscious decision to follow or reject them.

## **Setting Proper Expectations**

One should always keep in mind that there will always be certain limitations when working on a SAN, versus that of locally attached and networked movement of media. That's because certain applications have been optimized by making certain assumptions, and bypass standard disk access procedures. The fact remains that these limitations will differ with each workflow and the benefits of SAN based topology far out way that of conventional methods.

## Recommendations

### Volumes:

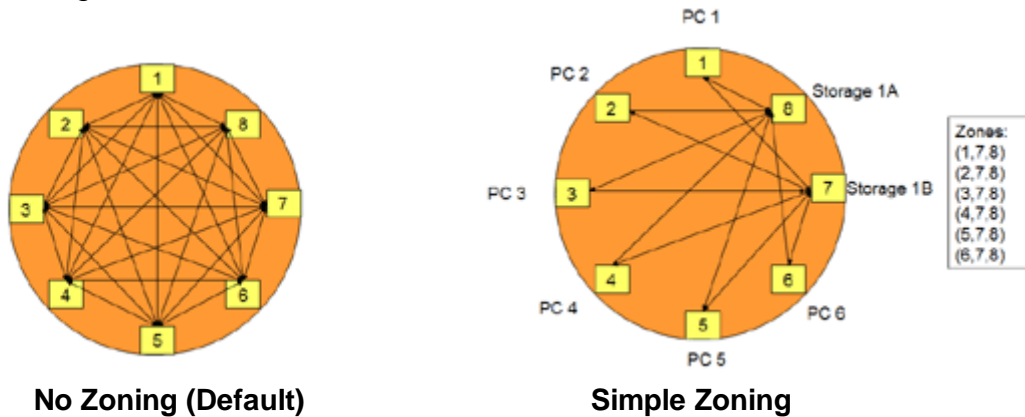
- 1) Avoid creating very large volumes. Keeping volumes no larger than 8-16TB will improve overall performance.
  - a) Mounting and parsing of the file system directory structures of a large volume takes more time
  - b) Volume maintenance and repairing of large(r) volumes will delay returning to operational state.
  - c) Turn to Tiger Technology poolIt for expanding or managing larger volumes.
- 2) Keep LUN size under 2TB for maximum compatibility.
  - a) Most 32-bit operating systems, such as Windows XP, cannot access LUNs larger than 2TB. It is therefore always best to stripe multiple 2TB LUNs together when creating large volumes. This recommendation applies to metaSAN, metaSAN iSCSI and metaLAN Server setup.
  - b) metaLAN clients do not see the physical LUNs and will therefore mount any supported SAN volumes.
- 3) A good rule of thumb for reliable SAN performance is to allocate 15 physical disks (capacity doesn't matter) for each 200MB/sec of effective SAN bandwidth required.
  - a) This might appear low when compared to the advertised (and even measured) performances of certain RAIDs, but keep in mind that most storage is optimized for serial I/O (single read or single write operations) and are not optimized to support multiple hosts performing concurrent I/O operations. In a SAN, this scenario is the norm and the raid controller performance (host port) will usually be dramatically less than the advertised performance. It is mostly the number of spindles (heads) that will compensate for performance loss associated with increased seek times.
  - b) If a host port is capable of theoretically 200MBs, the operating systems and driver overhead will remove 20% which might drop the host port performances to 160MBs.
- 4) Ensure the optimal RAIDs settings are used. These can be obtained from your storage manufacturer. If using Apple Xraid, use the same setting that Apple recommends for XSAN.
- 5) Format the volume using chunk sizes that match your application i/o pattern.
  - a) Applications performing SD accesses do not have the same requirements as those doing HD. If you do a mix of both, go for the most critical application.

