

High Availability and Disaster Recovery

DOMAIN 3.0

MODULE 12



High Availability and Disaster Recovery Topics

HA and DR Concepts

High Availability Mechanisms

Disaster Recovery Mechanisms

Facility and Infrastructure Support



HA and DR Concepts



High Availability

A system, network, or service that is continuously operational for a desirable length of time

Availability is measured in “9s”

- 90 % = “one nine” 36.5 days down in a year
- 99 % = “two nines” 87.6 hours (~3-1/2 days) down in a year
- 99.9% = “three nines” 8.76 hours down in a year
- 99.99% = “four nines” ~ 53 minutes down in a year
- 99.999% = “five nines” ~ 5 minutes down in a year

High Availability Mechanisms include:

- Fault tolerance
- Redundant components or systems
- Load balancing
- Clustering
- NIC teaming
- Port aggregation



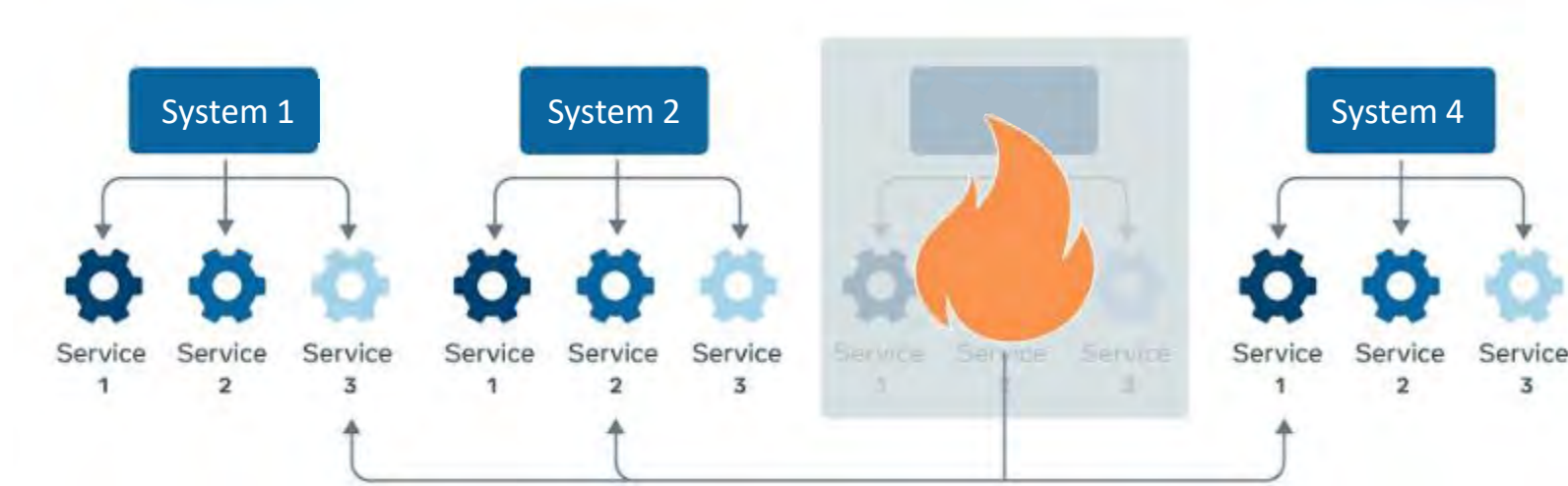
99.999%
uptime



Fault Tolerance

Capability of a system or network to provide uninterrupted service if one or more of its components fail

- No single point of failure
- Avoids losing data or connectivity
- Failover is rapid and automatic



Mean Time to Failure (MTTF)

One of many metrics used to evaluate the reliability of a manufactured product

- Usually published by the manufacturer

Used to estimate the normal lifespan of products that are not repairable

- Or repair cost exceeds replacement cost

Will be shortened by improper usage

Typically devices are field replaceable units (FRUs)

