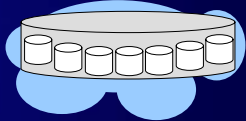


**COMPAQ**  
Inspiration Technology

## StorageWorks™ Enterprise Virtual Array

Virtualization Techniques



Siegfried Edelmann  
Windows NT & StorageWorks Support Engineer  
Compaq Customer Support & Solution Center München

**COMPAQ**

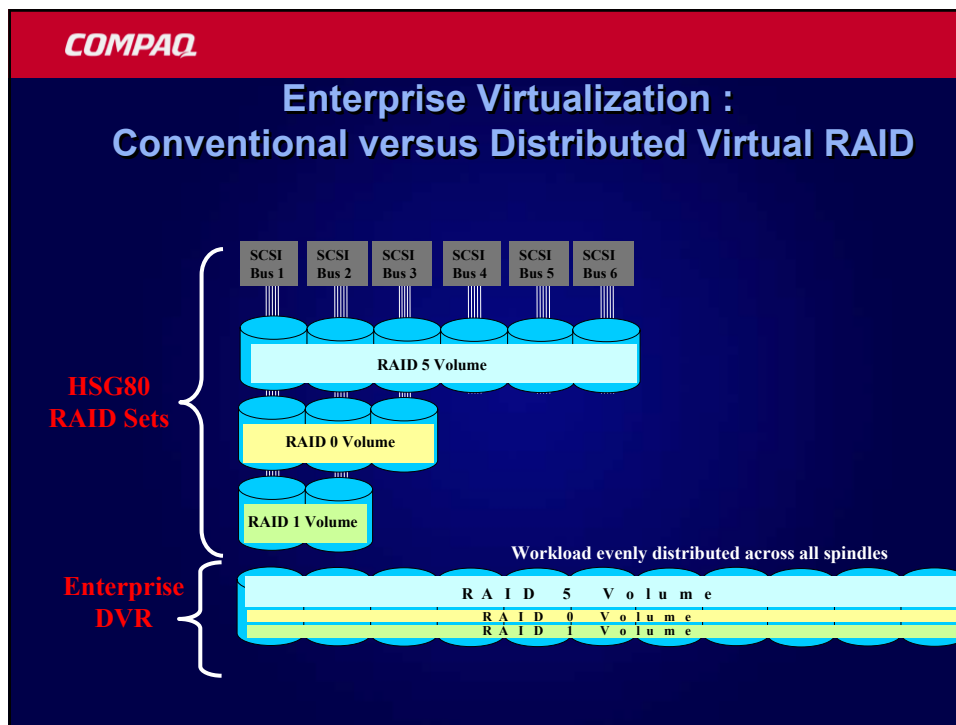
### Enterprise Virtualization : Subjects covered

