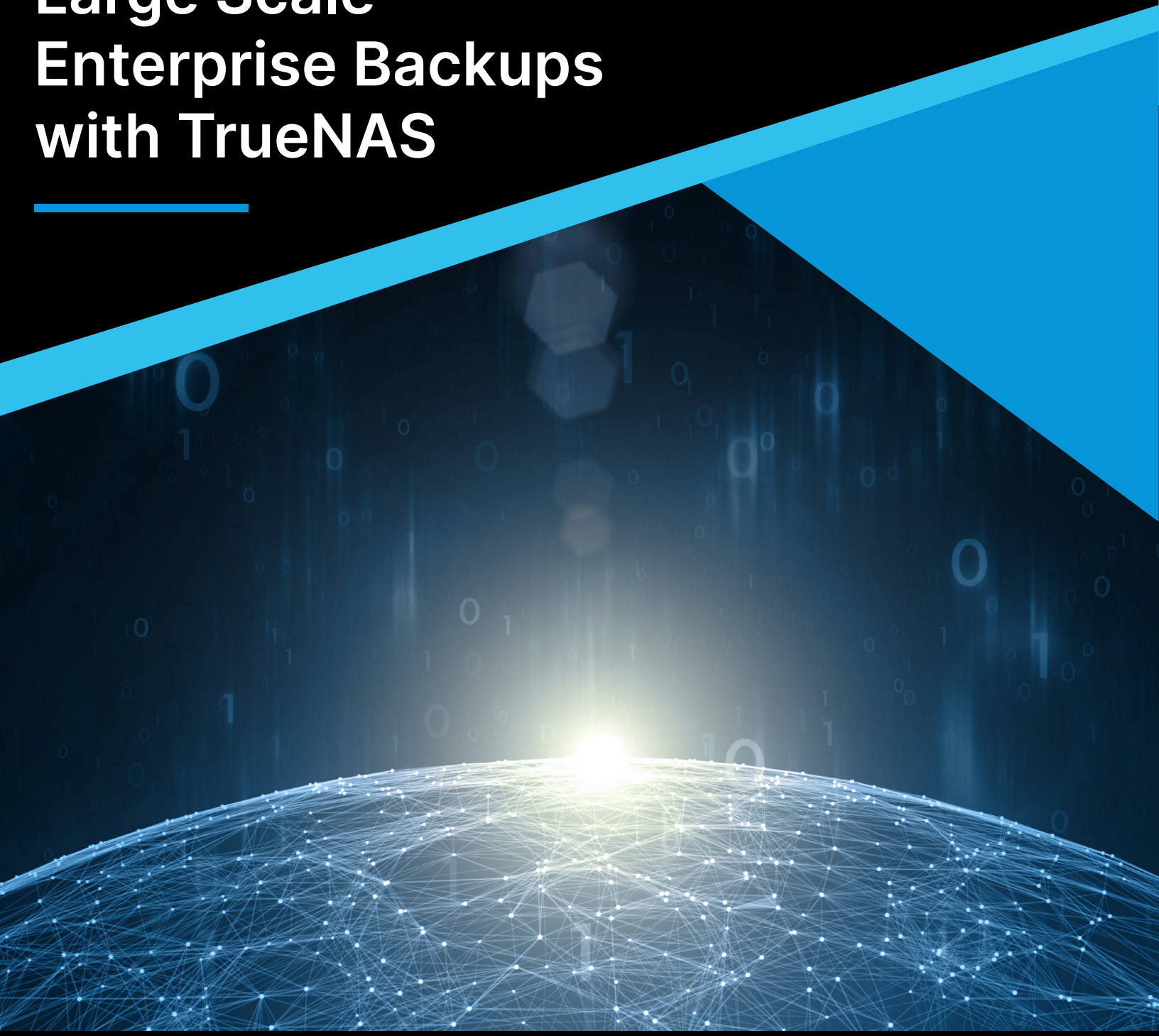
WHITE PAPER

Large Scale Enterprise Backups with TrueNAS



 | www.truenas.com | marketing@ixsystems.com

CONTENTS

1 Executive Summary

Organizations need a backup solution that is simple, secure, cost-effective, and most importantly, works seamlessly with existing applications and VMs. Learn how TrueNAS® from iXsystems protects business-critical applications and VMs, enabling business operations to quickly resume after any data disaster, ensuring business continuity, saving valuable time, and reducing data storage TCO.

