White Paper

SECURE, FLEXIBLE ON-PREMISE STORAGE WITH EMC SYNCPLICITY AND EMC ISILON

Abstract

This white paper explains the benefits to the extended enterprise of the on-premise, online file sharing storage solution from EMC Syncplicity and EMC Isilon. It also discusses solution deployment and load balancing options.

May 2013

EMC²

Part Number H11433.1

Table of Contents

Executive summary

Today's corporate employees expect to have access to data and services across all their devices as if they were working at the corporate office. Online file sharing (OFS) is growing in popularity because it helps users access their work files from any device. However, to achieve this, IT must adopt enterprise-grade solutions that give users the access they need while providing the security and controls that protect company data.

To deliver this level of control and protection, EMC® has launched an on-premise storage solution that combines the unmatched flexibility and ease of use of the EMC Syncplicity® cloud-based file sync and sharing technology with a secure, on-premise storage infrastructure based on EMC Isilon® scale-out network-attached storage (NAS). EMC Isilon provides a powerful, yet simple way to manage data and applications, enabling enterprises to expect much more from their storage. EMC uniquely provides enterprises a single vendor end-to-end solution to simplify deployments and to provide a single point of contact for your online file sharing needs. The benefits of this solution include:

- **Improved productivity:** Users can easily sync, access, and share files with anyone, anytime, anywhere and on any device for new levels of productivity and agility while giving IT the security, visibility, and manageability it needs.
- **Flexibility and ease of management:** Enterprises will experience the flexibility and continuous innovation that comes from deploying a cloud-based, online file sharing solution while retaining complete control over data and storage resources. The solution combines the Syncplicity enterprise-class security, compliance, and policy controls with the unmatched manageability, scalability, and resiliency of Isilon storage.
- **Reduced risk and increased security:** Enterprise information resides on Isilon storage, situated on-premise and subject to IT data governance and protection policies. Files are not replicated in the cloud. With the EMC solution, data objects stored on-premise, remain on-premise and within IT control.

This paper explains the solution and how Syncplicity can be configured with Isilon storage for an on-premise OFS solution your environment.

EMC Syncplicity with EMC Isilon On-Premise Storage

Syncplicity software includes four main components: clients (PC, Mac, iOS, Android, Web), orchestration, storage, and compute. The latter two make up the Syncplicity backend as shown in Figure 1:

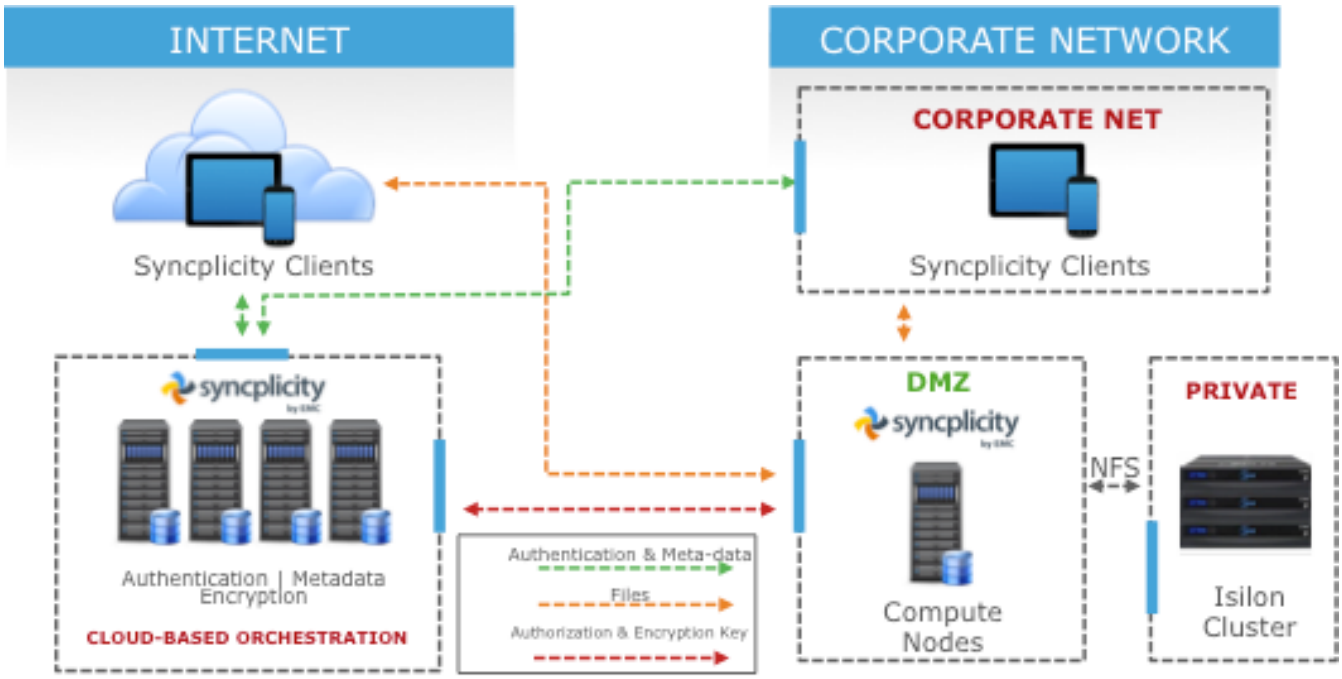

Figure 1. EMC Syncplicity and Isilon on-premise storage

Client

The client (end-user device) can be a mobile phone, tablet, or personal computer with the Syncplicity client software installed. It can also be a simple browser interface used to access Syncplicity files, and set permissions and policies as an IT administrator.

Orchestration

Orchestration includes authentication (unless delegated to SAML/OpenID SSO), authorization, account administration, metadata management, sharing and collaboration, the Web application, and the API (with the exception of file transfer). Orchestration is complex and benefits greatly from being offered as a "software as a service" (SaaS) application, absolving IT from deployment and maintenance headaches, delivering a constant stream of enhancements and innovations, and enabling seamless and straightforward enterprise collaboration across the extended enterprise.

Storage

The storage component stores file data in a highly available, redundant, and scalable backend designed for instant data access. The files that Syncplicity persists in the storage layer are stored in an opaque container on the Isilon system that is optimized for enterprise sync and share functionality.

Compute

Compute is a simple, easily deployable application. Enterprises opting for an on-premise storage solution have the ability to enable or disable encryption at the compute layer depending on administrative preferences and the security structure. The compute component exposes file transfer APIs, encryption, compression, file transfer resumption, and image thumbnail generation. The compute application is deployed on one or more virtual machines depending on load and availability requirements. All deployed compute application instances are completely stateless and independent of one another. Instances can be added and removed based on load requirements.

By design, clients may reach any instance in the pool with any request, including requests for various segments of the same file being uploaded. File transfers (uploads and downloads) are initiated by clients directly with the compute application. File data does not flow through the orchestration component at any time.

Compute application OS and application server requirements

Syncplicity recommends deploying at least two compute application instances on a minimum of two physical machines to assure basic levels of redundancy and availability. Syncplicity has certified and supports running the compute application on VMware vSphere Hypervisor (i.e., ESXi) 5.0 and 5.1.

Compute application hardware requirements

Compute application instances can run on virtual machines of nearly any compute capacity. For best performance with traditional workloads, Syncplicity recommends:

- 4 GB of RAM
- 4 virtual cores, Intel Xeon E5 Family processors, 2.20 GHz
- 500 GB HDD

The number of instances required depends on the load the deployment is expected to sustain and the exact hardware profile each instance enjoys.

Deployment

The compute application can be deployed in either of two network configurations, each with its own advantages and disadvantages. With both options, the compute application must be protected by a firewall that only permits incoming TCP connections on port 443 (HTTPS). Firewall services can be provided by the network or by the iptables firewall software on each compute application instance. Also, in both the cases the storage backend should be deployed according to each backend's best practices. Syncplicity clients never connect to the storage backend directly, either from the Internet or from within the corporate network. The storage backend must be reachable by all compute application instances via NFS (Isilon).

Deployment option 1 (recommended)

The recommended option is to deploy the Isilon storage system to be used for OFS in the DMZ within the organization's data center, while providing inbound access from Syncplicity clients to compute application instances directly from the Internet over port 443 (HTTPS). This option enables seamless, anywhere access from computers and mobile devices on any network, with no extra steps—such as establishing a VPN connection—required by the end user. At the same time, enterprise data assets reside on Isilon scale-out storage, situated on-premise, within IT control and protection policies.