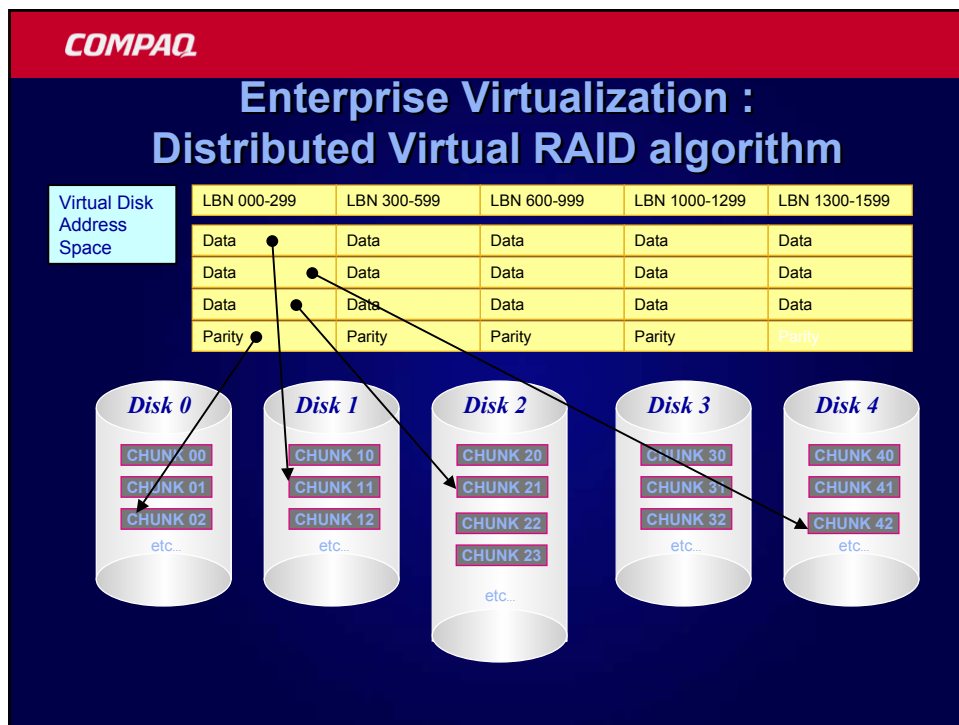
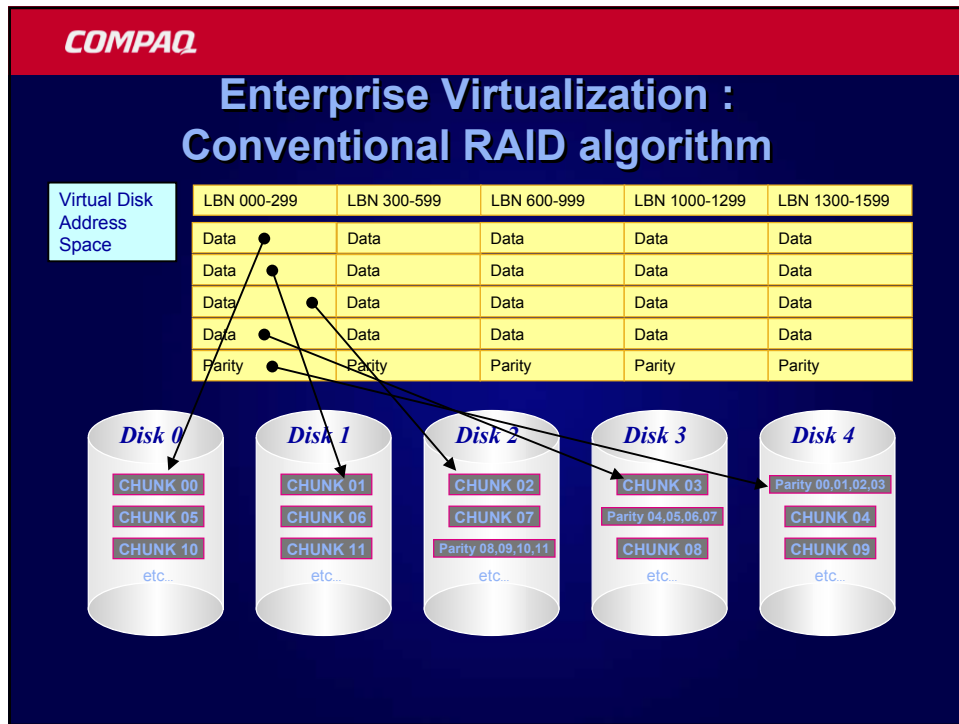
- Distributed Virtual RAID versus Conventional RAID.
- Disk Group Characteristics.
- Virtual Disk Ground Rules.
- Virtual Disk Leveling.
- Distributed Sparing.
- Redundant Storage Sets.
- Snapshot and SnapClone implementation.
- Configuration Remarks.

**COMPAQ**

### Enterprise Virtualization : Distributed versus Conventional RAID

Conventional RAID	Distributed Virtual RAID
<ul style="list-style-type: none"> <li>▪ Performance limited by # of disk drives in StorageSet</li> <li>▪ Possible to find customer data if one knows the LBN and chunksize.</li> <li>▪ Load Balancing required of application and Databases over available backend (SCSI) busses</li> <li>▪ I/Os balanced across StorageSet</li> </ul>	<ul style="list-style-type: none"> <li>▪ Performance limited by # of disk drives in Disk Group</li> <li>▪ Not possible to predict on which physical disk customer data resides</li> <li>▪ Eliminate load balancing procedures for applications and Databases.</li> <li>▪ I/Os balanced across disk group</li> </ul>





**COMPAQ**

## Enterprise Virtualization : Disk Group Characteristics

- Disk Failure protection level :
  - 0, 1, or 2 disk failures
- Minimum : 4 physical disk drives
- Maximum : Number of physical disk drives present
- Maximum number of Disk Groups : 16
- Net capacity :
  - $0.98 * (a - b) * c$  with
    - a = Number of disks in Disk Group.
    - b = Disk failure protection level.
    - c = Disk Drive Capacity.
- Chunk size : 2 MB (fixed).

