Global Information Assurance Certification Paper

Using AMANDA for High Performance Backups

by Laurence G. Guentert

I. Introduction

One of the key aspects of the computer security paradigm – Confidentiality, Integrity and Availability – is having the capability to easily recover from a security breach by restoring a system to a known good state. It is essential to have a reliable backup scheme in place to help recover from a security breach, data theft, a natural disaster, or system failure. Many commercial tape backup schemes at the departmental level are server centric, i.e., the goal is to back up files on a *server* or servers. This document describes a non-commercial tape backup program AMANDA, whose forte is backing up *many* computers on a LAN quickly. While describing every aspect of the program is beyond the scope of this article, a description is included as well as some key sample setup files that will greatly help configuring the program for the novice system administrator.

II. Description

What is AMANDA precisely? AMANDA is an acronym for the Advanced Maryland Automatic Network Disk Archiver - a program written and maintained by programmers at the University of Maryland. The program which is available for free (with a few restrictions) allows a LAN administrator to use a single host backup computer to backup multiple hosts to a single high-capacity tape drive. The program is very well thought out and covers about every imaginable backup scenario. Perhaps its key feature that greatly enhances performance over the server-centric model is that it uses a large holding disk drive that can be written to very quickly, and once the files are brought in over the wire to the holding disk, they are then dumped to tape at optimal speeds. Essentially the backup is done to disk very quickly, and then the host computer writes the disk to tape very quickly.

III. Key Features

- uses a large holding disk to greatly accelerate performance
- hosted by a Unix system but can back up **Unix and NT** platforms
- written in C and the source code is available
- has extensive logging features
- has a *pre-run* checker that will alert you of problems ***before*** backup begins
- reports results of operations by **e-mail** to administrators
- supports Kerberos security
- supports tape changers for DLT, DAT and other high-capacity drives
- can compress a backup dump with gzip or compress and compress the dump before or after sending it over the net.

- greatly configurable as it uses standard Unix backup software such as dump/restore and GNU Tar.
- gracefully recovers from typical errors such as tape full, hung clients, etc.

IV. A Typical Real-Life Configuration

The computer that hosts the tape backup drive and holding disk need not be the latest high performance computer. There are many anecdotal reports of system administrators who have converted an older Pentium computer to their backup system by installing some flavor of Linux on it and installing a tape drive unit, a large holding disk and AMANDA. This type of set up was recently installed at Purdue University in a Pentium 300 mhz computer with a DLT tape drive and a 70 gigabyte holding disk. After setting up AMANDA we saw backup times drop more than 50% from our previous Unix-based backup system which backed up more then 80 Unix and Windows NT workstations!

With the release of so many new flavors of Unix heralded by the release of Linux, there are many new system administrators who don't have a lot of expertise configuring a tape backup system. Some of the following example questions and real setup files are provided as a resource to the fledgling sys admin.

Some Typical Set-up Questions/Examples/Files:

1. How Do I schedule AMANDA to start a backup:

Using CRON, you would enter a line similar to the one below in your CRONtab file..

Example**:**

0 22 * * 1-5 /usr/local/amanda/sbin/amdump bio ; mt -f /dev/rmt/0hn rewoff

2. How do I list the systems/disks that I want to backup?

Example of disk list *file*:

```
#
# Biology department amanda disk list
#

#
# io is the Amanda master host
#################################################################################
######
```

```
# io.bio.purdue.edu     2000 May 24
io        c0t3d0s0        system-fs
io        c0t1d0s7        user-fs


DEC Alphas go here in alphabetical order...
###############################################################################
######

# alf.bio.purdue.edu    2000 June 21
alf       /               gnutar-system-fs
alf       /usr            gnutar-system-fs
alf       /usr/alfgnutar-user-fs

#
# RS/6000s go here in alphabetical order...
###############################################################################
######

# alaska.bio.purdue.edu                 10/27/98
alaska  hd4     system-fs
alaska  hd2     system-fs

# Suns go here in alphabetical order...
###############################################################################
######

# ansel.bio.purdue.edu                  8/14/96
ansel   c0t3d0s0        system-fs
ansel   c0t1d0s7        data-fs

#
# SGIs go here in alphabetical order...
###############################################################################
######

# 3dem.bio.purdue.edu           2000 February 23
3dem    root            system-fs
3dem    dks0d2s7        nocomp-data-fs
3dem    dks1d1s7        nocomp-data-fs
3dem    dks1d2s7        nocomp-data-fs
3dem    dks1d3s7        user-fs


# linix goes here in alphabetical order...
```

```
##########################################################################
######

# dai.bio.purdue.edu   28 July 1999
dai     /       system-fs
dai     /usr    system-fs
dai     /home   user-fs

#
# NT machines go here
##########################################################################
######

# sullivan (thru io) 2000 August 22
io      //sullivan/testamanda   gnutar-nocomp-data-fs
```

