

Zmanda: Open Source Backup

Chander Kant (ck@zmanda.com)

CEO

Zmanda, Inc.

www.zmanda.com



Agenda

- A Case for Open Source Backup
- Introduction to Amanda
- Amanda Enterprise Edition
 - Backup / Recovery / Monitoring / Reporting / Verification / Security
- Zmanda Recover Manager (ZRM) for MySQL
- Zmanda

A Case for Open Source Backup

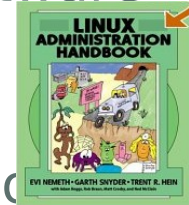
- All open source benefits apply to backup
 - High degree of freedom and flexibility
 - High quality code
 - Security
 - Significantly lower cost vs. proprietary solutions
 - Help and expertise provided by community
- Standard formats play especially important role for backup
 - No vendor lock-in
 - Recoverability of data from one platform to another
 - Protection for legacy and niche operating systems not served by proprietary vendors
 - Point other applications to backup archives (e-Discovery etc.)

Locking up your backup data in proprietary formats comes with a **huge** cost and pain

Amanda: Open Source Network Backup



- Most popular open source backup and archiving software
 - 2005 Linux Journal Readers' Choice Award for "Favorite Backup System"
- Developed at the University of Maryland. In public domain since 1992
- More than 500,000+ deployments around the world
- Distributed with Linux, Solaris, BSD distributions and live CDs



KNOPPIX

opensolaris™



- The only open source backup and recovery software with enterprise support

Amanda: Very Active Development Community



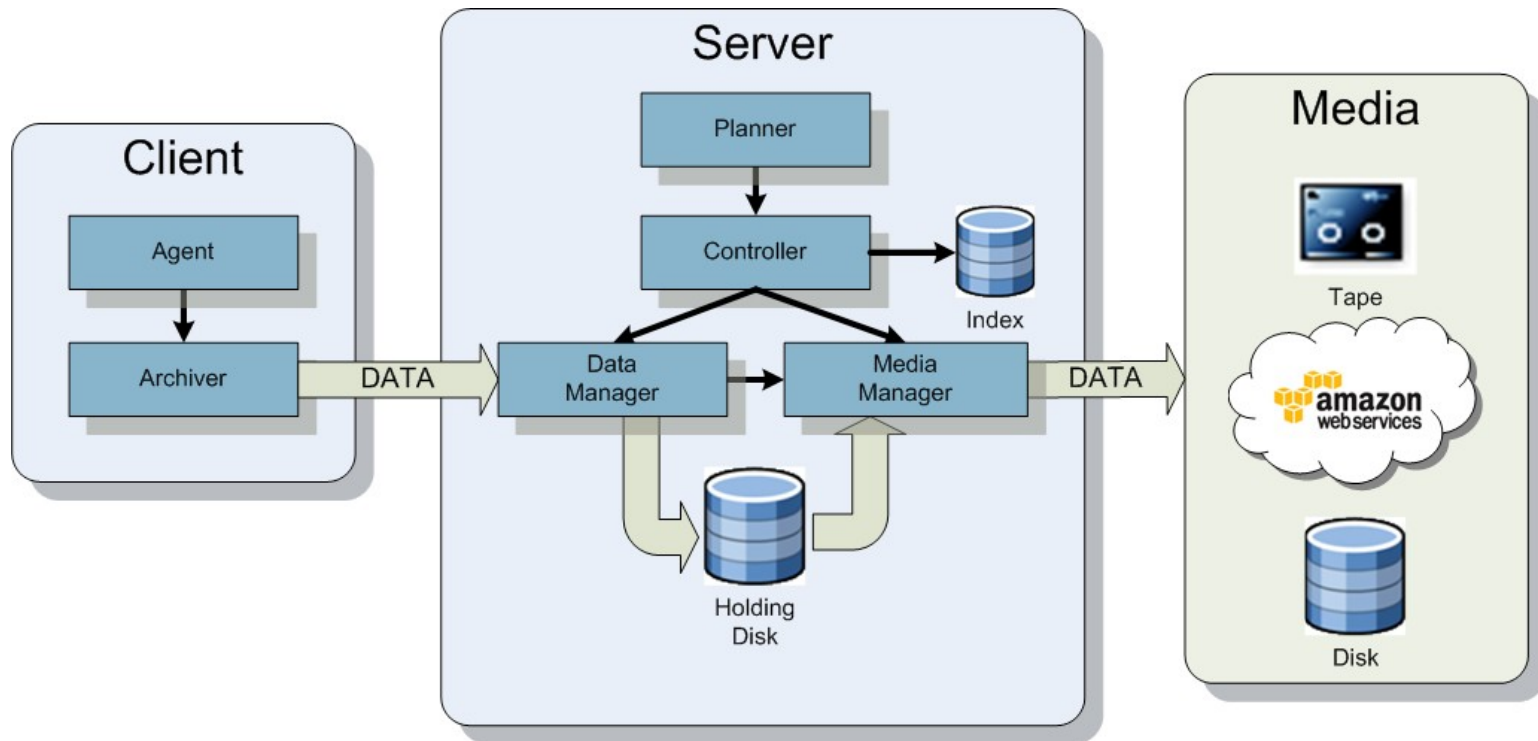
- Active SourceForge project with hundreds of developers and open development process
 - Project registered in '99 and has 99% activity rank
- Current production release: 2.6.0p1
 - Included in RHEL 5.2: Amanda 2.5.0p2
- Up-to-date documentation written by users for users on Amanda wiki <http://wiki.zmanda.com> and user-to-user help on forums <http://forums.zmanda.com/>
- Part of Coverity's open source quality project
 - Currently there are zero defects



