Reducing the Backup Window With Sun StorEdge™ Instant Image Software

Selim Daoud, Sun Professional Services, Switzerland

Sun BluePrints™ OnLine—July 2002

http://www.sun.com/blueprints

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95045 U.S.A.
650 960-1300

Part No. 816-5243-10
Revision 1.0, 7/15/02
Edition: July 2002

Please
Recycle

Reducing the Backup Window With Sun StorEdge Instant Image Software

This article discusses the advantages and methods of using a point-in-time (PIT) type of backup system versus a more traditional backup approach that requires extended downtime. This article is for anyone interested in reducing the *backup window* (improving the uptime of important applications) while backing up a system that is *nearly online*.

The process described here is called *nearly online* because you must take into consideration that to create a consistent snapshot of a file system, you must take the system offline to perform the following procedures:

- Quiesce I/O on the volume.
- Make sure the applications and data associated with this volume are in a good state (Oracle software must be in backup mode, file systems must be synchronized with utilities like `fsck`, and so on).
- Postpone all write operations until the snapshot is complete.

While the backup techniques described in this article are not focused on specific versions of backup software or hardware, this article uses the following products to provide realistic examples:

- Sun Solstice Backup™ software—Sun's OEM version of Legato's Networker software
- Sun StorEdge™ Instant Image software—Sun's software snapshot product

Almost any supported version of these two products can be integrated into your networked computing environment to gain the advantages of reducing the backup window and backing up a nearly online system.

Why Create a Snapshot?

Creating a snapshot of a volume (file system or raw partition) has several advantages:

- You obtain a PIT image of the data so all data is backed up at the same point-in-time, whatever the size of the master volume.
- The snapshot relieves the original volume (master volume) from excessive I/O during the backup.

Note – I/O is only relieved on the master volume when the snapshot is an independent snapshot as opposed to a dependent snapshot (also referred to as a copy-on-write). When performing a read I/O on a dependent snapshot, unmodified blocks are read from the master volume and modified blocks are read from the shadow volume. By contrast, read I/O from independent snapshots is performed entirely from the shadow volume, relieving the I/O on the master volume.

The latter point is true in the case where the snapshot is a full copy (independent snapshot) of a volume. As a consequence, a volume that has been copied can be made available to the original application. This has the following advantages:

- The window to backup this volume is very small.
- You can obtain several backups during the day.
- There is minimum disruption on the master volume.
- A copy of the primary volume is available for special applications (testing, data warehousing, and so on).

Comparison

To appreciate the benefits of a PIT backup system, consider the comparison between a PIT backup and a non-PIT backup.

Traditional Backup

In a traditional non-PIT file system backup, the backup process works schematically as shown in the following figure.

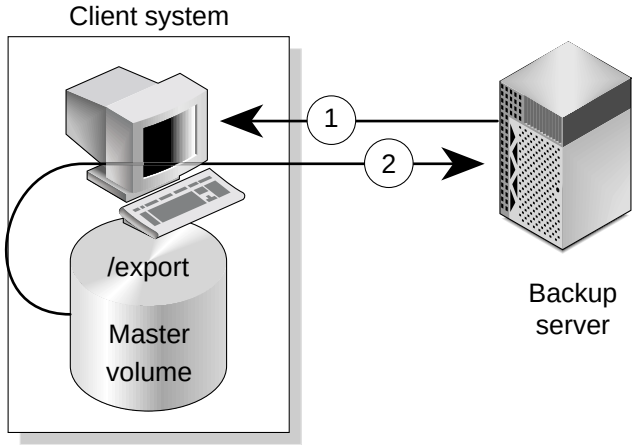

FIGURE 1 Non-PIT Backup

1. The backup server contacts the client, and arguments are passed defining what to save.
2. The client daemon selects all the files, and transfers them back to the server.

PIT Backup

When using the snapshot technology for a PIT backup, the backup process works schematically as shown in the following figure.

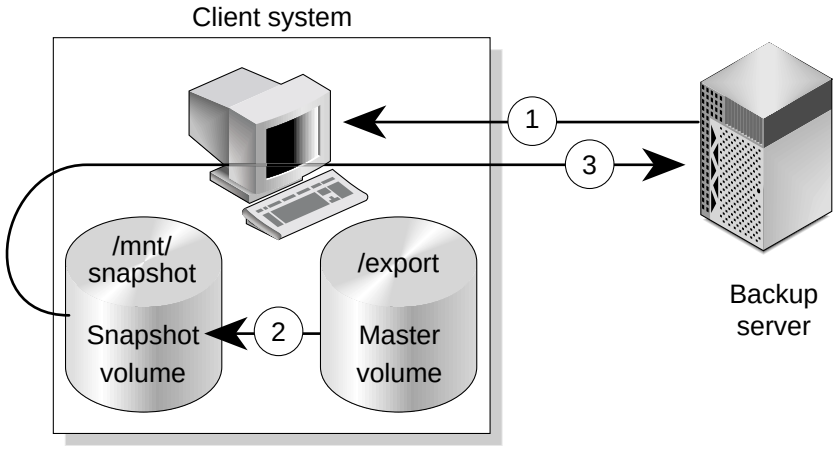

FIGURE 2 PIT Backup

1. The backup server contacts the client, and arguments are passed defining what to save.
2. The client creates the snapshot image (quick update).
3. The client daemon selects all the files on the snapshot volume and transfers them back to the server.

