Pulse2011

Best Practices for Migration to Tivoli Storage Manager for Your Data Protection Needs

Joe Cho
joecho@au1.ibm.com
Tivoli Storage Technical Specialist

Pulse2011

Analysis Engine Report (AER)

AER Overview

- Provides a comprehensive view of the current backup and recovery environment
- Visualise the size and complexity of your current environment
- Analyse the exposure to risk and test RTO and RPO objectives
- Based on real, empirical data not guesstimates

AER Process

1. Client buy-in
2. Complete Pre-requesite Questionaire (PRQ)
3. Collector created/delivered
4. Collection is performed on each backup server
5. Collection output submitted for analysis
6. Report produced and presented to client

Analysis Engin Pre-requisite Questionaire (PRQ)

- The PRQ is used at the initial stage to allow which Butterfly collectors are to be used
- The PRQ also gives data centre context around the backup infrastructure

Server name	OS	HW	SW	Ver	Location
SERVER1	Windows Enterprise Server 2003 SP2 64 bit	DL380 G4 x86	Symantec NetBackup	5.1	DATA CENTRE 1
SERVER2	Windows Enterprise Server 2003 SP2 64 bit	DL380 G4 x86	Symantec NetBackup	6.0	DATA CENTRE 2

PRQ Questions

- Please supply the number of Backup Servers
- Please supply the operating system and version of backup servers
- Please supply the hardware platform of backup servers
- Please supply the version of backup software on each backup server
- Please supply the physical location of each backup servers
- Please confirm CLI access to each backup server

Collector Delivery

- The collector will be constructed from the completed PRQ
- The required collector software will be automatically generated and placed in the Butterfly Portal
- This will occur within a 30 day period of confirmed receipt of the completed PRQ

Collector Execution

- The collector software must be placed in a temporary area and executed on the customer backup servers as defined in the PRQ.
- There must be the available space in the user home directory to capture the output data.
- The space required for the collector output could be up to 2GB per backup management server
- The collector must be executed by an Administrative user

Collection Process

- Collector execution will take up to 20 minutes on a mid sized system
- The collector progress bar will move as the collector phases complete
- The collector gathers all data and structures it in the encrypted output file
- On completion, the collector output and all traces should be removed from the backup server

Collector Output

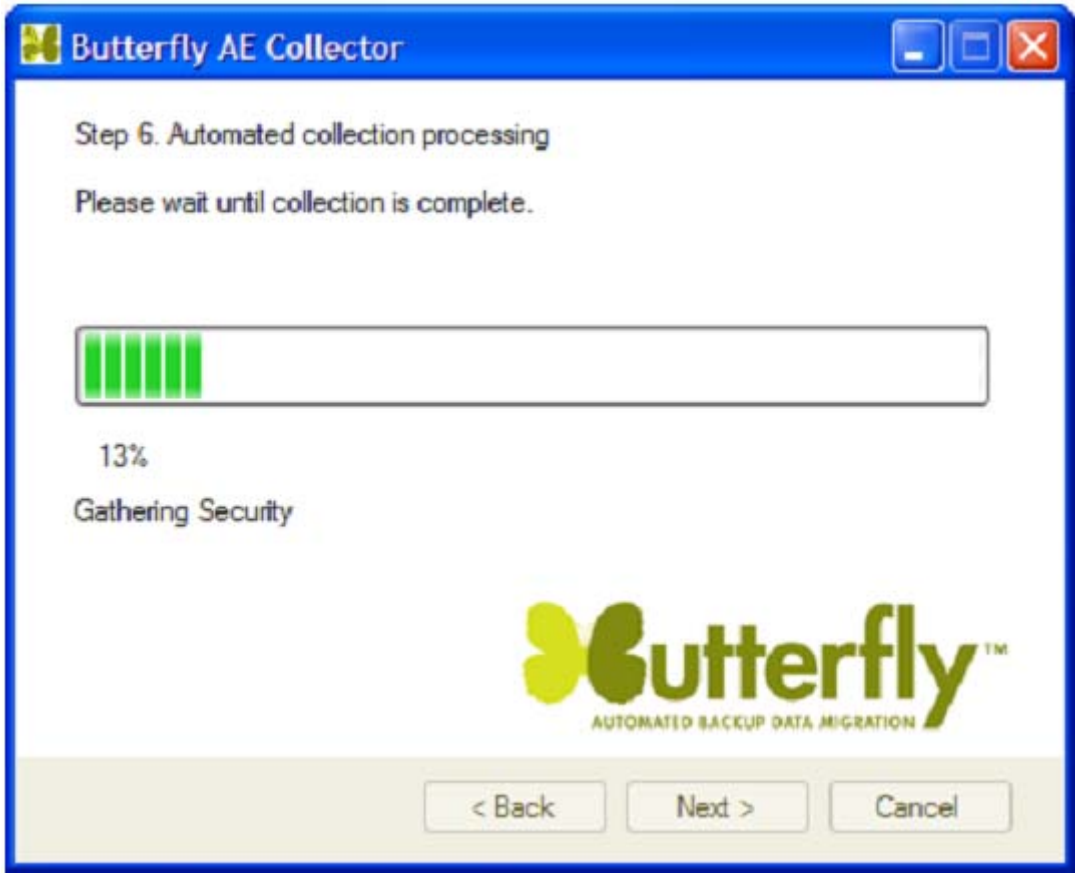
- This output file is encrypted allowing safe transition of the data to the portal
- The output files, including the collectior; should be completely removed from the backup server

Collection Process

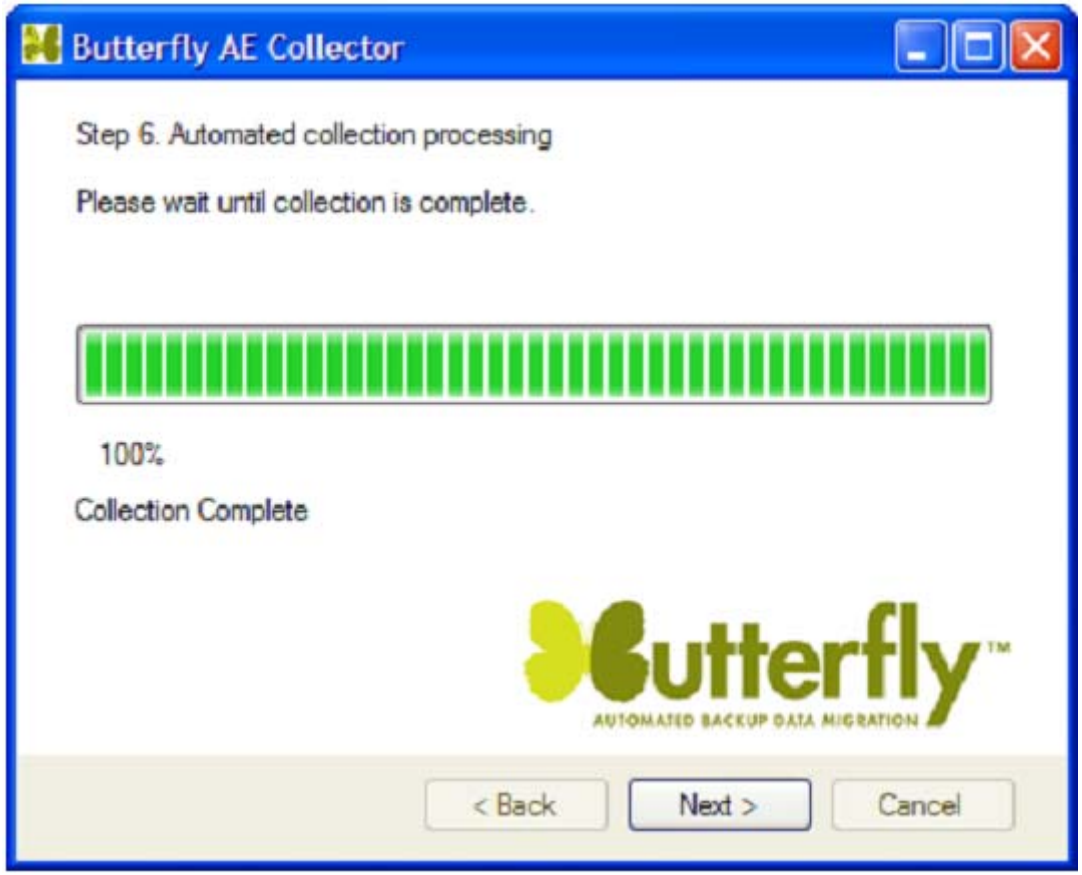
Collector Process - cont

Collector Process - cont

Collector Process - cont



Collector Process - cont



Collector Process - cont

Analysis Engine Report Delivery

- AER will be delivered within 5 business days of the confirmed receipt of the final collector output files
- AER will be uploaded to the portal and made available for download
- AER will be in PDF format
- It is advisable that this is printed by the representative on A0 size poster.

AER Presentation

- Presentation will be given by IBM Sales team
- AER format lends itself to the following presentation style:
  - Relaxed, consultative approach
  - Interactive session
  - AER on wall for discussion and interrogation
  - Not Power Point format!

Butterfly
Butterfly Analysis Engine- Report for Example Customer
SOURCE NetBackup Environment
Source Software Architecture
Source Hardware Architecture
Source Client Environment
Backup Cycle
Data Occupancy by Platform
Data Occupancy by Retention
Datacentre A
Media Servers
Master Servers
Datacentre B
IRON MOUNTAIN Offsite store
Manual Vaulting
TSM Library Manager
TARGET IBM TSM Environment
Indicative Target Architecture
Capacity Metrics
Top 20 Migration Sources by Volume
Daily Backup Volume DC-A
Daily Backup Volume DC-B
Daily File Count Summary
Occupancy by Backup Server
Butterfly AER Details
Document Name: Backup Migrator AnalysisEngine Example Customer
Author: Butterfly AnalysisEngine v1r (release 5.5)
Butterfly Differential Business Case
SOURCE Hardware Infrastructure
TOTAL COST OF OWNERSHIP 36 MONTHS
TOTAL MEDIA COUNT 36 MONTHS
Media Count
Occupancy
TSM Incremental Benefit
FIX FORWARD
FIX LEGACY
Management Summary
TARGET Hardware Infrastructure
Operational Issues Resolved
Infrastructure Issues
Operational Issues
Butterfly Migration Engine™ Planning
Operating Systems in Scope of Migration
Data Volume in Scope of Migration
Top 10 clients in Scope of Migration

AER – Source Environment

SOURCE NetBackup Environment

Source Software Architecture

- Source software environment based on **Symantec NetBackup 5.0**
- Software release date **May 2004**
- Environment been in production for **5 years 3 months**
- Existing environment based on dual **NBU** Master Servers per physical datacenter
- Multiple **media servers** to allow backup of mixed Operating Systems
- Primary and offsite tape pools are used for media and site protection
- FULL backup methodology and policy enforced throughout the environment
- Some adaptive differential Image backups
- 4* NBU Master Servers addressing a total of **454** clients
- Data retention policy varied from **30 days** to **7 years**. No standard offering is created.

Source Hardware Architecture

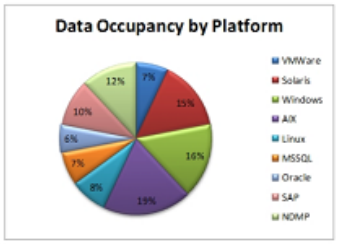
- Master Server technology based on **SUN SPARC** architecture
- Tape libraries in each site are **IBM 3584** with a mix of **LTO2** and **LTO3** generation tape drives
- Media in environment based on **LTO1**, **LTO2** and **LTO3** tape
- Total ONSITE Volumes **7734**
- Total OFFSITE Volumes **9507**
- Total environment library slot capacity is **9784**
- Primary and offsite tape pools are used for media and site protection
- FULL backup methodology and policy enforced throughout the environment
- The **IRON MOUNTAIN** offsite hosted vaulting is used for DR media
- Data retention policy varied from 1 month to 7 years. No standard offering is created
- Other unmanaged legacy standalone physical tape devices

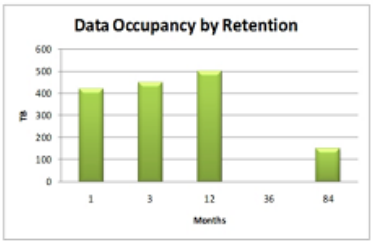
Source Client Environment

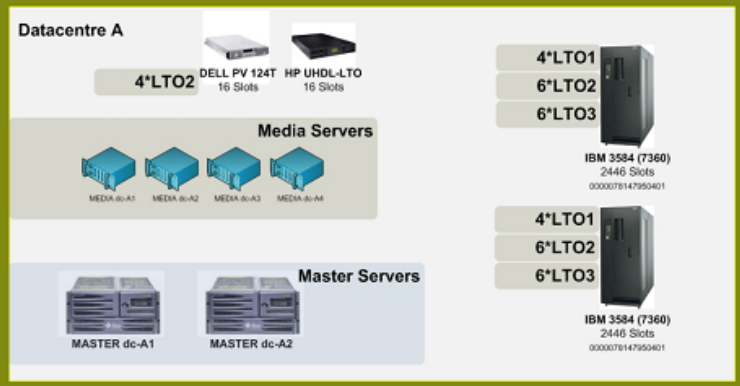
- Client operating system platforms include **Windows**, **Solaris** and **LINUX**
- Extensive use of **VMWARE ESX 3.5** for Windows virtualised environment
- Currently **454** Backup clients currently executing backup operations
- Client environment 32% VMWare, 47% Unstructured data, 21% Structured data
- Structured data types are **MS SQL**
- File services provisioned from Windows x86 based infrastructure
- Total client backup data occupancy is **1520 TB**

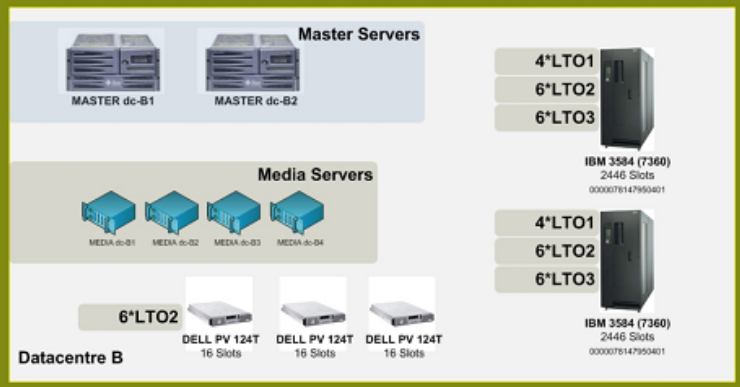
Backup Cycle

- FULL backup daily for 30 day retention
- DIFFERENTIAL weekly for 12 weeks retention
- FULL backup monthly for 12 month retention
- FULL backup yearly for 7 year retention
- ADAPTIVE DIFFERENTIAL cycles irregular





