Storage Analysis Engine Report - what is it?

30th March 2011

Version 1.0

Marcus Witts
Introduction .................................................................................................................. 3

Actual SOURCE Environment ....................................................................................... 4
  Source Hardware and Software Architecture ............................................................. 5
  Source Client Environment ......................................................................................... 5
  Capacity Information .................................................................................................. 5
  Storage Wastage Chart ............................................................................................... 6
  Capacity Metrics ......................................................................................................... 7

Indicative TARGET environment .................................................................................. 8
  Operational Issues Resolved ....................................................................................... 10
  Capacity Metrics ......................................................................................................... 11

Butterfly Differential Business Case ............................................................................. 12

Butterfly Migration Planning ......................................................................................... 13

Summary .......................................................................................................................... 13
Introduction
The Storage AER starts with an assessment of an existing disk storage environment. This assessment is then used to create a replacement target environment using the latest technology. The target environment will support 3 years growth, solve operational issues, have a lower TCO and improve storage utilisation efficiency.

Now let's look at each section in more detail.
Actual SOURCE Environment

We use light-touch agentless software technology - not a 20 day consultative engagement or a heavy weight agent based SRM tool. Using just two touch points; one that retrieves data from the disk arrays and another to retrieve data from backup servers we can create an accurate picture of a customers environment - in both a textual and schematic form.

To the right of the textual source environment information is a schematic representation of the customers environment: such as summary info on the disk arrays and the different structured and non-structured data types detected.

To the left is the textual explanation of the source environment.
Source Hardware and Software Architecture
Information includes disk array make and model, disk capacity and geometry, RAID type in use, remote and local replica information - all of this from a single non-invasive touch point.

Source Hardware Architecture
- MODULAR disk hardware based on HP Enterprise Virtual Array (EVA) disk arrays
- Disk array models in place are EVA 8100 (HSV21) and EVA 8400 (HSV450)
- 2 Disk arrays based in 2 data centre locations
- Microcode version on disk arrays based on 6200 and 9534
- Disk geometry based on HP EVA FATA 1TB 7200 rpm and HP EVA FC 450GB 15000 rpm
- Disk arrays with 4GB-8100 and 6GB-8400 of controller memory
- Array ports used for host attachment are 4Gb FC enabled
- ACTIVE data centre topology
- Snapshots and copy operations in use
- SAN fabric infrastructure in use for storage provision
- DUAL fabric infrastructure deployed for resilience

Source Software Architecture
- HP Storage Works used for array management
- Host multipath software in use for disk access
- ARRAY replication used for data protection
- DR_DESTINATION volumes are used for copy operations
- VRAID1 and VRAIDS in use for data protection

Source Client Environment
Analysis of the backup data is achieved via the other touch point - this gives us the ability to get detailed information on the structured and non-structured data types in the environment.

Source Client Environment
- Client operating system platforms include VMWare, VMS, Windows, Windows 2008
- Structured data types are Exchange 2007, Oracle and MSSQL
- Exchange 2007 running within VMWARE environment
- Boot from SAN capability
- VMS client capability
- Extensive use of clustering and multi-client allocations
- Disaster Recovery is done via HP EVA replication
- Synchronous vdisk replication in use
- Backup and Recovery provided by HP Data Protector

Capacity Information
On this section of the AER we summarize the total RAW capacity of the disk arrays, the actual amount of storage that’s been allocated and provide an insight into how inefficient the existing environment is with the Storage Utilisation percentage. Usually we find that the storage utilisation is less than 30%.

Capacity Information
- Total RAW storage deployed is 197 TB
- Predicted used storage is 53 TB
- Storage utilisation is 27%
- Calculated Growth of RAW capacity is 25%
- Total capacity based on FATA and FC disk technology
Storage Wastage Chart
This is a graphical representation of the wastage in the customers disk storage environment.

RAW - the total unformatted capacity of all disk drives
USABLE - the total usable space after array formatting, disk sparing and RAID overhead has been taken into account
ALLOCATED - total amount of space that’s been allocated to hosts, or used for local or remote replication purposes
DISCOVERED - a proportion of storage isn’t visible to hosts due to zoning issues or poor change management process
FORMATTED - host filesystem overhead
UTILIZED - by combining the two touch points - the backup server data and the array data - we can calculate how much of the host allocated storage is actually in use - remember all this is done without using any host agents
Capacity Metrics
The bar chart at the top of this section shows the top 20 hosts by allocated storage. By using the two touch points we can show how over provisioned these hosts are.