According to Gartner, "By 2028, large enterprises will triple their unstructured data capacity across their on-premises, edge, and public cloud locations compared to mid-2023.". Much of this growth will come from new applications that are not yet accounted for in most organizations. Critical data will propel business in the years to come, and protecting this data is a must. Think of data protection like insurance: it is only needed when things go wrong.

Additional studies have found that:

- 90% of companies that lost their data center for ten days or more due to a disaster filed for bankruptcy within one year of the disaster.
- Half of the businesses that did not invest in a comprehensive business continuance/disaster recovery (BC/DR) solution filed for bankruptcy immediately.[1]
- 99% or companies reportedly have backup strategies in place, however 26% of them acknowledged that they couldn't fully restore all their data or documents when recovering from a backup.[2]
- Cyberattacks of small and medium-sized businesses have surged to 61%.[3]

Maintaining a backup strategy for data is necessary to keep data safe, regardless of what happens to a server. Network or disk failures, viruses, or physical damage to hardware can corrupt data, making it unreadable or unusable. Having a way to get data back is crucial to disaster recovery.

Business stakeholders, systems integrators, and systems and storage administrators should read this whitepaper. It provides an introduction to data protection concepts and covers how TrueNAS can be used to protect data in a physical or virtual server disaster recovery plan. TrueNAS from iXsystems provides a comprehensive solution for backup. We show how the advice of "just restore it from backup" is actually true as TrueNAS detects and remedies storage issues before they become outages or unrecoverable data corruption problems.

1.1 Overview of TrueNAS

Most companies purchase data storage based on capacity and performance dictated by the needs of existing applications. As a result, businesses often end up with multiple classes of application-specific storage or "storage silos" including SAN, NAS, all-flash arrays, and many forms of direct attached storage (DAS).

TrueNAS offers many choices for data replication, including Syncthing data replication to safely replicate vital data to remote locations such as a disaster recovery data center. Other vendors charge extra for this feature, increasing the TCO for data storage. As a unified solution, TrueNAS supports multiple block and file protocols that make it possible to consolidate storage silos. It scales capacity to 25 PB without performance impact, balances price and performance, scales bandwidth to more than 20 GB/sec, and combines it all with top-rated support.

1 National Archives & Records Administration in Washington (unattributed, but often quoted)
2 https://www.helpnetsecurity.com/2022/06/20/it-decision-makers-backup/
3 TechRepublic https://www.techrepublic.com/article/cyberattacks-small-medium-businesses-data-exfiltration/

TrueNAS integrates with all major backup, virtual machine solutions, and cloud storage providers. This enables files of any size to be backed up or shared to hundreds or thousands of users. Backup more data and more operating environments on a single host from a single, hassle-free array. TrueNAS has been certified by Citrix, Veeam, and VMware and is integrated with VMware VAAI/Block, VAAI/NAS, DRS, and SIOC. TrueNAS snapshots are coordinated with vSphere snapshots. TrueNAS has a vCenter plugin and also integrates with Microsoft CSV, ODX, and VSS. TrueNAS provides instant and crash-consistent snapshots of any VMware VM, allowing you to replicate a VM and restart it.

These features make TrueNAS a core component of physical or virtual server disaster recovery plans. With TrueNAS in place, you can sleep well at night knowing that your applications and VMs are protected.

2 A Brief History of Data Protection

Modern data protection strategies started in the 1980s and have evolved along with business IT practices. At first, most computer systems were standalone, with all data residing on a single system. Networking and interconnectivity between systems was limited. Systems gradually grew in size and interconnections between them increased. As computer systems evolved, so did backup technologies. Options like shared tape backup and enterprise functions such as snapshots, replication, and cloud storage gradually appeared.

2.1 Tape Backup

In the 1980s and 1990s, tape was the prevalent interchange media for data and every major system had one or more tape drives. Tape technology evolved along with other computer subsystems. As computer systems became interconnected, it became more common to share a single tape drive or tape library between multiple systems.
The advantage of tape is that it is durable and that most backup products provide tape support for automated tape libraries, including Virtual Tape Libraries (VTL), and standalone tape drives. The disadvantages of tape became apparent as computer systems grew in capacity. As the amount of data grew, so did the number of tapes needed, the restore time required, and the backup time due to increased latency.

