# MULE ESB – High Availability (HA) CLUSTERING

## Availability, Reliability and Scalability

Abstract: Mule ESB offers a built-in active-active High Availability clustering capability. For applications that require an enhanced level of availability, this capability is a must. HA Clustering improves the availability, reliability and scalability of Mule applications. HA is most appropriate for large and mission-critical deployments. Learn more about Mule's HA Clustering, how it works, and its advantages.

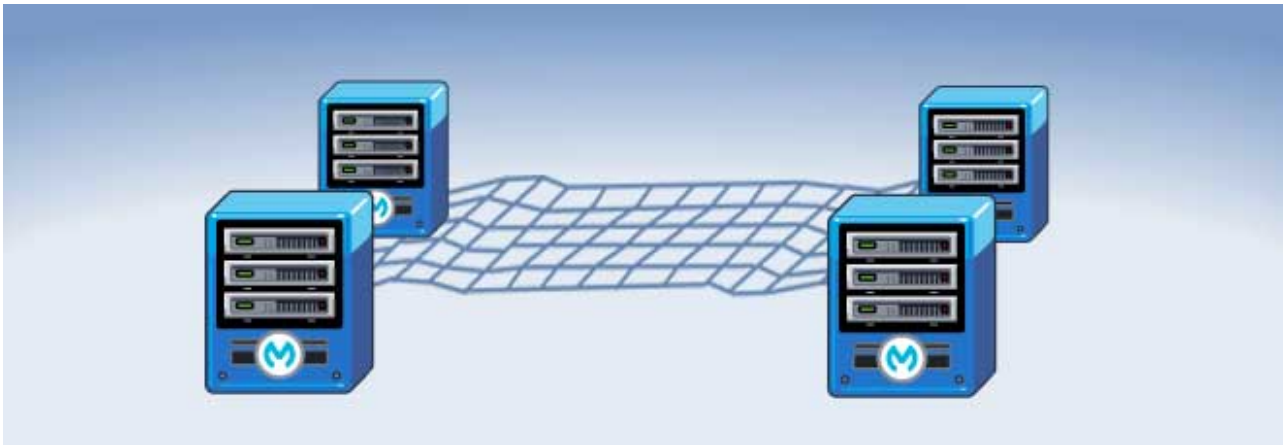
**www.mulesoft.com | info@mulesoft.com | 1-877-MULE-OSS**

# 1 Mule HA Clustering Overview

Mule has grown to be one of the most popular ESB platforms among developers due to its open source accessibility, lightweight, simplicity, and robust integration capabilities. It is used in over 2,500 production deployments by leading organizations such as Walmart.com, Nokia, Nestlé, Honeywell and DHL, as well as 85 in the Global 500 and 5 of the world's top 10 banks. It powers mission-critical application responsible for massive revenue streams for organizations ranging from major airlines to global banks.

For critical applications, maintaining availability of the application is paramount. While the core Mule ESB with its robust and scalable architecture is inherently reliable, problems can still arise. These not only include issues with Mule itself but also the underlying JVM and hardware. A failure of any of these can cause a downtime event. In mission critical applications the cost of such outages is exceedingly high as it implies lost revenue, productivity loses for business and IT, regulatory fines, penalties and/or damages to customer satisfaction.

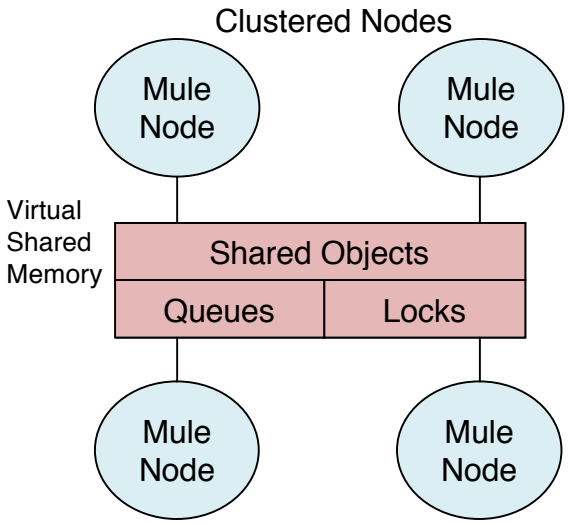
For this reason Mule ESB can be clustered for High Availability (HA). This is ideal not only for maintaining the availability of the total Mule environment, but also for making applications more reliable and more scalable. This paper provides an overview of Mule HA Clustering, how it works, how it is licensed and for what sorts of environments it is appropriate.