Notice that there is an additional step consisting of creating the snapshot volume. What happens is, in fact, a quick update because only modified parts (disk blocks) of the master volume are updated. Therefore, the creation of the snapshot is very quick.

Results

The following two graphs illustrate the I/O measurements that occur during the two types of backups.

In the first graph, the backup is performed without an Instant Image snapshot. The master volume experiences a sustained high level of I/O activity.

The second graph reflects the I/O measurements while performing a PIT backup using a snapshot with Instant Image software. You can observe that in terms of I/O, there is a load on the master volume at the beginning of the backup, but the duration is short when compared to the other I/O graph. This short period of master volume activity corresponds to the quick update phase of the backup. Following this, all the I/O is diverted to the shadow volume, relieving the master volume.

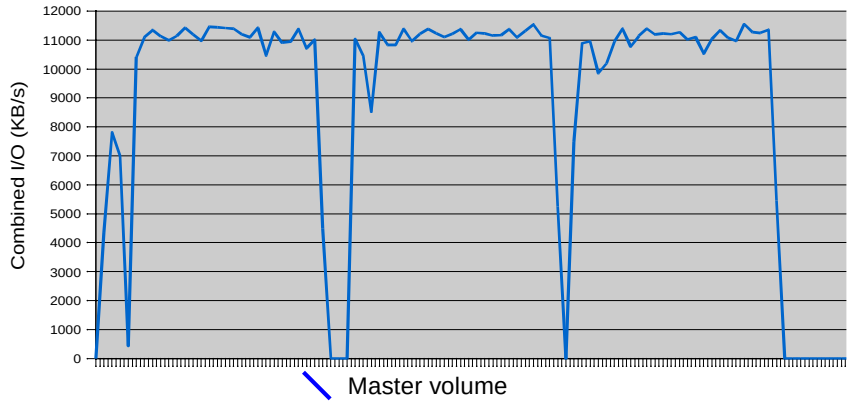

FIGURE 3 Backup I/O Statistics Without Instant Image Software

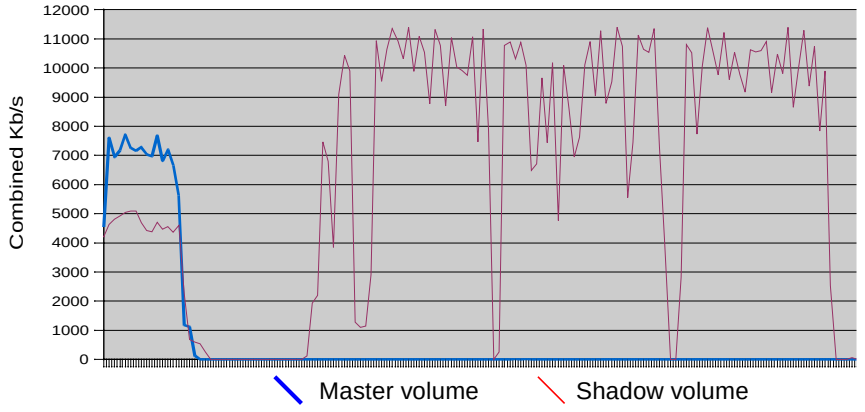

FIGURE 4 Backup I/O Statistics With Instant Image Software

Creating a Snapshot

Creating a snapshot of a master volume with Instant Image software (and generally with any PIT software) involves three components:

- Master volume
- Snapshot volume (also called the shadow volume)
- Bitmap volume (a raw partition)

To create an independent snapshot (a full copy, for example), the size of the master volume and the shadow volume are identical. The bitmap volume is much smaller and is equal to:

Bitmap Size = 8KB per GB + 24KB

The examples in this article use the following parameters:

Volume	Device	Capacity	Mount Point
Master volume	`/dev/rdsk/c0t0d0s6`	7 GB	`/export`
Shadow volume	`/dev/rdsk/c0t11d0s6`	7 GB	`/mnt/snapshot`
Bitmap volume	`/dev/rdsk/d0t11d0s0`	80 KB	
II Group name	`BACKUP`		
Solstice Backup Group name	`Default`		

Pre-Processing Script

A pre-processing script is created on the client, and initiated by the backup server. This script prepares the snapshot just before the backup operation, and consists of the following steps:

1. Synchronizing the master volume.
2. Locking the master volume against any modification.
3. Initiating the update.
4. Unlocking the master volume.
5. Mounting the snapshot image.