Another disadvantage came from the necessity to shut down or quiesce a transactional application to ensure data files were in a consistent state. The timeframe when applications were shut down came to be known as the backup window. Most companies today operate globally and can't afford to have their applications shut down at any time, motivating these companies to find other ways to protect their data. Restore times became a problem when the backup used multiple tapes. During a restore, each tape must be physically located, loaded, positioned to the data, and read. This process is repeated when restoring multiple files across different directories. This takes time, money, and effort to perform, and pressures businesses to find a better solution.

Backup vendors responded by creating alternate methods of backing up data to reduce the backup window and to ensure complete backups. These included technologies such as open-file backup, incremental or differential backups, using SAN or NAS APIs to access the data directly, and supporting faster-streaming tape drives. Despite these improvements, tape-based backups tend to remain inherently slow and cumbersome.

2.2 Snapshots/Clones

Some companies moved away from depending solely on tape backup, using a combination of snapshots and clones with tape backup. Snapshots and clones provide a crash-consistent point-in-time copy of a defined collection of data in its primary location: a file system, volume, LUN, or any other type of container supported by the system. In many systems, a clone is a full copy of the data and a snapshot is more like taking a picture of your data where a subsequent picture is only the difference from the previous one.

Snapshots provide a way to perform faster and more frequent backups, without the data changing under you. This translates to faster recovery from data corruption problems. Many businesses eliminate the restore latency of tape by maintaining a snapshot for use in the event of a disaster.

Snapshot Uses:

Business Continuance: Efficiently roll back data to a specific point in time. Most data recovery requests are for a single file or just a few recently changed files rather than an entire volume. Recovering files from snapshots resolves those requests quickly and efficiently.

Application Testing: Snapshots provide a way to perform test activities on a consistent and reusable copy of operational data while the source data remains unaltered. Application errors can be minimized by testing against a persistent image of real data before bringing new application changes online. Snapshots can also be utilized for testing upgrades against a snapshotted version of the application and data instead of the live data set.

Fast Data Copying: Snapshots provide a way to duplicate data quickly without disruption of data availability.

Data Sharing: Snapshots provide a way for multiple users to have a copy of production data and virtual machines without affecting the production copy.

Ransomware: A type of malicious software designed to block access to a computer system until a sum of money is paid. Snapshots can be used to return to a point in time before the attack.

Regardless of the implementation, snapshotting a network-attached file system is less disruptive, more reliable, and faster than traditional tape backup. TrueNAS ensures data integrity with instantaneous snapshots, clones, and integration with Windows Explorer Shadow Copies and ESXi snapshots.

Snapshots alone are not a backup strategy, but are the easiest way to quickly roll back data to a previous known- good point in time. Valuable data, including snapshots, remains on the storage system. But what happens when access to the storage system fails? Adding replication to the snapshot strategy addresses that issue.

2.3 Replication

The 1990s saw the advent of data replication, also referred to as file replication and storage replication. The expression can also refer to a program or a software suite that facilitates such duplication. Replication can be performed locally, inside the storage array, or remotely, either within a site or to a remote location. Remote replication duplicates data from a local storage array to a remote recovery storage system. Clients can access data on the remote storage array using standard file or block protocols.

Replication provides an extra measure of redundancy that can be invaluable if the main storage backup system fails. Immediate access to the replicated data minimizes downtime and associated costs. Replication streamlines disaster recovery by generating duplicate copies of all backed-up files. It provides excellent backup infrastructure for the efficient recovery from a natural or human-caused disaster such as a fire, flood, hurricane, virus, or ransomware. What would happen if all the infrastructure failed? Replicating data to a remote site or into the cloud addresses that issue.

2.4 Cloud Replication

The cloud is not a physical thing. Rather, the cloud is a network of servers, where each server has a different function. Some servers use computing power to run applications or "deliver a service." For example, Microsoft moved its Office software products to the cloud. Customers no longer buy the Office Suite (Word, Excel, PowerPoint, Outlook, Publisher, and OneNote) as a set of software. Instead, they pay a monthly subscription fee to use Office 365. Other servers are responsible for storing data. Cloud users don't know or control which server is used, they just get compute and storage resources.