- If possible, replace before actual failure

Field Replaceable Units

Examples of FRUs that should be replaced, not repaired:

- Hard drives
- Fans
- Video cards
- Motherboards
- CPUs
- RAM
- Surge protectors
- Power supplies
- Other peripherals such as monitor, keyboard, mouse, card readers, CD/DVD drives
- Cables (if they have reached expected lifespan and are starting to break down)

Mean Time Between Failures (MTBF)

Refers to repairable devices

- How long the device/system is expected to function until its first failure
- How long after first repair before device is expected to fail again

Can (hopefully) be extended by proper maintenance

Estimates only, but important in planning, implementation, maintenance, and future plans

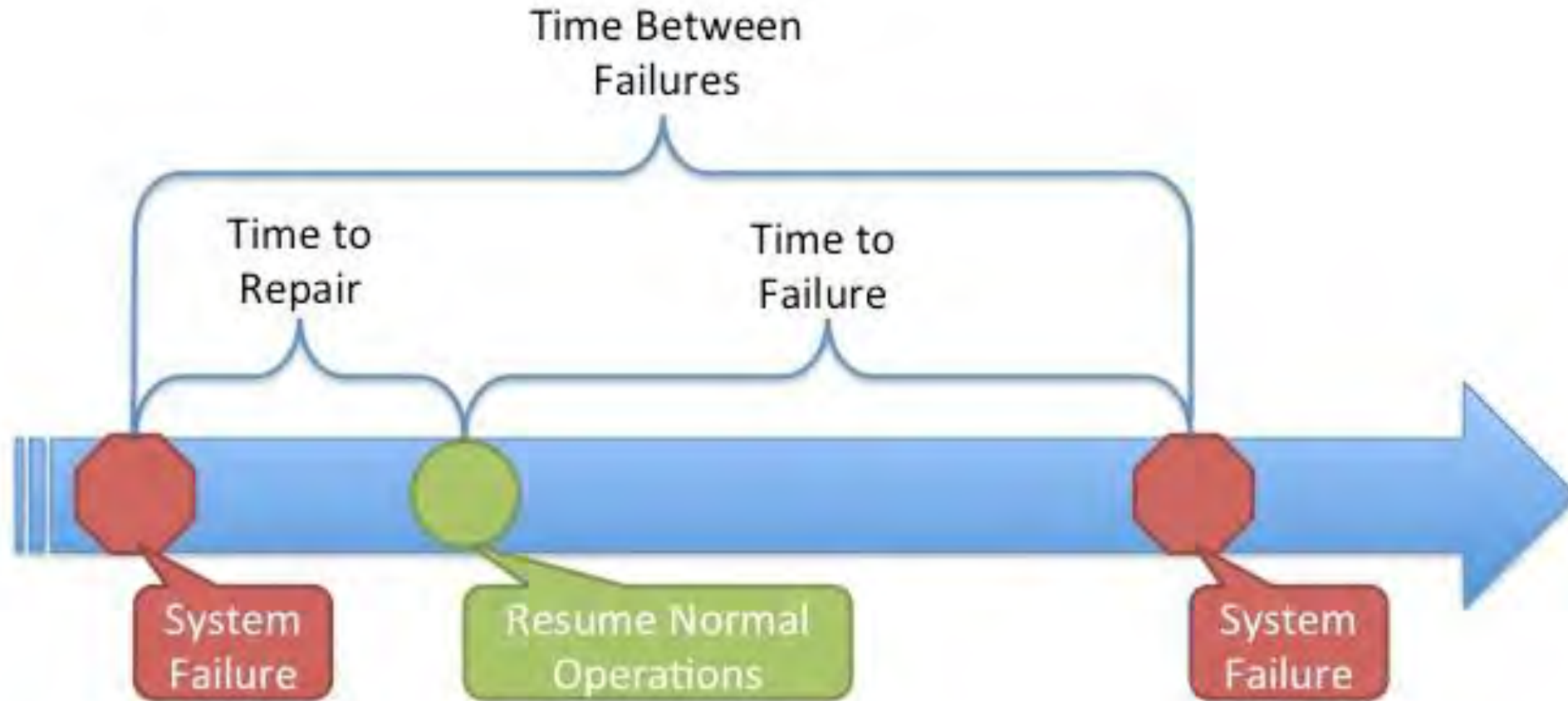
Mean Time To Repair (MTTR)

