NetWorker Usage Survey, 2013

Preston de Guise

NetWorker Information Hub

16 February 2014

NETWORKER INFORMATION HUB

Table of Contents

NETWORKER INFORMATION HUB

INTRODUCTION

Objective

The purpose of the NetWorker Usage Survey is to independently gauge key high level details of how EMC NetWorker is used within the community, and to report trends on that usage.

Survey Period

This survey was conducted between 1 December 2013 and 31 January 2014, and follows previous surveys conducted in 2012, 2011 and 2010 (June, November).

About the Author

Preston de Guise is a long-term data protection expert with a career focus on enterprise backup and recovery solutions. Preston is author of "Enterprise Systems Backup and Recovery: A Corporate Insurance Policy" (CRC Press, 2008, 978-1420076394) and of the upcoming title "Data Recovery: Preventing Data Loss in the Age of Big Data, Cloud, and Virtualization".

Preston has worked on and developed backup solutions in most industry verticals, covering the full range of businesses from SOHO through to Global Fortune 500 companies.

Preston is based on Melbourne, Australia.

NETWORKER SERVER VERSION

Responses

This question determined the version(s) of NetWorker server software being run by organisations. Note that it allowed for multiple datazones.
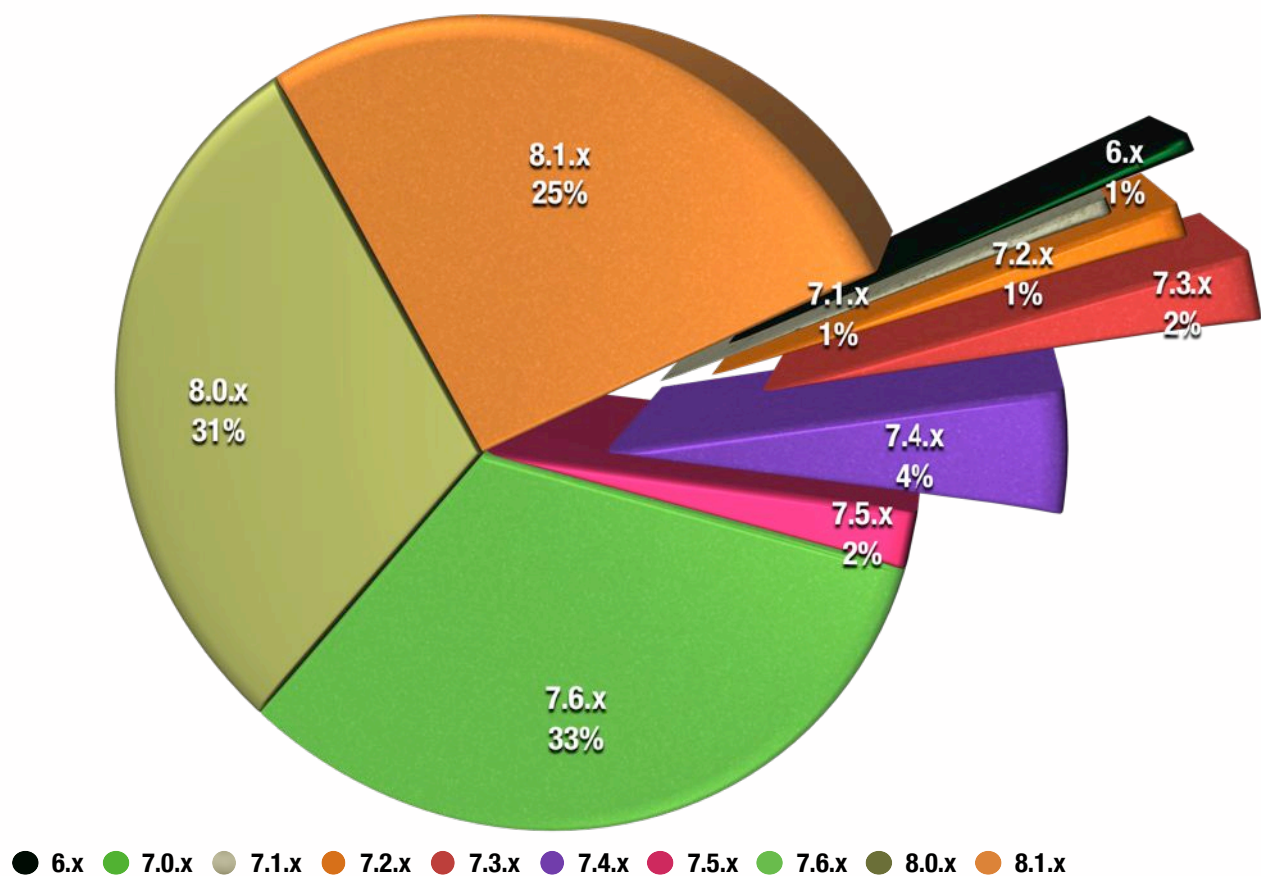
Findings

With significant feature enhancements first in the 7.6.x series, and more recently in the 8.x series, EMC has considerably moved its customer base along in NetWorker versions. Traditionally, NetWorker users have been reluctant to upgrade to the latest versions of NetWorker, but this tendency has been significantly countered, with the vast majority of users not only on supported versions of NetWorker, but *recent* versions.

With over 50% of users currently running NetWorker 8 or higher, it's clear EMC has succeeded in refocusing development on new and useful features without an impact to usability.

NUMBER OF DATAZONES

Responses

This question determined the number of actual datazones being used by EMC customers – i.e., the number of actual NetWorker *servers* within the business.
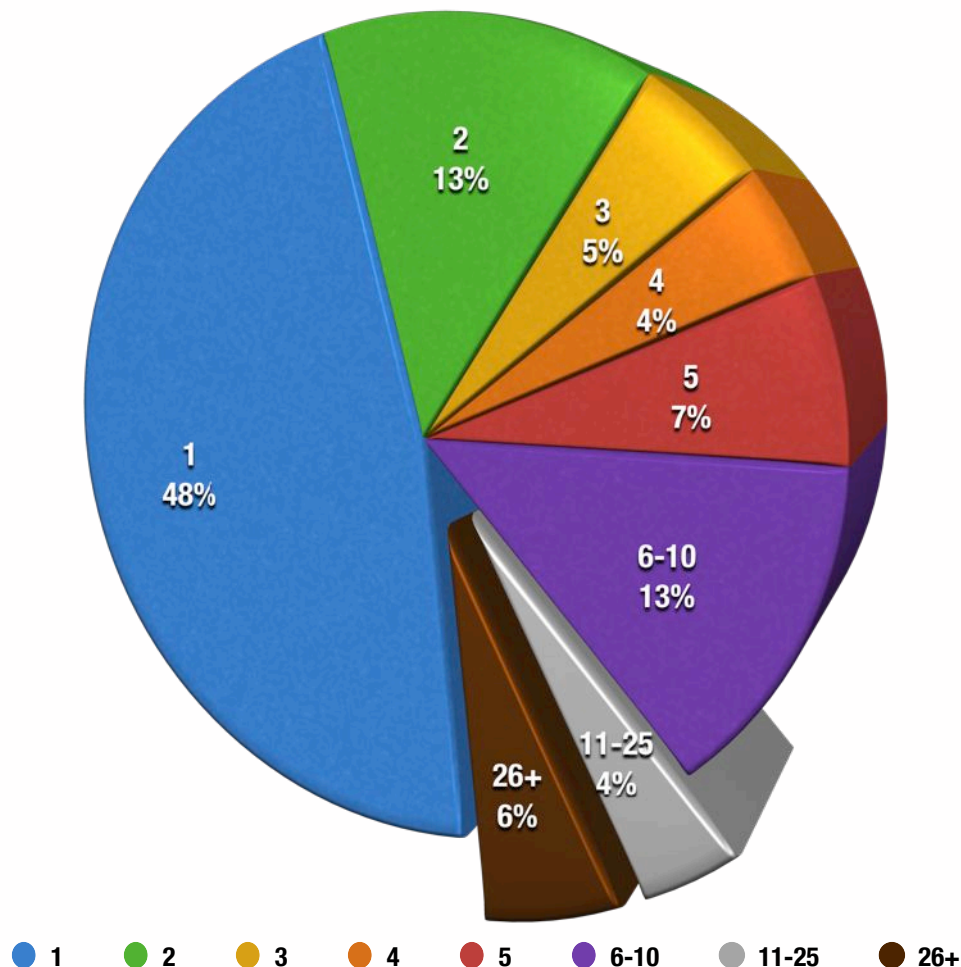
Findings

Similar to previous years, a large number of businesses only have a single NetWorker server within their environment – yet tellingly, that's not the actual *majority* of businesses.

It remains clear that the number of businesses running more than one datazone indicates a continued need for centralised control, reporting and monitoring. The default way to achieve this with NetWorker is NMC – NetWorker Management Console. Bundled with NetWorker and installable on practically any host within the environment, NMC offers a central reporting capability, as well as control, configuration and monitoring across all datazones it has visibility of.

While NMC comes free with NetWorker, EMC also sells Data Protection Advisor (DPA), which is an extended reporting suite capable of reporting against multiple backup products, storage systems and applications within an environment. DPA is also notoriously expensive and has a steep learning curve, something that puts it beyond the interest of many NetWorker customers.

More recently, EMC have developed and made (in limited availability) a new 'dashboard' style server, Backup and Recovery Manager (BRM), which provides very high level dashboard information on Data Domain, Avamar and NetWorker.

Reporting and monitoring of multiple NetWorker servers is an area where EMC is still sending a very confused message to the market place.

DPA is unquestionably the Swiss Army Knife of data protection reporting products, but that isn't necessarily a good thing. Quoting Steve Flynn, Director of Product Marketing for BRS at EMC in May 7 2013:

> *Let's face it – if you were out in the wild and had to cut wood for a fire, you'd want the proper tools for the job. Would you rather have a Swiss Army knife, with its mini-saw and variety of miniature utensils, none of which are particularly efficient – or a full-fledged survival kit with well-made, dependable implements?*
>
> *EMC Data Protection Survival Kit*
>
> *https://pulseblog.emc.com/2013/05/07/emc-data-protection-survival-kit/*

DPA is meant to be part of that "full-fledged survival kit", but it's like throwing a flat-pack IKEA rowboat at a drowning man who just needs a floatation device.

EMC clearly recognised those issues with the development of BRM, but that's arguably a tool that shouldn't need to exist. *Either* NMC needs to have 50% more capability *or* DPA needs a simpler mode of install and operation. Bridging products aren't always the solution.
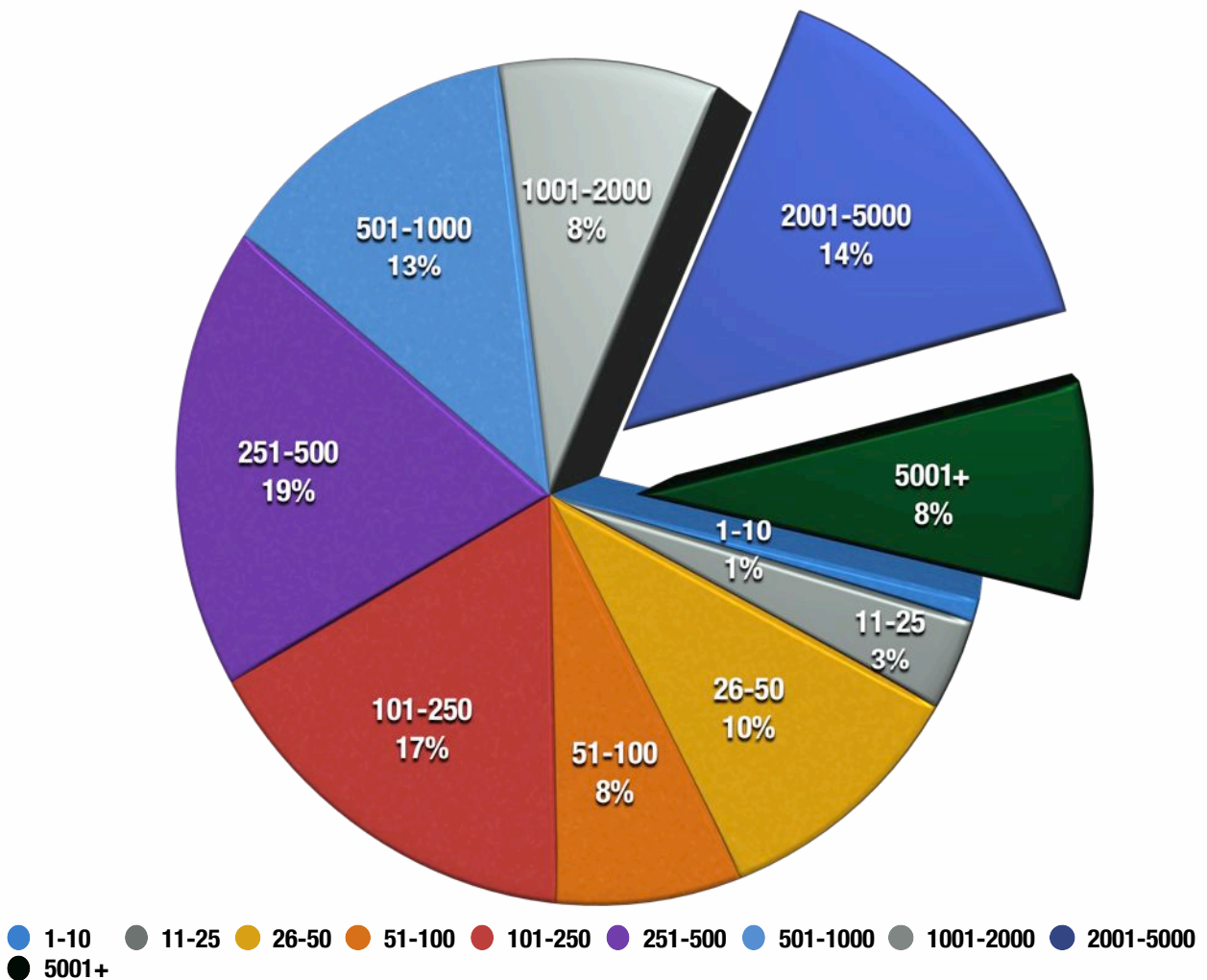
EMC needs to take a step back and evaluate what it's doing with reporting products for its data protection suite of products – at the moment it's unnecessarily messy.

