WHITE PAPER

CDW PEOPLE WHO GET IT™

REMOTE BACKUP IN THE CLOUD

For small businesses, remote backup offers a taste of cloud computing and peace of mind that critical data is protected.

Executive Summary

Cloud computing is the most dynamic technology area in IT today. While many executives see value in the technology, some don't know where to start assessing its potential. Cloud-based remote-backup solutions lend themselves well to such situations because they offer many of the benefits of hosted solutions while being comparatively easy to provision.

Through cloud-based remote backup, small businesses especially have a golden opportunity to leverage the scale of the cloud to build a resilient disaster recover solution without having to invest in hardware, complex software licensing or new, skilled personnel. Cloud-based backup offers an intrinsic benefit over onsite alternatives in that the physical storage is generally far removed from users and their source data.

Table of Contents

TWEET THIS!

As such, it is not likely to be affected by the same natural disasters and other localized outages that could hit a data center. Similarly, independent technologies and operational procedures provide additional redundancy, making backup data less susceptible to the type of human error that might leave primary data vulnerable.

That being said, the landscape of remote-backup solutions is diverse. For example, EMC's Mozy and Symantec Backup Exec.cloud offer familiar, personal backup features but target cloud storage. Other offerings leverage online backup and storage to help synchronize multiple user devices. Still others, such as Microsoft's Box for Office, Evernote and SkyDrive, double as collaboration tools.

Choosing a remote-backup solution requires an understanding of the technology's benefits and an analysis of an organization's specific requirements. Considerations include feature sets, service-level requirements, encryption and bandwidth. Some businesses also must factor in their need to comply with data protection rules or regulations. From there, getting started in cloud-based remote backup is a matter of differentiating between offerings and identifying the best solution to meet the organization's needs.

The Benefits of Remote Backup

Cloud-based remote backup offers benefits in four main areas: cost, risk, flexibility and quality. By backing up data in the cloud, businesses can reduce and simplify their cost structures while offloading some of their risk to a service provider.

They can leverage the elasticity of cloud computing to become more agile in responding to changing circumstances. And they can capitalize on outsourcing some of their IT responsibilities by redirecting precious resources to activities with more direct effect on the business, relying on the provider's core competency to achieve better backup services.

Cost: The most obvious benefits of cloud computing have to do with cost. Cloud-based remote backup can mean a significant reduction in upfront capital expenditures because there is no need to purchase extensive hardware infrastructure or software licenses. Instead, the organization can align its costs with actual usage.

Moreover, the IT team doesn't have to overprovision backup storage capacity to accommodate uncertain growth. And finally, remote backup can be configured as an automatic process both transparent to the user and requiring less IT administration and support.

Risk: Remote backup can offload risk from the customer to the service provider. By contractually stipulating provisions for data protection and disaster recovery that are tied to specific indemnities in the event of service failures, a company can further mitigate risk.

Remote backup also reduces the likelihood of under-provisioning data storage. During daily IT operations, data usage patterns can't always be predicted. There is the possibility of unanticipated spikes in storage requirements — some beyond the capacity physically available. Cloud providers offer elastic resource allocation and unlimited data retention, making it less likely that storage capacity will ever be exhausted.

Also related to risk is data security. In cloud computing, security is often considered a challenge because the organization is entrusting some of its IT systems or data to a third party. Nevertheless, there are several benefits that cloud computing may offer with respect to security.

Cloud providers typically undergo very strict security audits. In addition, precisely because they are entrusted with clients' information, providers typically employ best-available security procedures and solutions, and they keep those solutions up to date.

Furthermore, when using cloud for backup, it can be easier for a business to isolate customer and employee data if they are managed in different environments, mitigating the risk of inadvertently mixing the two and potentially revealing one or the other. To increase overall data security, it might make sense to segregate data — housing customer information in the cloud and staff information on an internal network.

Flexibility: A cloud infrastructure adds considerable flexibility and agility to enterprise architecture. As cited earlier, the scalable, elastic storage capacity afforded by a cloud service means that rapid data growth is easily accommodated. At the same time, the metering capabilities of a cloud service ensure that an organization pays only for the storage it uses.

Flexibility also comes in the form of speed of execution. For example, companies can roll out cloud-based software as a service (SaaS) with little preparation or planning time. When it comes to remote backup, using a cloud means data backups can happen continuously, which supports superior recovery point objectives (RPOs).