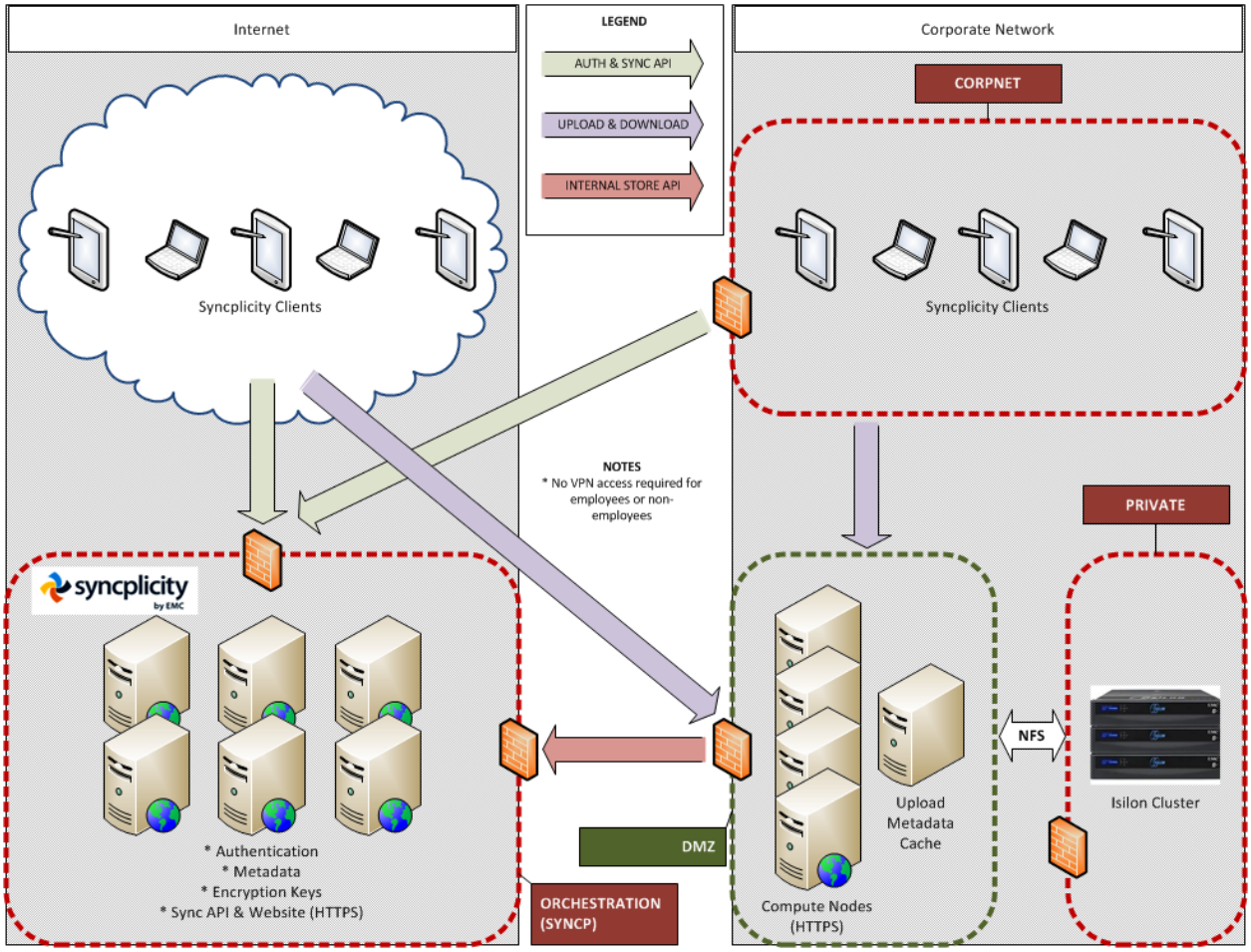

Figure 2. Deployment option 1 (with DMZ)

Deployment option 2

Another alternative is to utilize a private corporate network in the organization's data center, only accessible from within the corporate network or over a VPN connection. This option may be appropriate in cases where access to data stored in the on-premise storage and compute components must only be accessible by users with corporate network access privileges. While this option, shown in Figure 3, introduces another layer of protection, privacy, and security, it comes at the expense of user experience and extended enterprise sharing and collaboration.

