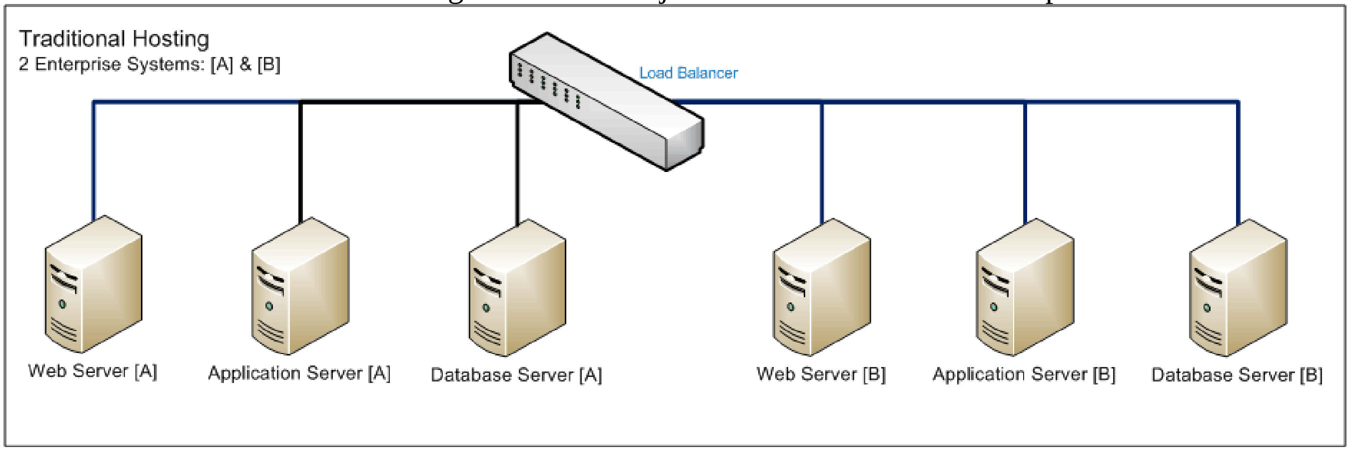
Virtualization Striping: Providing RAID-like Failover for Applications

By John C. Checco, CISSP

There is a growing dichotomy in the enterprise IT environment. It is the age old struggle between providing customers the best possible service – which translates directly into revenue – and keeping costs at a minimum. Technology has only made these issues more exciting and more complicated.

In the financial community, customer service means "always on" service. Any disruption of service could mean not only loss of revenue today, but loss of that customer altogether. So, to provide 24x7x365 accessibility usually means more redundant servers, more instances of applications in a clustered environment, and then the entire deployment is duplicated in a hot failover site for BC/DR.

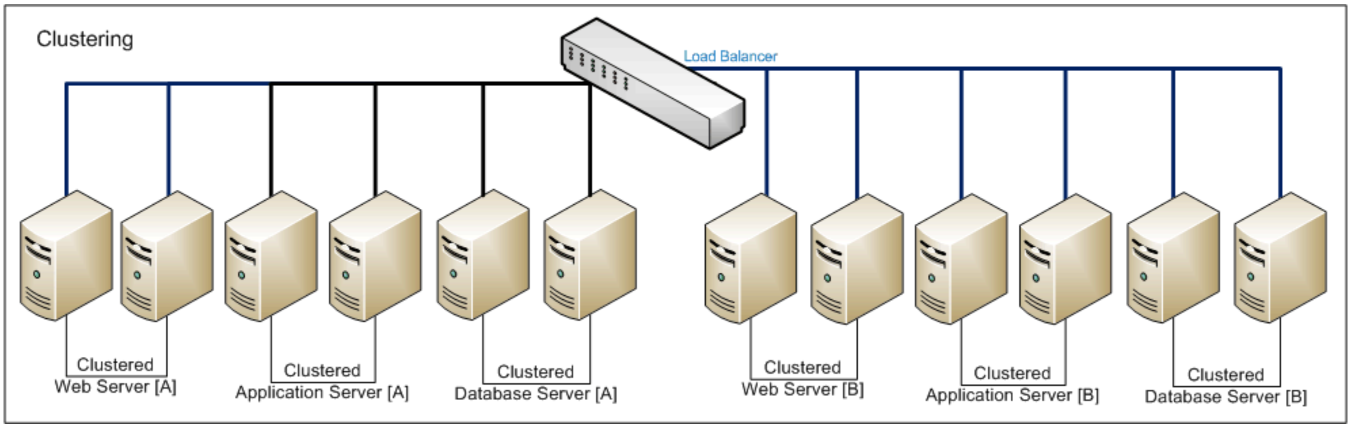
By the same token, the struggle to control datacenter costs relies on the reduction of servers and duplicate instances of applications. Power costs are no longer static overhead items; they are volatile and unpredictable because of market fluctuations in oil and electricity. As technology allows more CPUs to be placed in smaller spaces, the increased density of server racks presents cooling problems that traditional datacenter cooling solutions do not handle efficiently; so cooling costs increase whether a datacenter introduces a new cooling mechanism or just throws more BTUs at the problem.



