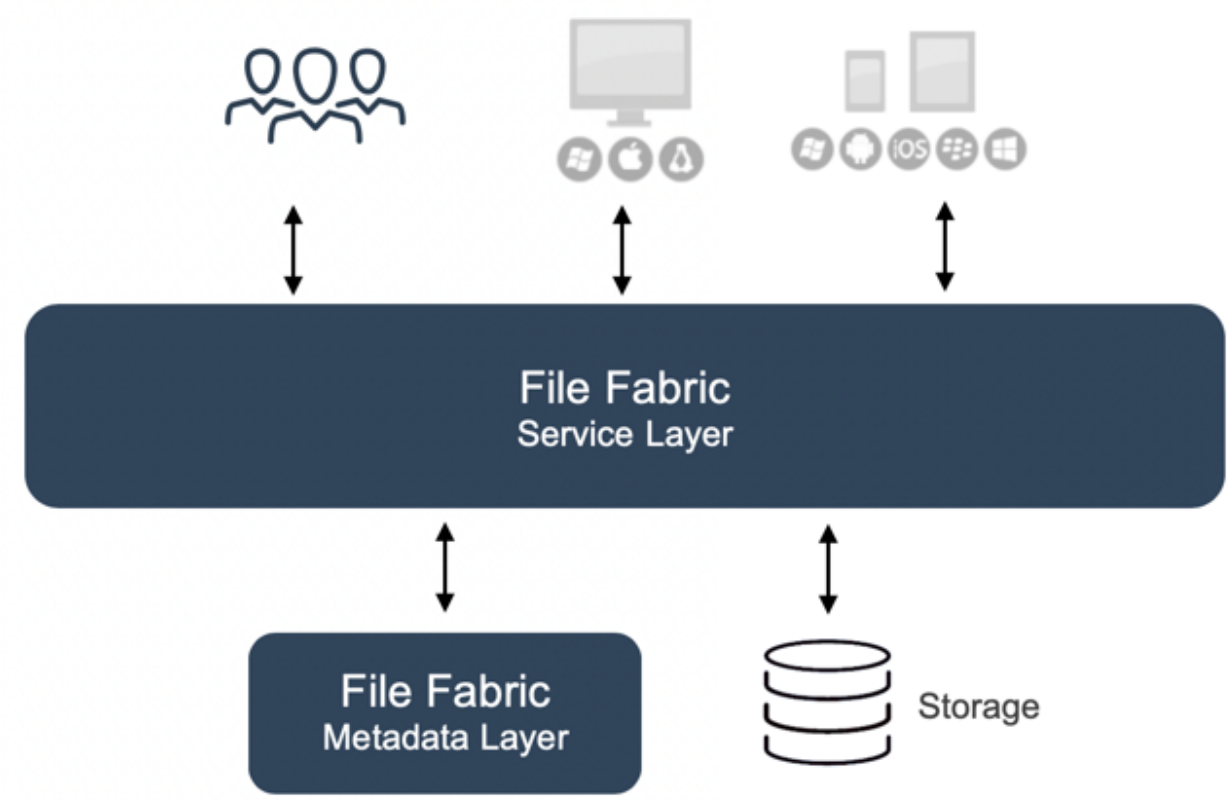
# File Fabric Components

The File Fabric can be logically divided into 2 layers of services.




## File Fabric Service Layer

The Service Layer provides client/frontend access services for end users to the following File Fabric features:

- File Fabric Mobile Clients (iOS, Android, Blackberry, Windows Phone)
- File Fabric Desktop Client (CloudEdit, Drives, Mac, Linux & Windows, Mail & Office Plugins)
- Shared links
- Online editing (Microsoft Office Online, Zoho)

The Service Layer includes the following services:

- File Fabric REST API
- File Fabric Web Application

- CloudFTP (FTP, FTP w/TLS, SFTP, FTPS)
- CloudDav (WebDAV)
- Antivirus Scanning
- Auditing
- Storage Provider Connectors

The Service Layer supports over 60 backend storage providers and applications. The service layer is stateless. It can be scaled out as necessary for high availability purposes or to increase throughput and concurrency.

## File Fabric Metadata Layer

The Metadata Layer is composed of stateful data services that record and map access to user data. These services include the following:

- MySQL Compatible Database
- Memcache
- Redis
- Content Search (Apache Solr)

Metadata Layer Services can be deployed on separate instances (decoupled) from the service layer. This is most often done to create an Highly Available (HA) deployment with no single point of failure. Separating the Metadata Services can also increase deployment security in the case that a load balancer or proxy is not available to be deployed in a DMZ. In this case the Services Layer will be deployed to the DMZ and the Metadata services to a more secure internal network.

