Global Information Assurance Certification Paper

Amanda, the Advanced Maryland Automated Network Disk Archiver

Drew Einhorn
December 19, 2000

Introduction

Amanda is designed to backup large computer networks, but also works well for a single computer. Clients and servers run on nearly any contemporary unix platform. Amanda can also support MS Windows clients via Samba.

Amanda is designed to run unattended late at night from a cron job.

History

Like most open source projects Amanda began when a user "wrote some code" to solve a local problem. Around 1992 James da Silva of the University of Maryland Computer Science Department was faced with a large number of client workstations drives he needed to backup using a single server with a large tape drive. As time went by other folks began using the software, and other developers joined the team contributing their enhancements to the software. Currently development is hosted at SourceForge and the team consists of 11 developers. Based on the membership of Amanda related mailing lists the user community is estimated at over 1500 users.

Amanda Stability

Amanda has been very stable as shown by the following table of recent Amanda releases:

11/21/98	2.4.1p1	Previous Production Release
12/16/99	2.4.2	Beta 1
10/25/00	2.4.2	Beta 2
11/25/00	2.4.2	Production Release

Over the two-year life of 2.4.1p1 only a handful of patches were issued. We see evidence of a mature, stable software package that is thoroughly tested prior to its release.

Amanda Information

The most recent production release is always available at the Amanda home page. A snapshot of the current development version is available for anonymous CVS from SourceForge.

Information regarding the Amanda mailing lists: amanda-announce, amanda-users, and amanda-hackers is available at the Amanda home page. Members of the Amanda development team monitor the amanda-users list and provide excellent support.

John R. Jackson, a very active member of the Amanda development team, and frequent responder to questions in the Amanda-users list, has written a chapter on Amanda for W.

Curtis Preston's book: *UNIX Backup & Recovery*, published by O'Reilly and Associates. The chapter is available online at BackupCentral. The Amanda chapter covers the current release, 2.4.2.

BackupCentral is also an excellent source for information about other backup software, backup hardware, and the related topic of disaster recovery.

Features

Configuration

Amanda is configured using text configuration files. There is no GUI.

Tape Handling

Amanda supports any tape drive supported by the underlying operating system and many tape changers. The interface for the tape changer hardware is well documented and it is relatively easy to add support for a new changer.

Amanda writes header records on each tape. Amanda checks these headers and will not overwrite an “active” tape. Amanda will mount each tape in a changer looking for tonight’s tape.

Amanda normally runs unattended via a nightly cron job. A diagnostic job can be scheduled to run during normal business hours. It will check to insure that all systems are up and functional, and that the correct tape is in the drive or changer, and send Email notification of any errors to the system administrator. Problems can be then corrected, before the nightly run.

Holding Disk

Amanda can be configured to concurrently dumps multiple computers/partitions to a holding disk. Then completed dump images are copied from disk to tape in full speed "streaming" mode.

Data Compression

Amanda can also use tape hardware compression, but gets better time/tape use estimates with software compression. A tape written without hardware compression is more likely to be readable on a different tape drive.

When Amanda is configured for software data compression, each filesystem can be configured independently for compression at the client, server, or no compression.

Data compression at the clients Amanda distributes CPU load and reduces network traffic.

Amanda can perform compression on the server for old clients lacking the necessary CPU horsepower.

You can select your own compression software, and may get much higher compression rates for special data file format/compression method combinations.

Fails gracefully

If a host or a portion of the network is down when the backup runs, the failure will be logged, reported, and the backups will be rescheduled for the next run. If the tape drive fails the error will be logged and reported. Amanda will then fall back to all "incremental" backups to the holding disk. A manual procedure can flush the backups from disk to tape before the next scheduled run.

Security.

Amanda can be configured for Kerberos IV or can use an .amandahosts configuration file similar to .rhosts. However Amanda uses its own protocols, does not fork general-purpose shells, and is not subject to the abuses of .rhosts files.

Future releases will support ssh, ssl, Kerberos V and other mechanisms for secure communication between client and server.

Client Dump Programs

The current release supports gnu tar, proprietary vendor versions dump/restore, and samba tar.

Future releases will support cpio, star, and possibly other backup programs. Future releases will also support pre- and post- dump procedures to deal with locking, database hot backups, Windows registry exports, etc.

