

Gaining Maximum Value from Information and Data Managing Intellectual Property Assets

Contacts:

Clive Longbottom
Quocirca Ltd
Tel +44 118 948 3360
clive.longbottom@quocirca.com

Dennis Szubert
Quocirca Ltd
Tel +44 1753 855794
dennis.szubert@quocirca.com

Vee Baker
EqualLogic, Inc.
Tel +44 20 7556 7878
vbaker@equallogic.com

Mid-sized organisations have similar issues to large organisations when it comes to the storage and management of information and data, yet they often have limited capability to attract, retain and maintain the skills that large organisations will use in managing their data assets. There is a strong need for more effective tools that will enable mid-sized organisations to manage their data assets and gain the greatest value from them.

- **Mid-sized organisations face many of the same storage problems as large organisations**
Storage volumes continue to grow at a rapid rate, and mid-sized organisations need to be able to meet this growth in a fully manageable manner. Existing heterogeneous storage approaches with a mix of direct attached storage (DAS) and shared folders on servers do not provide the flexibility and opportunities required for ongoing market competitiveness.
- **Storage management needs are similar for mid-sized and large organisations**
Business continuity, disaster recovery, and data archiving are all major issues that mid-sized organisations have to deal with. With existing storage being spread across a range of different data silos, it is difficult to respond rapidly to data loss; and the impact of data loss or non-compliance with information governance requirements can be catastrophic to a mid-sized organisation.
- **Storage area networks (SANs) are the optimum solution for managing such needs**
SAN technology has been proven as a strategic solution in large organisations, and a SAN approach offers several advantages to mid-sized organisations. The abstraction of the storage layer means that response times are more predictable and upgrading storage is more flexible and does not impact the running of the business.
- **High-end SAN skills are at a premium**
SANs based on Fibre Channel (FC) technology require specific skills for implementation and ongoing maintenance. These skills are not easily available within the mid-market, and a perception of complexity and high cost has grown up around the usage of SANs outside of the largest organisations.
- **IP-based SAN storage brings SAN capability within the reach of the mid-market**
By utilising standard Ethernet-based technology, IP-based SANs lower the skills requirement for the implementation and running of SANs, making them a prime solution for the mid-market. Bringing high-end SAN functionality into the reach of the mid-market yields direct business value through higher storage utilisation rates, less downtime, greater storage flexibility and the capacity to report across multiple data stores in an easier manner.
- **Virtualisation provides functionality for flexibility**
The capability to create a single virtual view of all storage assets, combined with the capability to use logical partitions to provide flexible “buckets” of storage for applications, allows for high levels of flexibility for organisations.

Conclusions

SANs help to create a very flexible storage solution that provides a high degree of future-proofing for an organisation's storage needs. However, existing perceptions of complexity and high cost have historically relegated SAN technology to only a few companies with large IT budgets in the mid-market. The advent of IP-based SAN technologies and the combination of standard Ethernet-based SAN connectivity with the decreasing cost of storage components now provide mid-sized organisations with the capability to easily enter the SAN world. For those who choose a solution complete with management tooling around virtualisation, partitioning, provisioning, backup/restore and so forth, IP-based SANs will provide immediate business value and higher levels of efficiencies and effectiveness in their market.

REPORT NOTE:

This report has been written independently by Quocirca Ltd to address certain issues found in today's organisations. The report draws on Quocirca's extensive knowledge of the technology and business arenas, and provides advice on the approach that organisations should take to create a more effective and efficient environment for future growth.

During the preparation of this report, Quocirca has spoken to a number of suppliers and customers involved in the areas covered. We are grateful for their time and insights.

CONTENTS

| | |
|--|----|
| 1. INTRODUCTION..... | 3 |
| 2. THE INTELLIGENCE AGE - DATA, INFORMATION AND KNOWLEDGE..... | 3 |
| 3. STORAGE REQUIREMENTS | 4 |
| 4. BASIC STORAGE APPROACHES..... | 5 |
| 5. THE ROLE OF VIRTUALISATION | 5 |
| 6. MAKING STORAGE AN EFFECTIVE BUSINESS ASSET | 6 |
| 7. STORAGE CASE STUDIES..... | 7 |
| 8. CONCLUSIONS AND RECOMMENDATIONS | 8 |
| ABOUT EQUALLOGIC..... | 10 |
| ABOUT QUOCIRCA | 11 |
| APPENDIX A: RAID AND DISK TECHNOLOGIES | 12 |
| APPENDIX B: STORAGE ASSET GROWTH, REUSE AND TIERING..... | 13 |

1. Introduction

Data growth shows no sign of slowing, and the increasing use of new data types (particularly voice and video) is stressing many organisations' approach to information storage and management. The need to integrate an organisation's information into a single resource pool, rather than as discrete pools of information, is growing faster than the underlying storage rates. Also, the speed of storage system response, reporting and data recovery degrades as the volumes of data grow. With these main worries for mid-sized organisations, it becomes clear that carrying on as we are is not a real option. Data retention laws mean that it is increasingly difficult for an organisation to minimise the amount of data and information that it stores, yet the business demands that the information required for decision making is easily and rapidly available – and action has to be taken now, before the problem gets any worse.

