Whitepaper

Software-defined Storage

SUSE® Enterprise Storage for IBM Spectrum Protect – Technology, Performance, Case Studies

By Jon Giffard

Seamlessly scalable, flexible and cost-efficient – these are the demands that must be met by a future-proof backend that performs backups using IBM Spectrum Protect. SUSE Enterprise Storage is a certified solution that combines software-defined storage with affordable off-the-shelf server hardware. This document outlines the solution and the way that it works. It explains how SUSE Enterprise Storage interacts with Spectrum Protect, and cites figures to demonstrate its performance as storage for Spectrum Protect. Case studies are provided to corroborate the wide range of ways in which SUSE Enterprise Storage can be used for backups.

Future-Proof Backups for Agile Companies

The volumes of data handled by companies and other organizations are rising sharply in the wake of the digital transformation. Although this problem is becoming an increasingly prominent topic of discussion as far as primary storage is concerned, backups and disaster recovery are often the poor relations when it comes to budget planning. Yet capacities must be boosted in this area too. This creates an urgent need for cost-saving and flexible solutions. The IT industry is prone to being influenced by trends that are more or less short-lived. In an ideal scenario, however, backups should form an integral part of a company's long-term infrastructure while also being flexible enough to adapt to any new demands that may arise in connection with the introduction of new services and technologies, the use of cloud structures or the expansion or alteration of the company's business model.

It is not enough for a company simply to continue on its current trajectory. Over the years, events frequently occur that entail sudden changes to the quantity and nature of the data to be stored. Such events may include mergers and take-overs, or alternatively a move into new fields of business. The users of backup systems expect them to be adaptable to both continuous and disruptive change processes, without gobbling up extra resources during periods of restructuring that are likely to be turbulent enough as it is.

The Need for Flexibility and Scalability

In order to be ready for future demands—some of which may not even be known at present—the design of backup systems should therefore be as universal as possible. This presupposes a certain level of complexity. Once a company reaches a certain size, an enterprise backup system is therefore a must-have. Many large companies and organizations rely on Spectrum Protect to handle all their backup needs. They either implement this solution themselves in their own computing center or procure it from a provider (Backup as a Service). According to IBM, Spectrum Protect's built-in deduplication, incremental forever backups and scalability right up into the petabyte range mean that it can keep pace with a company's growth while reducing the TCO of the backup infrastructure by up to 53%.

As a server-based system, Spectrum Protect's architecture is suitable not only for applications and databases but also for backing up unstructured data. Backups of virtual environments or cloud connections are also possible with the right configuration or with upgrades such as Spectrum Protect Plus. Spectrum Protect therefore caters for just about every functional demand, and also allows software cost savings to be made by choosing an appropriate license model or a custom-tailored configuration.

The Hidden Costs of Storage Backend

Yet there are often hidden costs associated with storage backends. Conventional enterprise storage combines hardware and software into integrated building blocks. At first glance, this would appear to make the task of backup server integration more straightforward; in reality, however, it often results in high follow-up costs for storage expansions and a vendor lock-in for the storage technology. This can often be a problem, particularly in the case of long-term projects such as data archiving. The customer must either go along with all of the manufacturer's technology-related decisions year after year, or engage in parallel efforts to acquire knowledge of alternatives with a view to expanding or switching the underlying backend technology.

SUSE Enterprise Storage

The alternative? Software-defined storage (SDS), which involves separating the software from the hardware, and often—but

unfortunately not always—using inexpensive standard server hardware. Spectrum Protect is hardware-agnostic as far as the backend is concerned, making it an ideal candidate for SDS. An open-source-based SDS system with widespread support among the IT community is recommended for effective investment protection. From an operational safety perspective, however, a system must also come with appropriate enterprise support. SUSE Enterprise Storage combines both of these attributes and offers many different advantages, just a few of which are listed below:

- *unlimited scalability*
- *no migration when upgrading to new technologies*
- *low administrative effort*
- *extremely high reliability thanks to the fact that the level of redundancy can be chosen freely*
- *attractive pricing*
- *no reliance on a single hardware manufacturer*
- *based on stable and high-performance open source technology*
- *easy management thanks to a graphical user interface*
- *enterprise support*

SUSE Enterprise Storage is also the perfect choice for Spectrum Protect because its architecture (with a database-based index and storage pools) can be accurately mapped to the backend architecture, ensuring optimal utilization of the backend performance and simplified administration of the entire backup system.