Similarly, data restoration can take place on demand, supporting excellent recovery time objectives (RTOs). Both RPO and RTO are important components of an organization's disaster recovery planning.

Finally, a globally replicated cloud infrastructure can facilitate data access from anyplace using any device at any time, which contributes to greater user flexibility and productivity. This cloud advantage is even more important when a company needs to integrate its business processes with those of its suppliers, partners and customers.

Quality: Quality of service (QoS) is clearly a major concern. Cloud service providers offer great economies of scale and specialization. They have developed rigorous processes and procedures to maximize uptime and optimize performance.

They run best-available software to monitor and manage the infrastructure, and they employ some of the most skilled practitioners to oversee management tools.

An on-demand, cloud-based service model can also be better than on-premises purchased and installed software because the service provider is able to distribute new functionality transparently without IT department intervention. As a result, users often receive more frequent and more timely updates as well as new functionality without the company's own IT staff having to test, schedule and provision each rollout.

Backup Needs: Which Cloud Option Is Best?

To assess its backup requirements and arrive at a viable solution, an organization must analyze its internal information landscape and make fundamental choices about what type of cloud service will accommodate its needs. First, it must determine how much storage it requires, then how best to procure that space in a cloud.

Data sizing: The most basic question to answer is how much data the company needs to back up. Start with how much data is currently stored, taking into account storage systems and devices as well as local data on servers, desktops, notebooks and even mobile devices.

There isn't a one-to-one proportional relationship between actual storage and required backup capacity. This fact introduces additional questions: How often does the data change? How often does it need to be backed up? How many versions must be held? How should backup data be retained?

It is risky to base a backup solution on today's requirements alone. Organizations typically commit to using backup solutions for three to five years, so it is important to forecast data requirements carefully. One standard benchmark estimates that the amount of data an organization stores, on average, triples every three years. But if a company has a higher retention level (the length of time it keeps data) or dynamic data, then its growth rate may be significantly higher.

Most remote-backup providers price their services in terms of gigabytes and offer them at different tier levels. Considering data growth, it may be advisable to overestimate usage to get the best price for the long term.

Cloud delivery mode: Cloud computing is often categorized by three models — infrastructure as a service (IaaS), platform as a service (PaaS) and SaaS. PaaS targets primarily application developers and software vendors and is of limited interest for data backup. IaaS and SaaS, on the other hand, can both be used for backup solutions, albeit with very different approaches.

IaaS is the most flexible option. It can accommodate almost any application that can run on a physical computer. In fact, cloud-based block storage can be used by legacy applications, including backup programs, with relative ease. IaaS leverages few of the benefits of an Internet-based, virtualized, utility-priced delivery model.

10 Questions to Ask a Potential Cloud Vendor

- Can the provider offer references comparable to your organization?
- Are strong user identification and encryption in place?
- Where is data hosted? Consider geographic risks and laws governing privacy rights.
- Who has access to your data?
- If you have users or customers with regulatory requirements, is the provider compliant?
- What's the latency of accessing data?
- What's the backup-and-restore process in case of a disaster?
- Does the provider offer a guaranteed service-level agreement?
- What equipment does the provider use?
- What other organizations are in this cloud? Is your data totally segmented from theirs? Remote backup in the cloud need not be an all-or-nothing proposition.

In general, if there is a standard-offering SaaS solution available that meets an organization's requirements, is priced appropriately and doesn't pose significant security concerns, it is usually the most compelling choice among remote-backup options. It's important to understand, however, that there also may be off-the-shelf solutions an organization can purchase or license to run on a cloud platform or infrastructure. SaaS doesn't always equal cloud and vice versa.

The key question regarding a SaaS offering is whether it will require significant customization. The more necessary a custom backup solution is for an organization, the less attractive SaaS may be. In that case, companies would want to consider an off-the-shelf commercial offering that either runs on a cloud platform or can be virtualized and run on cloud infrastructure (see *Storage Virtualization: A First Step* sidebar).

An environment that is owned and controlled by the consumer allows greater flexibility for customization. But keep in mind, doing so is likely to result in additional cost.

What is unique about data backup is that it entails both infrastructure (where data is stored) and software (the interface for data backup). Some solutions run both components in the cloud, while others do not (see *A Hybrid Option* sidebar).