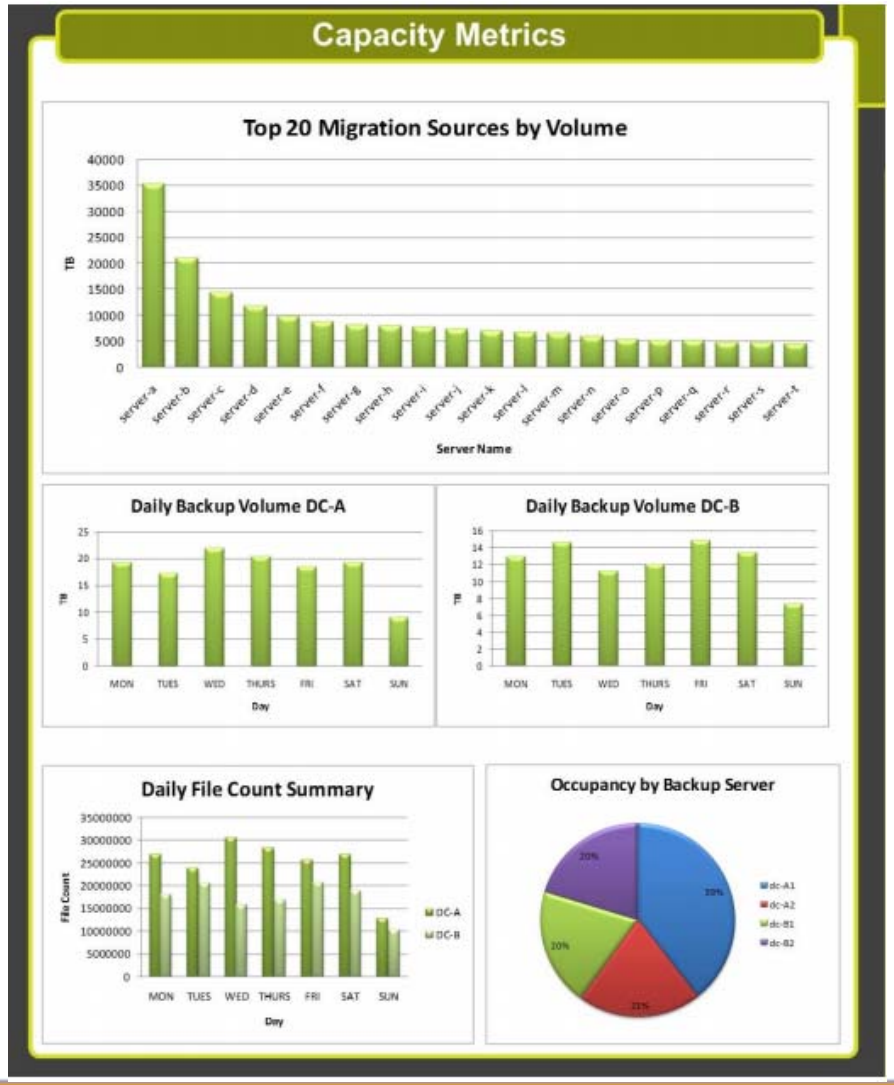




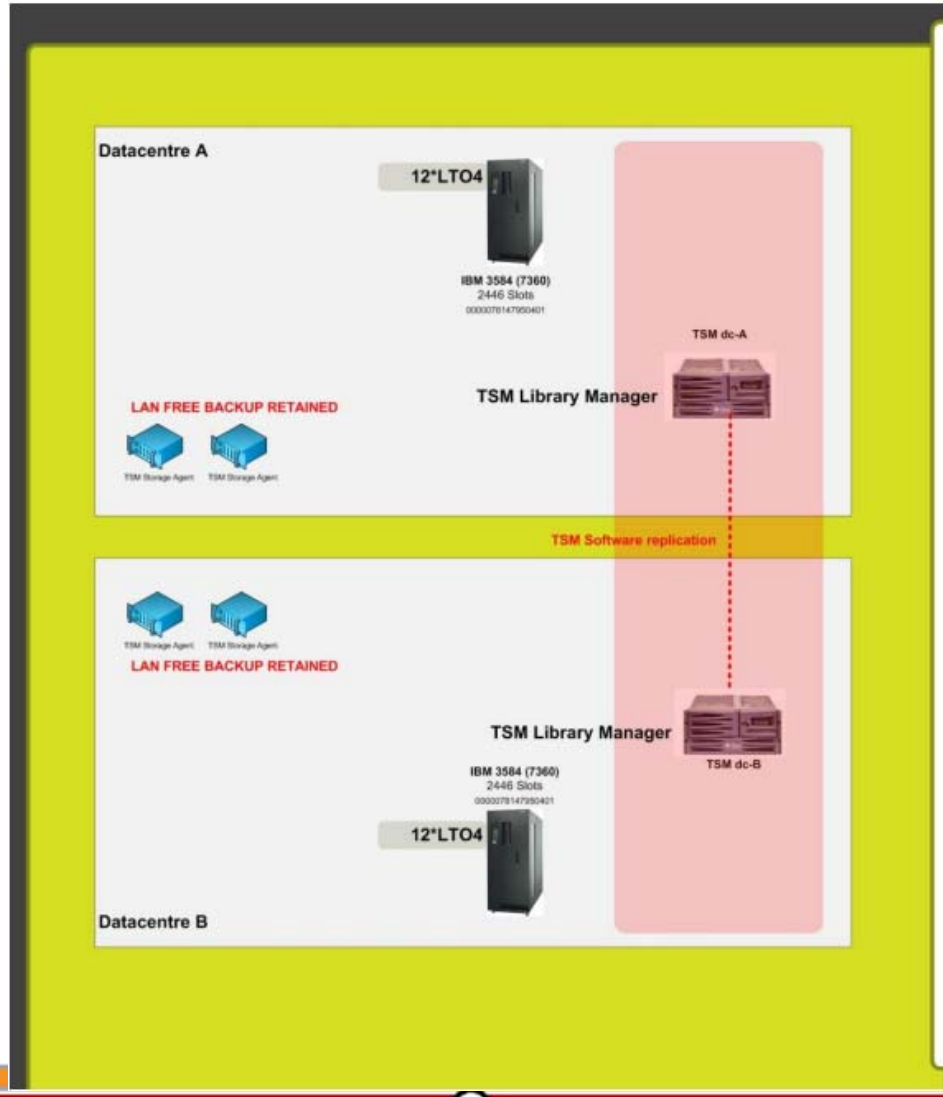
9507 Tape Volumes

Manual Vaulting

AER - Source Environment Cont



AER – Target Environment



TARGET IBM TSM Environment

Indicative Target Architecture

Hardware Infrastructure

- Consolidated server infrastructure and removal of media mount entities
- Removal of all stand alone physical tape infrastructure
- **Total tape storage consolidation**
- TSM infrastructure managing replication of backup data
- **Total removal of offsite manual vaulting**
- Removal of all manual tape operations

TSM Software Design

- **TSM v6** to be used as the management server
- TSM VCB integration to be used for VMWare backups
- TSM disk storage pool and VTL based tape storage pools to be used
- TSM incremental backup to be used
- Image backup for fast File server backup and recovery and increased deduplication benefits

Tape Library/Media Infrastructure

- Re-use of existing TS3500 chassis
- Media refresh for LTO4 tape media
- Library consolidation and TSM Library management

Architectural Changes

- Data backup to be incremented by TSM software
- Vaulting to be automated and kept within data centres
- Incremental backup methodology to reduce network/SAN load and data occupancy by 70%
- TSM Deduplication to **reduce total occupancy** of legacy data solution
- Consolidation resulting in **simplified operation and management**
- **Decommission existing widespread legacy tape infrastructure**

Operational Benefits

- Reduction of number of backup and recovery entities and associated headcount
- Simplification of management of disparate technology elements
- Total consolidation of physical tape
- Total removal of high-risk manual offsite vaulting
- Total removal of manual tape movements
- Reduced media numbers due to increased data density and incremental impact

Environmental Benefits

- Data centre space savings
- Data centre associated power and heat dissipation benefits

Environment Capability

- Backup data Primary and Secondary copy retained in secure data centre locations
- Automated online vaulting of backup data

AER – Operational Issues

Operational Issues Resolved

Infrastructure Issues

- 46 **Frozen physical tape volumes** due to excessive read and write errors
- **LTO interoperability** issues within the library frames
- High volume of physical offsite tape media with associated **offsite storage costs**
- Offsite volumes in primary library slots
- **Failed physical tape drives**: Drive03, Drive 12 in 3584 Library Data centre A
- Un-configured tape drives, installed but not used
- LTO1 tape drives with no LTO1 primary volumes
- Backup processes for FC backup hindered by **insufficient physical tape mount points**
- Restore processing delayed due to physical tape mount oversubscription
- Network and Fabric usage resulting in **inefficient tape throughput**- less that expected figures

Operational Issues

- Source software environment based on **legacy** Symantec NetBackup code version
- Software **upgrade required** to retain the support contract on the environment
- Infrastructure age means **increased failures** on library robotics
- Physical nature of backup servers has resulted in **unbalanced server infrastructure**
- Over running backup window due to insufficient mount points
- **Manual reruns** of failed backups consuming resources
- Manual Start Of Day checks required due to **unstable physical tape** infrastructure

AER – Business Case

Butterfly Differential Business Case

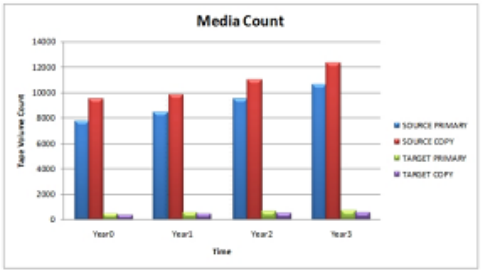
SOURCE Hardware Infrastructure

SITE	INFRASTRUCTURE	VENDOR	MODEL	QTY	KVA
A	Library	DELL	124T	1	3.2
A	Library	HP	UHDL	1	2.9
A	Library	IBM	3584	2	6.4
A	Master Server	SERVER	MASTER	2	14
A	Media Server	SERVER	MEDIA	4	17.6
A	Drive	IBM	LTO1	10	1
A	Drive	IBM	LTO2	16	1.6
A	Drive	IBM	LTO3	12	1.2
B	Library	DELL	124T	3	9.6
B	Library	IBM	3584	2	6.4
B	Master Server	SERVER	MASTER	2	14
B	Media Server	SERVER	MEDIA	4	17.6
B	Drive	IBM	LTO1	10	1
B	Drive	IBM	LTO2	18	1.8
B	Drive	IBM	LTO3	12	1.2
			TOTAL		99.5

TOTAL COST OF OWNERSHIP 36 MONTHS

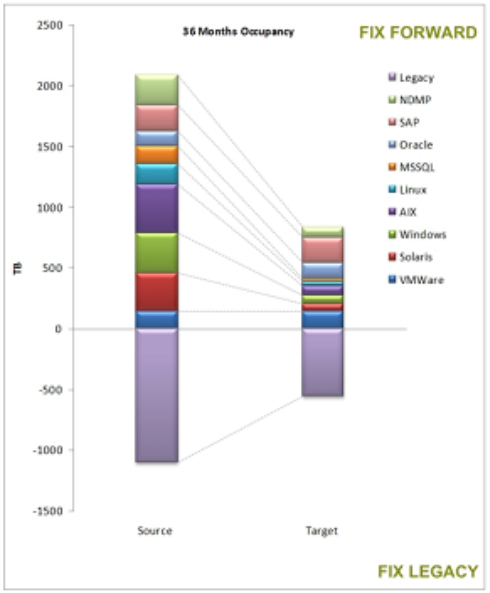
INFRASTRUCTURE	SOURCE UNITS	COST	TARGET UNITS	COST
TAPE VOLUMES	5722	$ 257,467.78	1221	$ 54,940.81
VAULT SLOTS	12334	$ 99,559.90	0	$ -
LIBRARY	9	$ 181,316.40	2	$ 70,274.56
TAPE DRIVES	78	$ 241,647.12	24	$ 179,352.96
MASTER SERVER	4	$ 192,451.20	2	$ 146,225.60
MEDIA SERVER	8	$ 181,510.08	0	$ -
FTEs	4	$ 1,500,000.00	3	$ 1,125,000.00
TOTAL		$ 2,653,952.48		$ 1,575,793.93
OCCUPANCY (TB)	2089		837	
POWER (KVA)	100		23	
Total Saving		$ 1,078,158.55		

TOTAL MEDIA COUNT 36 MONTHS



Occupancy

TSM Incremental Benefit

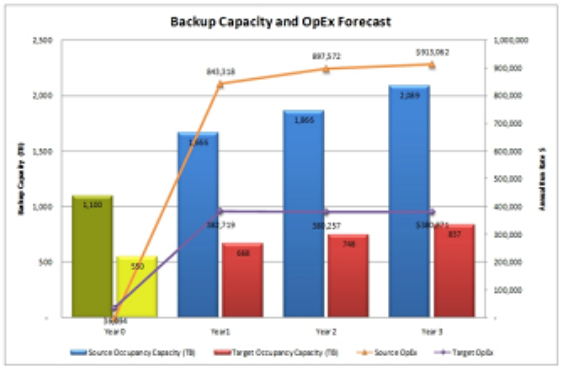


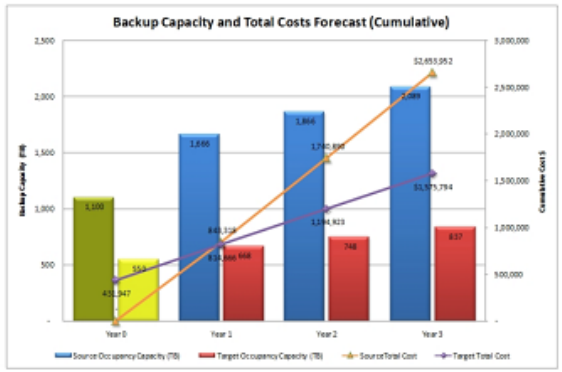
Management Summary

- FIX FORWARD Occupancy reduced by **1252** TB
- FIX LEGACY Occupancy reduced by **550** TB
- **TOTAL 36 MONTH Occupancy reduced by 1802 TB**
- Incremental ratios calculated on DATA TYPE, SCHEDULE, BACKUP TYPE, BACKUP FREQUENCY, PLATFORM TYPE and industry metrics
- Tape volume saving of **2324** volumes
- Saving of **70** KVA in dual data centre
- Daily change rate is reduced by **22 TB** allowing a calculated incremental ratio
- Improved **RECOVERY** and **DR** times from the removal of offsite vaulting
- Improved **SECURITY** by keeping all data onsite.
- Network load reduced by **59.92%**
- FTE reduction by **25%**
- Host data volume Growth Rate applied **10%**

TARGET Hardware Infrastructure

SITE	INFRASTRUCTURE	VENDOR	MODEL	QTY	KVA
A	Master Server	SERVER	MASTER	1	7
A	Library	IBM	TS3500	1	3.2
A	Drive	IBM	LTO4	12	1.2
B	Master Server	SERVER	MASTER	1	7
B	Library	IBM	TS3500	1	3.2
B	Drive	IBM	LTO4	12	1.2
			TOTAL		22.8





AER – Business Case Cont.