SUSE Enterprise Storage (SES)

SUSE Enterprise Storage is a software-defined storage product based on Linux and the open-source project Ceph. Among other things, the latter provides object storage and interfaces for block-based and file-based protocols. SUSE supplements the system with a web-based graphical management front end and a number of extra features, the most important of which is enterprise support. Both SES and Ceph have clear and transparently communicated roadmaps that can be used as a basis for storage expansion planning.

Object Storage with Ceph

Object storage is a form of storage in which data are represented as objects, and not as files or blocks as is the case with NAS or SAN. This has inherent advantages; every file or each block can be represented as an object, but not every object is a file or a block. Object storage is therefore a universal storage concept that can represent all current and future storage types. This universality comes at a cost. It means additional complexity that must be hidden from the user. It also means additional overheads that may have a negative impact on performance.

Ceph is one of the few software-defined storage systems that has largely resolved these problems. File-based and block-based storage can be mapped to the native reliable autonomous distributed object store (RADOS) with full transparency. In this context, autonomous means that the store organizes itself for the most part. Algorithms are responsible for placing the storage objects on the storage media and for carrying out repairs in the event that individual disks or entire servers fail, without any need for intervention by an administrator. The same applies to storage expansions. Introduction of the BlueStore backend in 2016 marked such a massive leap forward in terms of performance that any preconceptions about slow object storage can be consigned to the past.

Architecture and Mode of Operation of Ceph and SUSE Enterprise Storage

By way of contrast to many other storage types, Ceph has only one level instead of a hierarchy. Each object has a unique identifier that is managed centrally and automatically. Objects are allocated to placement groups (PG) and placed on the storage subsystems. This process is fully automatic and carried out by means of algorithms. There is no central database and no central

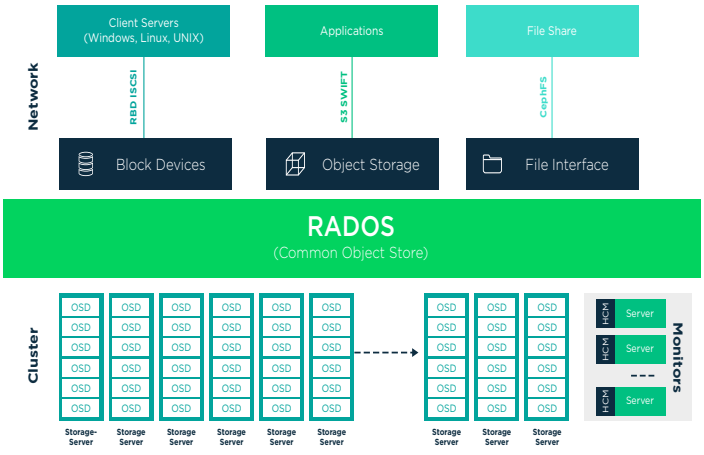

Figure 1: *Architecture of Ceph/SUSE Enterprise Storage*

hash table, meaning that there is also no single point of failure. A PG can be interpreted as a higher level of management that makes it easier to place the objects and balance the cluster. Should a failure occur, for example, all the objects within a PG will be placed accordingly (rather than individual objects). As far as the number of PGs is concerned, it is therefore important to avoid extremes (a single PG, or as many PGs as there are objects). As a general rule, one object storage daemon (OSD) exists for each hard disk or solid-state disk in the storage network. The primary OSD is responsible for continuing to place the objects via the cluster network or for delivering the object to the client. The CRUSH algorithm (controlled replication under scalable hashing) controls which objects are stored and where. By also using the compact CRUSH map, the object name and the pool name, the client is able to calculate the location of individual objects very quickly, and determine which OSD is primarily responsible for reading from and writing to these objects.