On some platforms vendor dump/restore programs may be faster than gnu tar, and may support platform vendor specific enhancements, such as access lists.

But in many cases the cross platform portability of gnu tar archives makes it a better choice. With gnu tar you can recover data on a box with lots of available disk space, even if it is a different hardware platform or operating system, investigate the extent of the damage, and then move the necessary files to the damaged system. Gnu tar does disturb access dates, may be a problem in some situations.

With gnu tar Amanda uses an amandadates file similar to dump/restore's dumpdates file and is able to do incremental backups.

Amanda Recovery

Amanda maintains a database of backup tape contents. Amanda will prompt the administrator to load the required tapes for a recovery.

We can still recover even if this database is not available. The essential information can be recovered from headers recorded on the backup tapes.

Here is a Bourne shell fragment I used to recreate an index of filesystems backed up on each tape, when recovering from a loss of the filesystem containing the Amanda database and configuration files.

```
TAPEDEV=/dev/rmt/1lbn
mt -f $TAPEDEV rewind
while true ; do
  dd if=$TAPEDEV bs=32k count=1 | head -1
```

```
    sleep 1
    mt -f $TAPEDEV fsf 1
  done
```

See the Amanda chapter for information on how to recover from these tapes without the Amanda software.

But beware! This script has serious problems!

It reads past EOT (End Of Tape). This can have catastrophic effects on some ancient tape drives. Modern tape drives usually handle this gracefully. Find out about your tape drive using a test backup tape, not a critical production backup tape. You probably want to know if your tape drive has a pathological response to this error.

The good stuff goes to stdout; dd spews a bunch of crud to stderr.

My tape drive seemed to need a second for the hardware to settle after reading, before forward spacing to the next file.

This script was good enough to get the job done, the first time I ran into this problem. I haven't needed it, or worked on it since. Your mileage may vary.

Practice, Practice, Practice

Periodically test your recovery procedures. Chose difficult recovery scenarios, the loss of the drive containing the Amanda software, the loss of your operating system boot drive. BackupCentral has a section on the bare metal recovery.

In a real situation where you have lost you OS boot drive, consider cheating: Add an extra drive to a different system, partition the drive, create the filesystems, recover the data, set the boot sectors, using a functional OS, instead of doing it with your arms tied behind your back on the crashed system. You will have more flexibility if you used gnu tar for you backups, instead of a proprietary dump/restore.

The time to learn difficult recovery techniques is when you are practicing, not when you are desperate.

Scheduling Backups

Amanda automatically schedules full/incremental dumps to maintain balanced daily runtimes and tape usage.

Some administrators find it difficult to accept this feature since it is contrary to the common practice of performing full backups for all systems on a designated day of the week and incremental backups on other days.

In a large network with many backup clients the predictable tape use and runtime is far more important. The backups must be complete by the time folks come in to work in the morning. The ability to automatically reschedule full backups that failed because of hardware or network problems is also important.

Amanda Weaknesses

Currently Amanda cannot handle a backup image larger than a single tape. This will be

addressed in a future release. Gnu tar has features that will allow backing up subdirectories of a filesystem.

Amanda cannot append to an existing backup tape. This is the topic of a constantly recurring discussion in the amanda-users mail list. Reliably positioning at the end of tape is extremely difficult with some tape drives. For example, some scsi drives will rewind, in response to a scsi reset, possibly caused by powering up/down some other piece equipment on the scsi bus. The Amanda developers are more worried about the protecting the backups at the beginning of the tape, than they are about using the empty space at the end of the tape. I would not hold my breath waiting for this feature to be implemented.

References:

SourceForge Projects Page, URL: http://sourceforge.net/projects/amanda/ (8 Dec 2000).

The Amanda Home Page. URL: http://www.amanda.org/ (8 December 2000).

Eggert, Paul. Gnu Tar 17 Jan 2000. URL:http://www.gnu.org/gnulist/production/tar.html (8 Dec 2000).

Jackson, John R. "Using Amanda" *UNIX Backup & Recovery* November 1999. URL : http://www.backupcentral.com/amanda.html (8 December 2000).

Preston, W. Curtis *UNIX Backup & Recovery* November 1999. O'Reilly and Associates.

Bare-metal Recovery, URL: http://www.backupcentral.com/bare-metal-recovery.html (8 December 2000)