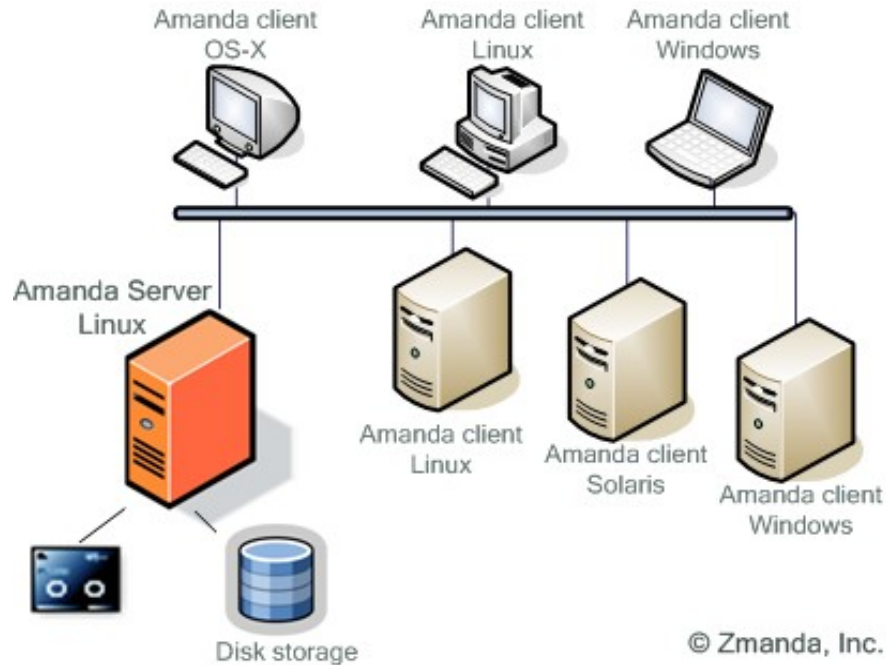
Amanda Features

- Client-Server architecture
 - Single server can backup multiple hosts to various media.
- Robust & Scalable (up to 1000s of client file systems)
- Uses native platform tools
 - Open and standards based data formats on the media.
- Well defined media interface
 - No media device driver dependencies.
- Provides most of the functionality provided by large proprietary network backup software such as Veritas NetBackup and EMC/Legato NetWorker.