How long it will take to repair a device, system, or component that is down and bring it back online

Assumption is that the device/system can be repaired

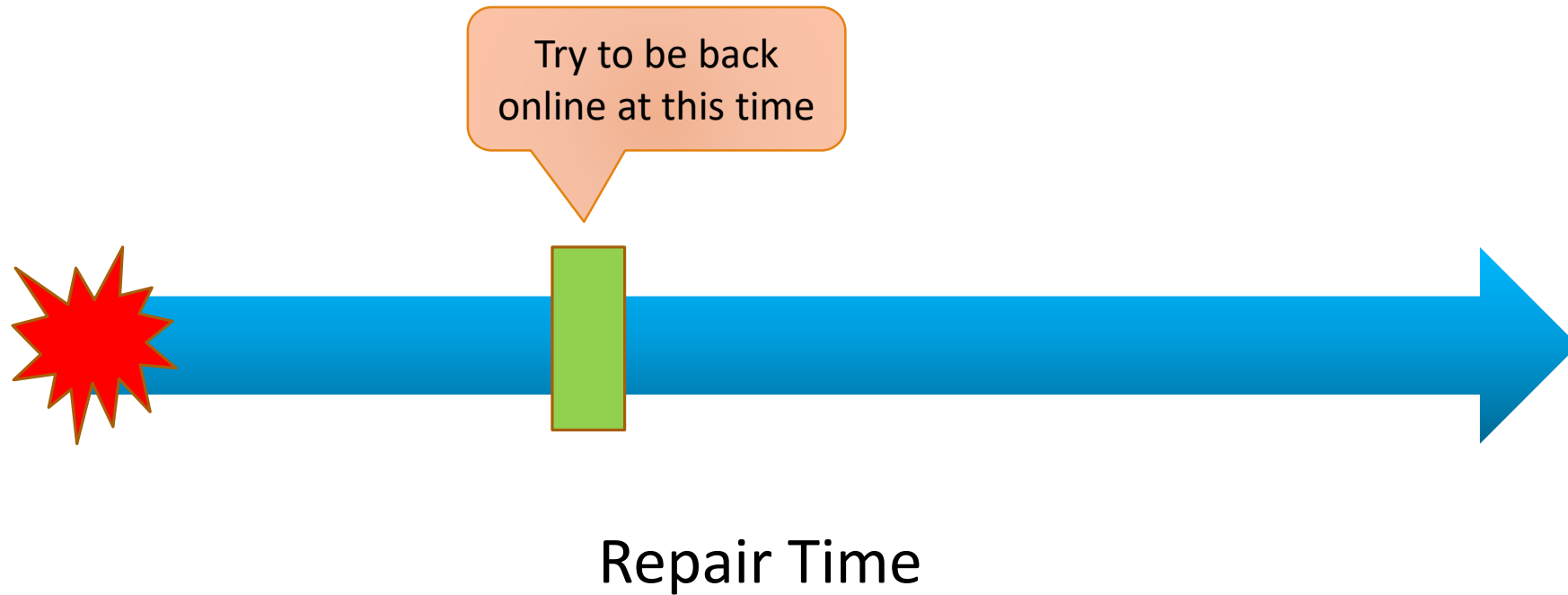
Critical metric in planning data center/cloud/system configuration and future configurations

Differentiating Between MTBF and MTTR



Recovery Time Objective (RTO)

