Data Protection & Utility Computing

Protecting Data in a Flexible Infrastructure
Important Notice

This report contains data and information up-to-date and correct to the best of our knowledge at the time of preparation. The data and information comes from a variety of sources outside our direct control, therefore Butler Direct Limited cannot give any guarantees relating to the content of this report. Ultimate responsibility for all interpretations of, and use of, data, information and commentary in this report remains with you. Butler Direct Limited will not be liable for any interpretations or decisions made by you.
EXECUTIVE SUMMARY

The history of the IT industry is littered with failed initiatives and models that failed to reach their full potential. Upon examining the reasons for this, one item stands out; solution providers and software vendors have always been happier to talk about the possible benefits and less comfortable when some of the practicalities were raised.

Utility computing is the latest model to excite the marketing machine of all the major vendors and, true to form, we are inundated with the benefits and less than informed when some of the possible hurdles that need to be overcome are mentioned.

VERITAS, long recognised as a major player within the data storage and management space, understands the dichotomy that exists within utility computing as it relates to Enterprise Data Protection (EDP) and is attempting to provide software that will address this dichotomy. VERITAS’ aim is to build upon its storage management tools and expertise to address other issues in the whole data storage/management space and beyond, as detailed in the Future Strategy Section of this White Paper. At first glance it might appear that EDP stretches beyond the bounds of storage management, and this is certainly the case. However, when one comes to examine EDP more closely, especially the requirements of new and upcoming legislation, it becomes obvious that the only place to implement such a solution is at the lowest level. What is required is an integrated solution that marries data storage into an overarching EDP implementation.

The reason for this is that the new legislations, some of which are discussed within this White Paper, will force a dramatic increase in data growth that has to be used for proving the control of such data. If we should attempt to place these as two separate items within the infrastructure, we run the risk of doing little more than increasing the burden of administrative cost.

What is really required is a way of siphoning off the administration cost of data management and re-applying it to the EDP element. This is precisely what VERITAS is aiming to achieve. Continuing its pragmatic approach to EDP and utility computing, VERITAS would be the first to admit that not every part is yet in place. What is clear, when one reads the section within this White Paper that covers the VERITAS solution and the roadmap for the future, is that at least VERITAS understands the problem.

It is, of course, far easier to talk about utility computing and EDP as if they were two synergistic elements rather than dichotomous ones. It is with some certainty that we can state that there are vendors who will be quite happy to do this. VERITAS, on the other hand, has accepted that problems will occur if only the synergistic view is taken, and this has become a major driver within its strategy.

The road to true utility computing will not be an easy one. It will fail, as so many other models have failed in the past, if we rely on the solution providers who are only prepared to talk of the benefits and who will not recognise the problems that will be met along the way. VERITAS does not have all the answers for utility computing; what it does have, however, is the ability to help users create a utility computing model by building on existing infrastructure.

Furthermore, the products provided will also give users the flexibility to respond to issues that impact storage management, but which also extend beyond a narrow view of storage management. One such element, EDP, is the case in point for this White Paper.
The idea of utility computing is one exciting much interest in the market place, and with obvious reason. The ability to utilise resources, of any description, in a highly dynamic manner makes sound economic sense. At least it is possible to make a strong economic case for the utility computing model, which is slightly different. The difference, as it exists, is one of degree and range. Utility computing, as with many other proposed models, considers the micro as opposed to the macro effects, costs, and benefits.

In no area is this requirement for a clearer understanding of the macro view more easily explained than when one comes to consider utility computing as it impacts Enterprise Data Protection (EDP). Before we come to examine this argument in more detail, we need to define the term ‘utility computing’ in order to create the context for the further discussions.

VERITAS describes utility computing as:

…the IT utility provides a fast, reliable service to end users, and can adapt to changing business needs without human intervention. At the same time, the IT utility makes use of shared hardware resources, which maximizes utilization and reduces overall costs.

Whilst not the only definition available, it does not differ in any great degree from those issued by other major players within the utility computing space, or those promoting utility computing solutions, and therefore stands as a usable and viable definition.

It fails to describe, however, the impact that such a model might bring to a wider range of issues, and corporate positioning. In fairness, it must be noted that this is not a failing by VERITAS, or an attempt to disguise the full impact of utility computing, it is simply a demonstration of how limiting a few phrases can be when describing any complex subject.

Utility computing is being promoted under a variety of names; perhaps the most highly visible being IBM’s ‘On Demand’ strategy. The importance of utility computing can be seen in the very fact that one of the world’s largest organisations is moving its whole strategy to this model.