Cloud delivery source: In addition to considering the cloud delivery mode, organizations need to consider the delivery source. In its earliest definition, cloud computing referred to solutions in which resources were dynamically provisioned over the Internet from an offsite, third-party provider who shared resources and billed on a fine-grained, utility-computing basis. Known today as the public cloud model, this approach offers many advantages in terms of cost and flexibility, but it has governance and security drawbacks.

Many companies have looked at ways to leverage some of the benefits of cloud computing while minimizing the drawbacks. Their efforts have led to a more restrictive private cloud model. Typically, a private cloud is hosted on-premises, scales "only" into the hundreds (or perhaps thousands) of nodes, and runs over private network links instead of the public Internet.

Along those same lines, a community cloud caters to a group of organizations with a common set of requirements or objectives. The most prominent examples are government clouds open to federal and municipal agencies. Similarly, major industries may have incentive to work together to leverage common resources, including remote-backup storage.

The categorization of cloud services into public, private and community clouds is a simplification. Not only is there no clear boundary between the three delivery models, but customers also aren't likely to confine themselves to any one approach. Instead, companies can expect to see a variety of hybrid constellations and consider tailoring their backup solutions to any combination of them.

Features of Backup Solutions

Once a company has chosen its basic approach to remote backup, it can begin looking closely at the different solutions on the market. Choices offer widely diverse features, which can make comparing them difficult. But diversity also means a slew of available options, which makes it easier for an organization to find a service that meets all its unique requirements.

When considering features, the areas that organizations should focus on evaluating include user experience, security, archiving and network load.

User experience: When a company considers a new way of doing things, it should be sensitive to user acceptance. One way of maximizing acceptance is to focus on the backup features that users will see, such as the software interface that runs on the client system.

Although there are some backup services that extract data over the network or directly from storage, the majority of popular solutions use a local agent. Some run on Microsoft Windows, while others support Apple Macs or Linux distributions. A few even have versions for mobile devices.

Multiplatform file synchronization services represent one architectural choice. Selective synchronization allows the user to select the files and folders that need to be backed up, set some of the backup parameters and, when necessary, restore files that may have been deleted or corrupted locally.

Imperfect Encryption

Although encryption is a vital component for securing data in a public environment, it doesn't guarantee that the information is safe. Its efficacy depends on the attack vector.

For example, it is a sound recommendation to encrypt SQL databases that are stored in the cloud. Then, if someone is able to copy the repository to another system, the data will still be protected. Otherwise, if a hacker is able to launch an attack using SQL injection, the application engine will decrypt the data and serve it back to the hacker in plain text.

One significant advantage of a cloud-based backup solution is its support for mobile users, who are almost always connected to the Internet. That makes it possible to implement continuous data protection so that files are immediately backed up as they're changed.

Another area to consider when it comes to user experience (though not end-user experience) is the administrative console. Does it allow for role delegation, advanced reporting or the definition of user groups with custom configurations?

There may also be a need to integrate the service into an internal infrastructure to enhance user experience. For example, Active Directory integration may facilitate single sign-on and allow a synchronization of user-based policies. On the management side, automated invoicing can help centralize billing for all departments and ease the burden on the IT team.

Security: Cloud-based backup solutions rely heavily on resource pooling and public Internet connections, so there is a risk that sensitive data could be compromised. To address this concern, it is important to look carefully at the security policies, processes and offerings of the provider.

This means evaluating the operational procedures in place for hiring personnel, enforcing physical security and implementing role-based segregation of duties. It also means verifying that the provider maintains necessary certifications and proactively submits to publicly verified audits.

In a virtual environment, the most impermeable boundary is a cryptographic one. Encrypting all data is the only secure option for sensitive information. But the task is more difficult than simply deciding to encrypt. The customer should indicate, or at least approve, the algorithm, key length and key management procedures (see the "Encrypting Data" section).

Archiving: One reason why backup has become a monumental task for many organizations is that it attempts to simultaneously address a number of requirements. In addition to facilitating both short-term data restoration and disaster recovery, it can be used to ensure compliance with local laws. However, it can be an overly blunt instrument in trying to achieve these objectives.

Data archiving offers a streamlined mechanism for complying with e-discovery regulations. It also reduces internal storage requirements and facilitates classification and enterprise search across deduplicated and compressed repositories of critical information.

It doesn't remove the need for backups, but it ensures that online and offline data are reduced to their minimum and makes them as accessible and usable as possible. Data archives, which are indexed and searchable, shouldn't be confused with data backups, which are essentially copies of data.

