

Ensuring High Availability



**Monitoring Clusters and
Load Balancers**

Ensuring High Availability- A White Paper

Monitoring Clusters and Load Balancers

Introduction

Enterprises employ several mechanisms to optimize the performance of their networks in order to ensure high availability. Clustering for failover (Active /Passive mode) and load-balancing (Active / Active mode) is a commonly adopted technique that supports redundancy, session or database replication, and load balancing requests across the servers in the cluster. Some networks have software or hardware load-balancers even outside the cluster to increase horizontal scalability.

Most businesses happen over the Internet and Enterprises prefer clusters or invest in load-balancers because of the fault-tolerant architecture. Critical servers and applications such as database servers, exchange servers etc are hosted in clustered environment for high availability. Further, load balancers like Big IP are plugged into the network to distribute the load on servers to address the primary need of un-interrupted availability at all times. Service interruptions, planned or unplanned, are costly affairs and are unacceptable for businesses of any size. Administrators exhaust almost every resource within their means to keep their networks happy and available.

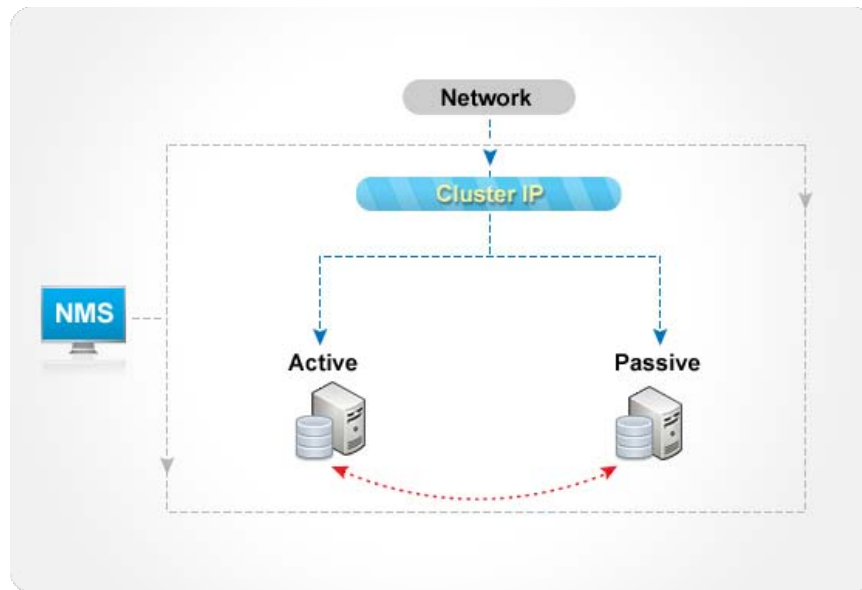
Service outage and its business impact

We are all huge consumers of various services the internet supports. Imagine a situation where you are accessing your bank account to transfer funds to a friend's for an emergency. And the site simply says, 'Oops! Sorry, the service is unavailable!'. The message concludes politely requesting you to try accessing after a little while. As end users, we would hate to be in this situation. That you get to call the customer service and bombard at random and look out for an ATM or a branch office is some respite. The damage caused to the bank is beyond economic repairs and the man behind the show, the administrator/network engineer gets to listen to variety music for this huge goof-up. This despite having invested in a cluster for a reliable service!

Forewarned is forearmed

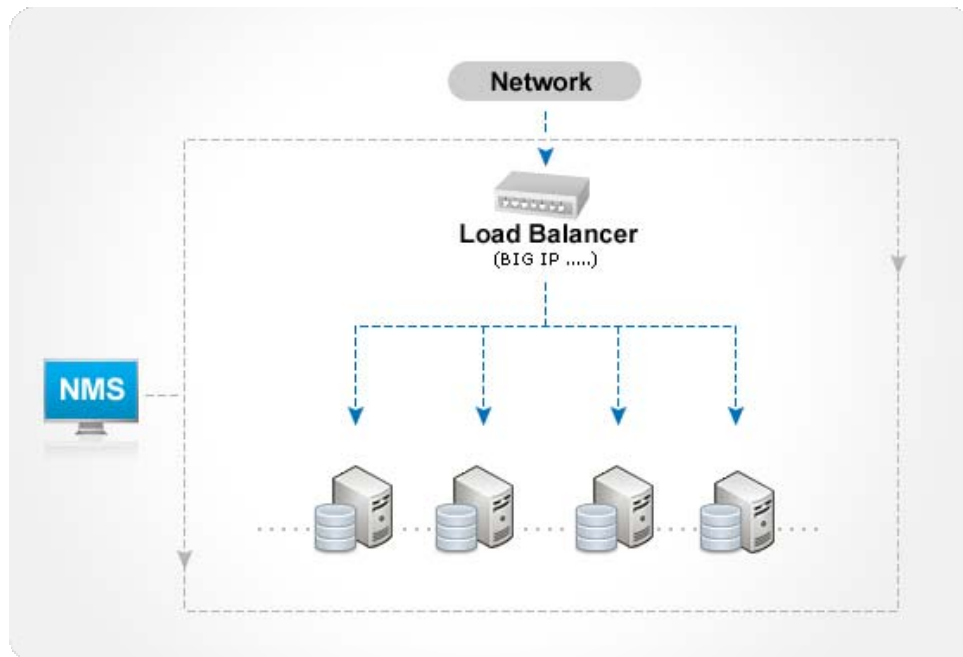
The job is only half-done by employing clustering for failover or load-balancing. Unless an administrator has a clear visibility of the good, bad, and the ugly components of his network including the ones in *that* critical cluster, or the important resources on the expensive load-balancers, its impossible to have an alert-free holiday!