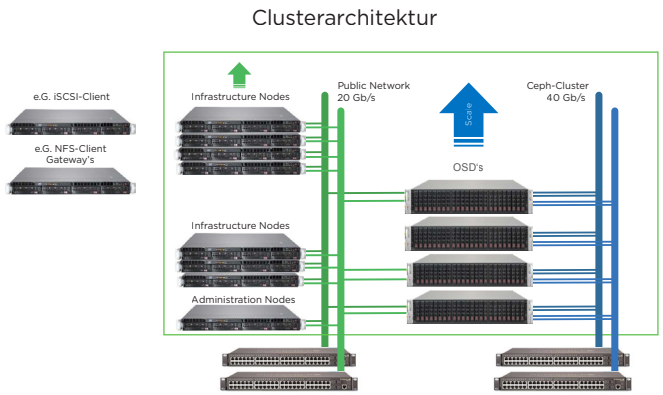

Figure 2: *Set-up of an SES Cluster*

The pools form the uppermost logical level of the cluster. Examples of structures within these pools include block devices or S3 buckets. This is therefore the level at which the administrator determines which data belong together in organizational terms, the type of storage (block, S3 buckets, file system) to which they are assigned or the level of redundancy they have.

Monitoring services (MON) act as the brain of the cluster and are up-to-date with the cluster's configuration at all times. They are the first point of contact for clients seeking information as a basis for calculating which OSD they need to establish a connection with in order to write to or read from objects. MONs also monitor the status of the cluster and take decisions according to the majority principle (simple quorum) in the event of failures or inconsistencies. The number of MON servers must therefore always be odd (at least three for a highly available implementation). MON services can also be operated on the OSD nodes without significant availability losses, to avoid the need for additional hardware. Separate MON nodes are recommended for larger installations, however.

Ceph often uses data reduction based on erasure coding for simple replication. This saves storage space but is more demanding in terms of CPU performance. For example, a net data volume of 100 TB would only require gross storage space of 150 TB rather than 300 TB (based on a standard replication factor of 3). The data replication type (erasure coding or n-fold replication) is specified at the level of individual pools.

Interfaces

Access to RADOS is possible in three basic ways:

- *RADOS block device (RBD) for direct access or via iSCSI*
- *Ceph object gateway for object storage (Swift and S3)*
- *CephFS for files*
- *NFS for files.*

All three access types can be combined with each other on the same storage cluster, and use the same internal storage mechanisms. The RBD interface is used if SAN storage needs to be replaced or emulated. This type of access is typically also used for enterprise backup applications. The object-oriented gateway tends to be used in cases where cloud storage compatibility is required or where large volumes of rarely used data need to be saved for posterity. Data archiving is a typical use case.

The Ceph Universe

In the medium term, an open source technology promoted by a single company is not significantly more useful than closed source software. The source code may be open, but maintaining independence from the service provider or manufacturer will come at a high cost in terms of internal development. Alongside the technical maturity of the product, a decisive criterion for the in-enterprise use of any piece of open source software is therefore a heterogeneous, active and well-organized community. Ceph meets this criterion. Although the majority of core developers are now employed by the (future) IBM subsidiary Red Hat, the Advisory Board includes technical experts from Intel, Cisco, Fujitsu, Canonical, CERN and SUSE. The Advisory Board

adopts long-term development goals and roadmaps for Ceph.

SUSE Enterprise Storage Appliance

In collaboration with the German server manufacturer Thomas-Krenn and other hardware manufacturers, SUSE has developed an appliance that is ideally suited for the initial migration of backups to SUSE Enterprise Storage:

- preconfigured for rapid commissioning
- four OSD nodes, one admin node
- hardware certified for the SUSE Enterprise Server
- 10-GB network including switch
- performance with Spectrum Protect tested against IBM references (blueprints)

Hardware Configuration of the OSD Nodes:

Mainboard: Supermicro X10DRI
CPU: 2x Intel Xeon E5
Intel 10GB Quad Port Network Interface
2x 400GB Intel NVMe (DC P3700 Series)
12x 4TB HDDs (SAS, HGST 512e)
Redundant 920 W power supply unit
Main storage 64 GB

Clusterarchitektur + Anbindung TSM

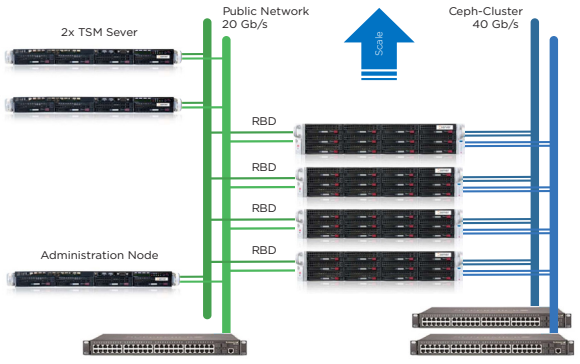

Figure 3: *Connection to Spectrum Protect*

The level of performance of the OSD nodes means that MON servers can also be operated on three out of the four nodes. The built-in storage consists of fast NVMe SSDs as well as high-capacity HDDs. From the perspective of operating Spectrum Protect, this means that the index database and the WAL are located on the NVMe storage and the containers are located on the hard disks. These requirements are persistent, meaning that auto-tiering is unnecessary.