Organizations started deploying applications and data in the cloud. This is great for collaboration, but not so good for data as people commonly think that files in the cloud do not need backup. Unfortunately, the cloud is not a silver bullet solution that addresses backup requirements, and many businesses have found that cloud files without a backup are not protected unless explicitly guaranteed by a service level agreement.

The cloud is not under your control. Cloud providers have a strategy to protect their servers, but if disaster strikes, your business must wait for the cloud provider's employees to make things right again, if they ever do. Cloud outages can leave business-critical data unavailable for hours. Without a service level agreement with the cloud provider on data availability, it might never be restored. The cloud provider may refund for the downtime, but your data is still gone.

Many organizations have concerns using the cloud for their data:

Cloud is not free, it costs time and money

Do not be misled by the free entry-scale compute and storage options that many cloud providers offer. Free is not free, as it takes time and money to add and retrieve cloud data. Without a dedicated high-speed connection to the cloud provider, writing and reading terabytes of information can take days or even weeks. Also, retrieving data from the cloud can be exorbitantly expensive.

Cloud increases TCO

Keeping data in the cloud for extended periods of time means paying for the operating costs of the cloud storage for multiple years, increasing TCO. If a business scales rapidly, it may experience unplanned increases in cloud storage costs. Keeping data on-premise can provide huge speed and cost savings and predictable benefits.

Security of data

Cloud environments face many of the same threats as a traditional corporation, but the concentration of data makes them an attractive target.

Cloud providers have a strategy to protect their servers from a data breach, but that does you no good if your data is exposed. You must be responsible for protecting your own data in the cloud. Without a service level agreement with the cloud provider on data protection, your data may be exposed. The cloud provider may refund for the exposure, but your data is still exposed. Be sure to check your service contract for file copying, restoring, security, and recovery to be aware of all the potential costs and guarantees

All major cloud service providers have suffered outages and exposures as well as issues that caused data loss. Unless you protect your data in the cloud, there is the risk of losing all of it in an instant, forever. Because of these exposures, many are deploying a "private cloud", to ensure their data is local.

TrueNAS can be used instead of an expensive backup service for a cloud deployment. Deploying public cloud storage seems inexpensive at first, where you pay a monthly fee to lease storage services instead of deploying your own equipment and software. However, TrueRack's converged approach reduces the TCO, complexity, risk, and time to deployment for virtualization, private clouds, big data, web scale, HPC, and virtually any hyperscale and computing application. It does not lock you into one design; as needs change you can change the infrastructure allocated to compute, storage, and networking.

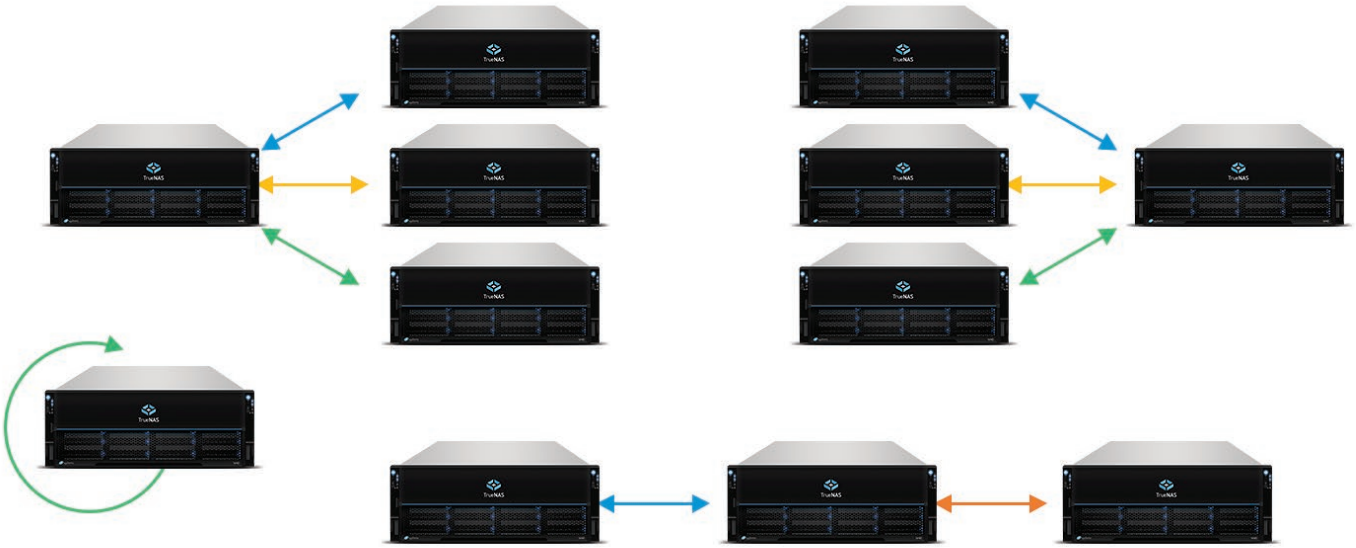
3 Using TrueNAS to Protect Mission-Critical Data

