REYNOLDS BUSINESS SYSTEMS
682 N. Brookside Rd. Allentown, PA 18106
P: (610) 398-9080 <> F: (610) 398-9752
www.reynoldsbusiness.com

SECURE, ENTERPRISE FILE SYNC AND SHARE WITH EMC SYNCPLICITY UTILIZING EMC ISILON, EMC ATMOS, AND EMC VNX

Abstract

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

Published August 2014

EMC²

Part Number h11980

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® Syncplicity® has launched an on-premise storage solution that combines the unmatched flexibility and ease of use of EMC Syncplicity's cloud-based file sync and sharing technology with a secure, on-premise storage infrastructure based on EMC Isilon, EMC Atmos and EMC VNX storage. EMC VNX 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.

EMC Syncplicity Overview

Syncplicity is an easy-to-use, enterprise-grade file sync and share solution. Our vision is to redefine "Files" for the mobile workforce. Unlike other solutions:

- **Users get** improved productivity from access and sharing of *all* their local files from all their devices automatically, with no extra steps.
- **IT gets** control over all of the content that currently exists in unmanaged locations like email attachments, local desktop folders, and consumer cloud services.
- **IT gets** strong security, controls and storage flexibility to protect corporate files and adhere to compliance requirements.

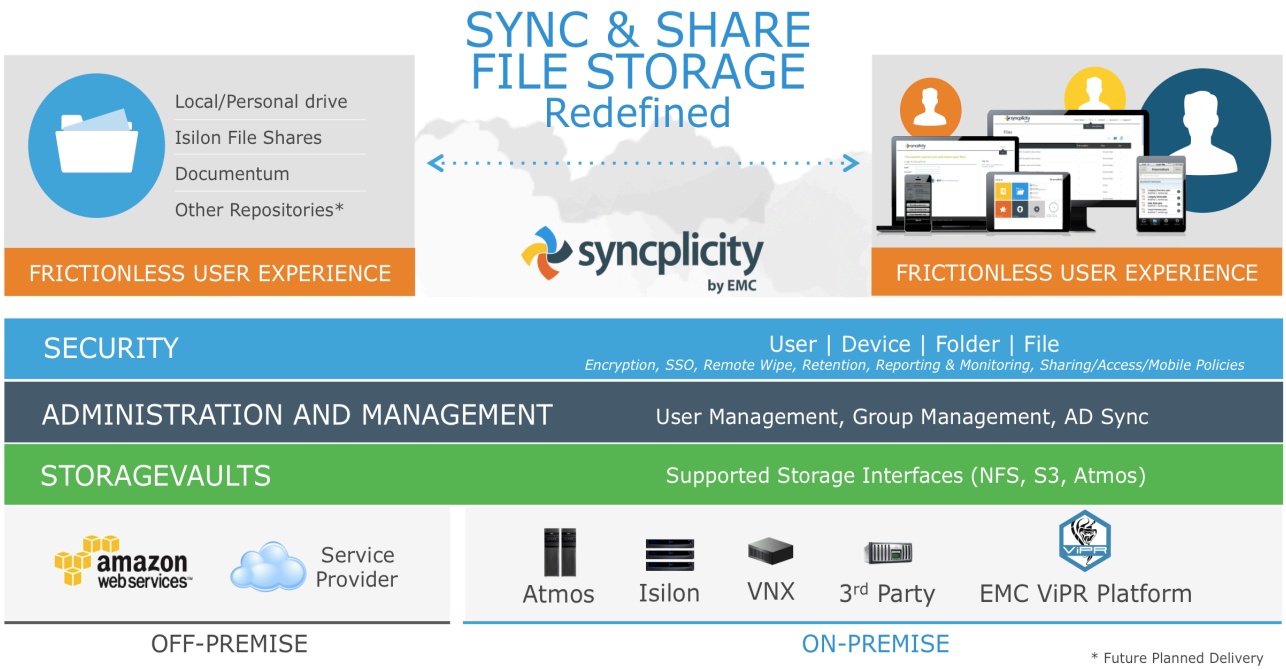

Figure 1. Syncplicity Functional Architecture

Key components of the Syncplicity functional architecture are shown in Figure 1 and include:

- **A "frictionless" user experience.** Syncplicity allows users to easily access and share files from all of their devices with a highly-optimized and native user experience on every major device platform. This is critical for driving end-user adoption and improving organizational security by reducing dependence on email attachments and consumer-grade online file sharing solutions.

