

**Company85**

# **Customer L Data Classification Report**

Mark Worts



# Engagement Objective

- Objective

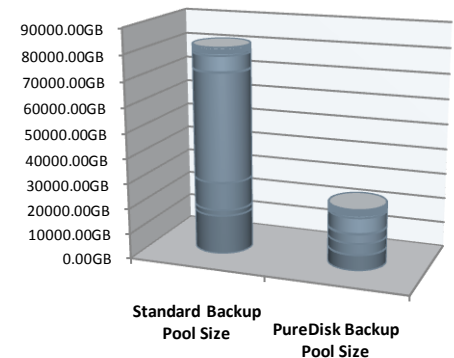
- To identify how much storage space is required for the next 3 (2011) years for:
  - Online storage
  - Data Protection
    - Backup storage (Tape and/or Disk)
    - Snapshots
    - Disaster Recovery
  - Archive
- to identify the required level of service from storage
- to ensure alignment with Business Requirements

# Management Summary

- Mail and Document Management System are the critical business systems and the big storage consumers
- Online data requirement is expected to grow from 4.3TB to 15TB by 2011
- Backup policies need to change to meet desired RPOs
- Backup Storage Pool is significantly reduced by PureDisk
- Archiving 7 year old data will have limited impact at the current time

Enterprise Vault – 0.8TB  
Worksite – 1TB

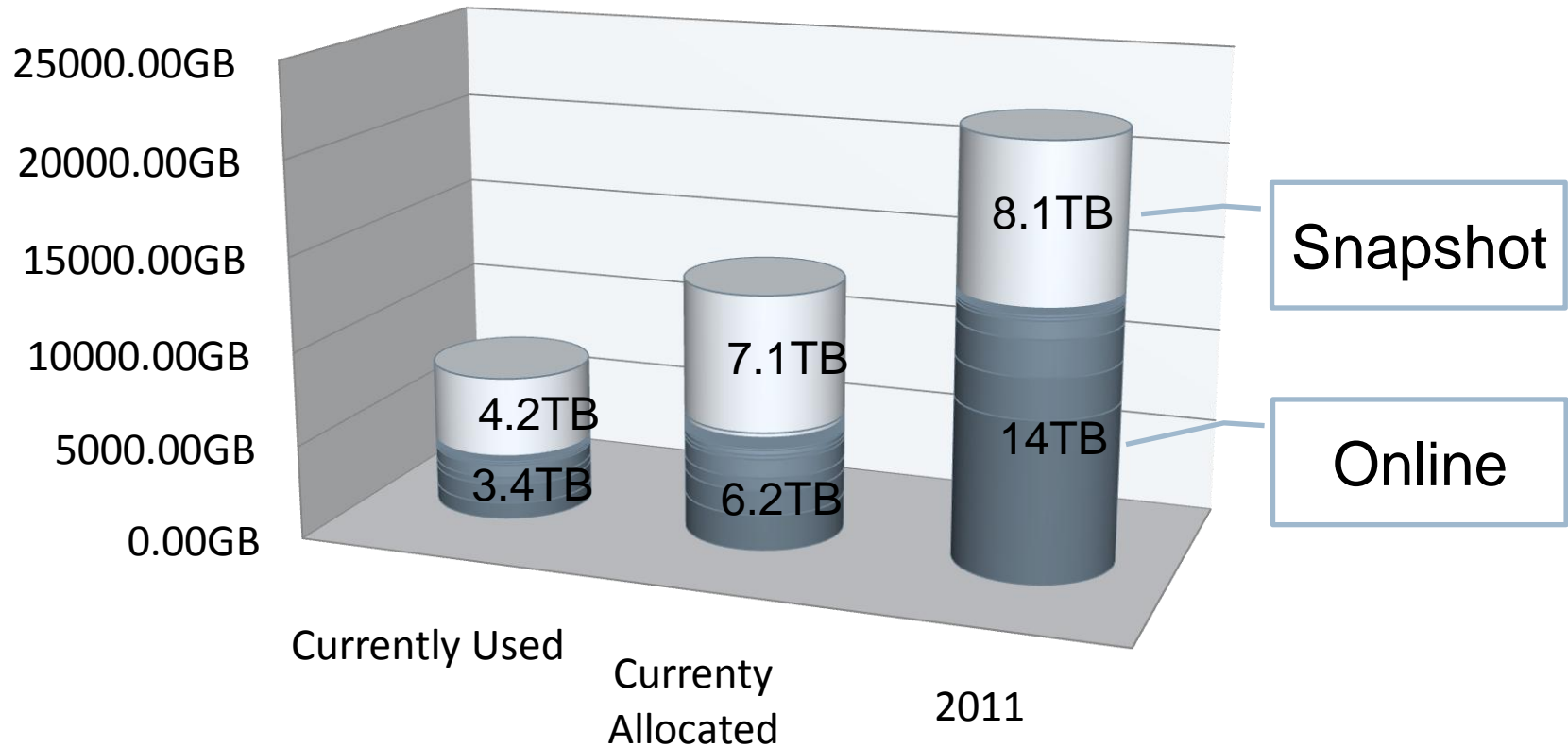
Enterprise Vault – 8.5TB  
Worksite – 2TB



