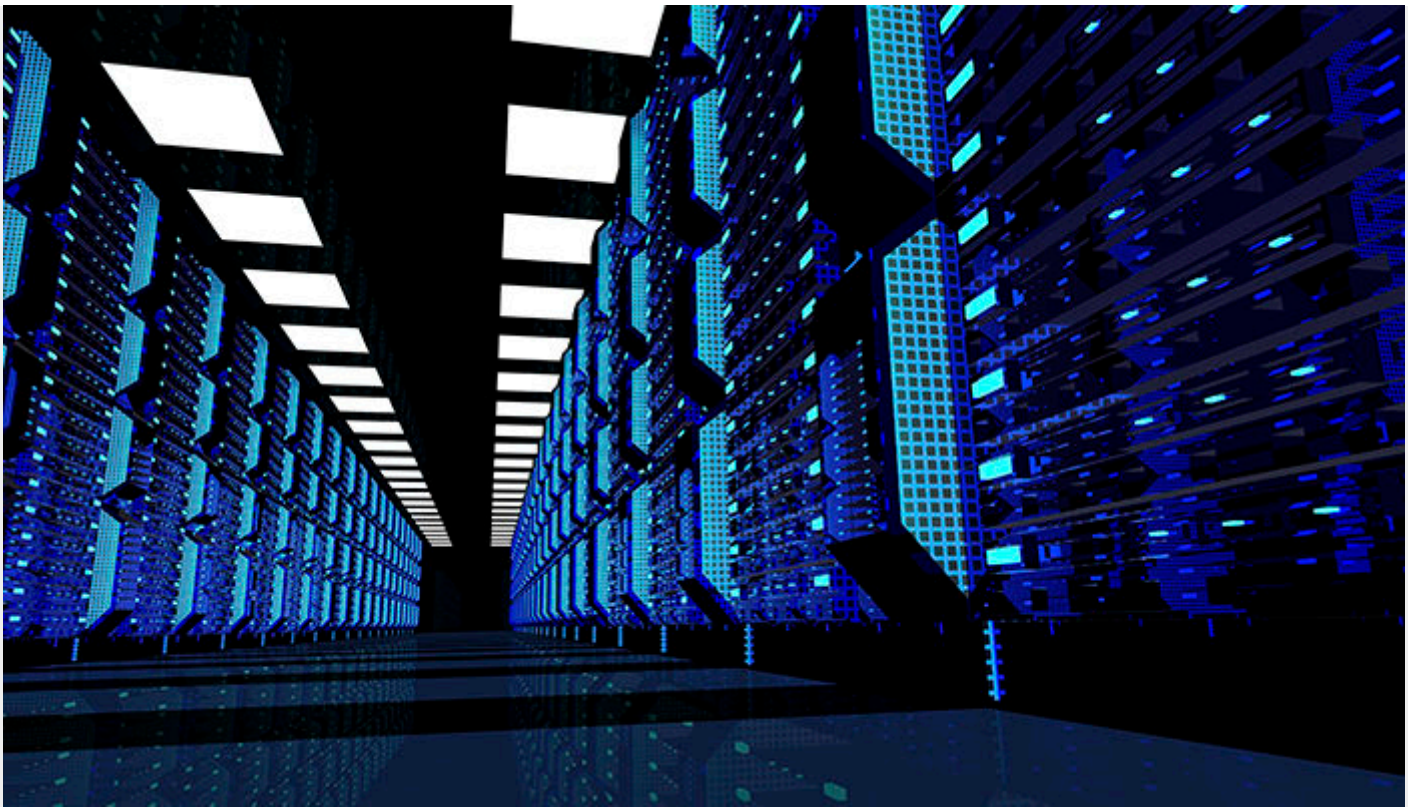
USING REDUNDANCY TO REMOVE RISK FROM IT OPS



Redundancy is good in Data Center (DC) design and management. I repeat: redundancy is good in Data Center design and management. When doing just about anything regarding your Data Center, redundancy is your friend.

Why is Data Center redundancy so valuable?

Redundancy is a core concept when architecting and running a Data Center. Backup capabilities for critical Data Center functions are essential to the smooth management of a Data Center. If a redundant primary service, function, feature, electrical service, or telecommunications line goes down, your backup can easily pick up the slack, allowing maximum uptime until the failing item is fixed.

And don't forget about redundancy when you're running in a managed service provider environment or running critical services in the cloud. Be sure to understand exactly what redundant features are provided in your hosting contract, and then get those services in writing.