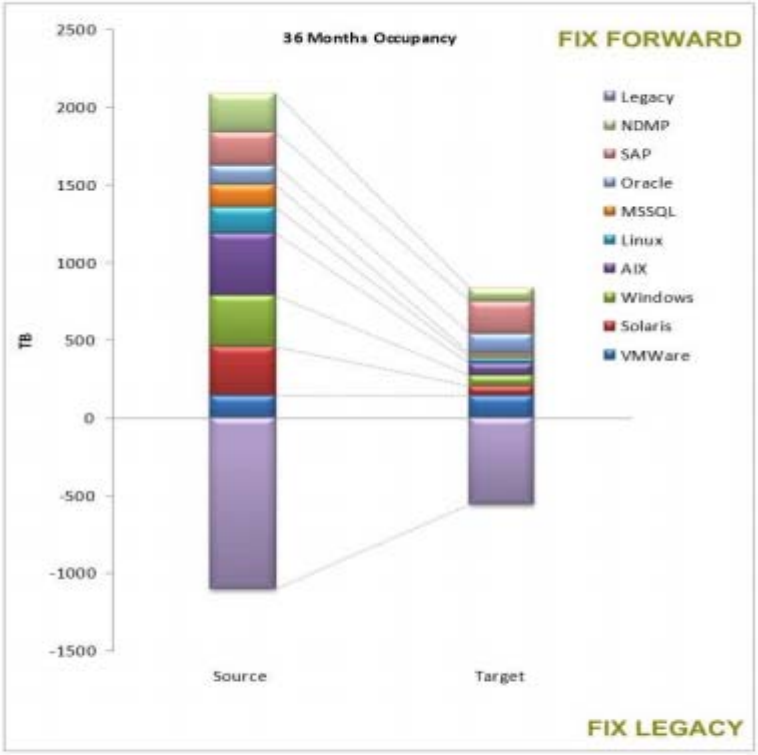
SOURCE Hardware Infrastructure

SITE	INFRASTRUCTURE	VENDOR	MODEL	QTY	KVA
A	Library	DELL	124T	1	3.2
A	Library	HP	UHDL	1	2.9
A	Library	IBM	3584	2	6.4
A	Master Server	SERVER	MASTER	2	14
A	Media Server	SERVER	MEDIA	4	17.6
A	Drive	IBM	LTO1	10	1
A	Drive	IBM	LTO2	16	1.6
A	Drive	IBM	LTO3	12	1.2
B	Library	DELL	124T	3	9.6
B	Library	IBM	3584	2	6.4
B	Master Server	SERVER	MASTER	2	14
B	Media Server	SERVER	MEDIA	4	17.6
B	Drive	IBM	LTO1	10	1
B	Drive	IBM	LTO2	18	1.8
B	Drive	IBM	LTO3	12	1.2
				TOTAL	99.5

TARGET Hardware Infrastructure

SITE	INFRASTRUCTURE	VENDOR	MODEL	QTY	KVA
A	Master Server	SERVER	MASTER	1	7
A	Library	IBM	TS3500	1	3.2
A	Drive	IBM	LTO4	12	1.2
B	Master Server	SERVER	MASTER	1	7
B	Library	IBM	TS3500	1	3.2
B	Drive	IBM	LTO4	12	1.2
				TOTAL	22.8

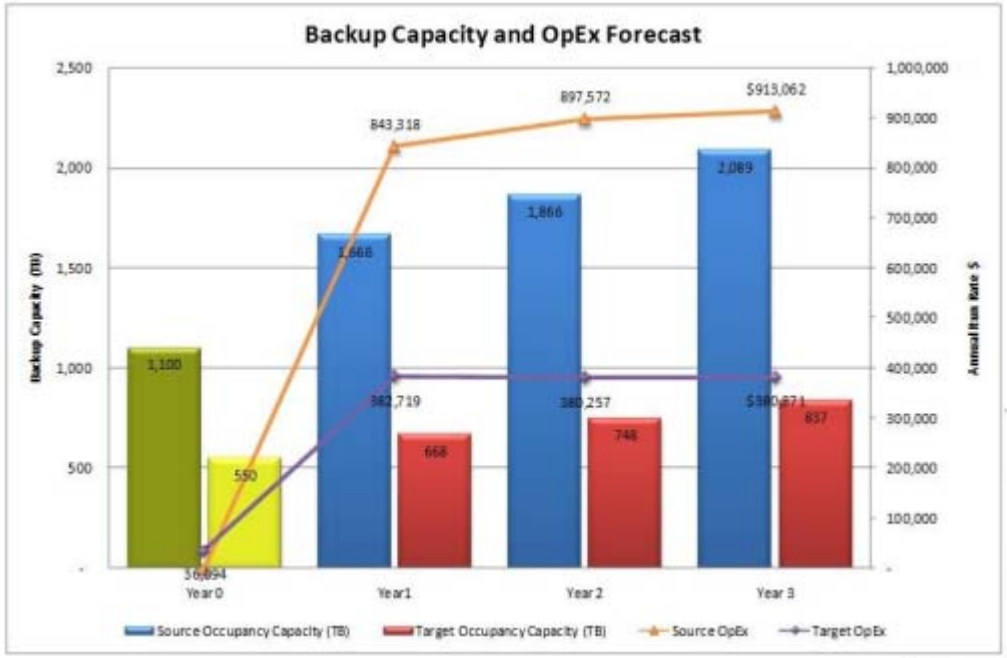
Occupancy

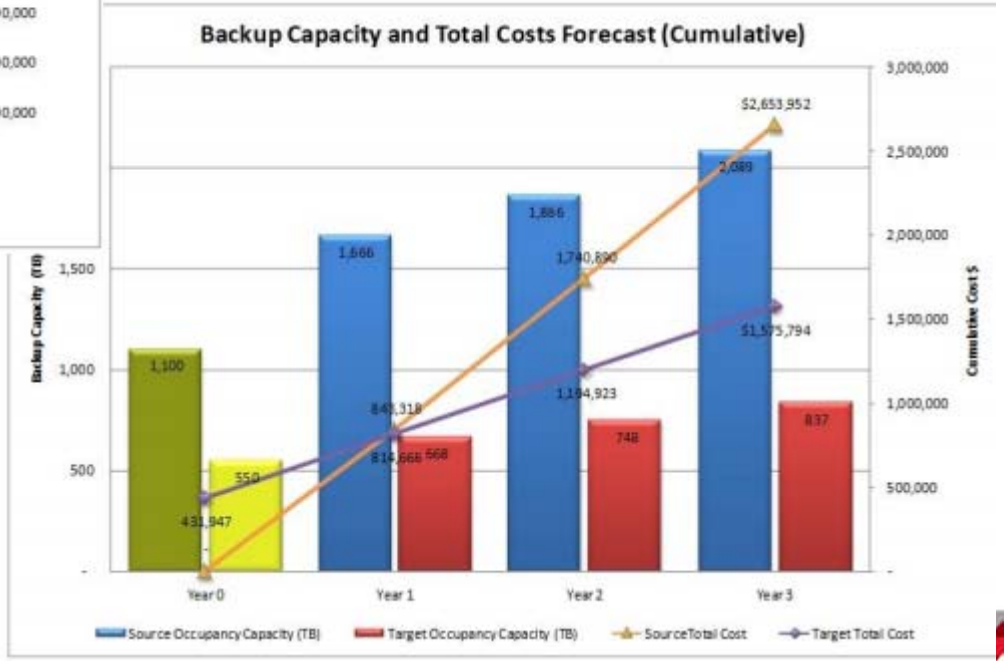


Management Summary

- FIX FORWARD Occupancy reduced by **1252** TB
- FIX LEGACY Occupancy reduced by **550** TB
- **TOTAL 36 MONTH Occupancy reduced by 1802 TB**
- Incremental ratios calculated on DATA TYPE, SCHEDULE, BACKUP TYPE, BACKUP FREQUENCY, PLATFORM TYPE and industry metrics
- Tape volume saving of **4501** volumes
- Saving of **70** KVA in dual data centre
- Daily change rate is reduced by **22 TB** allowing a calculated incremental ratio
- Improved **RECOVERY** and **DR** times from the removal of offsite vaulting
- Improved **SECURITY** by keeping all data onsite.
- Network load reduced by **59.92%**
- FTE reduction by **25%**
- Host data volume Growth Rate applied **10%**

AER - Business Case Cont.

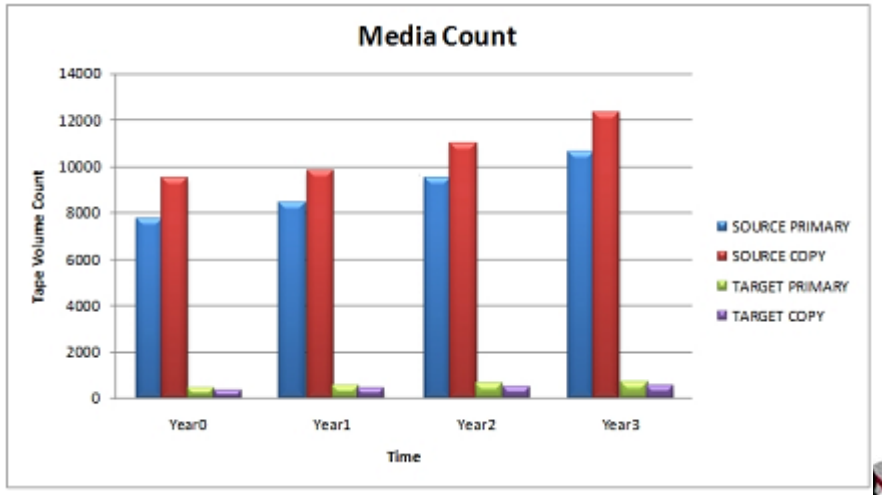




AER – Business Case

- Total Cost of Ownership over 36 months

INFRASTRUCTURE	SOURCE UNITS		COST	TARGET UNITS		COST
TAPE VOLUMES	5722	$	257,467.78	1221	$	54,940.81
VAULT SLOTS	12334	$	99,559.90	0	$	-
LIBRARY	9	$	181,316.40	2	$	70,274.56
TAPE DRIVES	78	$	241,647.12	24	$	179,352.96
MASTER SERVER	4	$	192,451.20	2	$	146,225.60
MEDIA SERVER	8	$	181,510.08	0	$	-
FTEs	4	$	1,500,000.00	3	$	1,125,000.00
TOTAL		**$**	**2,653,952.48**		**$**	**1,575,793.93**
OCCUPANCY (TB)	2089			837		
POWER (KVA)	100			23		
Total Saving		**$**	**1,078,158.55**			