The RADOS block device is used as an interface to the Spectrum Protect Server. The storage is therefore addressed in the same way as a traditional SAN. As many LUNs as required can be created in the pools and then presented to the Spectrum Protect server as file systems.

Performance with Spectrum Protect

IBM publishes blueprints with minimum values for key performance parameters. These can be used as a basis for the sizing of storage systems for Spectrum Protect. IBM also provides test scripts to measure the performance of storage systems. A proof-of-concept (PoC) exercise was carried out with the aim of achieving at least the small system reference values for all parameters. However, even the large system reference values were far exceeded in the case of most parameters. This became particularly clear in relation to the IOPS values for the database (Figure 4). The appliance can therefore handle a wide range of demands, even in the standard configuration with a simple 10 GB network and relatively little NVMe storage. No performance losses are to be anticipated for technological reasons, even in the event of a massive scale-out using additional nodes. Ceph systems scaled up to several hundred petabytes have been operating in production environments for many years.

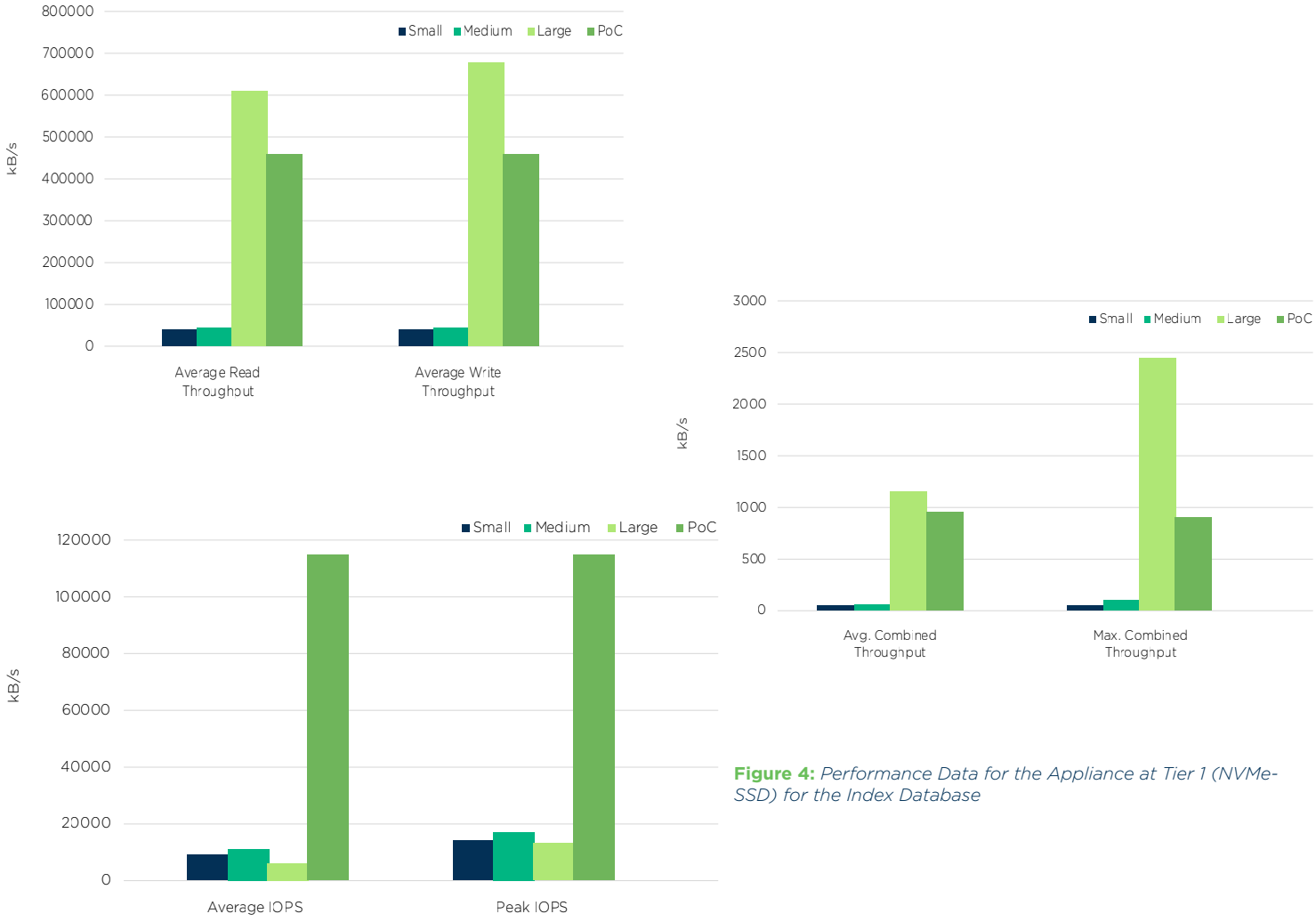

Figure 4: *Performance Data for the Appliance at Tier 1 (NVMe-SSD) for the Index Database*

Case Studies

The following examples are intended to illustrate how SUSE Enterprise Storage can be used together with Spectrum Protect in various situations to reduce costs, increase availability or reduce infrastructure complexity. The use cases are fictional but based on real scenarios.

Consolidation of Backups