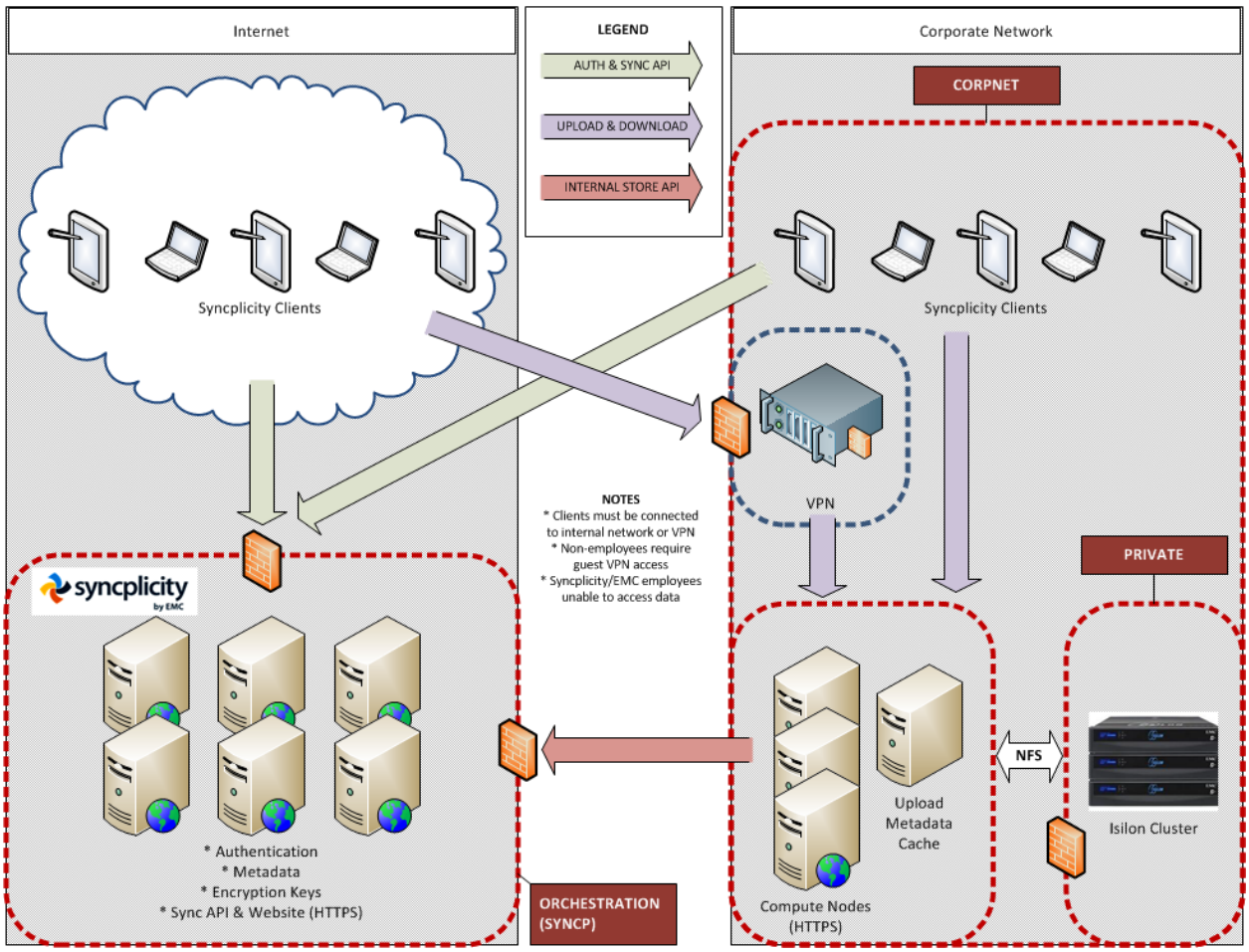

Figure 3. Deployment option 2 (using VPN)

Compute application data flow

Data flows to and from the compute application in three simple ways:

I. **INBOUND** from Syncplicity clients to compute application instances, over TCP port 443 (HTTPS).

II. **OUTBOUND** from compute application instances to storage backend, over port 2049 (NFS) with Isilon.

III. **OUTBOUND** from compute application instances to Syncplicity orchestration layer, over TCP port 443 (HTTPS).

Figures 4 and 5 show a detailed flow of the data across authenticated upload and download of data.