TrueNAS provides a better option than tape-based backup for data protection of applications, VMs, and their data. In addition to its support for the OpenZFS self-healing file system, unified file and block protocols, non-disruptive upgrades, capacity scalability, snapshots, replication, and performance, backups are a popular use case of TrueNAS.

An IDC survey of over 3,000 IT decision makers from small and large businesses stated that most of the participants were concerned about losing data due to a fire, flood, earthquake, or hurricane. However, IDC also stated that only 10% of disasters are due to “Mother Nature”.

A TrueNAS storage array includes high-performance, comprehensive file and block data protection using near- instant snapshots and TrueSync bidirectional data replication. These features do not need to be purchased separately, as with traditional storage vendors. It lets you establish an automatic protection schedule to meet your SLAs, or run snapshots or replication manually. The TrueNAS web-based graphical user interface is used to control every aspect of snapshots and replication. TrueNAS from iXsystems lets you survive a disaster of any type that could result in data loss due to:

- Hardware failure or power loss
- Application failure
- Loss of availability or data corruption
- Accidental or deliberate actions
- Theft



Local replication, usually written as A→A or A→B.

A and B are at the same location and can be on the same TrueNAS storage array, usually written as A→A, or on two different TrueNAS storage arrays, usually written as A→B. Having A and B on two TrueNAS storage arrays provides protection from a storage array outage of either A or B. However, local replication does not protect from a site-wide outage.

Bidirectional remote replication, usually written as A→B, B→A.

Protects data from a site-wide outage. It allows a user to perform production changes using the data on the target site and then replicate those changes to the original site. It also allows production on A and B. B is the DR target for the production on A, and A is DR target for B’s production. In the event of an outage at A or B, all of the production is done at the other site, and when it is online again, all of the changes at the DR target are sent to the other site.

Multi-node remote replication, usually written as A→B, A→C.

A major use of this is when system B is in the same general location as A, say on the same campus, and C is off-site at a remote location. The data is simultaneously replicated to B and C. During a physical outage of A, B can be used. If there is a site-wide outage that disrupts A and B, C can be used.

Multi-Hop remote replication, usually written as A→B→C.

Allows a company to have a DR site for their DR site, here called C. The data is first replicated to B and then replicated to C. With TrueNAS, different schedules can be used for B→C and for A→B. This means that you can frequently replicate from A to B, say once an hour, and replicate less frequently, say once per day, to site C.

Multi-Site remote replication, usually written as A→B, C→B.

Allows for a DR site that is used by sites A and C. If A has an outage, site B is used for data. If site C fails, site B is also used. If both A and C have an outage, site B can be used for both.

3.1 Fast Local Replication

TrueNAS local replication uses its instant snapshot capability to create an exact point-in-time image of a file system. The snapshot is then sent to another file system on the same TrueNAS storage array for data protection purposes.

3.2 Fast Remote Replication

Remote replication is the process of copying snapshots of production data to a TrueNAS storage array at a remote location for data protection and disaster recovery purposes. The remote TrueNAS storage array does not have to be the same model or have the same configuration. For example, you can choose to replicate between a TrueNAS hybrid and a TrueNAS all-flash array.

TrueNAS replicates snapshots, which means that the file system as it existed at the time of the snapshot is replicated. This ensures a consistent image of the file system, even if the replication itself takes hours. When a file system is initially replicated, called synchronization, all of the data is sent. This can take a significant period of time, from many hours to possibly days.

Since the initial synchronization may take a significant period of time, many customers invest in a high-speed network connection or put the target physically in the same datacenter as the production system and use the local network infrastructure for the synchronization, then move the replication target to the DR site. Subsequent replications take far less time, as only modified data is replicated.