Firmware Archive - 0.2TB

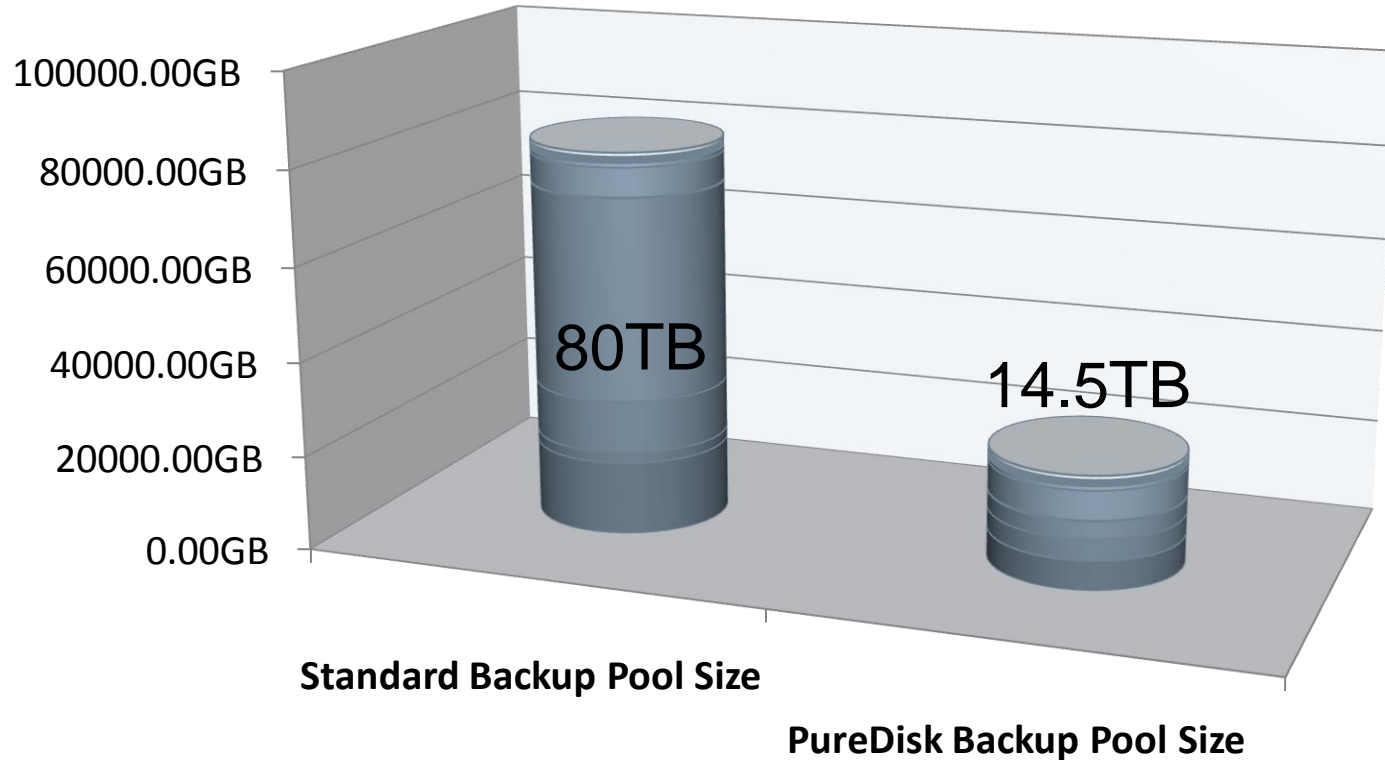
# Production Online Storage Summary

## Online Disk & Snapshot Capacity



# Backup Storage Pool Size Summary

## Backup Storage Pool Size

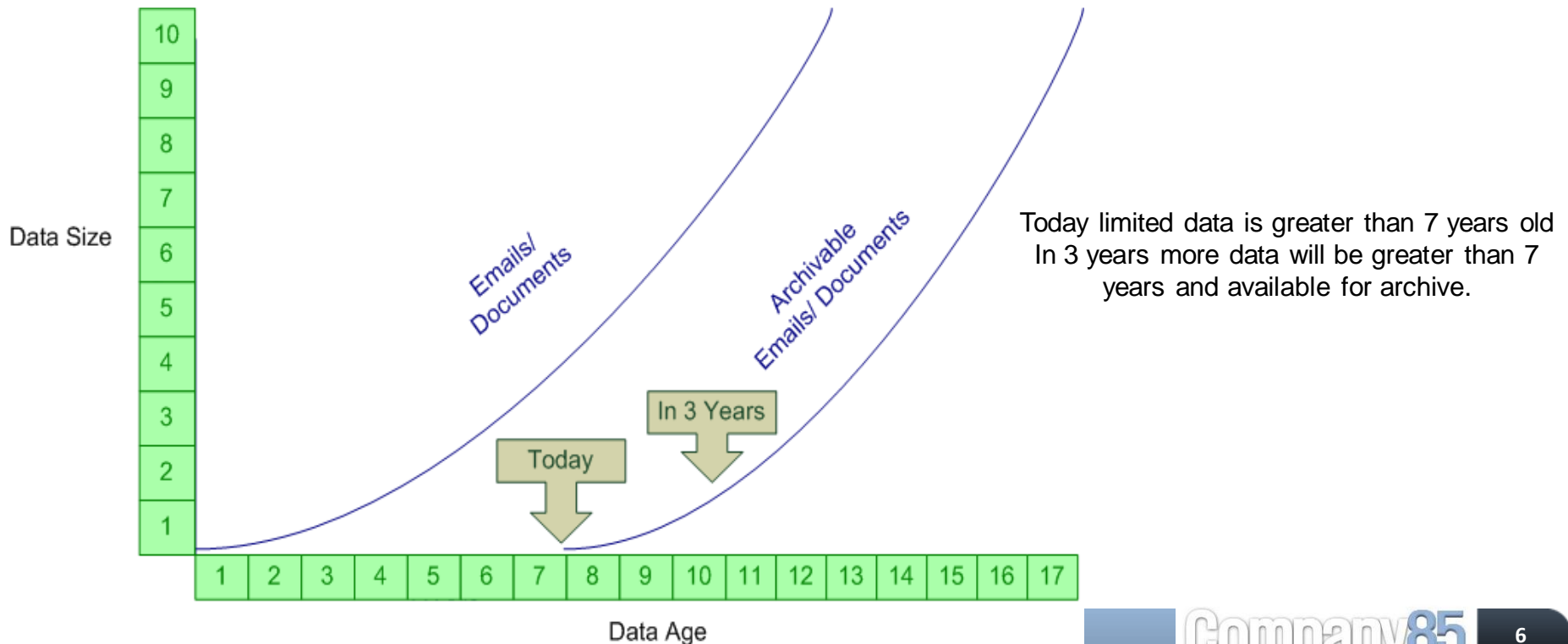


Predicted 2011 figures.

The graph shows how much backup storage space is required to maintain the retention policies defined with and without implementing PureDisk.

# Archive Headlines

- Email Archive is currently providing the appropriate level of service
- Users are happy with archiving data over 7 years old to slow recall storage
  - However this will currently have no impact on today's storage





# Data Classification (Anonymised)

Online Storage Services	Availability	Performance	Provisioning	Extended Functionality
Class 1H+	High	High + isolated	Weeks	Copy Services Snapshots Replication
Class 1H		High	< 1TB Days	
Class 1S		Standard	>1TB Weeks	
Class 2H+	Standard	High + isolated	Weeks	NA
Class 2H		High	< 1TB Days	
Class 2S		Standard	>1TB Weeks	

Online Storage Services Catalogue

Data Protection Service	Data Recoverability		Disaster Recoverability	
	RTO	RPO	RTO	RPO
Class 1	4 hours	4 hours	4 hours	24 hours
Class 2	16 hours	24 hours	48 hours	24 hours
Class 3	48 hours	24 hours	5 days	24 hours
Class 4	72 hours	24 hours	4 weeks	24 hours

Data Protection Services Catalogue

Service	Class 1 – SAN			Class 2 – SAN		
	H-P + Isolated	High Performance	Standard Performance	H-P + Isolated	High Performance	Standard Performance
Availability	High Availability 99.99%			Standard Availability 99.95%		
Response Time	Fast Large Cache High Bandwidth Dedicated Processors			Standard Reduced Cache Generic Processor		
Extended Services	Business Copies, Snapshots, Replication			None		
Through-put	RAID 1+0 Small / Fast Disks		RAID 5 Larger/ Slower Disks	RAID 1+0 Small / Fast Disks		RAID 5 Larger/ Slower Disks
Segregation	No shared resources	Shared resources		No shared resources	Shared resources	
Array Fan In/Out	Dedicated Ports	Low Port Density	High Port Density	Dedicated Ports	Low Port Density	High Port Density
Provisioning Times	Weeks	Days		Weeks	Days	
SAN Fan In/Out	Core Switch	Edge Switch		Core Switch	Edge Switch	
SAN Availability	Dual Fabric					

# Headline Assumptions and Observations

- Assumptions

- Business Growth of 10% year on year
- New Backup Policy as defined is implemented
- Projections are for Data sizes not raw space
- HP EVA snapshot technology is used as defined
- PureDisk savings are conservative estimates

