Backup with Data Deduplication Workshop – Tools and Methods for Starting

Madis Pärn
Senior Technology Consultant
EMC
Email: parn_madis@emc.com

Tools

- Avamar Sizing Configurator
- DataDomain Sizer
- Avamar Assessment Tool
- Data Profiling Assessment
- Avamar Virtual Demo (AVD)

The problem

- Prospects show interest in redundant data elimination
- De dupe rate depends on many factors like:
  - Type of data (Avamar has it's sweetspots)
  - Change rate (over time)
  - Application (file system vs. databases)
  - Growth rate
- PoC is one option to find these out
- Downside:
  - Hardware must be invested
  - Professional services required for installation and configuration
  - Time consuming

The solution – Avamar Assessment Tool (avasst)

- Avasst components:
  - Wrapper around avtar
  - Leverages avtar
  - Single toolset archive
    - Supports Windows – extract with Winzip
    - Support Unix/Linux – extract with tar
    - Three (four) components - easy to deploy:
- No Avamar server required
- Non intrusive to an existing Avamar installation
- Non intrusive to an existing NetWorker installation
- No Avamar skills required
- Download:
  - ftp://avamar_ftp:anonymous@ftp.avamar.com/software/avasst/avasst_1.6.tar

Operation

- Installs in seconds
  - Avasst and avtar (and all run time libraries) must be installed in the same directory
- Runs in seconds – single click (at least for Windows)
- Can be scheduled for recurring runs
  - Create file, named “silent”, in avasst directory or in any other directory and call avasst –s <pathname to silent file>
  - Add one single entry per line, e.g path or drive letter
    - Be sure to use a “simple” text editor (word or notepad may scramble characters)
- Full support for avtar.cmd flag file
  - can be used to apply debug option (--debug)
  - can be used to increase cache size
  - can be used to apply Office 2007 document optimization
  - see avtar technical addendum documentation in (Powerlink->Products->Avamar->document library)

Options

- Creates C:\avasst or /avasst (by default)
- Regular mode
  - Two line statistics
    - # of files, overall volume processed , new files, new bytes, percentage of new bytes, number of directories/folders (avasst 1.5)
    - Duration
  - Statistics written to stats.csv
  - Excel import allows for professional reports
    - Direct import may depend on excel version
- Options
  - –d : run in debug/verbose mode
  - -i : clear out existing cache files
  - -w DIRECTORY : do not use avasst directory, but DIRECTORY to store cache files
  - -o: override lockfile created by another avasst instance
  - -v : print version
  - -s PATH to Silent file, including file name: do not use the default „silent“ file, but another file .
  - -h : print help
  - Must use cli
- Can be run in silent mode
  - Create “silent” file in avasst directory or pick your own “silent file” and apply with option –s
  - Scheduling
    - Create task for avasst in Windows Task Manager
    - Create avasst cron job for Unix/Linux
    - Leverage option –s to use separate „silent“ files per client

Supported Environments

EMC² where information lives®

- Tested platforms:
  - RHEL 3 (x86) , RHEL (x64)
  - Suse 8 , 10 (x86)
  - Solaris 9, 10 (Sparc)
  - HP-UX B11.11 (PA-Risc)
  - Windows XP (x86), W2k3 (x86, x64)
  - Windows 2008 (x86, x64)
  - AIX 5.2, 5.3
- Support for other OSes by network mount
  - Microsoft Volume Shadow Copy Service will fail for CIFS/NFS network
  - Workaround for NAS devices -
  - Workaround for Vmware VMDKs
  - Workaround for NAS appliances (NDMP)
- Celerra
  - No VSS support for network backup
- Disclaimer
  - No extensive error checking
  - As is … you cannot call support

Testing

- Backup
  - Initial backup
  - Create new file
    - Verify commonality rate
  - Modify a file
    - Verify commonality rate
  - Move cache files to another server
    - Simulates backup to Avamar server
    - Create and export a network share and invoke avasst –w <network share>
      - Avasst locking will prevent cache file corruption
- First backup will report compression only but no deduplication
  - This is expected behavior, as no root hashes will exist

Avamar Virtual Demo (AVD) 1.0 - Limitations

- Avamar Virtual Demo is designed to provide a basic functional demonstration of Avamar software. The virtual machine is designed to run on a desktop or laptop using VMware Workstation or VMware Player.
- AVD is designed to deliver a single server demonstration
- Avamar maintenance processes have been disabled and data cannot be deleted from the system.
- AVD can store up to 10 GB of unique data before the system enters readonly mode.
- Included with AVD are backup agents for Microsoft Windows

AVD - Requirements

The following are the technical requirements for the AVDT VM.

Type	Requirments
Processor	1 CPU (3.0GHZ or greater)
Memory	1024 MB
Network	1 GB Network Connection
Software	VMware Workstation 6.5 Vmware Player 6.5


EMC²
where information lives®