ADVISORY REGARDING SAS® GRID MANAGER WITH ISILON

ABSTRACT
This document outlines the best practices regarding SAS® Grid Manager with Isilon Storage
November 2015
## CONTENTS

Overview ..................................................................................................................................... 4  
Version Specificity ..................................................................................................................... 4  
SAS® Grid Manager ................................................................................................................... 4  
EMC Isilon ................................................................................................................................ 4  
SAS® Grid Manager on Isilon ...................................................................................................... 4  
General Considerations ................................................................................................................. 5  
Fitness and Scope ..................................................................................................................... 5  
Tuning Suggestions .................................................................................................................... 5  
Operating System ..................................................................................................................... 5  
NFS Mount Concepts .................................................................................................................. 5  
  Mount Options for Shared Data ............................................................................................... 6  
  Mount Options for Non-Shared Data ....................................................................................... 6  
  SetUID ..................................................................................................................................... 6  
SAS® on Isilon .......................................................................................................................... 7  
  32Bit File ID’s for SAS 9.3 ...................................................................................................... 7  
NFS Exports ............................................................................................................................ 7  
  Network Interface Tuning for Isilon 10gb .................................................................................. 7  
Load Balancing ........................................................................................................................ 7  
Data Access Pattern .................................................................................................................. 7  
SSD Strategy ............................................................................................................................ 8  
  Data Protection Modes ............................................................................................................ 8  
Conclusion ................................................................................................................................... 8  
Additional Resources ................................................................................................................. 8  
Contact Information .................................................................................................................... 9
OVERVIEW

VERSION SPECIFICITY

The information represented in this document pertains to the following product versions:

- SAS® 9 – all products including SAS® Grid Manager.
- EMC Isilon OneFS 7.2.0.3 or higher (Please do not use a version of OneFS prior to 7.2.0.3)

SAS® GRID MANAGER

This document is focused on SAS Grid Manager, given that it is more popular for the sort of Big Data Analytics environments in which Isilon’s feature set is of the greatest interest. Some references may be made to ‘non-Grid’ SAS environments to distinguish them from SAS Grid Manager environments where the distinction is pertinent. SAS provides scalability through parallel processing and the ability to manage, access and process data in a distributed environment. With its ability to work with scalable procedures and I/O engines, it gives applications unmatched potential to scale up in SMP environments and scale out on the network at the same time. Because SAS is so analytically powerful, many SAS applications tend to be very data and/or compute intensive. As a result, the performance of these SAS applications may be improved by running in a Grid Manager environment.

EMC ISILON

EMC Isilon’s network-attached storage (NAS) systems deliver breakthrough ease-of-scale, ease-of-administration, and scalable NAS performance using a clustered architecture. Powered by Isilon’s OneFS™ Operating system, Isilon clusters provide multi-protocol access to large volumes of data over NFS, SMB, HTTP, FTP, and HDFS protocols. OneFS™ was originally developed to meet market demands for easy-to-deploy scale-out capacity, performance, and operational simplicity in “Big Data” markets including Media and Entertainment, Healthcare, Life Sciences, High Performance Computing (HPC), and the burgeoning Data Analytics market. Continuous refinements to Isilon’s hardware and OneFS software continue to extend the reach of OneFS into numerous new markets, enabled by a growing set of Enterprise Data Management features and wide compatibility with leading Hadoop distributions.

SAS® GRID MANAGER ON ISILON

From a pure performance perspective, NFS is not necessarily the highest-performing option for SAS applications. While many SAS and SAS Grid Manager applications can benefit enormously from Isilon’s streaming sequential performance and scaling (core strengths of the Isilon architecture), the I/O demands and performance requirements of various SAS-based applications can be quite diverse.

Obtaining the best possible results with SAS products on EMC Isilon NAS storage requires some consideration at every layer between SAS and the storage, including client system tuning and configuration factors, network infrastructure factors, and OneFS-specific performance options. This document summarizes known SAS-on-Isilon Best Practices and Guidelines and provides links to other documents for details on many specific areas.
GENERAL CONSIDERATIONS

FITNESS AND SCOPE

Storage utilized by SAS can be broadly divided into two categories; Shared and Non-Shared:

- **Shared** generally refers to data areas that are shared between clients on a shared filesystem such as SASDATA.
- **Non-Shared** generally refers to data areas that are exclusive to, and not shared between, individual clients such as SASWORK and UTILLOC.
- SAS Server Logs (metadata, object spawner, etc..) should be placed on local filesystems if the SAS install resides on OneFS.

Shared and Non-Shared Data each have very different performance requirements and constraints. It is most common to host these two categories of SAS storage on different storage solutions which match the respective performance requirements of each. (For example, EMC’s XtremIO, VMAX, VNX, and ScaleIO are all highly-favored for the most-demanding Non-Shared performance requirements.) It is SAS preference Non-Shared workloads reside on high speed block storage.