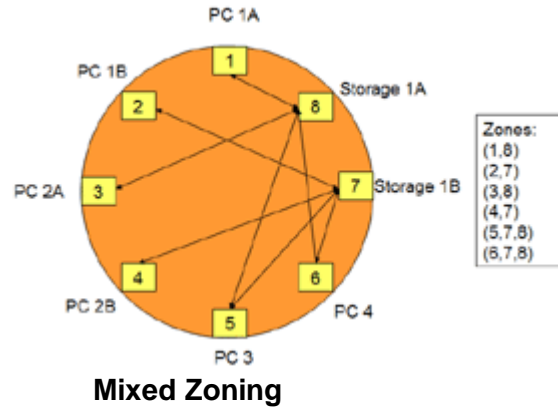
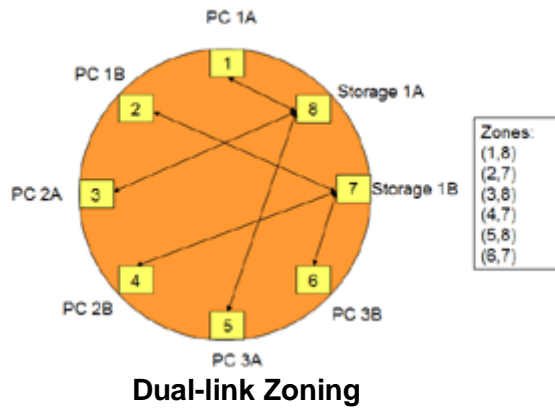
- b) MetaSAN offers a Dynamic Disk Optimization Tool for NTFS file system that enables you to format stripe sets using chunk sizes greater than the normal Windows 64KB size.
- 6) When formatting an NTFS file system using dynamic disk on the SAN
- a) Ensure there are no other dynamic disks in any of Windows workstations that connect directly to the SAN disk as this could lead to the loss of your stripe set information.
  - b) Make a backup of the LDM database after creating a software stripe set using Windows. Many people rely on the fact that their RAID is 50 protected and neglect to safeguard the striping information. Without this information, Windows might not be able to reconstruct its stripe set (after power failure, etc.). MetaSAN offers an LDM backup utility to take care of this.
- 7) When formatting an HFS+ volume >2TB, pay special attention to the MetaSAN installation notes
- 8) Do not fill volumes to more than 80% capacity.
- a) Most file system providers have recommendation about how full you may safely fill the system, each one differs, as a rule of thumb, maintaining at least 20% free space is highly recommended by all file system providers
  - b) Maintaining at least 20% free space will help maintain desired performance as performance on average drops significantly after that threshold.
  - c) There are real risks of file system corruption when HFS+ volume gets too full.
  - d) Reviewing file systems documentation from the provider and reviewing new groups on these topics will give some feel for best practices
- 9) Avoid keeping a huge amount of files in a single folder. This will improve the file system response time to both read and write commands.
- a) Try to keep an entire project within the same folder and subfolders. This also includes any audio that is imported to be used within that project.
  - b) A well defined tree structure helps as all the file systems are optimized for handling complex tree structures.
  - c) Avoid folder structure more than 5 directories deep. It has been proven many times that having deep file system structures will cause performance lags even on local file systems.
- 10) Run check disk and defrag on a regular basis to ensure file system integrity, especially after metadata master failures. Application crashes or systems that are not shutdown cleanly will leave some files open. This can lead to corruption of the file system. A little bit of corruption will evolve quickly deteriorate into lots of corruption (like a cancer).

- a) Be aware that the automatic checkdisk routine is disabled during boot sequence (a workstation can't perform checkdisk on the SAN while others are accessing the storage).
- b) You should schedule routine maintenance of your SAN volume. This is especially important when relying on a floating master.

**Fibre Channel Network:**

- 1) Select HBA based on your storage vendor's recommendation. MetaSAN is hardware agnostic and doesn't have any impact on the FC protocol, but hardware devices that communicates together do. Whenever possible, choose your HBA and switches from same manufacturer to guarantee maximum compatibility.
  - a) Proper zoning improves overall performances as each host is isolated from the traffic taking place between the other hosts and the storage. There is therefore no need to analyze those packets.
  - b) Windows OS cannot resolve dual-link configurations without proper zoning. It is recommended to uninstall MetaSAN PRIOR to zoning your switch. You should also zone your switch PRIOR to creating NTFS dynamic disks in order to avoid changes in LUN configuration.
  - c) Ensure all host systems are zoned to target (storage). Each zone should include all the storage you want that Fibre channel port to see and only one host. The following diagrams depicts a variety of single and dual-link configurations:





- 2) Hard-set the speed on the FC network including, switch ports, HBA's and storage
  - a) You'll get better performance (no negotiation required) and increased reliability
- 3) Avoid using a combination of 2Gb and 4Gb devices whenever possible.
- 4) If using a Qlogic switch, make sure to enable **"I/O Stream guard"** on host ports (workstations), and disable it on target ports (storage).
  - a) Ensures that each host system does not have interrupted I/O when another system is shutdown or reboots.

## Ethernet Network:

- 1) Use static IP addresses to ensure a more reliable setup
- 2) Take advantage of IEEE 802.3ad Link Aggregation (AKA trunking) to increase bandwidth delivered to LAN clients from your data servers.
  - a) Many popular multi-port NICs and manageable switches now support this standard
  - b) Aggregate ports on your NIC through the driver settings, and aggregate the equivalent ports on your switch
  - c) Client workstations can still connect using regular Gb connections
  - d) Combining multiple ports does not always provide a linear performance increase, but definitely remains the best way to achieve the optimal performance
    - i) Since aggregation is done in software, it requires CPU resources. Performance will therefore increase if your workstation has plenty of processing power.
- 3) Implement a dedicated network for Metadata controllers and SAN clients.
  - a) While the traffic on the office LAN network may not be extremely high, having potential network hangs and glitches that are common on LAN's must be avoided.
  - b) This network should have a unique IP stack and subnet.
  - c) A dedicated VLAN or a dedicated switch is recommended.
  - d) Make sure that all the ports are part of the same VLAN in the switch.
  - e) This dedicated network should not have an internet connection.
  - f) If re-configuring your SAN from using a public LAN to a dedicated one, consider creating a whole new SAN Definition
- 4) The office LAN network should be used for all other LAN traffic (Internet, file sharing, etc.).
  - a) If a dedicated network is not an option, analyze network traffic and check for collisions.
- 5) If using iSCSI, a separate network for payload is also recommended (replaces the FC connection)
- 6) When using jumbo frames, make sure it is enabled on ALL NICs and is the exact same size on all the systems. System will slow down if you are off by even 1 byte. You also need to make sure that both the TX and RX frames sizes are set to jumbo frames, on some drivers the jumbo frame size is set as a universal setting and on other the TX and RX are set individually.

