SAS® Visual Analytics: Distributed Environment Deployment Prerequisites

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2013. *SAS® Visual Analytics: Distributed Environment Deployment Prerequisites*, Cary, NC: SAS Institute Inc.

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

June 2013

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at **support.sas.com/bookstore** or call 1-800-727-3228.

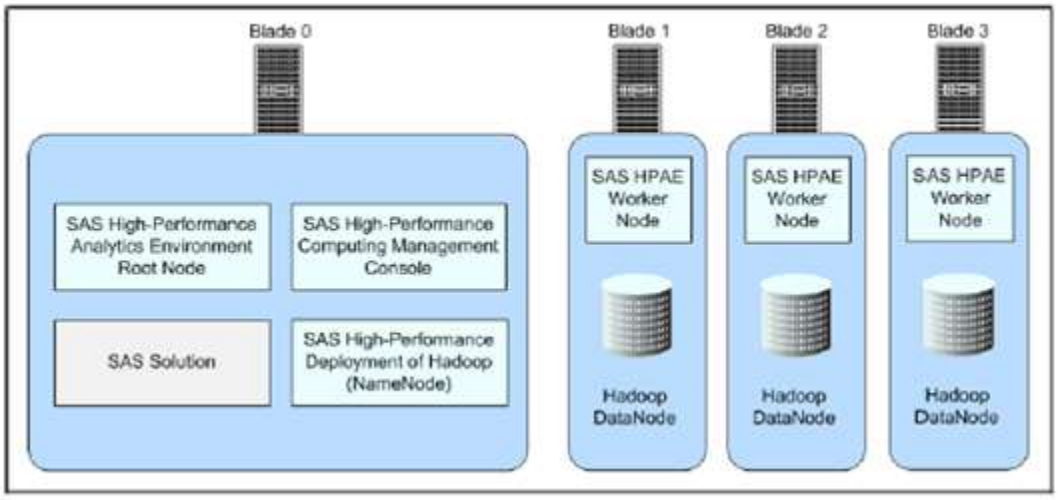
Table of Contents

Chapter 1 — Overview

SAS Visual Analytics can be installed in two modes: distributed and non-distributed. Non-distributed mode resembles the deployment used by most other SAS offerings, with the required software being centralized. Non-distributed mode requires no more documentation than those other offerings. Distributed mode, on the other hand, describes a deployment that distributes the work of the SAS High-Performance Analytics Environment and Hadoop across several machines. This document details the specific operating system and SAS software setup prior to be completed prior to deployment in order for SAS Visual Analytics to function properly in distributed mode.

A distributed SAS Visual Analytics hardware configuration consists of the SAS node, the SAS High-Performance Analytics Environment root node, the SAS High-Performance Analytics Environment worker nodes, the Hadoop NameNode, and the Hadoop DataNodes. Although it is not the only option, it is recommended that the SAS node, the SAS High-Performance Analytics Environment root node, and the Hadoop NameNode be placed on the first blade.

Note: *When the SAS node, SAS High-Performance Analytics Environment root node, and Hadoop NameNode are placed on the same machine, that machine is commonly referred to as the* SAS server node*. The combination of the SAS High-Performance Analytics engine worker node and Hadoop DataNode is referred to as the* SAS worker node*. This document uses these terms with these definitions in mind.*



Nomenclature

Many of the sections in this document are broken into two steps: **Validate** and **Perform**. The steps are signified by the arrow bullet as seen below:

- **Validate**
- **Perform**

The **Validate** step indicates that the customer should run the associated command or commands in order to validate a setting by comparing the output returned to the sample output provided in the document.

Note: *Not all output differences indicate that a setting is incorrect. For example, differences in dates are expected. Differences in paths, names, and versions should be validated with a SAS representative.*

If a **Validate** step is successful, the output should be copied to the Actual Output field in this document for SAS to review. If a **Validate** step is successful, the **Perform** step can be skipped.

If a **Validate** step is unsuccessful, the customer should follow the instructions in the accompanying **Perform** step. Upon completion of the **Perform** step instructions, the customer should retry the **Validate** step.

Code

All code provided by SAS is "AS IS" without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. By using this code, the customer acknowledges and agrees that SAS shall not be liable for any damages whatsoever arising out of their use of this material. In addition, SAS will provide no support for the materials contained herein. In addition, the customer is responsible for applying any necessary security patches and specific customizations to meet the end customer needs.

Chapter 2 — Core Requirements

SAS Software Depot

The customer must use electronic software delivery to download the software over the Internet and make the SAS Software Depot available to the SAS Visual Analytics environment. For instructions on acquiring software via electronic software delivery, refer to the *QuickStart Guide to Your SAS® 9.4 Planned Deployments using Electronic Software Delivery* at

http://support.sas.com/qsgesdplan94.html

The SAS Download Manager is required to access the SAS Electronic Software Delivery system. The SAS Download Manager can be found at

http://support.sas.com/downloadmgr

The downloaded SAS Software Depot should be made available to the SAS node and to the desktop used to perform the install.

SAS Software Order Information

A software order has been generated for the customer, and a Software Order E-mail was sent to the site's SAS Representative. The Software Order E-mail contains important information about obtaining SAS software and instructions for preparing to install it. The SAS Installation Data file, which contains the license required to run the SAS software, is included with the software itself, not the e-mail.

The following information is also contained in the Software Order E-mail. Record the following:

Site Number	
SAS Installation Key	
Order Number	

Chapter 3 — File Systems

This section provides a recommended structure for the file system. The customer does not have to follow this configuration verbatim; instead, it is provided as a starting point. If the customer has a preferred file system structure, it can be used in place of the configuration provided below. Customer-preferred file system structures should be shared with the SAS team to ensure there are no issues.

Note: *The file system on which the SASHOME will be located (typically,* `/opt/`*) should not be mounted with the* `nosuid` *option.*

SAS Server Node

The following volumes will be created:

Partition	Size
SWAP	32768 MB
/boot	500 MB
pv.01	Rest of disk

The Physical Volume pv has one logical volume defined where the following partitions are created:

Mount Point	Size
`/`	21 GB
`/opt`	60 GB
`/SASWORK`	Rest of disk
`/tmp`	8 GB
`/opt/sas/hadoop`	8 GB

SAS Worker Nodes

The following volumes will be created:

Partition	Size
SWAP	32768 MB
/boot	500 MB
pv.01	Rest of disk

The Physical Volume pv has one logical volume defined where the following partitions are created:

Mount Point	Size
/	21 GB
/opt	6 GB
/tmp	8 GB
/opt/sas/hadoop	Rest of disk

Validate

Use the following command for all blades and add it to the Actual row in the table below.

```
df -h ; mount
```

Note: *The file system on which the SASHOME will be located (typically, /opt/) should not be mounted with the* nosuid *option.*

Host	Source	Output
SAS Server Node	Sample	Filesystem Size Used Avail Use% Mounted on /dev/mapper/vg00-rootvol 21G 3G 18G 16% / tmpfs 48G 0 48G 0% /dev/shm /dev/sda1 500M 37M 463M 7% /boot /dev/mapper/vg00-opt 60G 49G 7.7G 87% /opt /dev/mapper/vg00-tmpvol 8G 1G 7G 12% /tmp /dev/mapper/vg00-saswork 900G 1G 899G 0% /saswork /dev/mapper/vg00-hadoop 8G 146M 7.9G 2% /opt/sas/hadoop /dev/mapper/vg00-rootvol on / type ext4 (rw) proc on /proc type proc (rw) sysfs on /sys type sysfs (rw) devpts on /dev/pts type devpts (rw,gid=5,mode=620) tmpfs on /dev/shm type tmpfs (rw,size=8G) /dev/sda1 on /boot type ext4 (rw) /dev/mapper/vg00-hadoop on /opt/sas/hadoop type ext4 (rw) /dev/mapper/vg00-tmpvol on /tmp type ext4 (rw) /dev/mapper/vg00-saswork on /saswork type ext4 (rw) /dev/mapper/vg00-opt on /opt type ext4 (rw) none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw) sunrpc on /var/lib/nfs/rpc_pipefs type rpc_pipefs (rw)
	Actual	

<table>
<tr><td>SAS Worker Nodes</td><td>Sample</td><td>
<pre>
Filesystem   Size Used Avail Use% Mounted on
/dev/mapper/vg00-rootvol
                      21G   3G   18G   16%   /
tmpfs   48G   0   48G   0%   /dev/shm
/dev/sda1   500M 37M   463M   7%   /boot
/dev/mapper/vg00-opt 60G   49G   7.7G   87%   /opt
/dev/mapper/vg00-tmpvol
                      8G   1G   7G   12%   /tmp
/dev/mapper/vg00-hadoop
                      900G 1G   899G   0%   /opt/sas/hadoop

/dev/mapper/vg00-rootvol on / type ext4 (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
devpts on /dev/pts type devpts (rw,gid=5,mode=620)
tmpfs on /dev/shm type tmpfs (rw,size=8G)
/dev/sda1 on /boot type ext4 (rw)
/dev/mapper/vg00-hadoop on /opt/sas/hadoop type ext4 (rw)
/dev/mapper/vg00-tmpvol on /tmp type ext4 (rw)
/dev/mapper/vg00-opt on /opt type ext4 (rw)
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)
sunrpc on /var/lib/nfs/rpc_pipefs type rpc_pipefs (rw)
</pre>
</td></tr>
<tr><td></td><td>Actual</td><td></td></tr>
</table>

Chapter 4 — Operating System

`/etc/gridhosts` *File*

SAS LASR Analytics Server, SAS Visual Analytics Hadoop, and SAS High-Performance Management Console require a file, `/etc/gridhosts`, which identifies all hosts in the configuration. Check to see if the `/etc/gridhosts` file exists on the SAS server node.

