



Go further, faster™

# High End Backup Environments

Peter Buschman  
Data Protection Consultant  
NetApp

Bacula Conference, Cologne  
September 21, 2011





# Agenda

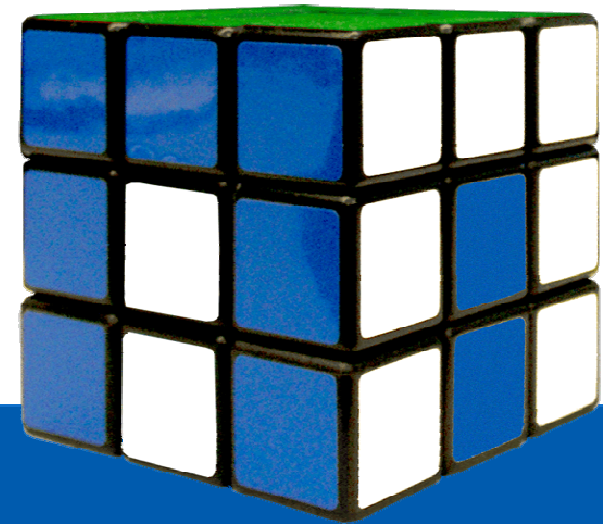
- Who am I?
- What is high-end?
- Lies my storage vendor told me
- Observations and industry trends
- Conclusions



# Who am I?

- Just another IT worker / techie / hacker / geek
- Grew up in the U.S.
- Living in Frankfurt, Germany since 2000
- Working in Unix and data backup since 1995
- Bacula user since 2006
- Consultant with NetApp since 2008
- Focus on data protection in very large environments

## What is High-End?





# What is High-End?

You might be high-end if...

- You backup more than a petabyte of data every month
- The number of systems you backup is in the thousands
- You have dedicated staff called backup administrators
- You have tape libraries big enough to walk around inside
- You know what tape drive streaming is and lose sleep worrying about it.
- You worry about backup windows and don't mean a Microsoft product.
- You backup to primary and secondary locations for redundancy
- Your backups have backups
- Your data is growing faster than your storage budget
- You have one of every client operating system your backup vendor supports (and several they don't).
- Your storage salesperson bought a new Porsche after your last order.
- You're not (yet) using Bacula (that you know of).



# What is High-End?

But what we are effectively saying is...

- High-end backup environments are big because the organizations they serve are big.
- Size brings complexity. You have more of everything.
- Complexity means more things that can break.
- Size means more bytes to backup. You need more space, more throughput, and more resources to do it all.
- Ensuring successful backups requires being an expert in all levels of modern IT infrastructure: storage, operating systems, networks, and especially applications. Each application has its own unique challenges (SAP, Oracle, Exchange) that must be addressed if a backup is to be recoverable.



# What does this have to do with Bacula?

A lot can be learned by surveying the competitive landscape and studying how other backup applications work at a large scale...

- The product with the largest marketshare still uses a flat-file database as a catalog.
- Filesystem backups (as opposed to plugins) still make-up the bulk of every backup environment, large or small.
- Resource sharing and load balancing is very different when you are running 10,000 backups a night vs. 100 backups a night.
- No vendor has the management problem completely solved. A lot of scripting is still required to automate routine tasks.
- At the end of the day, optimizing a backup system for a large environment has more to do with selecting the right hardware and operating system and configuring them properly than it does with the software in use. Backup applications are just big data movers.



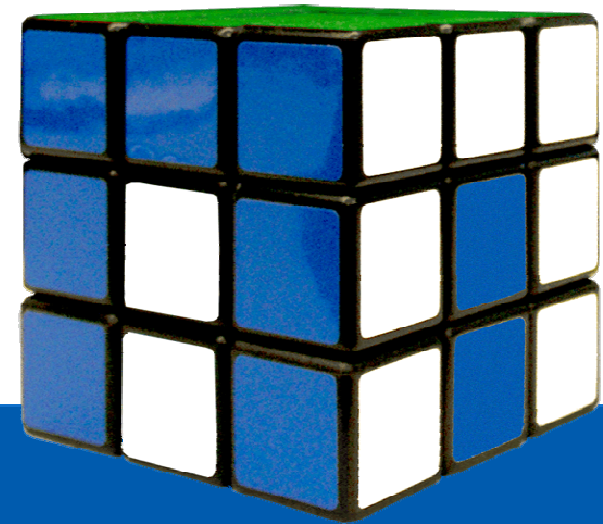
# Backup multiplies your data

- Each backup of a file that I make takes-up additional storage behind the scenes. The more versions I retain, the greater my storage requirements.
- This isn't of much concern when you don't have much data but there are real costs associated with each copy when you get into the petabytes.