- Observations

- Users are happy with current level of service
- Project Management and tidying up are not evident
- There is a good IT strategy
- Technical Management Information is lacking
- Applications do not have high performance demands and increasing the number of slower disks should be considered

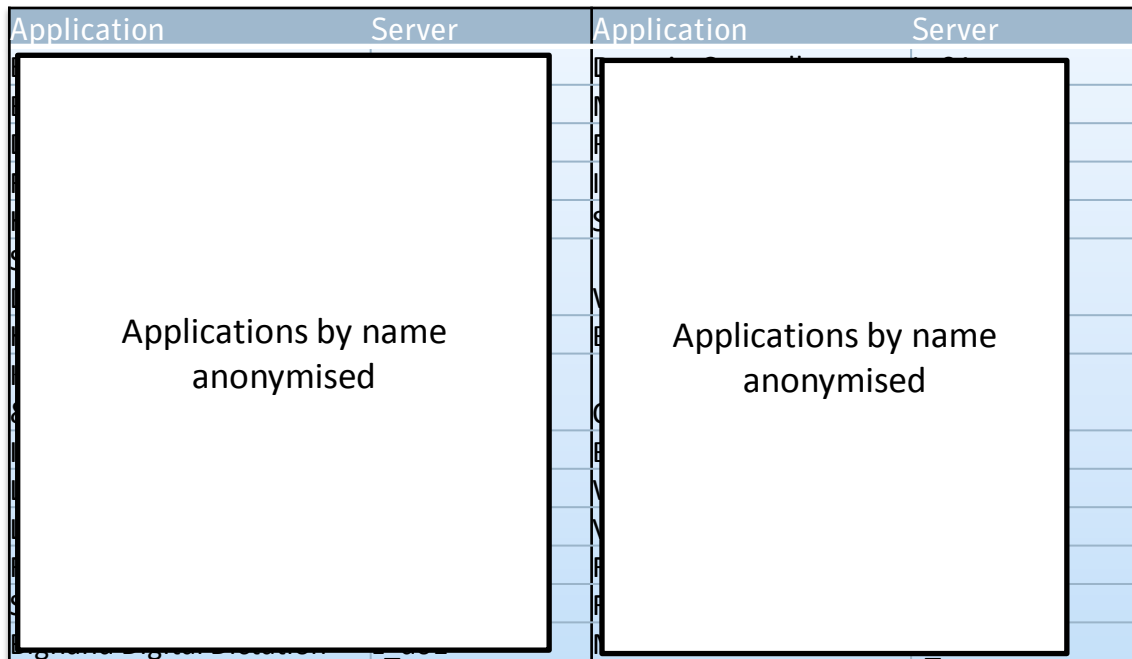
# Details

# Contents

- Introduction
- Business Requirements
- Online Storage Requirements
- Data Protection Requirements
- PureDisk Savings
- Snapshot Requirements
- Archive Opportunities
- Mail & File System Usage
- Observations

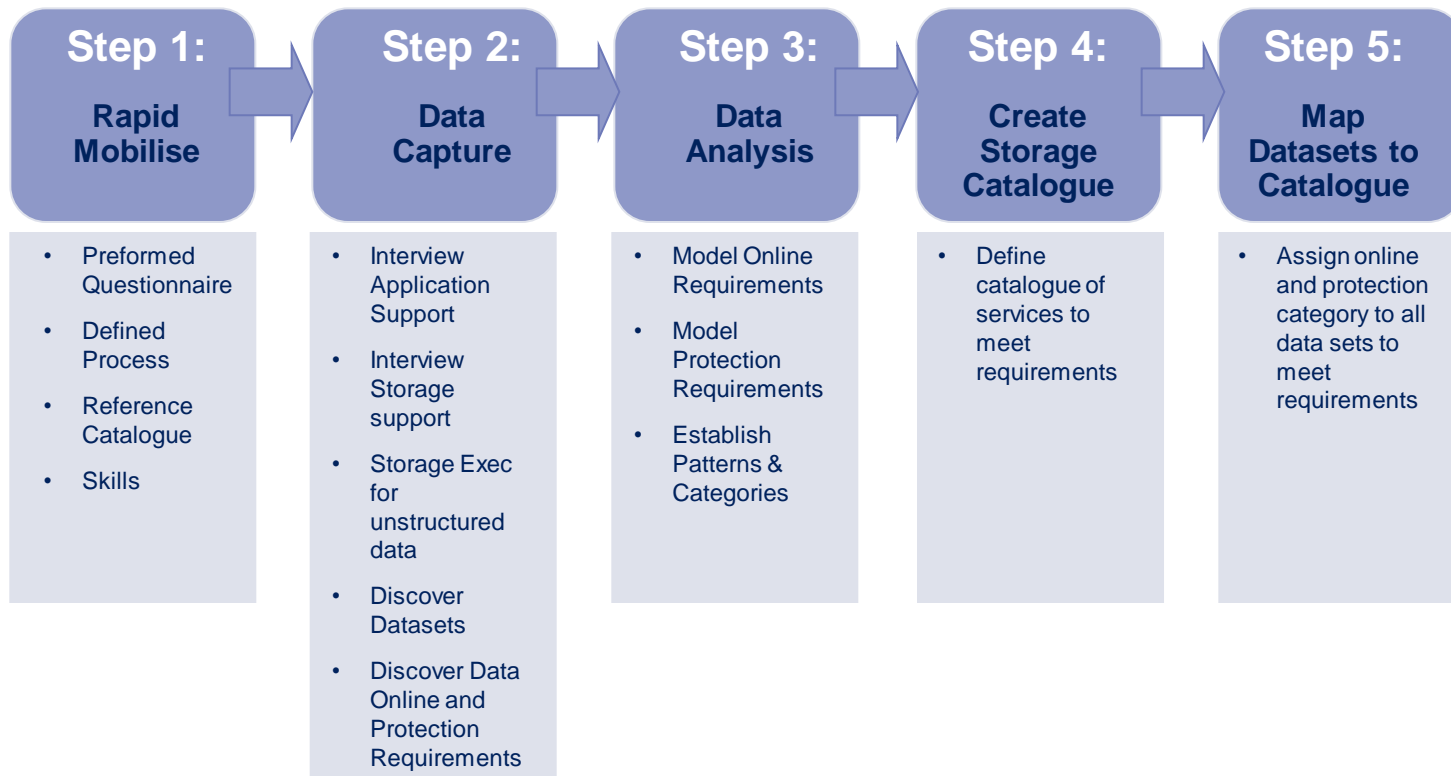
# Scope

- Applications in scope



- SAN and Local disk
- Other systems with storage were not asked to be in scope

# Engagement Process



# Technical Input

- Storage Usage Areas Covered
  - Application Description
  - Current application service level
  - Component Data Sets
  - Data Sizes
  - Application importance
  - Data retention
  - Prevention of data loss solutions
  - Data availability solution
  - Data growth
- Technical Interviews
  - Storage Manager
  - Windows Admin
  - IT Manager
  - Sharepoint Support
  - Messaging Manager
- Strategy Interviews
  - IT Director
  - CTO