Many cloud backup providers, such as Symantec, offer archiving solutions for e-mail, file shares and portals.

Storage Virtualization: A First Step

Storage virtualization is often a precursor to remote backup because it offers many benefits and helps achieve location independence for the data that may eventually move to the cloud. The technology presents a logical space for data storage and handles the process of mapping it to the actual physical location. Virtualization software redirects incoming requests based on a logical disk location and translates them into new requests referencing a physical disk location.

This abstraction makes storage consolidation much easier because it is possible to migrate data without disrupting access. The host only knows about the logical disk, so the physical data can be moved or replicated to another location without affecting the operation of any client.

Resource pooling can increase utilization. The physical storage is logically aggregated into pools. As a result, additional storage systems can be added as needed and the virtual storage space will scale up transparently. This allows users to avoid overbuying and overprovisioning storage solutions.

Finally, storage virtualization software can act as a centralized console for managing all volumes in the environment. Multiple dispersed and independent storage devices appear as a single monolithic storage device.

In summary, storage virtualization offers several benefits that stand on their own merits. But optimized efficiency, centralized management and storage abstraction are also key factors in facilitating a smooth move to the cloud.

Organizations that have strict compliance needs or plans to simplify their e-mail storage and search capabilities are well advised to consider an archival solution instead of (or in addition to) a remote-backup service.

Network load: One of the biggest concerns about cloud-based backup is that if large amounts of data need to be stored, they need to pass over the network. Solutions on the market employ a variety of techniques to minimize the effect on the user.

Most services make only differential or incremental backups of files that have changed. So a small change in a large file triggers only a small amount of network traffic. Furthermore, data that is transferred is often compressed to minimize the upload. Many solutions allow users to manage network load though intuitive sliders and other interface features.

Along those lines, bandwidth-throttling techniques allow for an adjustable data transfer rate. In some cases, the user can specify when to reduce the backup speed to prevent interference with other applications. In other cases, the system will detect when the network connection is slow or has a high cost associated with it, and will throttle or delay the process itself.

Network Connections and Bandwidth

High-performance, robust connectivity is vital for remote backup because it relies on the network to transport and restore data. The connection is generally not dedicated to backup, so it is not possible to plan in isolation. The network infrastructure for the entire IT architecture must be evaluated while taking backup requirements into consideration.

When it comes to remote backup, poor network planning can manifest in two ways. First, insufficient network capacity may mean that a backup cannot run in a reliable and timely manner, leaving some data unprotected – at least for a period of time. Second, the backup may encroach on the network needs of other, potentially mission-critical, applications and cause their failure.

The only way to understand the full network requirements of the organization is to inventory all applications and users, and then map out which applications use which storage. This exercise allows the IT shop to identify all the data paths and chart both their requirements and characteristics.

The next step would be to describe the requirements for each connection. Some of the questions an organization might ask about its network include:

- What applications and traffic will run over it?
- How critical are these applications to business operations?
- Are there any bandwidth/latency constraints?

- Who can access the connection? Is it secured?
- How reliable is it?
- Can any failover or redundancy be implemented?

Focus particularly on bandwidth and latency. Network planning also may lead to QoS requirements, whereby higher-priority traffic runs over the network before lower-priority traffic. And dealing with especially sensitive data may require one type of secure connection over another, which can further affect network performance.

The bandwidth requirement a company specifies will range from average bandwidth to peak load and depend on the criticality of the data. Anything less than average bandwidth requirements will create a bottleneck. Allocating above the expected peak load (including some margin for error and growth) may not present a technical problem, but also may not make sense economically.

When considering remote backup, the peak load is usually the initial full backup. Subsequent incremental changes are relatively small. There are two ways to mitigate the impact of the initial load.

One approach is to use a form of offline backup to bootstrap the system; some providers allow customers to ship tapes or disks that they will import for a fee. A second approach is to stagger initial backups by system or device. By dividing them into phases, the total network traffic will remain the same, but it will be spread over a longer period and therefore require less network capacity at any one time.

Latency, which is largely independent of bandwidth, represents the delays that occur between sending a packet from the source and receiving it at the destination. It is one of the most critical factors affecting user experience and speed of transactions.