TUNING SUGGESTIONS

OPERATING SYSTEM

Operating Systems settings are distributed to meet a wide and varied range of environments. However it is sometimes necessary to deviate from the standard OS settings to obtain the best performance possible under a given workload. Complete tuning guidelines can be found in the Additional Resources section of this advisory. As always it is a best practice to test these parameters and consult solution vendors on potential use before deploying into a production environment.

NFS MOUNT CONCEPTS

As discussed previously there are two sets of recommendations around NFS, with respect to NFS3, related to locking behaviors and SAS Grid Manager. As NFSv3 and NFSv4 have different locking behaviors it is important to use one version of the protocol for consistency.

Non-Shared data is only mounted to a single host and is not shared between hosts. As such coordinated file locking between SAS client systems is not required. The NFS mount option to use 'local' or 'nolock' reduces unnecessary overhead for Non-Shared data. Shared data, such as with permanent SAS data files, requires NFS-based locking in order for SAS to coordinate data access and prevent data corruption.

For non-SAS-Grid deployments, local locks will improve SAS performance. **One must assure that shared locks are not required** to coordinate file access with other applications which may be accessing the same data from other NFS clients. Failure to do so may result in data corruption.

The next section contains a short summary of the main NFS mount option considerations for SAS. For a definitive discussion of the options available on your specific NFS client platform, please consult your version-specific documentation such as the Unix/Linux nfs(5) manual pages.
MOUNT OPTIONS FOR SHARED DATA

For Shared data areas:

- acdirmin=0,acdirmax=0 - Preferred options to accommodate the needs of SAS products.
- Alternatives such as noac and actimeo=0 have significantly-higher negative performance consequences and are therefore not recommended.
- noatime
- nodiratime
- rdirplus - improves ‘ls -l’ performance. This is normally the default, but re-stating it should not be a problem.
- vers=3

MOUNT OPTIONS FOR NON-SHARED DATA

For Non-Shared data areas that are only mounted to a single host:

- nolock - Enables client-local lock management for greatly-improved performance. Inasmuch as more modern versions of SAS do not do file-locking operations in Non-Shared areas, one may see no benefit to this setting with specific workloads. However, since Non-Shared areas do not require inter-client coordination, it remains as a Best Practice to use this setting for Non-Shared storage areas.
- noatime
- nodiratime
- actimeo=86400 - That’s one day in seconds, but any reasonably-large value should suffice
- rdirplus - Improves ‘ls -l’ performance. This is normally the default, but re-stating it should not be a problem.
- nocto - This option may provide some performance benefit, but may also result in cascaded performance issues with memory management. Its utility must be evaluated on a case-by-case basis by empirical testing.
- Make sure to explicitly define the host IP or Name that can mount, and have root access, for each export to reduce the risk another host can mount the Non-Shared data.
- vers=3

SETUID

Some SAS executables use ‘setuid root’. Information on these files and what they do can be found in this linked document. If these files are accessed via NFS, the following constraints apply;

- The client’s mount point must not use the ‘noexec’ or ‘nosuid’ mount options.
- The server’s export must not suppress the root user’s identity. (i.e, On Isilon do not map ‘root’ to ‘nobody’. )
SAS® ON ISILON

Isilon offers SAS interesting feature sets uncommon in many other deployments. One main benefit of Isilon is ease and simplicity to scale the solution according to demand. It is generally much easier to increase the size of network attached exports than it is to manage LUNs on a typical array. Isilon also offers some very beneficial options around data protection and file layout documented further below. Several important settings for optimal configuration are listed below by category.

32BIT FILE ID’S FOR SAS 9.3

When using SAS 9.3 or earlier the filesystems needs to support 32-bit file ID’s. Isilon supports 32bit file ID’s but they are disabled by default. To enable 32bit File ID’s:

    isi nfs exports modify <file system ID number> —return-32bit-file-ids=yes

NFS EXPORTS

Since most NFS performance factors are associated with client mounts and client mount options, one could use a single NFS export to facilitate all of the mounts associated with a given SAS landscape. However, it is recommended to segregate Shared and Non-Shared Data exports to enable changing their server-side performance and security factors independently. Each grid server should have a unique set of exports for Non-Shared data.