But what items and components need to and should have redundancy built-in? Here's my list of the most important Data Center items that should and should not have built-in redundancy.

The Redundancy List

These items should have redundancy built in. Not every Data Center or Managed Service Provider implements each redundant item. However, the more redundancy you can build into your Data Center infrastructure, the better. And the fewer outages or problems you'll experience.

Servers

Production servers need redundant backups, but that redundancy can take several different forms. Here are a few examples of the ways you can provide redundancy for critical Data Center servers.

- **Redundant domain, front end, and validation servers** can be used for load-balancing to insure customers can always be serviced. A secondary DNS server (for example) can still resolve addresses if the primary server goes down or is busy. The same goes for Windows AD servers that validate user access to the domain. Load-balanced application servers, such as Web servers or front-end servers, can provide redundancy and improve throughput.
- **Replicated servers**, where a production server is paired with a replicated backup. Changes are replicated to their backup servers automatically, using one of several different software-based or hardware-based technologies. In the event of a server failure, the replicated server can be pressed into service.
- **Disaster recovery servers** are semi-hot spares that, in the event of a disaster, can have backup files quickly restored and restart processing.

Backups

Redundant backups can be performed in a variety of different ways. Tape backups are frequently supplemented by disk backups. Disk backups can be stored on a local backup server as well as on a remote backup server, or as a cloud backup. Multi-tiered backups can be implemented where some data is saved locally while other data is stored off-site (usually to a geographically distant area), allowing you to quickly restore data locally while also protecting your backups in the event of a regional disaster.

Disk drives

Hot spares should always be available in your critical servers so that if one disk drive in a RAID set goes bad, another drive can immediately replace it. Using hot spares, your RAID disk sets can continue processing when there's a single disk failure.

Power supplies

For critical servers, you should always implement redundant power supplies so that if one power supply fails, your server will continue to run. Redundant power supplies may also be available for non-server critical equipment and should be ordered and installed, if available.

UPS systems

Each server rack should have at least two redundant UPS systems powering it, that will continue supplying power in the event of a short-term brown-out or a short power outage (i.e., your servers

won't crash if the lights flicker). Each of the two redundant power supplies on your servers and other critical equipment should be plugged into a different UPS system, to prevent the equipment from going down in the event of a power failure and to guard against a UPS failure. All Data Center equipment should be plugged into at least one UPS, no exceptions.

Electrical circuits

Each of the redundant UPS systems in the same rack should be plugged into a different circuit breaker, ideally on different electrical boxes. This guards against taking down your entire rack in case there's an electrical problem with one of your circuits or electrical boxes.

Electrical supply/Generator

For added electrical redundancy, some shops (and many hosted Data Centers) provide different electric supplies for their redundant circuits, sometimes powered by different electrical providers. Many companies also install a generator that will take over powering the Data Center in the event of a longer power outage.

Telecom lines

Some in-house Data Centers and many hosted DCs have redundant telecommunications lines, usually provided by different telecommunications companies, using different feeder lines into your organization. Again, if one of the telecom lines fails (or somebody cuts it with a back hoe), network traffic can be shifted to the non-damaged telecom line.

Things that may not redundant

Unfortunately, you won't be able to apply redundancy to everything in your Data Center. There will be items that can be considered single points of failure for which redundancy is not possible. Here's a list of Data Center items that are resistant to redundancy and some suggestions for how to work around these items in the event of a failure.

Network ports, cabling, and patch panel ports

There is usually only one cabled data path from any location to the Data Center. Your entire cabling structure will seldom fail but occasionally, a cable run from one specific location to the Data Center will be accidentally cut, broken, or bent, causing that location to fail. When a cable run breaks, you can usually do one of the following:

- Repair the damaged cable, or run a new cable.
- Attach a small Ethernet switch to a nearby working line and share the working line with the equipment plugged into the failing location.
- Use two wireless access points as a bridge from the failed network line location to a working network line. This can be an expensive fix, but it can work very well in a pinch.

Phone systems

Except for Voice over IP (VoIP) systems, there is generally no good redundant solution for a physical

phone system, except to keep it in good working order and make sure you have a good maintenance provider.

Switches/routers/firewalls

In my experience, it's more difficult to provide hot swap or auto-fail over capability for these items. My best advice is to have current backups of your router/switch/firewall configurations and to have backup equipment available, in case your equipment breaks down.