3. How do I know what files are on what tape(s)?

The a*madmin* command can show you on a per filesystem basis which dumps are on which tapes. For example:

```
# amadmin bio info bilbo /home

Current info for bilbo /home:
  Stats: dump rates (kps), Full:  546.0, 500.0, 238.0
                 Incremental:  160.0, 215.0, 172.0
         compressed size, Full:  55.0%, 58.2%, 56.2%
                 Incremental:  13.7%, 25.7%, 25.9%
  Dumps: lev datestmp  tape         file   origK   compK secs
         0  20001229  VOL008          212 3862780 2125824 3889
         1  20010101  VOL009           77    7019     960    6
         2  20001213  VOL102          253  160519  147520  529
         3  20001220  VOL001          256  166940  150400  711
         4  20001228  VOL007          190  130142   33408  155
```

On top is the stats for the last three fulls and incrementals. Next is the list of active tapes. From this list you can see that bilbo's /home was full dumped to tape "VOL008" on Dec. 29 and file 212 on that tape.

4. How do I do a restore a specific machine's backup from the full backup set?

The a*mrestore* command allows you to grab a specific dump off of a tape (or pipe it into another command). If I were going to do a restore of computer bilbo's /home from the

full backup, I could start by putting tape "VOL008" in my tape drive and issuing the following command:

amrestore -c /dev/rmt/0hn bilbo /home

The "-c" flag tells amrestore to restore the compressed file rather than uncompressing in as it pulls it off of the tape.

5. How do I restore a specific file?

The amrecover command has an interface like the UNIX *restore* command and you can go and query it to find where a specific file is and run the restore itself. This only works if you have been using the index feature of amanda to build up the tape index files and you have the index server set up on the amanda backup server. Here's an example:

```
[root@bilbo /etc]# cd /home
[root@bilbo /home]# /usr/local/amanda/sbin/amrecover -s io -C bio
AMRECOVER Version 2.4.1. Contacting server on io ...
220 io AMANDA index server (2.4.1p1) ready.
200 Access OK
Setting restore date to today (2001-01-02)
200 Working date set to 2001-01-02.
200 Config set to bio.
200 Dump host set to bilbo.
$CWD '/home' is on disk '/home' mounted at '/home'.
200 Disk set to /home.
/home
amrecover> cd ginn
/home/ginn
amrecover> add playlist
Added /ginn/playlist
amrecover> list
TAPE VOL008 LEVEL 0 DATE 2000-12-29
        /ginn/playlist
amrecover> quit
200 Good bye.
```

6. What does a Tape.conf file look like

As this file is rather lengthy, see **Appendix 1.**

7. What does some output of the pre-run checker program that is e-mailed to the sys admin look like?

See **Appendix 2**

V. Conclusion

The tape backup program AMANDA is a free, powerful and fast, Unix-based tape backup program that can greatly ease a system administrator's mind by helping them perform system backups and restores on multiple computers in a department, and thus ensuring the integrity of the data to which they have been entrusted.