Even for large organisations, maintaining consistent knowledge of what is happening in the storage world is proving difficult. Storage management skills are increasingly being attracted to system integrators and storage specialists, leaving little in the way of skills for the medium and small organisations.

However, the problems around storage management are similar for the large and the medium sized organisations – only the scale of the problem is different. Mid-sized organisations have to be flexible and able to respond to internal and external market forces, plus the technologies chosen must be able to support and facilitate the needs of the business.

This paper examines these business and technology needs that a mid-sized organisation is battling on a daily basis and provides insights in to how such organisations should approach these issues.

2. The Intelligence Age - Data, Information and Knowledge

Main Findings:

- The rate of growth in the volume of data and information continues to increase
- Lack of effective control over information assets leads to poor decision making

Data growth continues to accelerate beyond levels that can be easily managed – indeed, the overall quantity of data stored in electronic format is estimated to be doubling at a rate of less than every six months. A large proportion of this is driven by consumer storage of digital pictures, music and videos; but organisations are also seeing rapid growth in volumes of commercial data. This corporate data growth, often complicated by the need for internal and legal compliance, remains a major concern and has driven organisations to consider information storage and lifecycle management from a business point of view, rather than from a purely technical one.

For mid-sized organisations, this growth in data and information volumes runs the risk of being uncontrolled – not only will there be corporate information that needs managing, but also a mix of personal data, such as letters, photos and mp3s that may be being saved to central shared drives by end users. This blurring between the personal and

the business is something that is difficult for a mid-sized organisation to effectively manage – a highly proscriptive approach can rapidly lead to employee demoralisation and to ineffective and inefficient work processes.

Many mid-sized organisations are simply drowning in data. While up to 80% of stored information may have little to no direct business value, it has proven difficult to identify which 80% this applies to. Data and information is stored on a “just in case” basis – to show compliance, good governance or the capability to trend information. Large organisations will often filter data as it is being stored and will utilise advanced search and retrieval tools to facilitate better information identification. For the mid-sized organisation, many of these tools are perceived to be out of reach from both a financial and complexity perspective – and yet the problems are just as acute for them as they are for larger organisations.

In Figure 1, we see an information triangle. Individuals and groups within an organisation need to pull out information from its underlying data assets – either through running reports against formal data stored in databases, the creation of ad-hoc written reports based on employees' work and understanding or on the aggregation of views based on a mix of internal and external information. From this information, an executive or knowledge worker within the organisation can then make decisions based on a degree of knowledge that they would not have had previously.

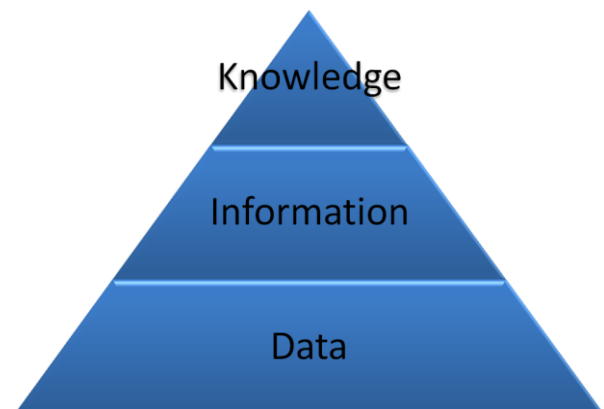


Figure 1

Although organisations would like to be in the position where these decisions are made based on a view of all the information available, the reality is that they are far more likely to be in the position shown in Figure 2 – there are plenty of disparate data sources available to them, but these are stored across many different environments and technologies. This makes extraction of information from the data difficult, even with powerful data mining and business intelligence solutions to hand. Indeed, many decisions are being based upon a subset of the available data, leading to a dangerous “perception of knowledge”. Here, information that would change the decision may not have been available through the chosen search and reporting mechanism acting only against certain data stores.

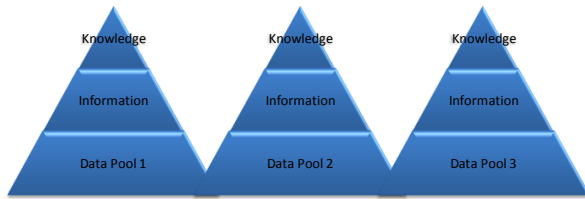


Figure 2

Although many see this as a problem only affecting large organisations, the mid-market is increasingly coming up against the same issue – as they implement multiple storage assets, including file servers, email data stores and more formal databases, all are being held on completely separate storage assets and managed in completely different ways.

Even those who have looked at creating a common view across existing data storage assets have run into other problems – response times slow as data sources are sequentially searched and information is aggregated and correlated before it can be reported on.

To address this issue, it is necessary to adopt a solutions-centric approach to data storage – how to effectively optimise storage to respond to the demands of business-critical applications and composite solutions. Here we need to consider response times and how the provision of data at the optimum rate can impact the overall performance of an application and process. For this, data needs to be a single virtualised asset that can be more easily managed and interrogated as a single resource which can then facilitate the increasingly frequent changes in business processes that are seen in today's dynamic markets. For mid-sized organisations, this all has to be done within the constraints of the financial and human resources available.