A company in the service industry (Company A) merges with a competitor (Company B). Company A has used Spectrum Protect to date. A strategic decision is taken to consolidate the IT systems of the two companies. Enterprise backups will continue to be performed using Spectrum Protect. The necessary storage capacity within the existing SAN exceeds the current capacity by 50%. The company's CIO assumes that data volumes will increase by an annual figure of around 20% over the next two years. The plans are also subject to a high level of uncertainty.

Project goals

- consolidation of backups to Spectrum Protect
- affordable storage backend
- guaranteed scalability without interruptions
- long-term cost savings thanks to cheap storage and intelligent licensing.

At the needs analysis stage, the consultants working on the job discover a high proportion of redundant data. There is therefore a great deal of potential for saving volume-based licensing costs on the Spectrum Protect side through the reorganization of backups (introduction of container pools and data deduplication).

The entire data volume (around 130 TB after deduplication) is significantly lower than the 4 PB that an individual Spectrum Protect instance can manage. The Spectrum Protect server hardware must be upgraded, since deduplication takes a heavier toll in terms of CPU and RAM. But cost estimates show that ROI will be achieved after just two years thanks to the savings achieved in the area of licensing costs.

An SES cluster will replace the previous SAN in the storage backend. Three monitoring nodes will be operated on separate hardware that will be freed up as a result of the IT consolidation. Five new OSD nodes with a gross capacity of 400 TB are procured. If necessary, capacity can be increased yet further by adding extra nodes without interrupting operations. The existing tape library will continue to be operated.