VI. Additional Internet Resources

Internet download Sites for AMANDA::
ftp://ftp.cs.umd.edu/pub/Amanda
http://http://www.amanda.org/pub/amanda/

AMANDA's Main Website is:
http://www.amanda.org/

Information on Tape types:
http://www.cs.columbia.edu/~sdossick/amanda/

Internet Program Review and Compile Examples:
http://filewatcher.org/sec/amanda/int_1week.html

Internet Information on DDS3 Tape Types:
http://www.control.auc.dk/~magnus/Mailboxe/amanda-archive.1/1546.html

Internet Help Regarding SCSI Error:
http://web.gnu.walfield.org/mail-archive/linux-kernel/2000-January/0636.html

Internet Tip on Help with 4MM Tape:
http://lists.openresources.com/FreeBSD/freebsd-scsi/msg00036.html

Other Common Questions:

I. What is the current version:
version 2.4.2

II. How can I receive announcements about AMANDA?
send e-mail to: Amanda-announce-request@amanda.org

III. Is there an AMANDA users group?
send e-mail to: Amanda-users-request@amanda.org

VII. Acknowledgements

Thanks to Dwight McKay for his help in supplying example files.

Appendix 1 – Example AMANDA.CONF File

```
#
# biology department amanda configuration file.
#
# test amanda.conf file
#
# Original by Meng
# Modification history:
# 6/5/00
#       Moved to io and make initial changes. --ddm
#

org "bio"               # your organization name for reports
mailto "bio-amanda"  # the mailing list for operators at your site
dumpuser "root"             # the user to run dumps under

inparallel 10           # maximum dumpers that will run in parallel
netusage  8000          # maximum net bandwidth for Amanda, in KB per sec

dumpcycle 4 weeks   # the number of weeks in the normal dump cycle
runspercycle 20             # the number of amanda run in dumpcycle days
                        # (4 weeks * 5 amdump runs per week -- just weekdays)
tapecycle 107 tapes     # the number of tapes in rotation

bumpsize 20 MB              # minimum savings (threshold) to bump level 1 -> 2
bumpdays    2           # minimum days at each level
bumpmult    2           # threshold = bumpsize * (level-1)**bumpmult

# tpchanger "chg-generic" # the tape-changer glue script
# tapedev "/dev/nrst8"# or use the (no-rewind!) tape device directly

tapedev         "/dev/rmt/0hn"# no-rewind, high density
```

```
tapetype DLT7000      # what kind of tape it is (see tapetypes below)
labelstr "^VOL[0-9][0-9]*$"  # label constraint regex: all tapes must match

####NEW#####
runtapes 1 #number of tapes to be used in a single run of amdump

##old## diskdir "/usr/amanda"          # where the holding disk is
##old## disksize 16600 MB                    # how much space can we use on it
holdingdisk hd1 {
        comment "main holding disk"
        directory "/export/amanda"
        use 39 Gb
        chunksize 1 Gb ##so we have 25 dump files
        }

# Amanda needs a few MB of diskspace for the log and debug files,
# as well as a database.  This stuff can grow large, so the conf directory
# isn't usually appropriate.  We use /usr/adm.  Create an amanda directory
# under there.  You need a separate infofile and logfile for each
# configuration, so create subdirectories for each conf and put the files
# there.  Specify the filenames below.

infofile "/usr/adm/amanda/bio/curinfo"      # database filename
logfile  "/usr/adm/amanda/bio/log"    # log filename
indexdir "/usr/adm/amanda/bio/index"    # index directory

# tapetypes
#
# Define the type of tape you use here, and use it in "tapetype" above.
# Some typical types of tapes are included here.  The tapetype tells amanda
# how many MB will fit on the tape, how big the filemarks are, and how
# fast the tape device is.
#
# For completeness Amanda should calculate the inter-record gaps too, but it
# doesn't.  For EXABYTE and DAT tapes this is ok.  Anyone using 9 tracks for
# amanda and need IRG calculations?  Drop me a note if so.
#

# latest adjustment 6/11/97 --ddm

define tapetype EXB-8500 {
   comment "Exabyte EXB-8500-alike drive on IBM RS/6000-570"
   length 4000 mbytes
   filemark 60 kbytes
   speed 474 kbytes
```

```
}

# DLT 7000 with CompactTape IV cartridges
#       first approximation 5/5/98 --ddm
#       updated 2000 April 20 --ddm

define tapetype DLT7000 {
   comment "Quantum DLT 7000 drive on a SPARCstation 5 w/ SunSwift adapter"
   length 33500 mbytes
   filemark 8 kbytes
   speed 2500 kbytes
}

# dumptypes
#
# These are referred to by the disklist file.  The dumptype specifies
# certain "options" for dumping including:
#       compress-fast  - (default) compress on the client using fast algorithm
#       compress-best - compress using the best (and slowww) algorithm
#       no-compress    - don't compress the dump output
#       record         - (default) record the dump in /etc/dumpdates
#       no-record      - don't record the dump, for testing
#       no-hold                - don't go to the holding disk, good for dumping
#                        the holding disk partition itself.
#       skip-full      - Skip the disk when a level 0 is due, to allow
#                        full backups outside Amanda, eg when the machine
#                        is in single-user mode.
#       skip-incr      - Skip the disk when the level 0 is NOT due.  This
#                        is used in archive configurations, where only full
#                        dumps are done and the tapes saved.
#       no-full        - Do a level 1 every night.  This can be used, for
#                        example, for small root filesystems that only change
#                        slightly relative to a site-wide prototype.  Amanda
#                        then backs up just the changes.
#
# Also, the dumptype specifies the priority level, where "low", "medium" and
# "high" are the allowed levels.  These are only really used when Amanda has
# no tape to write to because of some error.  In that "degraded mode", as
# many incrementals as will fit on the holding disk are done, higher priority
# first, to insure the important disks are dumped first.

#####NEW---Indexing for all dumptypes defined after global#####