Amanda Architecture



Amanda Configuration



- Operating systems support for Amanda server:
 - Linux and Solaris
- Operating systems support for Amanda clients:
 - Linux
 - Mac OS X
 - UNIXes
 - Windows

All configurations are done on the server

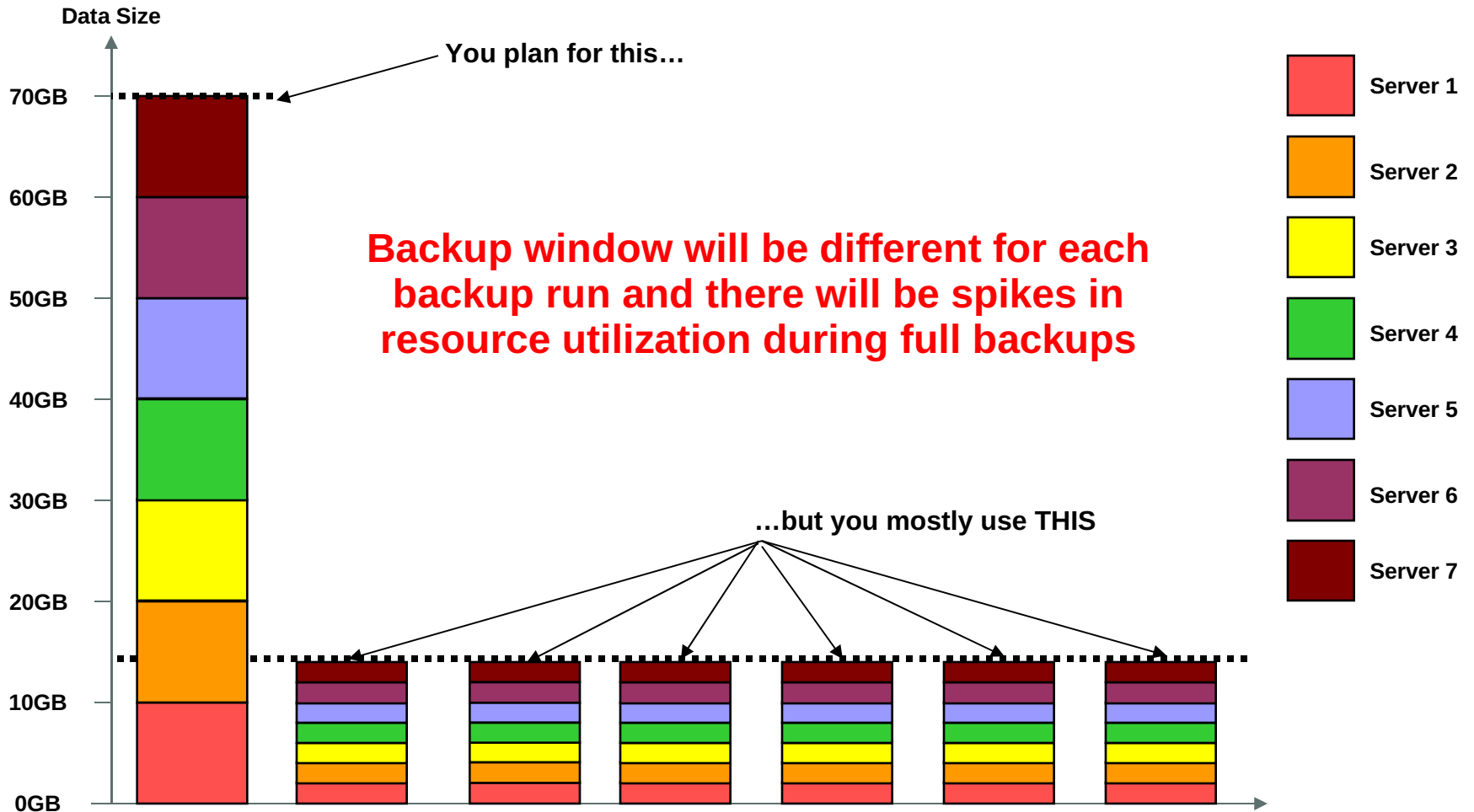
Amanda: Unique Features

- Automatic backup level selection
 - Consistent backup window.
 - Supports multiple backup levels.
 - Provides overrides for backup levels.
- Uses native backup/restore tools
 - Supports filesystem backup tools, GNU tar commands.
 - Uses the same format on the media.
 - Restores from the backup media can be done without using Amanda tools.