3. Storage Requirements

Main Findings:

- Data and information assets have to be seen as a single corporate resource
- Virtualisation provides the basis for a cohesive information management architecture
- Information assets have to be secured and have suitable business continuity protection solutions applied

Basic storage needs

At a basic level, all information needs to be stored in a secure and resilient manner. However, these days this is not enough, and organisations have to ensure that information is easily available to meet the needs of individual applications, composite applications based around web services, reporting and business information tools and the needs of individuals within the organisation. The need for a “360 degree” view in many cases requires a means of bringing data together to ensure that decisions are built upon a contextual view of all the information available to an organisation – and not just on the data belonging directly to an application at any one point in time.

The main problems that mid-sized organisations find include: that they have had to overprovision their individual technical environments to meet expected peak loads, that multiple instances of applications have been introduced to meet

specific needs, that information is being stored in multiple islands and silos and that far too much human resource is being utilised to manage and maintain the environment.

The first step for many mid-sized organisations is to look at how existing storage needs can be rationalised and consolidated. Through consolidation, the over-provisioning of the past can be brought under control – but only if the consolidation is carried out in a fully planned manner. Here, rationalisation to a minimum set of storage assets and to a common set of storage and information management tools will ensure a fully optimised solution.

To this end, organisations should consider a virtualised storage infrastructure that makes all their main storage look and act like a single pool of resource.

This use of virtualisation provides the underpinnings for a far more flexible infrastructure: applications can be abstracted from the underlying physical storage, and the applications can then be updated and moved far more easily than when the application and data were tightly coupled to physical devices. Scalability becomes an issue of the past – as more storage is required, it can be introduced as part of the resource pool and rapidly made available to the application needing it. The server infrastructure becomes more dynamic – applications that need to have more storage can have their needs fulfilled on an incremental basis, with logical units of storage being served from the resource pool.

Resilience and data protection

Although improvements in the manufacture of disk drives have led to a dramatic lowering of failure rates, as they are mechanical devices disks will still fail over time. If steps have not been taken to ensure that the data on a disk has been secured, then such a failure will lead to irrevocable loss of that data.

Simple backup/restore, while having its place in the overall data policies of an organisation, does not provide for business continuity and could heavily impact a business for several days while full and incremental restores are being carried out. However, by utilising storage techniques that write data across multiple disks such that there is always more than a single copy of any item of data, we can ensure that any failure of a single disk will not impact the business. To do this, we need to look at the use of a Redundant Array of Inexpensive Disks, or RAID.

For those who would like a deeper look at RAID and at disk technology, a primer can be found in Appendix A.

Storage asset protection

As a mid-sized organisation grows, it must also allow for flexibility and for the optimisation of its investments in existing technologies. With many old-style approaches to storage, any significant growth required a “fork-lift” upgrade, with existing storage assets being removed and replaced with the new. This is not only inefficient at a cost level, but also severely diminishes an organisation's ability to continue with its business as storage assets are being replaced and data is being backed up and restored.

Newer approaches can allow existing storage assets to be seamlessly expanded through the addition of new physical assets, and also for older, slower assets to be designated as second tier or archival storage. This seamless growth, reuse and “tiering” of assets is covered in more detail in Appendix B.

4. Basic storage approaches

Main Findings:

- An immediate priority is to move away from direct attached storage (DAS) based solutions
- Although network attached storage (NAS) has a part to play in certain storage environments, modern storage area network solutions (SAN) provide the best approach for mid-sized organisations

Storage types

• Direct Attached Storage (DAS)

DAS is the most basic form of storage and is the type of storage found within desktops and basic servers. The storage is essentially dedicated to the one machine and is difficult to share in any meaningful manner. Even the use of folder “shares” (where a user gains access to a remote folder held on another machine’s disk) does not markedly increase utilisation rates, nor does it provide the levels of flexibility and data resilience that mid-sized organisations require.

• Network Attached Storage (NAS)

NAS takes the storage assets and attaches them as a separate resource to the existing network, accessible to multiple servers and/or users via the network. All data being written and retrieved traverses the same network as other network traffic. As a simple means of adding storage resource to a network, NAS is a relatively cheap approach, but the overall impact on the performance of the underlying network has to be borne in mind.

• Storage Area Network (SAN)

SANs take a similar approach to NAS, offloading the storage medium to a separate environment, but they also offload the majority of the data transport away from the main network on to a separate network. For highly data-intensive environments, SANs provide the optimum means of ensuring data response; but for many companies, the need for specific skills in areas such as traditional Fibre Channel technology and the costs associated with SAN deployment and management have held them back.

SAN storage comes in two main types – Fibre Channel and IP-based storage.

Fibre Channel Approach

Fibre Channel uses optical technology to gain very fast data transfer rates on the storage area network. However, the network interface cards (NICs) required to support Fibre Channel remain at a high cost compared to standard Gigabit Ethernet controllers.

Fibre Channel also requires different skill sets for laying the requisite cable and for managing the environment. For example, terminating Fibre Channel is a specialised job, where a marginally misaligned optical joint in a termination will either make a connection unreliable or will stop it from working at all. As another example, the angles that fibre can be bent round are far less than can be done with today’s Ethernet cables, with fractures of the optical fibre or light leakage being relatively common occurrences where the requisite care has not been taken.