➢ **Validate**

Determine if `/etc/gridhosts` exists by using the following command. If `/etc/gridhosts` exists, list the contents in the following table.

```
cat /etc/gridhosts
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 sasva001 sasva002 sasva003	

`/etc/gridhosts` should be owned by root with the permissions 644. Issue the following command in order to verify permissions are correct:

```
ls -al /etc/gridhosts
```

Host	Sample Output	Actual Output
SAS Server Node	-rw-r--r--. 1 root root 48 Dec 2 20:22 /etc/gridhosts	

➢ **Perform**

As root, create the `/etc/gridhosts` file and enter the host name for each blade on a separate line using the command below. The LASR root node must be the first in the list. Ensure there are no leading blanks, no trailing blanks, and no blank lines.

```
vi /etc/gridhosts
```

After all nodes have been entered, save the file.

SELinux

By default, Red Hat Enterprise Linux has a security subsystem called SELinux, which is enabled. SELinux is not compatible with SAS Visual Analytics when it is being installed, so SELinux must be made permissive on all nodes. Setting SELinux to “Permissive” is enough. SELinux can also be set to "Disabled".

Note: *It is possible to re-enable SELinux post-deployment. The instructions to do so are available in the “Deploying on SELinux and IPTables” appendix of the* SAS High-Performance Analytics Infrastructure: Installation and Configuration Guide.

➤ **Validate**

Issue the following command to display the current setting for SELinux on all nodes:

```
hostname;getenforce
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 Permissive (or Disabled)	
SAS Worker Node 1	sasva001 Permissive (or Disabled)	
SAS Worker Node 2	sasva002 Permissive (or Disabled)	
SAS Worker Node 3	sasva003 Permissive (or Disabled)	

➤ **Perform**

To disable SELinux, use vi to edit `/etc/sysconfig/selinux` on all nodes with the command below. Change the setting of enforcing to permissive. **It is very important that this change is made very carefully. Failure to modify exactly as specified may result in an inability to connect to the machine via SSH and prevent the system from booting correctly.**

```
vi /etc/sysconfig/selinux
```

Change

```
SELINUX=enforcing
```

to

```
SELINUX=permissive
```

In order to validate these settings, a reboot is required. Execute the following command successively on all hosts:

```
shutdown -r now
```

Root Requirements

Many of the prerequisites of SAS Visual Analytics requires root access. Commands in the following sections will need to be run as root unless otherwise specified.

Password-less SSH for Root

SAS Server Node in the SAS Visual Analytics environment contains the SAS tier, LASR root node, and the Hadoop name node. The root account is recommended to have password-less access to the remaining nodes to speed up the setup of the environment.

➤ **Validate**

As the root user, issue the following command to verify that password-less SSH is configured between all nodes:

```
for hst in `cat /etc/gridhosts`; do ssh $hst hostname ; done
```

If this is the first attempt to use SSH from the first node to the other nodes, you will likely receive a fingerprint message, to which you should respond "yes". This does not constitute a failure of the **Validate** step.

If, however, you are prompted to enter a password as in the bolded text in the example below, then password-less SSH has not been configured properly and you should go to the **Perform** steps.

```
[root@sasva000 ~]# for hst in `cat /etc/gridhosts`; do ssh $hst hostname ; done
The authenticity of host 'sasva000 (10.96.8.96)' can't be established.
RSA key fingerprint is 81:a8:a5:19:e5:5f:5c:22:47:b3:87:85:46:2b:10:bc.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'sasva000,10.96.8.96' (RSA) to the list of known
hosts.
root@sasva000's password:
sasva000
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 sasva001 sasva002 sasva003	

➤ **Perform**

To create the initial SSH key files for root, use the following command:

```
ssh-keygen -q -t rsa -N '' -f /root/.ssh/id_rsa
```

Note: *The argument following the* `-N` *is two single quotes with nothing between them.*

Add the public key for root to the authorized_keys file for root with the following command:

```
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

Use the following command as root to copy the authorized keys file to all blades in the environment:

```
for hst in `cat /etc/gridhosts`; do scp -r ~/.ssh/ $hst:~; done
```

You will be prompted to accept the server public key and provide the password for the root account on each blade.

Operating System Version

SAS Visual Analytics requires Red Hat Enterprise Linux 6 (64-bit version) or SLES 11 SP2 and higher on the x64 chip family.

➤ **Validate**

Confirm the exact version of the operating system using the command:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname; cat
/etc/redhat-release; uname -m "; done
```

or

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname; cat /etc/SUSE-
release"; done
```

Host	SAS Server Node

Sample Output	`sasva000`
	`Red Hat Enterprise Linux Server release 6.4 (Santiago)`
	`x86_64`
	`sasva001`
	`Red Hat Enterprise Linux Server release 6.4 (Santiago)`
	`x86_64`
	`sasva002`
	`Red Hat Enterprise Linux Server release 6.4 (Santiago)`
	`x86_64`
	`sasva003`
	`Red Hat Enterprise Linux Server release 6.4 (Santiago)`
	`x86_64`
	OR
	`sasva000`
	`SUSE Linux Enterprise Server 11 (x86_64)`
	`VERSION = 11`
	`PATCHLEVEL = 2`
	`sasva001`
	`SUSE Linux Enterprise Server 11 (x86_64)`
	`VERSION = 11`
	`PATCHLEVEL = 2`
	`sasva002`
	`SUSE Linux Enterprise Server 11 (x86_64)`
	`VERSION = 11`
	`PATCHLEVEL = 2`
	`sasva003`
	`SUSE Linux Enterprise Server 11 (x86_64)`
	`VERSION = 11`
	`PATCHLEVEL = 2`

Actual Output	

Note: *The operating system must be exactly the same on all hosts in the environment.*

Groups

To deploy SAS on UNIX, an operating system group should be created and assigned as the primary group for the UNIX SAS Installer user. Members of this group are given access to certain directories and files created by the SAS Deployment Wizard. The SAS Spawned Servers account should also be made a member of the `sas` group. The GID should be same across all nodes.

Description	**Group Name**	**Example GID**	**Members**
The `sas` group is used to access the software and files necessary to install and run SAS Visual Analytics.	`sas`	502	`sas`, `sassrv`

The `hadoop` group is the primary group for the hadoop account.	`hadoop`	503	`hadoop`
The `sasusers` group is used for the other sas accounts.	`sasusers`	504	`sasdemo`, `lasradm`

➤ **Validate**

Check to see if the `sas`, `sasusers`, and `hadoop` groups exist with the following command.

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;egrep
'sas|hadoop' /etc/group | grep -v saslauth; echo"; done
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 sas:x:502: hadoop:x:503: sasusers:x:504: sasva001 sas:x:502: hadoop:x:503: sasusers:x:504: sasva002 sas:x:502: hadoop:x:503: sasusers:x:504: sasva003 sas:x:502: hadoop:x:503: sasusers:x:504:	

➤ **Perform**

Add the `sas`, `sasusers`, and `hadoop` groups on all nodes using your preferred method. Alternatively, you can use something like the following command. Before using the command, ensure the GIDs 502, 503, and 504 are free.