One of the issues about utility computing that has to come under serious consideration concerns heterogeneity. Utility computing solutions have to address the fact that, as we are talking about creating a usable whole from a disparity of resources, they must have complete independence from the underlying infrastructure. Utility computing cannot work if platform independence is not built into the solution.

Apart from taking a strong position with utility computing, VERITAS is also a leading proponent of EDP, and a provider of solutions in this area, where its platform independence answers this requirement of utility computing. Although the implementation of a utility computing model and the requirements for EDP are not completely diametrically opposed, there are enough points of conflict to raise concerns that must be answered.

Utility computing is about availability of infrastructure; switching resource in a dynamic manner. Data does not fit easily into this model. Apart from EDP (which can be viewed as mainly led by legislative forces), there are many other requirements for data availability (such as data positioning for ease of re-use), which also have to come under consideration. Flexibility of data management/use cannot occur on an inflexible infrastructure, but matching flexibility at one layer with flexibility at another raises its own problems, and these need to be addressed if utility computing is to become a viable model.
ENTERPRISE DATA PROTECTION

As with utility computing, it is important that the terms of reference are clearly stated before undertaking any further discussion. EDP can be distilled down to one statement:

_An EDP solution or implementation has to exist or be carried out with the one basic aim of ensuring compliance with any and all Governmental rules and/or legislative requirements for the storage, maintenance, and utilisation of data with due regard to the differences that exist with such Governmental rules and/or legislation across geographical boundaries._

A recent problem that has arisen within the airline industry illustrates the point most effectively. Legislation within the USA requires airlines to provide comprehensive data on each individual passenger flying to the USA. However, the Data Protection (DP) laws that exist within the USA allow that data to be used for purposes other than those stated as the primary purpose (in this instance, the combating of terrorism). This comes into direct conflict with DP laws that exist within the European Union (EU). As one example, and to highlight the depth to which this conflict goes, one of the requirements is for data on individual's meal requirements. Ignoring the argument as to whether or not this piece of data could be used to combat terrorism (points for and against could be taken on both sides), it is the further use to which this data could be put that is the stumbling block.

There is a fundamental difference between DP laws in the USA and the EU, which make it perfectly acceptable to utilise this data for multiple or extended purpose in the USA and at the same time illegal in the EU. Therefore the airlines are caught between dual compliance requirements.

Although VERITAS does not have an answer to this problem (resolution of conflicts between Governments is beyond the scope of any IT software vendor) it does demonstrate two points. Firstly, any EDP solution has to address the issue of multiple levels of compliance. Secondly, there is an essential dichotomy between utility computing, with its promise of utilisation of shared resources; and EDP, which has, by its very nature, to take a more tight controlling position.

Not all EDP issues are as complex as the one highlighted, but even more mundane ones, such as ensuring compliance with the storage/availability of data to comply with current and changing regulations has an impact on every organisation. Flexibility of storage management infrastructure is the only way in which these issues can be handled. If a change in legislation, for example an increase in the time that data has to be stored on-line for instant retrieval, comes about, then no organisation would want to bear the additional cost of having to change its whole storage infrastructure.

In implementation terms, and with regard to compliance, the primary requirement is to have a solution that ensures the correct controlling procedures, based upon codified policy (most likely abstracted from specific legislative imperatives), are in place and are followed. Simplistically, the best defence for any possible infringement is to demonstrate that the utmost care was taken, allied to full auditability of data existence and utilisation – data lifecycle management.

VERITAS cannot supply all the answers to all the problems, but it can supply software that helps bridge the gap that exists between the required flexibility of the control layer and flexibility of the infrastructure.
JOINING EDP & UTILITY COMPUTING

Although we have demonstrated that conflict points exist between implementing EDP and utility computing, we also stated that the two are not completely diametrically opposed. We now need to consider how the opposition between the two forces can be managed.

Firstly, we need to ensure the implementation of a system that can automate and manage the process of data selection for storage based on a DP policy. Automation has two key benefits. Firstly, it fits well with the ability to provide proof of compliance. If the selection process is automated (against policy) then once the policy is approved, the process itself gains automatic approval. This would not necessarily be the case if we relied upon human interpretation on an ongoing basis.

Secondly, automation decreases the administration cost, without any increase in risk of non-compliance. This reduction in cost can form the basis for redirection or re-utilisation of resource to other areas. Specifically, in this instance, we would consider re-allocation of resource away from administration and towards a higher level of EDP.

Redirection of resource away from simple management that can be automated to a dedicated EDP management policy will have no negative budgetary impact, and will have a positive impact in terms of compliance and stakeholder perception of organisational excellence.