- **A comprehensive set of security features and controls.** Security and controls at the user, device, folder, and file level give IT the tools and infrastructure integration to deploy the solution with confidence and maintain control of and visibility into large-scale file sharing.

- **Enterprise-grade administration and control features.** Administration, support, and reporting features give IT the tools they need to deploy and support Syncplicity at scale.

- **Syncplicity StorageVaults provide flexibility to ensure security and compliance.** Syncplicity StorageVaults provide a policy-driven hybrid cloud that gives IT the control it needs over data storage and residency to meet internal and industry regulations for file handling and data residency. Using StorageVaults, organizations can configure Syncplicity to store file versions and history in multiple storage arrays at the same time, based on user, group, and folder policies.

Syncplicity Architecture

The Syncplicity architecture includes three primary components:

- A cloud-based **orchestration layer** that controls the sync process, enabling sharing of files and folders between users and devices. This is a multi-tenant cloud-based service that is common across all Syncplicity customers.

- **Compute nodes** control where files (and past file versions) are stored in Syncplicity. This layer is single-tenant for customers that choose on premise storage and is multi-tenant for customers using Syncplicity's public cloud storage.

- The **storage layer** is the actual physical storage that the compute nodes point to and where files are actually stored.

IMPORTANT: When a user or device needs to receive a file, the file is sent directly from the storage and compute layers to the device, not through the orchestration layer.

Figure 2 illustrates these components and how data/files flows between them.

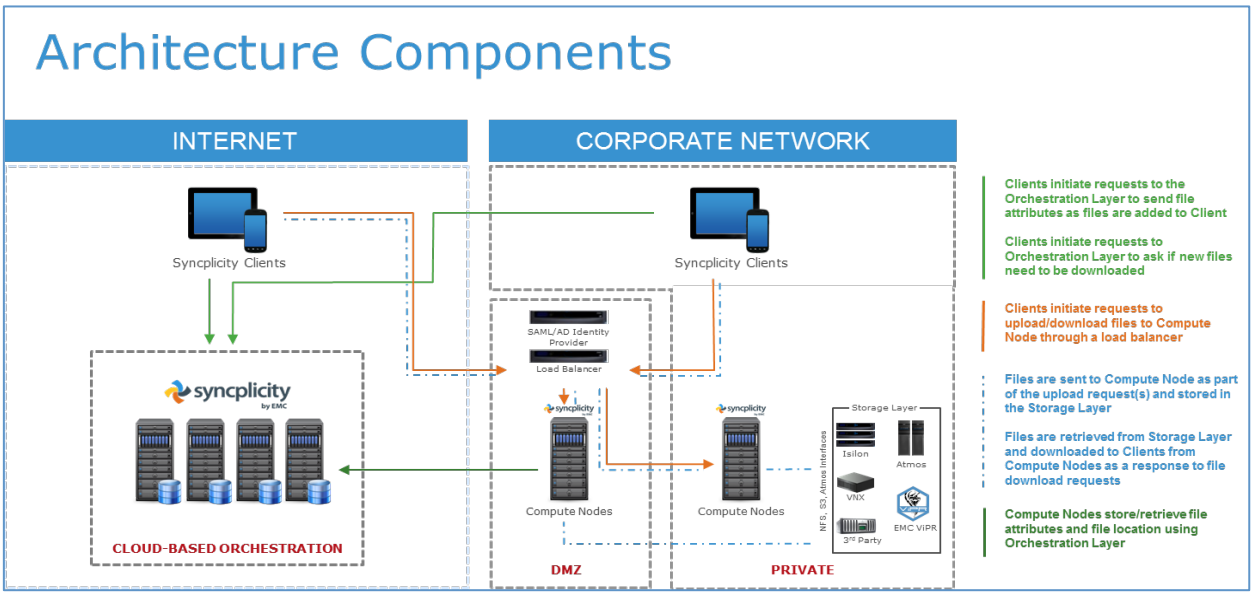

Figure 2. Syncplicity Architecture

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 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.