# 2 How Mule High Availability (HA) Clustering works

The Mule HA solution is based on clustering of multiple Mule nodes. The nodes of a cluster act as a logical unit handling a common load. Should a node fail, any other node within the cluster is capable of take over the fail node's workload. This type of architecture is referred to as an active-active approach as all nodes are actively processing information at all times. Unlike many other approaches, the Mule HA solution does not require an application server to achieve high availability. Many traditional HA approaches relied on the scalability and reliability of an application server.

In the Mule HA solution, the cluster nodes are typically run on multiple physical servers. The nodes communicate and share information through a distributed shared memory data grid. This allow high messages throughput while maintaining availability. If one machine has a downtime event or a performance issue another can take automatically take over the load.




To deliver a fully reliable solution, Mule's high availability solution has several inherent capabilities. Mule clusters enable automatic coordination of access to share resources such as files, FTP, and JDBC. It also load balances processing within the cluster. Should a failover event occur, it automatically redirects message processing to one of the other nodes in the cluster.

Mule clusters are also easy to manage and maintain. Mule offers alerts on node failure, performance monitoring of both nodes and entire clusters, and cluster lifecycle management and control. Mule's clusters are managed through the familiar Mule management tools with which Mule administrators and developer are already familiar.

# 3 Increased reliability and availability

High Availability Clustering provides reliability and increases availability. It is important to understand the distinction between availability and reliability. While interrelated concepts, availability is actually a prerequisite for reliability. Availability refers to maintaining uptime on the ESB itself. By ensuring that if a node becomes unavailable due to failure or planned downtime, another node in the cluster can assume the workload and continue to process existing events and messages, Mule clustering can guarantee the reliability of a Mule ESB environment.

Reliability refers to the correct operation of the individual applications running within Mule. Mule clustering is a key component of a reliable and fault tolerant infrastructure. However, reliability implies an overall design that anticipates all possible issues that may arise in production. When building critical applications the entire architecture must be designed with fault tolerance in mind.

To achieve reliability, the individual components within Mule must be designed to be fault-tolerant as well. For transports such as JMS, VM and JDBC, transactions are used. For non-transactional transports, such as File, FTP, HTTP, and TCP, Mule provides Reliability Patterns that ensures that either that a message is accepted and successfully processed or that it generates an "unsuccessful" response allowing the client to retry. However, for either of these approaches to be successful the underlying ESB needs to remain available. This is why clustering is a prerequisite for any reliable architecture.