NETWORK INTERFACE TUNING FOR ISILON 10GB

<table>
<thead>
<tr>
<th>sysctl OID</th>
<th>Recommended</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>kern.ipc.maxsockbuf</td>
<td>16777216</td>
<td>2097152</td>
</tr>
<tr>
<td>net.inet.tcp.recvspace</td>
<td>524288</td>
<td>131072</td>
</tr>
<tr>
<td>net.inet.tcp.recvbuf_inc</td>
<td>32768</td>
<td>16384</td>
</tr>
<tr>
<td>net.inet.tcp.recvbuf_max</td>
<td>2097152</td>
<td>262144</td>
</tr>
<tr>
<td>net.inet.tcp.sendspace</td>
<td>524288</td>
<td>131072</td>
</tr>
<tr>
<td>net.inet.tcp.sendbuf_inc</td>
<td>16384</td>
<td>8192</td>
</tr>
<tr>
<td>net.inet.tcp.sendbuf_max</td>
<td>2097152</td>
<td>262144</td>
</tr>
</tbody>
</table>

LOAD BALANCING

Isilon operates a DNS responder-based load balancer to distribute traffic across the Isilon cluster nodes. This feature is known as Smart Connect Advanced. It is recommended to use SmartConnect Advanced and Dynamic Access Zones for NFS. Dynamic Smart Connect Zones are required for IP failover functionality. Support for NFSv4 graceful failover arrives with OneFS 8.0.

DATA ACCESS PATTERN

Isilon can optimize disk layout and cache algorithms based on expected access patterns. Automation of these parameters is available through SmartPools. By Default, and without SmartPools, there is a pool default access pattern. Administrators manually denote the access pattern of specific files or directories. SmartPools enabled both file policies as well as a policy engine. These policies are defined at the file or folder level and help determine the access pattern of any files written to a location among various other options. For Shared data paths the recommended I/O optimization is “streaming” since the expected access pattern is large block contiguous I/O. The streaming I/O optimization increases the number of disks incorporated into file layout and enables a
more aggressive cache pre-fetch algorithm. For Non-Shared data paths the recommended I/O optimization is “concurrency” as these regions exhibit both random and sequential I/O patterns.

SSD STRATEGY

Isilon can alter the behaviors around metadata handling to improve performance. By default, the Isilon uses an SSD Strategy known as L3 and this works well for most use cases. Alternatives are Metadata Read Acceleration and Metadata Read and Write acceleration. Whereas Metadata Read acceleration writes an extra mirror of the metadata to SSD, Metadata Read Write Acceleration makes every attempt to write all metadata onto SSDs. This accelerates metadata operations for both reads and writes but also uses much more SSD capacity. If Isilon is on a dedicated cluster, for use by SAS Grid Manager, then use Metadata Read and Write acceleration. Otherwise evaluate all workloads on the Isilon cluster before considering a change so as to strike the right balance across workloads. Metadata Read, implemented slightly differently in L3, is incorporated into L3 SSD Strategy. Metadata Read and Write acceleration is not available with the L3 SSD Strategy.

DATA PROTECTION MODES

The Default Data Protection Policy in OneFS is 2:1. This means, for any one file, Isilon sustains two disk failures or one node failure and maintains service for that files, and others that cross the fault domain. Data Protection Policies are applied per file, directory, or pool. Alternatively using SmartPools, mentioned above, Data Protection Policy can be altered automatically based on file policy.

Hybrid Data Protection Policies (such as 2:1) leverage Forward Error Correction algorithms to protect data. A recommendation for SAS Non-Shared (SASWORK/UTILLOC) workloads is to use one of the mirroring protection policies. While these policies incur higher capacity overhead they are significantly less taxing on CPU for high write workloads. An example of manually setting 2x protection policy:

```
isi set -p 2x <isilon path to SASWORK folder>/*
```

For Shared data the default protection policy, or any other hybrid protection policy, are applicable so that capacity is optimized. Employing this approach balances the needs for fault tolerance and performance with respect to capacity.

CONCLUSION

As a reminder, heavy WRITE intensive workloads, such as exist in Non-Shared (SASWORK, UTILLOC, Metadata Logs), should receive careful consideration before using NFS Storage for performance reasons. In addition, the SAS Metadata and SAS Object Spawner logs should be placed on local filesystems if the SAS install resides on OneFS.

It is crucial to work with your Isilon engineer to plan, install, and tune your host utilizing the cluster to achieve maximum performance. The guidelines listed in this paper are beneficial and recommended. Your individual experience may require additional guidance by Isilon depending on your host system, and workload characteristics.

ADDITIONAL RESOURCES

These following references make occasional mention of NFS and NAS, and provide some background information about SAS I/O and performance in general. For operating SAS on Isilon NFS storage, this document is intended to supersede all others. Additional references will be added here as they become available. The following SAS-developed documents should be reviewed independently for additional background and insight regarding SAS I/O requirements;

Best Practices for Configuring IO for SAS

FAQ for Storage Configuration

For Operating System tuning guidelines, please go to http://support.sas.com/kb/53/873.html
CONTACT INFORMATION

B. Scott Cassell  
EMC Corporation  
176 South Street  
Hopkinton, MA  01748  
508-435-1000  
BScott.Cassell@emc.com

Margaret Crevar  
SAS Institute Inc.  
100 SAS Campus Dr  
Cary, NC 27513-8617 United States  
919-531-7095  
margaret.crevar@sas.com

Special thanks to Tony Brown from SAS, Brian Radwanski from EMC, and Bob Sneed from EMC for their review and input.