# High Level Storage Strategy

- All online storage is presented over a SAN from HP EVAs as FCAL RAID2
- HP EVA Replication across SAN is used to provide storage DR
- Standalone servers use local boot disks
- VMware Hosts boots of local mirrored disks
- VMware Datastores are provisioned by the SAN
  - VMware Guests, OS and data, disks are provisioned by the SAN Datastores
- SAN replication supports VMware DR solution
  - Virtual hosts are brought online at DR site using replicated OS and data Datastores
  - DR for individual Guests is supported by separated Datastores
- NAS (File Shares) is not supported to servers or users.
  - Strategic to enforce use of DMS
- Desktop / Laptop data storage is not supported
  - All data to be stored in DMS which supports offline access
- All servers are backed up using Symantec Netbackup

# Business Requirements

# Business Input

- Business input regarding data storage requirements was sort on the following areas:
  - Applications Used
  - Current application service level
  - Application importance
  - Data retention requirements including archiving potential
  - Prevention of data loss requirements
  - Data availability requirements
  - Data growth
- Business interviews:
  - Head Of Business Units
  - User A
  - User B
  - User C
  - User D

# Business Availability Requirements

- System Criticality:

- High

- Mail (MS Exchange)
- Document Management System

- Medium

- Applications by name anonymised
- 
- 
- 
- 

- Low

- Applications by name anonymised
- 
- 
- 
- 

- Data Loss

- High

- Desired
  - Range - no loss to 2 hours
  - Average – No loss
- Acceptable
  - Range – no loss to 24 hours
  - Average – 2 Hours

- Medium & Low

- 24 hours and above

- Down Time

- High

- Desired
  - Range – immediate – 4 hours
  - Average 2 hours

- Medium & Low

- 24 Hours and above

- Disaster Recovery

- High

- Desired
  - Range – immediate – 24 hours
  - Average 2 hours

- Medium & Low

- 24 Hours and above

# Business Storage Requirements

- Data Retention

- Minimum – 6 Years
- Maximum – Permanent
- No identified requirement for monthly point in time copies

- Data Growth

- Business Growth
  - 10% organic growth is expected
  - No known changes to current predictions
- Email
  - Individual usage may now have reached peak
  - More and more business conducted by email
- Documents
  - Will grow with business size
  - All documents now produced online
  - Possible document scanning increase

- Archiving

- Acceptable

- Worksite
  - Flagging
  - Folder based rules
  - 24 Hour max recovery
- Mail
  - All mail accessible
  - Current recall time good
  - Longer ok for >7 yrs
- Home & Shared Folders
  - Not used for long term storage
- Firmware
  - Old month ends should be archived

- Not Applicable

- All other applications

# Online Storage Requirements

# Storage Water Fall

- Disk Space

- Physical

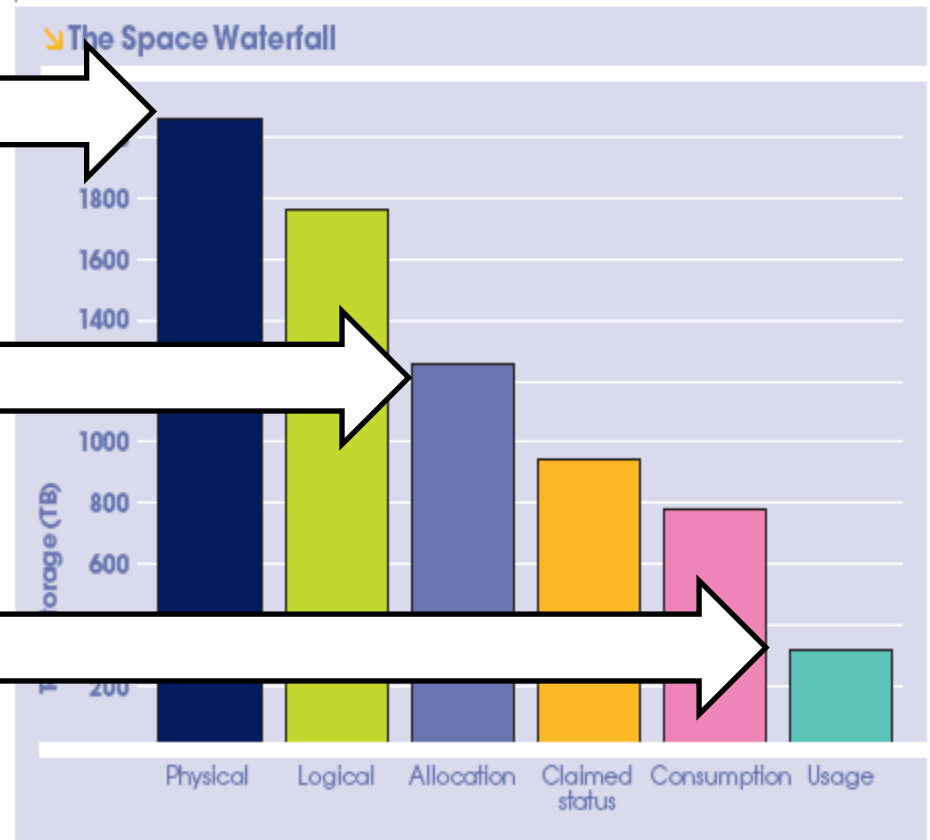
- SAN – 17.5TB
- Local – 4.8TB

- Scoped Allocation

- SAN – 3.7TB
- Local - 4.8TB

- Dataset Size

- SAN – 2.3TB
- Local – 2TB



# Growth Calculation

- Observed
  - Enterprise Vault – 6.3GB per day then 10% year on year
  - Worksite (WIP, Archive) – 1GB per business day then 10% year on year
  - Firmware (Live & Reporting) – 3GB per annum then 10% year on year
- Unmeasured
  - 10% growth year on year assumed

# Online Storage Growth Current vs 2011

Application	Current Used Size	Estimated Used Size in 2011	Current Allocations	3 Year Utilisation against allocated
Applications by name anonymised	1028.05GB	1964.73GB	1308.00GB	150%
	901.00GB	1199.23GB	1390.00GB	86%
	856.92GB	8502.69GB	1906.50GB	446%
	697.40GB	928.24GB	987.70GB	94%
	390.00GB	1905.35GB	684.00GB	279%
	134.00GB	178.42GB	230.00GB	78%
	127.00GB	148.18GB	268.00GB	55%
	67.51GB	89.86GB	125.90GB	71%
	34.40GB	45.79GB	40.00GB	114%
	15.20GB	16.92GB	27.00GB	63%
	15.00GB	19.97GB	86.00GB	23%
	15.00GB	19.97GB	52.40GB	38%
	13.62GB	18.13GB	17.00GB	107%
	13.00GB	13.03GB	20.00GB	65%
	10.00GB	13.31GB	15.00GB	89%
	9.10GB	12.10GB	33.50GB	36%
	9.00GB	11.98GB	18.00GB	67%
	8.00GB	10.65GB	68.00GB	16%
	8.00GB	8.00GB	34.00GB	24%
	6.90GB	6.90GB	67.70GB	10%
6.59GB	8.77GB	16.90GB	52%	
6.20GB	6.27GB	20.00GB	31%	
4.80GB	4.97GB	20.00GB	25%	
3.09GB	3.16GB	8.94GB	35%	
0.19GB	0.26GB	243.00GB	0%	
0.00GB	0.00GB	11.69GB	0%	
0.00GB	0.00GB	986.00GB	0%	
<b>Grand Total</b>	<b>4379.97GB</b>	<b>15136.87GB</b>	<b>8685.23GB</b>	