define dumptype global {
        comment "Global definitions"
        index yes
}
```

```
#####NEW---Allow server side compression#####
define dumptype server-compress {
        global
        comment "Compression on server side"
        compress server fast
        priority medium
}

define dumptype comp-user {
        global
        comment "Non-root partitions on reasonably fast machines"
        ##options compress-fast
        compress fast
        priority medium
}

define dumptype nocomp-user {
   global
   comment "Non-root partitions on slow machines"
   ##options no-compress
   compress none
   priority medium
}

define dumptype holding-disk {
   global
   comment "The master-host holding disk itself"
   ##options no-hold, no-compress
   holdingdisk no
   compress none
   priority low
}

define dumptype comp-root {
   global
   comment "Root partitions with compression"
   ##options compress-fast
   compress client fast
   priority low
}

define dumptype nocomp-root {
   global
   comment "Root partitions without compression"
   ##options no-compress
```

```
    compress none
    priority low
}

define dumptype comp-high {
    global
    comment "very important partitions on fast machines"
    ##options compress-best
    compress client best
    priority high
}

define dumptype nocomp-high {
    global
    comment "very important partitions on slow machines"
    ##options no-compress
    compress none
    priority high
}

define dumptype nocomp-test {
    global
    comment "test dump without compression, no /etc/dumpdates recording"
    ##options no-compress, no-record
    record no
    compress none
    priority medium
}

define dumptype comp-test {
    global
    comment "test dump with compression, no /etc/dumpdates recording"
    ##options compress-fast, no-record
    record no
    compress client fast
    priority medium
}

define dumptype user-fs {
        global
      comment "Biology User filesystem"
        ##options compress
        compress client fast
        priority high
}
```

```
define dumptype gnutar-user-fs {
        global
      comment "Biology User filesystem"
        ##options compress
        compress client fast
        priority high
        program "GNUTAR"
}

define dumptype gnutar-nocomp-user-fs {
        global
      comment "Biology User filesystem"
        ##options compress
        compress none
        priority high
        program "GNUTAR"
}

define dumptype noidx-user-fs {
        global
      comment "Biology User filesystem"
        ##options compress
        compress client fast
        priority high
        index no
}

define dumptype nocomp-user-fs {
        global
      comment "Biology User filesystem, nocompression"
        ##options no-compress
        compress none
        priority high
}

define dumptype system-fs {
        global
      comment "Biology System filesystem"
        ##options compress
        compress client fast
        priority medium
}

define dumptype gnutar-system-fs {
        global
      comment "Biology System filesystem"
```

```
        ##options compress
        compress client fast
        priority medium
        program "GNUTAR"
}

define dumptype noidx-system-fs {
        global
    comment "Biology System filesystem"
        ##options compress
        compress client fast
        priority medium
        index no
}

define dumptype nocomp-system-fs {
        global
    comment "Biology System filesystem, no compression"
        ##options no-compress
        compress none
        priority medium
}

define dumptype data-fs {
        global
    comment "Biology Data Collection filesystem"
        ##options compress
        compress client fast
        priority low
}

define dumptype gnutar-data-fs {
        global
    comment "Biology Data Collection filesystem"
        ##options compress
        compress client fast
        priority low
        program "GNUTAR"
}

define dumptype gnutar-nocomp-data-fs {
        global
    comment "Biology Data Collection filesystem"
        ##options compress
        compress none
        priority low
```

```
        program "GNUTAR"
}

define dumptype noidx-data-fs {
        global
      comment "Biology Data Collection filesystem"
        ##options compress
        compress client fast
        priority low
        index no
}

define dumptype nocomp-data-fs {
        global
      comment "Biology Data Collection filesystem, no compression"
        ##options no-compress
        compress none
        priority low
}

define dumptype test {
        global
        comment "Biology test configuration"
        ##options no-record
        record no
        priority low
}
```