**COMPAQ**

## Enterprise Virtualization: Virtual Disk Ground Rules

- Virtual Disk redundancy:
  - **None**: Data is striped across all physical disks in the Disk Group.
  - **Moderate** : Data is striped with parity across all physical disks in the Disk Group. Always 5 (4+1) physical disks per stripe are used.
  - **High** : Data is striped mirrored across all physical disks in the Disk Group. 2 physical disks per stripe as used.

**COMPAQ**

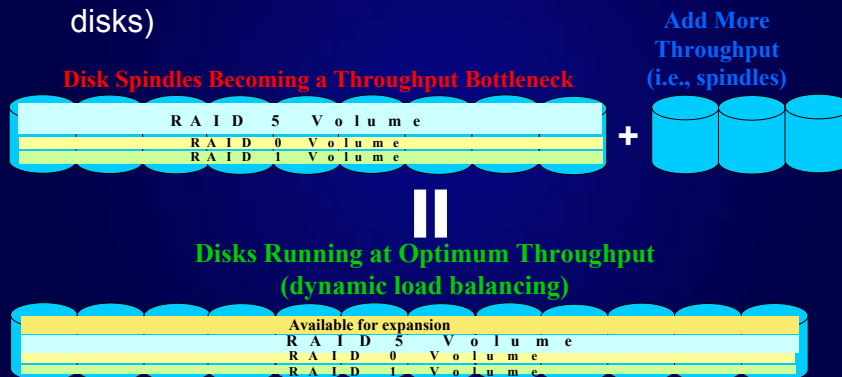
## Enterprise Virtualization Virtual Disk Leveling

- Goal is to provide capacity leveling across all disk drives within the Disk Group.
  - Example 1: Disk Group = 100 drives, all 18G
    - All disks will contain 1% of the Virtual Disk.
  - Example 2: Disk Group = 100 drives , 50\*18G, 50\*36G
    - Each 36G disk will contain > 1% of the V.D.
    - Each 18G disk will contain < 1% of the V.D.
- Load Balancing is achieved through capacity leveling

**COMPAQ**

## Enterprise Virtualization : Virtual Disk Leveling

- Dynamic pool capacity changes
- Pool capacity can be added in small increments (3 disks)



**COMPAQ**

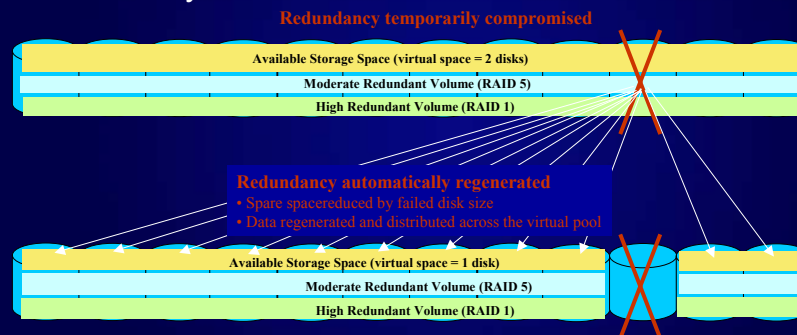
## Enterprise Virtualization : Distributed Sparing

- Chunks allocated, but not dedicated as spares, on all disk drives to survive 1 or 2 disk drive failures.
- Example :
  - Disk Group : 8 \* 36G, Protection level : 1 disk.
    - Total Disk Group size : 8 \* 36 = 288 GB.
    - Maximum size for Virtual Disks (8-1) \* 36 = 252 GB.
  - Disk Group : 120 \* 36G, Protection level : 2 disks
    - Total Disk Group size : 120 \* 36 = 4.32 TB
    - Maximum size for Virtual Disks (120-2) \* 36 = 4.25 TB

**COMPAQ**

## Enterprise Virtualization : Distributed Sparing

- Virtual Disk blocks automatically regenerated to restore redundancy.



**COMPAQ**

### Enterprise Virtualization : Redundant Storage Sets

- Reduces chance of Data Loss in Large (> 12 physical disks) Disk Groups.
- Not visible on User Interface.
  - Complete managed by the HSV controllers.
- Typical size : 6 - 12 physical disks.
  - As soon as Redundant Storage Sets exceeds 12 physical disks, it splits into 2 RSSs of 6 disks each.
- Failed disk drive recovery restricted to affected Redundancy Storage Set only.