TOTAL CLIENTS – ALL DATAZONES

Responses

This question gauged the total number of clients that were being protected by NetWorker servers across *all* datazones within a business.

Legend:
● 1-10   ● 11-25   ● 26-50   ● 51-100   ● 101-250   ● 251-500   ● 501-1000   ● 1001-2000   ● 2001-5000   ● 5001+

Findings

Under a quarter of all businesses surveyed use NetWorker for environments with 100 clients or less. Significantly, 30% of businesses use NetWorker for protecting more than 1,000 clients.

Evaluating the higher client counts of 2000+, we see a significant percentage of those businesses running quite a few datazones:
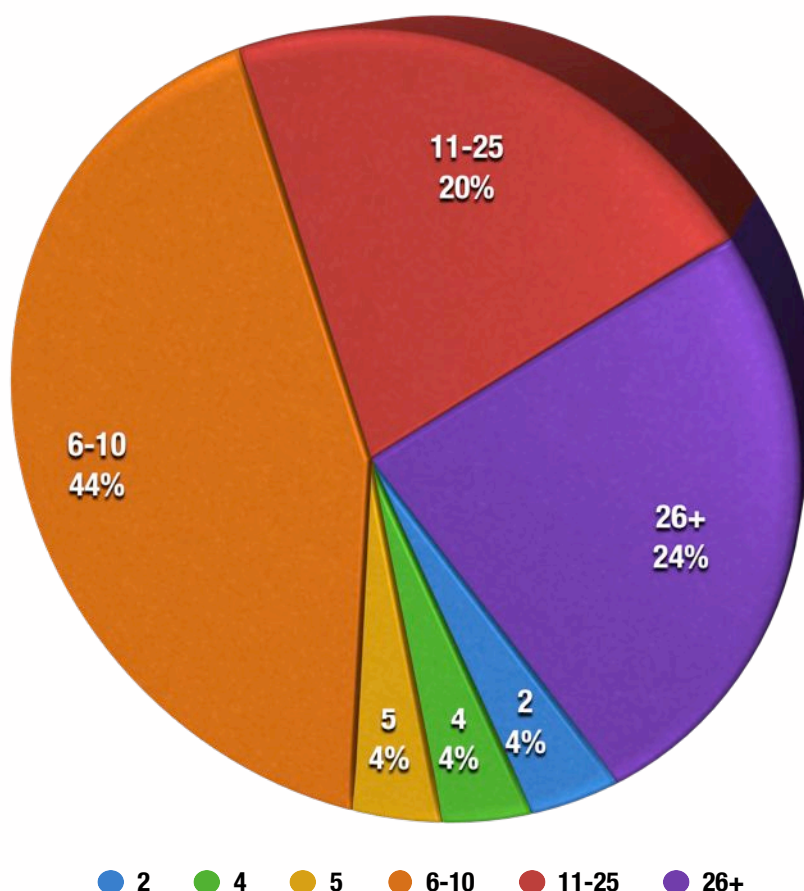
While insufficient data was gathered to accurately state reasons behind the number of datazones, historically a few reasons can typically be noted, viz.:

• Geography

• Scaleability

• Security

Larger, multi-national businesses, and businesses with relatively low-speed links between disparate datacentres will usually run multiple datazones – centralised control otherwise often becomes problematic.

These days NetWorker scales excellently, capable of handling high numbers of clients, but older versions did have various scaling issues that made splitting large numbers of clients into more datazones to be a sensible approach. Further, the number of physical *devices* being backed up to (especially in tape environments), could substantially impact the responsiveness of a backup server under load as the number of concurrent streams increased. While this has been substantially improved, and tape continues to exit the day-to-day backup space, EMC provides no wizard or easy option for merging datazones, meaning the number of datazones does not tend to shrink often.

Up until NetWorker v8, there was no support for multi-tenancy, and it's fair to say that what is currently defined in NetWorker is more of the format of *secured sub-datazones* than true multi-tenancy. In particular, there is still

far too much information leakage about client names, etc., between the sub-datazones for NetWorker to claim true multi-tenancy, something EMC will need to work on.

Evaluating the number of datazones where there are between 501 and 2000 clients shows a significant number of single datazone environments, but still a reasonable number of environments with multiple datazones:
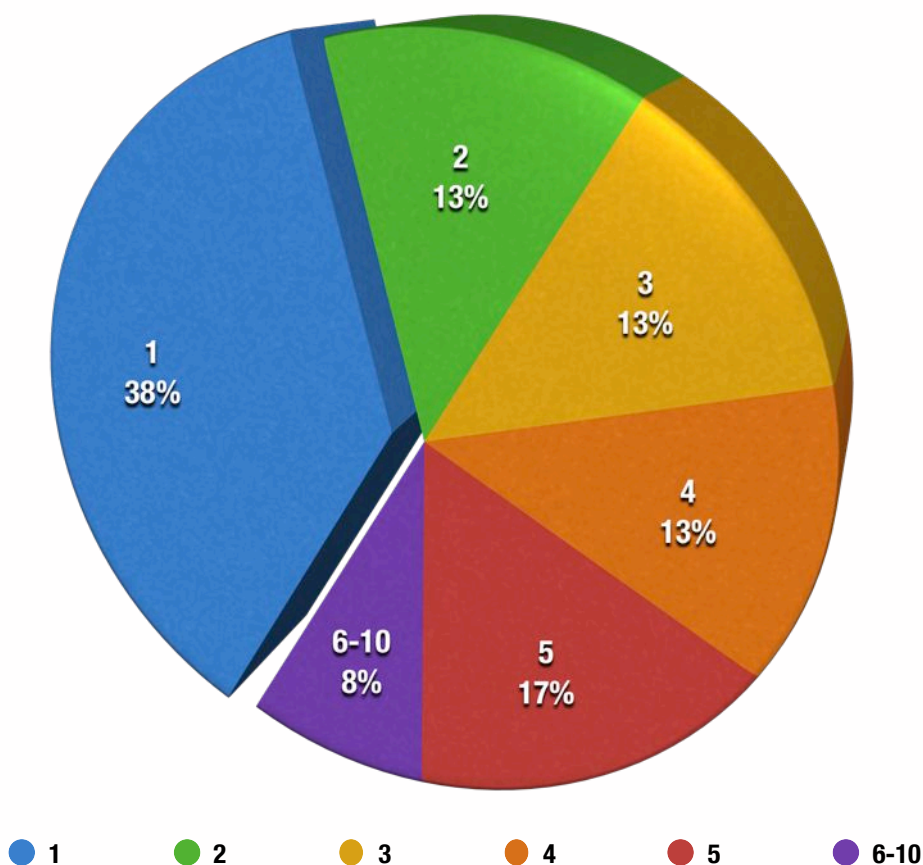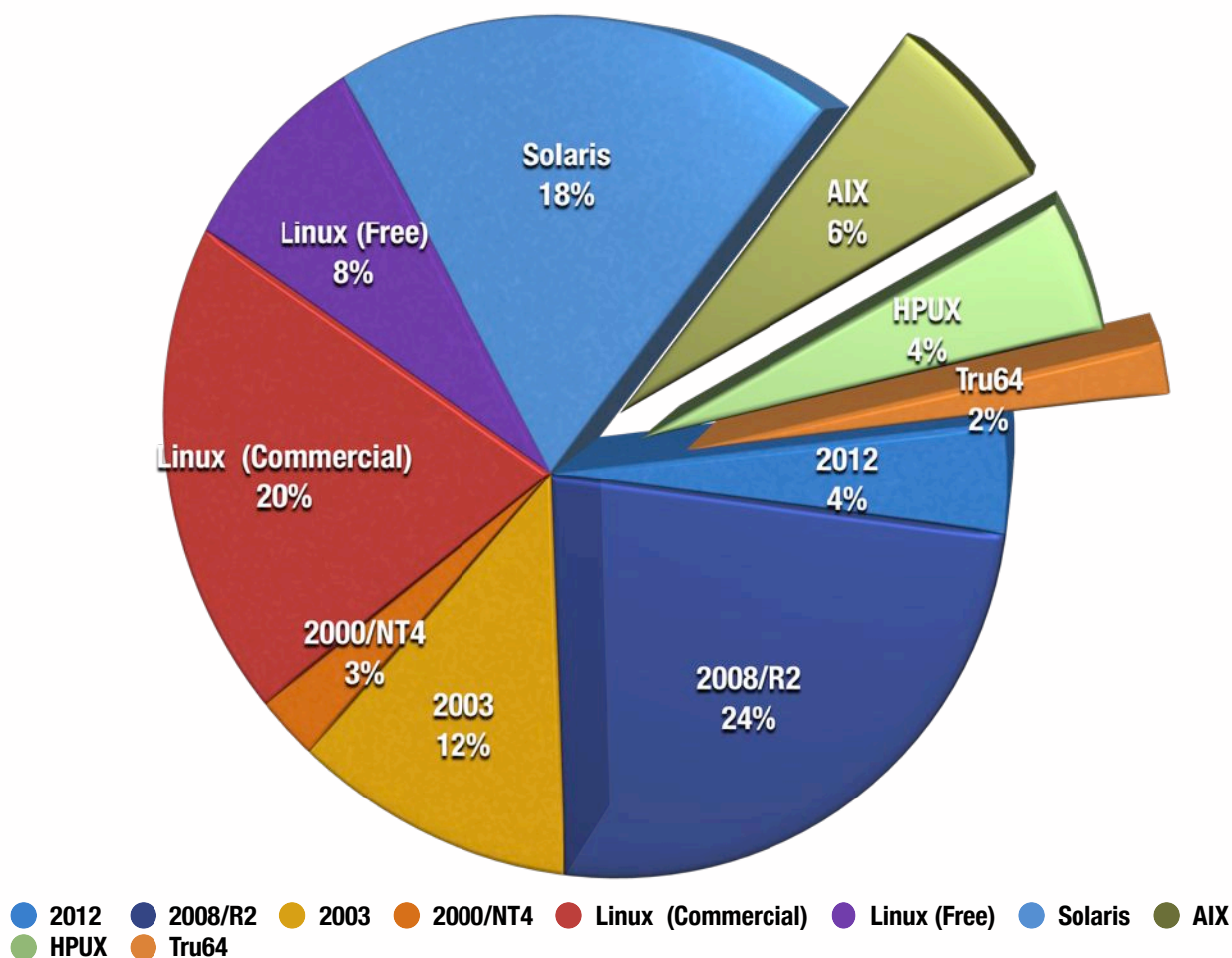

The prevalence of multi-datazone businesses, and the number of clients in those datazones, further goes to show EMC needs to come up with a cohesive and cost-justifiable reporting story for data protection.

NETWORKER SERVER OPERATING SYSTEM

Responses

This question determined the operating systems currently being used to run a NetWorker datazone.

Findings

Clearly the dominant platform for use as a NetWorker server is Windows. When the above data is simplified, it becomes even more evident – the Windows platform as a whole accounts for 42% of the NetWorker server share, Linux for 28% and Solaris for just 18%. All other potential server operating systems occupied the remaining 12%.
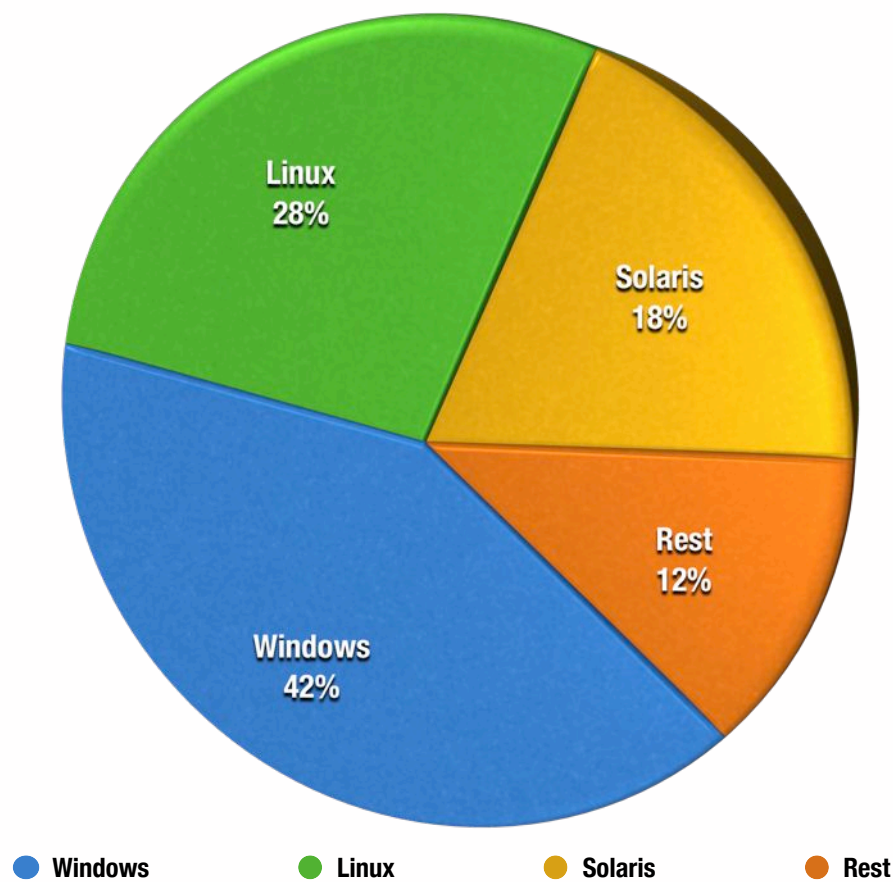
Under the awkward stewardship of Oracle, Solaris continues to fall as a NetWorker server platform of choice.