Increasing Accessibility through Clustering

In the early days of the internet, increasing availability meant simply duplicating content to edge servers or third-party accelerators. This worked well because most web sites hosted static pages. In today's world where the web has become the thin client for many enterprise applications, content is extremely dynamic and requires replication. Whereas duplication copies data from a central location out to the edge, replication allows edge changes to the data to be integrated back into the central location.

Advanced technologies such as clustering and collaborative computing do not merely replicate business logic and enterprise data, but manage diversification of computing resources across multiple physical systems to achieve increased performance, increase uptime and optimized CPU utilization.

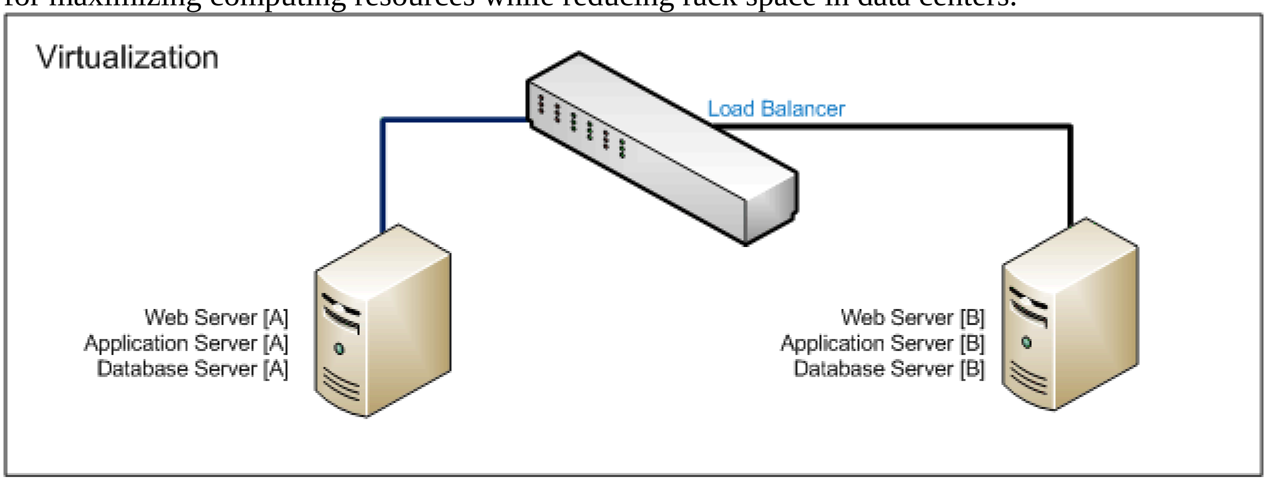


From an IT perspective, instantiating a distributed technology such as clustering into an enterprise process is no simple task. There is an enormous amount of complexity in designing (and redesigning) each application component to operate well in a distributed infrastructure. Distributed systems also introduce another level of monitoring and troubleshooting.

Despite all the complexity and obstacles in converting a linear business process to a distributed computing model, the benefits of a well-designed and well-implemented system can be measured long-term as the system grows. In addition, the development skill sets necessary for such systems is beneficial to other development projects as well.

Decreasing Costs through Virtualization

It is a well-known fact that most corporate servers use approximately 30% of the total computing resources available on a single physical machine.[1] It has also become apparent to most enterprises that the greatest costs in their datacenters are power, cooling and space. Virtualization, the process of creating many instances of an operating system on a single physical machine, has been the driving force for maximizing computing resources while reducing rack space in data centers.



With the explosive growth in virtualization efforts, there have been almost as many virtualization strategies as there are virtualization technologies: hardware virtualization, OS virtualization, application virtualization and even data virtualization. Each strategy has its own benefits and risks, so a great deal of knowledge and planning needs to done to ensure the most efficient use of resources and the least dependencies.