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname; groupadd -g 502
sas; groupadd -g 503 hadoop; groupadd -g 504 sasusers; echo"; done
```

Note: *The GID for the group must be exactly the same on all hosts in the environment.*

Users

Operating System Users

The following operating system accounts are required for the SAS Visual Analytics deployment. The UIDs should be consistent across all the nodes, as well as their primary group, home directories, and shells. The following table describes the recommended user setup:

Username	Hosts	Example UID	Shell	Example GID	Password
sas	All	502	/bin/bash	502	SASVA123
sasdemo	SAS Server Node	503	/bin/bash	504	SASVA123
sassrv	SAS Server Node	504	/bin/bash	502	SASVA123
hadoop	All	505	/bin/bash	503	SASVA123
lasradm	All	506	/bin/bash	504	SASVA123

Record the actual user setup below:

Username	Hosts	Example UID	Shell	Example GID	Password
sas					
sasdemo					
sassrv					
hadoop					
lasradm					

Note: *The UID for each user must be exactly the same on all hosts in the environment.*

➢ **Validate**

Display the defined user IDs with the following command on the SAS server node:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;egrep
'sas|lasradm|hadoop' /etc/passwd | grep -v saslauth; echo";done
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 sasdemo:x:503:504::/home/sasdemo:/bin/bash sas:x:502:502::/home/sas:/bin/bash sassrv:x:504:502::/home/sassrv:/bin/bash hadoop:x:505:503::/home/hadoop:/bin/bash lasradm:x:506:504::/home/lasradm:/bin/bash sasva001 sas:x:502:502::/home/sas:/bin/bash hadoop:x:505:503::/home/hadoop:/bin/bash lasradm:x:506:504::/home/lasradm:/bin/bash sasva002 sas:x:502:502::/home/sas:/bin/bash hadoop:x:505:503::/home/hadoop:/bin/bash lasradm:x:506:504::/home/lasradm:/bin/bash	

	sasva003 sas:x:502:502::/home/sas:/bin/bash hadoop:x:505:503::/home/hadoop:/bin/bash lasradm:x:506:504::/home/lasradm:/bin/bash	

➢ **Perform**

Use the preferred method to create the users as described earlier in the section. Alternatively, the following instructions can be used.

Issue the following commands on the SAS server node to define user IDs required for SAS Visual Analytics. Since no shell is specified for these users, the default login shell specified by the SHELL variable in `/etc/default/useradd` will be used.

```
useradd -g 504 -u 503 sasdemo
useradd -g 502 -u 504 sassrv
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname; useradd -g 502
-u 502 sas; useradd -g 504 -u 506 lasradm; useradd -g 503 -u 505
hadoop; echo"; done
```

Be sure to set the password for each user created.

Verify `umask`

On Red Hat Linux, if a user's primary group has the same name as the user (like hadoop/hadoop), the `umask` is set to 002. Otherwise, it is set to 022. This can cause issues when creating the `authorized_keys` file during the password-less SSH setup. For that reason, it is important to check that all umasks are set to 022.

➢ **Validate**

Execute the following command for each account (SAS installer, LASR admin, and Hadoop admin) on all blades:

```
umask
```

➢ **Perform**

The `umask` can be set by either modifying the content of `/etc/profile`, which will apply to all users, or appending a `umask` entry to the `~/.bashrc` file of a particular user.

SSH Keys

SSH Keys are required to enable password-less SSH for some users of the SAS Visual Analytics environment. The following users must have SSH public/private keys created and distributed to all hosts in the SAS Visual Analytics environment:

- SAS Installer (`sas`)
- LASR Admin User (`lasradm`)
- Hadoop Admin User (`hadoop`)

It is important to note that the keys created should be RSA keys, and not DSA since DSA keys can potentially lead to issues in the deployment.

➢ **Validate**

Determine if password-less SSH for these user IDs is configured. If password-less SSH exists for these accounts, you should not be prompted for passwords.

Execute the following command for each account (SAS installer, LASR admin and Hadoop admin):

```
ALLHOSTS=`cat /etc/gridhosts | tr '\n' ' '`
for hst in $ALLHOSTS; do ssh -o StrictHostKeyChecking=no $hst "
 echo '--From:'; hostname
 echo '--type of key' ; cat ~/.ssh/id_rsa | grep BEGIN
 echo '--Password-less SSH as ' \`whoami\` 'to:' ; for hst2 in
$ALLHOSTS ; do ssh -o StrictHostKeyChecking=no \$hst2 'echo
\`hostname\` ' ;done
 "; done
```

Host	Sample Output	Actual Output
SAS Server Node (sas user)	--From: sasva001 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as sas to: sasva001 sasva002 sasva003 sasva004 --From: sasva002 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as sas to: sasva001 sasva002 sasva003 sasva004 --From: sasva003 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as sas to: sasva001 sasva002 sasva003 sasva004 --From: sasva004 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as sas to: sasva001 sasva002 sasva003 sasva004	

SAS Server Node (lasradmuser)	--From: sasva001 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as lasradm to: sasva001 sasva002 sasva003 sasva004 --From: sasva002 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as lasradm to: sasva001 sasva002 sasva003 sasva004 --From: sasva003 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as lasradm to: sasva001 sasva002 sasva003 sasva004 --From: sasva004 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as lasradm to: sasva001 sasva002 sasva003 sasva004	
SAS Server Node (hadoop user)	--From: sasva001 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as hadoop to: sasva001 sasva002 sasva003 sasva004 --From: sasva002 --type of key -----BEGIN RSA PRIVATE KEY-----	

	--Password-less SSH as hadoop to: sasva001 sasva002 sasva003 sasva004 --From: sasva003 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as hadoop to: sasva001 sasva002 sasva003 sasva004 --From: sasva004 --type of key -----BEGIN RSA PRIVATE KEY----- --Password-less SSH as hadoop to: sasva001 sasva002 sasva003 sasva004	

➤ **Perform**

To generate a public/private key pair, use the following command for each account:

```
ssh-keygen -q -t rsa -N '' -f ~/.ssh/id_rsa
```

Note: *The argument following the* `-N` *is two single quotes with nothing between them.*

This command will create the private (id_rsa) and public (id_rsa.pub) keys in the `.ssh` subdirectory of the user's home. The public key should then be added to the `authorized_keys` file using the following command for each account:

```
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

Finally, the content of the user's `.ssh` directory should be copied to each host in the environment. This is performed using `scp` for each account. Log on to the first node with each account (lasradm, sas, hadoop) and execute the following command:

```
for hst in `cat /etc/gridhosts`; do scp -pr ~/.ssh $hst:~ ; done
```

Note that it is normal to be prompted for passwords at this stage.

Alternatively these accounts can be created using the SAS High-Performance Computing Management Console which will generate and populate SSH keys automatically.

SAS Internal Metadata Users

In addition to the operating system users, some additional user accounts are created and exist only within the SAS Metadata Server. These user accounts will be created by the SAS configuration process. However, a password needs to be specified for each user.

Use the following table to record the appropriate passwords:

Description	User ID	Password
SAS Administrator— The user account that has privileges associated with the SAS Metadata Unrestricted Users role.	sasadm@saspw	
SAS Trusted User— The user account that can impersonate other users on connections to the metadata server. Some SAS processes use this account to communicate with the metadata server on a client's behalf.	sastrust@saspw	
Search Interface to SAS Content User— The user account that permits access to SAS content that is supplied to SAS Information Retrieval Studio for indexing.	sassearch@saspw	
SAS Environment Manager Service Account— The user account that is used by SAS Environment Manager agents to poll SAS Metadata	sasevs@saspw	

Other Internal Users

Record the following:

Description	User ID	Password
User ID and password for administering the SAS High-Performance Computing Management Console.	hpcadmin	
User ID and password for accessing the SAS Web Infrastructure Platform Data Server. This is the owner of all databases.	dbmsowner	
User ID and password for accessing the Administration database used by the SAS Environment Manager	adminowner	
User ID and password for accessing the EVManager database used by the SAS Environment Manager	EVManager	
A valid key used for encrypting and decrypting the SAS Environment Manager EVManager database user password. The key must be at least eight characters long and can contain letters and numbers only.	Database Encryption Passphrase	
User ID and password for accessing the database used with SAS Visual Analytics Services	vatadm	
User ID and password for accessing the database used with SAS Visual Data Builder	vdbadm	
User ID and password for accessing the SharedServices database used by the SAS Web Infrastructure Platform	SharedServices	

Required OS Configuration

Host Name

➤ **Validate**

The host name for each host in the environment must be the same as the DNS name for the host. Confirm the name for every node using the following command:

```
for hst in `cat /etc/gridhosts`; do ssh $hst hostname ; done
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 sasva001 sasva002 sasva003	

Red Hat Network

➤ **Validate**

All hosts in the environment must be registered with the Red Hat Network to enable operating system updates. Use the following command to confirm each host is registered:

```
for hst in `cat /etc/gridhosts`; do ssh $hst ls -la
/etc/sysconfig/rhn/systemid; done
```

Host	Sample Output	Actual Output
SAS Server Node	-rw-------. 1 root root 1321 Mar 5 2012 /etc/sysconfig/rhn/systemid -rw-------. 1 root root 1321 Mar 5 2012 /etc/sysconfig/rhn/systemid -rw-------. 1 root root 1321 Mar 5 2012 /etc/sysconfig/rhn/systemid -rw-------. 1 root root 1321 Mar 5 2012 /etc/sysconfig/rhn/systemid	

➤ **Perform**

Configure the servers so that they are registered with the Red Hat Network.

Network Configuration

The SAS Visual Analytics environment can leverage the separation between a private and public network. The private network is used by the SAS software to communicate within the environment while the public network is used to connect to the customer's network. The private network should be defined in the `/etc/hosts` file for each blade and the corresponding host name set.

Note: *If the value returned by the host name command (such as* `sasva000`*) is on the same line in the hosts file as the loopback address (127.0.0.1), it will create issues with the SAS deployment.*

➤ **Validate**

The following table provides a concise view of the network configuration for the SAS Visual Analytics platform. Enter the primary host names and associated IP addresses.

Host	Sample Host Name	Actual Host Name	Example IP Address	Actual IP Address	Actual Secondary IP Address (optional)	Role
SAS Server Node	sasva000		10.0.0.1			SAS Node/ LASR Root/ Hadoop Name
SAS Worker Node 1	sasva001		10.0.0.2			LASR Worker Node/ Hadoop Data Node
SAS Worker Node 2	sasva002		10.0.0.3			LASR Worker Node/ Hadoop Data Node
SAS Worker Node 3	sasva003		10.0.0.4			LASR Worker Node/ Hadoop Data Node

The SAS node will host the traditional SAS 9.4 servers. This will be the first blade in the environment with a host name of sasva000.