**COMPAQ**

### Enterprise Virtualization : Snapshot and SnapClone implementation

- **Snapshot** : Data is copied from Virtual Disk to Snapshot on demand.
  - “Preferred” by the same controller, as the original Virtual Disk is “preferred” to.
  - Space efficient :
    - chunks will be allocated in the Disk Group on demand.
    - Snapshot removed if Disk Group becomes full.
  - Space inefficient :
    - chunks are allocated in the Disk Group at moment of Snapshot creation.
    - Snapshot remains available if Disk Group becomes full.

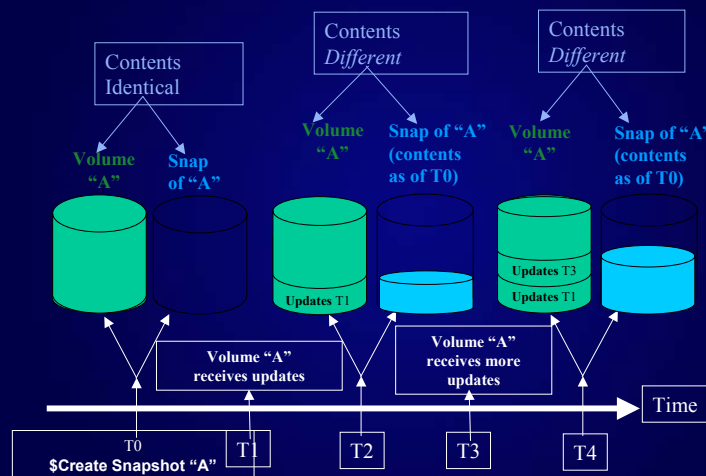
**COMPAQ**

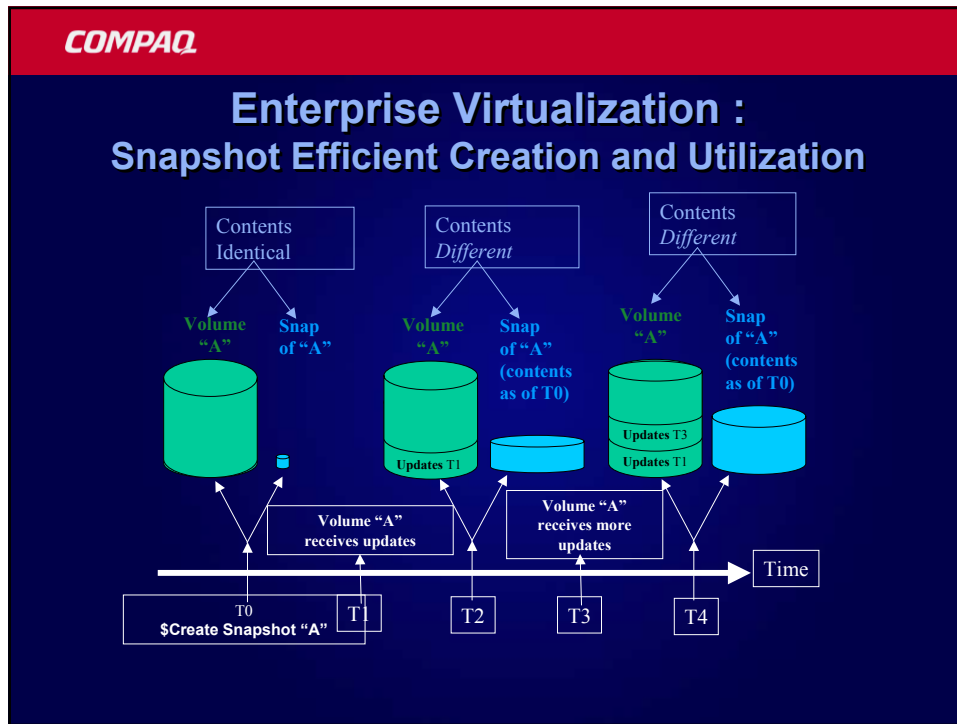
## Enterprise Virtualization : Snapshot and SnapClone implementation

- **SnapClone** : Data is copied from Virtual Disk to SnapClone in background
  - chunks are allocated in the Disk Group at moment of SnapClone creation.
  - Clone may be set to a different "Preferred" controller as the original Virtual Disk
    - Only after background copy process is finished.
- No relation between redundancy level of Snapshot or SnapClone and original Virtual Disk
  - All combination of snapshot/clone redundancy levels (None, Moderate, High) and Virtual Disk redundancy levels (None, Moderate, High) are possible and allowed.