There are four possible alternatives that can be undertaken, and which can be codified quite simply:

• Remain with the administration overhead as it exists and keep the spend on EDP at its current low level. This will mean almost certain non-compliance penalties for insufficient EDP, but keep spend at its current level – this equates to one negative item, and one of equality.

• Keep the administration as it exists and increase the spend on EDP. This will reduce the risk of non-compliance penalties, but increase the spend on administration/management – this equates to one positive item, and one negative item.

• Move towards an automated solution in areas where automation can be implemented and keep the current spend on EDP. This will reduce the overall cost, but increase the risk of non-compliance – providing one positive item, and one negative.

• Move towards an automated solution in areas where automation can be implemented and increase the current spend on EDP. This will reduce the overall cost (perhaps not as much as for the previous option), and reduce the risk of non-compliance – this provides two positive elements.

Some additional weighting has to be given to this simplified codification. Administration cost has a correlation with the amount of data stored, the type of data required, and the longevity of that data. Compliance requirements have a three-pronged knock-on effect:

• The amount of data stored will need to be increased. Much of the choice of which data is relevant and which is not will be taken away by compliance requirements.

• In a similar manner, much of the choice regarding the life of data will be removed or be subsumed by the EDP policy.

• Finally, there will be an increasing need to include new file types within the EDP policy. There will be a blurring of the distinction that currently exists between data and information. As one example, e-mails that are currently discarded may now require storage for compliance purposes.
There is no linear correlation between data volume and/or data life, and administration cost. However a correlation does exist, and without automation of some parts then administration will increase, even if there is no formal recognition or budgetary spend on EDP. Effectively, this negates the first two items on the codified structure; keeping administrative spend at the current level is a fast-disappearing option.

In reality this only leaves the last two options. For those who do not fear the outcomes of non-compliance, then the obvious choice is for reducing administration cost by automation and not spending the saved cost on higher levels of EDP. For all others apart from that sad and soon to be sadder individual we are really only left with the final option.

This only leaves the issue of the level of cost shift/saving to be placed against administration/EDP. Whilst this will depend on a number of organisational-specific factors, and therefore impossible to quantify. However, there is one aspect that can be seen to reduce any possible total cost base, and that is by having a synergic system that covers EDP and data storage and management.

If we can link EDP and data storage and management into an integrated whole, we also open up the possibility of creating an infrastructure that goes beyond budgetary balancing and moving towards providing service-based functionality around EDP; a single functional piece of utility computing.

## DATA PROTECTION AND DISASTER RECOVERY

An EDP strategy is not just concerned with managing data from a data-instance viewpoint; it also has to incorporate a higher level of EDP, which revolves around disaster recovery. This can be considered as EDP at the macro level.

One factor that is often overlooked, but which is handled by the VERITAS solution is the amount of data that exists on the desktop/laptop. Current estimates show that around 55% of corporate data exists on these devices. Within that 55% there will be a certain proportion that is replicated to the data centre, but this will still typically leave in excess of 40% for which there is little or no control.

Not only is this plainly unacceptable, it brings into disrepute any disaster recovery strategy. There is an inherent relationship and interdependence of corporate data that is not always explicitly expressed within informational models. If the disaster recovery solution cannot take account of this then the implicit relationships between data elements will be lost; degrading the data cleanliness.

Again, this can be brought back to the requirement to have a fully integrated system of data management, administration, and data protection at both the micro and macro level. Any disjoin between the implemented solutions will inevitably lead to gaps within the whole EDP infrastructure.

## DATA PROTECTION AS UTILITY

The conflict points that exist with utility computing and EDP can be reduced or minimised by implementing the correcting overarching solution. This also creates an infrastructure that can lead to EDP as a service offering.
Outsourcing EDP has the usual benefits associated with this model; economies of scale, reduction in administration cost, etc. However, outsourcing has associated risks attached. With other parts of the operation, the risk/benefit equation can be heavily weighted towards the benefit side. This is not the case with EDP. Risk mitigation with taking EDP as a service or utility has to receive the highest weighting. It has to be accepted that ultimate responsibility lies with the service consumer and not the provider.

This almost demands that the service provider has to take charge of the whole data storage infrastructure in order to apply meaningful and workable EDP policies. Additionally, the provider has to demonstrate that this infrastructure is one that can handle all the requirements detailed previously in this White Paper.

In the final section of this White Paper we consider the solutions that VERITAS has to offer in this space and, equally importantly, the roadmap that shows the progression across other areas of IT, to bring to fruition the total infrastructure requirements, and the control processes that will make utility computing a viable option for corporate use.

