FileCloud on AWS – High Availability Architecture

FileCloud hosted on Amazon infrastructure provides the best of two worlds: the complete control, flexibility, data separation and customization of FileCloud; and the scalability, resiliency of the AWS infrastructure

WWW.GETFILECLOUD.COM

Note: This white paper is intended to provide an overview and does not constitute legal advice. For more comprehensive information on regulations and their implications, please consult your legal counsel.

Copyright Notice

Any questions regarding this document should be forwarded to:

CodeLathe Technologies, Inc.
13785 Research Blvd, Suite 125
Austin TX 78750, USA

Phone: U.S: +1 (888) 571-6480
Fax: +1 (866) 824-9584
Email: support@codelathe.com

Introduction

FileCloud enables organizations to run their own file share, sync, and remote access solution. With FileCloud, organizations can have dedicated instance customized and configured to meet their requirements and regulations. Key advantages of using FileCloud over any public SaaS solution include:

- **Control:** Provides useful management control options for system administrators to manage and secure organization data.

- **Customization:** Unlike public cloud solutions, such as Dropbox, FileCloud offers unparalleled branding and customization options to broadcast customer's brand.

- **Integration:** Integrates easily with existing corporate IT systems (Active Directory, NTFS File permissions, ADFS, NTLM SSO) and third party software, such as Office.

- **Value:** Delivers higher value for customers by providing simple pricing, product
- innovation, and technical customer support.

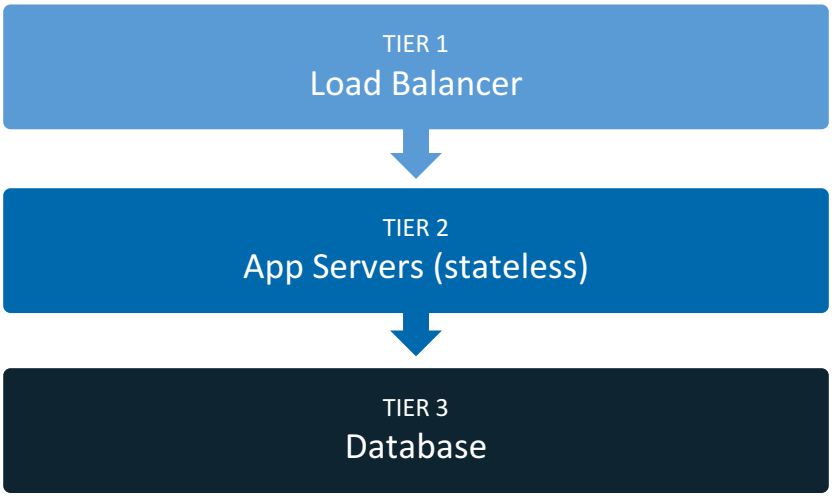
FileCloud is used by over 1000 organizations across 90 countries, including world leading government organizations, Fortune 500 enterprises, universities, and research organizations. FileCloud for AWS will allow businesses to host FileCloud in the Amazon cloud infrastructure. Customers will experience the flexibility and total control of FileCloud and the scalability and reliability of AWS.

Using AWS infrastructure (EC2, EBS, S3) customers can jumpstart their own branded, file storage solution in a few minutes at a compelling price point. With FileCloud's AMI, organizations can host FileCloud for their organization in less than 10 minutes.

The FileCloud system is versatile and can be deployed on commodity physical servers or on a virtualized environment (VMware, XEN). The FileCloud system can also use any NFS, CIFS, SAN appliance, or s3 compatible object storage systems for file storage. This paper explains the requirements for implementing FileCloud using high availability architecture on AWS.

High Availability Architecture

FileCloud solution can be implemented using the classic three tier high availability architecture. The first tier will comprise the load balancer and access control services. Tier 1 will be a web tier made of load balancers. Tier 2 will be stateless application servers, and for FileCloud implementation, this layer will comprise Apache nodes and helper services. Tier 3 will be the database layer. The advantage of this architecture is separation of stateless components from state full components, allowing great flexibility in deploying the solution.



Tier 1 – Web Tier

Tier 1 is the front end of the deployment and act as the entry point to all external clients. The components in Tier 1 are stateless and primarily forward the request to the webservers in Tier 2.

Scaling of the web tier can be done by adding and removing load balancer instances, since they are stateless. Each webserver node can handle any request.