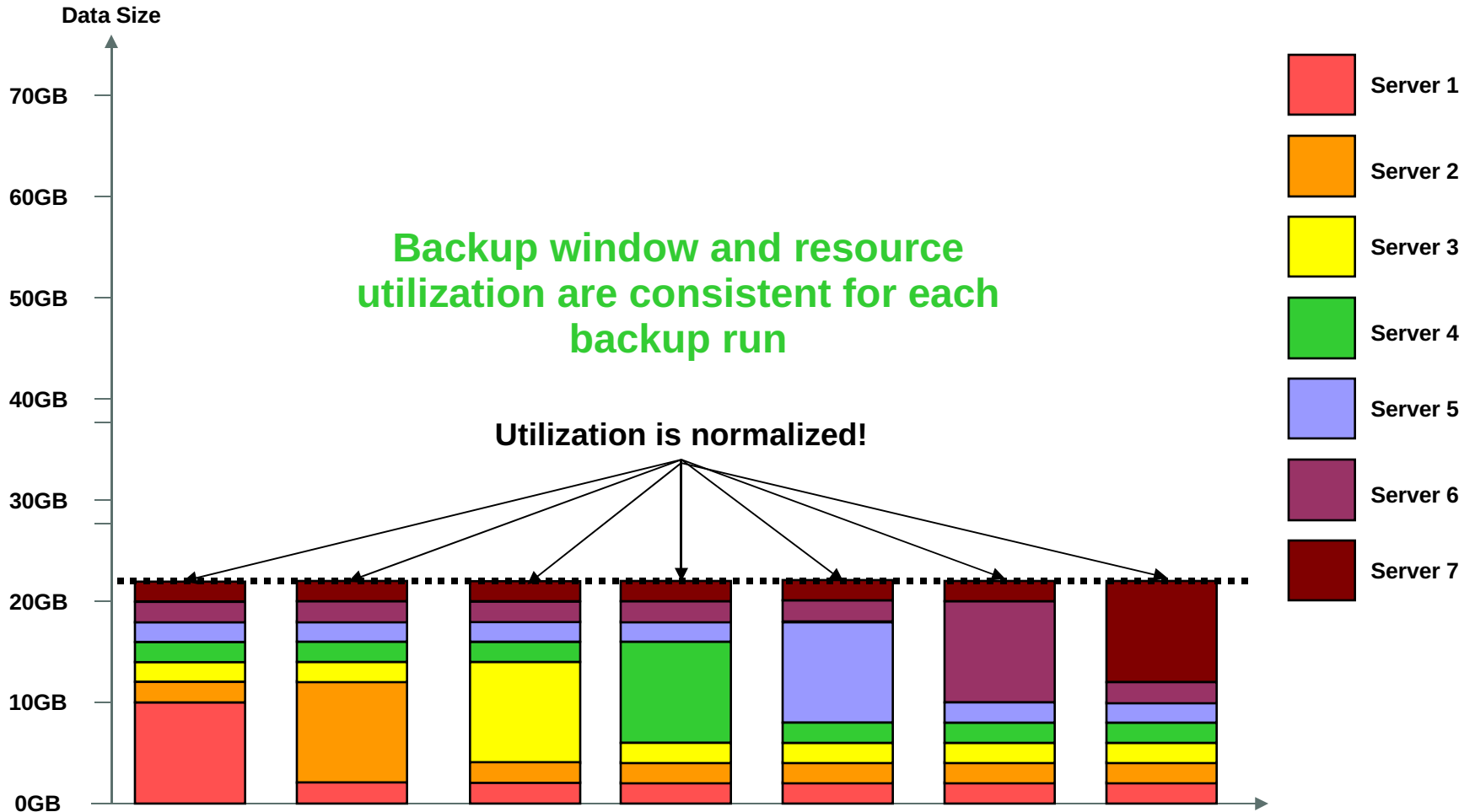
Backup to disk in Amanda

- D in Amanda is for disk
- **vtapes** as files on any type of disk
 - Local disk
 - RAID
 - NAS
- **RAIT** – Redundant Array of Inexpensive Tapes
 - Data is striped over multiple tapes
 - Tapes or **vtapes** could be used
 - Ability to write the very same data to tape (for off-site) and disk (for on-site) at the same time is unique functionality

Traditional Backup Scheduling



Amanda's Intelligent Backup Scheduler



Amanda intelligent scheduler vs. traditional approach



Traditional

- Backup administrator specifies which backup level should be used for each filesystem for each day.
- Administrator has to estimate the rate of change for each filesystem.
- Must have sufficient media space available to backup all data in the backup run (backup job).
- The amount of media required for backup will be different for each backup run.
- **Backup window will be different for each backup run and there will be spikes in resource utilization during full backups.**