Survey	Windows	Linux	Solaris
March 2010	29%	22%	43%
November 2010	29%	19%	43%
June 2011	38%	23%	33%
December 2012	41%	30%	20%
December 2013	42%	28%	18%

Once the dominant server platform for NetWorker, Solaris is now a shadow of its former self, and likely to continue to shrink. While Linux initially gained share against the fall of Solaris, it is possible this has now plateaued. Though too early to tell, anecdotal evidence suggests that businesses are just as willing to switch from Solaris to Windows as they are from Solaris to Linux.

With such a dwindling number of systems running AIX and HPUX as their server platform, it seems that 'conventional' Unix, in the NetWorker space at least, has lost the race.

CLIENT/STORAGE NODE OPERATING SYSTEMS

Responses

While not all operating systems can run as a NetWorker storage node, the two types were combined to avoid confusion between servers and storage nodes.
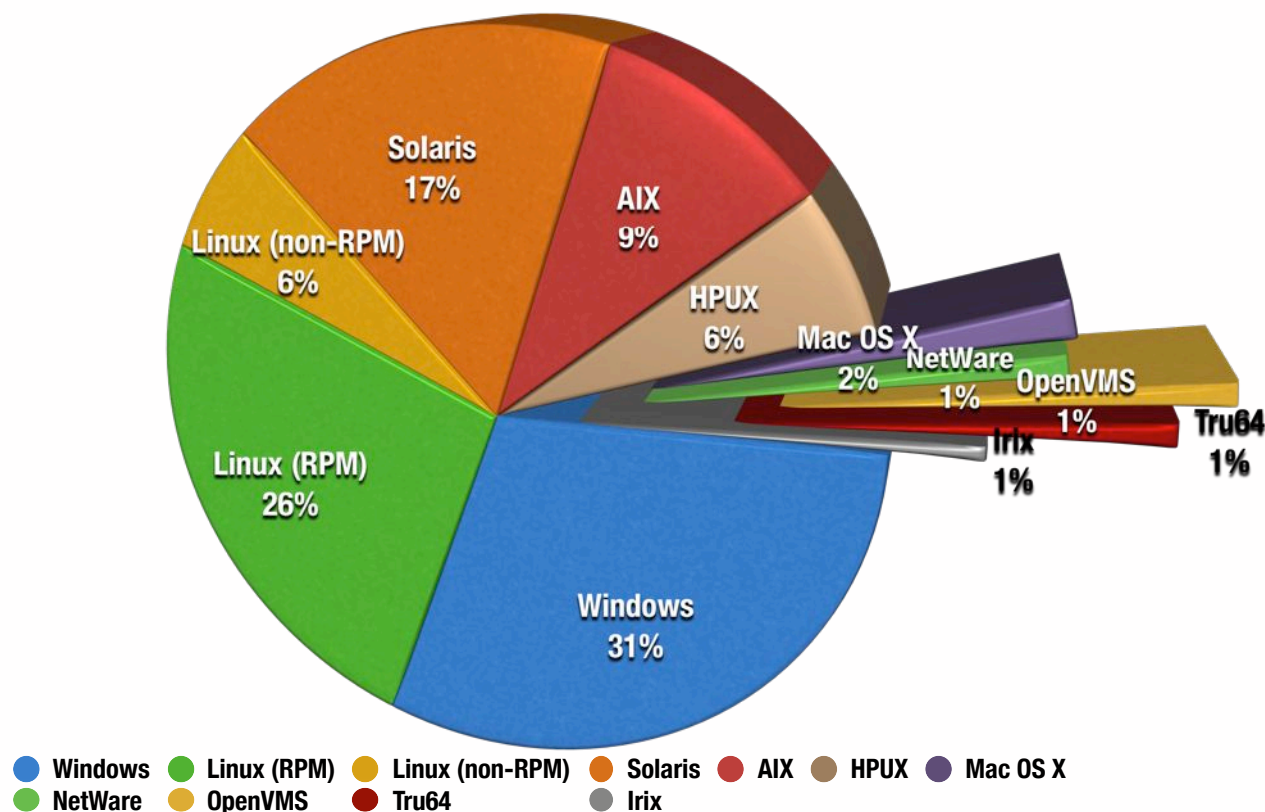
Findings

Although Linux and Windows represents the lion's share of the client distribution for NetWorker, other operating systems collectively still represent just under 40% of the client OS base. The substantial consolidation of preferred platforms for the NetWorker server platform may point to eventual development savings, but continued support for a wide variety of client operating systems will work in EMC's favour for some time to come.

Survey	Windows	Linux	Solaris	Other
March 2010	28%	23%	21%	28%
November 2010	28%	23%	22%	27%
June 2011	27%	22%	20%	31%

Survey	Windows	Linux	Solaris	Other
December 2012	29%	29.5%	17.5%	24%
December 2013	31%	32%	17%	20%

Measuring overall server market share is somewhat challenging – the combined figures are not readily obtainable and many results focus only on those servers which are internet visible, which becomes weighted towards particular operating systems.

That being said, the breakup of operating systems for NetWorker clients is skewed a little further towards non-Windows operating systems than traditional estimates on Windows server usage, and may be because of higher competition for Windows backup software amongst workgroup products, and the historical deployment pattern of NetWorker.

BUSINESSES USING DEDUPLICATION

Responses

This question determined how many respondents are using deduplication technology, and broadly, where that technology was deployed.
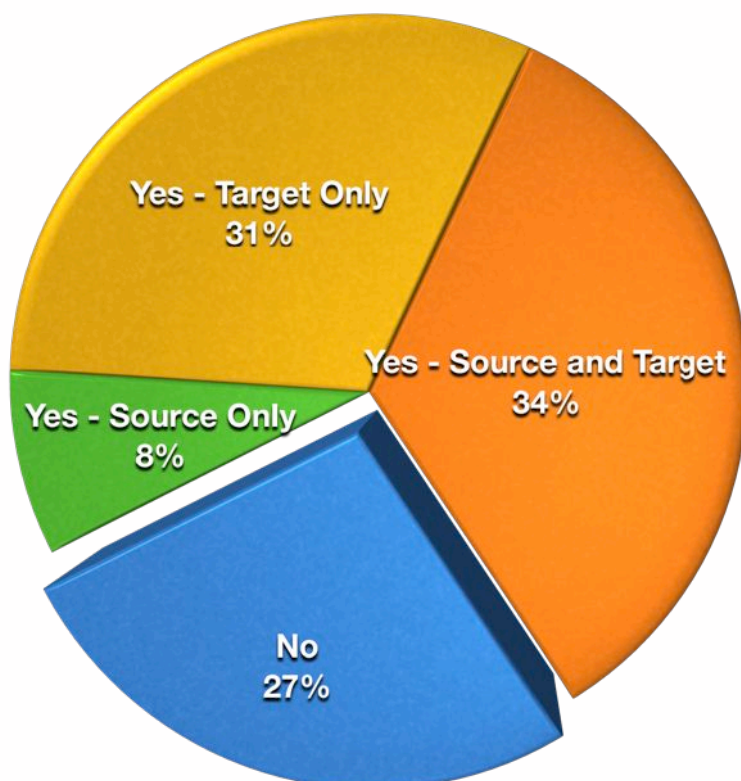
● **No**   ● **Yes - Source Only**   ● **Yes - Target Only**   ● **Yes - Source and Target**

Findings

Since only warranting a whisper 5 years ago, deduplication technology has rapidly evolved into a highly utilised feature in backup environments. Barely more than a quarter of respondents were making no use of deduplication technology whatsoever – 73% of respondents have deduplication within their environments.

Survey	No Deduplication	Source Only	Target Only	Source and Target
November 2010	68%	4%	20%	7%
June 2011	64%	5%	27%	4%
December 2012	37%	5%	31%	27%
December 2013	27%	8%	31%	34%

In a relatively short number of surveys, deduplication has gone from 68% **not** used to 73% **used**, which indicates a large number of businesses can see (and are getting) the benefits out of using deduplication technology. Indeed, the remaining relatively low percentage of those *not* using deduplication would correlate reasonably well to businesses that are in the strongly conservative part of the technology adoption curve and those who have relatively stringent and long budgetary cycles.

The growth of combined source and target deduplication within NetWorker is in no small part due to the successful integration of Data Domain Boost into NetWorker's processes. With Boost initially integrated to allow for storage node offloading of deduplication, and more recently client based offloading of deduplication, Data Domain has largely left Avamar adoption within NetWorker environments as an also-ran.

When we go onto consider *only* those businesses using deduplication, we see a telling break-down of integration methodologies:
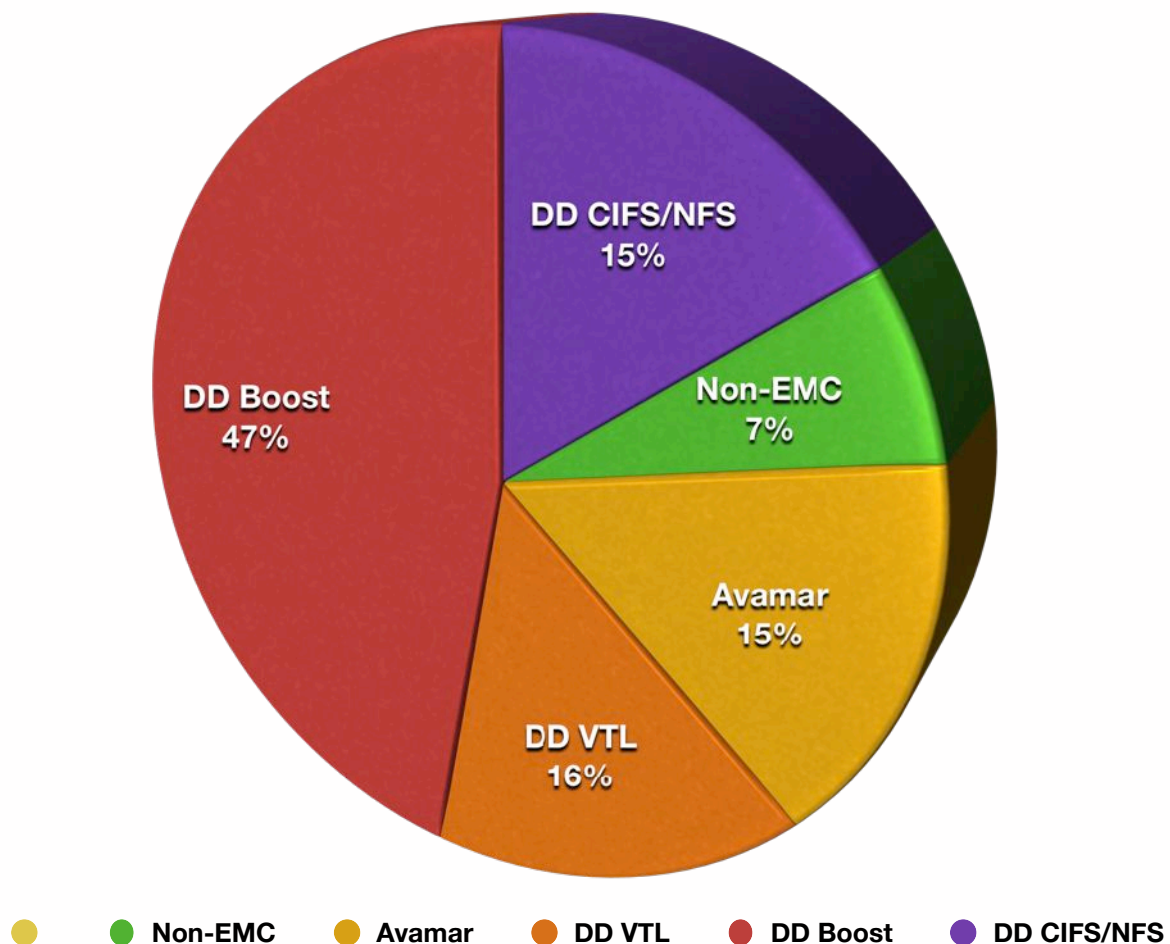
Non-EMC deduplication technology accounts for less than 10% of the deduplication integration within NetWorker. This would typically account for backup to a deduplication VTL or file-share from vendors such as HP and Quantum. Data Domain, however, is the elephant in the room.

Total Data Domain usage accounts for 78% of deduplication options within NetWorker based data protection environments, while Avamar accounts for only 15%.