```
ALLHOSTS=`cat /etc/gridhosts | tr '\n' ' '`
for hst in $ALLHOSTS; do ssh $hst "
 echo '------------------START------------------' ;
 echo '--The actual hostname is:'; hostname
 echo '--the contents of /etc/hosts is '; cat /etc/hosts
 echo '--the contents of /etc/gridhosts is '; cat /etc/gridhosts
 echo '--The ip returned by hostname -I is' ; hostname -I
 echo '--The hostname, as returned by hostname -A is' ; hostname -A
 echo '--This machines pings the others as:' ; for hst2 in $ALLHOSTS ;
do ping -c 1 \$hst2 | grep PING ;done
 echo '--This machines traceroute the others as:' ; for hst2 in
$ALLHOSTS ; do traceroute \$hst2 |grep traceroute ;done
 echo '--------------------END----------------' "; done
```

Host	SAS Server Node
Sample Output	Sample Output [root@sasva000 ~]# ALLHOSTS=`cat /etc/gridhosts | tr '\n' ' '` HOSTS; do ssh $hst [root@sasva000 ~]# for hst in $ALLHOSTS; do ssh $hst " > echo '------------------START------------------' ; > echo '--The actual hostname is:'; hostname > echo '--the contents of /etc/hosts is '; cat /etc/hosts

```
> echo '--the contents of /etc/gridhosts is '; cat /etc/gridhosts
> echo '--The ip returned by hostname -I is' ; hostname -I
> echo '--The hostname, as returned by hostname -A is' ; hostname -A
> echo '--The nslookup of the main IP (hostname -i) returns' ; nslookup
`hostname -i`
> echo '--This machines pings the others as:' ; for hst2 in $ALLHOSTS ; do
ping -c 1 \$hst2 | grep PING ;done
> echo '--This machines traceroute the others as:' ; for hst2 in $ALLHOSTS ;
do traceroute \$hst2 |grep traceroute ;done
> echo '--------------------END----------------' "; done
------------------START------------------
--The actual hostname is:
sasva000
--the contents of /etc/hosts is
127.0.0.1 localhost.localdomain localhost
::1 localhost6.localdomain6 localhost6
10.0.0.1 sasva000
10.0.0.2 sasva001
10.0.0.3 sasva002
10.0.0.4 sasva003
--the contents of /etc/gridhosts is
sasva000
sasva001
sasva002
sasva003
--The ip returned by hostname -I is
10.0.0.1
--The hostname, as returned by hostname -A is
sasva000
--The nslookup of the main IP (hostname -i) returns
Server: 149.173.5.54
Address: 149.173.5.54#53

237.7.96.10.in-addr.arpa name = sasva000.customer.com.

--This machines pings the others as:
PING sasva000 (10.0.0.1) 56(84) bytes of data.
PING sasva001 (10.0.0.2) 56(84) bytes of data.
PING sasva002 (10.0.0.3) 56(84) bytes of data.
PING sasva003 (10.0.0.4) 56(84) bytes of data.
--This machines traceroute the others as:
traceroute to sasva000 (10.0.0.1), 30 hops max, 60 byte packets
traceroute to sasva001 (10.0.0.2), 30 hops max, 60 byte packets
traceroute to sasva002 (10.0.0.3), 30 hops max, 60 byte packets
traceroute to sasva003 (10.0.0.4), 30 hops max, 60 byte packets
--------------------END----------------
------------------START------------------
--The actual hostname is:
sasva001
--the contents of /etc/hosts is
```

```
127.0.0.1 localhost.localdomain localhost
::1 localhost6.localdomain6 localhost6
10.0.0.1 sasva000
10.0.0.2 sasva001
10.0.0.3 sasva002
10.0.0.4 sasva003
--the contents of /etc/gridhosts is
cat: /etc/gridhosts: No such file or directory
--The ip returned by hostname -I is
10.0.0.2
--The hostname, as returned by hostname -A is
sasva001
--The nslookup of the main IP (hostname -i) returns
Server: 149.173.5.54
Address: 149.173.5.54#53

237.7.96.10.in-addr.arpa name = sasva000.customer.com.

--This machines pings the others as:
PING sasva000 (10.0.0.1) 56(84) bytes of data.
PING sasva001 (10.0.0.2) 56(84) bytes of data.
PING sasva002 (10.0.0.3) 56(84) bytes of data.
PING sasva003 (10.0.0.4) 56(84) bytes of data.
--This machines traceroute the others as:
traceroute to sasva000 (10.0.0.1), 30 hops max, 60 byte packets
traceroute to sasva001 (10.0.0.2), 30 hops max, 60 byte packets
traceroute to sasva002 (10.0.0.3), 30 hops max, 60 byte packets
traceroute to sasva003 (10.0.0.4), 30 hops max, 60 byte packets
--------------------END----------------
------------------START------------------
--The actual hostname is:
sasva002
--the contents of /etc/hosts is
127.0.0.1 localhost.localdomain localhost
::1 localhost6.localdomain6 localhost6
10.0.0.1 sasva000
10.0.0.2 sasva001
10.0.0.3 sasva002
10.0.0.4 sasva003
--the contents of /etc/gridhosts is
cat: /etc/gridhosts: No such file or directory
--The ip returned by hostname -I is
10.0.0.3
--The hostname, as returned by hostname -A is
sasva002
--The nslookup of the main IP (hostname -i) returns
Server: 149.173.5.54
Address: 149.173.5.54#53
```

	237.7.96.10.in-addr.arpa name = sasva000.customer.com. --This machines pings the others as: PING sasva000 (10.0.0.1) 56(84) bytes of data. PING sasva001 (10.0.0.2) 56(84) bytes of data. PING sasva002 (10.0.0.3) 56(84) bytes of data. PING sasva003 (10.0.0.4) 56(84) bytes of data. --This machines traceroute the others as: traceroute to sasva000 (10.0.0.1), 30 hops max, 60 byte packets traceroute to sasva001 (10.0.0.2), 30 hops max, 60 byte packets traceroute to sasva002 (10.0.0.3), 30 hops max, 60 byte packets traceroute to sasva003 (10.0.0.4), 30 hops max, 60 byte packets --------------------END---------------- ------------------START------------------ --The actual hostname is: sasva003 --the contents of /etc/hosts is 127.0.0.1 localhost.localdomain localhost ::1 localhost6.localdomain6 localhost6 10.0.0.1 sasva000 10.0.0.2 sasva001 10.0.0.3 sasva002 10.0.0.4 sasva003 --the contents of /etc/gridhosts is cat: /etc/gridhosts: No such file or directory --The ip returned by hostname -I is 10.0.0.4 --The hostname, as returned by hostname -A is sasva003 --The nslookup of the main IP (hostname -i) returns Server: 149.173.5.54 Address: 149.173.5.54#53 237.7.96.10.in-addr.arpa name = sasva000.customer.com. --This machines pings the others as: PING sasva000 (10.0.0.1) 56(84) bytes of data. PING sasva001 (10.0.0.2) 56(84) bytes of data. PING sasva002 (10.0.0.3) 56(84) bytes of data. PING sasva003 (10.0.0.4) 56(84) bytes of data. --This machines traceroute the others as: traceroute to sasva000 (10.0.0.1), 30 hops max, 60 byte packets traceroute to sasva001 (10.0.0.2), 30 hops max, 60 byte packets traceroute to sasva002 (10.0.0.3), 30 hops max, 60 byte packets traceroute to sasva003 (10.0.0.4), 30 hops max, 60 byte packets --------------------END---------------- [root@sasva000 ~]#
Actual Output	

Directory Locations

SAS Server Node

Create the following directory locations:

Component	Owner	Group	Permissions	Default Directory	Actual Directory
SAS Software Depot	sas	sas	755	/mnt/depot	
SAS Root Directory	sas	sas	775	/opt/sas	
SAS Configuration Directory	sas	sas	755	/opt/sas/config	
SAS Installation Directory	sas	sas	755	/opt/sas/sashome	
SAS Third Party Directory	sas	sas	755	/opt/sas/thirdparty	
SAS Work Directory	sas	sas	777	/saswork	
SAS High-Performance Node Installation Directory	sas	sas	755	/opt/sas/TKGrid_2.5	
SAS High-Performance Management Console Installation Directory	root	root	755	/opt/sas/hpcmc/webmin	
RDBMS Client Installation Directory				/usr/odbc	
SAS Hadoop Installation Directory	hadoop	hadoop	755	/opt/sas/hadoop/	
License location for SAS middle tier server	sas	sas	755	/etc/opt/vmware/vfabric	This location cannot be changed

SAS Worker Nodes

Create the following directory locations:

Component	Owner	Group	Permissions	Default Directory	Actual Directory
SAS High-Performance Node Installation Directory	sas	sas	755	/opt/sas/TKGrid_2.5	
SAS Hadoop Installation Directory	hadoop	hadoop	755	/opt/sas/hadoop/	
SAS Environment Manager Agent location	sas	sas	755	/opt/sas/config /opt/sas/sashome	

Required Operating System Settings

SSH

SAS Visual Analytics requires a number of settings for the SSH Daemon (SSH Server process) on all hosts in the environment. Use the root account to display and modify the SSH configuration file.

Parameter	Default	Required Value
MaxStartups	10	1000

➤ **Validate**

Use the following command to determine the number of startups:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname ; egrep
MaxStartups /etc/ssh/sshd_config" ; done
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 MaxStartups 1000 sasva001 MaxStartups 1000 sasva002 MaxStartups 1000 sasva003 MaxStartups 1000	

➤ **Perform**

Use an editor, such as vi, to edit the sshd configuration file as root. The following command will open the sshd configuration file in vi. Modify each setting as specified in the table above on the first node.