**Lies my storage vendor  
told me**



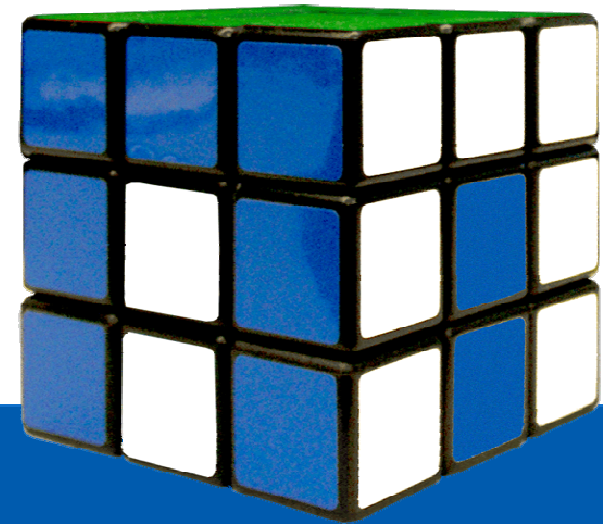


# Lies my storage vendor told me

When was the last time you heard...

- Tape is dead.
- Disk storage makes sense for storing long-term backup data.
- Deduplication is the answer to all backup storage problems.
- High dedupe ratios (> 20:1) are desirable.
- Backups aren't necessary if you take snapshots of your data.
- Virtualization will make backups easier.
- Our solution is so easy it requires no administration.

## Observations and industry trends





# Observations and industry trends

Enterprise backup environments are evolving in some ways that benefit Bacula...

- Linux is routinely replacing other Unix platforms as the operating system of choice for backup servers.
- Linux-based storage appliances are very wide-spread and accepted (Data Domain, Quantum, even NetApp..)
- High availability of the backup system (director-equivalent) is increasingly done using virtualization. The combination of a virtual director plus physical storage daemons is not unusual.
- High inertia / resistance to total change in the environment produces increased tolerance of point solutions where they solve a particularly painful problem.
- No signs that the market is consolidating around any single backup and recovery software vendor. If anything, the number of large players has increased from 3 to around 6 with many smaller vendors and marketshare more evenly distributed.



# Observations and industry trends

In other ways, they are not evolving at all...

- Traditional full / incremental backup concepts continue to be the bread-and-butter of even the largest backup environments. Despite predictions to the contrary, these methodologies will continue to be used for a long time to come.
- Automation still largely done with command-line scripting.
- Capacity planning still done with Excel.
- Throwing hardware and consulting at the problem is still the most common solution to major problems.
- License costs remain high even for functionality that has long-since become a commodity.
- The rule „all backup products suck except the one you happen to know best“ continues to apply.
- The more things change, the more they stay the same...



# Conclusions

- There are no technical reasons why Bacula cannot evolve in similar ways that competing frameworks have. They share a similar architecture and added functionality incrementally as they grew over time.
- In some ways, Bacula has a significant development advantage. On the server side (director, storage daemon), it can take advantage of the richness of the Linux ecosystem without having to worry about maintaining the same breadth of compatibility that legacy vendors do. Given the amount of development taking place on Linux in the areas of databases (catalog), filesystems (deduplication, replication), and performance, this means Bacula can develop new features rapidly.
- Bacula has a bright future. From a small start, it has already shown tremendous scale. Although not yet able to handle the largest types of environments described here, Bacula has shown it is the best-of-breed backup solution in the open-source world.



# Thank You

[peter.buschman@netapp.com](mailto:peter.buschman@netapp.com)

+49 151 527 555 24