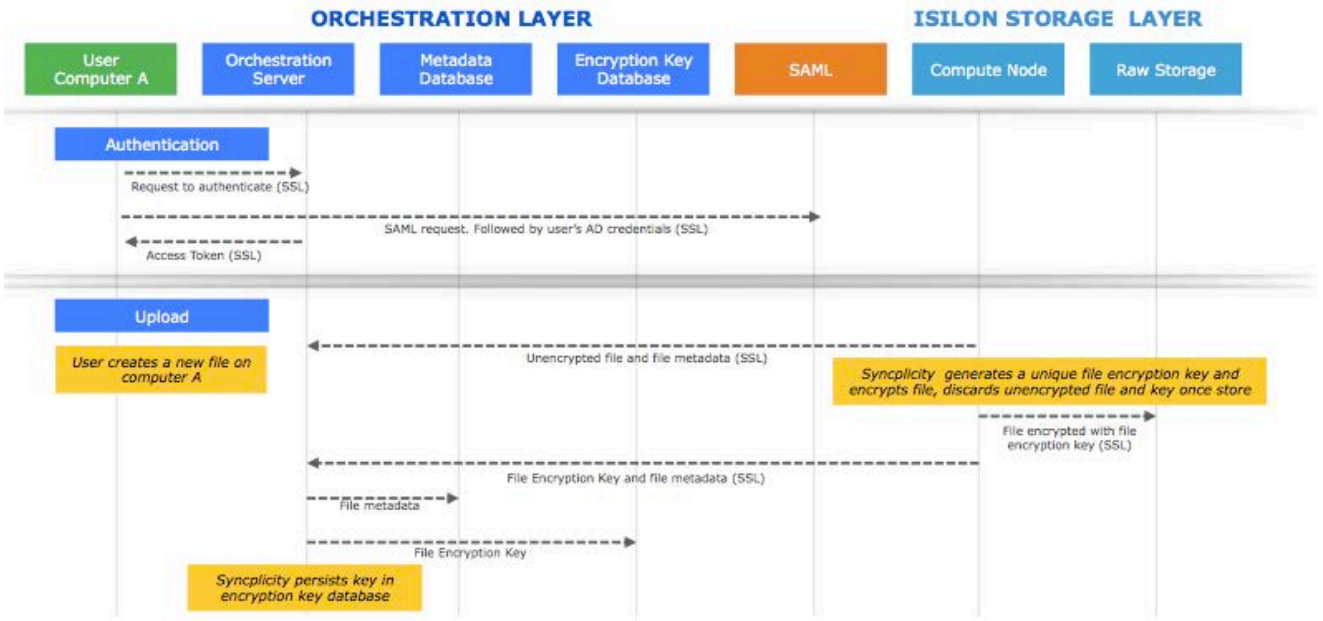

Figure 4. Syncplicity data flow—authentication, upload process

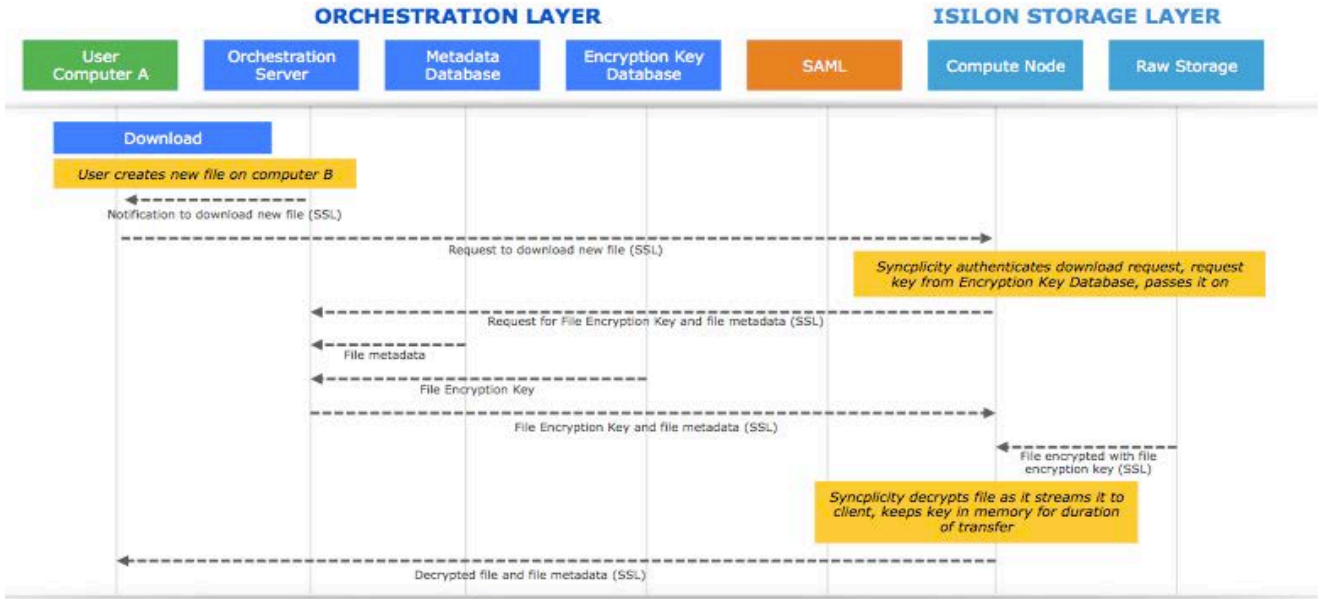

Figure 5. Syncplicity data flow—download process

Load balancing

With a set of compute application instances deployed inside the enterprise data center, file transfer traffic from Syncplicity clients must be evenly distributed to ensure proper utilization of available resources. Here are three common load balancing options.

Syncplicity-driven application-level load balancing

An administrator provisions one or more compute application instances, assigns a unique hostname to each one, and registers all hostnames with the Business Edition account. The administrator may use the administration console or the Syncplicity API to register and deregister hostnames. Syncplicity subsequently distributes a random hostname to each connected Syncplicity client. Whenever a client is unable to reach a compute application instance behind a specific hostname, the client will re-query the orchestration component for a new one. This ensures that a client always has a working compute application instance to work with.

When Syncplicity is in charge of load balancing across the enterprise compute application instances, the list of hostnames available to it must always be up to date. This means that manual or automated (via the Syncplicity API) intervention is required anytime instances are added or removed from the on-premise deployment.

DNS round-robin load balancing

An administrator provisions one or more compute application instances and configures a Domain Name System (DNS) record for the same hostname that resolves to a list of IP addresses for all the provisioned instances. The administrator subsequently registers the singular hostname with the Business Edition account. When Syncplicity clients perform a DNS resolution against the singular hostname, the DNS server will return one of the IP addresses in its list in a round-robin manner, assuring proper distribution of load.

As instances are added to or removed from the on-premise deployment, the DNS record must be kept up to date. Administrators must keep in mind DNS caching performed by intermediate DNS servers and ensure that the DNS refresh interval is set appropriately.

On-premise load balancing