# Overview / Sizing

The Storage Made Easy® (SME) File Fabric™ is a scale out solution for data governance, secure file access/sharing and collaboration. The File Fabric can be be deployed in a number of environments, both on-premises and in the cloud. This guide is designed to assist users with the proper sizing of servers to run the Enterprise File Fabric.

## Testing / Running Proof of Concept (POC)

During most Proof of Concept (POC) implementations, it is best to run all services on a single appliance for simple deployment and testing. When moving to production, a Storage Made Easy architect can aid in design and deployment.

The Enterprise File Fabric is available in the following formats:

- OVA for VMware Environments
- QCOW2 for KVM / Openstack Environments
- VHD for Microsoft Hyper-V and Azure Cloud
- AMI for AWS Marketplace
- Google Marketplace
- Azure Marketplace

For POC's we recommend the following resources:

- 4 vCPUs
- 6 GB RAM
- 60GB for OS Volume
- 100GB for DB Volume on SSD or performant storage
- Static IP, DNS registered FQDN, SSL certificate

# File Fabric Production Deployments

File Fabric components can be collocated for small installations or separated into multiple instances for larger instances or HA deployments.

## All-in-One Deployments:

All services running on a single instance are most appropriate for smaller installations. In this topology HA can still be configured or it can be provided by the Hypervisor or Backup/Replication system.

Recommended resource allocations:

- 6 vCPUs
    - Max 12vCPUs (if using M-Stream™ accelerated transfers)
- 16 GB RAM
- 36GB for OS Volume
- 40GB for MySQL on SSD or performant storage
    - Add 1GB of storage per 1 Million indexed records over 20 Million
- 40GB for Content Search (if enabled) on SSD or performant storage
    - Add 1GB of storage per 1 Million indexed records over 20 Million
- Static IP, DNS registered FQDN, SSL certificate

## Separate Instances:

The Enterprise File Fabric can also be deployed as separate instances for Web Services and Metadata Services.

## File Fabric Service Layer:

The Service Layer may be deployed as separate instances to accommodate more users or for High Availability. A stateful web load balancer will be required to distribute user load between Service Instances.

Sizing for these nodes is based on the number of concurrent logged in users If all users will not be actively accessing resources, architect for peak concurrent active users.
The sizing guidelines below are for typical office workers averaging 2-4 open/save/update file requests per hour.

Recommended resource allocations:

- vCPUs
    - 6 vCPUs for 100-1000 Concurrent Users
    - 8 vCPUs for 1000-2000 Concurrent Users
    - 10 vCPUs for 2000-3000 Concurrent Users
    - 12 vCPUs up to 5000 Concurrent Users or if using M-Stream accelerated transfers with files over 10GB
- 8 GB RAM
- 60GB for OS Volume
- Static IP, DNS registered FQDN, SSL certificate

## File Fabric Metadata Layer:

The Metadata Layer may be deployed as a separate instance to accommodate more users or for High Availability. When deploying for High Availability, follow the Guides available for High Availability deployment.

Sizing for these nodes is based on the total number of Indexed files across all storage providers.

### MySQL

Typical MySQL Disk Usage:

| # of Files | Database Size | CPUs | RAM | Disk Partition Size |
|---|---|---|---|---|
| Up to 5 Million | 3GB | 2 | 4GB | 40GB |
| 50 Million | 30GB | 4 | 8GB | 60GB |
| 100 Million | 60GB | 4 | 16GB | 100GB |
| 250 Million | 150GB | 6 | 32GB | 200GB |
| 500 Million | 300GB | 6 | 64GB | 500GB |
| 1 Billion | 600GB | 8 | 128GB | 1TB |

- 50GB for OS Volume
- **Database partition should run on SSD or high IOPS Storage. Default DB partition size is 40 GB**

## Apache Solr

Sizing for deep search depends on the number of user files that will be indexed and the total size of those files which qualify for indexing.

Solr indexing of Word, Excel, Powerpoint, PDFs generally utilizes more space than pictures, videos, zipped files. Space requirements can vary greatly and may be 1KB to 128KB per indexed file. Not all file types are indexed, and a limit can be placed on the max size of a file to be indexed.

| # of Index Files | Total Size of Index Files | CPU | RAM | Disk |
|---|---|---|---|---|
| Up to 5 Million | Up to 1TB | 4 | 8GB | 50GB |
| 50 Million | 10TB | 8 | 16GB | 500GB |
| 100 Million | 20TB | 8 | 32GB | 1TB |
| 250 Million | 50TB | 8 | 64GB | 3TB |
| 500 Million | 500TB | 8 | 128GB | 5TB |

- 36GB for OS Volume
- **Solr Indexes should run on SSD or high IOPS Storage**