**VERITAS SOLUTIONS**

An integrated approach is required in order to deliver a comprehensive EDP solution, and this is what VERITAS has successfully achieved. Three products provide the foundation of this solution:

- NetBackup Enterprise Server 5.0.
- Data Lifecycle Manager 5.0.
- CommandCentral Service 3.5.

All three are suited to a more service-oriented model, enabling the IT department to start to move towards becoming a value centre, and this is the first stage of VERITAS' utility computing model.

**NetBackup Enterprise Server 5.0**

One of the highlights of NetBackup Enterprise Server 5.0 is the inclusion of a Desktop and Laptop Option (DLO), which is an evolution of the NetBackup Professional product. Desktop PCs and laptops are generally inadequately protected, being left out of normal back-up routines. Many organisations try to get around this problem by introducing company policies whereby all corporate content has to be stored on the central servers, which is difficult to impose, particularly with a mobile workforce. If IT is to offer its resources as a service, then it should offer protection for all devices and not just servers. The desktop and Laptop Option provides the mechanism for including PCs and laptops in the back-up policy, by enabling the background synchronisation of data, with minimal impact on the end-user. This can be initiated in different ways. Either the user can trigger the process, or it can be automatically started when the user connects to the network.

DLO provides the ability for end-users to initiate the restore process without IT intervention, introducing an element of self-service into the protection of PCs and laptops. More importantly for business managers purchasing back-up resources from the IT department, it reduces the amount of back-up resource required. The option does not require an additional back-up server, nor is any additional software on the file server needed.

A pre-requisite of providing data protection as a service is the ability to offer different levels of protection according to the requirements of the customer. This needs to be at a granular level, allowing different types of data to be backed-up, with a choice of methods and frequencies. NetBackup 5.0 supports this through the ability to create Synthetic Back-ups. This is an innovative method of creating full back-ups from incremental back-ups, enabling IT to offer higher service levels at a reduced cost.
One of the problems with incremental back-ups is the length of time that it takes to restore data, as the last full back-up and all of the subsequent incremental tapes are required. The Synthetic back-up facility enables organisations to take much more frequent incremental back-ups, which are written to disk. A full synthetic back-up is then created by taking the last full back-up plus the incrementals, which is then written to tape.

Disk-to-disk copies can be created on-the-fly, which allows multiple copies to be made using one read and many writes. This pushes VERITAS into the area of compliance, as unalterable copies of data may be a requirement of regulations. Disk staging is also supported, which allows better tape utilisation to be achieved, which in turn results in faster recovery.

The Advanced Client provides a number of easy-to-use snapshot technologies, which are Wizard-driven and enable VERITAS to claim that only five clicks are required to configure a snapshot. A powerful feature is the ability to integrate different environments into a single view. The user is presented with a recommendation of the snapshot technique most suited to his or her environment, which can either be accepted or another method selected. This is another powerful enabler for utility computing, where the back-up method may be determined by the criticality of the data.

**Data Lifecycle Manager 5.0**

Data Lifecycle Manager 5.0 (DLM), formerly known as NetBackup Storage Migrator for Windows, addresses the requirements of compliance through the management of the complete lifecycle of data from creation to disposal. It is policy-based with data archive, retention, and retrieval functions. Data is managed and migrated across all types of media through policies defined according to a file’s age, size, retention requirements, and value to the organisation. The automation of the migration process helps to alleviate some of the pressures placed on IT departments required to manage ever higher volumes of data.

An important requirement for organisations with vast quantities of existing data, that is coming under the auspices of regulations, is the ability to archive that data, which is nigh on impossible if it has to be sorted manually. DLM includes the ability to build indexed archives from historical data held on back-up tapes, aided by tight integration with NetBackup and Backup Exec. Virtual data archives are created across on-line, near line, and off-line media, providing organisations with a variety of options. Archiving to non-alterable media such as WORM and optical devices is also supported; this is a key requirement of compliance. A further necessity for compliant data is the maintenance of an audit trail of all changes made to the archive.

High-speed search capabilities have been built into DLM to speed up the retrieval of data from back-ups and archives. These are based on mathematical algorithms rather than pre-configured lists of keywords.

The migration capabilities provide a powerful tool for IT departments. By automating data management with DLM, it is possible to automatically migrate little used data to more cost-effective media, delaying the need to acquire additional storage resource. A pointer to the data’s new location is left in its original location. When subsequent back-ups or restores are undertaken only the pointer is copied reducing the time taken to undertake these tasks. This has the added benefit of reducing the cost of the data’s storage and management to the line of business that owns it.