**COMPAQ**

## Enterprise Virtualization : Snapshot Inefficient Creation and Utilization





**COMPAQ**

### Enterprise Virtualization: Snapshot versus Snapclone

	Description	Pro's	Con's
Snapshot Space efficient	Pointer based Copy before Write Allocate space on demand	Space efficient (allocated on demand)	Overcommit problem
Snapshot Space inefficient	Pointer based Copy before Write PreAllocate space on creation	No Overcommit problem	Space inefficient (allocated right away)
Snapclone	Same as Snapshot space inefficient, but now with background process to separate VD.	No Overcommit problem Repeatable, separate VD's	Space inefficient Consumes some background process time

**COMPAQ**

## Enterprise Virtualization : Configuration Remarks

- Use even number of drives in Disk Group
  - if using High redundancy on VD level.
  - if using Moderate redundancy on VD level, optimum number is multiples of 10 disk drives per Disk Group
- Use Disk Groups to separate Virtual Disks that need to be separated.

**COMPAQ**

## Enterprise Virtualization : Configuration Remarks

- Disk Groups can be created in 2 ways:
  - **Vertically**, so one disk drive per shelf
    - *Pro* :
      - 100% guarantee that it will survive a shelf meltdown failure
    - *Con*:
      - Limited in Number of disk drives per Disk Group:
        - HSV110 16-18 disks
      - Performance (less spindles)
      - Cannot use full capacity due to loss in space reservation for (space efficient) Snapshots

**COMPAQ**

## Enterprise Virtualization : Configuration Remarks

- Disk Groups can be created in 2 ways (continued):
  - **Multiple disk drives per shelf**
    - *Pro* :
      - Unlimited Performance due to large number of disk drives in Disk Group:
        - HSV110 max 240 disks
      - No loss in space reservation for Snapshots and / or Snapclones
    - *Con*:
      - Isn't guaranteed to survive a shelf meltdown failure.

**COMPAQ**

## Enterprise Virtualization: Food for Thought

- A virtual disk is a disk drive that is imaginary, but operational. It isn't really there but it works just fine.
- And then there is a snapshot which is a non-existent, but operational copy of an imaginary (but operational) disk drive.
- And then there is the console LUN, which is a pretend virtual disk -- an imaginary disk drive that isn't operational but acts as if it might be.

**COMPAQ**

## Background Slides

**COMPAQ**

### Enterprise Virtualization : Disk Group and Virtual Disks map

- Map contains information which Virtual Disks use which chunks.
- Chunks within Disk Group are allocated on Virtual Disk creation time.
  - Allocated size = size of Virtual Disk.
- Allocated chunks are not contiguous.
  - They are within stripes (Moderate, High redundancy levels), but may not be between stripes
- Map is kept on 6 disk drives:
  - Quorum disks (slow blinking spindle LED).

**COMPAQ**

### Enterprise Virtualization : Disk Group and Virtual Disks map

- Map is read into HSV controllers during boot.
- Map is written on state change of Disk Group
  - Disk drive failure,
  - Disk drives added or removed.
  - Space inefficient Snapshot is created.
  - SnapClone fully copied,
  - Virtual Disk added or deleted.
  - Virtual Disk is enlarged.
  - On each write I/O to a Virtual Disk, which has a space efficient Snapshot.
    - If results in a new chunk
  - Others

**COMPAQ**

### Enterprise Virtualization : Disk Group and Virtual Disks map

Disk 1	Disk 2	Disk 3	Disk 4	Disk 5	
					Virtual Disk size:
Stripe1	Stripe2	Stripe3	Stripe4	Stripe5	VD_3 = 20MB
					VD_2 = 20MB
					VD_1 = 40MB
Mirror3	Mirror4	Mirror4	Mirror5	Mirror3	Virtual Disk redundancy level:
Mirror1	Mirror1	Mirror2	Mirror2	Mirror3	Stripe1 None
					Mirror1 High
					Data1 Moderate
Data2	Data2	Data2	Parity2	Data2	
Data1	Data1	Data1	Data1	Parity1	