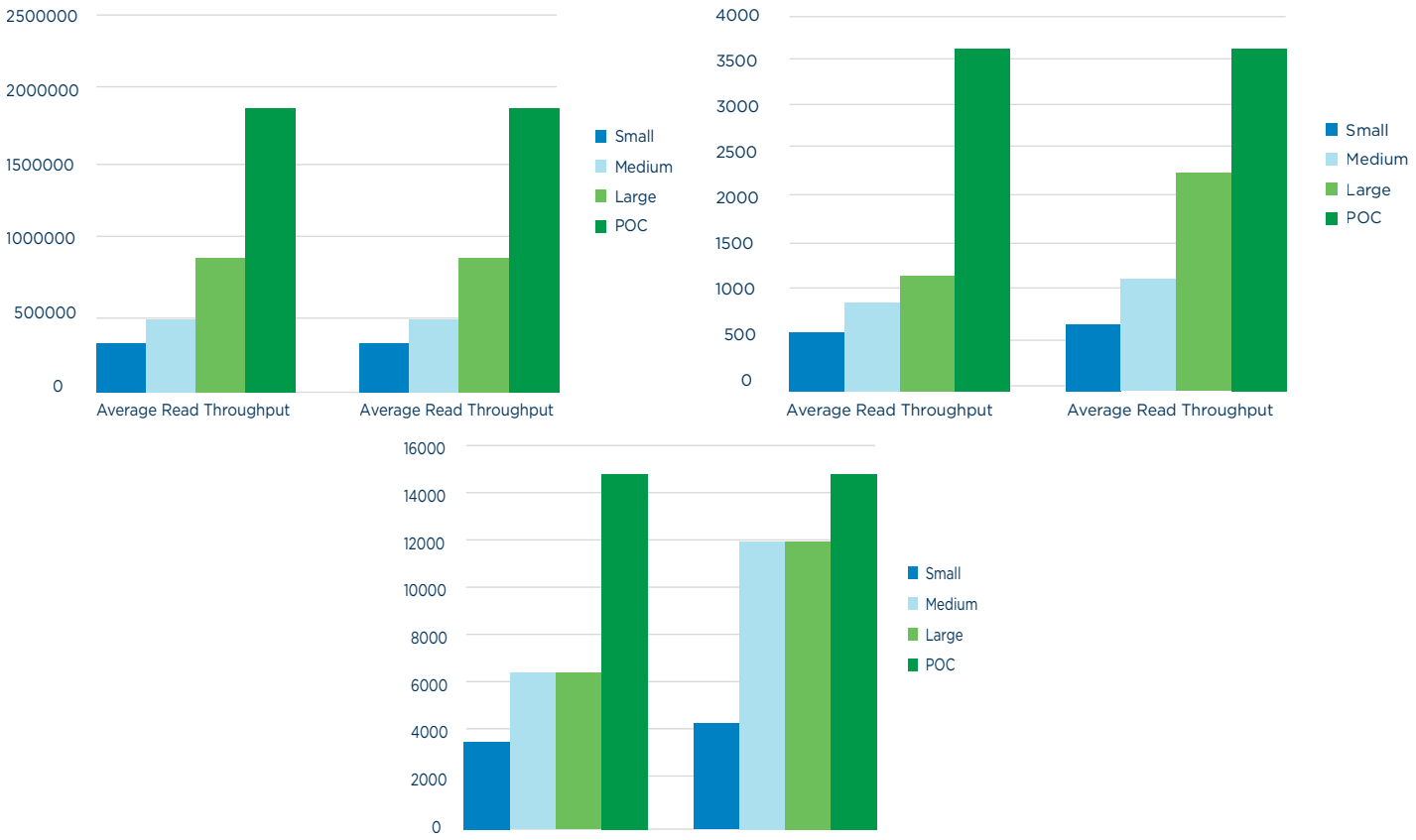

Figure 5: *Data Backup for Single-Site Operation with SES Cluster and Tape Library*

The second example relates to a financial service provider that uses S3 storage in the cloud for user data backups. During an internal audit, it is discovered that this violates the compliance guidelines. The company already operates a cluster with SUSE Enterprise Storage as primary storage.

Project goals
- achieving compliance
- integrating (private) cloud storage
- using existing storage systems.

The S3-based backup was configured in Spectrum Protect using cloud container storage pools. It could also be implemented in exactly the same way on the internal SUSE Enterprise Storage. This would make it possible to use the Ceph object gateway directly for S3 storage. Since the existing capacity is inadequate for the extra data, a new OSD node must be added to the cluster.

A two-tier solution with SSD and HDD will be implemented for the primary storage. The additional S3 storage will not require the fast SSD memory, meaning that HDD capacity will be sufficient. The capacity will be made available using an extra storage pool. The S3 buckets will be created there in a similar way to the public cloud. The only changes that need to be made in Spectrum Protect relate to login details and permissions.

Off-Site Backup

The last example relates to a manufacturer of consumer goods, whose production operations, logistics, administration, etc. are based at a central location. The company also operates an e-commerce app. Its server is located in a computer center 250 km away (colocation). Most of the IT systems at the central location are operated on an on-premise basis and the majority are virtualized. Spectrum Protect is used on-site at the central location to back up Microsoft Exchange, the ERP system, databases, snapshots of the virtual machines and unstructured employee user data.