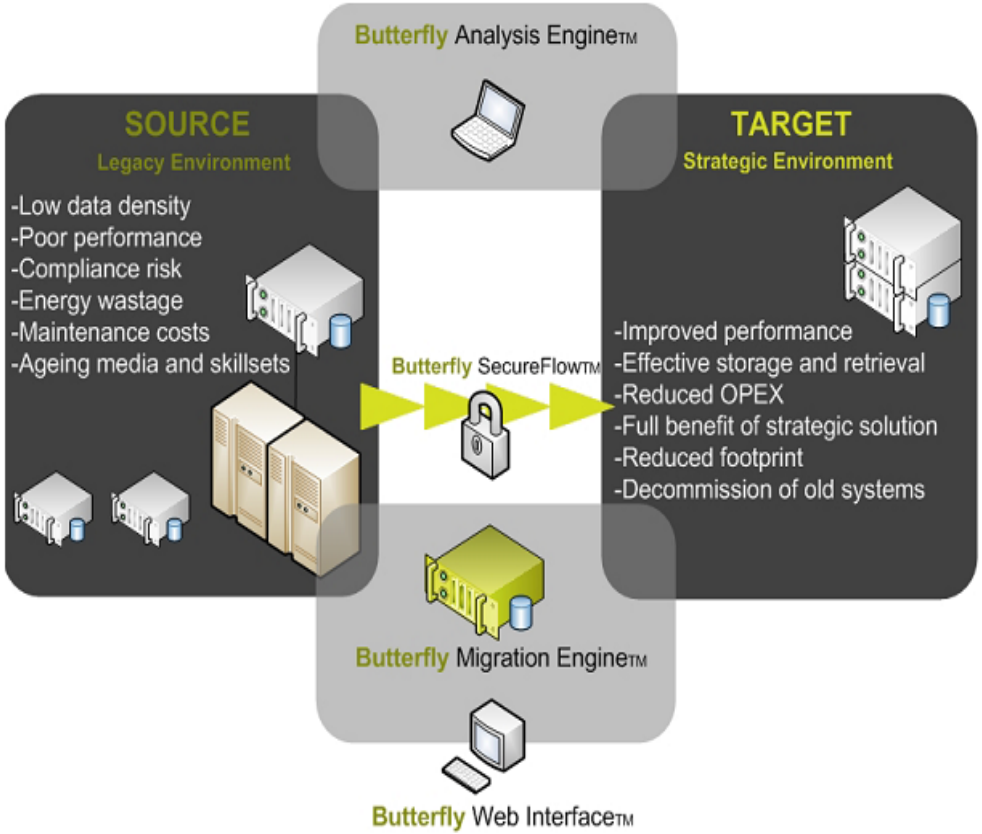
Migration Engine

Why Migrate

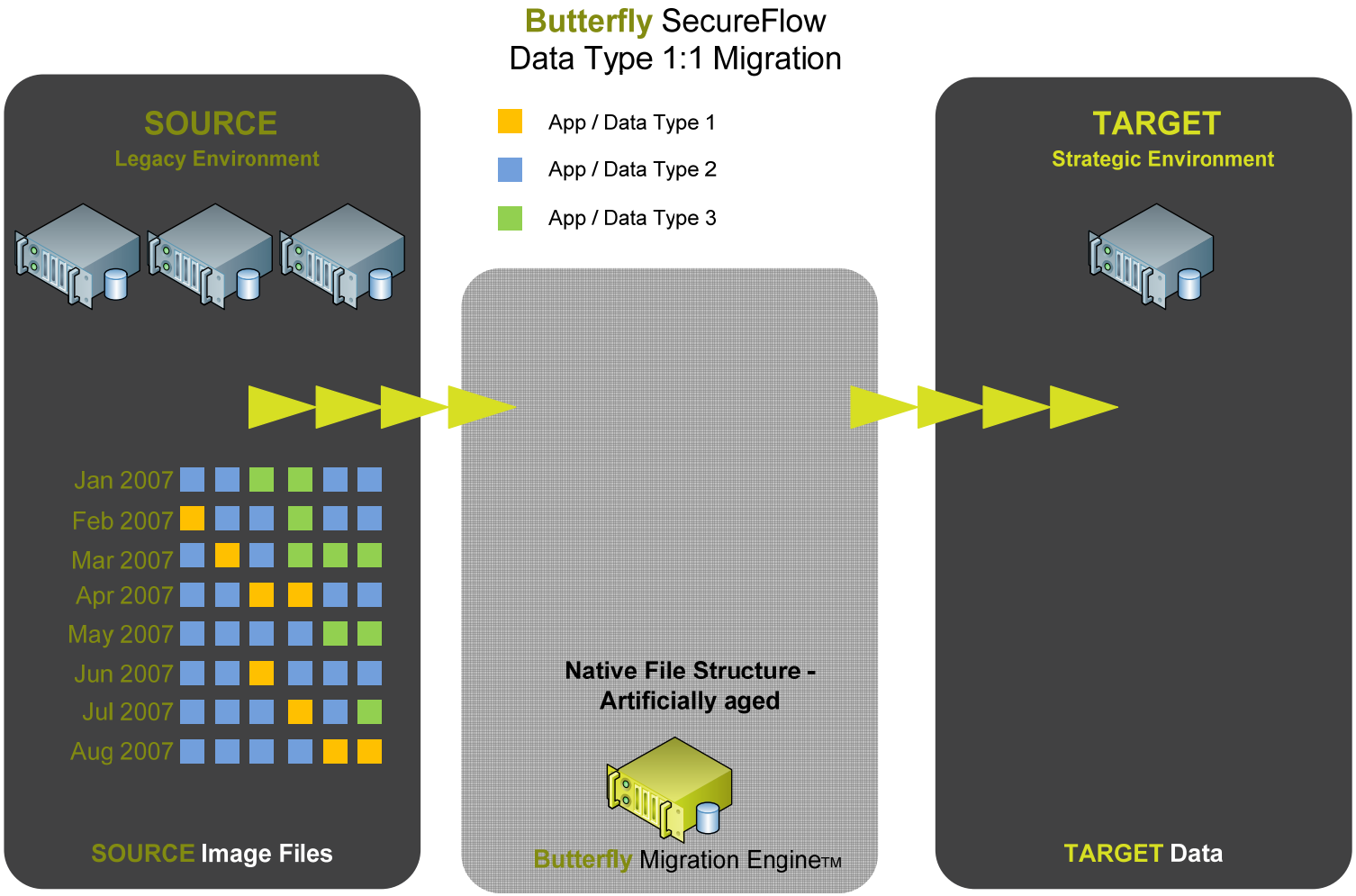
- Many orananisations have "moth balled" legacy environments resulting in cost, risk and operational issues. These include maintenance charges, legacy skills, data centre space, performance impact and media obsolescence
- Long term data retention does not fit with most technoloy lifecycle models. This results in several generations of infrastructure being required to retain backup data.
- Manual migrations can be costly and lengthy procedures and can introduce error and risk which means data is lost or corrupted during the migration phase.

Butterfly Migration Engine

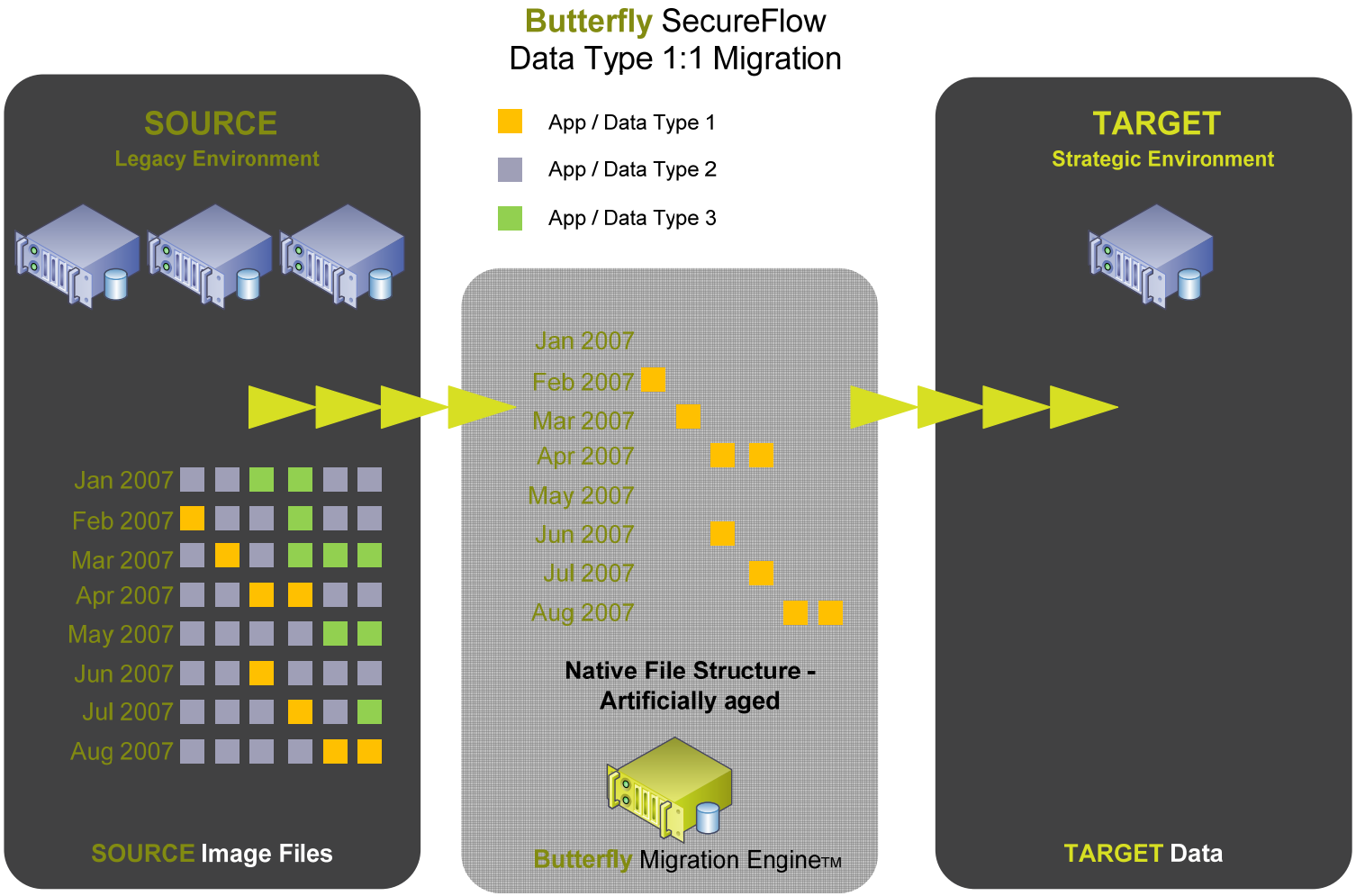
- A dedicated application server is provided with a single point-of-control web interface.
- The migration schedule is automatically generated from customer preferences.
- The Migration Engine holds source data securely until it is ready to be written to the target environment.
- The source data can be de-duped, encrypted, compressed or excluded to maximise the efficiency of the target environment.