Appendix 2 – Example AMANDA Error Report

Amanda Error Report

```
Amanda Tape Server Host Check
-----------------------------
/export/amanda: 41374449 KB disk space available, that's plenty.
ERROR: /dev/rmt/0hn: no tape online.
     (expecting tape VOL005 or a new tape)
NOTE: skipping tape-writable test.
Server check took 0.311 seconds.

Amanda Backup Client Hosts Check
--------------------------------
Client check: 94 hosts checked in 68.061 seconds, 0 problems found.
```

(brought to you by Amanda 2.4.1p1)

Amand Email Report
These dumps were to tape VOL011.
Tonight's dumps should go onto 1 tape: VOL012.

FAILURE AND STRANGE DUMP SUMMARY:
  laevo   / lev 1 STRANGE

STATISTICS:

	Total	Full	Daily
Dump Time (hrs:min)	5:03	2:14	0:56 (0:23 start, 1:30 idle)
Output Size (meg)	16547.8	13714.3	2833.5
Original Size (meg)	29915.3	23739.7	6175.6
Avg Compressed Size (%)	44.0	44.3	43.1
Tape Used (%)	49.4	40.9	8.5 (level:#disks ...)
Filesystems Dumped	284	28	256 (1:236 2:10 3:8 4:2)
Avg Dump Rate (k/s)	180.7	225.6	92.1
Avg Tp Write Rate (k/s)	1482.4	1746.3	856.3

FAILED AND STRANGE DUMP DETAILS:

```
/-- laevo      / lev 1 STRANGE
sendbackup: start [laevo:/ level 1]
sendbackup: info BACKUP=/bin/gtar
sendbackup: info RECOVER_CMD=/usr/bin/gzip -dc |/bin/gtar -f... -
sendbackup: info COMPRESS_SUFFIX=.gz
sendbackup: info end
? gtar: ./dev/log: socket ignored
| Total bytes written: 12912640 (12MB, 841kB/s)
sendbackup: size 12610
sendbackup: end
\--------
```

NOTES:
  planner: Incremental of leger:dks0d1s6 bumped to level 3.
  planner: Incremental of io:c0t1d0s7 bumped to level 2.
  taper: tape VOL011 kb 16954048 fm 284 [OK]

DUMP SUMMARY:

HOSTNAME	DISK	L	DUMPER STATS ORIG-KB	OUT-KB	COMP%	MMM:SS	KB/s	TAPER STATS MMM:SS	KB/s
3dem	dks0d2s7	1	32	32	--	0:09	3.7	0:02	26.7
3dem	dks1d1s7	1	32	32	--	0:01	23.7	0:02	26.6
3dem	dks1d2s7	1	32	32	--	0:02	18.1	0:02	26.7
3dem	dks1d3s7	1	18885	6144	32.5	0:34	183.1	0:08	795.8
3dem	root	1	6829	576	8.4	1:18	7.3	0:04	164.1
absaroka	dks0d2s7	1	3808	3808	--	0:20	186.8	0:07	559.3
absaroka	dks1d1s6	2	2624	2624	--	0:16	159.7	0:04	611.8
absaroka	root	1	4288	416	9.7	0:48	8.7	0:03	178.1
alaska	hd2	1	1622	160	9.9	0:41	3.9	0:02	78.8
arizona	hd4	1	715	128	17.9	0:15	8.8	0:02	67.0
arizona	hd9var	0	13420	2208	16.5	0:53	42.0	0:04	555.6