The following code sample is an example of a pre-processing script.

```
#!/bin/sh
# need to explicit binary path when running pres/post script
PATH=/sbin:/usr/local/bin:/usr/local/sbin:/usr/bin:/usr/sbin:/
usr/opt/SUNWesm/sbin
export PATH

# definition of volumes used in II
SHADOW_RAW=/dev/rdsk/c0t11d0s6
SHADOW=/dev/dsk/c0t11d0s6
PRIMARY_MNT=/export
SHADOW_MNT=/mnt/snapshot
II_GROUP=BACKUP

# we re-direct output in a logfile
exec >> /nsr/logs/snapshot.log 2>&1

# we lock the Master volumes so no I/O take place during "Quick
Update"
echo "Quick update started"
lockfs -f -w $PRIMARY_MNT
iiadm -g $II_GROUP -u s

if [ ! "$?" = 0 ] ; then
        echo "failed to update shadow volume"
        exit 1
fi
echo "Update finished."

# we wait for Instant Image to finish its operations
iiadm -g $II_GROUP -w

# Master Volume can be reactivated and ready for normal operations
lockfs -u $PRIMARY_MNT


# when doing PIT of a mounted volume, fsck must be executed on
snapshot
fsck -y $SHADOW
# we mount the snapshot volume, ready to be backed up
mount $SHADOW $SHADOW_MNT

echo "Pre-Script finished"
```

Post-Processing Script

A post-processing script is also created so that normal operations can continue when the snapshot is complete. The post-processing script is much simpler with one purpose, to unmount the shadow volume. The following code sample is an example of a post-processing script.

```
#!/bin/sh
# need to explicit binary path when running pres/post script
PATH=/sbin:/usr/local/bin:/usr/local/sbin:/usr/bin:/usr/sbin:/
usr/opt/SUNWesm/sbin
export PATH


SHADOW_MNT=/mnt/snapshot

II_GROUP=BACKUP

# we re-direct output in a logfile

exec >> /nsr/logs/snapshot.log  2>&1
echo " "
echo "Unmounting snapshot volume ($SHADOW_MNT)..."
echo "-----------------------------------"
umount $SHADOW_MNT
echo "Post-Script finished."
```

Putting Everything Together

1. **Configure Instant Image software.**

   In this example, Instant Image software is activated using the following command:

```
# iiadm -g BACKUP -e ind /dev/rdsk/c0t0d0s6 /dev/rdsk/c0t11d0s6
/dev/rdsk/c0t11d0s0
```

2. **Create the pre-processing and post-processing scripts on the client.**

   If you are using Solstice Backup software, the two scripts must be placed in the `/nsr/res` directory, along with a file called *SBU_Group*`.res`. The *SBU_Group*`.res` file defines the pre- and post-processing scripts.

   In our example, the *SBU_Group*`.res` file contains the following Solstice Backup software parameters:

```
# cat Default.res
type: savepnpc;
precmd: /nsr/res/start.snapshot;
pstcomd: /nsr/res/stop.snapshot;
timeout: “12:00pm”;
```

3. **Configure Solstice Backup software to activate the scripts.**

   For each specific client, in the corresponding Solstice Backup software group, modify the definition of this client so that the parameter Backup Command displays `savepnpc`, as shown in the following figure.