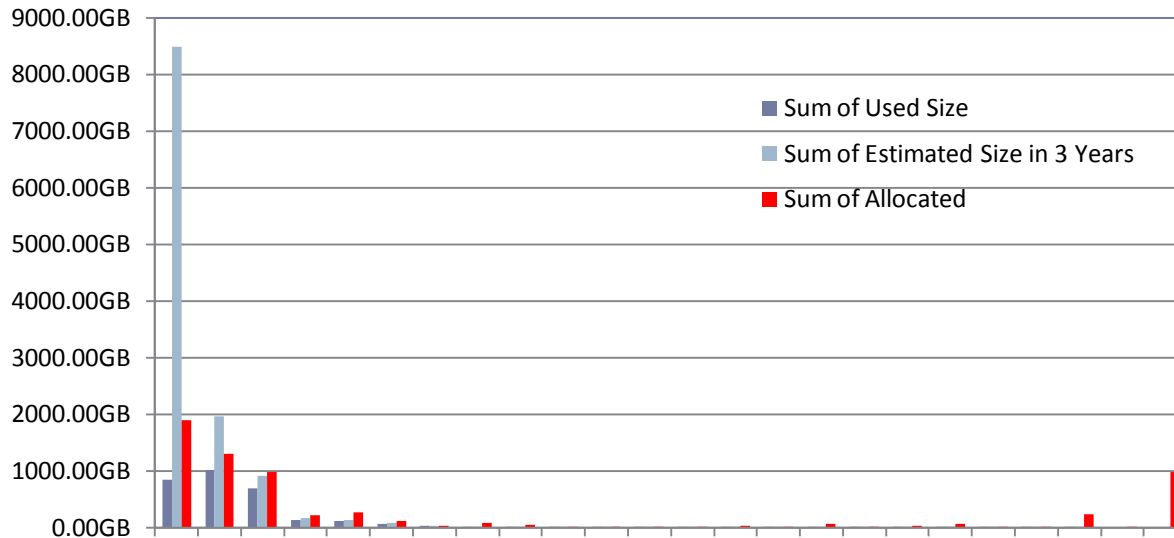
\* Due for decommission

Measured Growth

Interview Estimate

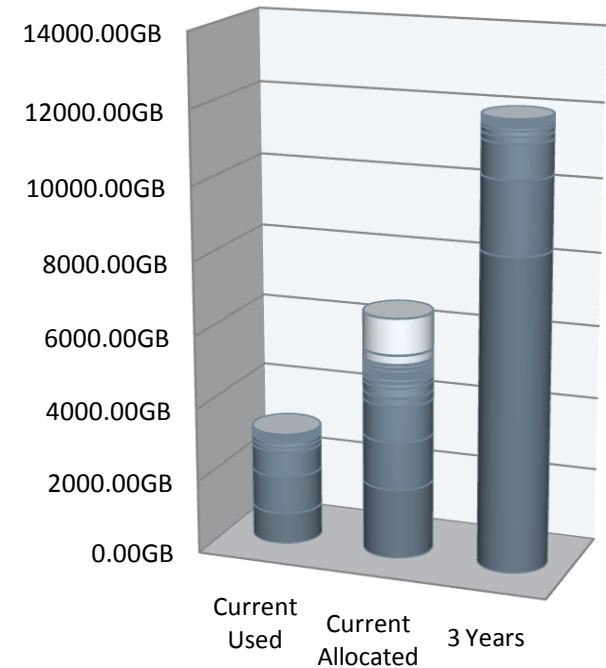
# Storage Usage

## Current and 3 year projection



Applications by name anonymised

### Totals



# Online Storage Requirements

- Applications identified as sensitive to storage performance
  - With High Performance Requirement
    - MS Exchange
  - With Medium Performance Requirement
    - Applications by name anonymised
  - With Low Performance Requirement
    - Applications by name anonymised
- The following datasets are opportunities to move to low performance disks:
  - Candidates
    - Applications by name anonymised
    - SMS
    - MS Exchange Backup Dump
    - Enterprise Vault
    - McAfee AV Server

# Additional Online Storage Requirements

- Applications to be migrated to the SAN

—		6.90GB
—		8.00GB
—	Applications by name anonymised	13.62GB
—		8.00GB
—		9.00GB
—	SMS	127.00GB
—	Enterprise Vault	856.91GB
—	Applications by name anonymised	6.59GB
—	Total	2026GB

# Data Protection Requirements

# Data Protection Requirements

- Currently production data is backed up to a 24 hour recovery point plan on an opt in basis
- There is a desire to meet a 2 hour recovery point objective (RPO) on critical systems
- No Recovery Time Objectives (RTO) are agreed
- Recovery to a DR site is required for high priority systems
- Zero Data loss Replication to DR is required



# Data Protection Observations

- The following application currently have a lower level of backup data protection than desired
  - Recovery Point Objective (RPO) of 2 Hours Can't be met:
    - Applications by name anonymised
    - Mail (Exchange & Blackberry)
  - Recovery Time Objective (RTO) of High Can't be met:
    - Applications by name anonymised
    - Mail (Exchange & Blackberry)
- Customer L propose snapshot technology to address the shortfall.



# New Backup Policies

## Critical Services Backup Policy

- 2 Hourly Snapshots
  - Retention 24 Hours
  - Copies 24
- Daily Full Backup
  - Retention 7 days
  - 6 Copies
- Weekly Full Backup
  - Retention 5 weeks
  - 6 copies
- Monthly Full Backup
  - Retention 12 Months
  - 13 copies
- 13<sup>th</sup> Month de-staged to long term storage

## Standard Services Backup Policy

- Week Day Incremental
  - Retention 5 copies
- Weekly Full Backup
  - Retention 5 copies