The pie charts in this section provide further insight into the nature of the customers disk infrastructure - we’ve broken down the storage distribution by disk tier, RAID level, state and location.
**Indicative TARGET environment**

First, there’s a schematic depicting the target design - this is bespoke to their actual requirements. The design can be influenced by an engineer making what we call a “target technology decision”. This may be necessary due to an existing sales process that the SAER is being used to reinforce. The software will determine the correct configuration for a given technology decision. There’s a breakdown of the new disk tiering model and capacities associated with each array.
To the right of this schematic is a breakdown of the target hardware and software products used, an explanation of the architectural changes and a concise list of the actual transformation benefits.

It provides the customer with the confidence that the vendor solution is the correct solution.

**Target Hardware Architecture**
- IBM StorWize V7000
  - Consolidated Enterprise class Storage infrastructure
  - Creation of tiered storage model with cost and performance benefits
  - Removal of legacy components and technology
  - **Consolidation & Centralisation** of primary data
  - Improved storage utilisation efficiency to 57%
  - IBM V7000 for tiered storage virtualisation and automated data tiering
  - **Up to 200% performance improvement with automatic migration to SSD**
  - Simplified provisioning and configuration
  - Non-disruptive migration capability
  - Thin provisioning support for dynamic application growth

**Target Software Architecture**
- IBM Systems Director
- IBM System Storage Easy Tier
- Metro/Global Mirror replication
  - **Tivoli Productivity Center Standard Edition**
    - Integrated management capability
    - Best in class array management with simplified provisioning and control
    - Best in class thin provisioning management and visibility
    - Data replication software for disaster recovery and remote replication
    - Automatic migration of frequently accessed data elements

**Architectural Changes**
- **SSD tier for highest performance storage**
- Automated tiering and migration based on application usage
- Simple, powerful management interface
- Thin provisioning for reducing footprint
- **Non disruptive migration of storage capability**

**Transformation Benefits**
- **Commercial, Technical & Operational benefits of a unified strategic storage platform**
- Simplified management and faster storage provisioning
- Cost and environmental benefits of thin provisioning deployment
- **Highest performance capability** with SSD dynamic Easy Tier
- Enterprise class storage availability and performance
Operational Issues Resolved
The software trawls through logs and configuration data to find problem areas - such as offline drive channels and failed disk drives. We also have a constantly updated database of best practice configuration rules that are checked for each array - this highlights other areas of concern such as insufficient quantity of hot spares.

Information presented here provides additional confidence that we understand the customers environment - customers often nod in agreement when these issues are discussed, they are equally likely to be surprised by some of the issues raised here.

Operational Issues
- **HP EVA 8100 in operational attention state**
- Largest workload array with least capability
- Very low storage utilization
- Performance in VRAID configuration
- High environmental impact for usable storage volume
- 8100-ITEX in attention operational state
- Exchange 2007 IO latency in disk controller
- IO response time and disk latency above expected figures
- Large physical footprint due to overheads
- A 450GB disk in FAILED condition (EVA 8100, \Disk Groups\Ungrouped Disks)

Performance Issues
- **Storage Processor** constraints
- No linear performance scalability
- Back end disk **scalability limitations**
- FATA back end disk IO throughput
- 2Gb loops within array contention
- Limited cache for read ahead and reuse
**Capacity Metrics**

These three graphs depict the client environment after the vendor solution has been landed.

There’s an objective view of the same top 20 hosts, but this time using thinly provisioned storage to improve utilisation efficiency to 67%.

Our software is aware of the thin provisioning friendly and non-friendly applications and filesystems - this is all taken into consideration during the calculations.

The changes in disk tiering and storage distribution are shown by the last two pie charts.
Butterfly Differential Business Case
The effects of 36 months growth of the current hardware infrastructure are compared against a more efficient target design.

The calculated source and target costs include SAN port count, disk cost and power requirements. These two figures are compared to show a differential business case.

We use current street pricing for our modeling. This has proven to be very accurate. During a recent large vendor bid our pricing was within 1% of the vendors calculations.

This business case is focused on the tangible hardware and software investment that a customer will make.
Butterfly Migration Planning
A high level project plan is created for use with Butterfly Migration Engine - using tools that will automate and manage the phasing of the deployment to the target solution.

Summary
The data collection process is simple and non-disruptive.

The Butterfly Analysis Engine quickly processes the collected data to present an easy to understand representation of the customers existing environment, a suitable target environment and a differential business case that is supported by empirical data.

Any questions?