Amanda intelligent scheduler

- Amanda will compute an optimized schedule based on amount of data changed for each client.
- Administrators specify a few parameters, e.g. how often to do a full backup.
- Distributes full backups over the backup cycle to balance the amount of data taking into consideration compression and availability of media.
- **Backup window and resource utilization are consistent for each backup run.**

Backup fault tolerance

- Skips the clients (such as laptops) that could not be backed up during the scheduled run.
- Backup media errors
 - Backed up data is kept in holding disk.
 - Holding disk backup data can be flushed when the media problem is resolved.
 - Does incremental backups to conserve holding disk space.
- Operations re-try in case of Client - Server communication timeouts.

Data recovery is simple and robust



- Amanda stores all media and backup information in a database that can be browsed using Amanda commands.
- Data can be recovered to any Amanda client.
- Data in the media is stored in native format and can be recovered using OS tools without Amanda commands.
 - Backup images spanning multiple media can also be restored using OS tools.
 - In the beginning of each tape we provide an exact command how to recover data

```
# mt -f <media_dev> fsf 1
```

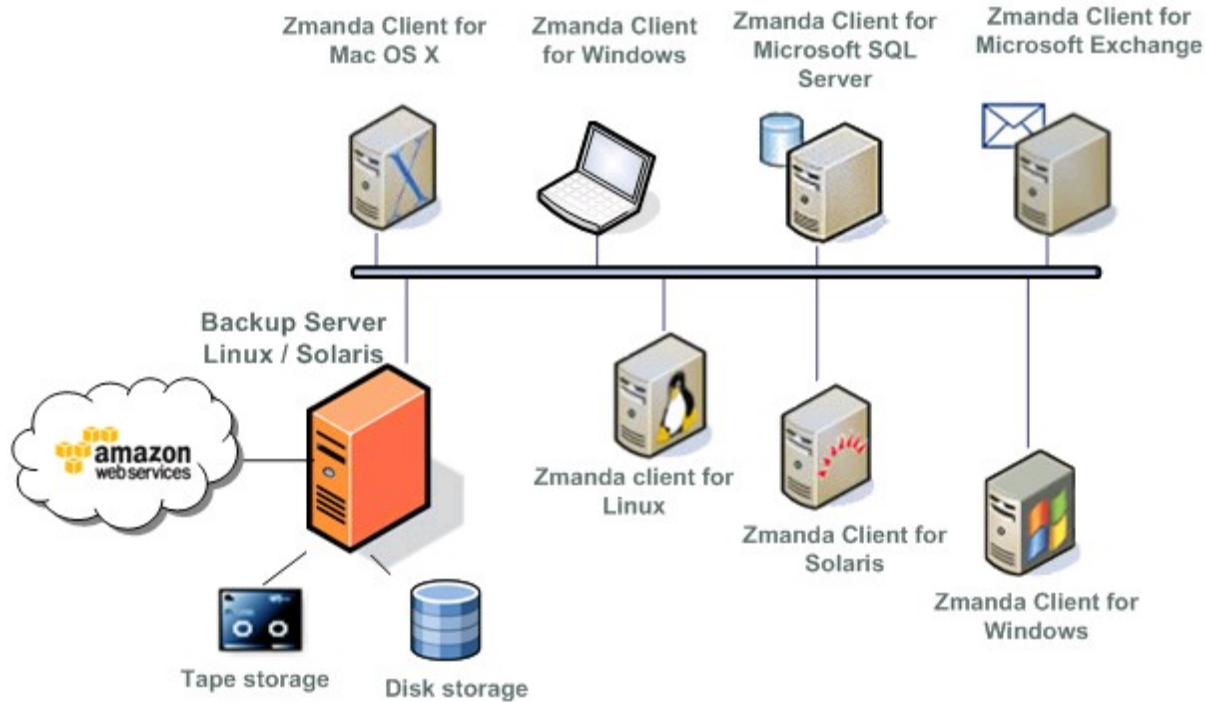
```
# dd if=<media_dev> bs=32k count=1
```

```
AMANDA: FILE 20060228 natasha /boot lev 1 comp N program /bin/gtar
```