When restoring a system, the maximum allowable time that can elapse before the system is available again



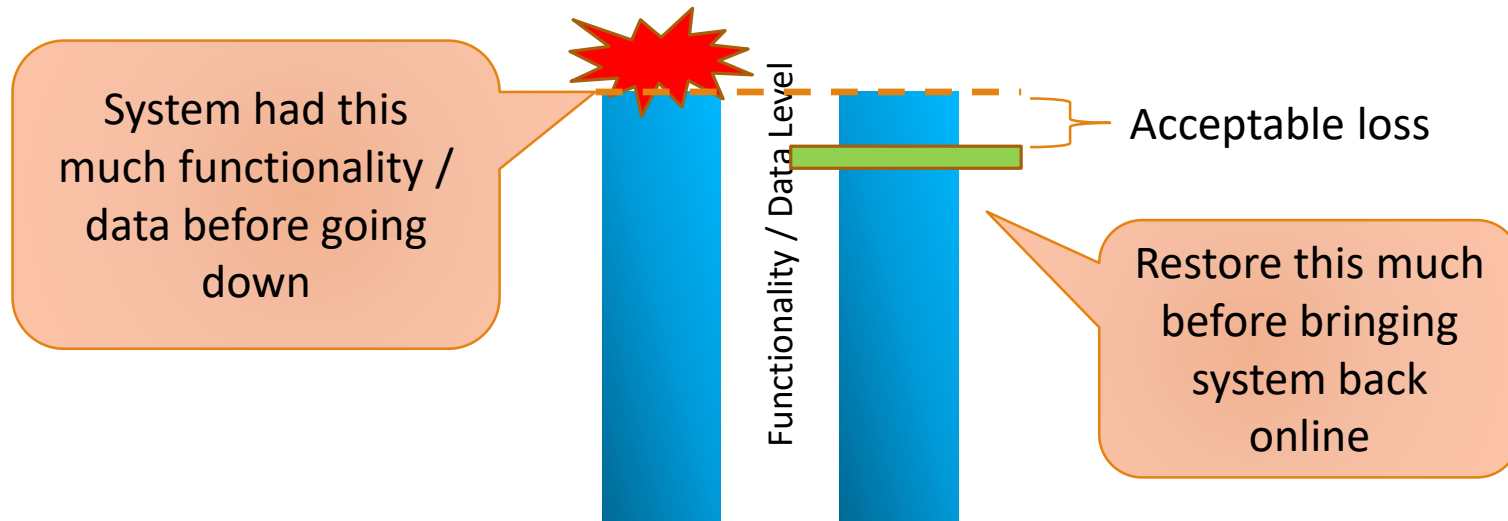
Recovery Point Objective (RPO)

When recovering a system, the level of original functionality to be restored before bringing the system back online

- Sometimes data is lost so you cannot fully recover the system to its original state
- Sometimes you sacrifice level of recovery in the interest of quickly making the system available again

RPO is often used in database recovery

- Identifies the last saved transaction to be restored before the database is made available again
- Any transactions after the RPO will have to be manually re-entered