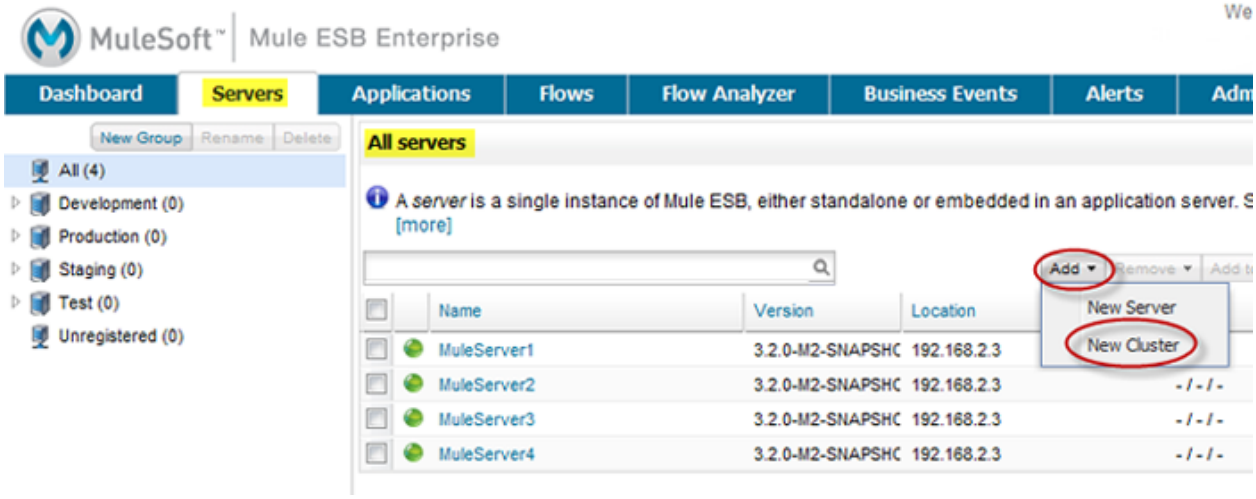
# 4 Performance and scalability implications

Another benefit of HA Clustering is improved performance and scalability. In a non-clustered environment only a single server handles load for any given application. While load balancing can be utilized for balancing single HTTP or TCP endpoints, load balancing alone fails to address processing within the Mule application itself.

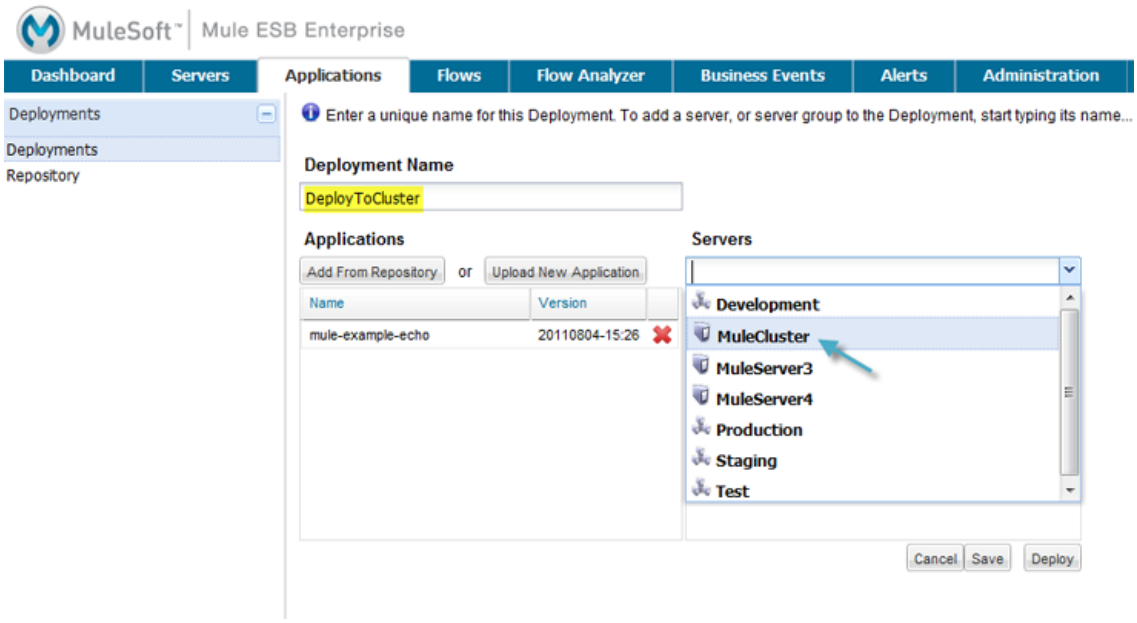
Clusters share the workload across multiple nodes and increase flexibility by allowing additional nodes to be easily added or removed as required. In this sense HA Clusters are ideal for applications that require flexible scaling to meet variations or spikes in transaction volumes.