A Fibre Channel environment will also require specific skills in technical management, with specific tooling to identify

root cause of any problems on the fibre network itself. For many companies, the cost of acquiring such technical skills and tooling makes the cost of a Fibre-Channel-based SAN too daunting.

IP-based Storage

Ethernet is now the ubiquitous network connection technology for organisations worldwide. From its origins as a simple hub-and-spoke alternative to older ring-based technologies, Ethernet is utilised in organisations of all sizes as well as many consumer environments. Skills are readily available, and improvements in technology have led to extremely cost-effective availability of network interface cards (NICs) providing data transport speeds of up to 1 Gb/s, with 10 Gb/s already available at a premium cost.

For the majority of mid-sized organisations, a SAN based on 1 Gb/s Ethernet provides an entry point into highly resilient, controllable SAN-based storage architectures at lowest cost and skills requirement, and it will deliver ample throughput for even the most data-intensive applications.

For many mid-sized organisations, the option of a SAN has been perceived as being too costly, too complex and requiring technical skills well outside of the organisation’s capabilities. However, modern approaches to SANs means that the high-end capabilities that have been available predominantly to large organisations are now increasingly available to the mid-market. Enterprise-class IP SANs are available at a suitable price point with simpler deployment, eliminating the need for in-depth skills and now requiring only those skills that can be easily found within a mid-sized organisation’s existing portfolio.

5. The Role of Virtualisation

Main Findings:

- Storage management is made easier by virtualisation
- Virtual SANs are an ideal match for virtual servers

Storage virtualisation uses a layer of abstraction between storage assets and server hosts to mask the underlying complexity of the physical storage from physical servers and administrators alike, allowing them to focus on the task in hand rather than the technology. It provides a dynamic pool of storage, from which disk space is presented to the operating system as a set of logical volumes that appear to be normal disks as far as the operating system and applications are concerned.

This technology should not be dismissed as merely of technical interest – it provides real-world benefits to users. Not only does it improve efficiency and productivity, but it also reduces the time, effort and knowledge needed to manage storage. In addition, it is a key enabler in the following areas.

Tiered Storage

Tiered storage provides a useful tool for creating a flexible and cost-effective storage environment. Within a storage pool, different tiers can be created, allowing a single pool to be composed of different disk or RAID types. Storage pools could be defined by department, for example, to prevent resource competition, or by status (production and archive), offering a “SAN within a SAN” functionality – all within a single SAN view. Virtualisation also enables movement of

volumes or arrays between storage tiers online, without disrupting applications. More can be read on tiering in Appendix B.

Automatic Load Balancing

With automatic data placement and optimisation in the SAN to ensure maximum utilisation of both capacity and performance, administrators can add disks or arrays online without disruption; and data is automatically load balanced across disks, controllers, cache and network ports.

Thin Provisioning

Thin provisioning is an advanced form of storage allocation for growing applications. Historically, all the storage an application will need over a period of time is allocated up front, avoiding the need for repeated data growth operations. With thin provisioning, the inefficiencies of over-provisioning can be avoided by limiting initial physical storage allocation to what is needed now, with further storage resources being automatically applied as the application grows.

Virtualisation

Using virtualisation, there are two distinct views of storage – a virtual view, where available storage size seen by the operating system, applications and users is based purely on what is allocated to them at any one time, and a physical view of the underlying storage assets for administrators. The former is highly flexible, can be changed at will and is only dependent on the overall limit of the physical assets. A key result is improved utilisation of physical storage resources.

In today's IT environment, with the increasing trend toward virtual servers and technologies that allow for the flexible movement of virtual machines between physical machines, virtual servers ideally should be matched with virtual SAN storage with automatic load balancing to provide the most flexible and fluid virtual infrastructure implementation.

6. Making Storage an Effective Business Asset

Main Findings:

- Storage assets need a comprehensive set of management tools to provide the optimum solution for the organisation
- A virtualised storage environment, combined with dynamic management capabilities, provides a highly flexible and cost-effective approach for an organisation

From the simple days of direct attached storage, where the main management requirements were to know how much space was available, what the overall health of the disk was and the availability of tools to carry out tasks such as defragmenting, the storage management arena has now become a market all to itself.

With today's storage environments, it has become incredibly important to be able to virtualise the storage environment and then to dynamically manage this resource pool. Within the management function, we need to look at how we implement areas such as backup and restore, how we approach business continuity, how we best utilise the storage resources we have available to us and so on.

This leads to the need for a well-integrated suite of management capabilities. Within the large organisation environment, these capabilities are generally provided through the integration of a combination of base-level systems management solutions with best-of-breed storage management and information management solutions. Often, these solutions address individual aspects of the problem, such as email management or document management, with different layers of tools being required to create the overall view of the information that is required by the business itself.

This is not a feasible approach for mid-sized organisations. The base-level cost combined with the cost of skills and of ongoing maintenance will essentially negate the overall business value of the approach.

Therefore, bearing in mind that the basic information-management needs for an organisation of any size remain constant, what are the differentiated needs for mid-sized organisations in information storage and management?

- **Effective cost of acquisition**

High-end solutions may well be required for the largest organisations, but mid-sized organisations require a manageable up-front capital cost. Buyers should be looking for a price that is all inclusive, with no hidden costs. Some storage vendors offer packaged functionality that comes in a complete product.