High Availability Mechanisms



Load Balancing

Two or more systems simultaneously provide the same service

If one node fails, the other node(s) continue to provide service

Especially good resilience against denial-of-service attacks

All systems have their own IP address, but share a common virtual IP address

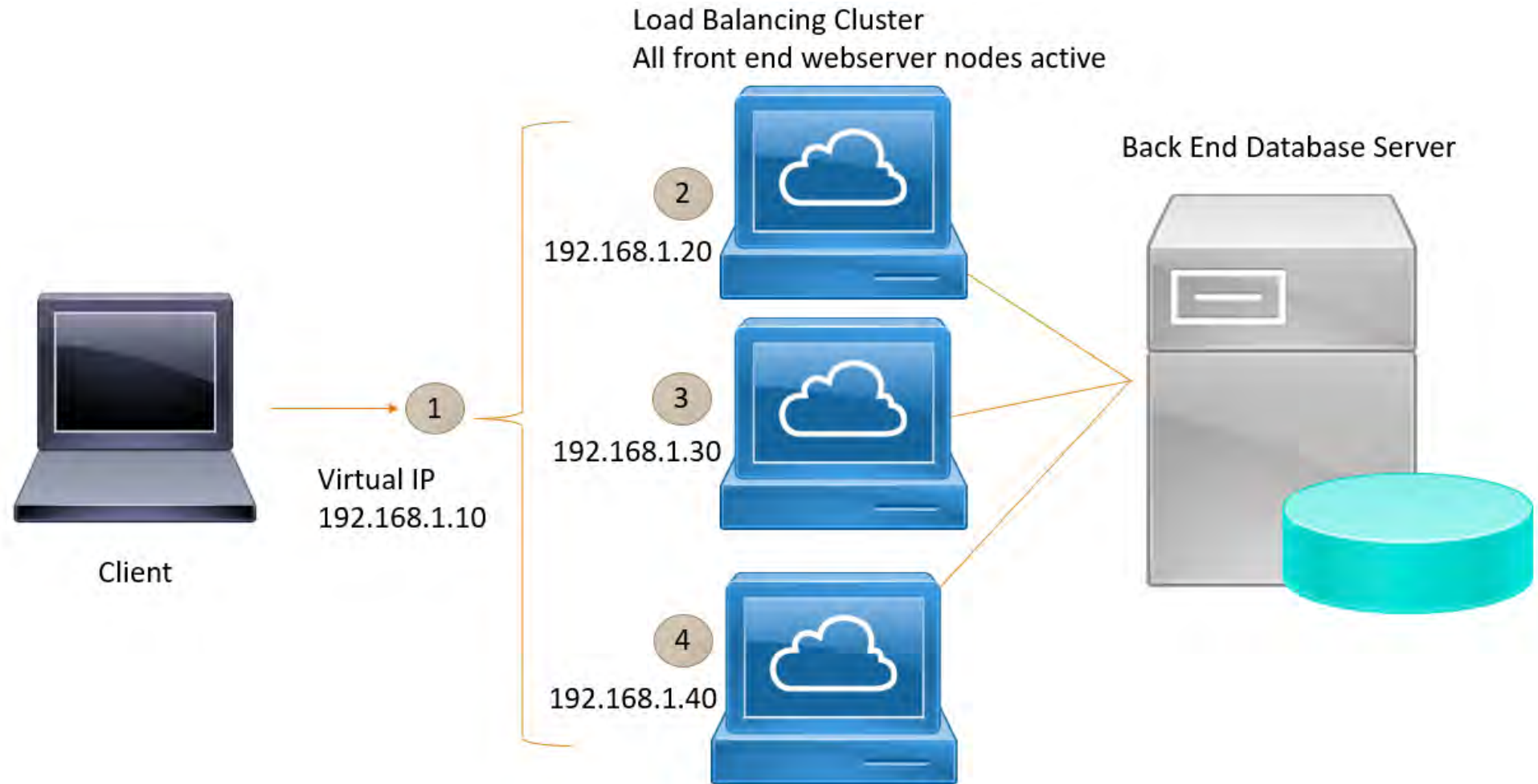
Clients connect to the virtual IP address

Systems do NOT share a common database/data files

Systems are typically “front end” web sites that “point” to a common back end database server