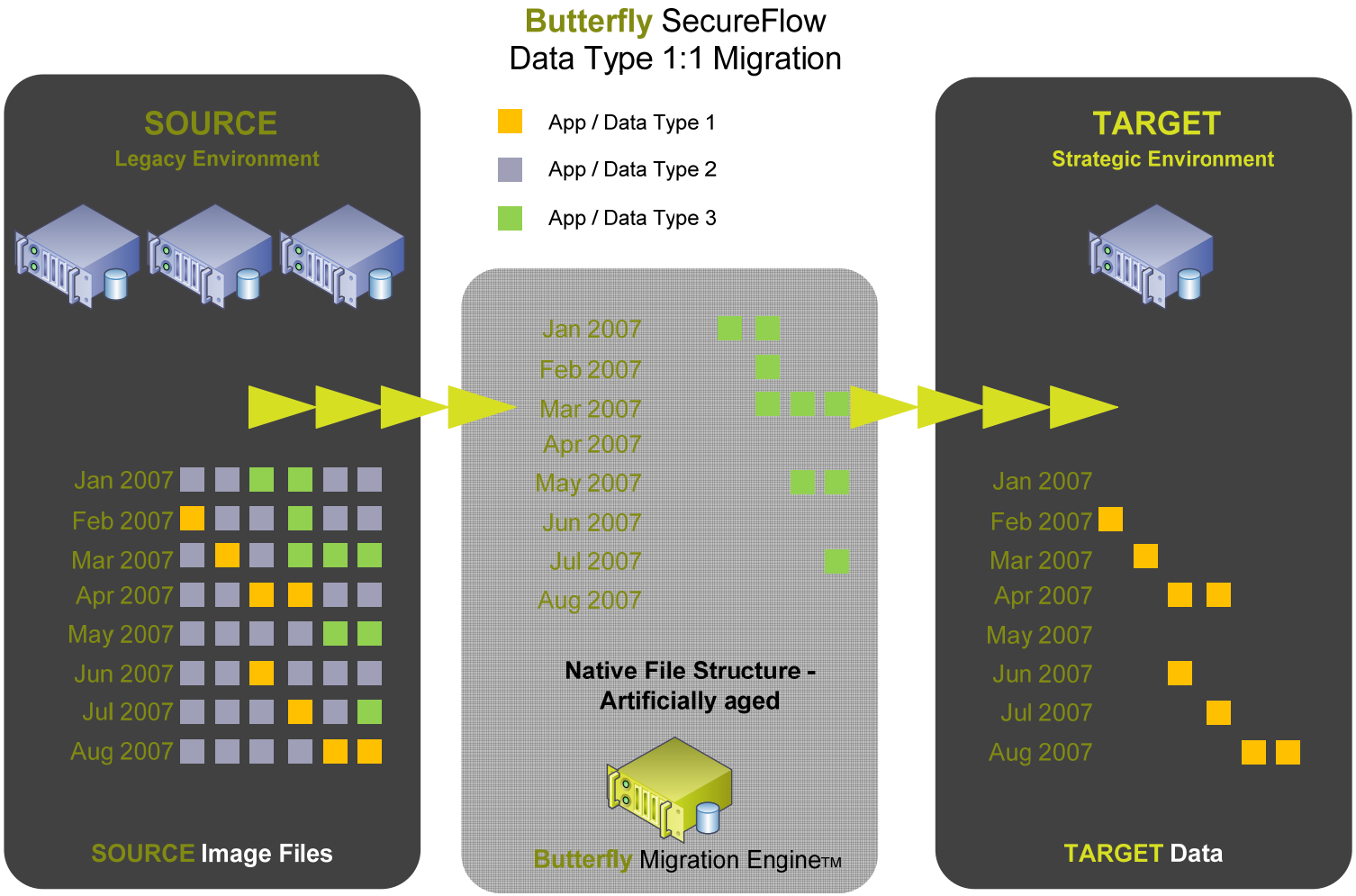
Initial Stage



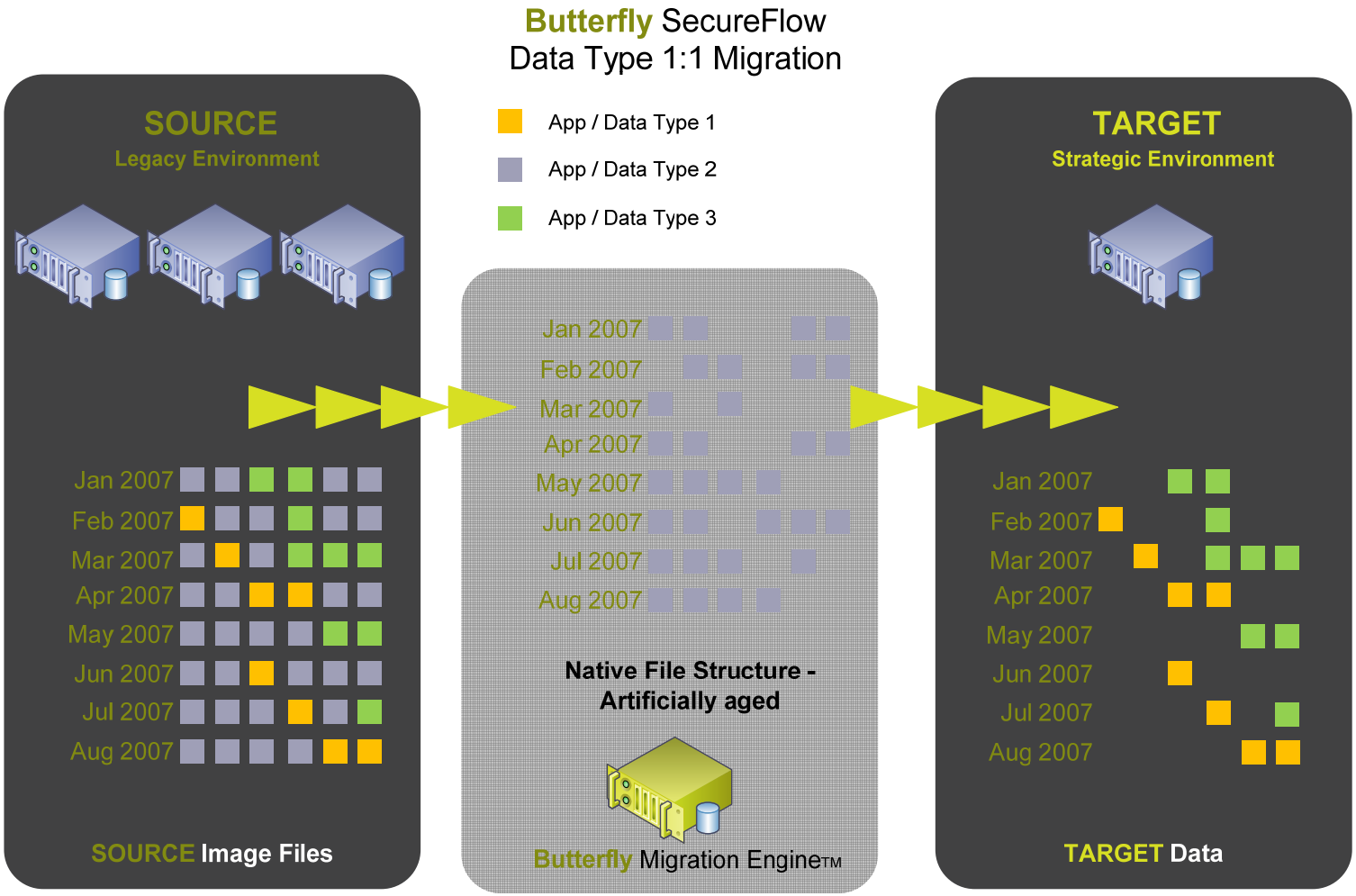
Stage 2



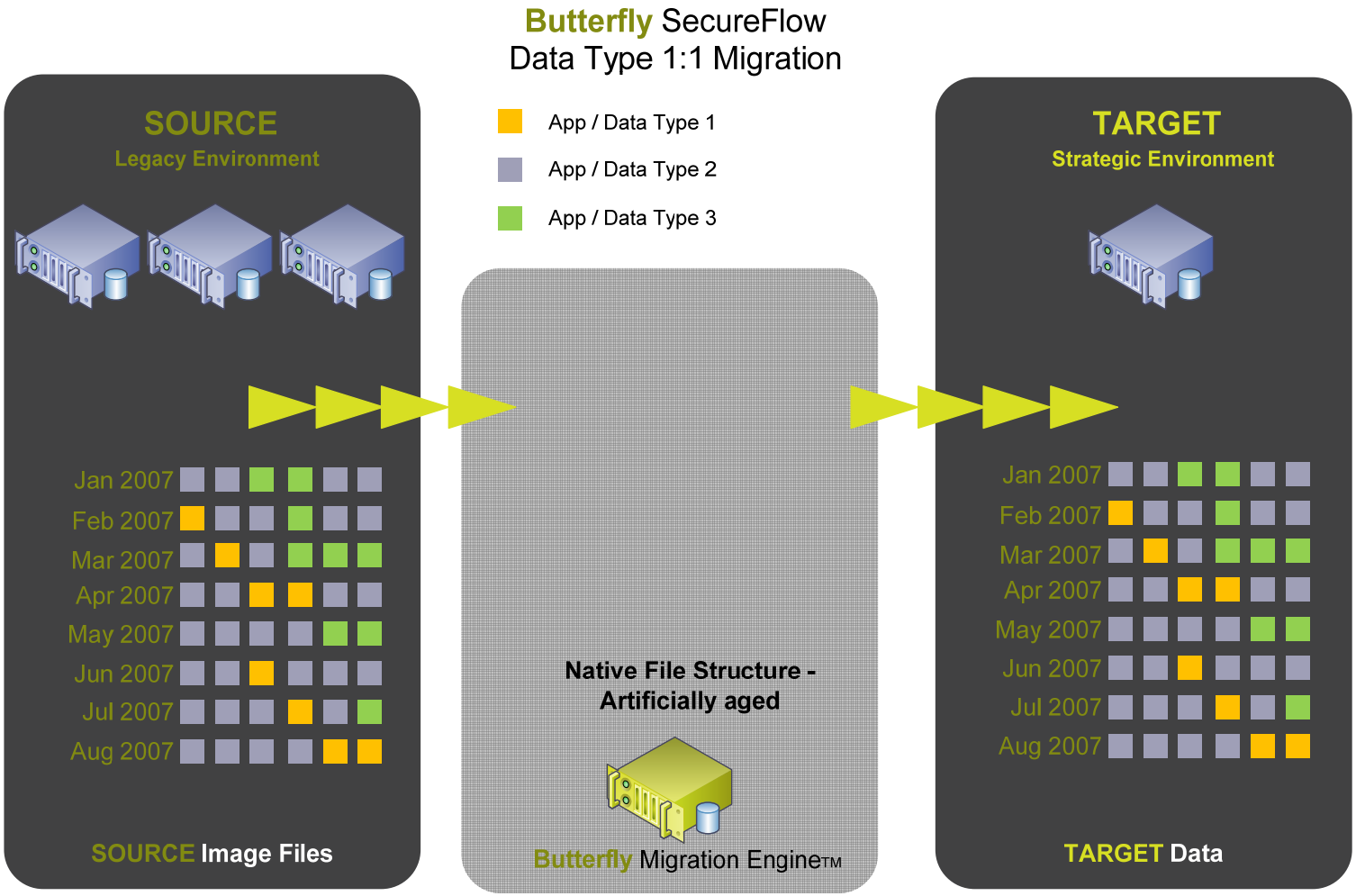
Stage 3



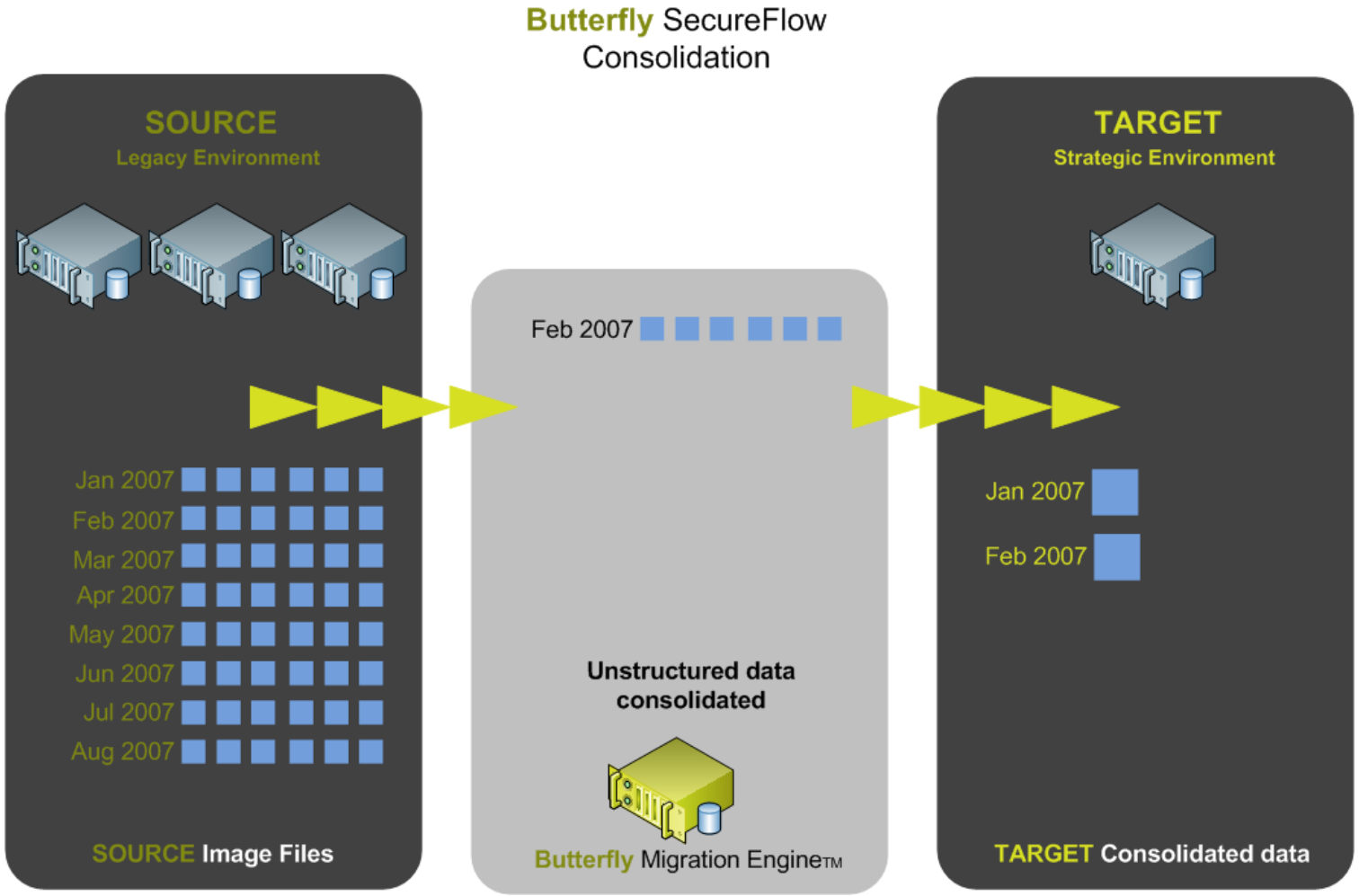
Stage 4



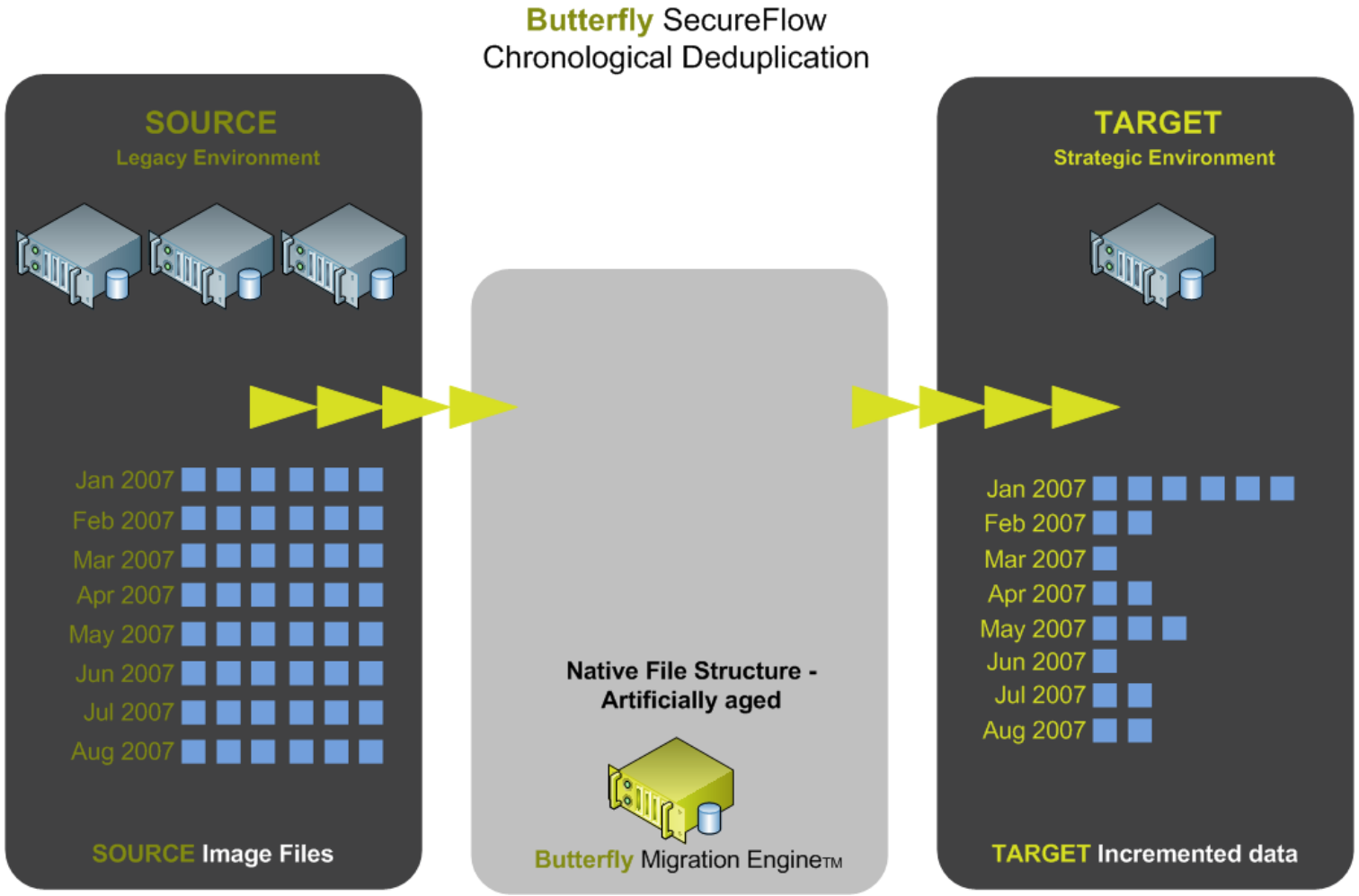
Completion



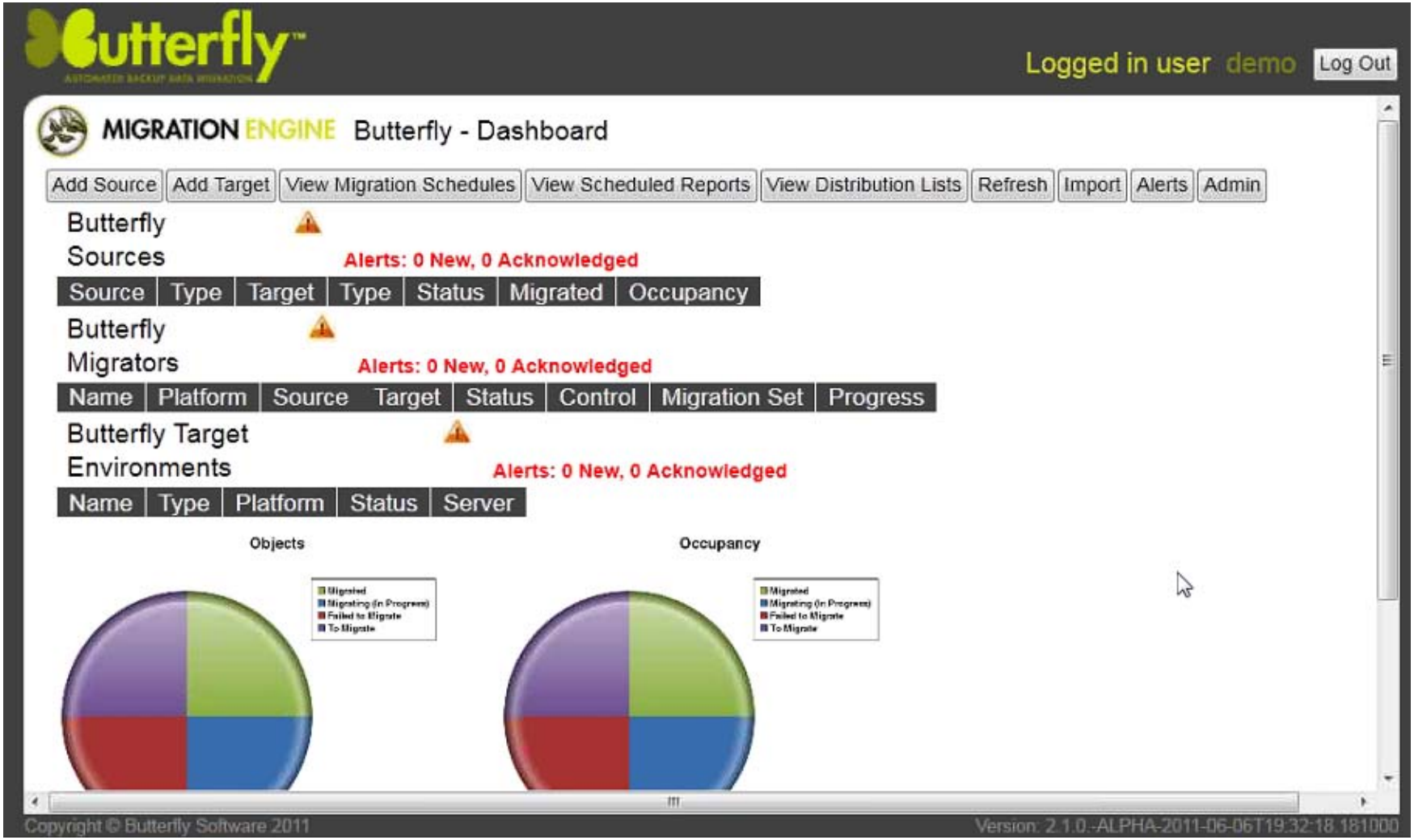
Migration Options - Consolidation



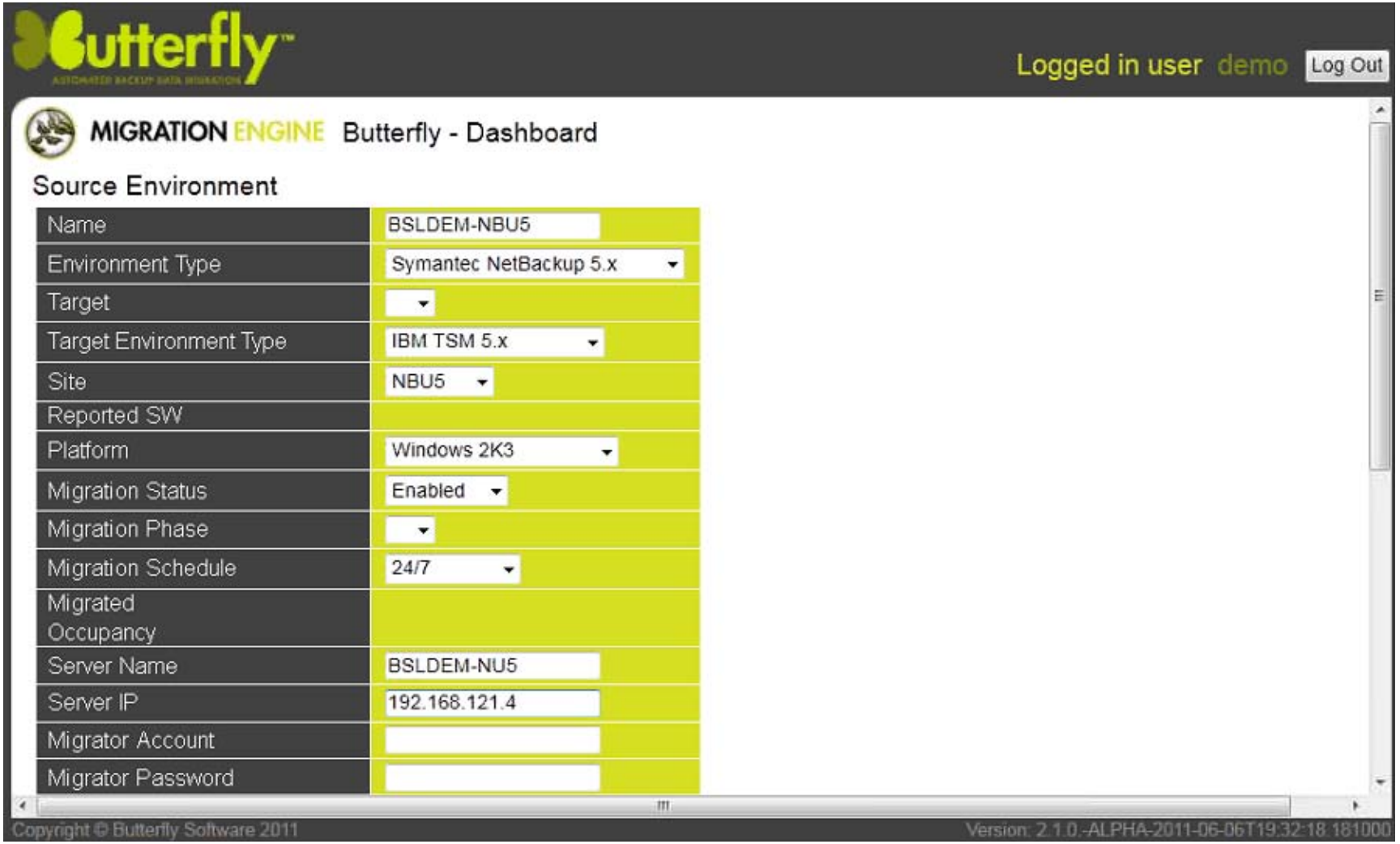
Migration Options - Deduplication



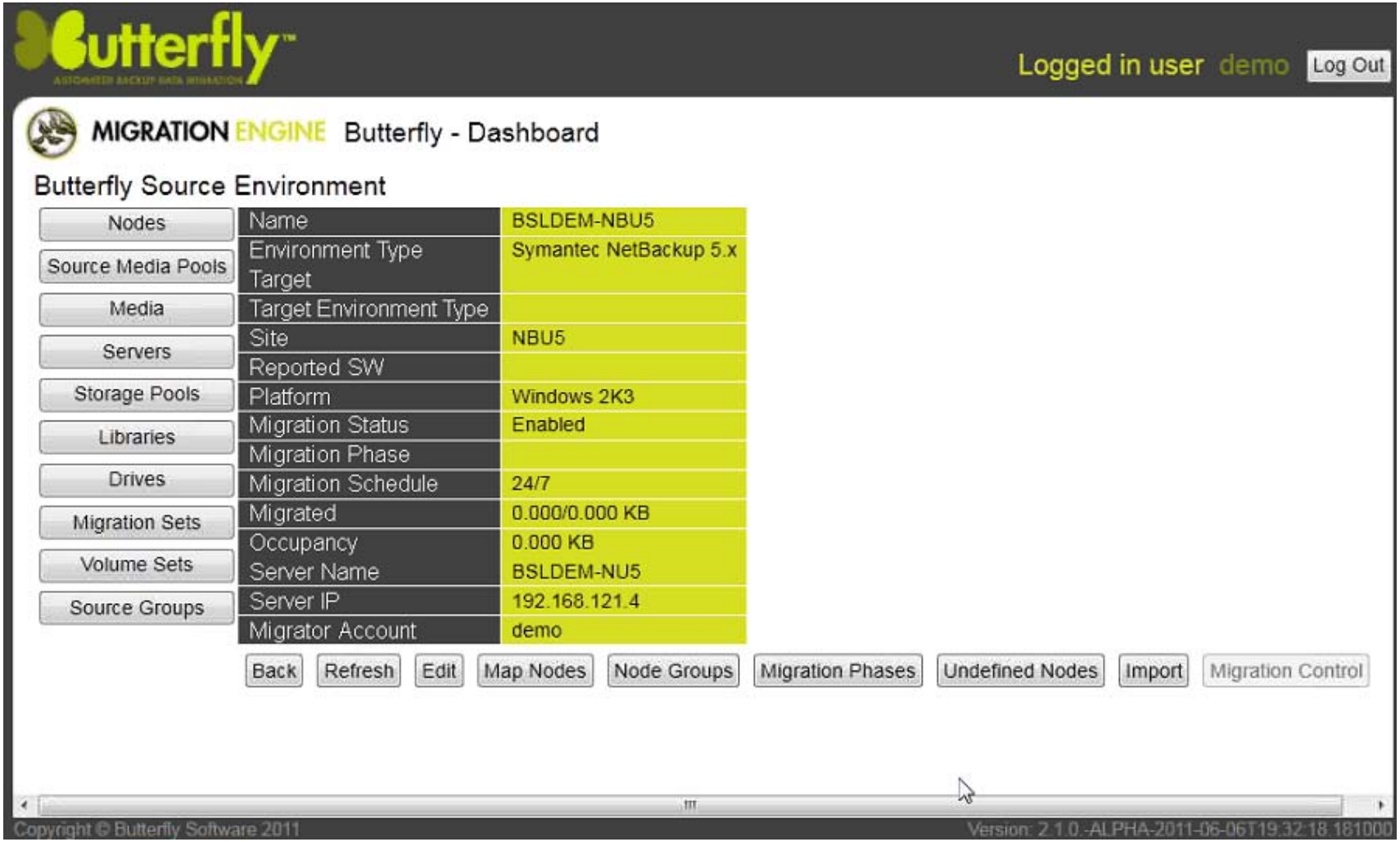
Migration Dashboard



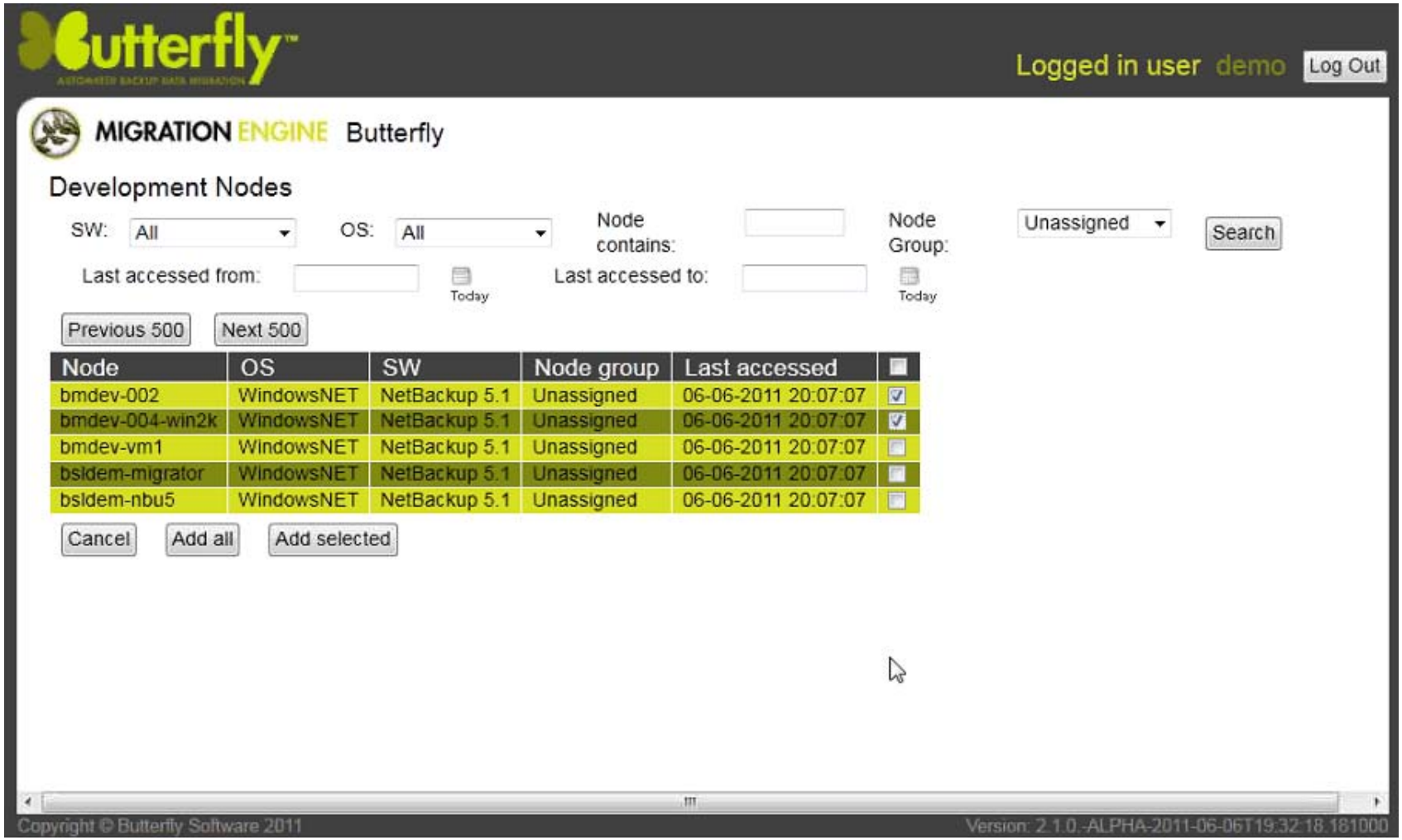
Identifying Source Environment



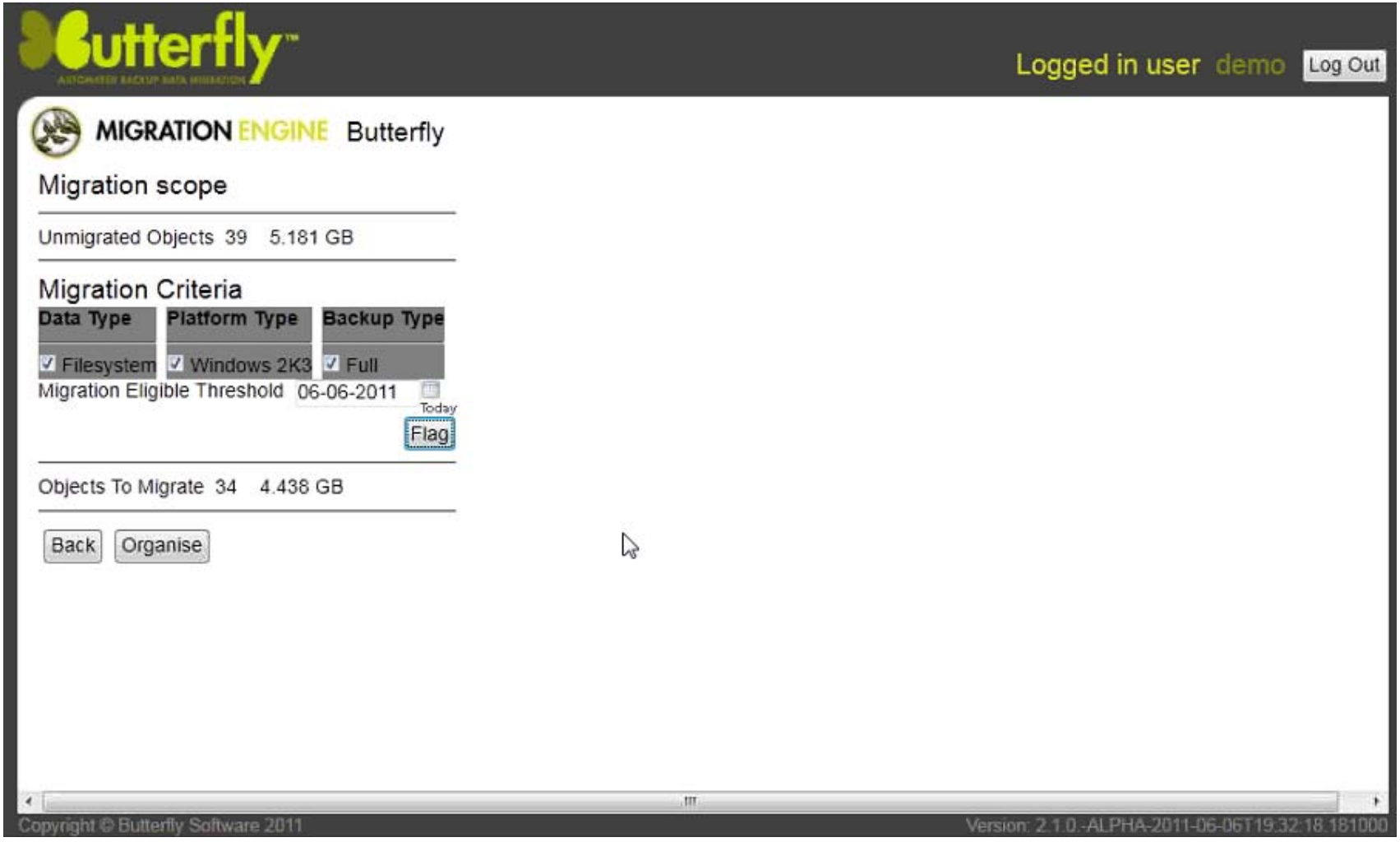
Import Source Information



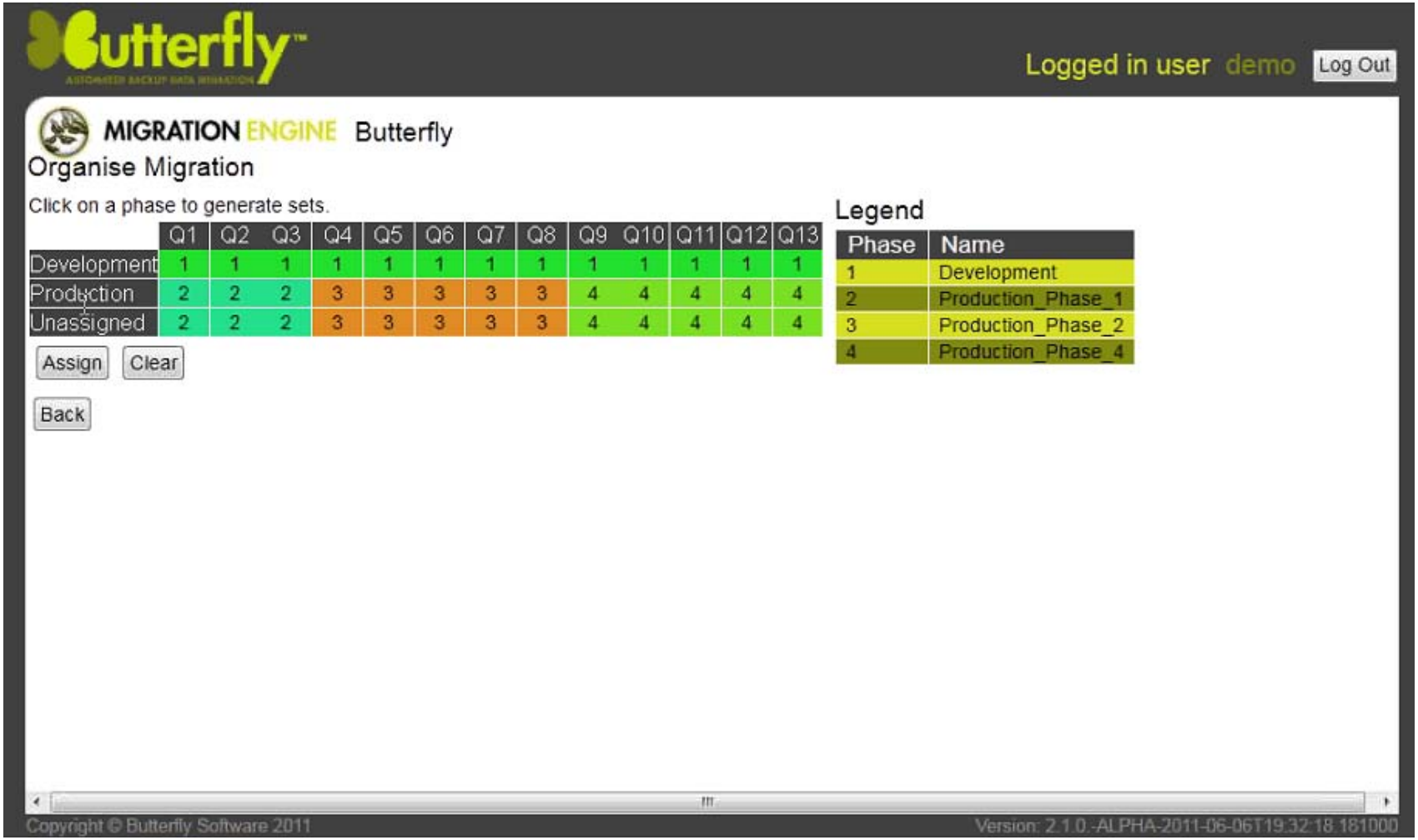
Define Node Groups



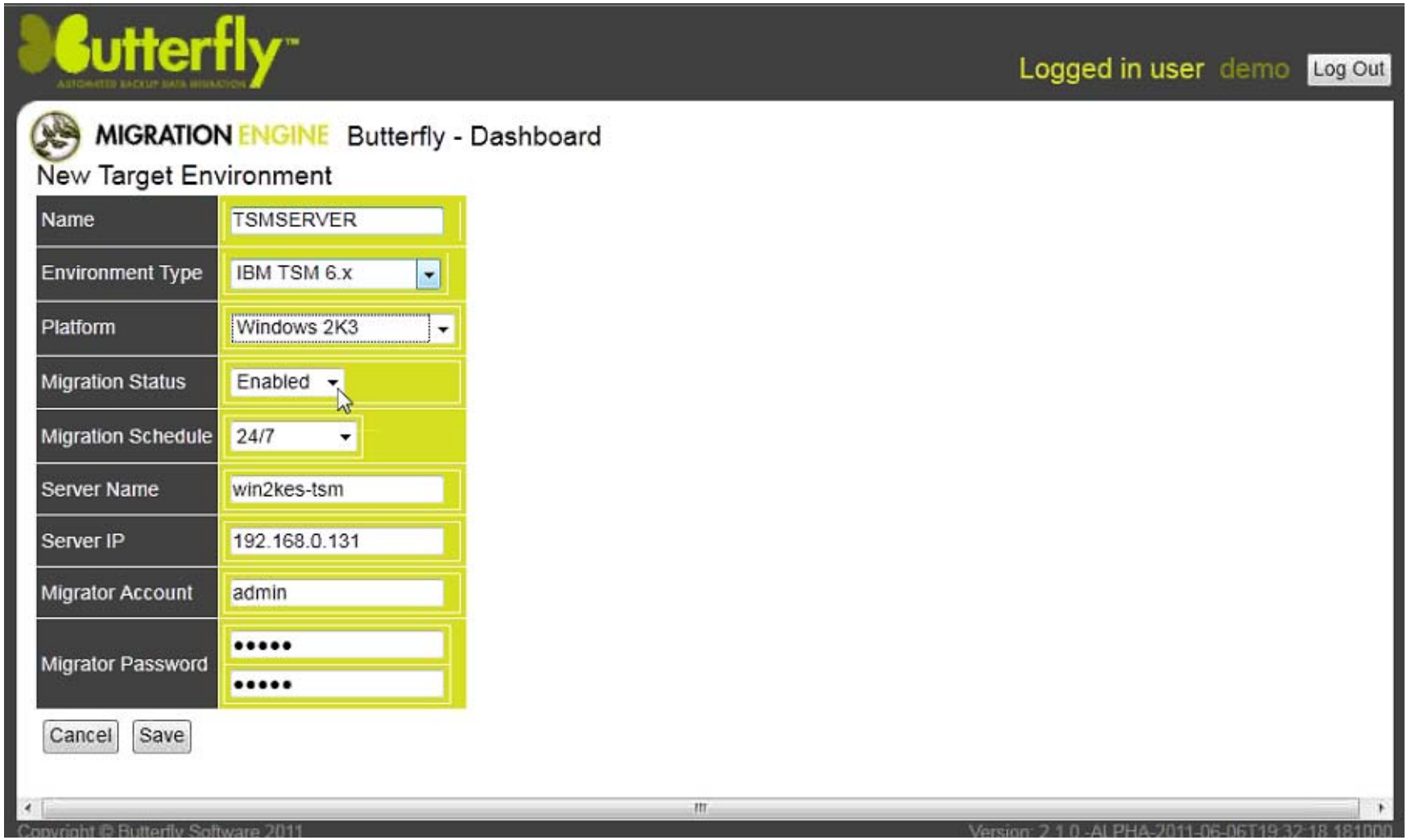
Migration Phases – Migration Scope



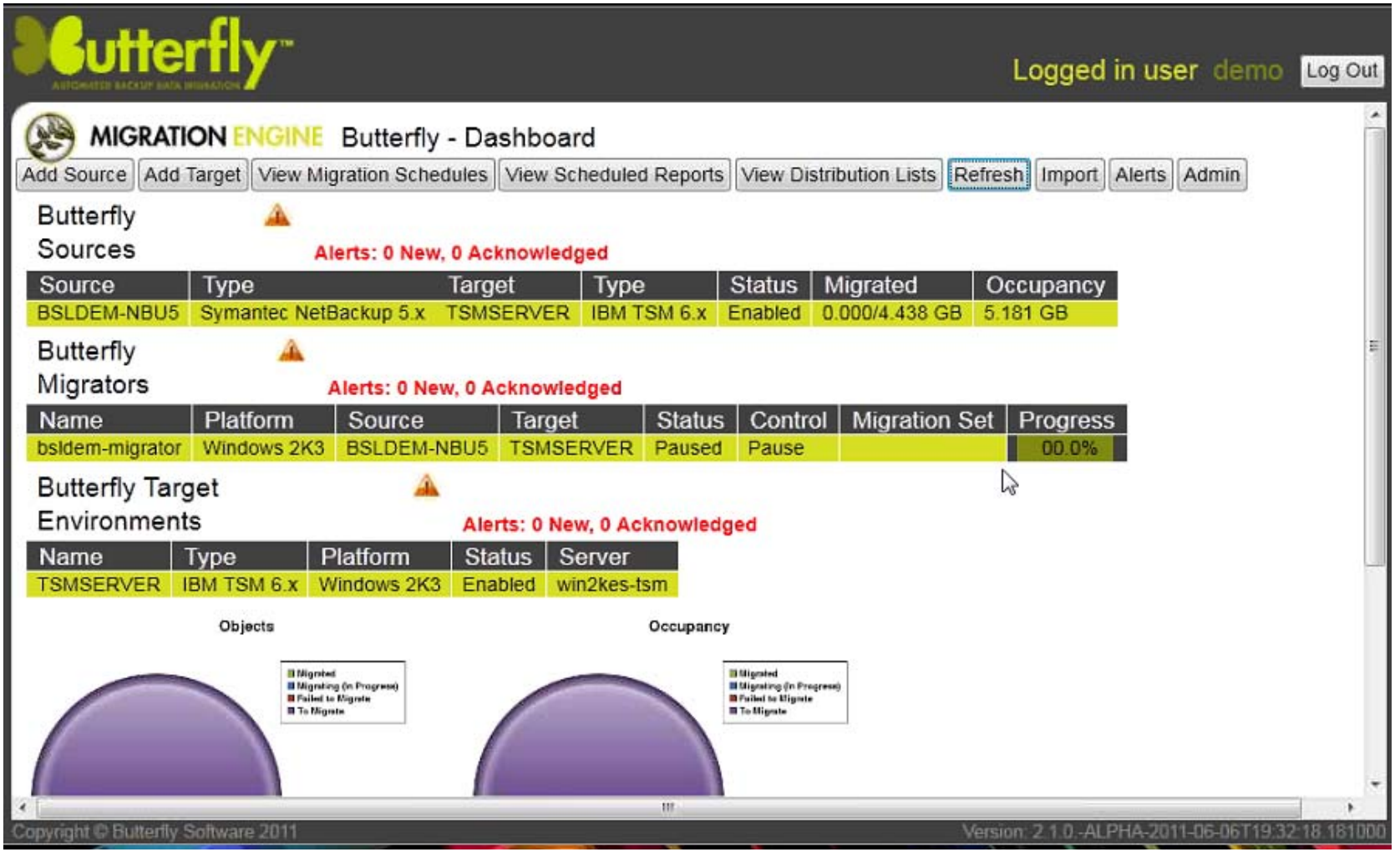
Migration Phases - Organise



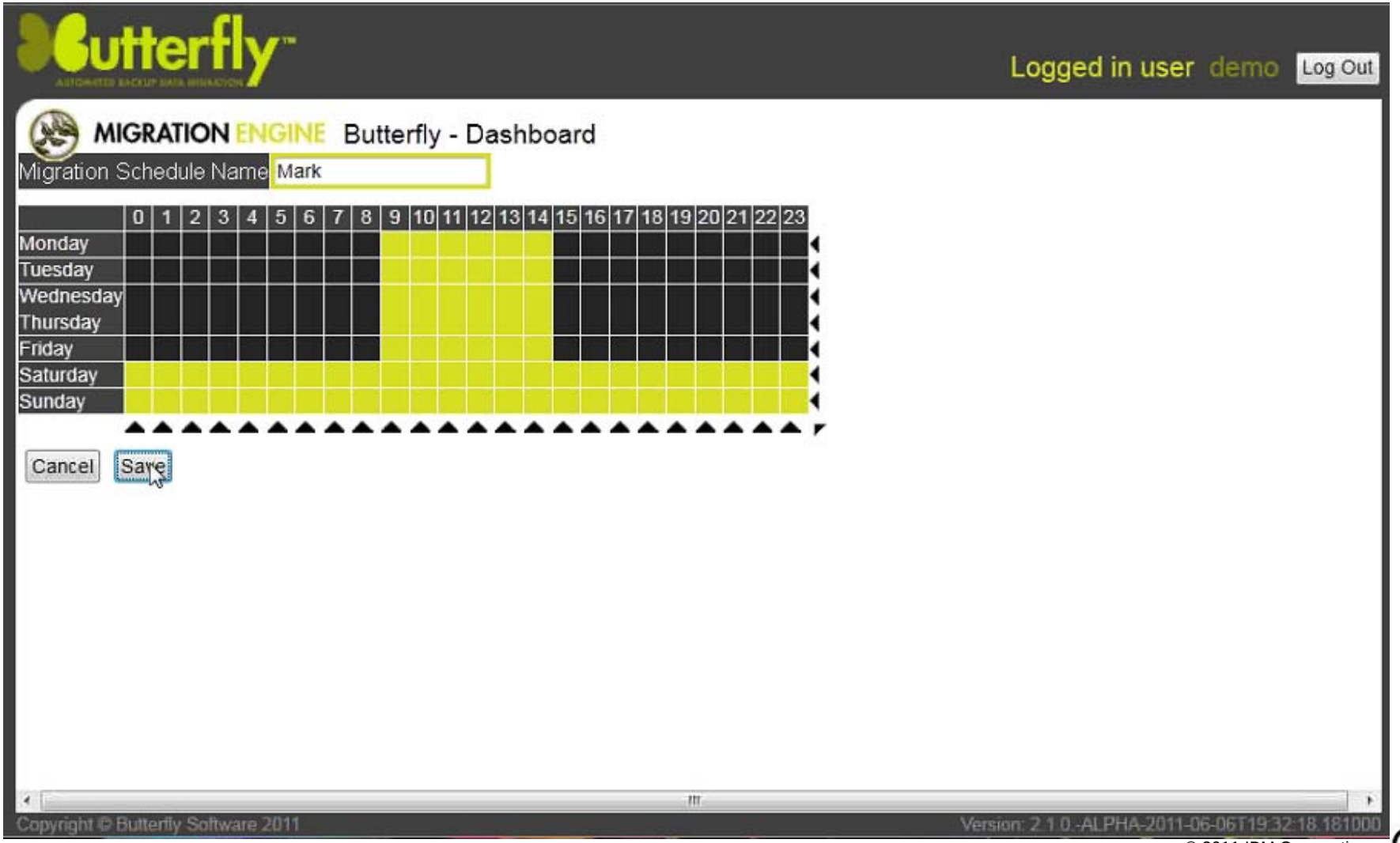
Define Target Environment



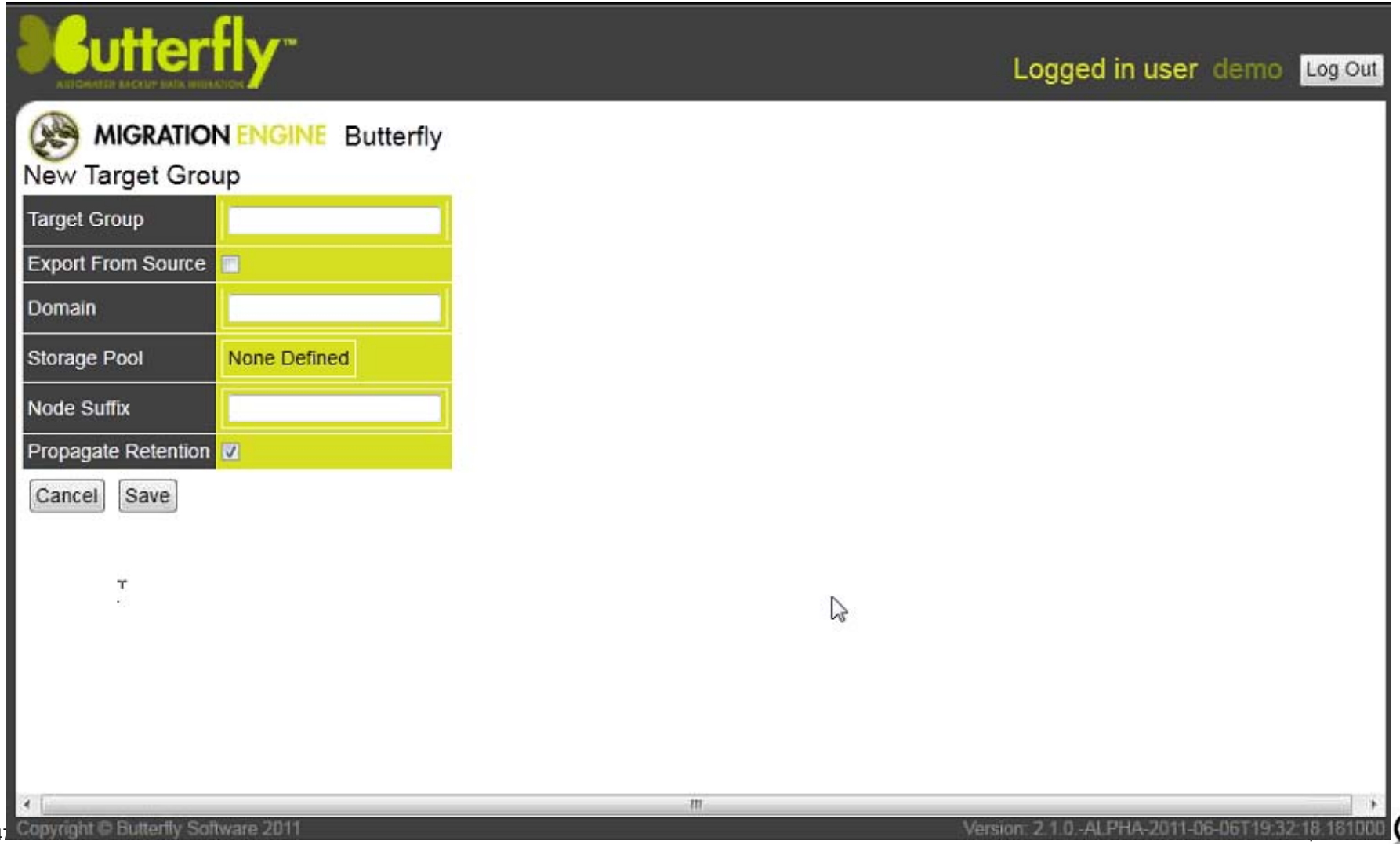
Configuration Complete



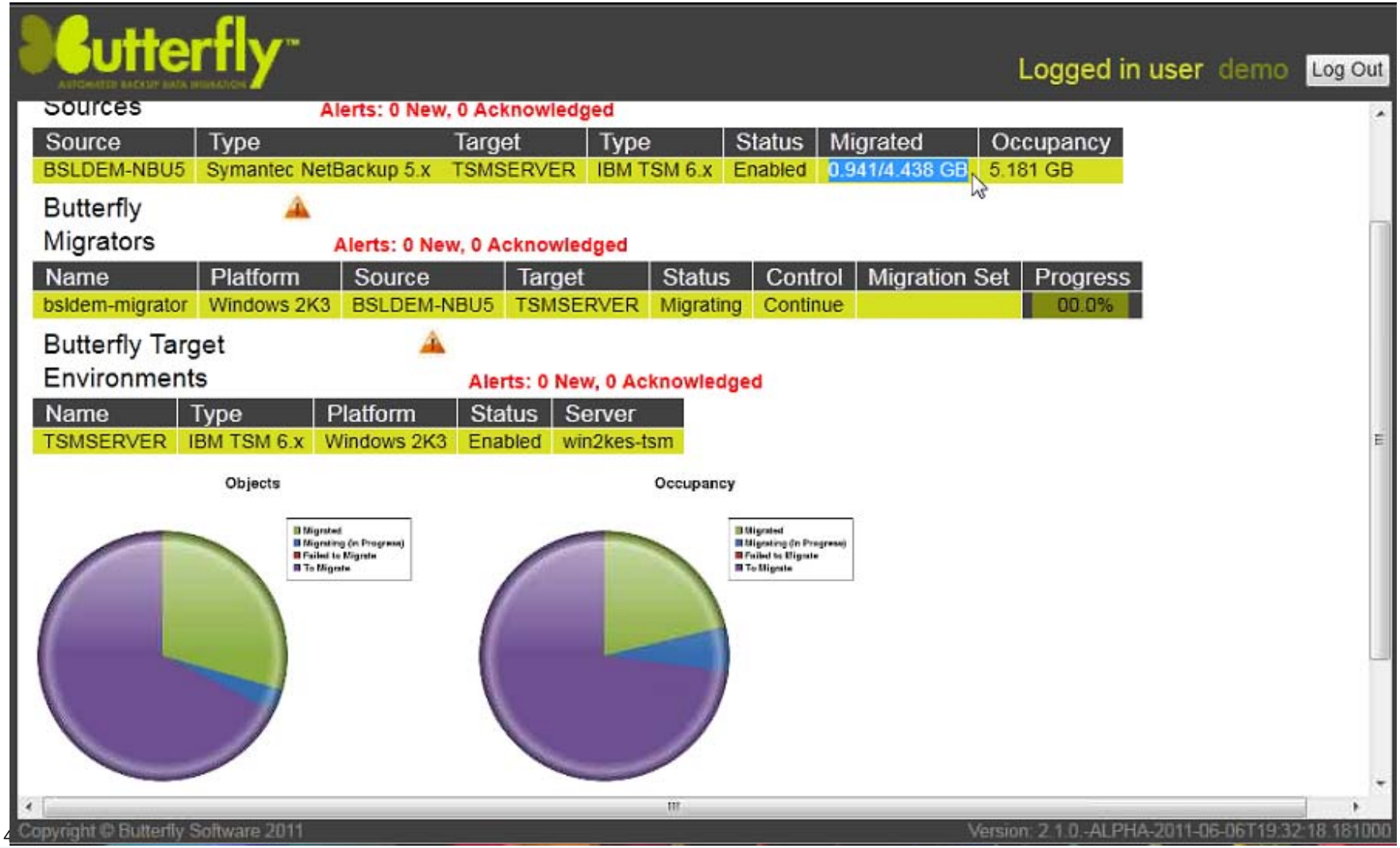
Schedule



Target Groups



Migration Status



Migration Status - Cont

Butterfly™ AUTOMATED BACKUP DATA MIGRATION

Logged in user demo Log Out

MIGRATION ENGINE Butterfly

VS5-PL12-PH1-FILESYSTEM Objects

50 << First < Prev Next > Last >> Page 1 of 1 Export CSV

Object	Size	Backed Up	Expires	Datatype	Level	Status	Migrate	Migrated