The common advantage among all the virtualization strategies is to reduce costs by consolidation and optimizing CPU utilization. There is also the added benefit of allowing quick duplication of system images when necessary, thereby reducing unexpected downtime. And virtualization has its benefits in areas that require dependent systems, such as rapid development environments and live CDs. Virtualization makes the best use of resources when each virtual instance is heavy on utilizing different resources. Virtualization of 20 web servers does not achieve performance increases if the first 5 instances consume all the physical network bandwidth. Also, virtualization of all similar services on one physical machine also creates a single point of failure for multiple applications. For example, putting all your application servers on one physical machine will cause all the applications to go down on one physical failure.

End-To-End Virtualization Sets

A virtualization strategy that includes a web server with its related application, database and supporting servers provides a fair mix of resources that can be optimized – and physical failures affect only a single application “silo”. In addition, by grouping related virtualized services on one physical machine, one is able to backup and restore an “end-to-end” system quickly and easily.

Like the “live CDs” that allows anyone to test an enterprise solution in a self-contained environment, the “end-to-end” virtual machine set provides a self-contained enterprise system with preset configurations, useful for disaster recovery or even quickly creating test and development environments.

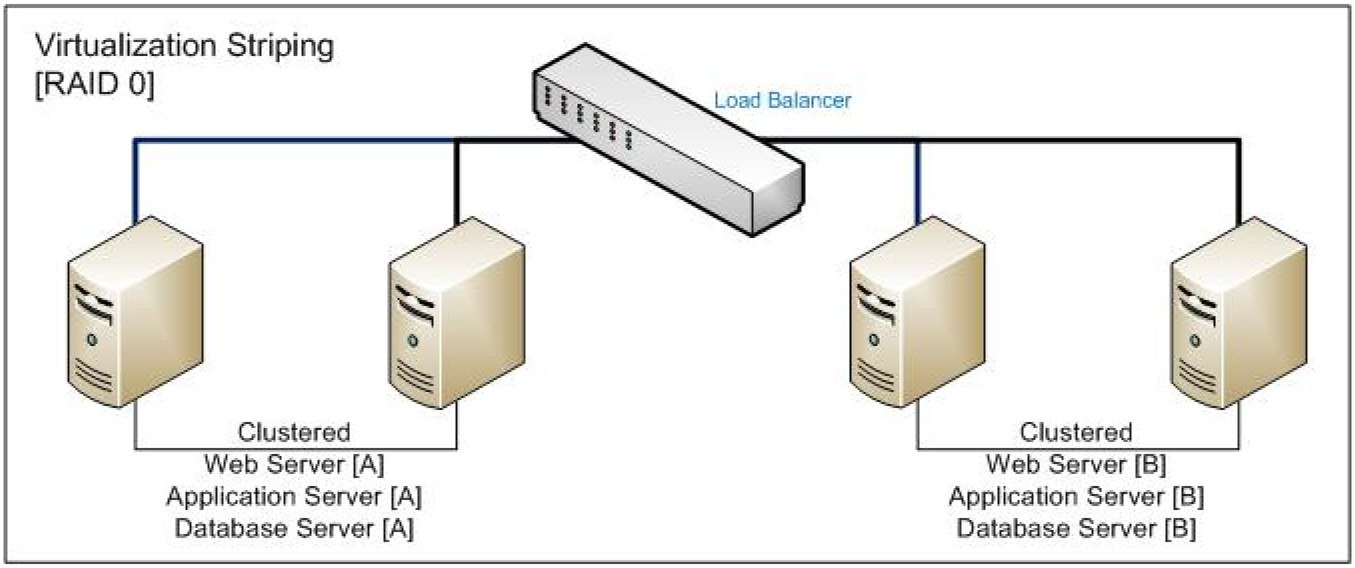
Clustering and Virtualization

At first look, these two paradigms seem to be not so much incompatible, but a zero-sum game. For example, take one physical system and place four virtualized systems on it. Then install in each of the virtualized systems a clustered or grid-based application. Intuition tells us this should result in a negative impact because the characteristics of virtualization combined with the characteristics of clustering nullify each other, notwithstanding the overhead for implementing both strategies in one physical machine. In this example, the errors in design were twofold: the use of just a single machine and the use of just a single clustered application. From a failover perspective, if one were to place all clustered virtual instances on one physical machine, all usefulness of clustering is lost.

However, there is a real use for this combined strategy known as virtualization striping. In its simplest terms, creating a matrix implementation of clustered virtualized applications across separate physical machines allows for five nines of uptime with the best possible CPU optimizations.