## Archive Services Backup Policy

- Week Day Full
  - Retention 2 copies

# Backup Policy and Storage Pool Size

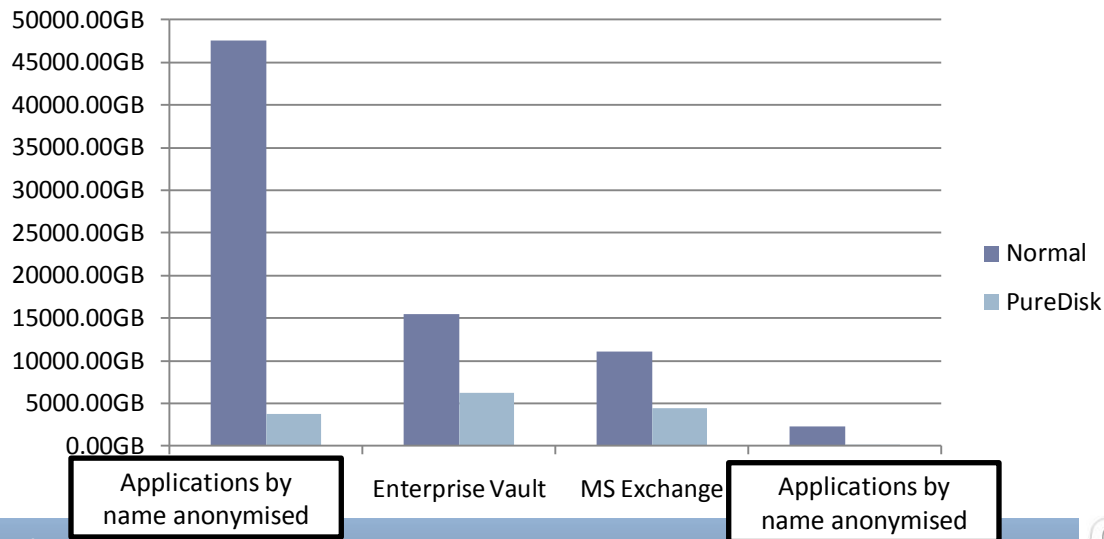
Applications	Backup Policy	Full Backup data in 3 years	Incremental backup date in 3 years	PureDisk backup data in 3 years	PureDisk Incremental backup data in 3 years
<b>PureDisk Clients</b>					
Applications by name anonymised	Archive	15518.20GB	0.00GB	6233.59GB	0.00GB
	Critical	2075.00GB	194.97GB	54.40GB	0.00GB
	Critical	11084.75GB	0.00GB	4442.00GB	0.00GB
	Critical	43217.91GB	18.69GB	3719.78GB	0.00GB
<b>Netbackup Clients</b>					
Domain Controller	Standard	100.00GB	100.00GB	100.00GB	100.00GB
File server	Critical	4600.00GB	1574.73GB	4600.00GB	1574.73GB
Applications by name anonymised	Critical	31.00GB	10.00GB	15.50GB	10.00GB
	Standard	891.77GB	30.06GB	891.77GB	30.06GB
	Standard	15.78GB	45.00GB	15.78GB	45.00GB
	Standard	66.55GB	1.54GB	66.55GB	1.54GB
	Standard	150.00GB	51.50GB	150.00GB	51.50GB
	Standard			0.00GB	0.00GB
	Standard	43.86GB	4.39GB	43.86GB	4.39GB
	Standard	0.83GB	0.85GB	0.83GB	0.85GB
	Standard	85.00GB	87.08GB	85.00GB	87.08GB
	Standard	100.00GB	1.00GB	100.00GB	1.00GB
	Standard	60.23GB	43.20GB	60.23GB	43.20GB
	Standard	34.50GB	0.63GB	34.50GB	0.63GB
	Standard	55.00GB	1.32GB	55.00GB	1.32GB
	Standard	28.33GB	1.83GB	28.33GB	1.83GB
	Standard	0.00GB	0.00GB	0.00GB	0.00GB
	Standard	100.17GB	216.60GB	100.17GB	216.60GB
	Standard	40.00GB	3.81GB	40.00GB	3.81GB
	Standard	60.00GB	13.07GB	60.00GB	13.07GB
	Standard	3.99GB	8.28GB	3.99GB	8.28GB
	Standard	755.00GB	11.09GB	755.00GB	11.09GB
Standard	95.00GB	1.66GB	95.00GB	1.66GB	
<b>Grand Total</b>		<b>79212.85GB</b>	<b>2421.28GB</b>	<b>21751.28GB</b>	<b>2207.64GB</b>

62TB Saving shown on next slide



# PureDisk Backup Pool Size Saving

Application	Current Full Backup Size	Full Backup Size in 3 years	Total storage for retention policy in 3 years	Total storage for retention policy in 3 years (PureDisk)	PureDisk Saving
Applications by name anonymised	927.21GB	1827.73GB	47523.05GB	3719.78GB	43803.28GB
Enterprise Vault	849.43GB	7759.10GB	15518.20GB	6233.59GB	9284.61GB
MS Exchange	350.80GB	460.03GB	11084.75GB	4442.00GB	6642.75GB
Applications by name anonymised	80.00GB	83.00GB	2269.97GB	54.40GB	2215.57GB
<b>Total</b>	<b>2207.44GB</b>	<b>10129.86GB</b>	<b>76395.97GB</b>	<b>14449.77GB</b>	<b>61946.20GB</b>



# Snapshot Storage Requirement

# Snapshot Requirements

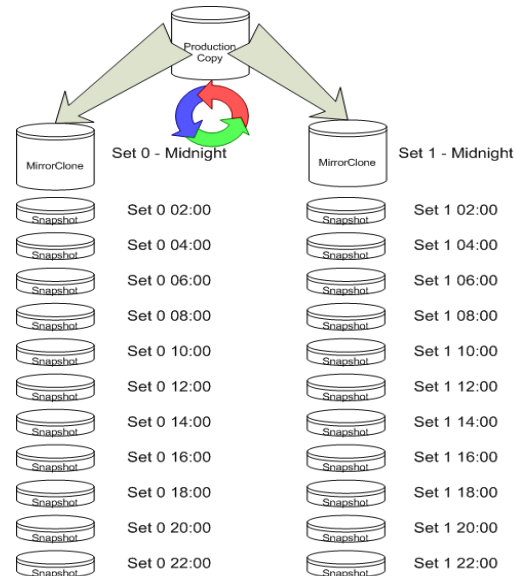
- There is a requirement to reduce data loss a maximum of 2 hours for the following applications:
  - Firmware
  - MS Exchange
  - Worksite
- i.e. an Recovery Point Objective of 2 Hours.
- To meet these requirements Snapshots are required

# Snapshot - Disk Requirements

## Snapshot Sizing

Row Labels	Currently Used Size	Estimated data Size in 2011	Current Allocated Size	Snapshot space for current used size	Snapshot Space for current allocation	Snapshot space for used size in 2011
Applications by name anonymised	6.20GB	6.27GB	20.00GB	19.84GB	64.00GB	20.05GB
	197.00GB	38.99GB	684.00GB	121.60GB	2188.80GB	124.78GB
MS Exchange	405.00GB	539.06GB	560.00GB	1296.00GB	1792.00GB	1724.98GB
Applications by name anonymised	719.10GB	1636.27GB	978.00GB	2301.12GB	3129.60GB	5236.08GB
<b>Grand Total</b>	<b>1327.30GB</b>	<b>2220.59GB</b>	<b>2242.00GB</b>	<b>3738.56GB</b>	<b>7174.40GB</b>	<b>7105.88GB</b>