(brought to you by Amanda version 2.4.1p1)

Error Message for No Tape

*** A TAPE ERROR OCCURRED: [no tape online].
*** PERFORMED ALL DUMPS TO HOLDING DISK.

THESE DUMPS WERE TO DISK. Flush them onto tape VOL005 or a new tape.
Tonight's dumps should go onto 1 tape: VOL006.

FAILURE AND STRANGE DUMP SUMMARY:
laevo / lev 1 STRANGE

STATISTICS:

	Total	Full	Daily
Dump Time (hrs:min)	3:25	0:00	0:00 (0:29 start, 2:57 idle)
Output Size (meg)	3637.2	0.0	3637.2
Original Size (meg)	7571.7	0.0	7571.7
Avg Compressed Size (%)	42.4	--	42.4
Tape Used (%)	10.9	0.0	10.9 (level:#disks ...)
Filesystems Dumped	284	0	284 (1:242 2:15 3:14 4:8 5:2 6:3)
Avg Dump Rate (k/s)	121.2	--	121.2

Avg Tp Write Rate (k/s) -- -- --

FAILED AND STRANGE DUMP DETAILS:

/-- laevo / lev 1 STRANGE
sendbackup: start [laevo:/ level 1]
sendbackup: info BACKUP=/bin/gtar
sendbackup: info RECOVER_CMD=/usr/bin/gzip -dc |/bin/gtar -f... -
sendbackup: info COMPRESS_SUFFIX=.gz
sendbackup: info end
? gtar: ./dev/log: socket ignored
| Total bytes written: 12697600 (12MB, 886kB/s)
sendbackup: size 12400
sendbackup: end
\--------

NOTES:
planner: Incremental of tennessee:dks0d2s7 bumped to level 3.
planner: Full dump of holly:dks1d1s0 promoted from 1 day ahead.
planner: Full dump of geffen:root promoted from 1 day ahead.

DUMP SUMMARY:

			DUMPER STATS				TAPER STATS	
HOSTNAME	DISK	L	ORIG-KB	OUT-KB	COMP%	MMM:SS	KB/s	MMM:SS KB/s
3dem	dks0d2s7	1	32	32	--	0:09	3.5	N/A N/A
3dem	dks1d1s7	1	32	32	--	0:02	21.0	N/A N/A
3dem	dks1d2s7	1	32	32	--	0:01	25.5	N/A N/A
3dem	dks1d3s7	1	11769	4160	35.3	0:33	124.9	N/A N/A
3dem	root	1	4229	448	10.6	1:13	6.1	N/A N/A
absaroka	dks0d2s7	1	6048	6048	--	0:13	479.4	N/A N/A
absaroka	dks1d1s6	2	2624	2624	--	0:13	206.7	N/A N/A
absaroka	root	1	3392	288	8.5	0:44	6.6	N/A N/A
alaska	hd2	2	1622	160	9.9	0:41	3.9	N/A N/A
alaska	hd4	1	3056	448	14.7	0:21	21.3	N/A N/A
alaska	hd9var	1	1540	256	16.6	0:11	22.5	N/A N/A
arizona	hd2	1	4213	960	22.8	0:59	16.2	N/A N/A
arizona	hd4	1	1012	160	15.8	0:16	10.1	N/A N/A
ewald	/usr/ewald3	1	300	32	10.7	0:05	5.8	N/A N/A
virion05	/usr	1	779	160	20.5	0:18	9.0	N/A N/A
virion06	/	1	20342	6048	29.7	0:24	250.8	N/A N/A

virion06	/home	1	257	32	12.5	0:04	7.2	N/A	N/A
virion06	/usr	1	779	160	20.5	0:18	9.0	N/A	N/A
virion07	/	1	21127	6144	29.1	0:13	465.3	N/A	N/A
virion07	/home	1	257	32	12.5	0:04	7.7	N/A	N/A

(brought to you by Amanda version 2.4.1p1)