This layer can also be configured to do SSL offloading, allowing lighter weight communication between Tier1 to Tier2.

This layer can also be configured to provide simple affinity based on source and destination addresses. The traffic will be forwarded to healthy application server nodes.

This layer also monitors available application servers and will automatically distribute the traffic depending on the load. Customers can use Amazon's Elastic load balancing or HA proxy for load balancing.

Tier 2 – Application Servers

Tier 2 in FileCloud deployment comprises the following services:

- Apache servers
- FileCloud helper
- Antivirus service
- Memcache service
- Open Office service

The apache servers in FileCloud store no state information and are therefore stateless. However, they cache data for faster performance (such as convert and cache documents for display). They primarily execute application code to service a request. All state specific data is stored in database tables and, therefore, are stateless.

If an application server node fails, the request can be handled by a different application server node (provided the clients retry the failing request). Capacity can be increased or reduced (automatically or manually) by adding or removing apache server nodes.

FileCloud helper service provides additional capabilities, such as indexed search, NTFS permission retrieval etc. FileCloud Helper is a stateless service and, therefore, can be added or removed as needed.

Similar to FileCloud helper service, the Antivirus service is also a stateless service, providing antivirus capability to FileCloud. Any file uploaded to Filecloud is scanned using this service.

Memcache service is an optional service required for local storage encryption. This service is also stateless and is required only if local storage encryption is required. This service is also started in the same node as the Apache service.

Open office service is an optional service required for creating document file previews in the browser. This server is stateless and is started in the same node as the Apache server.

Tier 3 – Database Nodes

Tier 3 comprises state full services. This comprises the following services:

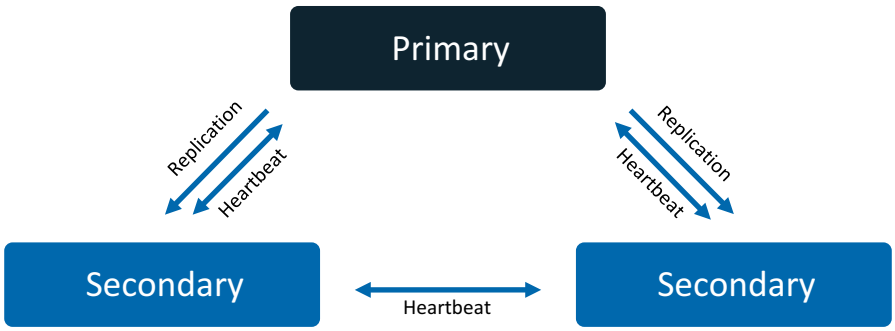
- MongoDB servers
- Solr Servers

The High availability for each server varies, depending on the complexity of the deployment. The failure of these services can have limited or system wide impact. For example, MongoDB server failure will cause FileCloud solution wide failure and is critical, while FileCloud helper server will only impact a portion of function, such as network folder access etc.

MongoDB server High Availability

MongoDB servers store all application data in FileCloud and provide High Availability using replica sets. The MongoDB replica set configuration provides redundancy and increases data availability by keeping multiple copies of data on different database services. Replication also provides fault tolerance against losing a single database server. It is also possible to configure Mongo DB to increase the read capacity.