- **Ease of implementation**

Many mid-sized organisations will not have the range of skills to enter into complex storage architectures with multiple layers of management tooling. Storage within this market must be easy to implement and should have easily implemented management tools associated with it.

- **Flexibility**

As an organisation grows, it needs to adapt its operations to the changing conditions of the marketplace. The speed of change in their markets drives the need for flexibility within such organisations. It is imperative that the storage capabilities within the organisation have the built-in flexibility to support critical responses to these changes.

- **Growth capability**

The vast majority of mid-sized organisations want to grow. Even for organisations with slow business growth, the volumes of data and information continue to grow rapidly, driving the need for storage solutions that can grow to meet these demands. However, this storage growth must be manageable within the cost and skills constraints of the organisation and also must be capable of being implemented without impacting the capacity for the organisation to continue its business. Look for backwards compatibility with older equipment from the same vendor, and ask for assurances for upwards compatibility with future equipment.

- **Manageability**

Unmanaged storage runs the risks of containing important information assets that cannot be easily retrieved for reporting purposes. Also, such storage may not be recoverable on failure, leading to an organisation being out of compliance. Look for storage that has simple yet effective management tooling that enables information and data to be protected, secured,

backed up and restored easily across all available storage assets.

- **Investment protection**

As time passes, today's top-tier storage assets will be seen to be slower and less effective than newer storage offerings. It is necessary to ensure that newer equipment can work alongside existing equipment to maximise the investments made in the storage environment. Users should look for products that allow component upgrades to maintain relevance and performance currency.

- **No special skills required**

Mid-sized organisations cannot always afford to attract and retain people with the level of skills required for the implementation and running of complex storage environments. Nor can they afford large numbers of IT management employees. However, by carefully choosing the storage solution, the level of skills and number of IT professionals required can be minimised. Tooling that graphically shows the state of all storage assets, that allows logical storage to be provisioned and re-provisioned at will and which enables an administrator to add physical storage easily can make enterprise storage capabilities valid for the mid-market. Combine this with easy tools for managing data backup and restore, applying policies around business continuity, setting storage limits for applications/users and reporting on events to provide proactive advice. Such a storage environment could be easily managed as part of an existing mid-market IT function.

Likewise, using a highly standardised means of attaching storage to the existing infrastructure (i.e. through the use of Gigabit Ethernet) makes implementation easier and can be carried out by existing resources within the organisation.

- **Open standards support**

Storage systems used to be massively proprietary, with each vendor having their own means of managing storage assets and the information held on them. However, through the use of open storage standards, systems can be far more intelligent, being more "aware" of their environment and capable of out-of-the-box integration with applications utilising the systems. For virtualisation, it is imperative that applications can utilise such virtual storage without the need for the applications to be changed in any way – the logical partitions must be able to be seen by the application as if they are standard physical units.

- **Low running costs**

A key area for the mid-market is the longevity of any storage solution. Organisations will want to sweat the storage assets for as long as possible, and this makes the ongoing running costs a major part of the overall cost of owning the storage. Therefore, not only should buyers look to the initial cost of acquisition, but also to ongoing maintenance charges, to the cost of power and cooling required for the storage assets, as well as the cost of human resources in managing the storage. In this area, the simpler the storage management is, the lower the costs will be.

- **Good power and thermal design**

Although many organisations are looking to demonstrate their green credentials through the optimisation of power utilisation, this is not the main reason why good power and thermal design is important within the mid-market. Equipment that runs hot tends to have a lower shorter life, with components failing faster than on cooler running equipment. Also, hot equipment needs external cooling, which is in itself a further cost. Similarly, high-power equipment costs more to run, but also needs more backup capacity in the case of electrical power source failure. The cost of uninterruptible power supplies, although having fallen over the past few years, is still a considerable cost, and ensuring that sufficient stand-by power is available to keep a business running through a power failure is becoming an issue for many mid-market organisations.

7. Storage Case Studies

As IP-based SANs have become available, some leading-edge mid-sized organisations and divisions of larger organisations have investigated how such SAN technology can help them, and they have implemented mission-critical solutions. Here, Quocirca looks at some organisations that have gained significant business value through the implementation of IP-based SAN technology.

Foster's EMEA

- **The problem**

Foster's EMEA, part of the global premium multi-beverage company, Foster's Group, needed to create a better environment for supporting its on-going sales activity across Europe. The main area of concern was within Foster's EMEA's importing, marketing and selling of its portfolio of wines.

Originally, Foster's EMEA had 30 servers running within its main European data centre. These servers hosted multiple customer applications, web servers, domain controllers, remote access and file-and-print servers. Further expansion was envisaged to give greater support for Foster's EMEA's mobile workforce.

Foster's EMEA has an increasing number of mobile and remote workers across Europe, and managing and maintaining laptops has proven to be problematic. With end users not making suitable backups of the laptop due to lack of available technology, any issues that impacted a laptop meant that as well as the loss of any data that was held on the device, a new device would need to be sent out to the user, causing further down-time issues and costs for that user.

- **The Solution**

Foster's EMEA decided to look at a fully virtualised environment utilising VMware for the servers and EqualLogic for storage. Virtualising the servers created a far more flexible environment for the mixed workloads found in the data centre, whereas the choice of several EqualLogic PS300E storage arrays provides a fully redundant, hot-swappable environment. Through the implementation of this storage environment, not only does Foster's EMEA gain a far more resilient storage system but also a better means of managing its remote laptop problem.

