

TSM

between the Skylla and the Charybdis

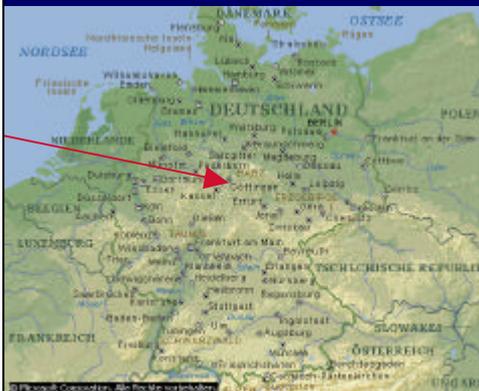
Gerhard Schneider

Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen
and University of Göttingen

September 20, 2001

gerhard.schneider@gwdg.de
<http://www.gwdg.de/~gshnei2>

Göttingen



1837: World's first data highway (1 bit/s)



Göttingen's Data Network for Science and Research

GWGD

3

The local high speed network GöNet, started in 1991, connects the following institutions at gigabit speed:

- **University of Göttingen**
 - URL: <http://www.uni-goettingen.de>
- **Max Planck Institute for Aeronomy**
(in Katlenburg-Lindau, some 30km outside of Göttingen – dark fibre)
 - URL: <http://www.mpae.gwdg.de>
- **Max Planck Institute for Biophysical Chemistry**
 - URL: <http://www.mpibpc.gwdg.de>
- **Max Planck Institute for fluid Dynamics**
 - URL: <http://www.gwdg.de/mpisf>
- **Institute for the scientific Film**
 - URL: <http://www.ifw.de>
- **Max Planck Institute for Historical Science**
 - URL: <http://www.geschichte.mpg.de>
- **Max Planck Institute for Experimental Medicine**
 - URL: <http://www.mpiem.gwdg.de>
- **GWGD**
 - URL: <http://www.gwdg.de>
- **City of Göttingen**
(since 3/99 – with the local schools)
 - URL: <http://www.goettingen.de>
 - URL: <http://www.goe.ni.schule.de>
- **Primatenzentrum**
 - URL: <http://www.dpz.de>

Carthard Schneider

Skylla

GWGD

4

- **IT is moving ahead – configuration of an average PC (supermarketstandard, summer 2001):**
 - 1,5 GHz, 128 MB, 60 GB for less than 1000,-
- **A typical research group has more than 1 TB of disk space**
 - i.e. just a few PCs, 10-15 say, and a little server
- **The demand for backup capacity and backup speed is growing fast**
- **And restore remains a problem:**
 - It requires the generation of lots of file handles....
 - Speed of restore is inadequate
 - Pumping 100GB through a modern 100 Mbit/s network takes at least 3 hours.
- **There is no point in limiting the backup capacity for a user**
 - In a science environment, users don't care about restrictions
 - Or they react with „No backup“ or homegrown/chaotic backup solutions
 - If there is a crash, the overall cost for a science institution is higher than with unlimited backup capacity

Carthard Schneider

Skylla

5

- **The situation (in Göttingen and elsewhere):**
 - A new 100 GB disk (plain IDE) costs 400,--
 - One 50 GB AIT-2-tape costs 75,--
 - so backup adds 40% - or 80% if you need two tape copies
 - Remember 1997?
 - The backup surcharge was 2-5% in those days using AIT technology
 - (and 19.3% using 3590 technology)
- **This ain't everything....**
 - Overall cost of a tape robot (Grau, 2000 slots), TSM server (IBM H80), AIT tape drives (3000 per drive), TSM software, TSM clients, etc.:
 - Around 150,-- per slot
 - Thus 50 GB AIT-2 storage really cost
- **And our new IDE disk requires an extra**

$$2 \times 75_{,-} + 2 \times 150_{,-} = 450_{,-}$$

Skylla

6

- **Data keeps pouring in...**
- **From experiments and computation – the well known front**
 - Computation is expensive, therefore we use backups extensively
 - Is this really true, given the cost of modern hardware?
 - Disk space is expensive, thus we use hierarchical storage management
 - Is this still true?
 - Why not keep your copies on cheap VOLKSRAID, i.e. low cost IDE-RAID-systems
 - transtec 5000 SCSI/IDE RAID-subsystem, 8 x 76 GB disk = 600 GB raw capacity costs 8000 or 13,50 per GB
 - But recall is instantaneous – compared to tapes
 - GRID computing requires large archives at the local sites
 - We still can't pump TBs over the atlantic because of lacking funds
 - How long will this be true?