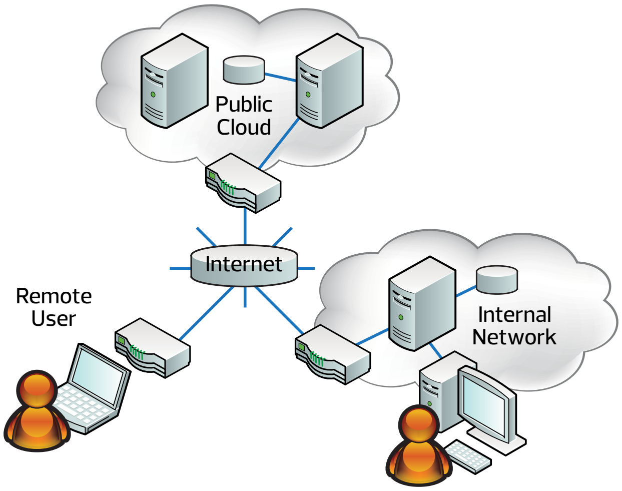
QoS addresses latency, as well as dropped packets and jitter (or, variance in latency). In general, browser-based HTTP access is largely resistant to minor transmission problems. The upper layers in the protocol stack can recover from errors without the user noticing a difference.

Latency and QoS are only noticeable in remote-backup solutions that operate synchronously, and most do not. Asynchronous backup has its own drawbacks, namely that it cannot always guarantee data integrity. But it greatly improves user response time in a high-latency environment and allows for offline operation.

Finally, if a network carries sensitive traffic, it should be secured. This doesn't necessarily mean it must be encrypted at the network or transport layer. Applications may use the Secure Sockets Layer (SSL), which will provide a similar level of protection. Nonetheless, the requirement should be documented so that it isn't inadvertently lost if the application provider assumes network layer security.

Most of these considerations apply to a company's intranet and Internet service provider (ISP) connection, the areas over which it has the most control. However, with remote backup, an organization needs to look at end-to-end connectivity to obtain a complete picture, because some of the network path to the cloud is outside its control.

Examining End-to-end Connectivity



Like companies, network carriers and cloud providers have finite network capacity and they must also manage their bandwidth. Not only do they need to plan for future growth in their subscriber base, but they also need to project the collective usage patterns of all their customers.

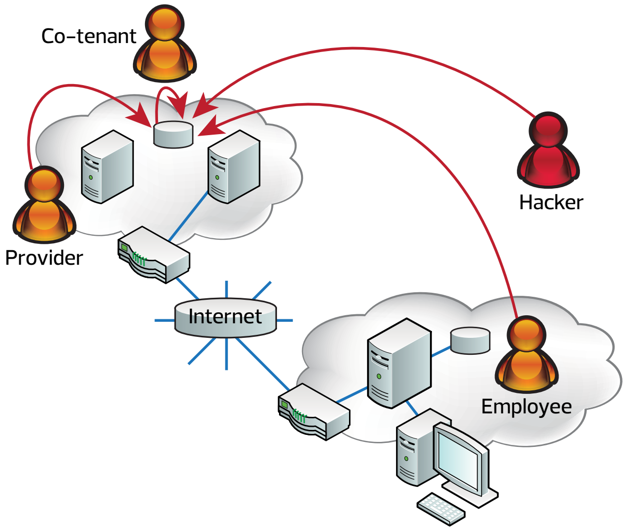
When considering a remote-backup solution, companies may want to ask questions about the service provider's bandwidth planning.

Encrypting Data

There are many reasons to encrypt data stored in a cloud, especially if it is delivered over a public network. Public network travel leaves it vulnerable to attacks from anonymous cybercriminals and rogue company or provider employees. Moreover, all communications between applications and system resources within a particular cloud may be vulnerable to interception by another cloud tenant.

To complicate matters further, cloud-based data storage itself may not be compartmentalized as well as it should be, providing an opportunity for another tenant (or a malicious provider employee) to copy and analyze sensitive information. Similarly, if data destruction policies are not enforced completely, there is always a possibility of data leakage. Encrypting all data in a remote-backup solution — whether it is in motion or at rest — is the only secure way to protect sensitive information.

Cloud Attack Vectors



When it comes to remote backup and data encryption, the biggest challenge is often determining how to manage cryptographic keys. If the provider takes responsibility for key management, a single key may be used for all customers.

It is in the interest of the customer to isolate encryption realms as much as possible and not share them with other tenants. But this increases the management burden for the provider and may become an item of negotiation. Regardless of the number of keys, it is critical to ensure that data and keys are kept separate and the chain of custody for keys is well documented.