bmdev-002_1306602059	270365	28-May-11 18:00	30-Nov-11 17:00	Filesystem	Full	Migrated	1	1
bmdev-002_1305997260	270365	21-May-11 18:01	23-Nov-11 17:01	Filesystem	Full	Migrated	1	1
bmdev-002_1305710342	2044	18-May-11 10:19	20-Nov-11 09:19	Filesystem	Full	Migrated	1	1
bmdev-002_1305535464	30	16-May-11 09:44	18-Nov-11 08:44	Filesystem	Full	Migrated	1	1
bmdev-002_1305392460	270365	14-May-11 18:01	16-Nov-11 17:01	Filesystem	Full	Migrated	1	1
bmdev-002_1305215992	2061	12-May-11 16:59	14-Nov-11 15:59	Filesystem	Full	Migrated	1	1
bmdev-002_1305213701	42358	12-May-11 16:21	14-Nov-11 15:21	Filesystem	Full	Migrated	1	1
bmdev-002_1305203904	123854	12-May-11 13:38	14-Nov-11 12:38	Filesystem	Full	Migrated	1	1
bmdev-002_1305202334	5287	12-May-11 13:12	14-Nov-11 12:12	Filesystem	Full	Migrated	1	1
bmdev-002_1305192734	22	12-May-11 10:32	14-Nov-11 09:32	Filesystem	Full	Migrated	1	1
bmdev-002_1304787657	270365	07-May-11 18:00	09-Nov-11 17:00	Filesystem	Full	Migrated	1	1
bmdev-002_1304692641	270354	06-May-11 15:37	08-Nov-11 14:37	Filesystem	Full	Extracted	1	0
bmdev-002_1304688527	2044	06-May-11 14:28	08-Nov-11 13:28	Filesystem	Full	Pending	1	0
bmdev-002_1304687432	2044	06-May-11 14:10	08-Nov-11 13:10	Filesystem	Full	Pending	1	0
bmdev-004-win2k_1306602062	64	28-May-11 18:01	30-Nov-11 17:01	Filesystem	Full	Pending	1	0