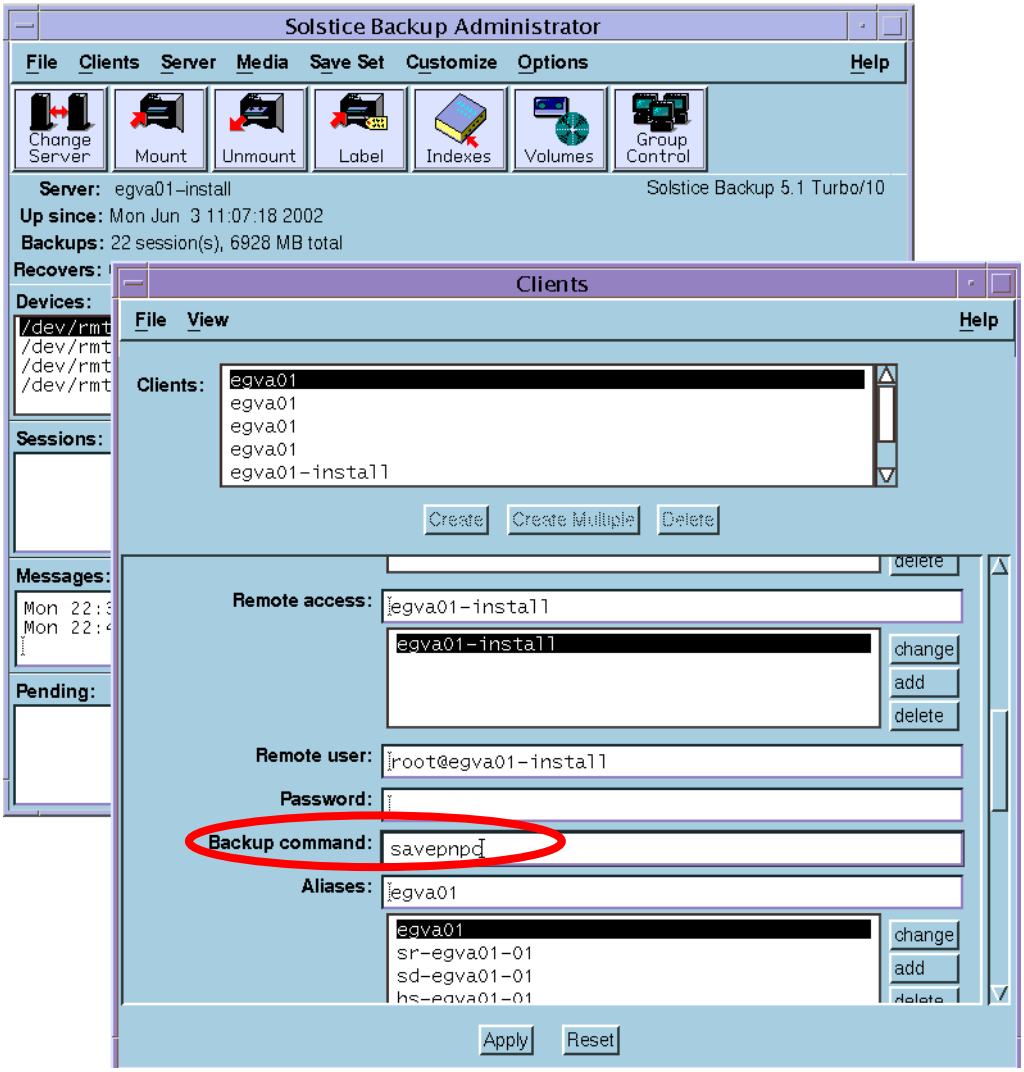

FIGURE 5 Modified Backup Command in the Sun Solstice Backup Client Window

Conclusion

While the examples in this article use Sun StorEdge Instant Image software and Sun Solstice Backup software, it is possible to obtain similar results with other backup tools such as Veritas Netbackup.

Veritas Netbackup also includes a snapshot option called FlashBackup. FlashBackup uses a dependent copy (or copy-on-write). Although this feature is an interesting way of obtaining a PIT copy (because it can be created quickly), it does not relieve the I/O on the master volume, and therefore the backup degrades performance of applications using this volume.

The PIT method described in this article is particularly applicable if the volume backed up is a UFS type file system. The main reason is that a snapshot is successful if the master volume can be quiesced during creation of the shadow volume. In our examples, we used the command `lockfs` to achieve this. This is not applicable for raw partitions, however. For raw partitions, other processes must be developed to make sure nothing is written on the volume (unmounting it, stopping all applications using it, and so on).

In conclusion, with a small investment in backup planning and configuration, you can implement backup strategies that effectively shorten the backup window, thus improving the availability of important applications such as databases, web servers, and home directories.

About the Author

Selim Daoud is a recognized leader in data storage and backup technologies for open systems. He obtained an MSc in computer science at the University of Wales (UK), and an MSc in applied mathematics in computing at Toulouse University in France.

Over the course of his career in the computer industry, Selim gained valuable experience working with data backup technology, storage system design (mainly RAID implementations), and UNIX systems administration. He managed a support organization (dealing with storage and backup technology) in London, served as a consultant specializing in backup systems in Paris, and was in charge of multiple migrations of backup systems and storage deployment for the European Organization for Nuclear Research (CERN) in Geneva, Switzerland.

Selim currently holds a project engineering position, specializing in computer storage and backup technology in the Sun Professional Services organization in Switzerland.

Ordering Sun Documents

The SunDocs℠ program provides more than 250 manuals from Sun Microsystems, Inc. If you live in the United States, Canada, Europe, or Japan, you can purchase documentation sets or individual manuals through this program.

Accessing Sun Documentation Online

The `docs.sun.com` web site enables you to access Sun technical documentation online. You can browse the `docs.sun.com` archive or search for a specific book title or subject. The URL is `http://docs.sun.com/`.

To reference Sun BluePrints OnLine articles, visit the Sun BluePrints OnLine Web site at: `http://www.sun.com/blueprints/online.html`