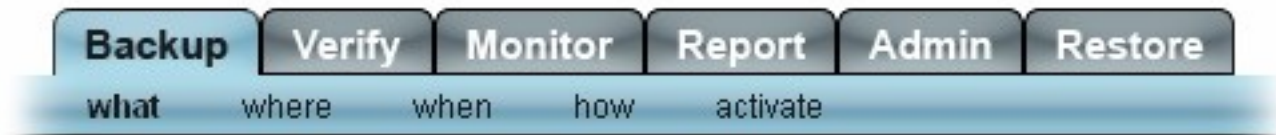
To restore, position tape at start of file and run:

```
dd if=<tape> bs=32k skip=1 | /bin/gtar -f... -
```

Amanda Enterprise Edition from Zmanda



Zmanda Management Console



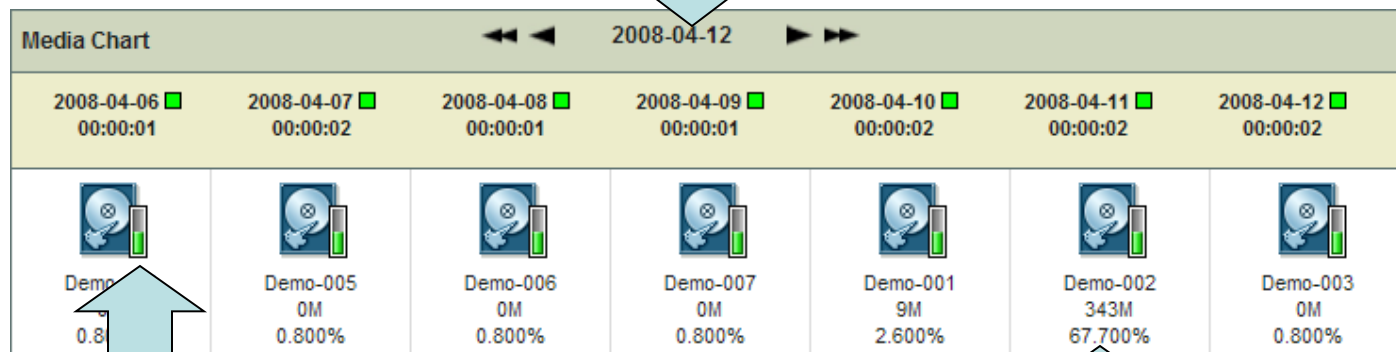
Reporting: Backup Calendar



- At-a-glance view of backup history
- Visual indicators of success, failure, or warning
- Interactive calendar
- Reports available online and delivered via email

Reporting: Backup Media

Week
at-a-glance



Media
utilization
gauge

Data size and
media use
percentage

Reporting: Backup Clients



Week
at-a-glance

Data Chart << 2008-04-12 >>

HostName	Directory	2008-04-08 00:00:01	2008-04-09 00:00:01	2008-04-10 00:00:02	2008-04-11 00:00:02	2008-04-12 00:00:02
client-linux	/home/demo					
client-win	s and Settings/Demo					

Name and
filesystem of
DLE

Backup
level
indicator

Backup Monitoring



Watch backups as they run...

Monitor Chart 12/12/2007 15:52:17

HostName	Directory	Level	Flush	Estimate	Holding	Media
ken-xp-vm	...uments and Settings					
rhel4-vm	/home					

- Live view of backup progress
- Data path visualization
- Backup run status on mouse-over
- Stored in report history

Backup Verification

- Amanda can validate backup data on the media.
 - ***Amverify*** reads an Amanda format tape and makes sure each backup image can be processed by ***amrestore***
- Amanda Enterprise provides easy tools to verify server and client configurations and media
 - Allows administrators to fix problems that can happen during backup run

Secure backup with efficient compression



- Data can be compressed on the client or server:
 - gzip
 - bzip2
 - custom compression algorithms
- Compression algorithm based on type of data being backed up. Different clients can use different compression
- Encrypted backups (symmetric and asymmetric key encryption).
 - Client/Server
 - SSL encryption (takes advantage of hardware support)
- Server/Client communication encryption & authentication.
 - OpenSSH, Kerberos
- Works in SE Linux environment (Strict Policy)