```
vi /etc/ssh/sshd_config
```

When the `sshd_config` file has been updated, use the following command to copy the first file to all other blades:

```
for hst in `cat /etc/gridhosts`; do scp /etc/ssh/sshd_config
$hst:/etc/ssh/sshd_config ; done
```

In order to validate these settings, restart the SSH daemon with the following command:.

```
for hst in `cat /etc/gridhosts`; do ssh $hst "service sshd restart";
done
```

Note: *The SSH Daemon settings must be exactly the same on all hosts in the environment.*

CPU speed

The Red Hat operating system has settings that throttle the CPU to save power. These should be disabled in order to get the most performance out of your hardware.

➢ **Validate**

In order to determine if the CPU has been throttled, use the following command:

```
for hst in `cat /etc/gridhosts`;do ssh $hst "hostname;grep GOVERNOR
/etc/sysconfig/cpuspeed "; done
```

Host	SAS Server Node
Sample Output	sasva000 ### GOVERNOR ### # - The GOVERNOR parameter is only valid on centrino, powernow-k8 (amd64) GOVERNOR=performance sasva001 ### GOVERNOR ### # - The GOVERNOR parameter is only valid on centrino, powernow-k8 (amd64) GOVERNOR=performance sasva002 ### GOVERNOR ### # - The GOVERNOR parameter is only valid on centrino, powernow-k8 (amd64) GOVERNOR=performance sasva003 ### GOVERNOR ### # - The GOVERNOR parameter is only valid on centrino, powernow-k8 (amd64) GOVERNOR=performance [root@sasva000 ~]#
Actual Output	

➢ **Perform**

In order to change the CPU speed setting, edit the following file with the following command:

```
vi /etc/sysconfig/cpuspeed
```

Revise the GOVERNOR line so that it reads as follows:

```
GOVERNOR=performance
```

Then copy the file from the first blade to all the others using the following command:

```
for hst in `cat /etc/gridhosts`; do scp /etc/sysconfig/cpuspeed
$hst:/etc/sysconfig/cpuspeed; done
```

In order to validate these settings, a reboot is required. However, the next section also requires a reboot, so the customer should decide if they want to reboot now or wait. If a reboot is desired, execute the following command:

```
for hst in `sort -r -k1 /etc/gridhosts`; do ssh $hst "shutdown -r now";
done
```

Disabling SUSE CPU Throttle

In addition, the SUSE operating system has settings that throttle the CPU to save power. These should be disabled in order to get the most performance out of your hardware.

- **Validate and Perform** (as root)
  1. Launch the interactive `yast` interface with the following command:

     ```
     yast
     ```

  2. Choose “System”.
  3. Choose “System Services (Runlevel).
  4. Disable “powersaved”.
  5. Save your changes and exit `yast`.

Limits

SAS Visual Analytics requires changes to the security limits placed on end users in order that the required number of processes and file allocations for SAS Visual Analytics function correctly. These settings must be exactly the same for all hosts in the environment.

- **Validate**

Run the following commands as the SAS installer ID in order to display the current hard and soft limits for max number of processes and open files:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;echo \"#soft
limits\" ; ulimit -n -u -s ; echo \"#hard limits\" ; ulimit -H -n -u -s
"; done
```

Host	Sample Output	Actual Output
SAS Server Node	sasva000 #soft limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240 #hard limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240 sasva001 #soft limits open files (-n) 350000 max user processes (-u) 100000	

	stack size (kbytes, -s) 10240 #hard limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240sasva002 #soft limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240 #hard limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240sasva003 #soft limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240 #hard limits open files (-n) 350000 max user processes (-u) 100000 stack size (kbytes, -s) 10240	

➢ **Perform**

Setting the limits requires editing two files on each host, using vi or any other text editor to make the changes. These changes must be made as root. First, edit the `/etc/security/limits.conf` file. The following command will open the file in vi:

```
vi /etc/security/limits.conf
```

Add the following lines to the bottom of the file just before the line "`# End of file`".

```
* hard nofile 350000
* soft nofile 350000
* hard nproc 100000
* soft nproc 100000
* hard stack 10240
* soft stack 10240
```

The second file which must be edited by root is `/etc/security/limits.d/90-nproc.conf`. The following command will open the file in vi:

```
vi /etc/security/limits.d/90-nproc.conf
```

The value of nproc needs to be changed from the default value of 1024 to 10240. After the changes, the file will contain the following:

```
# Default limit for number of user's processes to prevent
# accidental fork bombs.
# See rhbz #432903 for reasoning.

* soft nproc 100000
```

Copy the two files that were modified from the SAS node to all worker nodes using the following commands:

```
for hst in `cat /etc/gridhosts`; do scp /etc/security/limits.conf
$hst:/etc/security/limits.conf; done
for hst in `cat /etc/gridhosts`; do scp /etc/security/limits.d/90-
nproc.conf $hst:/etc/security/limits.d/90-nproc.conf; done
```

Note: *The limit settings must be exactly the same on all hosts in the environment.*

Software Firewall

It is strongly recommended not to have any firewall between the hosts in the SAS Visual Analytics environment. A firewall may be used to protect the SAS Visual Analytics environment, but there should be nothing limiting the communications between hosts within the SAS Visual Analytics environment.

Note: *It is possible to re-enable SELinux post-deployment. The instructions to do so are available in the "Deploying on SELinux and IPTables" appendix of the* SAS High-Performance Analytics Infrastructure: Installation and Configuration Guide.

➤ **Validate**

Issue the following commands as root to check if the software firewall is enabled:

```
for hst in `cat /etc/gridhosts`;do ssh $hst "hostname;/sbin/chkconfig -
-list iptables"; done
for hst in `cat /etc/gridhosts`;do ssh $hst "hostname;/sbin/chkconfig -
-list ip6tables"; done
```

Record the results in the following table:

Host	SAS Server Node
Sample Output	sasva000 iptables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva001 iptables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva002 iptables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva003 iptables 0:off 1:off 2:off 3:off 4:off 5:off 6:off
Actual Output	
Host	SAS Server Node
Sample Output	sasva000 ip6tables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva001 ip6tables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva002 ip6tables 0:off 1:off 2:off 3:off 4:off 5:off 6:off sasva003 ip6tables 0:off 1:off 2:off 3:off 4:off 5:off 6:off
Actual Output	

➤ **Perform**

In order to disable the software firewalls on each host within the SAS Visual Analytics environment, issue the following commands as root:

```
for hst in `cat /etc/gridhosts`;do ssh $hst "hostname;/sbin/chkconfig --level 0123456 iptables off; service iptables stop "; done
for hst in `cat /etc/gridhosts`;do ssh $hst "hostname;/sbin/chkconfig --level 0123456 ip6tables off ; service ip6tables stop "; done
```

Cron Scheduler

The SAS Installation account will have to be able to schedule tasks using cron in order to periodically update the search database. Moreover, it is possible that that using the AutoLoad LASR server also requires `sas` or `lasradm` to be able to use cron.