Project goals
- *shortening backup windows*
- *off-site disk backups for business continuity, disaster recovery*
- *no change to client-side Spectrum Protect architecture*
- *guaranteed replication, even with a slow WAN connection*

To improve business continuity, the tape libraries that have been used to date will be replaced entirely by disk backups, and replication of data from the main location will also be set up in the colocation computer center. As a first stage, the tape library is therefore replaced by a four-node cluster with SUSE Enterprise Storage. The connection is implemented in the form

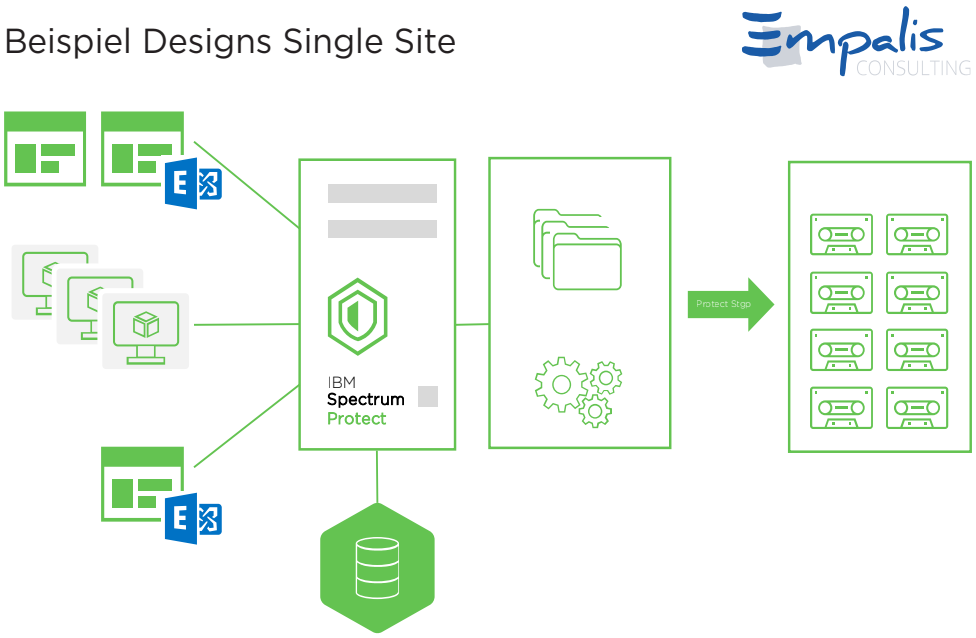

Figure 6: *Off-site Backup with Mirroring of the SES Cluster*

of block storage via iSCSI. As a second stage, a cluster for SUSE Enterprise Storage with approximately the same capacity will also be configured in the colocation computing center. Replication of Spectrum Protect itself could be used to connect the replicated cluster. However, this would require an additional server in the computing center.

Mirroring and asynchronous replication of the block storage directly into the SUSE Enterprise Storage is a much more straightforward approach. It ensures that slow WAN connections do not block the I/O of the instance to be replicated, as well as guaranteeing that temporary network failures are tolerated. Mirroring is set up on a pool basis in SUSE Enterprise Storage. Both all the block devices available in the pool and individual block devices can be mirrored asynchronously.

Empalis and SUSE – a Mutually Beneficial Partnership

The functional scope of Spectrum Protect makes it possible to implement almost any enterprise backup requirement, no matter how complex. As a flexible and low-cost storage backend solution, SUSE Enterprise Storage can go a long way toward future-proofing a backup infrastructure based on Spectrum Protect. The wide range of interfaces available mean that high-performance mapping of any scenario is possible, involving either block, file or object memory – with the highest possible level of reliability, low maintenance effort and the best possible use of existing resources. A rapid ROI is ensured through the use of standard hardware as well as through attractive and transparent pricing.

The partnership between SUSE and Empalis in the field of backups guarantees reliable advice no matter how complex the project as well as optimum data security on one of the most innovative storage systems on the market. Highly qualified individuals within both companies are available at all times to provide answers to any of your questions:

Markus Stumpf
Service Operations Manager, Empalis GmbH
markus.stumpf@emaplis.com

Sandra Zügner
Inside Account Executive, SUSE Software Solutions Germany GmbH
sandra.zuegner@suse.com

Additional contact information and office locations:
www.suse.com

www.suse.com