Each laptop now has an image of the laptop held on the storage array, which is updated automatically when the user touches the network. This provides a low-cost and highly effective disaster recovery solution for the mobile workforce.

- **The Benefits**

Foster's EMEA now have a far more flexible environment for supporting the business. The solution is easily managed and can be provisioned to support new workloads as required, while saving on power, cooling and physical space. As this was a test bed for a Foster's corporate strategy, the overall benefits to Foster's as a global organisation could be massive.

Schoolbank.nl

- **The Problem**

Schoolbank.nl is one of the Netherlands' largest social networking sites, allowing former classmates and friends to regain contact with each other. With more than three million registered users (nearly one in five of the total Dutch population), growth has been rapid and has stressed the original design of the underlying architecture.

The original design was kept as simple as possible – highly standardised, based on standard Wintel servers, with only three web servers, two database servers and a tape library. This platform was hosted by TelecityRedbus, a major co-location hosting company.

As with many social networking sites, the original idea of a textually based environment for friends to get back in touch soon changed into a site where people wanted to exchange more than just text. By July 2006, an increasing number of images were being loaded up onto the site, which led to an urgent need for the storage capacity of the site to be upgraded.

However, just upgrading the amount of direct attached storage (DAS) was not a viable alternative, as utilisation rates would still remain low and would only put off further painful decisions on what could be done about the need for further storage upgrades.

At an operational level, backing up all the extra image data was also becoming an issue. Increasing amounts of processor power had to be diverted to service the backups, rather than servicing the customers coming in through the web site. With the threat of a steep dive in customer response times, it was recognised that something had to be done rapidly.

Schoolbank.nl soon realised that the optimum solution would be a SAN – but initial investigations flagged Fibre Channel SANs as being too expensive for their capabilities and would also require skills that were not available within their workforce.

- **The Solution**

Further web-based research uncovered the possibilities of an iSCSI-based SAN. EqualLogic was chosen as the preferred supplier from a short list. Exhaustive testing of a PS100E array showed that this would provide a future-proofed solution that offered simple implementation and management, along with solving Schoolbank.nl's most pressing problems.

The EqualLogic solution meant that there was no intrusive impact on the existing network – connection was via a simple IP-based connection. This also meant that existing Ethernet-based skills could be utilised, rather than facing a need for retraining or skills acquisition. The chosen solution also means that extra storage capacity can be added as needed, without the need to take down the site. Performance was

also improved, and adding new storage assets gave immediate further improvements.

The PS100E's snapshot capability meant that storage backups could be carried out without any impact on the network itself, and any need for defining backup windows disappeared.

A major point for Schoolbank.nl was that implementation of the solution was so rapid. From the decision that something had to be done to full implementation and production running of the EqualLogic solution took three and half months.

- **The Benefits**

The EqualLogic solution not only solved Schoolbank.nl's immediate storage issues, but also provided a flexible platform for future growth at a reasonable cost with no requirement for specialist skills. The solution also dealt with the growing issue of how to carry out data backups.

Schoolbank.nl now wants to purchase an additional EqualLogic PS Series array to provide an offline copy of the data that TelecityRedbus holds, and also for development purposes through the use of the built-in virtualisation capabilities.

By freeing themselves from the constraints of DAS storage, Schoolbank.nl sees these initial implementations of EqualLogic storage as only the start.

8. Conclusions and Recommendations

Main Findings:

- The mid-market needs effective storage solutions just as much as large enterprises
- Simplicity is key – storage must be able to be managed within the constraints of the mid-market IT capabilities

As the need for information management grows, driven by governance and compliance, along with the requirement to gain greater understanding of the knowledge held within the vast amounts of data and information, the need for affordable, manageable storage hardware becomes an imperative.

Once only within the reach of the largest organisations, due to the cost and complexity of implementing and managing such solutions, storage area networks are now a realistic choice for the majority of mid-market organisations.

The immediate benefits gained through the use of suitable IP-based storage area networks are manifold:

- Ease of implementation and maintenance
- Flexibility of assigning storage to specific applications on an "as needs" basis – no more massive over provisioning at the commencement of implementation
- Business continuity capabilities through local and remote mirroring and backup/restore capabilities
- Incremental growth capabilities
- Investment protection through maximising the life of assets through planned aging

- Better visibility of information and data assets enabling better reporting and decision making

Planning for consolidation and rationalisation of existing storage is the precursor to implementing a single virtualized storage resource. In carrying out this process, a better understanding of the information assets that an organisation owns can be reached, and, from there, suitable policies for information storage, retention and overall lifecycle management can be created, which can then be easily implemented on a suitable storage architecture.

About EqualLogic

EqualLogic, Inc. is a leading provider of enterprise-class midrange storage area network (SAN) solutions that provide mid-size businesses and large enterprises with the best return on investment in the storage industry. The company's PS Series family of storage arrays offers an affordable and easy-to-manage alternative to traditional storage systems. Its unique dynamic virtual storage technology delivers simple setup, automated management, reliable data protection, and easy growth of storage capacity as needed. EqualLogic is pioneering a new model for businesses to buy storage, allowing incremental purchases, online movement of data, and complete storage management capabilities as a standard product offering.