Since remote replication is asynchronous, designed to work over long distances, and requires less bandwidth, it is often a better option for disaster recovery. The remote replication RPO is typically expressed in minutes or hours, depending on the importance of the data that is replicated. There may be different RPOs for different systems, determined by the nature of the data as well as bandwidth and cost limitations.

3.3 Quickly Restart any VM or Application

An application or VM can be restored from a backup stored on TrueNAS, but that increases the time needed before restarting these applications and VMs. Since snapshots and clones in VMware environments can be hard to coordinate between vSphere and a storage array, TrueSync replication is designed to be space-efficient and VM- consistent, allowing for instant recovery from data loss or corruption.

With TrueNAS, every snapshot is crash-consistent, providing a known recovery point for business-critical applications and VMs. This provides a quick restart of applications and VMs from a snapshot. With snapshots, TrueNAS provides VM-consistent, point-in-time copies of any VM or application, allowing it to be replicated and restarted seamlessly.

3.4 Improving Backup Performance, Reliability, and Redundancy

TrueNAS improves read and write performance, which reduces the time needed to backup and restore applications or VMs. Improving the time storage I/O takes to complete results in faster backup performance. TrueNAS uses TrueCache to combine RAM and flash memory with spinning disks for optimal I/O performance with the economics and capacity of hard disks.

The read caching used by TrueNAS uses predictive algorithms to copy data proactively from spinning disks to SSDs, enabling faster read times. When a restore operation starts, TrueCache accelerates the restore reads by an order of magnitude.

Since writing backup data directly to hard disk drives increases write latency dramatically, TrueNAS usesTrueCache technology to cache writes. This decreases backup write latency and ensures writes complete faster. When a backup writes, TrueNAS acknowledges the write once it hits the flash. TrueCache then asynchronously writes the backup data to the HDDs in the background, allowing for faster overall backup speeds. This also ensures backup writes will not be lost if a crash, power failure, or other system disruption occurs.

TrueNAS checksums each backup write and verifies those checksums with each restore operation. If checksum verification fails, an available redundant copy of the data is retrieved and used to correct the original error. All file system metadata is also checksummed, allowing periodic validation of the integrity of all data on disk in order to guarantee that it is not suffering from silent data corruption, also known as bit rot.

Reducing the RTO (Recovery Time Objective) of the data being protected also increases the storage used by the backup images. RTO means “how quickly can the business return to making money?”. Using TrueNAS as a backup repository allows you to scale backup storage as the RTO increases. TrueNAS grows to over 25 PB, letting you protect thousands of VMs.

When backup data grows, the expansion of TrueNAS storage is simple and non-disruptive. When a drive is added, its capacity is immediately available for use, allowing for capacity expansion without service interruption. TrueNAS also features the non-disruptive addition of expansion shelves to add more storage capacity when required.

TrueNAS ensures that backup data is always available. It has at least two network ports that support link aggregation and failover, providing fault-tolerance and high-speed multi-link network throughput. TrueNAS also supports active/passive failover between pairs of links.

For the ultimate in High Availability, any TrueNAS model can be upgraded to dual-storage controllers by adding a second storage controller. This improves the availability of restore operations without increasing the kept storage.

3.5 Graduate from Tape

A business is down longer when its backup recovery relies on tape instead of using TrueNAS. A tape’s read latency is measured in minutes or hours. When TrueNAS is used for backups, its throughput and latency are significantly faster. Here are more reasons to avoid tape:

- TrueNAS helps to avoid backup windows. Files on tape are processed sequentially while files on TrueNAS are processed individually. Restoring two files from tape involves the tape moving to access the second file, which can take several seconds or minutes. Since TrueNAS directly accesses the second file, its data is available faster.
- Backups and restores are more economical with TrueNAS. There are fewer moving parts in TrueNAS, which means significantly less heat generated, less power utilized, and less maintenance.
- If the files to be recovered reside at an offsite tape vaulting service, it takes much longer to recover that data. You can back up and store mission-critical data on-premise with TrueNAS from iXsystems.
- Leveraging the Remote Replication functionality of TrueNAS allows you to maintain site-level redundancy without the time delay, expense, and risk of physically transporting tapes.