For a while after their acquisition, EMC pushed Avamar hard as the ultimate backup product. Between its quirky interface, repugnant XML information dialogues, activity black-out window and high cost, it did not prove overly popular with NetWorker sites as an adjunct to provide deduplication. This was not the case with Data Domain.
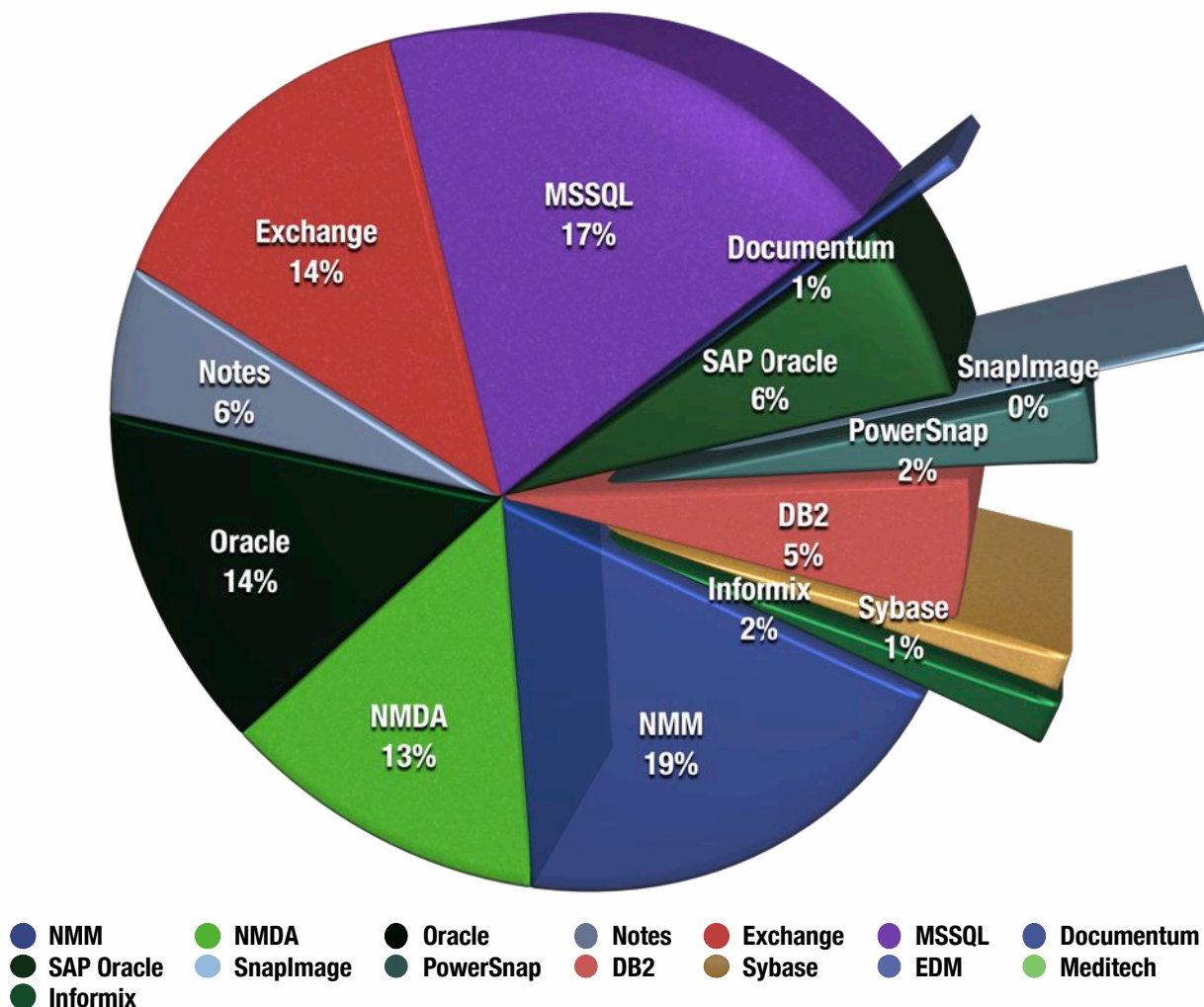
The flexibility of Data Domain ultimately led to its overwhelming success in the NetWorker environment. Pre-Boost, the VTL functionality was as good as other VTLs on the market place but provided considerable capacity savings with deduplication. The global deduplication storage pool has always been an attractive selling point for businesses where database administrators refuse to integrate their backup processes with the rest of the company, hence the continued use of CIFS/NFS shares for backup functionality to this day.

The ongoing development of Boost functionality has been of substantial benefit, with cloning controlled Boost replication and more recently Virtual Synthetic Fulls proving quite popular. By integrating with NetWorker's process flows rather than remaining an adjunct to them, Boost was practically guaranteed to dominate from the moment it appeared.

NETWORKER MODULE USAGE

Responses

This section covers the types and breakdown of database and application modules used within NetWorker.

● NMM	● NMDA	● Oracle	● Notes	● Exchange	● MSSQL	● Documentum
● SAP Oracle	● SnapImage	● PowerSnap	● DB2	● Sybase	● EDM	● Meditech
● Informix						

Findings

Modules for Microsoft application backups remain the largest player with NetWorker. Between NMM and the legacy Exchange/MSSQL modules, Microsoft application backups comprise 50% of the module usage. Conversely, traditional 'Unix'-centric databases, via Oracle, SAP on Oracle, DB2, NMDA, Sybase and Informix account for 41% of the module market, with the remaining products and features competing for the last 9%.

Despite EMC's continued investment in the module, Meditech barely rates a mention (in this survey there were zero respondents who were using it), so presumably EMC have at least one or two customers using Meditech who have significant investment in the EMC product suite. (Rumour has it that it is precisely this sort of high-

value customer that led to the development of a Mac recovery GUI for NetWorker, even when EMC was focusing recovery GUI development into NMC.)

It seems illogical and counter-intuitive that EMC have dumped support for Documentum, one of their own products. While Documentum is a complex system to deploy, each Documentum deployment represents substantial value to EMC compared to say, a single Microsoft SQL Server install. The high complexity in achieving stable, recoverable backups of Documentum undoubtedly make it challenging to develop, but pushing the onus for this backup out to customers and third party companies seems to be an odd direction. Currently EMC customers using Documentum need to turn to third party backup tools, such as HOTbackup from CYA Solutions Group. While CYA's HOTbackup solution is viable, and indeed recommended by EMC, that it is required remains puzzling.

With the recent update to the NetWorker Module for Databases and Applications (NMDA), EMC introduced support for open source databases via MySQL – of a kind. NMDA introduced support for MySQL Enterprise Edition, which is the Oracle produced version of MySQL, requiring costly licenses and reasonably stringent operating system requirements, all of which are out of alignment with most MySQL usage.

● **None** ● **PostgreSQL** ● **MySQL (Free)** ● **MySQL (Oracle Enterprise Edition)** ● **Firebird** ● **Other**
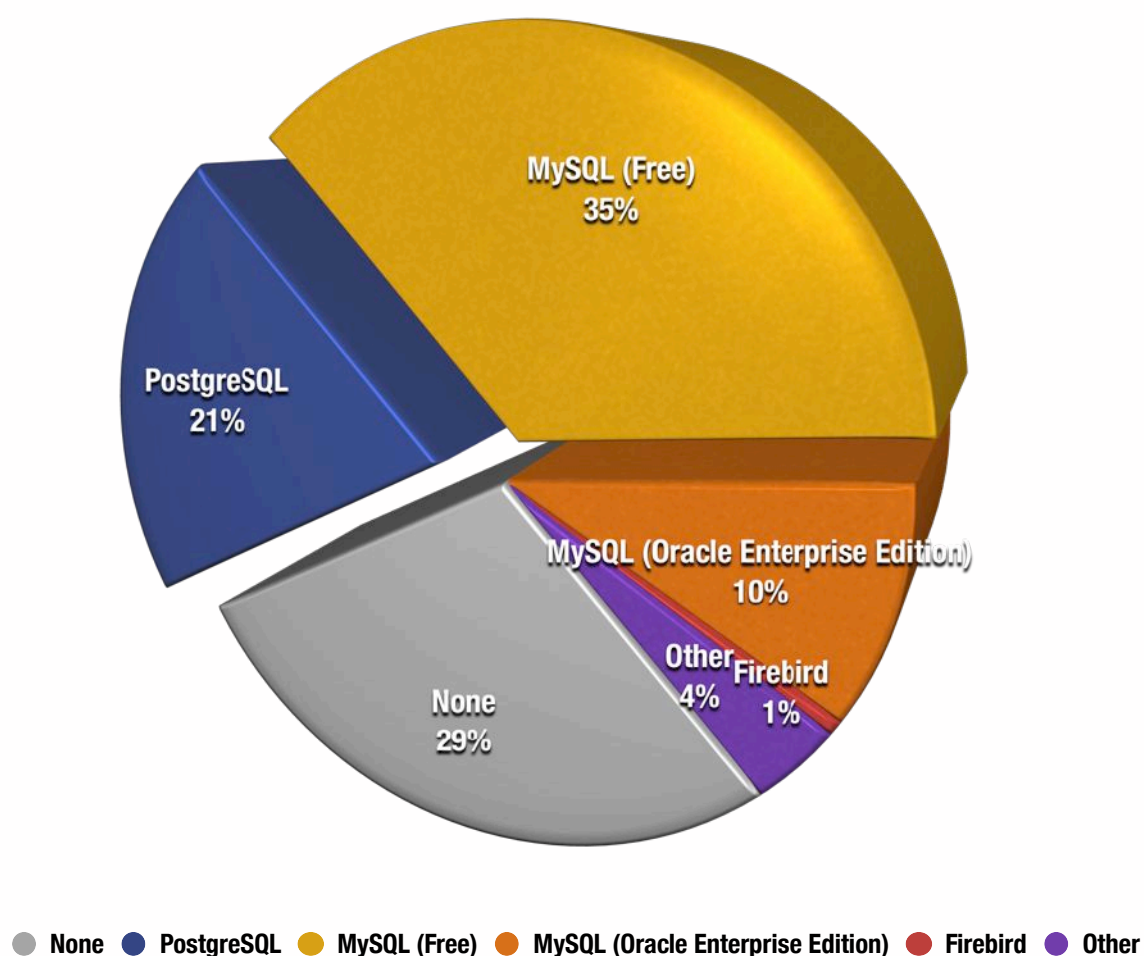
While 29% of respondents don't use Open Source Databases (OSDs) within the environment, amongst those who do, the MySQL Enterprise Edition remains relatively unpopular compared to free editions of MySQL and PostgreSQL.

● PostgreSQL   ● MySQL (Free)   ● MySQL (Oracle Enterprise Edition)   ● Firebird   ● Other
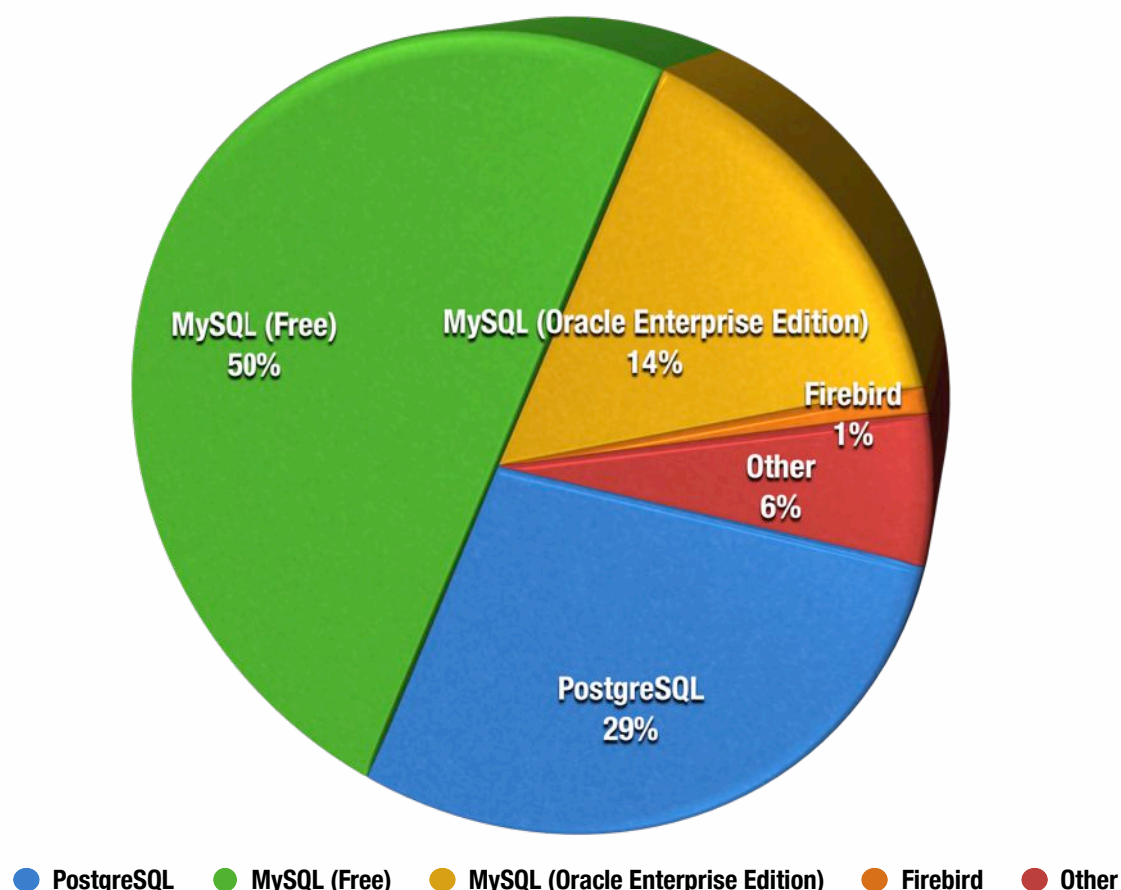
EMC's decision to support only the Oracle MySQL Enterprise Edition is understandable only from the perspective of limiting the development cost, since Oracle include a backup and recovery interface for Enterprise Edition which has similarities with RMAN.

Yet, that development cost seems practically the *only* reason for the release of the module, and if the aim of it was to gauge desire for a more generic open source database module, EMC may have created a situation where the only possible customer response will be 'not interested', despite the obvious interest amongst NetWorker users.