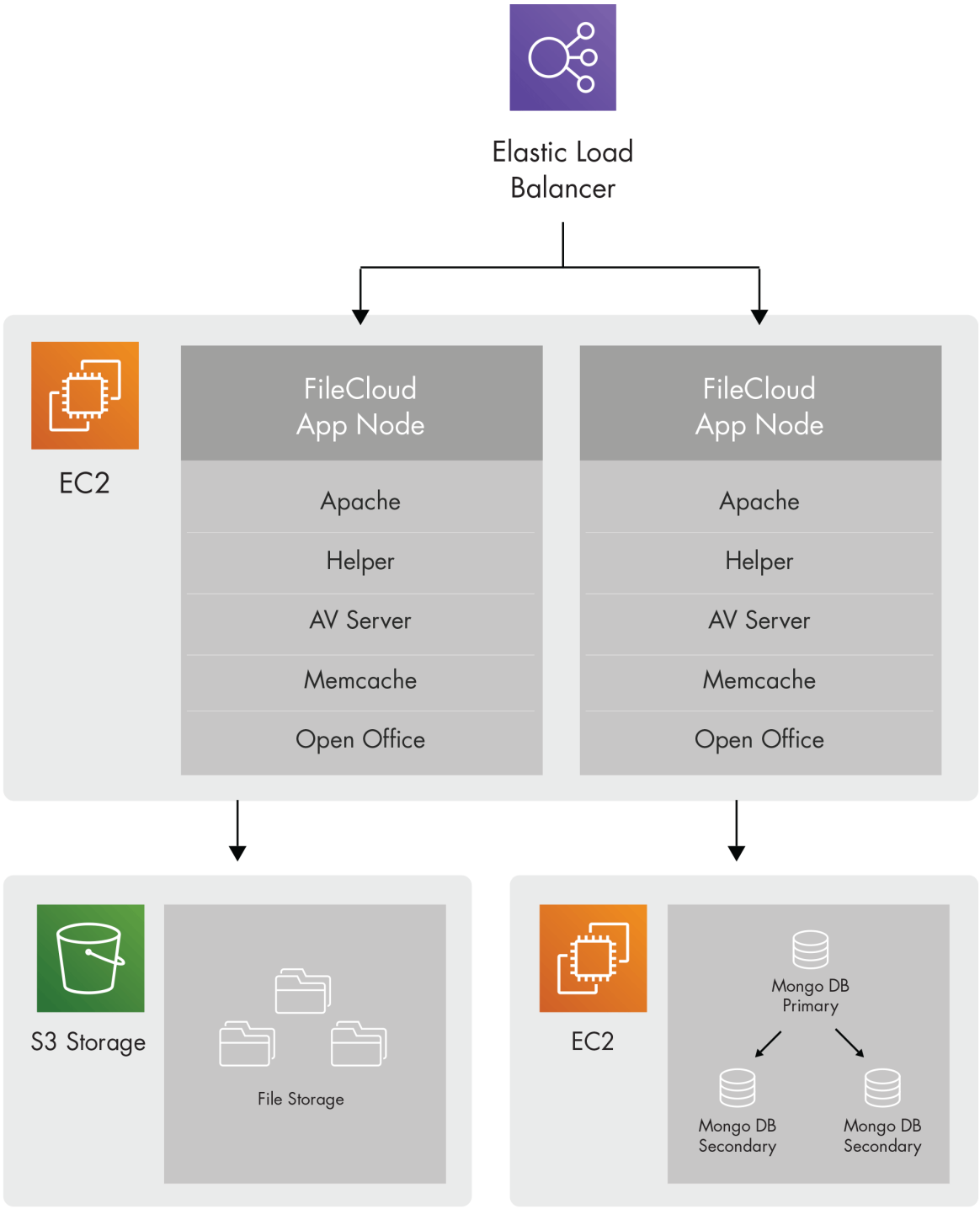
The minimum number of nodes needed for Mongo DBserver HA is a 3 node member set (It is possible to also use 2 nodes + 1 arbiter). In case of primary Mongo DB server node failure, one of the secondary nodes will failover and will become primary.



The heartbeat time frame can be tuned, depending on system latency. It is also possible to setupthe Mongo DB replica to allow reads from secondary to improve read capacity.

Putting It All Together

The three-tier structure for the FileCloud component is shown below. The actual configuration information is available in FileCloud support. This provides a robust FileCloud implementation with high availability and extensibility. As new services are added to extended functionality, the layer can be decided whether they are stateless or store state. The Stateless (Tier 2) nodes can be added or removed without disrupting service. Tier 3 nodes will store state and require specific implementation, depending on the type of service.



Winner of 2018 Gartner Peer Insights Customer Choice Award

CodeLathe's FileCloud is recognized in the July 2018 Gartner's Magic Quadrant for Content Collaboration Platforms

Mentioned in Forrester Wave, Enterprise File Sync and Share Platforms Hybrid Solutions, 2017

One of the Top 10 Hot Storage Companies to Watch

About Us

A privately held software company, headquartered in Austin, Texas, USA. Our company offers two products – Tonido for consumers and FileCloud for businesses – used by millions of customers around the world, ranging from individuals to Global 2000 enterprises, educational institutions, and government organizations, and managed service providers.

1M+ USERS

3000+ ENTERPRISES

100+ RESELLERS

90+ COUNTRIES

Headquarters:
13785 Research Blvd, Suite 125 Austin TX 78750

Email:
sales@codelathe.com

Website:
https://www.getfilecloud.com/aws

Phone:
+1 (888) 571-6480

Fax:
+1 (866) 824-9584