A cross-platform authentication engine provides the ability to set privileges for the management and migration of data for individuals and groups based on roles. Policies implemented can be automatically pushed out to laptops and desktop PCs, bringing data stored on these devices under the control of the enterprise.

A benefit for organisations is the fact that NAS, DAS, and SAN resources and licences are shared with NetBackup Media Manager and Backup Exec ADAMM.
CommandCentral Service 3.5

CommandCentral service is core to VERITAS' utility computing strategy, as it provides a portal to centrally manage IT services, and also reports and quantifies the results of expenditure for back-up and recovery and storage utilisation. It sits in the upper tier of VERITAS' management portfolio, leveraging the capabilities of NetBackup, Backup Exec, SANPoint Control, and Storage Reporter to provide physical, logical, and business management.

An important element of offering any IT function as a service is the ability to demonstrate that service levels are being met; and data protection is no exception. CommandCentral Service defines, measures, and monitors service levels across the storage hierarchy. Moving to a service oriented model is likely to meet opposition from business managers who suddenly find that IT resources must come from their own budgets. By tying this to service levels whereby managers are guaranteed the level of service appropriate to their line of business, which often means improvements to the old service, the pain of moving to this new model can be eased somewhat.

However, it is not sufficient for IT to claim that service levels are being met; business managers must be able to see this for themselves. The ability to report on usage and consumption in CommandCentral Service enables business managers to view how well the IT department is performing in delivering agreed service levels. It also lets them see how quickly they are consuming resources. The Web-based portal can be personalised to deliver information relevant to each user's requirements, with access strictly controlled.

When integrated with NetBackup and Backup Exec, CommandCentral Service provides resource discovery and media information. Reports highlight errors such as failed back-ups, enabling IT to react quickly to any problems.

VERITAS already has a chargeback mechanism in place that allows IT to charge for its services, and CommandCentral Service provides the cost value analysis that is required for chargeback to business lines. This information provides managers with the flexibility to select the lowest cost services that meet their requirements. VERITAS already has the chargeback mechanism in place to deliver data protection as a service.

CommandCentral Service offers a complete solution for the allocation of services to business lines, down to the incorporation of customisable workflow for the service request and approval processes. By adding workflow VERITAS has taken some of the pain out of what can be a long-winded process to get approval for expenditure.

The reporting capabilities provide business level information, which helps align IT resources with business priorities, and by examining historical data trends can be correlated between performance over time and cost and resource allocation metrics.

At present, CommandCentral Service is limited to providing storage resource allocation, but this will be extended to provide server provisioning, availability, and application performance in future releases.

Future Strategy

The three products highlighted above comprise the first stage of the VERITAS strategy to enable utility computing through software. This strategy is enabling customers of VERITAS and other vendors to take IT services such as Storage Management, Data Protection, Application Performance Management, Availability, and Server Provisioning and deploy these against defined and managed service levels, which can then be mapped to line of business requirements. The company still has a long way to go before it can claim to provide a full utility computing service, and provide the mechanism for the IT department to complete the transformation from cost centre to value centre.
However, it has successfully delivered the first stage of its strategy though the ability to monitor, deliver, provision, and chargeback Data Protection as a service. In 2004 VERITAS will expand this methodology to provide strong service level workflow and ticketing system integration. The company will also provide the ability to deploy high availability, application performance, and server provisioning as a service, through the integration of not just its own technologies, but also other leading software solutions.

**CONCLUSION**

It has been the intention of this White Paper to bring to the attention of the readers the possibilities that exist with utility computing, but to do so with reference to a conflicting aspect; how the dichotomy between the two can be reduced, or even eliminated.

It is our view that utility computing is a model that will receive much attention in the coming months and years, and which will be implemented in a large number of organisations. However, it is always easier to extol the benefits of a new model and gloss over any possible problems.

VERITAS does not have all the answers, but its products do allow both a structured approach to implementing a utility computing model and it does this with reference to the possible points of conflict within this model.

**CONTACT DETAILS**

VERITAS Software Corporation  
350 Ellis Street  
Mountain View  
CA 94043  
US  
Tel: +1 (0) 650 527 8000  
Fax: +1 (0) 650 527 2908  
www.veritas.com

VERITAS Software Ltd.  
350 Brook Drive  
Green Park, Reading  
RG2 6UH  
UK  
Tel: +44 (0)870 243 1080  
Fax: +44 (0)870 243 1081