## Snapshot Solution



To Support Recovery to any point in last 24 hours using 2 rolling full copies

MirrorClone 100% Production size

Snapshot average 10% of production Size

Disk Required = 100% + 200% + (22 \* 10%)  
= 420%

Production = 100%  
Data Protection = 320%

Note:  
Databases would have to be quiesced to provide recoverability

# Archive Requirements

# Archive Requirements

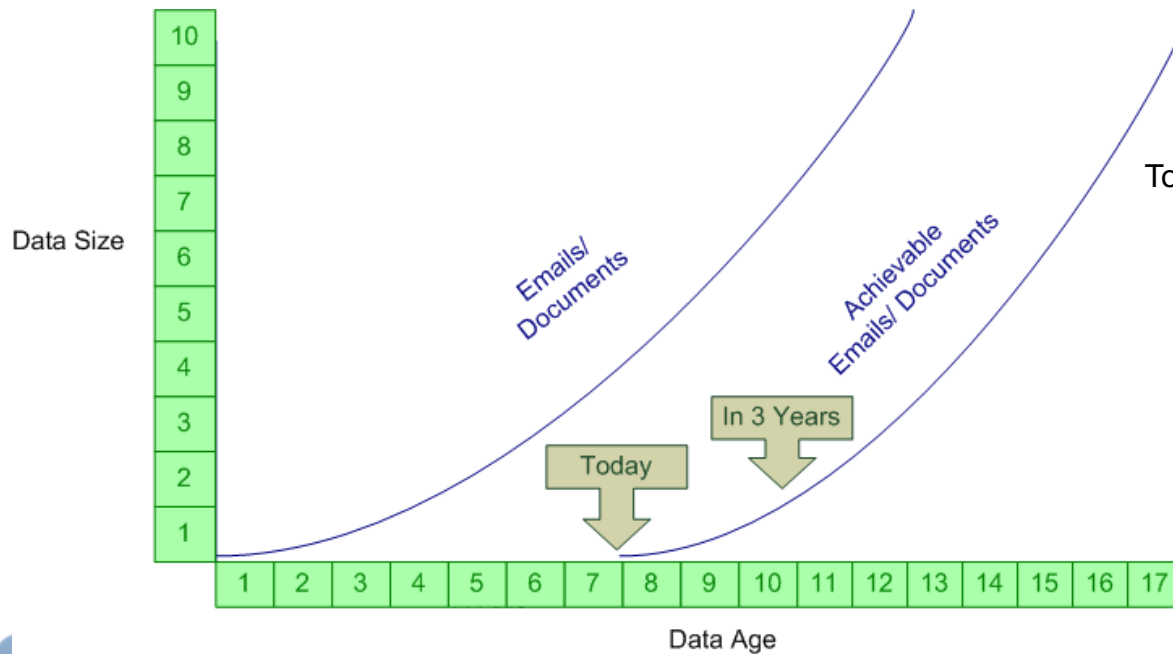
- There is a desire to archive all data over 7 years to a low cost storage medium to remove it from costly online storage
- The business were comfortable with this policy but could not define an online retention period for themselves
- Currently month end copies of the accounts database are kept online by an informal copy to a separate server. These could be archived.
- Due to their large volumes Mail and Worksite DMS have high requirements for archiving

# Archive Data (2011)

Row Labels	Current Used Size	Estimated ize in 2011	Currently available to Remove to archive	Estimated for Removal to archive in 2011
Applications by name anonymised	390.00GB	1905.35GB	193.00GB	1509.36GB
<b>Enterprise Vault</b>	856.92GB	8502.69GB		318.68GB
Applications by name anonymised	1028.05GB	1964.73GB		244.29GB

Archiving month end copies of Accounts

Based on 3/7 of current dataset size i.e.  
First 3 year of current data.



Today not much data is greater than 7 years old  
In 3 years more data will be greater than 7 years and available for archive.

# Assumptions

# Data Assumptions

Assumptions				
<b>Business Growth</b>	10%	1.10		Used for calculating growth when no other number available. Year on Year
<b>Full backup Retention</b>	5			
<b>Incre Backup Rentention</b>	5			
<b>Monthly's kept</b>	Yes			
<b>Amount of existing DMS documents that can be archived in 3 years</b>	43%			
<b>Amount of existing EV that can be archived in 3 years</b>	43%			
<b>Deduplication factor</b>	2	50.00%		
<b>Snapshot Space required (times production space)</b>	3.2			Based on Visio description of policy
<b>Exchange and Worksite Backup Proposal</b>				
<b>Daily</b>	6	On disk	Full	
<b>Weekly</b>	6	On disk	Full	
<b>Month End</b>	13	On disk	Full	
<b>Total Full Backup Copies</b>	25			
<b>Enterprise Vault Backup Proposal</b>				
<b>EV rolling 2 copies full</b>	2		Full	Archive to EV once every week so exchange grows towards end of month by 70GB

# Disaster Recovery

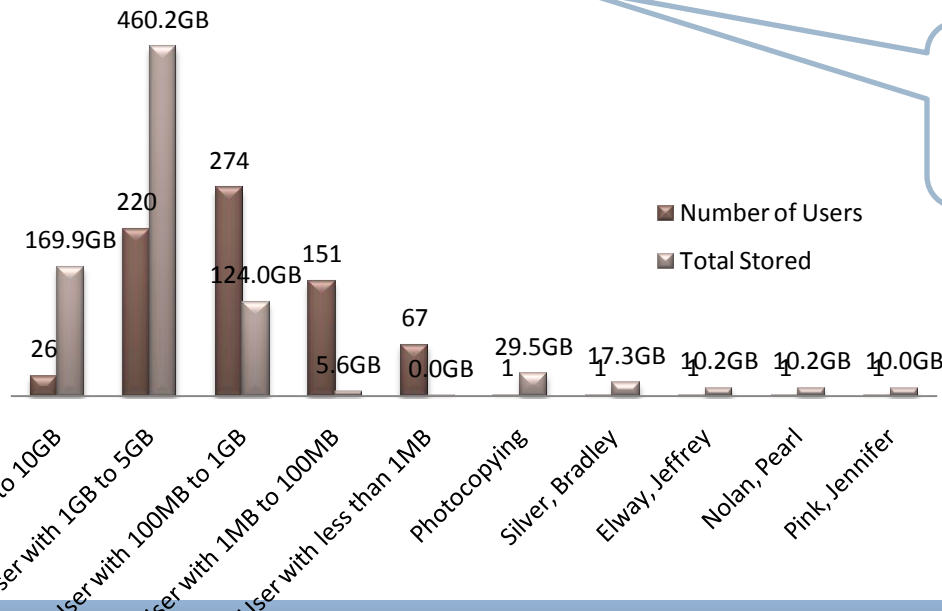
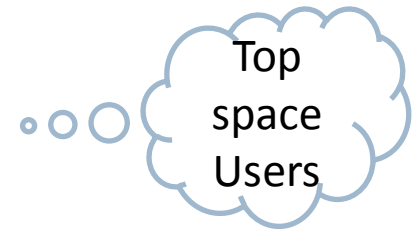
# DR Replication

Application	Criticality	DR Required	Current DR	Data Size
<b>MS Exchange</b>	High	Yes	SAN Replication	560GB
Applications by name anonymised	High	Yes	SAN Replication	1058GB
<b>Domain Controller</b>	High	Yes	Application	40GB
<b>Blackberry</b>	High	Yes	None	40GB
<b>VMware Templates</b>	High	No	SAN Replication	200GB
<b>Intranet Apps</b>	High	No	SAN Replication	52GB

# Enterprise Vault

# Enterprise Vault - Summary

Space Category	Number of Users	Total Stored	% of Stored	% of Users
User with 5GB to 10GB	26	169.9GB	20%	3%
User with 1GB to 5GB	220	460.2GB	55%	30%
User with 100MB to 1GB	274	124.0GB	15%	37%
User with 1MB to 100MB	151	5.6GB	1%	20%
User with less than 1MB	67	0.0GB	0%	9%
Photocopying	1	29.5GB	4%	0%
User names anonymised	1	17.3GB	2%	0%
	1	10.2GB	1%	0%
	1	10.2GB	1%	0%
	1	10.2GB	1%	0%
	1	10.0GB	1%	0%
<b>Grand Total</b>	<b>743</b>	<b>836.8GB</b>	<b>100%</b>	<b>100%</b>