➤ **Validate**

Issue the following commands, first as the `sas` account and then as the `lasradm` account, on the SAS node.

```
su - sas
cat <(crontab -l) <(echo "1 2 3 4 5 echo Cron Test") | crontab - ;
crontab -l ; crontab -r ; crontab -l


su - lasradm
cat <(crontab -l) <(echo "1 2 3 4 5 echo Cron Test") | crontab - ;
crontab -l ; crontab -r ; crontab -l
```

Host	Sample Output	Actual Output

SAS Server Node	`no crontab for sas`
	`1 2 3 4 5 echo Cron Test`
	`no crontab for sas`
	
	`no crontab for lasradm`
	`1 2 3 4 5 echo Cron Test`
	`no crontab for lasradm`	

SMTP Server

The SAS Deployment Wizard prompts you for an SMTP server that can be used by the deployment to send e-mails.

Record the host name and port to be used:

	Sample Values	Actual Values
Host name	Mail.customer.com	
Port	25	

In order to validate that your SMTP server is present and works, the following instructions can be helpful:
http://www.vsysad.com/2013/10/testing-smtp-server-from-the-command-line/

Rebooting All Blades

SAS recommends that the user reboot all the blades after changes are made to the SSH, CPU speed, limits, and SELinux.

Use the following command to reboot all nodes in reverse order.

```
for hst in `sort -r -k1 /etc/gridhosts`; do ssh $hst "shutdown -r now";
done
```

Software Packages

The following packages are required on all blades in the environment. Some are absolutely required and some are useful in certain scenarios. SAS recommends installing all of them.

Required Packages

Libnuma

SAS Visual Analytics requires libnuma to be installed on all hosts in the environment.

➤ **Validate**

Run the following commands as root in order to confirm the package is installed:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;rpm -q numactl" ; done
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	sasva000 numactl-2.0.3-9.el6.x86_64 sasva001 numactl-2.0.3-9.el6.x86_64 sasva002 numactl-2.0.3-9.el6.x86_64 sasva003 numactl-2.0.3-9.el6.x86_64	

➤ **Perform**

If the package is not installed, run the following command as root in order to install it:

(Red Hat) `for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;yum install numactl"; done`

(SuSE) `for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;zypper install numactl"; done`

Note: *The same package version must be installed all hosts in the environment.*

Glibc

SAS Foundation 9.4 requires GLIBC 2.12.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
hostname;rpm -q glibc
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	sasva000 glibc-2.12-1.132.el6.x86_64 glibc-2.12-1.132.el6.i686	

➤ **Perform**

If the package is not installed, run the following command as root in order to install it:

(Red Hat) `yum install glibc`

(SuSE) `zypper install glibc`

Java

The SAS High-Performance Deployment of Hadoop requires the Java Runtime Environment (JRE) to be installed on all the hosts.

Note: *SAS recommends the customer use one of the JRE versions listed on* http://wiki.apache.org/hadoop/HadoopJavaVersions

Ensure that a compatible version of the JRE is installed, even if it's not the default used by the operating system. Run the following command as root in order to confirm the package is installed:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "
 hostname;
 which java;
 java -version";
done
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	sasva000 /usr/bin/java java version "1.6.0_24" OpenJDK Runtime Environment (IcedTea6 1.11.11) (rhel-1.61.1.11.11.el6_4-x86_64) OpenJDK 64-Bit Server VM (build 20.0-b12, mixed mode) sasva001 /usr/bin/java java version "1.6.0_24" OpenJDK Runtime Environment (IcedTea6 1.11.11) (rhel-1.61.1.11.11.el6_4-x86_64) OpenJDK 64-Bit Server VM (build 20.0-b12, mixed mode) sasva002 /usr/bin/java java version "1.6.0_24" OpenJDK Runtime Environment (IcedTea6 1.11.11) (rhel-1.61.1.11.11.el6_4-x86_64) OpenJDK 64-Bit Server VM (build 20.0-b12, mixed mode) sasva003 /usr/bin/java java version "1.6.0_24" OpenJDK Runtime Environment (IcedTea6 1.11.11) (rhel-1.61.1.11.11.el6_4-x86_64)	

	`OpenJDK 64-Bit Server VM (build 20.0-b12, mixed mode)`	

Python

SAS Visual Analytics requires Python with a version between 2.6 and 3.0 to be installed on the first host in the environment, in order that SAS Information Retrieval Studio can run and update the search database.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
rpm -q python
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	`python-2.6.6-20.el6.x86_64`	

➤ **Perform**

If the required version of Python is not installed, run the following command as root in order to install it:

(Red Hat) `yum install python`

(SuSE) `zypper install python`

X11 Libraries

SAS Visual Analytics requires specific 64-bit libX packages to be installed on the first host in the environment in order that the SAS Deployment Wizard can run. Additionally, if the SAS Environment Manager Agent is to be deployed on all the nodes, libX will be required to run the SAS Deployment Wizard on all the nodes. Therefore, the X libraries should be installed on all nodes.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname ; rpm -q libXext
libXp libXtst xorg-x11-xauth " ; done
```

Record the following:

Host	Sample Output	Actual Output

SAS Server Node	`sasva000`
	`libXext-1.1-3.el6.x86_64`
	`libXp-1.0.0-15.1.el6.x86_64`
	`libXtst-1.0.99.2-3.el6.x86_64`
	`xorg-x11-xauth-1.0.2-7.1.el6.x86_64`	

	sasva001 libXext-1.1-3.el6.x86_64 libXp-1.0.0-15.1.el6.x86_64 libXtst-1.0.99.2-3.el6.x86_64 xorg-x11-xauth-1.0.2-7.1.el6.x86_64 sasva002 libXext-1.1-3.el6.x86_64 libXp-1.0.0-15.1.el6.x86_64 libXtst-1.0.99.2-3.el6.x86_64 xorg-x11-xauth-1.0.2-7.1.el6.x86_64 sasva003 libXext-1.1-3.el6.x86_64 libXp-1.0.0-15.1.el6.x86_64 libXtst-1.0.99.2-3.el6.x86_64 xorg-x11-xauth-1.0.2-7.1.el6.x86_64	

➤ **Perform**

If the packages are not installed, run the following command as root in order to install them:

(Red Hat) `for hst in `cat /etc/gridhosts`; do ssh $hst "hostname ; yum -y install libXext.x86_64 libXp.x86_64 libXtst.x86_64 xorg-x11-xauth.x86_64"; done`

(SuSE) `for hst in `cat /etc/gridhosts`; do ssh $hst "hostname ; zypper install libXext.x86_64 libXp.x86_64 libXtst.x86_64 xorg-x11-xauth.x86_64"; done`

Perl-Net-SSLeay

SAS Visual Analytics requires the Perl-Net-SSLeay module to be installed on the host running the SAS High-Performance Computing Management Console.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
rpm -q perl-Net-SSLeay
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	perl-Net-SSLeay-1.36-1.el6.rfx.x86_64	

➤ **Perform**

If the Perl-Net-SSLeay module is not installed, run the following command as root in order to install it:

(Red Hat) `yum install perl-Net-SSLeay`

(SuSE) `zypper install perl-Net-SSLeay`

at

The SAS Operating System Scheduler uses the `at` command and daemon. If you plan to use this scheduler, ensure that at is installed and configured.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
which at; service atd status
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	`/usr/bin/at` `atd (pid <nnn> ) is running...`	

➤ **Perform**

If `at` is not installed, run the following commands as root in order to install it:

(Red Hat) `yum install at`

`service atd start`

(SuSE) `zypper install at`

`service atd start`

libpng

Some parts of Base SAS may require the use of the libpng library.

➤ **Validate**

Run the following command as root in order to confirm the package is installed:

```
rpm -q libpng
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	`libpng-1.2.46-1.el6_1.x86_64`	

➤ **Perform**

If libpng is not installed, run the following command as root in order to install it:

(Red Hat) `yum install libpng`

(SuSE) `zypper install libpng`

JUnit

JUnit is required for validation of some components. Without JUnit these products may operate properly but you will not be able to validate them using SAS Deployment Tester. See the third party software support page for details:
http://support.sas.com/resources/thirdpartysupport/

Description	Default Value	Actual Value
JUnit JAR file location	/opt/sas/thirdparty/junit	
Product Version	4.8.1 or greater	

Recommended Packages

General Packages

SAS recommends that your Red Hat Enterprise Linux operating system have the following packages installed on all hosts in the environment.