Technology

Based on a vision of simplifying networked storage, EqualLogic's peer storage architecture combines advanced software features with fully-redundant, hot-swappable hardware integrated in a modular chassis design. This architecture delivers an affordable, enterprise-class storage system that is easy to install and manage and which is always available. By replacing cumbersome, labour-intensive administrative tasks with automatic, self-managing intelligence, PS Series storage arrays enable fast, flexible storage provisioning and dramatically reduce the time and costs required to manage and maintain a SAN environment.

Target Market

EqualLogic has enabled Fortune 100 companies and medium-size businesses in multiple industries – including application service providers, financial services, government, healthcare, higher education, legal services, manufacturing, and retail – to migrate from the limitations of direct-attached storage (DAS) to the simplicity, efficiency, and cost-savings of consolidated storage. EqualLogic arrays are a proven, primary storage solution for a broad range of departmental applications, including customer relationship management (CRM), enterprise resource planning (ERP), database, e-mail, file-and-print, web services, data warehousing, streaming media, geophysical analysis, and online transaction processing. The PS Series works seamlessly in mixed operating system (OS) environments that include Windows®, UNIX®, Linux, Solaris™, HP-UX, AIX®, NetWare®, VMware® ESX, and Mac OS® X.

Product

EqualLogic PS Series storage solutions deliver enterprise performance, scalability, and reliability – all within an intelligent, automated management framework that eliminates tedious administrative tasks while enabling effortless best-practice storage management. The PS Series is a family of self-managing storage arrays designed to fit a variety of SAN and network attached storage (NAS) environments and are completely interoperable and upgradeable without disruption of service.

EqualLogic PS Series storage arrays deliver true customer choice in storage by offering a variety of capacity, performance and price point options tuned to meet the needs of business critical applications.

About Quocirca

Quocirca is a primary research and analysis company specialising in the business impact of information technology and communications (ITC). With world-wide, native language reach, Quocirca provides in-depth insights into the views of buyers and influencers in large, mid-sized and small organisations. Its analyst team is made up of real-world practitioners with first hand experience of ITC delivery who continuously research and track the industry in the following key areas:

- Business process evolution and enablement
- Enterprise solutions and integration
- Business intelligence and reporting
- Communications, collaboration and mobility
- Infrastructure and IT systems management
- Systems security and end-point management
- Utility computing and delivery of IT as a service
- IT delivery channels and practices
- IT investment activity, behaviour and planning
- Public sector technology adoption and issues
- Integrated print management

Through researching perceptions, Quocirca uncovers the real hurdles to technology adoption – the personal and political aspects of an organisation’s environment and the pressures of the need for demonstrable business value in any implementation. This capability to uncover and report back on the end-user perceptions in the market enables Quocirca to advise on the realities of technology adoption, not the promises.

Quocirca research is always pragmatic, business orientated and conducted in the context of the bigger picture. ITC has the ability to transform businesses and the processes that drive them, but often fails to do so. Quocirca’s mission is to help organisations improve their success rate in process enablement through better levels of understanding and the adoption of the correct technologies at the correct time.

Quocirca has a pro-active primary research programme, regularly surveying users, purchasers and resellers of ITC products and services on emerging, evolving and maturing technologies. Over time, Quocirca has built a picture of long term investment trends, providing invaluable information for the whole of the ITC community.

Quocirca works with global and local providers of ITC products and services to help them deliver on the promise that ITC holds for business. Quocirca’s clients include Oracle, Microsoft, IBM, Dell, T-Mobile, Vodafone, EMC, Symantec and Cisco, along with other large and medium sized vendors, service providers and more specialist firms.

Sponsorship of specific studies by such organisations allows much of Quocirca’s research to be placed into the public domain at no cost. Quocirca’s reach is great – through a network of media partners, Quocirca publishes its research to a possible audience measured in the millions.

Quocirca’s independent culture and the real-world experience of Quocirca’s analysts ensure that our research and analysis is always objective, accurate, actionable and challenging.

Quocirca reports are freely available to everyone and may be requested via www.quocirca.com.

Contact:

Quocirca Ltd
Mountbatten House
Fairacres
Windsor
Berkshire
SL4 4LE
United Kingdom
Tel +44 1753 754 838

Appendix A: RAID and Disk Technologies

RAID comes in many different flavours, the main ones of interest for data performance and resilience being:

- RAID 0

Uses data striping, where information is written sequentially, generally across two matched disks. Although this increases storage response, there is no data redundancy built in and the failure of any disk within the array will render the entire storage array inoperable. However, all the capacity of the disks is available for utilisation.

- RAID 1

Uses data mirroring, usually across two matched disks. Every item written to one disk is written to the other disk, so that on the failure of one disk, the other can take over without any data loss. However, RAID 1 provides no performance benefits, and the overall performance will be the same (or slightly less) than that of a single disk. Also, only the capacity of a single disk is available.

- RAID 5

Uses data striping against a set of disks (minimum of three) and introduces distributed information parity. Here, each piece of information is written to more than one disk, so that we gain information redundancy and performance improvements. On the failure of any single disk, the other disks take over with little loss in performance. The failed disk can be replaced, and the array re-constitutes the information that was held on the failed disk. Failure of two disks at the same time will, however, still lead to all data on the array being lost. Capacity is generally $n-1$, where n is the number of disks in the array, so that a five-200Gb-disk RAID 5 array would provide 800Gb of storage capacity.