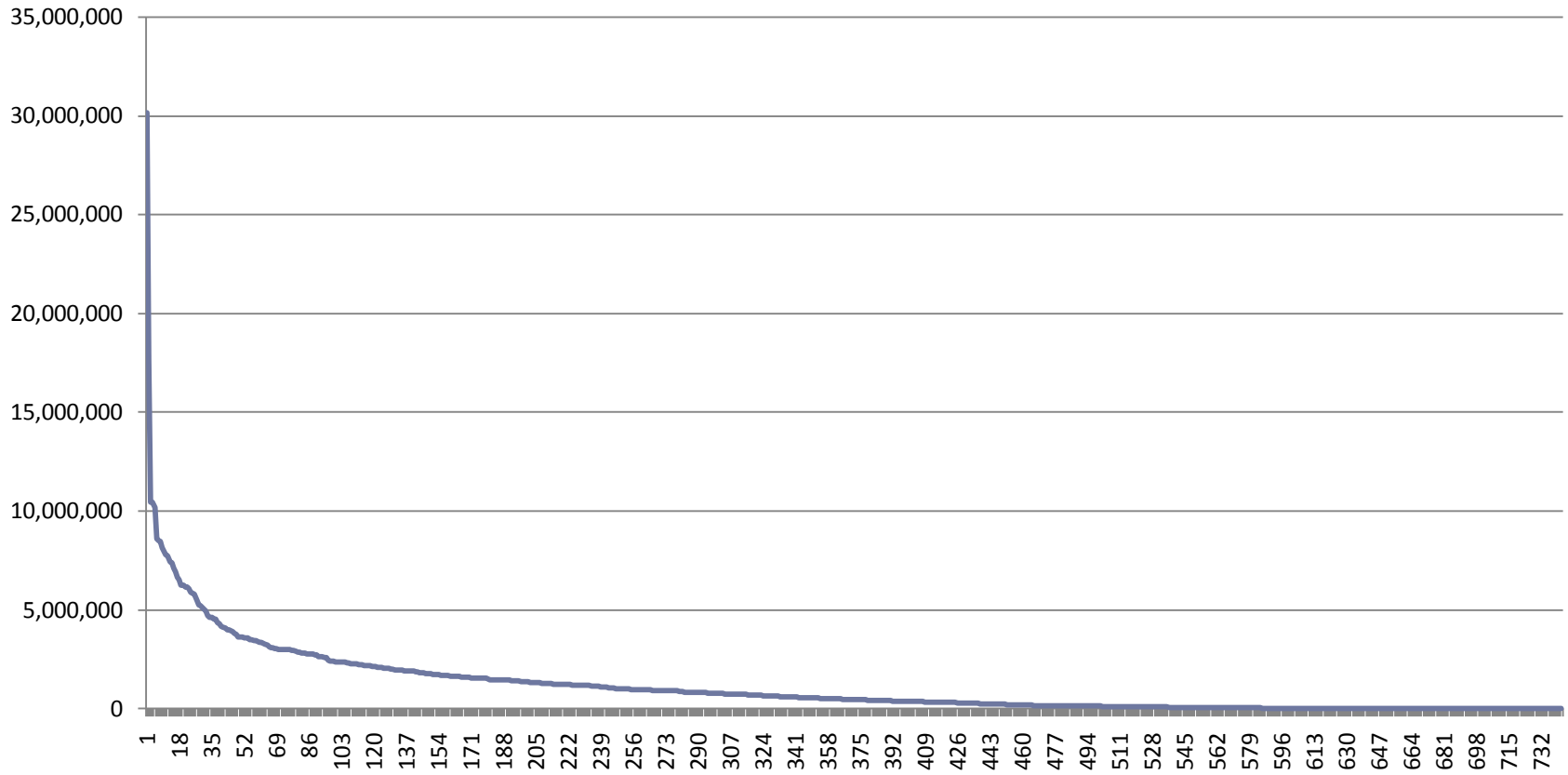


Total User Data

\*Additional 108GB Used by System

# Enterprise Vault – Usage Distribution

Total Size (KB)



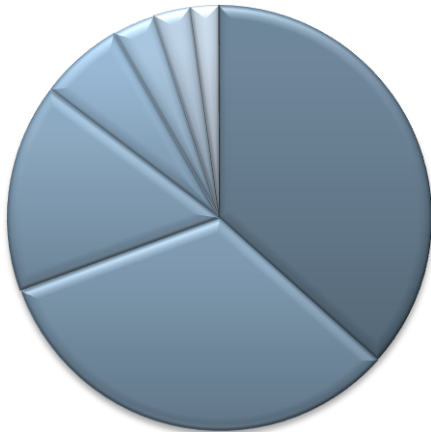
# Home Directories

# User Home Directory Usage

	Number of Users	Total Stored	Percentage Of Users	Percentage of Storage
F:\HomeDir\as	1	5094.9MB	0%	3%
F:\HomeDir\scanning	1	5918.9MB	0%	3%
F:\HomeDir\SGK	1	11604.2MB	0%	6%
F:\HomeDir\njf	1	66181.2MB	0%	37%
User storage <1MB	493	78.2MB	64%	0%
User storing 1MB to 100MB	171	3866.6MB	22%	2%
Users storing 100MB to 1GB	76	28702.6MB	10%	16%
Users storing >1GB	29	59226.8MB	4%	33%
<b>Grand Total</b>	<b>773</b>	<b>180673.4MB</b>	<b>100%</b>	<b>100%</b>



## Total Stored



- \HomeDir\njf
- Users storing >1GB
- Users storing 100MB to 1GB
- \HomeDir\SGK
- \HomeDir\scanning
- \HomeDir\as
- Users storing 1MB to 100MB
- Users storing <1MB

Total User Data

# File System Scan

- User home folders
  - dominated by one user at this stage, using 40% of the disk approximately 60GB.
  - Top ten users utilize 70% of the file server scanned, the remainder of the users which store data on this server utilize 60GB.
  - PST files utilize approximately 40% of the file server. It is recommended that they are imported into Enterprise Vault.
  - Graphics files is the next largest file group. 30%
  - Office files utilise the least amount of disk approximately 20%.
- Space allocation.
  - From the sample reports collected it has been recommended that disk allocation should be reinvestigated; some servers have been allocated large amounts of disk while the utilisation on these boxes is still quite low.

# Observations

# Observations

- Users are generally happy with level of service.
  - However some do have very demanding expectations
- Desktops
  - performance of Worksite is an issue.
  - not locked down and venerable to data loss
- Backups
  - fail frequently
  - High risk “opt in” policy
- Data Retention
  - No policy. Work in progress.
  - Mail archive backup policy does not protect against undetected data corruption as it is not being written to a worm device.
  - Accounts month end policy required.
- Mail is biggest storage problem.
  - Archiving disk solutions

# Observations

- Rate of change is not controlled
  - Too many things not being finished. Databases all over the place.
  - Project management is not evident
  - Good local product skills but lack of documentation
- Management Information weak
  - No tools & reporting weak
  - No overall shared view of the estate
- Disaster Recovery
  - Everybody has own idea of priorities.
  - No recovery plan.
  - Appear to be in line with business.
- Strategy
  - Have a strategy.
  - Solid VMware strategy.
  - Storage – everything on the SAN
  - Migrating to central database cluster.