When we move on to consider respondents who are just using some form of MySQL, it's a somewhat bleak picture:
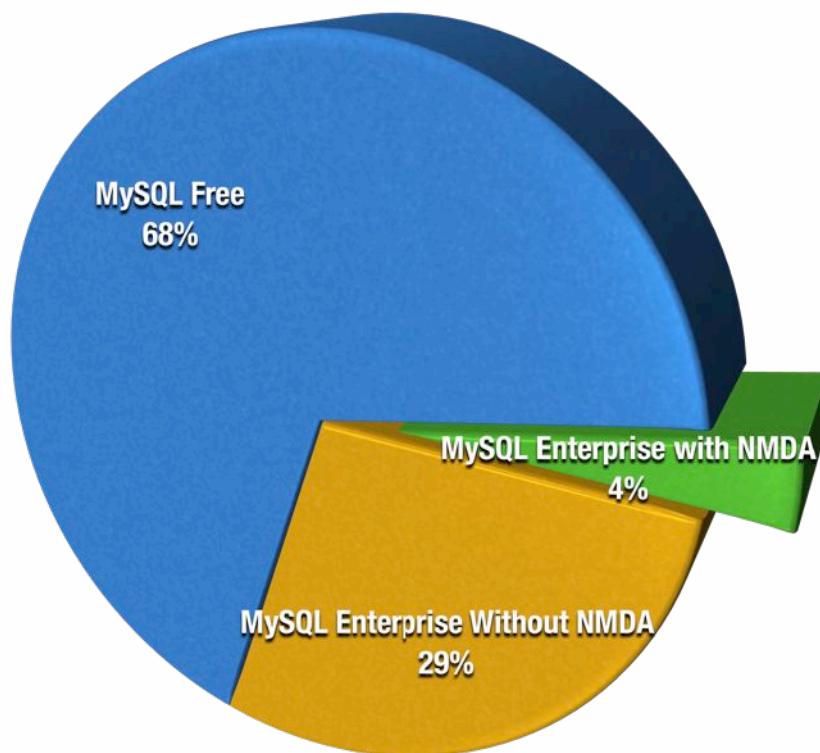
● **MySQL Free**   ● **MySQL Enterprise with NMDA**   ● **MySQL Enterprise Without NMDA**

While a relatively new module, amongst those with MySQL, only 4% of respondents are making use of *both* MySQL Enterprise Edition and the NMDA module that supports it. Even if the other 29% of respondents with MySQL Enterprise Edition were to take up NMDA, EMC are still leaving far too many apples on the tree.
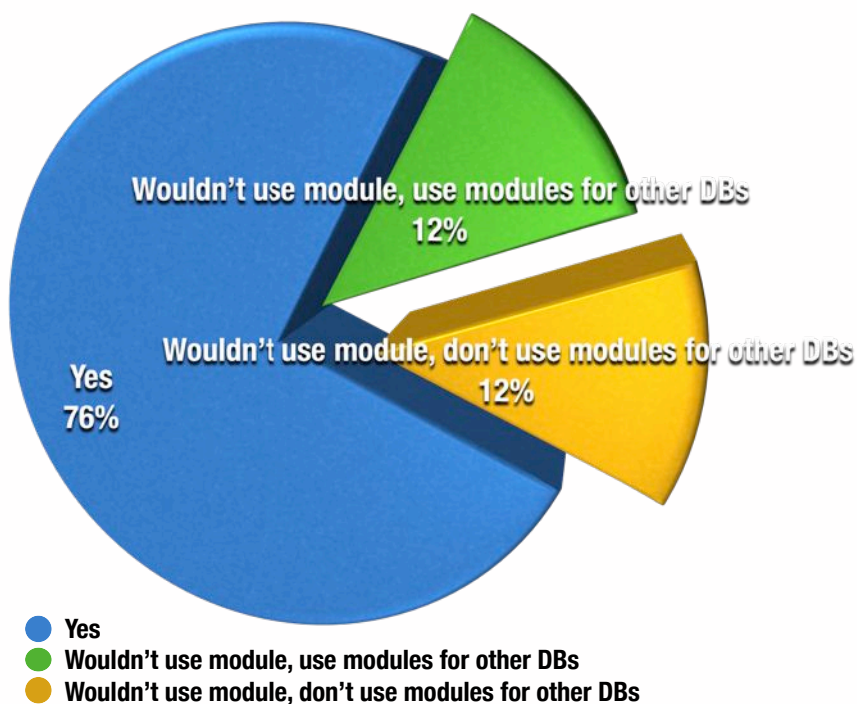

● **Yes**
● **Wouldn't use module, use modules for other DBs**
● **Wouldn't use module, don't use modules for other DBs**

While 24% of respondents wouldn't make use of an OSD module (with it being a 50/50 split between those who don't use modules at all, and those who do), 76% of respondents remain interested in an OSD module, which indicates EMC still has to adjust its outlook.

Particularly with SAP's ongoing attempt to distance itself from an increasingly competitive former partner, Oracle, it is more so than ever the time for EMC to do what Legato had long resisted – develop a generic framework API that can be called by other developers to properly integrate their data stream into a NetWorker saveset. The precedent, of course, has already been set – NetWorker supports NDMP, which effectively allows a black box unit to send data through to NetWorker which NetWorker doesn't understand, yet still store it in a standard saveset format.

Long ago when NetWorker was still a Legato product, there were regular calls by NetWorker users to provide a *pipe* module – one which would allow a third party application to pipe data to the *save* utility (or more generically, wherever it *needed* to go), which would allow for successful integration of products not immediately supported by Legato into the NetWorker backup process. This would result in a fictional database called "HuzzahDB" configured in NetWorker with an appropriate backup command and saveset(s), successfully feeding its data through to NetWorker without NetWorker understanding the actual data content.
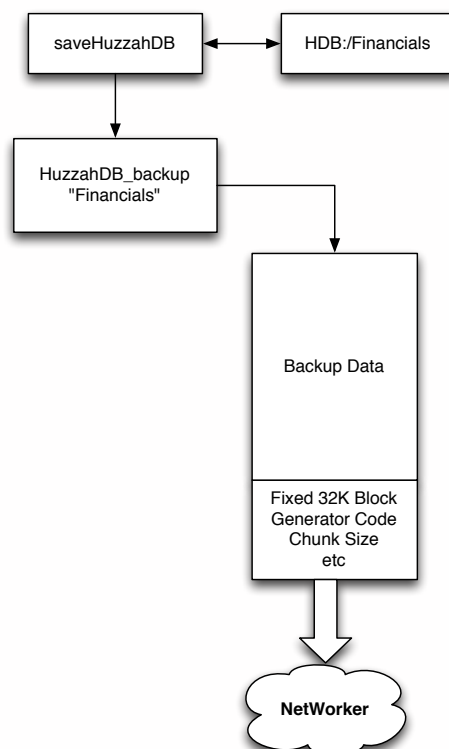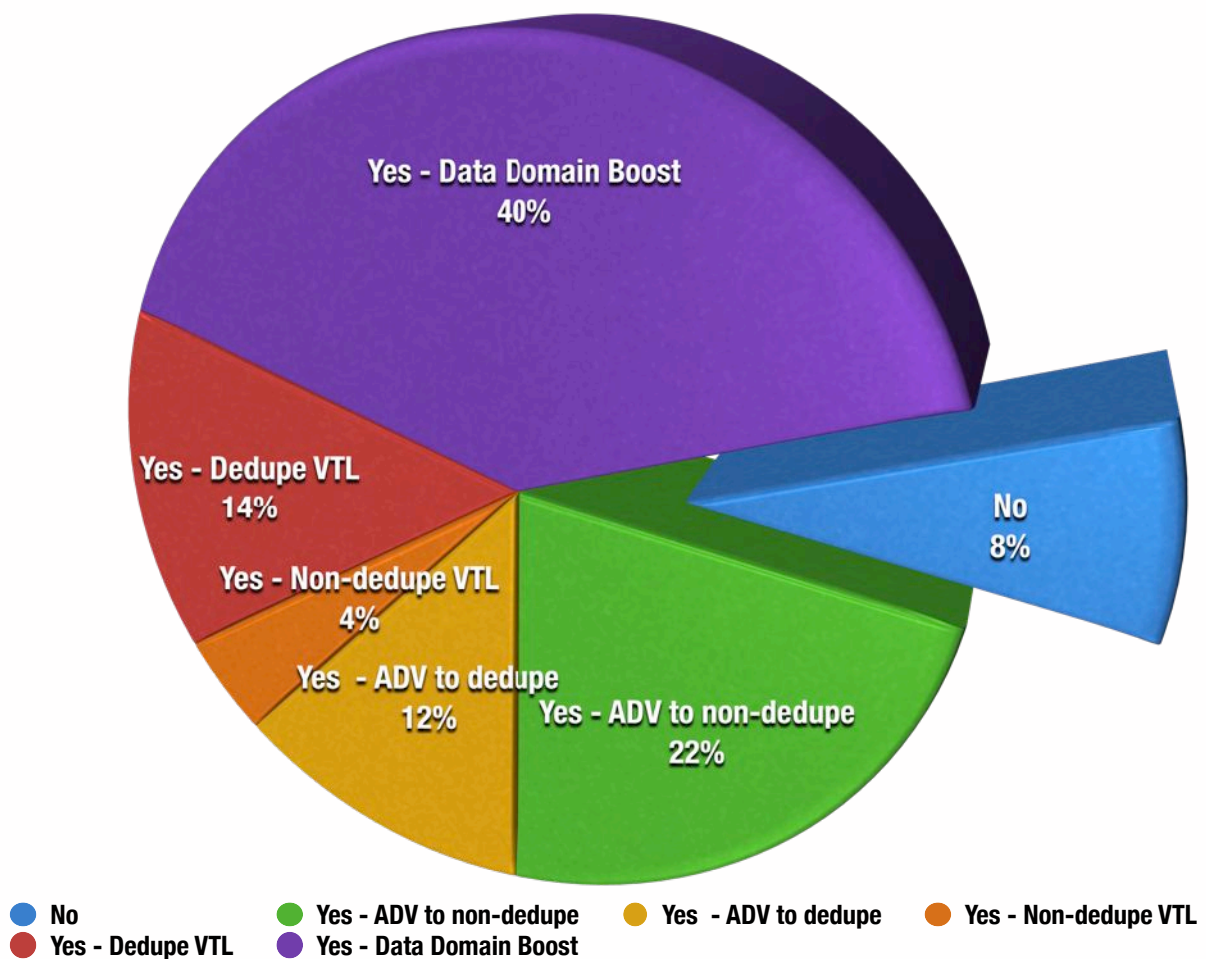

More generically than an OSD module, there is significant logic in EMC developing an appropriate *and supported* save and recover API mechanism for NetWorker.

BACKUP TO DISK TECHNOLOGY

Responses

Continuing the trend of deduplication awareness, this question broke out backup to disk technology further than previous surveys.

Findings

These days it's practically unheard of for a business to still be directly backing up to tape. Typically this occurs in only one of two scenarios: staunchly conservative businesses and businesses whose data sets are so large that disk-based backup presents higher cost than return.

Survey	No	Yes - ADV_FILE	Yes - VTL	Yes - DD Boost
November 2010	16%	52%	32%	N/A
June 2011	15%	47%	33%	5%

Survey	No	Yes - ADV_FILE	Yes - VTL	Yes - DD Boost
December 2012	8%	38%	24%	30%
December 2013	8%	34%	18%	40%

Data Domain Boost continues to grow in use with NetWorker, reducing both VTL and ADV_FILE usage. While initial surveys had shown a drop in *not* using backup to disk, this has plateaued at 8% and may be indicative of the percentage of businesses that will continue to use tape for some time to come.

The difference between deduplication usage amongst VTL and ADV_FILE users is interesting, but understandable. As VTLs matured, deduplication of some form or another (i.e., at rest vs inline) became a more common selling factor, and NetWorker 7.6.x and lower proved popular for deployment with Data Domain VTLs, even before EMC acquired Data Domain.

Conversely, many ADV_FILE deployments are either to cheaper disk in SAN/NAS deployments, 'retired' SANs, and even DAS in larger servers (the Sun "thumper" series, for instance, proved particularly possible for ADV_FILE based backup environments).

Clearly however, any business that has stayed away from backup to disk waiting for the market to mature should take those figures as evidence that it's time to make the move.

