Is High Availability Right For You?

View All Posts By Jake Fellows

7-9 minutes

In recent years, the idea of "high availability" infrastructure has gained more and more mindshare in the hosting space. But as with most technology, a common understanding or agreed definition is a little elusive. Especially around high availability clustering.

In this piece, we will **examine what High Availability Architecture actually is, what it is not, and how it works.**

After learning the intricacies of high availability and examining some reasons or use cases for HA, you'll be able to assess whether this new view on infrastructure is the right fit for your organization or project."

First and foremost, it is important to recognize that all people and companies using hosting services would benefit from high availability. That is to say that everyone would like their websites and applications to be fully functioning and operational online as much as possible.

[Downtime is bad](), and the difference between a lot of downtime, a little downtime, and no downtime at all usually comes down to what infrastructure fits the budget and how that infrastructure is deployed.

In recent years, however, the notion of high availability has moved past the simple want for high quality hosting and into a more

technical realm where advanced solutions are actually built with downtime prevention front and center.

What is a High Availability Cluster?

High Availability Clustering is the use of multiple web-servers or nodes to ensure that downtime is minimized to almost zero even in the event of disruption somewhere in the cluster.

High availability clustering uses a combination of software and hardware to:

1. Remove any one single part of the system from being a single point of failure

2. Increase performance throughout the cluster

   Whereas a single shared hosting account or VPS might be felled by the outage of a network switch or a break in the power supply, High Availability clusters have redundancies built in that remove the chance of any single break disrupting service."

By leveraging shared resources across many machines, no one part of the cluster can break the cluster as a whole.

Software components of the cluster detect if an application or service within the cluster is experiencing an issue, and automatically restart the service or application elsewhere in the cluster.

This process, known as [automatic failover](), prevents services, sites, or applications from being unavailable until a server is fixed, as is the case in traditional hosting.

How High Availability Clustering Works

In the most commonly deployed version of HA clustering, a load balancer distributes traffic between a number of web nodes that are all serving the same application or website to

users.

This [load balancing](#) reduces the strain on any one server so that each in the cluster can perform at an optimal level.

Even better, if one of the nodes is compromised or starts to fail, the load balancer can route traffic away from it and to the healthy nodes. As a result, the chance of downtime is significantly reduced and overall performance increases.

HA Clusters are essentially infrastructure that does not depend on any single web-server, power supply, network switch, or in some cases, even a single data center to perform. That way, if any singular event impacts a component of the cluster, the other components can then step in and pick up the slack until the issue is fixed.

Use Cases For High Availability

Highly Available Clusters typically cost more than standard infrastructure, which is why not everyone uses them. That said,

there are a growing number of use cases wherein HA makes sense.

eCommerce

For example, consider an eCommerce site that runs regular flash sales. These sales generate often unpredictable traffic surges, and downtime of even a few seconds can result in lost orders and damaged brand reputation.

By investing in HA clustering, even the most popular eCommerce store can be reasonably certain that their website will stay online even during a [massive traffic spike](#).

Financial Institutions

Also, consider a financial institution that counts on having its customer portal available 24 hours per day. If the website and application are offline for too long, customers may think there is reason to panic with regard to their security. At a minimum, they may be inconvenienced enough to consider switching to a competing bank.

As a result, an investment in highly-available web hosting via HA clustering makes a great deal of sense for that financial institution.

Ridesharing

Think about a ridesharing company that needs to power connection between riders, drivers, and headquarters 24 hours per day. If the mobile application ceased working, then riders might not have a way to order service, drivers wouldn't know where to go, and the company itself would lose scores of useful data…not to mention revenue.

In this case, highly available hosting isn't a luxury; it's a requirement.

Is High Availability Right For You?

In the strictest sense, **HA infrastructure is for everyone**.

Who wouldn't want to have the most robust and reliable infrastructure possible? Even casual site publishers and hobbyists aren't pleased when their website is offline for even a few minutes.

> But as a practical matter, HA Clustering can be expensive, as each component typically carries additional cost. In many environments, a two-node cluster costs twice that of a one node environment, just as a four-node cluster costs twice as much as a two-node cluster."

And as the clusters grow, the complexity in managing and monitoring the clusters increases as well. As a result, not all businesses or individual users can balance the expense of a cluster with the benefit of reduced downtime.

In your specific case, the best way to evaluate if highly available hosting is the right choice is to [quantify the cost of downtime](#).

In some cases, like eCommerce sites that literally cannot generate revenue if they are offline, this calculation is simple. Calculating how much revenue can be expected each minute is a pretty good guide to determine how much revenue could be lost for minutes or even hours of downtime.

In other cases, quantifying lost reputation, lost contacts, or lost content is a little harder. However, if your business is dependent on the website to generate leads or revenue, or dependent on web-based applications to drive growth, then the more robust hosting provided by a high available cluster is worth exploring further.

High Availability at Liquid Web

At Liquid Web we've been providing highly-reliable, highly-secure highly available infrastructure for more than twenty years.