An alternate approach is for an organization to implement its own encryption to protect sensitive data. This is advisable even if the cloud provider offers basic encryption of its platform. Nonetheless, this approach also has drawbacks: Not only must the company implement its own key management solution, but it also will not benefit from any value-added services because it has made the data opaque to the provider.

Choosing a Provider

Having determined that cloud-based remote-backup service suits its needs, matched delivery models to internal requirements, and settled on which backup features are most important, it's time for the organization to choose a provider. There are four main categories of evaluation criteria for selecting a cloud provider: environmental, contractual, technical and financial.

Environmental: At the highest level, the key environmental factor is location. Given the potential effects of a natural disaster, a company should be aware of any cause for concern in the geographic area where the provider concentrates its resources. For instance, the area's susceptibility to electrical storms might increase the chances of a power outage, which would warrant additional safeguards and precautions in the provider's data center.

Reliable power is essential to a hosted data center. Ideally, a provider should be able to tap into more than one power grid. Furthermore, it should have considerable power redundancy via alternate sources, from batteries to generators.

Similarly, the provider's heating, ventilation and air-conditioning should be able to sustain any single point of failure. Finally, given the dependency on connectivity, a provider should be able to guarantee a level of network resilience through multiple ISPs.

Physical access control is also an environmental consideration. Does the facility have solid security procedures that prevent unauthorized access to the building and computer equipment? These controls might include everything from biometric and multifactor authentication systems to electrical fences and state-of-the-art surveillance systems.

Contractual: The contract between an organization and its remote-backup provider may involve a number of legal subtleties. Some of the most important considerations include metering, billing and audits/certification. But for organizations that expect their cloud-based backup resources to operate reliably around the clock, availability is key.

One of the most important components of the terms of service is the service-level agreement (SLA), which focuses on availability — the percentage of time a particular resource or service is in a usable state over a measured time interval. Ideally, all resources would be available 100 percent of the time. In practice, there's always some downtime due to technical problems or planned maintenance.

It is common to describe approximate uptime in terms of the number of "nines" of availability. Before an organization becomes too focused on the nines, consider that there is less than 0.1 percent difference in uptime between three nines and 10 nines.

If a company has stringent availability needs, it can opt for premium SLAs from the service provider or implement redundancy and failover measures itself. All of these options have associated costs, so it is imperative to understand the business justification for higher availability.

Technical: From a technical viewpoint, the most important considerations are required features and functionality, as well as security and the provider's ability to fulfill the service requirements.

Security criteria might include Payment Card Industry (PCI) compliance, which includes security provisions that range from physical to operational security. In a multitenant environment, the degree of tenant isolation and the mechanisms used to enforce it are important. If an organization will be relying on a provider for encryption, key management must be spelled out.

Service availability is largely a contractual issue. But if the service fails, a contract can only compensate for loss – it cannot prevent it. So when choosing a provider, a company should gain insight into the current capacity and utilization of a provider's data centers, whether in terms of floor space, power or bandwidth. If the organization is not comfortable with expected trends, it should investigate whether a provider already has engaged in building out new data centers.

Financial: Among the most important criteria when selecting a service provider should be its likelihood of remaining in business over the long haul (a real concern). If a backup provider goes bankrupt, its customers may be in deep trouble.

There are technical precautions an organization can take, such as maximizing interoperability, so it's easy to retrieve data from a fledgling provider, maintaining a secondary provider or backing up data locally. But to minimize the extra measures, it's best just to understand the financial condition of a vendor and to ask for references from other customers.

On a more practical level, commercial backup services have an associated price. Every prospective customer needs to match cost with the perceived value of the functionality delivered. Even if cloud-based backup solutions offer great appeal in ease of use and flexibility, few companies will want to invest in them if they don't make a compelling business case.

TWEET THIS!

The information is provided for informational purposes. It is believed to be accurate but could contain errors. CDW does not intend to make any warranties, express or implied, about the products, services, or information that is discussed. CDW®, CDW·G® and The Right Technology. Right Away® are registered trademarks of CDW LLC. PEOPLE WHO GET IT™ is a trademark of CDW LLC.
All other trademarks and registered trademarks are the sole property of their respective owners.
Together we strive for perfection. ISO 9001:2000 certified
108161 – 120917 – ©2012 CDW LLC

CDW® PEOPLE WHO GET IT™