CLONING OF BACKUPS

Responses

While there are other ways to achieve backup redundancy, the most compatible method with NetWorker is cloning, which provides a duplicate copy fully accounted for within NetWorker's media database.
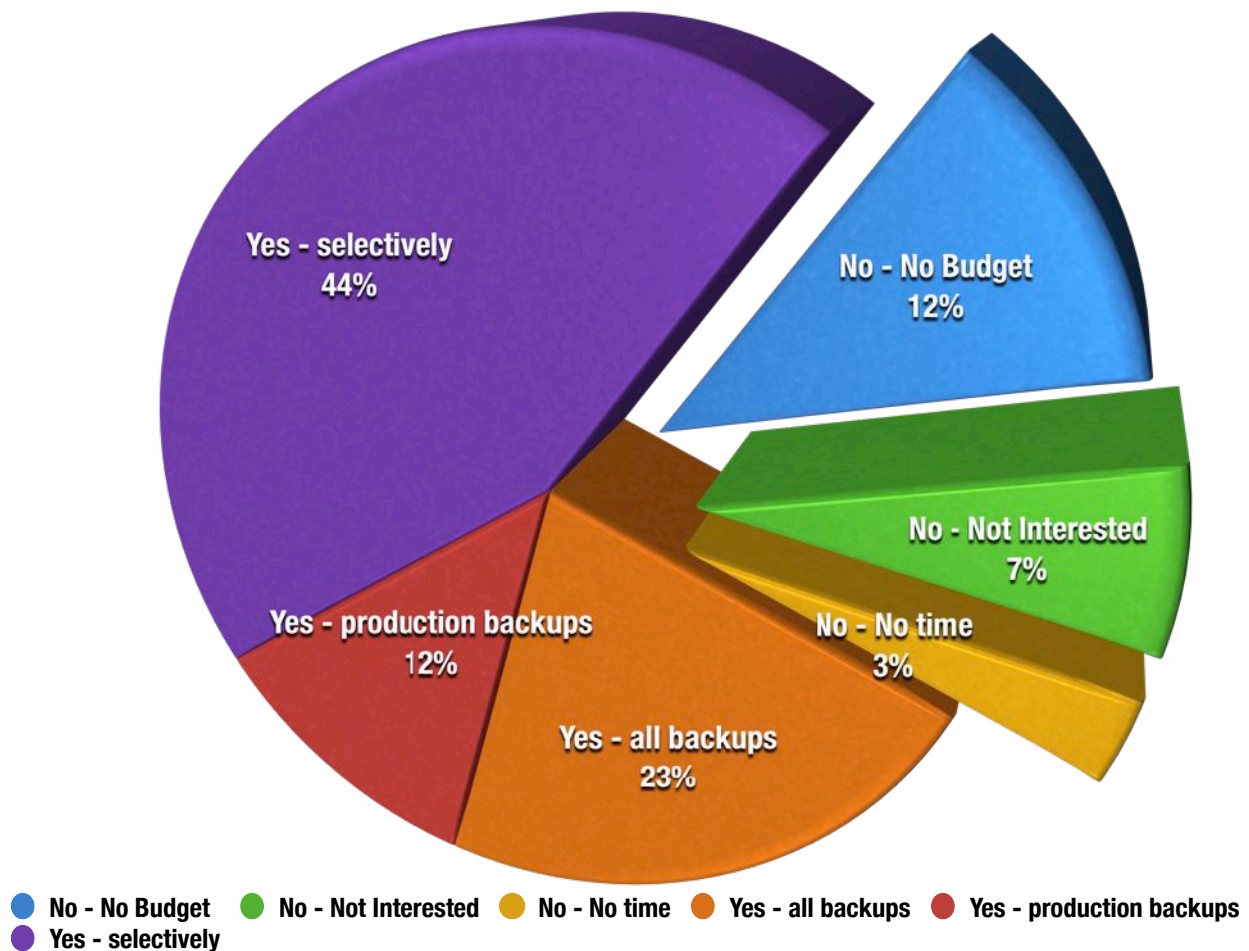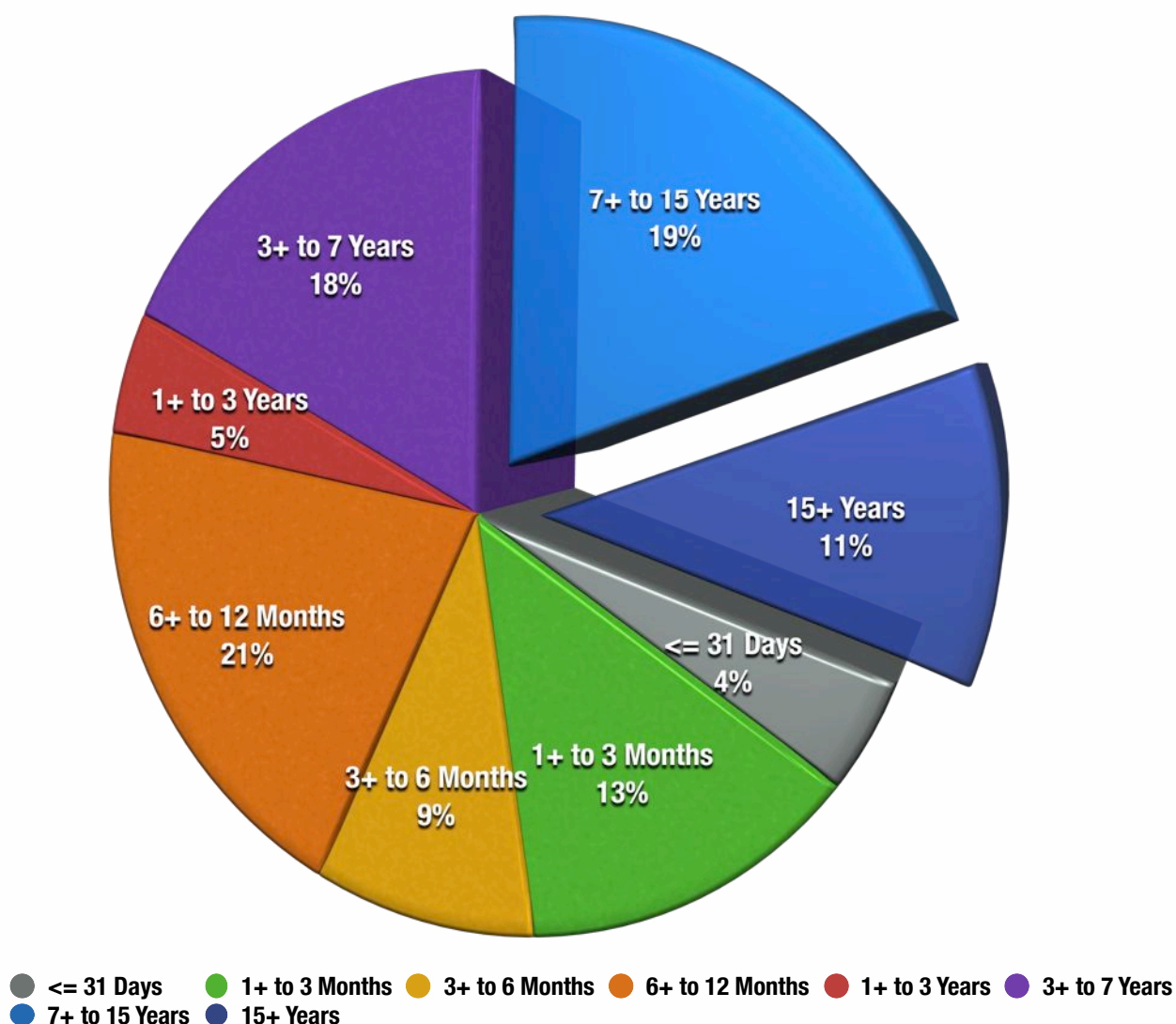

Findings

The overwhelming number of respondents indicated that they clone at least *some* of their backups. 23% indicated they cloned all backups, while 56% indicated they cloned more selectively. Only 22% of respondents indicated they don't clone their backups, with the majority reason being a lack of budget.

It nevertheless remains disturbing that a reasonable number of businesses elect not to clone their backups. Unless the backups themselves are pushed out to secondary use via other data protection strategies (often not the case in those businesses that choose not to clone), they become a potential single point of failure for production systems.

LONGEST RETENTION TIME

Responses

This question gauges the longest period businesses keep their backup data for, once generated.

Legend:
- <= 31 Days
- 1+ to 3 Months
- 3+ to 6 Months
- 6+ to 12 Months
- 1+ to 3 Years
- 3+ to 7 Years
- 7+ to 15 Years
- 15+ Years

Pie chart values: 7+ to 15 Years 19%, 15+ Years 11%, <= 31 Days 4%, 1+ to 3 Months 13%, 3+ to 6 Months 9%, 6+ to 12 Months 21%, 1+ to 3 Years 5%, 3+ to 7 Years 18%

Findings

For over 50% of respondents the longest retention time exceeded 1 year, with 30% reporting 7 or more years.

Typically this covers monthly backups, though some businesses do still separate off monthly and yearly backups, keeping monthlies for less than two years and retaining yearly backups the longest.

These lengthy retention times speak to where tape is still likely to be used within many organisations – as a means of putting long-term backups out of day-to-day circulation in an economical way.

Despite what many disk vendors would have us believe, long-term retention on disk is still a higher cost technique.

For daily/weekly backups, backup to disk (particularly with deduplication technologies) is rapidly becoming the primary backup method:
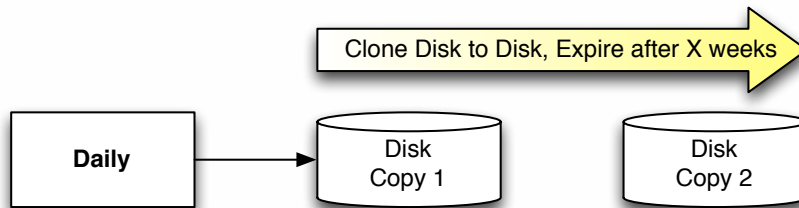
In such a scenario, daily/weekly backups are written to the first disk copy, then cloned to the second disk copy (usually at another location), and expired after the standard retention period for short term backups. Increasingly, such backups are no longer copied to tape at all – redundancy comes simply from secondary data copies. While in theory this is arguably riskier than keeping a tape copy (both disk copies are likely to be online), for many businesses that risk is acceptable.

The longer term backups are now seeing a mix of disk and tape, viz.:
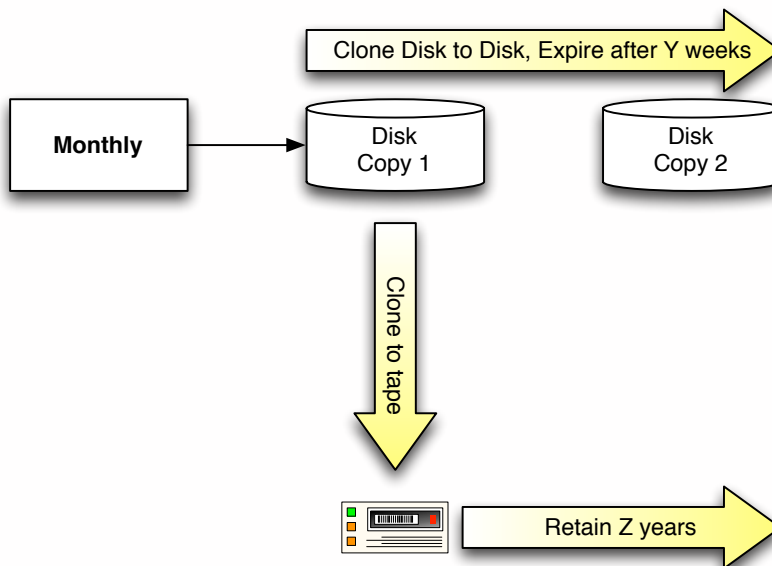
In this scenario, monthly/long term backups are generated first to disk, then cloned to an another disk. This allows standard short term recoveries to still be facilitated from disk.

For longer term storage, once written to disk, the backups are cloned efficiently out to tape. As an added bonus, this can be done either from the primary or secondary disk copy. At some point after the clone to tape, the copies on disk are allowed to expire in order to reclaim space for standard growth.
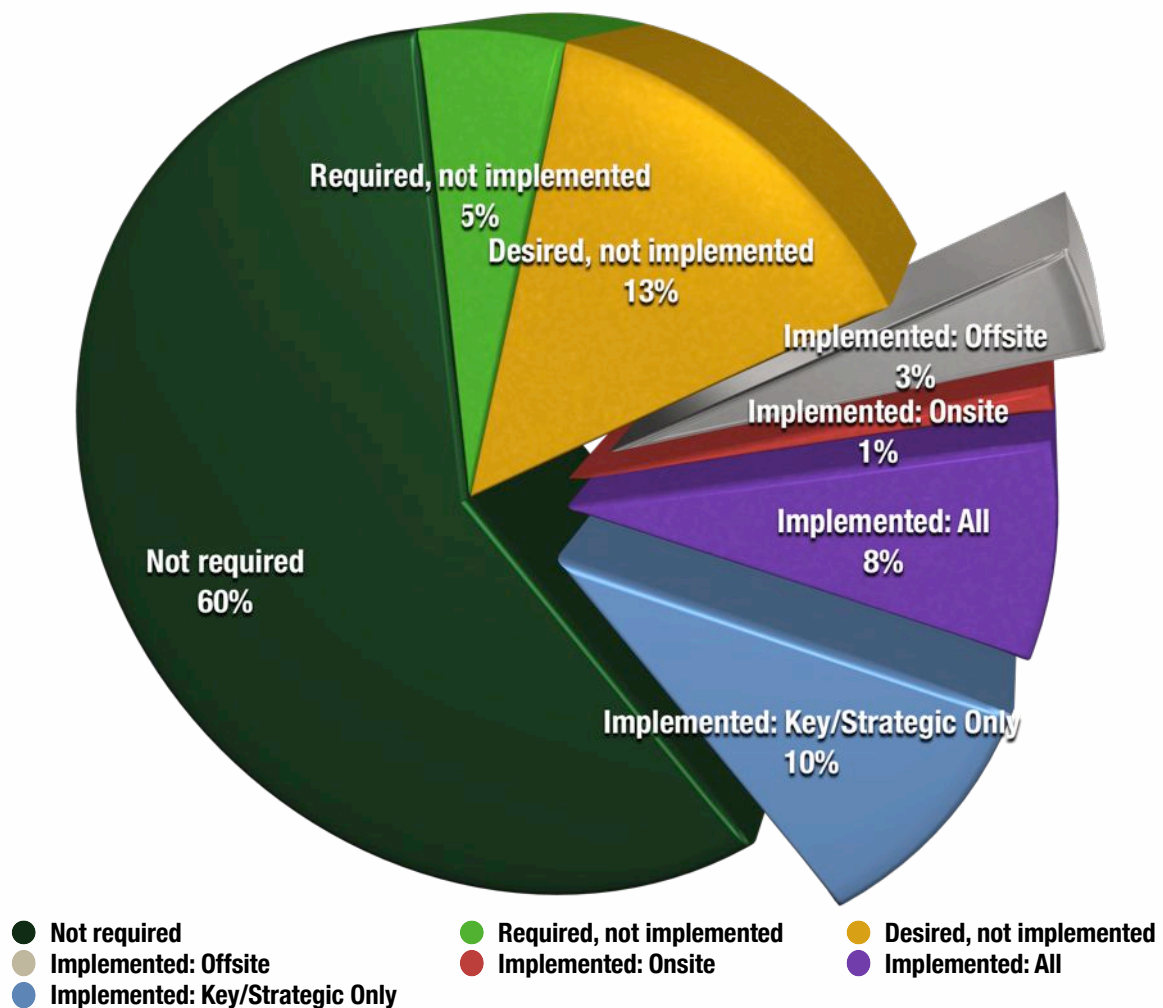
Such strategies maximise ROI and usefulness of disk backup solutions without impacting the economics of long term retention.

ENCRYPTION OF BACKUPS

Responses

A perennial topic for backup solutions is the encryption of data, yet despite availability of various encryption options, it has not yet become a primary consideration for respondents:


Findings

It remains the case that a significant number of respondents do not yet have a business requirement to encrypt backup data. (Uptake in Cloud technologies is likely to present a different requirement over time.)