- RAID 6

Uses data striping against a minimum of four disks with dual distributed data parity, and provides an extension to RAID 5 capabilities. RAID 6 provides fault tolerance against the failure of up to two disks. As disk drive storage has increased, the time it takes for a failed disk to rebuild has also grown. With RAID 5, the whole array is at risk until the failed disk has been rebuilt. With RAID 6, the data is still guaranteed, even if a drive fails during the failed disk rebuild. Capacity is generally $n-2$, where n is the number of disks in the array, so that a five 200Gb-disk RAID 6 array would provide 600Gb of storage capacity.

- RAID 10

A combination of RAID 1 and RAID 0, this provides both striping and mirroring, so providing a degree of resilience and performance improvement. Available storage capacity will be half the overall physical capacity.

- RAID 50

A combination of RAID 5 and RAID 0, providing very high levels of resilience as well as good levels of support for services requiring high I/O and data transfer rates.

Other RAID levels (e.g. RAID2, 3, 4) exist, but are rarely used outside of highly specific environments.

Disk Types

Historically, disk drives came with one of two main interfaces – either the Integrated Drive Electronics (IDE) connector or the Small Computer System Interface (SCSI). IDE drives are also known as Parallel Advanced Technology Attachment (PATA) drives.

SCSI drives tended to be used in intensive, mission critical areas, due to longer life times and higher performance. The SCSI interface also supported a larger command set, meaning that more information on the performance and setup of the disk drive could be exchanged between the device itself and any management software.

Lately, a new interface has come to the fore, through the use of the Serial Advanced Technology Attachment (SATA) standard. These drives provide far faster performance than the older PATA drives. The current version for SATA is SATA II, which has improved data throughput considerably, and work is ongoing for further improvements.

Also, SCSI drives have moved forwards to support the Serial Attached SCSI (SAS) standard. SAS drives are still aimed at the high-end performance market, being generally of higher speed than SATA drive, but of lower capacity. The main advantage of a SAS controller is that it can also support SATA drives. Therefore, a storage subsystem could start off with lower-priced SATA drives and could be migrated to SAS as required, or could hold a mixed environment of physical tier 1 and tier 2 storage which could be virtualised as needed.

Appendix B: Storage asset growth, reuse and tiering

Although storage costs have fallen continuously and rapidly over the past few years, data growth has meant that the overall cost of acquisition of required storage assets has not fallen but, in fact, continues to grow. When combined with the increasing costs of managing existing storage systems, it is increasingly important to successfully sweat existing storage investments wherever possible.

For those organisations finding that existing storage space is running out, the purchase of more storage has historically been a choice between adding extra incremental storage and manually managing how this interacts with existing storage space or buying a storage asset that is big enough to hold the existing data plus enough free space to cover expected growth for a period of time. The first approach can lead to slow response and even catastrophic data loss if managed incorrectly; whereas the second approach leads to heavy application downtime while the existing data is backed up, the new storage is introduced and the data is then restored. Also, as storage needs have proven difficult to predict in the past, this replacement approach generally means that this has to be done on an increasingly frequent basis.

What mid-sized organisations really need is the capability to add storage assets in an incremental basis that does not involve application down time, while maximising the return on their existing storage investments. By taking a virtualised approach to storage, organisations can ensure that the physical and logical storage assets remain separate – and this creates a platform that opens up new opportunities for how information can be managed.

For example, let us consider a storage environment within a mid-sized organisation. The organisation takes the decision to create a modern, virtualised storage infrastructure. It puts in place new high-speed disk systems and creates logical partitions against this resource pool such that different applications have access to different logical drives. However, as the storage system is a single logical pool, it can be simply managed through a single management interface.

As time progresses, the organisation will find that it requires more storage. Since the storage pool is virtualised, it is a simple matter of adding physical storage units to the pool and then growing the logical partitions as required to create the needed headroom for the individual applications.

At some stage, faster storage systems will become available. If the organisation had not taken a virtualised approach, it would have had to rip out existing storage resources and replace them with the new to gain the benefits of the higher speeds. However, through the use of a virtualised storage environment, the new high-speed disk can be introduced easily alongside the existing slower storage assets, and existing data can be transferred while the applications are accessing the new disks.

This then leaves us with a set of slower disk assets. The option may be to retire these, but the organisation can decide that such disks can be used for applications or storage where speed is not so much of an issue. Through the use of virtualisation, this is a valid and relatively easy approach and is known as storage “tiering”.

The single virtualised pool of mixed speed storage assets can be subdivided up into areas as required through standard partitioning, with specific applications being pointed to specific virtual storage units. These virtual storage units can be based on the real properties of the underlying physical assets. For example, all high-speed disks can be virtualised as a single resource pool and allocated as “tier 1” storage against the needs of primary applications; and slower disks can be virtualised and declared as “tier 2” storage, where less time-dependent applications can store information. Further, the archiving tier can be written off to physical tape (or to optical disk, an option increasingly being seen as cost effective within the mid-sized organisation) as required for long-term storage.