bmdev-004-win2k_1305997263	64	21-May-11 18:01	23-Nov-11 17:01	Filesystem	Full	Pending	1	0
bmdev-004-win2k_1305392463	64	14-May-11 18:01	16-Nov-11 17:01	Filesystem	Full	Pending	1	0
bmdev-004-win2k_1304787660	64	07-May-11 18:01	09-Nov-11 17:01	Filesystem	Full	Pending	1	0
bmdev-004-win2k_1304689195	13873	06-May-11 14:39	08-Nov-11 13:39	Filesystem	Full	Pending	1	0

Back Refresh

Copyright © Butterfly Software 2011 Version: 2.1.0-ALPHA-2011-06-06T19:32:18.181000

Interoperability Matrix

			Backup Management Server																		
Operating System	OS Version	Hardware	Symantec NetBackup			CA ArcServe			EMC Legato Networker			HP Data Protector		Symantec Backup Exec			NetVault BakBone		CommVault Simpana	
			6.5	6.x	5.x	12.5	12.0	11.0	7.5	7.4	7.3	6.1	6.0	13	12.5	11d	8.5	8.0	8	7
AIX	5L 5.2 (32/64 bit)	Power4, Power5	X	X																
	5L 5.3 (32/64 bit)	Power4, Power5	X	X																
	6L 6.1 (32/62 bit)	Power5, Power6																		
Solaris SPARC	8	SUN SPARC sun4u	X	X	X		X	X		X	X									
	9	SUN SPARC sun4u	X	X	X		X	X	X	X	X									
	10	SUN SPARC sun4u	X	X			X		X	X	X									
Solaris x86/x64	8	32 bit (x86)	X	X	X		X	X												
	9	32 bit (x86), 64 bit (x64)	X	X	X		X	X												
	10	32 bit (x86), 64 bit (x64)	X	X			X		X	X										
Windows	Server 2003 R2 (all SP)*	32-bit (x86)	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X