Example 1: A two-node SQL cluster in an enterprise:



The reasons for an Active node to failover to a standby could be either a system or an application failure, or both. An administrator must have a clear visibility into the system and application performance, which is possible only when proactively monitored. In the scenario discussed above, it is possible that the clustering controller instance has failed causing the whole system to fall apart! Or, despite the Active node successfully failing-over to the Passive, the Passive too fails due to insufficient resources! This mix-up could have been avoided or identified a little earlier to reduce the damage by monitoring the basic resources on these systems.

Example 2: A load-balancer distributing requests across a few servers:



Despite redundant servers set up for load-balancing, imagine a hardware resource failure on the load-balancer leading to the service unavailability! The user requests never make it to the server even when the service is up and running fine!

Reduce the damage

The purpose of clustering is lost if the resources are not constantly monitored. Even as the administrator tries to ensure that the end-users do not 'feel' any service failure, he must quickly identify the cause for the failover from active to the passive node, or why the load exerted on a particular server is on the high. So, all the components or resources that need to run for the clustering to work well, needs monitoring. This includes the Cluster service on the nodes, the dependent services, the system resources on the load-balancer, response time of the individual devices etc. Contant automated monitoring of key components helps reduce the damage and helps realize the goal of ensuring high availability at all times.

The key resources include:

Availability of the nodes: A detailed availability report indicating if the node is unavailable due to a dependent device failure or if the node is pulled down for maintenance.

Name	Up	On Hold	Maintenance	Dependent Unavailable	Down	Availability(%)
Active-w2k31	1 Day	0 Sec	0 Sec	0 Sec	0 Sec	100
Cluster1	6 Hours 19 Mins	0 Sec	0 Sec	0 Sec	17 Hours 40 Mins	26.34
Standby-w2k32	1 Day	0 Sec	0 Sec	0 Sec	0 Sec	100

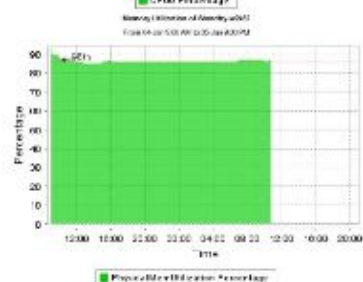
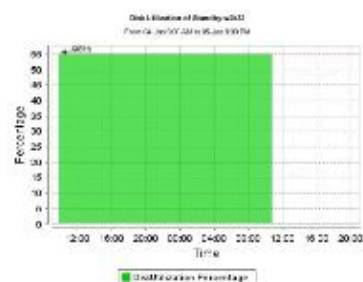
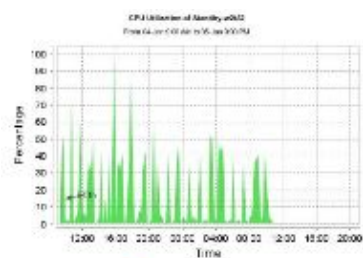
Response time of the nodes: Response time of the nodes at any given time, and its average response time indicating the load on it.

Name	Min	Max	Avg (ms)
Standby-w2k32	1	7	1.08
Active-w2k31	1	3	1.00
Cluster1	1	1	1.00

Services availability and response time: Availability of the cluster service and its related services on the nodes.

Name	Service Name	Min	Max	Avg (ms)
Standby-w2k32	DNS	1	16	2.51
Standby-w2k32	MSSQL	1	16	1.94

System Resource utilizations: A constant check on the performance of the hardware resources because the last thing you want is insufficient resources rendering a critical service unavailable!



Service Parameters: Critical parameters of a service that can lead to a potential failure.

DEFAULT instance of Standby-w2k32 (192.168.118.222)

MSSQL Server ❗ DEFAULT instance of SQL Server has 5 problem(s).