3.6 Non-Disruptive Backup Data Testing

TrueNAS enables comprehensive and regular tests of application or disaster recovery data without any disruption to production servers or backups. These tests, known as fire drills in the recovery industry, are important to ensure the integrity and effectiveness of recovery capabilities in light of the server, system, and network changes that occur over time.

Performing a fire drill with tape or the cloud can take lots of time and planning, so many users put off or never perform this necessary activity. Performing a fire drill with TrueNAS allows the user to access their backup data instantly.

3.7 Object Storage

Object storage has emerged as a modern backup target storage for several enterprise data protection software suites such as Veeam, Veritas and Commvault. TrueNAS provides an S3-compatible object store powered by MinIO. For enhanced security, the optional Object Retention feature can be used to make object-based backups immutable, increasing resilience against inadvertent deletion or modern malware threats that target backup repositories.

Backup datasets on object storage are also easier to move and replicate. Cloud Sync enables secure transfer of data between premises and clouds, facilitating data restoration and a wider range of recovery capabilities.

3.8 Remote Storage Management

TrueNAS can be easily managed from across the room or across the globe, making it ideal for backup and DR and eliminating the need for specialized, remote office IT staff. Commands are available through the web GUI, a programmatic REST API, or the TrueCommand management software, allowing it to fully integrate into a backup and DR infrastructure. Backing up with TrueNAS is safe, reliable, and simple.

3.9 Up to 24/7 Award-Winning Support for Storage

Backups are a critical asset with the success of a business depending on a backup's performance, availability, and reliability. TrueNAS includes predictive and proactive white-glove support, allowing iXsystems to identify and address a problem before it impacts your success. Opening a support issue is easy and can be done from the TrueNAS GUI. In a recent survey, people preferred TrueNAS support over the competition. Read more about our top-rated support here.

3.10 Using Third-Party Backup Software

Backup administrators can connect to the TrueNAS enterprise storage array using iSCSI, letting them keep their backup data on an enterprise storage array instead of relying upon a server with local disks. This supports maintaining multiple concurrent backup servers as well as providing redundancy and fault tolerance through backups.

3.10.1 Veeam

iXsystems is an alliance partner with Veeam software. TrueNAS appliances from the X, M, and R-series have been certified as Veeam Ready Repository solutions, resulting in joint product offerings for mutual customers.[4]

Using TrueNAS with Veeam gives administrators 3x-4x faster VM backups vs. traditional NAS solutions. Combining TrueNAS with Veeam Availability Suite gives customers an enterprise storage array with all the features and performance of other alternatives at less than half the price of competing storage arrays.

4 https://www.veeam.com/alliance-partner-technical-programs.html

In addition to both SMB and iSCSI, TrueNAS can also receive Veeam backups as object-based storage through MinIO, an S3-compatible object-based storage provider. Backups to an object-based storage target can be marked as immutable, helping to secure the backup datasets from unsolicited modifications and encryption, such as ransomware. TrueNAS S3 target is powered by MinIO, and well-positioned as a backup target for Veeam to backup VMware ESXi VMs, Office 365, Oracle, and SAP HANA.

3.10.2 Asigra

TrueNAS natively runs Asigra Cloud Backup™ DS-System. An integrated appliance, the Asigra TrueNAS solution offers backup and recovery for physical and virtual servers, files, databases, enterprise applications, endpoint devices, and even SaaS and PaaS sources, such as MS Office 365, G Suite, Salesforce, AWS, and Azure. Uniquely, the backup solution also helps prevent attack loops with malware scans on backup data at the time of backup and time of recovery.

3.10.3 Other Backup Products

Some customers use other backup products, such as Bacula, Commvault, NetBackup, and Yosemite, which are fully supported by TrueNAS.

4 Conclusion

A business needs a fast way to back up their data against unintentional alteration. Using TrueNAS for backups gives fast local and remote data protection, making it a core component of application and VM backups. Use TrueNAS instead of depending on tape or keeping data in the cloud.

With the ability to backup and replicate petabytes of application and VM data, there's no question that TrueNAS is hands-down the best value in enterprise data storage. We encourage you to experience TrueNAS for yourself today (download TrueNAS here), listen to what our customers are saying on Gartner's Website, or if you're storing critical data, feel free to schedule a call with one of our helpful Solution Advisors.

If you want to talk about your backup needs, contact us at truenas.com/get-quote, hello@truenas.com, or call 1.855.GREP.4.IX. Our product experts look forward to hearing from you.