	Server 2003 R2 (all SP)*	64-bit (x64)	X	X					X	X		X	X	X	X	X	X	X	X	X
	Server 2008**	64-bit (x64)	X												X		X		X	
	Server 2008**	32-bit (x86)	X			X						X	X		X		X		X	
	2000 SP4	32-bit (x86)	X	X	X	X	X	X							X		X		X	

Interoperability Matrix - cont

			Backup Management Server																	
Operating System	OS Version	Hardware	Symantec NetBackup			CA ArcServe			EMC Legato Networker			HP Data Protector		Symantec Backup Exec			NetVault BakBone		CommVault Simpana	
			6.5	6.x	5.x	12.5	12.0	11.0	7.5	7.4	7.3	6.1	6.0	13	12.5	11d	8.5	8	8	7
Red Hat Enterprise LINUX	5.x Intel x86	32 bit (x86)	X	X	X				X	X	X									
		64 bit (x64)	X	X	X				X	X										
	5.x x64	64 bit (x64)	X	X	X				X	X										

* **Enterprise and Datacenter Editions**
** **Standard, Enterprise and Datacenter Editions**

Trademarks and disclaimers

The customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list pricesand performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Prices are suggested U.S. list prices and are subject to change without notice. Starting price may not include a hard drive, operating system or other features. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Photographs shown may be engineering prototypes. Changes may be incorporated in production models.