## Dedicated metadata master:

- 1) There is always one (and only one) machine mastering a given SAN volume. The master receives requests and grants permissions to other SAN members.
  - a) The machine must be of the native file system format. For instance, a Windows machine is required to master an NTFS volume, and a Mac is required to master an HFS+ volume. As such, failover of an HFS+ volume cannot take place on a Windows machine, and vice-versa.
  - b) While the mastership role doesn't require a lot of CPU processing, it does need responsiveness.
  - c) The more files a master must serve or the more fragmented a file system is, the more CPU cycles will be consumed by MetaSAN to serve requests from client machines. More resources are necessary when gathering cluster chains for fragmented files or finding suitable space to allocate for new files on a fragmented file system.
  - d) The master requires a direct connection to the storage.
  
- 2) By default, the first machine to boot becomes the SAN master. It is however possible to define mastership priorities among the various SAN members (it is also possible to assign different masters to different SAN volumes).
  - a) A master will retain mastership of the SAN until failover occurs:
    - i) The workstation fails (hangs or shutdown)
    - ii) The workstation gets manually demoted
  - b) Failover will cause mastership to be transferred to the next available workstation (based on assigned priorities) or to the manually selected workstation in the case of a promotion
  
- 3) Failover will automatically occur if the current master stops responding (after adjustable timeout period). While any "native" member can be elected as a floating master, there are instances where a "dedicated" machine is recommended as each failover occurrence will cause the SAN to become unresponsive for a few second. The use of a dedicated metadata controller is therefore recommended in the following scenarios:
  - a) When failover disruption is not acceptable to users (mission critical environments).
  - b) When CPU availability of the floating master cannot be guaranteed
    - i) If the machine acting as master is busy rendering a job, there might not be enough CPU left for metaSAN to perform mastership of the SAN, therefore causing the SAN to be sluggish.
  - c) When stability of the floating master cannot be guaranteed
    - i) If master hangs, failover will occur, causing some disruption of service.
  - d) When the Virtualization for Avid feature is enabled
    - i) In this case, a dedicated master is always recommended because of complex permissions settings.

- 4) As a rule of thumb, any setup with more than 4 SAN members should have a dedicated master.

### **MetaSAN Configuration:**

- 1) Multiple stations writing a SAN will cause each system to interleave the write commands, thus leaving segmented data across the drives and file system. MetaSAN lets you set pre-allocation size for write operations:
  - a) The default setting is 100MB and equates to 3 seconds of SD material or .6 seconds of HD material.
  - b) For HD work, increase the MetaSAN pre-allocation size to 500 MB or higher. This can be done from the MetaSAN UI Settings tab.
  - c) Rendering applications might perform better with 0MB as they tend to write small chunks of data. In this case, pre-allocation might introduce an additional delay.
- 2) In a cross-platform setup, consider configuring for Linux style shared access. Contact Tiger support team for details
- 3) Even if a workstation can be dedicated to act as a primary master, you should choose "floating master" during the SAN installation to allow failover to take place. Simply ensure that your primary workstation has the highest priority on the list of members

### **Other:**

- 1) Remove all unnecessary and non production software from SAN systems, including metadata controllers (especially resident programs).
  - a) Little used or non production software takes away vital systems resources from applications required for production, including MetaSAN.
  - b) Avoid using unnecessary workstation and server ports, such as com and parallel ports. This will release system resources to overall increase system performance and stability.
- 2) Avoid using anti-virus software on SAN workstations. They tend to sit at the same level and can interfere with MetaSAN. At the very least, they should not manage the SAN volume.
- 3) Backups should be completed during non production hours.

- a) Breakup backup sets into multiple segments as oppose to one large continuous set.
- 4) Backup software and hardware will typically best work when installed on the master machine. That's because backup software tend to play at low level, much like metaSAN.
  - a) Compatibility with various backup software is not guaranteed. If unsure, please check with support.
- 5) Always update your Mac OS X using a combo update (as oppose to automatic updates). This will involve repairing permissions on the system disk. Keep in mind that you may have to re-install MetaSAN afterwards (simply running the MetaSAN installer and choosing the update option).
- 6) Disable the MAC OSX UpdateDB weekly cron job on all Mac machines. The UpdateDB has been known to cause issues on both local and SAN based topologies. Please contact our support department for further information on this topic. Spotlight should be used instead.
- 7) It is recommended to disable Spotlight on SAN volumes because it can reduce performance dramatically while indexing contents. Keep in mind that disabling Spotlight for a SAN volume also disables searching on this volume.