Data stored in the orchestration layer is minimal, and includes:

- File name, creation date, size, SHA-256 hash
- Storage information (# of chunks, chunk size, encryption key, data length after encryption/compression)

- Virtual path relative to the sync point root
- Full path to where synched and shared folders are mapped on the user's machines

Information such as file size, SHA-256 hash, and encryption keys are stored in separate databases from all user-specific information.

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.

Storage

The storage component (StorageVaults) 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 storage array that is optimized for enterprise sync and share functionality.

Syncplicity StorageVaults give IT departments more control and flexibility around where their data is stored – using a policy-driven hybrid model to store content in a public cloud or private cloud on-premise based on user, group, folder file, or content type. With Syncplicity StorageVaults enterprise, customers can:

- Comply with data residency/sovereignty requirements
- Apply appropriate security and control for each type of content
- Set policies at the user, group or folder level
- Tiered "policy sets" and "storage sets" allow organizations to gain maximum control (e.g., marketing content may go in to the public cloud, but German user content, even for marketing users, could go to a Isilon array in Germany)
- Leverage EMC storage, EMC VIPR, and third-party storage solutions (utilizing NFSv3) for unparalleled flexibility
- Manage compliance in highly regulated industries

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:

- 8 GB of RAM
- 8 virtual cores, Intel Xeon E5 Family processors, 2.20 GHz
- 25 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.

On Premise Compute and Storage Deployment

The compute application can be deployed in three possible network configurations, each with its own advantages and disadvantages (illustrated together in Figure 2). 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 NFSv3 (Isilon, VNX, or third-party storage solutions) or HTTPS (Atmos).

Deployment Option 1: Compute in DMZ (Recommended)

The recommended option is to deploy the compute node(s) application instances 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 EMC or 3rd partystorage, situated on-premises, within IT control and protection policies.

Deployment Option 2: Compute Behind Firewall

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, introduces another layer of protection, privacy, and security, it comes at the expense of user experience and extended enterprise sharing and collaboration.

Deployment Option 3: Hybrid DMZ/Firewall Configuration

The final deployment option is to use a hybrid of compute nodes in the DMZ and behind the firewall. In this case a load balancer (discussed below) can be used to direct traffic to compute nodes behind the firewall if a user is on the corporate network. Likewise, users outside the corporate network would access compute nodes in the DMZ to avoid any degradation of user experience.

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 VNX or 3rd party NAS storage or port 443 (HTTPS) with Atmos

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

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 Enterprise Edition account. The administrator may use the administration console 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 Orchestration) 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 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 Enterprise 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 Enterprise 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 nodes from a standardized VMware .ova

template. Subsequently, an administrator will add the new VM instance to the pool of available instances either via Syncplicity, DNS, or an on-premise load balancer.

Scalability and EMC Isilon

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 3 node configuration with 18 TB of capacity to a 144 node cluster configuration with over 20 petabytes of capacity.

Scalability and EMC Atmos

Atmos provides a highly scalable, object-based cloud storage platform to store, archive and access unstructured content as scale. It offers unimaginable room for growth of your online file sharing storage solution. EMC Atmos is architected for cloud and provides Enterprises and Services Providers with operational efficiencies at scale, instant access from any device, and options to deliver public, private and hybrid clouds.

Scalability and EMC VNX

VNX provides a highly scalable, modular storage architecture that can grow easily with the needs of your business. VNX scale-up NAS offers file-based storage to accommodate your online file sharing storage solution. For added efficiency and business optimization, since Syncplicity only leverages the file-based NFS interfaces for VNX, a system can be configured for Unified (block and file) storage access thereby simultaneously supporting other block-based workloads.

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's on-premise storage solution 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, Atmos, and VNX storage. 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 and the available deployment options, please visit www.emc.com or www.syncplicity.com Request a Quote in the EMC Store: https://store.emc.com/syncplicity