sim3d – General Storage Data Deduplication Analysis

Service Offering – accurately determining actually possible Data Deduplication Ratios

Highlights

- Simulating Data Deduplication

- Determining accurate Ratios, even in complex environments (TSM Progressive Incremental, RMAN etc.)

- No additional Hardware required

- No adaption of backup procedures required

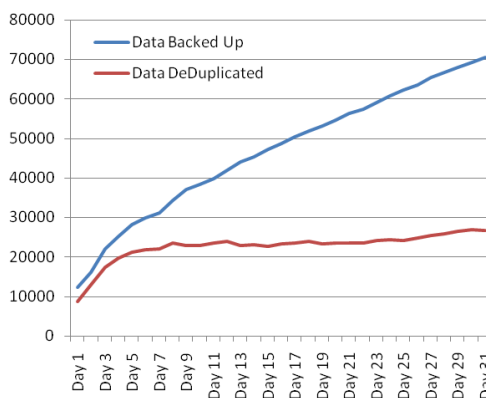- Service Offering incl. Installation, Analysis, Interpretation and Presentation of Results

Simulation Deduplication

Data Deduplication is one of the most heavily hyped technologies in the storage arena today.

"Intelligent" Appliances and Software are supposed to shrink mountains of data, minimize cost and solve performance problems along the way.

On the surface, Data Deduplication is a relatively simple process, which identifies redundant chunks of data and eliminates them. Since data redun-dancies are primarily being caused by backing up, Data Deduplication should, in theory, have the largest effect when replacing traditional backup media.
Depending on the scale of reduction, even large tape libraries could be substituted by relatively modest disk setups and deduplicating appliances at competitive cost levels.

Deduplication Simulation (TSM Server)

Obviously, whether or not the advertized magnitudes of Data Reduction can be reached in the real (your) world, heavily depends on a whole range of factors:

- How often do you back up?
- What do you back up?
- How do you back up?
- How long and in how many versions do you keep backups?

And of course:
- By what factor can your specific data be reduced?

Naturally, misjudging any of these factors by even a small margin may lead to buying appliances magnitudes too small or too big. In order to reduce these risks, General Storage has developed the Data Deduplication Simulation "sim3d".

sim3d is a small utility, running either on your backup clients, the backups servers or on any other machine having read access to the data in question.
According to the frequency of your backups, it scans through data byte by byte, catalogs its structure and summarizes it in a way that allows us to calculate its actual "deduplicatability".

The detailed analysis includes "what if" scenarios, covering different Dedu-plication solutions, effects of variations in versioning and retention and useful details for planning like how you should spread specific data across more than one Deduplication VTL in order to maximize overall reduction.

sim3d is a Service Offering, including the installation of software components, analysis, interpretation and presentations of results. For further details, please contact support@general-storage.de.

Tivoli. software

IBM **Business Partner**