Skylla

7

- We have trained our users to use our services – and now they do.
 - Do we feel like the sorcerer's apprentice?



- Will a computer centre end up buying nothing but tapes?
 - Do we want to introduce a „backup levy“ on each device bought by the university?

Acknowledgement - The cartoon on this slide is the copyright of the Disney Corporation

Carthard Schneider

Skylla

8

- Are tapes still appropriate?
 - What about an array of large „switch on when needed“ disks?
 - Also less volume per byte
 - Our ftp-server (ftp.gwdg.de – a big and well known Linux site) is no longer in TSM.

Instead we do our backups onto a second machine with large IDE drives via good old (i.e. old fashioned) *rsync*.
 - In case of failure, the backup machine (although slow, but in a different location) can take over immediately
 - Whereas a restore of several hundred GB and millions of files from tape is an interesting experience
 - No cheat – the ftp-server is a Linux based PC – fast and cheap
 - MTBF works in our favour: disks don't crash, even when hitting the floor
 - Thanks to notebook technology

Carthard Schneider

Charybdis

- **New data keeps pouring in...**



- **There are new devices around:**

Acknowledgement: The cartoon on this slide is the copyright of the Disney Corporation

9

Charybdis

- **mobile storage devices:**
 - USB-based disk drives are attached to different computers – thus they may experience being backed up from multiple machines
 - „machine based“ backup is out of date – backup has to be device based!
 - Thumb drives keep user data – but are they ever backed up?
 - Trigger backup when it is inserted, but don't do multiple backups (consistency)
- **new devices:**
 - digital cameras
 - Or do you want to lose the photos of your kids?
 - More seriously : digital imaging is penetrating science
 - So what about a TSM client for cameras?
 - Cameras are network centric these days (USB, Bluetooth)
 - MP3 players
 - DVD ripping
 - PDAs
 - Now I hotsync but as they become more powerful and independent, they need their TSM clients

10

Charybdis

11

- *New applications:*
 - Digital libraries
 - We don't want to leave our university library out in the rain?
 - Electronic books, scans, etc – and scans are large
 - What is an electronic book?
It requires special software, just keeping the bits is pretty useless.
 - Data bases with multimedia data – the real digital library
 - Only the data base knows the connection between meta data and the files
 - Without the meta data, the files are useless
 - i.e. more care is required than just doing backup/archiving/HSM
 - Knowledge about the internals is required and necessary when moving to a new system
 - Backup of a digital library is more than keeping bits – it has to be kept alive

Charybdis

12

- *New demands:*
 - The discussion about „Fraud in Science“ puts scientists under pressure:
 - Keep your data to prove your innocence
 - An interesting new principle, but apart from that, keeping data is a good idea
 - The Max Planck Society has just issued a guide of conduct:
primary data should/must be kept for at least 10 years
 - A new experience for a computer center:
having to worry about data whose owner is no longer around
 - Try asking for a new tape robot for your 1000 TB of un-owned data
 - The CC is now in charge of the data, like a notary
 - Current setup in Göttingen: data is handed over to a sysadmin who generates a few meta data and archives it under his name
 - This person may retire in 8 years...
 - Clearly this method does not scale well

Charybdis

- **What is a good setup in TSM to make it work as a notary?**
- **Make sure that data cannot be manipulated**
 - By hackers
 - By the author
 - By „loving“ colleagues
- **We need a „one-way“ backup**
 - Once backed up, it cannot be manipulated, not even by the sysadmin
- **How do we (dis-)prove manipulation?**
 - Göttingen setup: an MD5 code is generated for each primary data file and is kept on a separate machine

Charybdis

- ***User based encryption***
 - My plea (since 1995) has partly be heard
 - Necessary for external services (outsourcer) and internal services
 - Administration has to store critical (i.e. personal) data in a safe place
 - As TSM is not legally safe (thanks to sysadmin), they do it on ZIP drives
 - At a cost of 80 per GB...
- ***New concepts like SAN need fast backup and restore***
 - Problem: you don't know what is in your SAN without looking into the server
 - How do we do incremental backups on raw disks? Logging of sectors?
- ***New user groups***
 - Ever thought about your students?
 - Thousands of high end PCs

Outlook

- **We must find proper solutions for these issues, although – unlike the sorcerer’s apprentice – the flow of data will not stop**
- **After all it is a wonderful feeling to see a grateful user whose data has been restored**
 - It might be the future president of your University ☺
 - Test drive our robot: <http://www.gwdg.de/~dross/GWDGBackup.html>