

Experiences with backup in the Petabyte range

Thomas Eifert, Daniel Stanek

RWTH Aachen University, Center for Computing and Communication

1. Backup Scales

RWTH Aachen University is among the four largest technically oriented University in Germany and the largest in the State of North-Rhine Westphalia. Almost consequently, the amount of scientific data at the institutes and departments as well as at the Centre of Computing and Communication has reached the Petabyte range since a while. To handle a backup of this data which change with a rate of about 4 TB per day, our installation contains a 16-frame 3494 tape library and 18 “Jaguar” drives.

On the other hand, the University of Paderborn (UPB) has an estimated backup data of about 30 TB in total, an amount where a local TSM team won't operate at a competitive price, compared to commercial service providers.

UPB decided to outsource the backup service, but not to a commercial provider but to RWTH Aachen University.

RWTH already has some experiences in this field: Some other universities do TSM-Server-to-server backup to our server for enhanced data protection, University of Applied Sciences (FH) Aachen has no “own” backup at all. Instead, their clients are connected to the RWTH server as well as from our institutes – no problem due to direct fibre connection. However, the primary backup service over a distance of about 250km was something new.

2. The configuration

We decided to place a TSM server with disks as the primary backup server at the compute center of UPB. By this, we have a low-latency connection between the clients and “their” server, no matter of WAN availability or performance. This server then copies its data to the tape library at RWTH by TSM-Server-to-server backup, thereby realizing an offsite backup.

The operation of the TSM server itself is managed remotely by the RWTH team while handling the clients is done by the helpdesk at UPB. This distribution ensures that end users still can have personal contact to a supporter.

3. The Experiences

The new backup server just went into service, but at a first glance the users at UPB experience backup and restore at a completely new performance level. Since they still are in contact primarily with the support people at UPB, virtually nothing has changed.

However, we now monitor a strong oscillation in WAN performance which is currently under investigation.

4. Extrapolation to local configuration

For the backup at RWTH we run with collocation enabled. For a large number of disaster restores where a whole local server's data was lost collocation was the way to restore data at terabyte level within a few hours (which would have been impossible with data spread across many tapes). However, this was feasible at a time where mainly the departmental servers were clients to the backup system. In the meantime, quite a large number of PCs has become backup nodes. One problem is the rapid growth of allocated tapes which fill up the library and causes many mounts during backup. The other problem is that with the amount of data from such a client, the advantage in media costs of tape over disk is lost. This issue will become even more urging with our next procurement when most probably TeraByte tape cartridges will be available.

The solution of collocation groups didn't seem attractive to us since the grouping of clients is not only dependant of the client itself but also of its "colleagues" belonging to the same collocation group.

Due to the success of the configuration at UPB we considered backup to disk, thus getting rid of the need for collocation, but a disk pool of about 600 TB isn't currently feasible at our site.

5. Classification of Clients

A look at the distribution of client sizes revealed a very large number of very small clients (with respect of their amount of backup data) and a quite small number of large clients. With this distribution of client sizes we are now able to shift many clients from tape to disk while having a need of a quite small amount of disk space. Currently, we are just trying this idea within our current installation, but due to the procurement in spring 2008 these experiences are necessary for precise specifications.