# 5 Managing clusters

Mule's Enterprise management console is used to fully manage all clustered resources. It allows you to create clusters and add or remove additional nodes to a cluster. You are able to easily create a cluster, manage the cluster, monitor the cluster and deploy your applications to the cluster directly from the management console. This simplifies the installation and management of a Highly Available solution greatly.




Once a HA Cluster has been created, within Mule's management tools, it appears as a single logical entity. Using this feature an application can be deployed with one click to all nodes in the cluster, or easily rolled-back using the consoles Application Deployment repository.




# 6 HA Clusters and Load Balancing

When Mule HA Clusters are used to serve socket-based IP based requests, which includes SSL, TCP, UDP, HTTP, and HTTPS, load balancing is needed to distribute incoming requests among the available clustered instances. Messaged based requests do not require any external load balancing as their architecture inherently shares load between multiple loads. Mule clusters automatically coordinate

access to resource based such as JDBC, File and FTP. Within the cluster, Mule automatic load balances processing. Therefore, load balancing is used exclusively for use with incoming socket-based connections.

There are various software and hardware load balancing alternatives that can be used for this purpose. Some alternatives include, but are not limited to the following products;

Software Load balancers:

- Nginx
- HA Proxy

Hardware Load balancers:

- Big IP “F5” load balancer
- Citrix Netscaler

While load balancing is a necessary component of a highly available environment, Mule does not enforce any restrictions on which product(s) are chosen.

# 7 Licensing HA Clustering

Mule’s HA Clustering capabilities are available in the Mule Enterprise Edition. In order to take advantage of these capabilities, each node within the cluster must be licensed with a Mule Platinum subscription.

# 8 Alternative Approaches

In the past some customers with unique high availability needs chose to run Mule ESB embedded within an application server. Mule fully supports such a configuration. However, the exact feature set and availability design can differ substantially from one application server to another. Some customers have reported having to purchase further software to deliver true reliability. Often custom development is needed to fully support a high availability environment when run in such a configuration. This makes this approach quite expensive and difficult to maintain.

Running Mule in an application server is a fairly heavyweight approach to dealing with availability requirements and has several disadvantages. On top of requiring an additional product, when deployed on an application server, Mule users are no longer able to take advantage of hot deployment and server management (start/stop/restart) capabilities of Mule. Management of the servers and high availability must be managed separately outside of Mule. This makes this approach much more difficult than when Mule is run standalone. For these reasons, in most scenarios running Mule in standalone mode is the recommended configuration.

Mule HA Clustering offers superior total cost of ownership and superior performance to many alternative approaches and is a fully supported configuration that is backed by the Mulesoft the company behind the worlds most popular Open Source ESB..

# 9 Summary

As the leading open source enterprise service bus, Mule ESB is stable, robust, and mature. Mission critical applications that require an enhanced level of availability can utilize Mule's active-active HA Clustering capabilities to maximize uptime in their environments. HA Clustering improves the availability, reliability and scalability of Mule applications.

### About MuleSoft

MuleSoft provides the most widely used integration platform for connecting SaaS and enterprise applications in the cloud and on-premise. With the rise of cloud and mobile, enterprises face a choice: become overwhelmed by the resulting explosion of end points or seize the opportunity to gain competitive advantage. Founded on the idea that connecting applications should not be hard, MuleSoft lets organizations harness the power of their applications through integration. MuleSoft's Anypoint™ technology eliminates costly, time-intensive point-to-point integration, enabling business agility. Delivered as a packaged integration experience, CloudHub™ and Mule ESB™ are built on proven open source technology for the fastest, most reliable integration without vendor lock-in. Supporting billions of transactions per day, MuleSoft is used in production by thousands of enterprises, including Walmart, MasterCard, Nokia, Nestlé and Honeywell, and powers integrations with leading SaaS vendors such as Salesforce.com, NetSuite, Workday, Intuit and Zuora.

**For more information: www.mulesoft.com, or email info@mulesoft.com.**

**Download Mule ESB: http://www.mulesoft.com/download/**