➤ **Validate**

Verify that these packages are on all nodes using the following command:

```
rpm -q nfs-utils.x86_64 nfs-utils-lib.x86_64 firefox compat-libstdc++-
33 libuuid libSM libXrender fontconfig libstdc++ zlib apr ksh
```

Package	Sample Output	SAS Server Node Output	SAS Worker Node Output
nfs-utils.x86_64	nfs-utils-1.2.3-7.el6_1.1.x86_64		
nfs-utils-lib.x86_64	nfs-utils-lib-1.1.5-3.el6.x86_64		
firefox	firefox-3.6.24-3.el6_1.x86_64		
compat-libstdc++-33	compat-libstdc++-33-3.2.3-69.el6.x86_64		
libuuid.i686	libuuid-2.17.2-12.el6.x86_64		
libSM.i686	libSM-1.1.0-7.1.el6.x86_64		
libXrender.i686	libXrender-0.9.5-1.el6.x86_64		
fontconfig.i686	fontconfig-2.8.0-3.el6.x86_64		
libstdc++.i686	libstdc++-4.4.5-6.el6.x86_64		
zlib.i686	zlib-1.2.3-25.el6.x86_64		
apr	apr-1.3.9-3.el6_1.2.x86_64		
ksh	ksh-20100621-16.el6.x86_64		

➢ **Perform**

Install any of the missing packages using the appropriate command:

(Red Hat) `yum install packagename`

(SuSE) `zypper install packagename`

NTP

We strongly recommend that you have an accurate and consistent time across the different servers that make up a SAS Visual Analytics environment. While failure to do so may not cause issues with the software itself, it may lead to unnecessary confusion when looking at time stamps from different machines.

➢ **Validate**

```
for hst in `cat /etc/gridhosts`; do ssh $hst "hostname;rpm -q ntp;
service ntpd status" ; done
for hst in `cat /etc/gridhosts`; do ssh $hst "date" ; done
```

Record the following:

Host	Sample Output	Actual Output
SAS Server Node	sasva000 ntp-4.2.6p5-1.el6.x86_64 ntpd (pid 1809) is running... sasva001 ntp-4.2.6p5-1.el6.x86_64 ntpd (pid 1886) is running... sasva002 ntp-4.2.6p5-1.el6.x86_64 ntpd (pid 1901) is running... sasva003 ntp-4.2.6p5-1.el6.x86_64 ntpd (pid 1882) is running... Mon Apr 14 11:49:07 EDT 2014 Mon Apr 14 11:49:07 EDT 2014 Mon Apr 14 11:49:07 EDT 2014 Mon Apr 14 11:49:07 EDT 2014	

SAS Software Depot

The SAS Software depot must be available to the SAS Server Node of the environment. It can either be mounted or copied onto it.

Note: *When transferring the depot using FTP or SCP, binary mode is required.*

The Software Depot should be owned by the SAS Installation account (`sas`).

➢ **Validate**

The following commands are run as root to confirm the depot is available and belongs to the appropriate account:

```
ls -al /mnt/depot/
```

Record the following:

Host	SAS Server Node
Sample Output	total 820 drwxr-xr-x. 14 sas sas 4096 Jan 21 11:40 . drwxr-xr-x. 4 root root 4096 Jan 21 11:36 .. -rwxrwxr-x. 1 sas sas 12 Jan 21 11:37 cd.id -rwxrwxr-x. 1 sas sas 1192 Jan 21 11:37 depotsummary.html drwxrwxr-x. 3 sas sas 4096 Jan 21 11:37 hot_fix drwxrwxr-x. 3 sas sas 4096 Jan 21 11:37 install_doc -rwxrwxr-x. 1 sas sas 124882 Jan 21 11:37 linemode.rexx drwxrwxr-x. 2 sas sas 4096 Jan 21 11:37 media_data drwxrwxr-x. 3 sas sas 4096 Jan 21 11:37 order_data drwxrwxr-x. 2 sas sas 4096 Jan 21 11:37 plan_files drwxrwxr-x. 621 sas sas 40960 Jan 21 11:37 product_data drwxrwxr-x. 606 sas sas 40960 Jan 21 11:40 products -rwxrwxr-x. 1 sas sas 7 Jan 21 11:40 sassd.txt -rwxrwxr-x. 1 sas sas 800 Jan 21 11:40 setup.dat -rwxrwxr-x. 1 sas sas 501136 Jan 21 11:40 setup.exe -rwxrwxr-x. 1 sas sas 35115 Jan 21 11:40 setup.rexx
Actual Output	

➢ **Perform**

Choose your preferred method to obtain the desired result.

Chapter 5 — Network

DNS

Successful resolution of both fully qualified and short names is crucial for the successful deployment of SAS Visual Analytics. Names need to be resolvable by all hosts within the SAS Visual Analytics environment. All hosts within the SAS Visual Analytics environment need to reside in the same DNS domain and sub domain.

DNS Alias

In order to present a friendlier name to users accessing the SAS Visual Analytics environment, it is possible to create a DNS alias. If the customer decides to create an alias, it must be resolvable by all desktops and iPads capable of accessing the environment. The SAS Node of the environment must also be able to resolve the alias.

This alias should resolve to the SAS Node of the environment.

Description	Default Value	Actual Value
Alias Names	`sasva.<company name>.com` `sasva`	
IP	192.168.0.1	

Ports

The following list indicates what ports are used for SAS by default, and gives you a place to enter the port numbers that you will actually use. SAS recommends that you document each SAS port that you reserve in the standard location on each machine, `/etc/services`. This practice will help to avoid port conflicts on the affected machines.

Port Description	Default Port Number	Port Type	Actual Port Number	Traffic Direction
SAS Environment Manager Agent	2143-2153	TCP		Inbound
SAS Deployment Agent	5660-5669	TCP		Both
Event Broker Admin Port	6050 - 6059	TCP		Inbound
SAS Environment Manager Dashboard Port (HTTP)	7079-7088	TCP		Both
IP Multicast UDP Scheduler Main Channel Port	7450 - 7459	UDP		Both
IP Multicast UDP Scheduler Main Channel Port	7460 - 7469	UDP		Both
SAS/CONNECT Server and Spawner	7550 - 7559	TCP		Both
Event Broker HTTP Port	8110 - 8119	TCP		Inbound

Operating System Services scheduler	8450 - 8459	TCP		Inbound
SAS Metadata Server	8560 - 8569	TCP		Inbound
SAS Object Spawner - Operator Port	8580 - 8589	TCP		Inbound
SAS Object Spawner: pooled workspace server port bank 1	8800 - 8809	TCP		Inbound
SAS Object Spawner: pooled workspace server port bank 2	8810 - 8819	TCP		Inbound
SAS Object Spawner: pooled workspace server port bank 3	8820 - 8829	TCP		Inbound
SAS Workspace Server	8590 - 8599	TCP		Inbound
SAS Stored Process Server: Bridge connection	8600 - 8609	TCP		Inbound
SAS Stored Process Server: load balancing connection 1 (MultiBridge)	8610 - 8619	TCP		Inbound
SAS Stored Process Server: load balancing connection 2 (MultiBridge)	8620 - 8629	TCP		Inbound
SAS Stored Process Server: load balancing connection 3 (MultiBridge)	8630 - 8639	TCP		Inbound
SAS Pooled Workspace Server	8700 - 8709	TCP		Inbound
SAS Deployment Tester - Server	10021 - 10029	TCP		Inbound
SAS Information Retrieval Studio - Server Admin Port	10650 - 10659	TCP		Both
SAS Information Retrieval Studio - Proxy Server Port	10660 - 10669	TCP		Both
SAS Information Retrieval Studio - Proxy Server Admin Port	10670 - 10679	TCP		Both
SAS Information Retrieval Studio - Proxy Server Web Admin Port	10680 - 10689	TCP		Both
SAS Information Retrieval Studio - Pipeline Server Port	10690 - 10699	TCP		Both
SAS Information Retrieval Studio - Pipeline Server Admin Port	10700 - 10709	TCP		Both
SAS Information Retrieval Studio - Pipeline Server Web Admin Port	10710 - 10719	TCP		Both

SAS Information Retrieval Studio - Index Builder Port	10720 - 10729	TCP		Both
SAS Information Retrieval Studio - Query Server Port	10730 - 10739	TCP		Both
SAS Information Retrieval Studio - Query Statistics Port	10740 - 10749	TCP		Both
SAS Information Retrieval Studio - Query Statistics Port	10740 - 10749	UDP		Inbound
SAS Information Retrieval Studio - Crawler Admin Port	10750 - 10759	TCP		Both
SAS Information Retrieval Studio - Query Web Server Port	10760 - 10769	TCP		Both
Multicast Port A multicast group communications protocol is used to communicate configuration among middle tier SAS applications in a single SAS deployment (the set of applications connected to the same SAS Metadata Server). Therefore, the combination of multicast IP address and multicast UDP port should be different for each SAS deployment and also different from those used by other multicast applications at your site. The multicast port defaults to the same value as the metadata server port, but there is no conflict since it is a UDP port vs. a TCP port. Extra care should be taken in the assignment of this port, due to the SAS Deployment Wizard's inability to detect usage conflicts with the entered value.	8560 - 8569	UDP		Both
SAS Deployment Agent	5660-5669	TCP		Both
Event Broker Admin Port	6050 - 6059	TCP		Inbound
Mail Server	25	TCP		Outbound