Just under a quarter of respondents have actually implemented encryption of backups.

To understand how encryption is being used within environments, the next question on the survey focused on just those businesses who are encrypting.
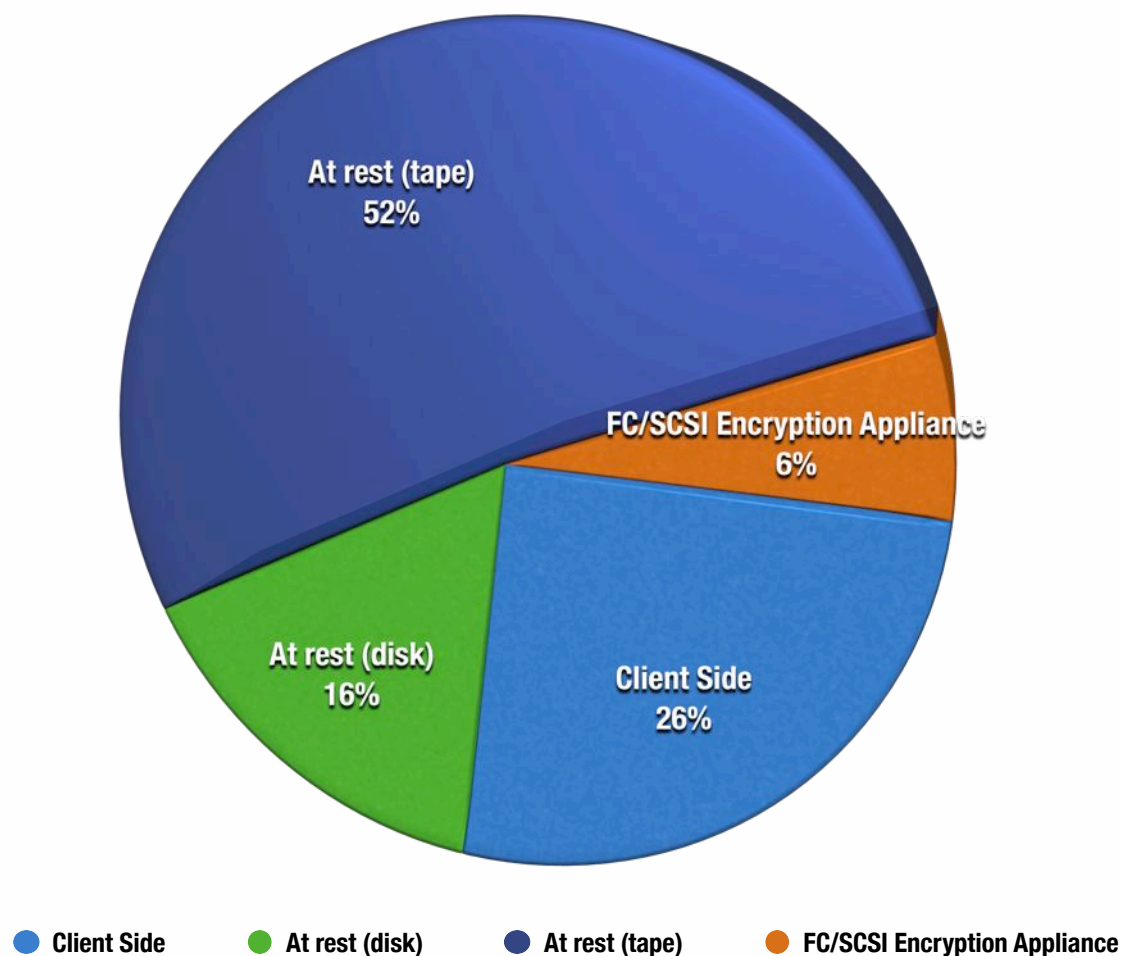
The majority of encryption is currently being achieved at-rest via tape technologies – the enabling of encryption in the LTO standard has undoubtedly been a driving force on this front. Before the advent of the LTO encryption standard, most businesses evaluating backup encryption were considering costly and problematic encryption routers/appliances that would intercept the data between the backup server and the target media. This potentially introduced bottlenecks into the environment that were more difficult to control.

A surprising number of respondents indicated they were engaging in client side encryption – something which is notoriously slow and impactful on the storage process if used with deduplication technologies.

Even with Data Domain and Avamar supporting at-rest encryption of backup data, the number of businesses performing encryption at a disk level is relatively low at 16%. This is perhaps best understood by appreciating that higher physical security applies to disk backup storage. Typically kept within a data centre, disk backups will mostly need the removal of the entire appliance to be recoverable, and *then* the application of any appropriate encryption key. As such, businesses focus efforts on encrypting media which is more readily removable – physical tape.

LONGEVITY OF NETWORKER USE

Responses

Particularly for larger businesses, switching between backup products is usually a costly and challenging activity. This question aimed to determine how long respondents have been using NetWorker within their environments.

- Less than a year
- 1-2 Years
- 2+ to 5 Years
- 5+ to 10 Years
- 10+ Years
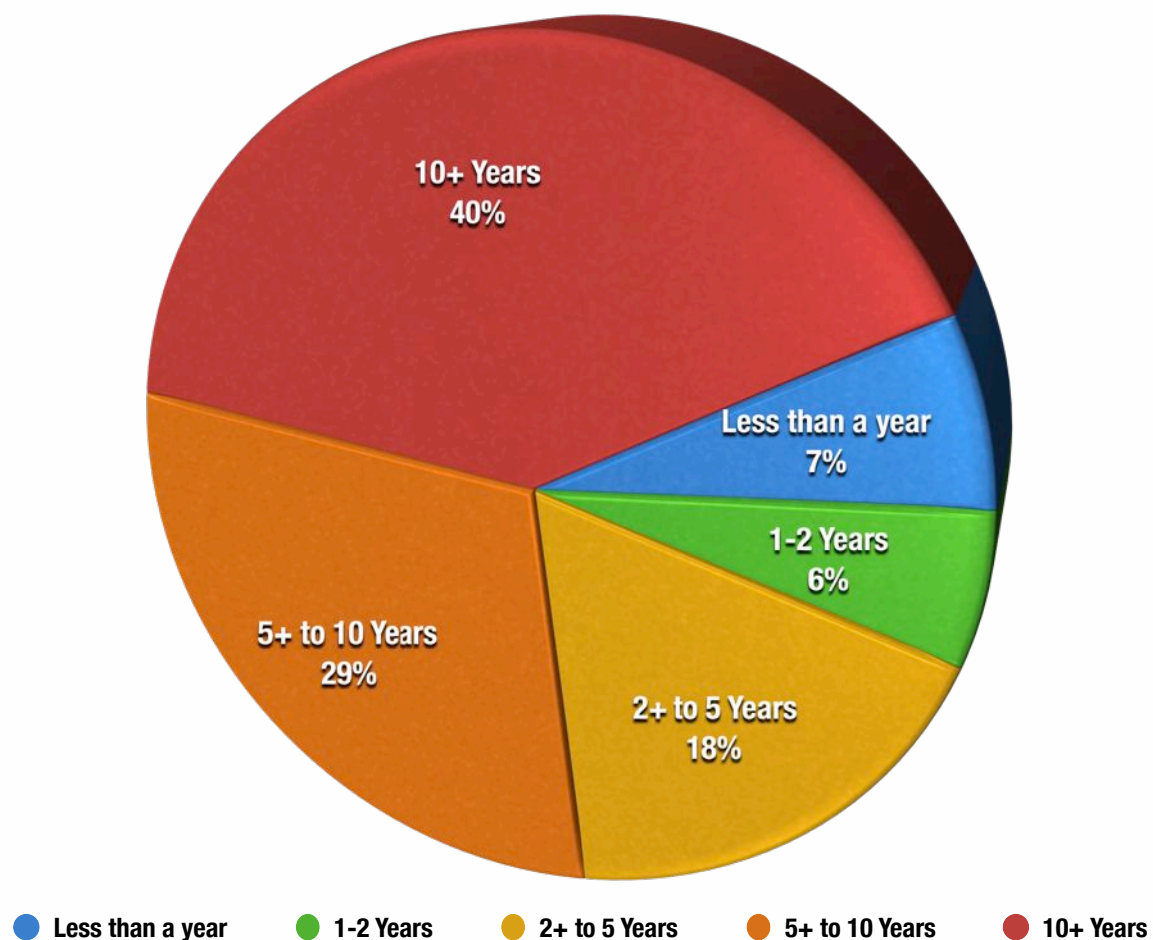
Findings

The numbers are telling – NetWorker has a highly loyal customer base. Given most companies refresh major IT budget cycles in periods of approximately 5 years, the fact that 69% of businesses have been using NetWorker for 5 or more years, and 40% for more than 10 years speaks volumes of the reliability of the product.

Consider, for instance, just those businesses who have been using NetWorker for 10 or more years:
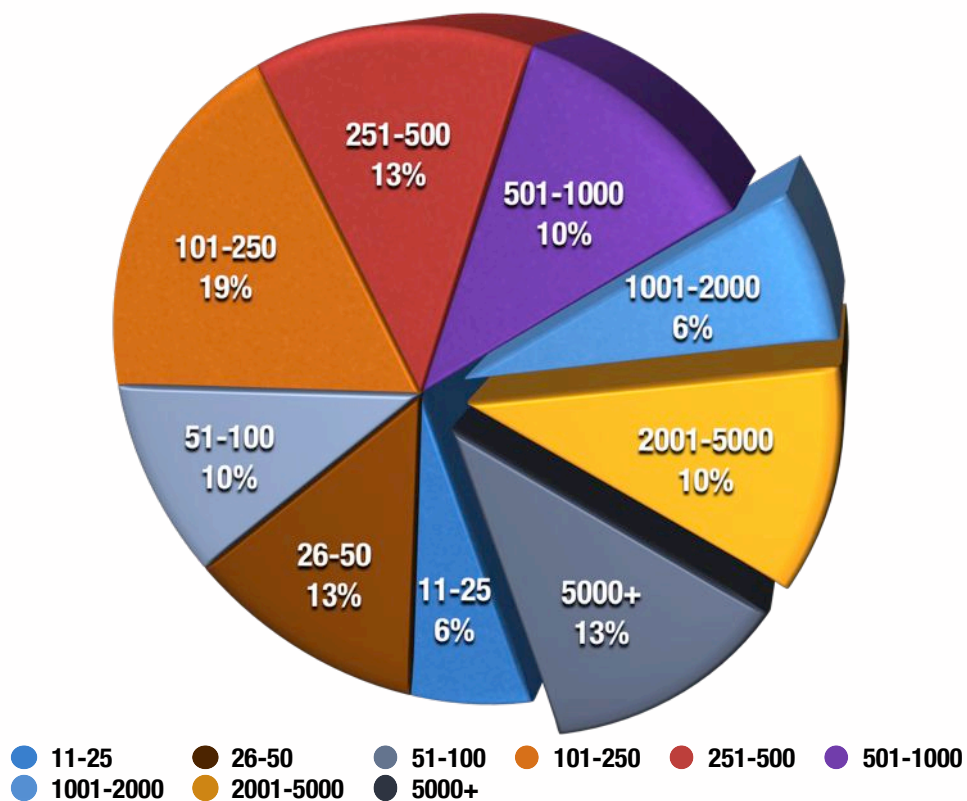
More than a quarter of those businesses are protecting their systems with a product first deployed over a decade ago, arguably when NetWorker was still the flagship product of Legato Systems Inc.
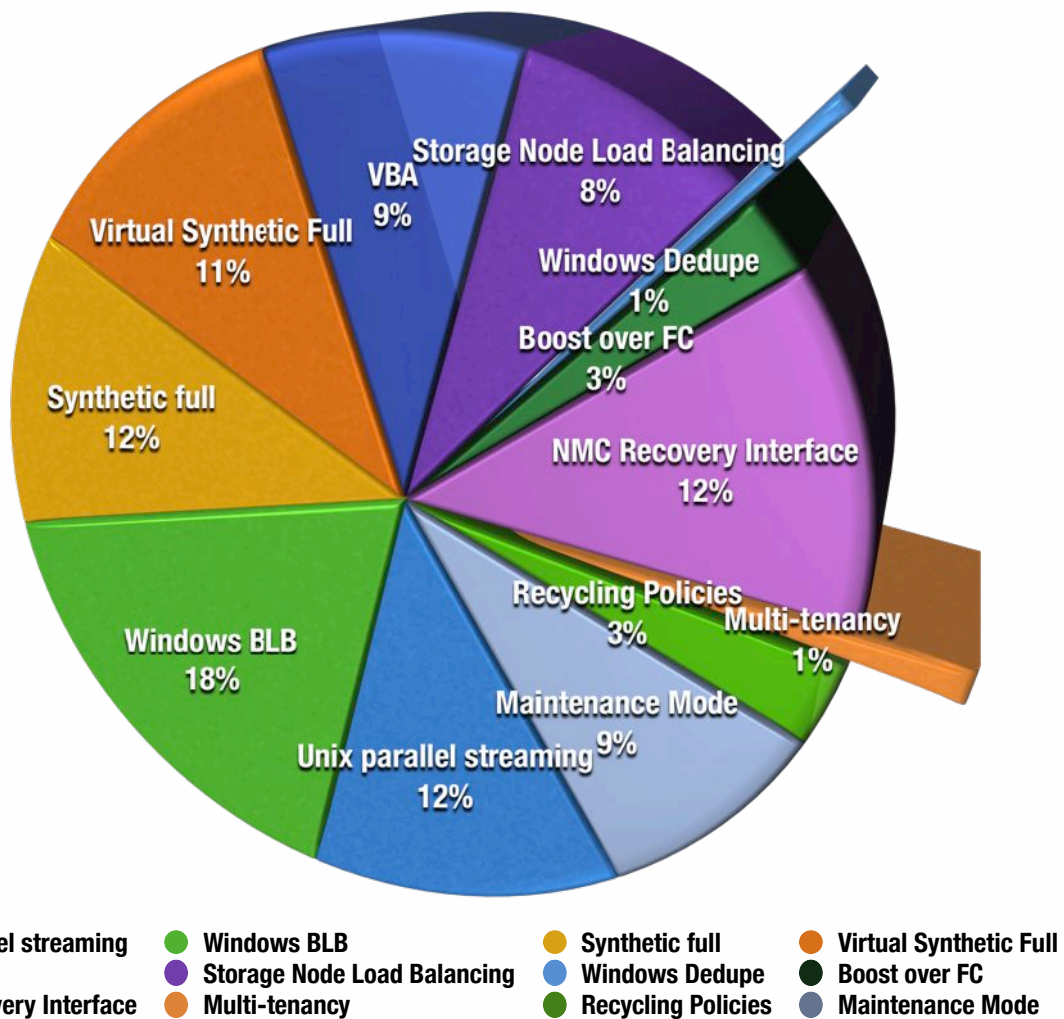
Looking at the other end of the spectrum, examining businesses using NetWorker for 2 years or less, the product is still deployed in a wide variety of client numbers:
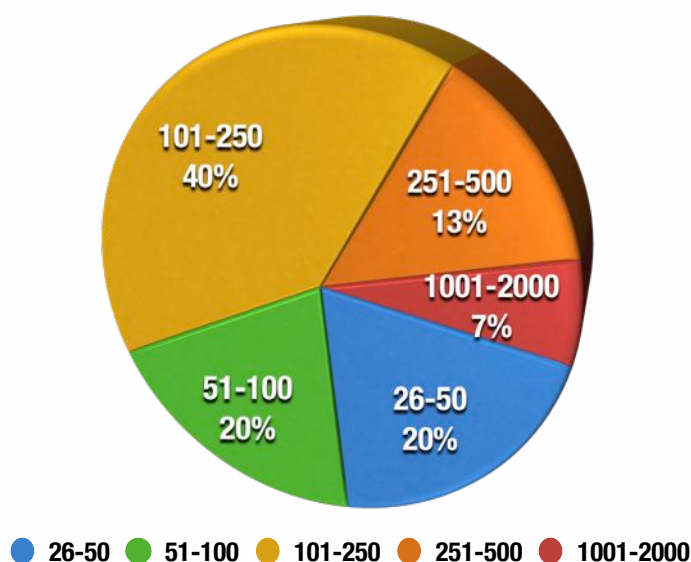
NETWORKER 8+ FEATURES