An administrator provisions one or more compute application instances and places them all behind a hardware or software load balancer. The load balancer's IP address is registered with a singular DNS hostname, and the hostname is registered with the company's Business Edition account. All traffic flows through the load balancer. The load balancer is in charge of distributing traffic evenly across the set of available instances, ceasing to forward traffic to downed instances, and beginning to flow traffic to newly created instances.

The unique upside with this option is the lack of reliance on any external factor or mechanism for load balancing. The load balancing process is invisible to Syncplicity clients, Syncplicity orchestration, and the DNS server. The internal load balancer can react immediately to up/down state changes of the backend instances. Nevertheless, this option requires the most work and expense on the part of the network administrator.

Scalability

Scalability is very easy to achieve within the compute application. Due to the stateless nature of each compute application instance, instances can be dynamically added and removed as load requirements change. To scale the compute application, an administrator will deploy the compute application .WAR file on an additional machine with available resources to spare. Subsequently, an administrator will add the new instance to the pool of available instances either via Syncplicity, DNS, or an on-premise load balancer.

Isilon provides a highly scalable, modular storage architecture that can grow easily with the needs of your business. Isilon scale-out NAS offers unimaginable room for growth of your online file sharing storage solution. Isilon clusters can scale from a three-node configuration with 18 TB of capacity to a 144-node cluster configuration with over 20 petabytes of capacity.

Conclusion

EMC is the first and only vendor to provide a single end-to-end solution from consulting to financing to software to storage to global support. This provides ease of acquisition and deployment, peace of mind and the freedom to focus on your business rather than managing multiple vendors. EMC Syncplicity is the on-premise storage solution that combines the unmatched flexibility and ease of use of cloud-based file sync and sharing with a secure, on-premise storage infrastructure based on EMC Isilon scale-out NAS. For the extended enterprise, this solution offers:

- Improved productivity
- Flexibility and simplified management
- Reduced risk and increased security

To learn more about the EMC Syncplicity on-premise storage solution with EMC Isilon scale-out NAS, please visit www.emc.com/isilon or www.syncplicity.com