Practical Backup Security

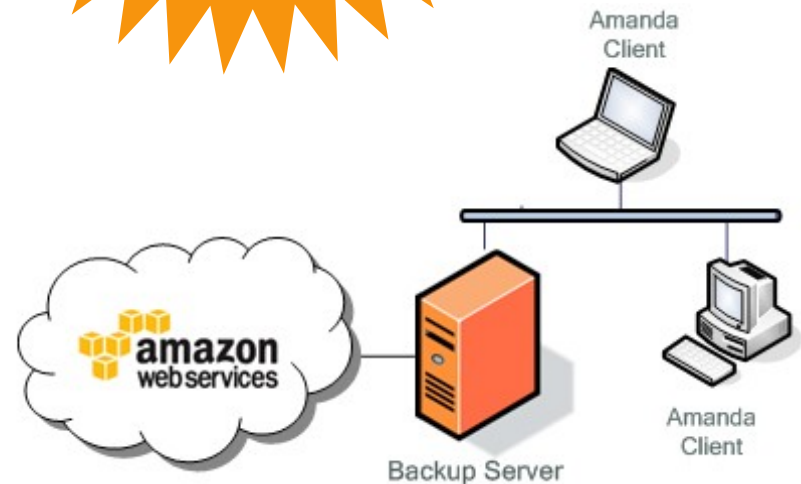


- Amanda Enterprise already separates privileges
 - Runs as a separate user (amandabackup)
 - Only elevates permissions when necessary
- The Zmanda Management Console adds another layer
 - Administrators have full control of the console
 - Operators only act on backup sets
- Powerful role-based access controls (RBAC's)
 - Separation of administrator from operator
 - Operators only see what they own
 - Allows for sharing of resources while keeping data security

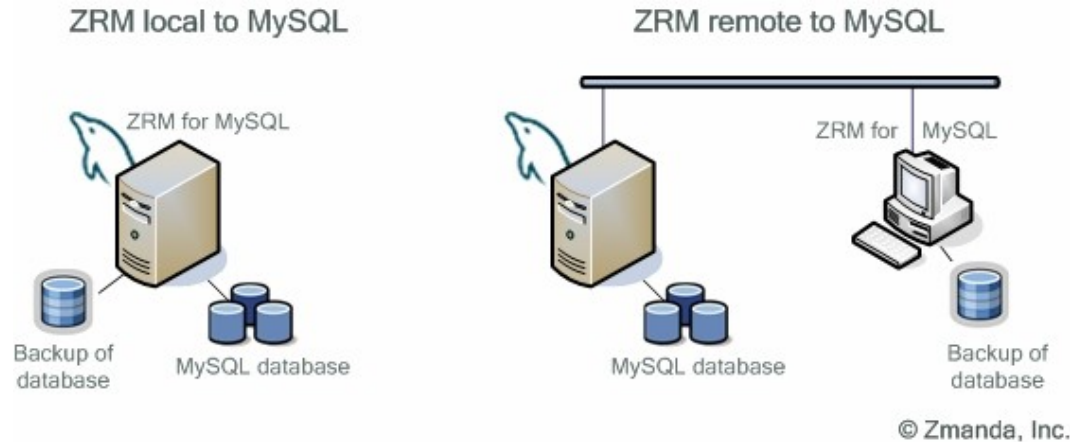
Zmanda Internet Backup



- Utilizes Amazon S3 to provides scalable, reliable, fast, and inexpensive data storage
- Uses strong authentication to ensure that data is kept secure
- Data cached on Amanda Backup Server



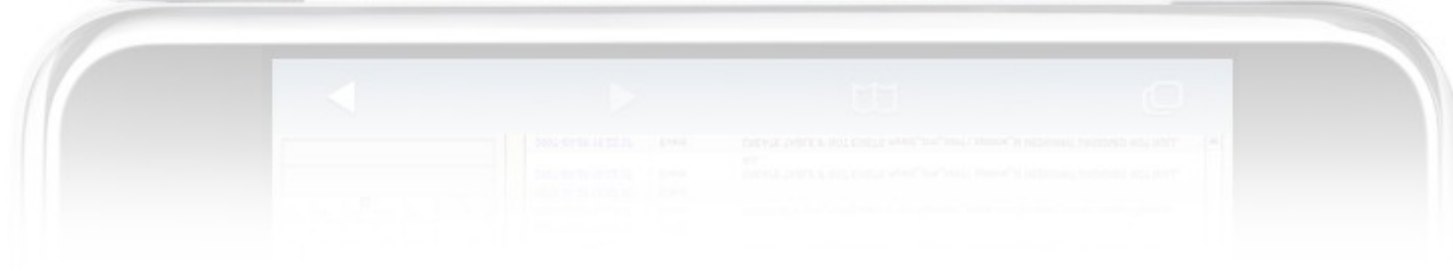
MySQL Backup: Zmanda Recovery Manager for MySQL