Responses

NetWorker 8 introduced a raft of new features, and this continued into NetWorker 8.1. The purpose of this question was to gauge which new features are being used by businesses.

Legend:
- Unix parallel streaming
- Windows BLB
- Synthetic full
- Virtual Synthetic Full
- VBA
- Storage Node Load Balancing
- Windows Dedupe
- Boost over FC
- NMC Recovery Interface
- Multi-tenancy
- Recycling Policies
- Maintenance Mode

Findings

Most of those on NetWorker 8.x or higher are making use of at least some of the new features. Surprisingly, few are citing the use of NMC's new recovery interface, despite it being promoted to the primary recovery interface outside of the command line by EMC.

Synthetic fulls, either standard or virtual (Boost) are popular, comprising 23% of responses. Equally as companies look towards taming dense filesystems, both parallel streaming of individual Unix savesets and block level backups for Windows (the successor to SnapImage) have proved very popular, comprising 30% of respondents.

As would be expected, other options that are either more specialised or less likely to be discovered without a comprehensive reading of the release notes, such as recycling policies and maintenance mode aren't as widely used, but still have some uptake.

Given the relative newness of Windows 2012's deduplication filesystems, it seems reasonable that optimised backups for such filesystems is still relatively low. (Introduction of deduplication options directly into base operating systems seems to be a likely trend in future, and hopefully NetWorker will adapt to offer optimised backups to those operating systems too.)

Equally unsurprisingly, multi-tenancy has not been a hit. While the concept is good, and the wizards available to help set it up are useful, the lack of true multi-tenancy makes it less appealing to administrators.

BACKUP ADMINISTRATORS

Responses

As per the previous survey, this question asks whether or not respondents have dedicated backup administrators in their environment.
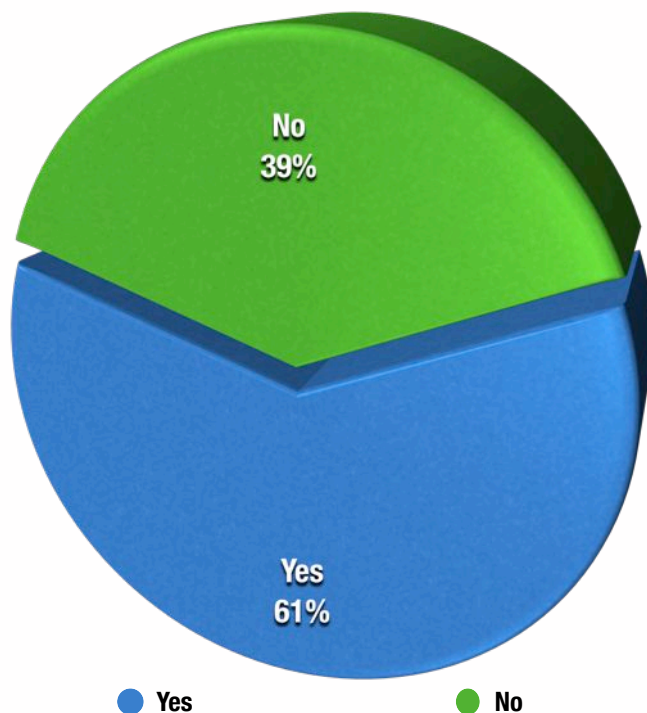
Findings

This represents a drop in dedicated backup administrators between the previous survey and the current, a trend that will be watched.
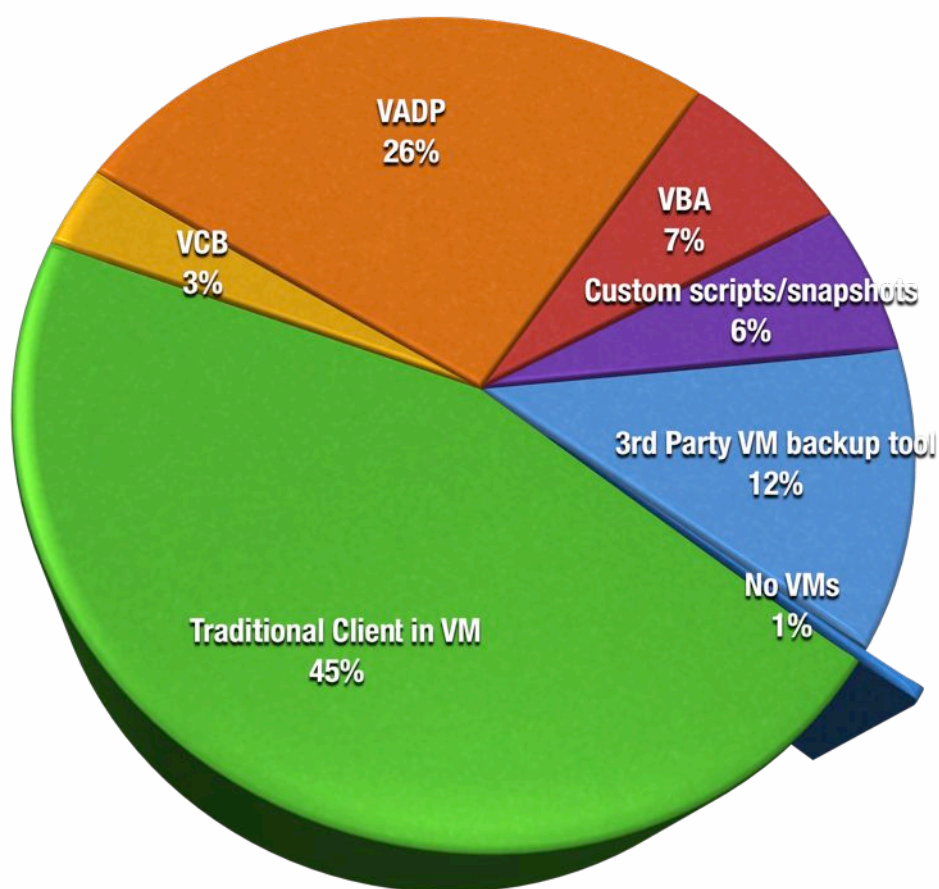
In reality, dedicated backup administration teams are a feature of late 1990's/early 2000's approaches to data protection. The growing complexity created by storage, virtualisation and data protection is best tackled by businesses being prepared to combine those three teams to provide dedicated *service layer* staff, in much the same way as systems administration teams and application administration teams were centralised in earlier decades.

In short, data protection as a rule is no longer something that can be handled in isolation by a single team. It must take storage and virtualisation into account in order to be efficient and maximise return on investment for a business.

VIRTUAL MACHINE BACKUP

Responses

Increased consolidation of computing and application workloads to virtualised farms continues to create backup management overheads for most organisations. VMware's evolving mechanisms on preferred techniques and a continued lack of comprehensive operating system support for file level recovery has not exactly cleared the process up as yet.

● No VMs   ● Traditional Client in VM   ● VCB   ● VADP   ● VBA
● Custom scripts/snapshots   ● 3rd Party VM backup tool

Findings

At the moment the most popular way amongst NetWorker users to backup virtual machines is to install an agent within the virtual machine itself, leading the way at 45%. (Down from last year's 61%.)

That's not to say ~~VADP~~VBA is home and hose yet. Indeed, despite the popularity of NetWorker 8.1, few businesses have moved to VBA. While VBA has presented some better options in the virtual machine backup

space (most notably file level recovery from Linux backups), the lack of a mechanism to seamlessly switch from the old to the new, and the restrictions of VBA has made it more applicable to greenfield application of Virtual Machine image level backups rather than an immediate replacement to VBA installs.

Third party backup tools such as Actifio, Veeam and vRanger are popular with businesses regardless of whether they're using NetWorker, NetBackup or a variety of other conventional backup products. Despite their various disadvantages, the inconsistent approach (thus far) to virtual machine backups by VMware has led to a sophisticated third party tool set which the market is actively exploring – and one which is unlikely to go away any time soon. These backup platforms are here to stay, and will just as likely eat into the traditional guest agent as will VADP and VBA.

CONCLUSIONS

NetWorker continues to be used in a wide variety of environments, with clients numbering from the tens to the thousands, and a fairly broad range of datazones. That being said, Windows and Linux are becoming the dominant platforms for NetWorker servers – the significant failings of Linux with tape is less of an issue these days given the prevalence of deduplication devices such as Data Domain.

The size and scope of NetWorker deployments demonstrates an urgent need for EMC to step back and come up with a consistent cost effective approach to monitoring and reporting on data protection activities. Somewhere between the base level functionality available in NMC, the Swiss Army knife approach of DPA and BRM, a central dashboard and reporting system must be produced which is useful and a minimum cost to businesses.

Solaris continues to decline as a NetWorker server platform, and it is seemingly only a matter of time before the server OS groupings are reduced to "Windows", "Linux" and "Other". Oracle's bellicose pricing models and attitudes towards entrenched Sun customers has seen a substantial dwindling of what was once a significant platform in the data centre to a ghost.

While the virtualisation backup market is maturing, it's not necessarily maturing in a way which is immediately favourable to NetWorker. EMC took a long time getting VADP ready for NetWorker compared to competitors, and seemingly so soon after VADP became available, VMware switched to the VBA methodology. Third party backup tools, for all their limitations can sometimes present a more consistent backup mechanism independent of changes in VMware's strategy. In order to consolidate image level backup of virtual machines back to NetWorker, consistency of approach and improved functionality will remain a key requirement for some time. (The instant-on functionality recently introduced into Avamar 7 when combined with virtual machine backup to Data Domain will undoubtedly make VBA more compelling, should it appear in NetWorker.)

Consistency remains a challenge with virtual machine backups in another form – *application level* consistency. At the moment the most reliable mechanism to generate a backup of a database or complex application, such as Exchange, Oracle, SQL Server, etc., is to deploy an agent within the guest operating system. This allows the databases or files used by the application to be either quiesced or (more usually) processed with appropriate logging such that a subsequent recovery generates usable data. (Hence the database terminology of first *recovering* the data, then *restoring* the database to consistency.) Between VMware and Hyper-V, achieving this application level consistency from an image level backup is something which is at best only in its infancy. Until this is something readily available with image level backups of *most* databases and applications it could be argued that the image level virtual machine backup market remains incomplete.

Big data, cloud and a continuing complexity within the application space is pointing towards a demonstrable need for EMC to release a module-style API for NetWorker. The relatively limited coverage offered by NMDA's MySQL functionality serves as a weighty example on this front.

Data Domain is most definitely the jewel in EMC's BRS crown, and has partly been responsible for a resurgence in NetWorker over the last several years. The ease at which Data Domain can be integrated into either an existing *or* greenfield NetWorker site synergistically offers considerably enhanced backups and excellent return on investment. Ongoing development of Data Domain's Boost functionality will only make this more so over time.

EMC should be congratulated on the rapid up-take of the NetWorker 8.x series. More so than anything else it goes to prove that when care and attention to detail is taken with product development, even in conservative markets, users will jump to a better iteration of a product.

The first usage survey conducted similarly followed the recent release of NetWorker 7.6, and at the time there was a relatively small uptake in that version, despite its advances. Anecdotally, many NetWorker sites had found a version in the 7.x tree they were happy with and reluctant to move onto new versions until they were well proven (or at all). There is no doubt this attitude has been reversed with the NetWorker 8.x tree.

The next survey will be conducted in December 2014.