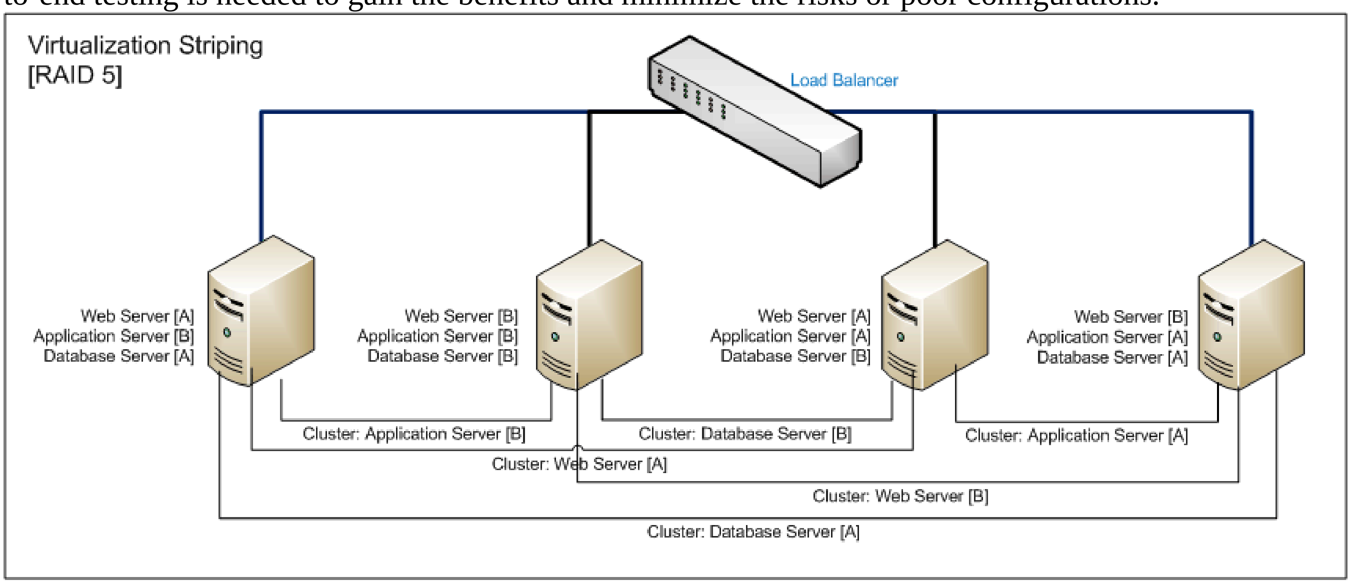
Virtualization Striping

Virtualization striping creates an advantage when end-to-end virtualized systems are clustered against one (or more) similar setups on separate physical machines. When one physical machine encounters a hardware failure, only one application silo is affected; and with clustering against other similar setups, the effect is minimal as the other virtualized systems are acting as failover.



The virtualized striping technique needs to "undo" one major aspect of end-to-end virtualization sets: tightly coupled configurations. Configurations with clustered systems usually operate with a defined pool of resources or defer configuration to a load-balancer. Both types of configuration lie outside the physical realms of the virtualization set. The loss of self-contained features is an acceptable trade-off for flexibility and failover.

Like traditional RAID setups, virtualization striping can be done in 2 flavors: mirrored and striped. The same complexity applies with system configurations as with disk RAID configurations, so careful end-to-end testing is needed to gain the benefits and minimize the risks or poor configurations.



The Downside

Of course, no solution is a panacea, and virtualization striping has its share of pain points as well. Because many systems are so inter-related, many difficult assessments need to be made. Which applications are most tightly coupled? Which logically belong together? Which combinations make the best use of virtualized resources? Which applications are in the critical path of other systems? Which non-clustered applications need to be considered? How does authentication and access control affect the virtualization strategy?

In addition, there is the reality of implementing the virtualization set. There have been many instances where certain applications did not "virtualize" nicely, either because of legacy support, expected exclusive access to hardware devices or just plain old gremlins in the code.

A final point to be made is with use of end-to-end virtualization set used for development environments – locking development into an environment. Once the virtualization environment is properly setup and working, new development must ensure that code does not take advantage of the virtualization features directly… if the systems were ever to be deployed in a non-virtualized environment, they must be guaranteed to work.

Executive Summary

Virtualization striping allows enterprises to gain the advantages of CPU optimization, server consolidation and first-level failover strategy. Virtualization and clustering technologies are more closely aligned than you think.

John C. Checco [CISSP] is a member of Infragard NY, $(ISC)^2$, ASIS NY, *WSTA Advisory/Content Committee* and president of Checco Services Inc. Feel free to send comments to "john.checco" at "checcoservices" dot "com".

[1] Reference: http://h71028.www7.hp.com/enterprise/downloads/5983-0505EN.pdf