- Recover database easily to any point-in-time
- Centralized global backup management
- Schedule full and incremental, logical or raw, backups of MySQL database
- Perform backup that is the best match for the specific MySQL configuration
- Blazing fast snapshot-based backup
- Reporting and Monitoring
- Continuous Data Protection (CDP) for MySQL

<http://www.mysql.com/zrm>

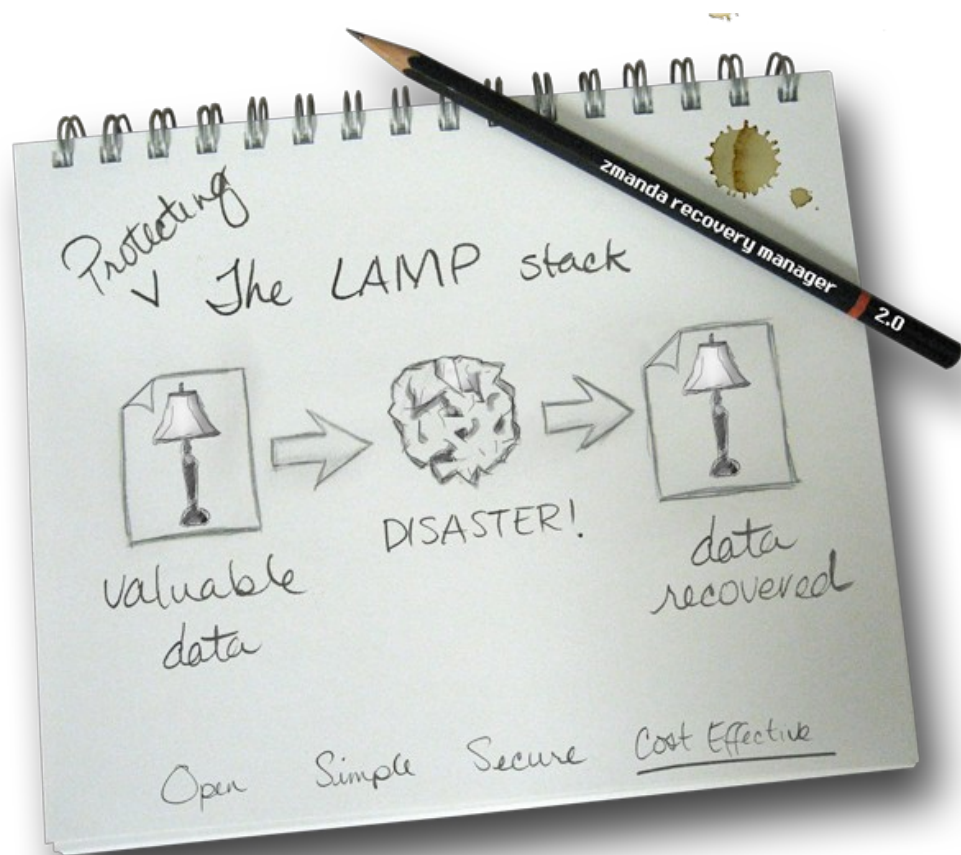
Anytime, Anywhere



Zmanda and the LAMP Stack



- Amanda is the leading backup and recovery solution for Linux file systems
- ZRM for MySQL is built to address backup requirements of DBAs
- Use LAMP in our own products



Protected By Zmanda



Manufacturing & Services



Federal & Government



Web and Media



Research & Education



Telecom and ISV



More information

- Packaged Amanda for downloading:
<http://www.zmanda.com/downloads.html>
- Amanda Forums:
<http://forums.zmanda.com/>
- Amanda Documentation:
<http://wiki.zmanda.com>
- ZRM-MySQL Documentation
<http://mysqlbackup.zmanda.com/>

Backup and restore with confidence...

- Simple installation
- Web-based interface
- Highly flexible and configurable
- Modern and modular security
- Subscription pricing model
- Open media and device formats
- ***ABSOLUTELY NO*** vendor lock-in