General		Memory	
Active database connections	12	Excess Total committed memory needed	21816 KB
Logins per sec	0 /Sec	Granted WorkSpace Memory	0 KB
Page Splits per sec	0 /Sec	SQL Cache Memory	1472 KB
SQL Compilations per sec	0 /Sec		

Lock Monitors		Buffer	
Number of Deadlocks per sec	0 /Sec	Buffer Cache Hit Ratio	100 %
Average Lock Wait Time	0 ms	Page reads per sec	0 /Sec
		Lazy Writes per sec	0 /Sec

Cache	
Cache Hit Ratio	82 %

Database						
Database Name	Active transactions	Database Size	Logfile Used %	Log Cache Hit Ratio	Transactions Per Sec	
VoIPReports	0	37 MB	32 %	0 %	0	
rejoefinalsanity	0	16 MB	38 %	0 %	0	
dbconfiguration	0	7 MB	33 %	0 %	0	
msdb	0	11 MB	35 %	0 %	0	
master	0	17 MB	36 %	0 %	0	











System Events pertaining to the cluster: Keeping a tab on the system events including the application events so that there are no sudden surprises and all avenues of fault are watched.

Source	Alarm Message	Status	Technician	Category	Date / Time
Standby-w2k32	ID=1053 Source=Userenv Type=1 Message=Windows cannot determine the user or computer name. (The system detected a possible attempt to compromise security. Please ensure that you can contact the server that authenticated you.) Group Policy processing aborted.	Critical	UnAssigned	Server	Jan 05,2010 03:27:36 PM

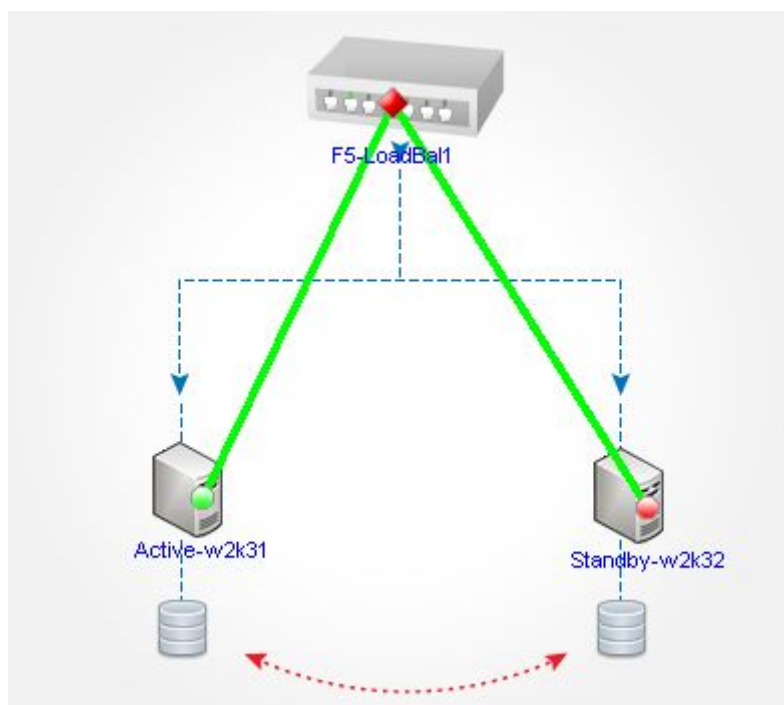
Availability and Performance of the load balancer: Ensuring the basic availability and responsiveness of the load balancer.



System resources on the load balancer: Monitoring the critical resources on the load-balancers to identify and problem indicators much ahead.

<u>ClusterMember State</u>	5	Not Enabled	
<u>Dropped Packet(s)</u>	5	Not Enabled	
<u>Established Connection(s)</u>	5	Not Enabled	
<u>Global TM PoolMember State</u>	5	Not Enabled	
<u>Global TM VirtualServer Status</u>	5	Not Enabled	
<u>HTTP Request(s)</u>	5	Not Enabled	
<u>Half Closed Connection(s)</u>	5	Not Enabled	
<u>Half Opened Connection(s)</u>	5	Not Enabled	
<u>Incoming Packet Error(s)</u>	5	Not Enabled	
<u>Local TM PoolMember state</u>	5	Not Enabled	

Cluster Groups (Business Views): A holistic view of the nodes in a cluster with an ability to drill down to the root cause. This provision to visualize a cluster helps to understand the health of the cluster at a glance.



Summary

ManageEngine OpManager is a [network monitoring software](#) that monitors all the resources on your LAN and WAN. The [performance](#) and [fault management](#) capability of OpManager helps identify performance bottlenecks quickly. Its ability to drill-down to the root cause of a fault and the huge custom-capability, makes OpManager a preferred solution among thousands of network administrators world-wide. A few useful plug-ins and add-ons such as [NCM Plug-in](#), [NetFlow Plug-in](#), [VoIP add-on](#), and the provision to easily integrate with other applications in the management suite such as [ServiceDesk Plus](#), makes it a one-stop shop for all your network and IT management needs. Visit www.opmanager.com to test drive your 30 days free trial versions of ManageEngine OpManager.