SAS Environment Manager Dashboard Secure Port (HTTPS)	7442-7050	TCP		Both
SAS Web Application Server JMX 1 Port	6969 (or 6969-6973 for vertical clustering)	TCP		Inbound
SAS Web Application Server JMX 2 Port	7069 (or 7069-7073 for vertical clustering)	TCP		Inbound
SAS Web Server HTTP Port	7980	TCP		Both
SAS Web Application Server HTTP Server 1 Port	8080 (or 8080-8084 for vertical clustering)	TCP		Both
SAS Web Application Server HTTP Server 2 Port	8180 (or 8180-8184 for vertical clustering)	TCP		Both
SAS Web Server HTTPS Port	8343	TCP		Both
SAS Web Application Server HTTPS Server 1 Port	8443 (or 8443-8448 for vertical clustering)	TCP		Both
Operating System Services scheduler	8451			Inbound
SAS Web Application Server HTTPS Server 2 Port	8543 (or 8543-8548 for vertical clustering)	TCP		Both
SAS LASR Analytic Server Monitor RMI Port	9270			Inbound
SAS High-Performance Configuration Management Console Server	10020			Inbound
SAS Visual Analytics Hadoop	15452			Inbound
Cache Locator Port	41415	TCP		Both
Hadoop JobTracker	50030			Inbound
Hadoop TaskTracker	50060			Inbound
Hadoop Name Node Web Interface	50070			Inbound
Hadoop SecondaryNameNode	50090			Inbound
JMS Server Port	61616	TCP		Both
Hadoop DataNodes	50010, 50020, 50075			Inbound

Firewall

For SAS Visual Analytics environments behind a firewall, a number of ports will have to be opened to allow for deployment and usage of the environment. It is the customer's responsibility to make the necessary requests to open the firewall according to the information below.

No firewall should be configured between the SAS Visual Analytics nodes.

Inbound Access

Source	Client	Target	Component accessed	Port	Direction
User/Installer Desktop	SAS Clients (SAS Management Console or SAS Enterprise Guide)	SAS Server Node	Metadata Server	8561	Both
			Object Spawner	8581	Both
			Workspace Server	8591	Both
User/Installer Desktop	Browser	SAS Server Node	High Performance Computing Management Console	10020	Both
			Hadoop Web Interface	50070	Both
			SAS Web Application etc…	7980	Both
User/Installer Desktop	Putty X server	SAS Server Node	SSH	22	Both
			X display over SSH	NA	
iPad connected to internal network	SAS Mobile App	SAS Server Node	SAS Web Application etc…	7980	Both

Note that mobile devices require access to the SAS node. They are expected to be connected to the network via VPN or through an internal company WIFI network. By this, we mean that the whole mobile device is connected to the customer network.

Note that client-less or web-based VPNs are not acceptable for mobile devices because the SAS node is accessed through a native mobile device application and not a web browser.

Mobile devices might not be in the same VLAN as the desktops.

Blade Outbound Access

Source	Client	Target	Component accessed	Port	Direction
SAS Node	SAS Web Application Server	opnsta.sas.com opnstb.sas.com opnstc.sas.com	OpenStreetMap server	80/443	Outbound

The SAS node requires access to the OpenStreepMap servers hosted by SAS in order to issue HTTP or HTTPS requests that are required for the drawing of the map tiles. The host names of the OpenStreetMaps servers are listed in the Target column of the table.

For the Twitter stream import function (which may or may not be used), a proxy is not optional, even if the SAS node has a direct access to the internet.

The customer should provide SAS with the proxy host and port information to be used for the outbound connections, if required.

Proxy Host	Proxy Port

➤ **Validate**

Verify that the SAS server node is able to access the URL of a given tile by running the following commands:

Without proxy:

```
wget http://opnsta.sas.com/mq_lite/0/0/0.png/0
```

With proxy:

```
export http_proxy=http://<your_proxy_host>:<your_proxy_port>
wget http://opnsta.sas.com/mq_lite/0/0/0.png/0
```

Record the following:

Host	SAS Server Node

Sample Output	Without Proxy:
	
	`--2013-11-06 09:44:43-- http://opnsta.sas.com/mq_lite/0/0/0.png/0`
	`Resolving opnsta.sas.com... 149.173.160.5`
	`Connecting to opnsta.sas.com|149.173.160.5|:80... connected.`
	`HTTP request sent, awaiting response... 200 OK`
	`Length: 24917 (24K) [image/png]`
	`Saving to: â0â`
	`100%[======================================>] 24,917 --.-K/s in 0.006s`
	`2013-11-06 09:44:43 (3.80 MB/s) - â0â saved [24917/24917]`
	
	With Proxy:
	
	`--2013-02-15 07:57:02-- http://opnsta.sas.com/mq_lite/0/0/0.png/0`
	`Resolving <your_proxy_host>... 10.11.12.13`
	`Connecting to <your_proxy_host>|10.11.12.13|:<your_proxy_port>... connected.`
	`Proxy request sent, awaiting response... 200 OK`
	`Length: 24917 (24K) [image/png]`
	`Saving to: â0â`
	`100%[======================================>] 24,917 --.-K/s in 0.006s`
	`2013-11-06 09:44:43 (3.80 MB/s) - â0â saved [24917/24917]`

Actual Output	

Chapter 6 — Third Party Software

RDBMS Client

Ensure that the required DBMS client software is installed on the SAS application server if SAS/ACCESS features are to be used.

Chapter 7 — Installation Client

A client PC is required for the SAS Installation Team to access the SAS Visual Analytics environment. The SAS Installation Team uses the client PC to run the various SAS installation tools.

The client PC must be able to initiate a SSH connection to the SAS Visual Analytics environment. An example SSH client is PuTTY.

The client PC must have an X11 display manager installed. MobaXterm is also an acceptable solution. More information can be found in:

http://support.sas.com/kb/50/454.html

and

http://support.sas.com/resources/thirdpartysupport/v94/othersw.html#tab_x11

The client PC must have the SAS client software installed, including the SAS Management Console. The SAS Installation Team requires administrative access to the client PC to complete this installation.

The client PC also must have a browser with Flash installed on it.

The SAS Software Depot should be available to the workstation.

Chapter 8 — Optional

The following sections are optional. Please check with the SAS team whether they apply to you or not.

Storage for Central Vault backup storage

SAS recommends that you have a network-accessible resource in order to store the automated environment back-ups. This location should be mounted all the machines in the same location and has to writeable by the `sas` user.

HTTP(S) Proxy Host

SSL Certificate

To provide encryption, a Secure Sockets Layer Certificate is required. This SSL Certificate is signed by a Certificate Authority (CA) to prove its identity. A client accepting the SSL Certificate will need to trust the Certificate Authority which has signed the SSL Certificate. SSL Certificates can be purchased from trusted third parties or generated in-house. A purchased third-party certificate will by default be trusted by a wide range of clients, while a SSL certificate generated in-house will require additional steps to provide this trust.

Certificates generated in-house are called “self-signed” certificates. Implementing self-signed certificates requires access to the Certificate Authoritiy’s public key. This public key needs to be added to client application’s trusted keystore. Within the SAS configuration, a number of different processes are considered clients for the SSL transaction, including Java processes as well as iPad applications.

Configuring an HTTP proxy to encrypt network communications back to the SAS Web Server may require two SSL certificates. If the Apache HTTP Server is on a separate host to the SAS Web Server, a certificate will be required for each host.

Once the certificates are generated, record the following:

Description	Default Value	Actual Value
SAS Web Application Server Certificate: CN	Blade0.hostname	
SAS Web Application Server Certificate: Issuer	CN=xxx, DC=xxx, DC=xxx	
SAS Web Application Server Certificate: Valid to		
SAS Web Application Server Certificate: Key Usage	Digital Signature, Key Encipherment (a0)	
SAS Web Application Server Certificate: Password		
SAS Web Application Server Certificate: Java Keystore	`server.keystore`	

SAS Web Application Server Certificate: CA Keystore	`Trusted.keystore`	
SAS Web Application Server Certificate: CN	HTTPD.host name	
SAS Web Server Certificate: Issuer	CN=xxx, DC=xxx, DC=xxx	
SAS Web Server Certificate: Valid to		
SAS Web Server Certificate: Key Usage	Digital Signature, Key Encipherment (a0)	
SAS Web Server Certificate: Password		
SAS Web Server Certificate: Certificate File	`webserver.crt`	
SAS Web Server Certificate: Private Key File	`webserver.key`	
SAS Web Server Certificate: CA Certificate File	`CAroot.crt`	

SAS is the leader in business analytics software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions, SAS helps customers at more than 65,000 sites improve performance and deliver value by making better decisions faster. Since 1976 SAS has been giving customers around the world THE POWER TO KNOW®.