Can be hardware or software solution

Load Balancing Example




Multipathing

A generic term for any redundant network path

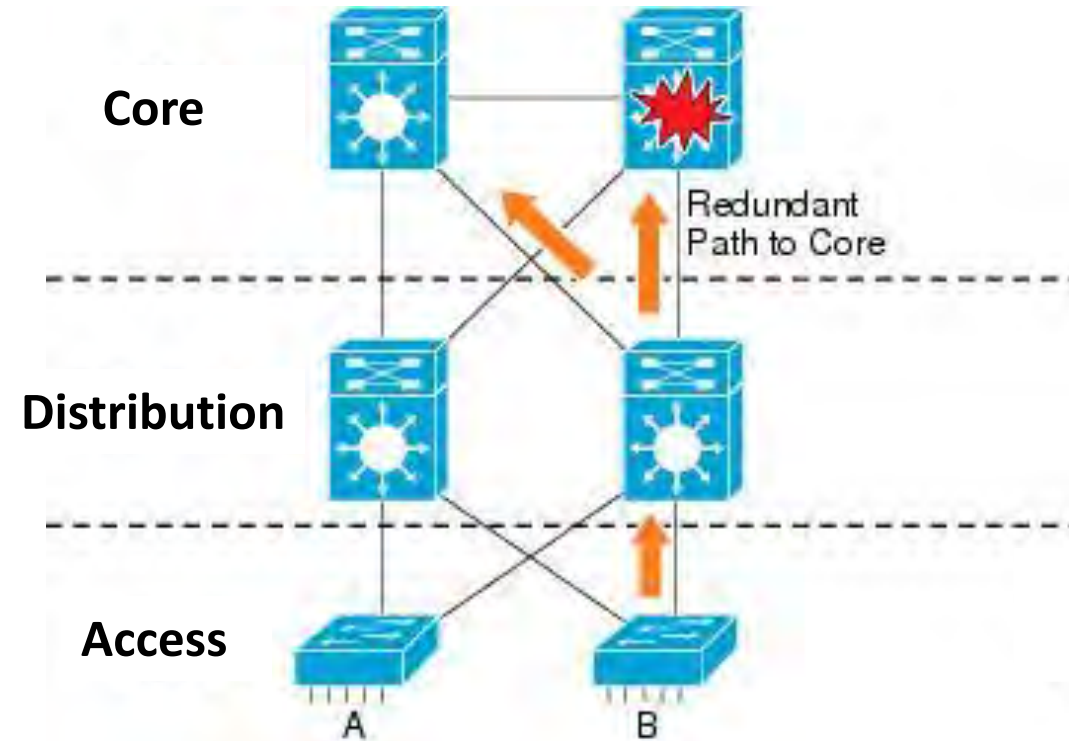
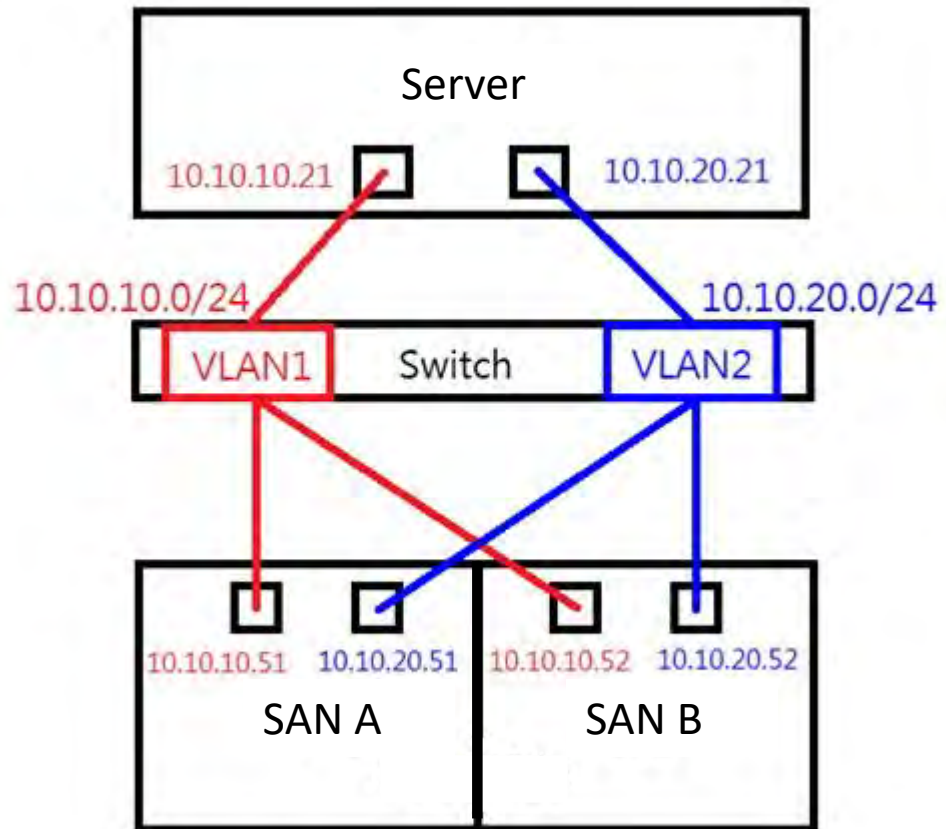
Can refer to:

- Multiple links between servers and SAN storage
- Redundant links at the same network layer
- Redundant links between network layers
- Multiple links to the same ISP
- Multiple ISPs



Most common
usage

Multipathing Examples

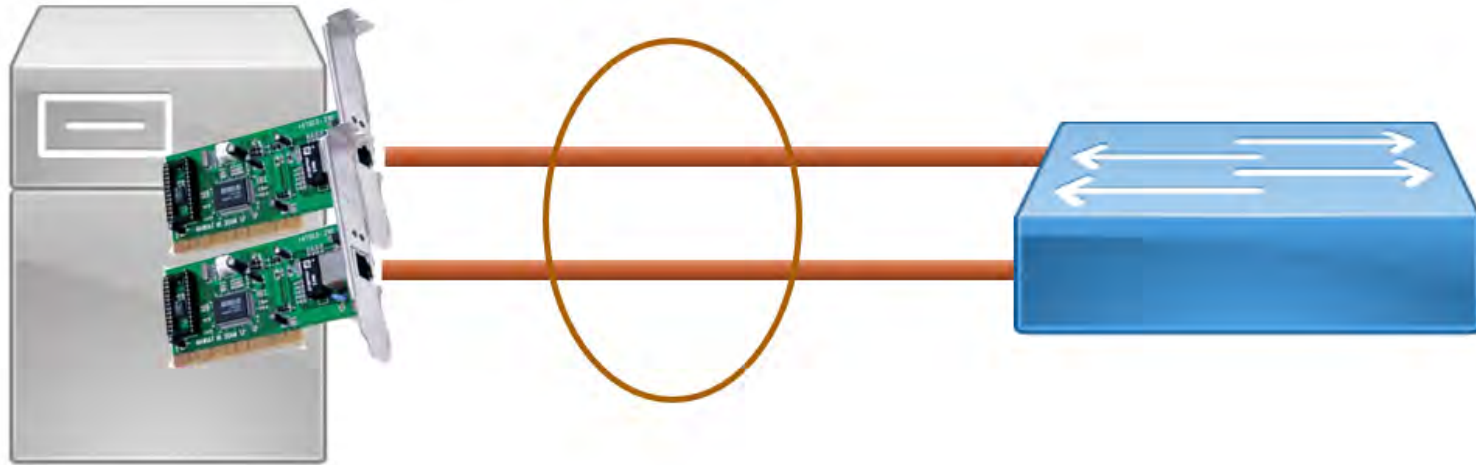


NIC Teaming

Network Interface Card teaming combines multiple NICs/connections to create a single “link”

- Aggregates bandwidth
- Increases performance
- Provides fault tolerance

Also known as aggregation, balancing, and bonding



Clustering

Redundant hardware acting together as a unit

- Two or more systems provide a single service
- Systems typically share a common database/files
- All systems have their own IP address, but also share a common virtual IP address
- Clients connect to the virtual address

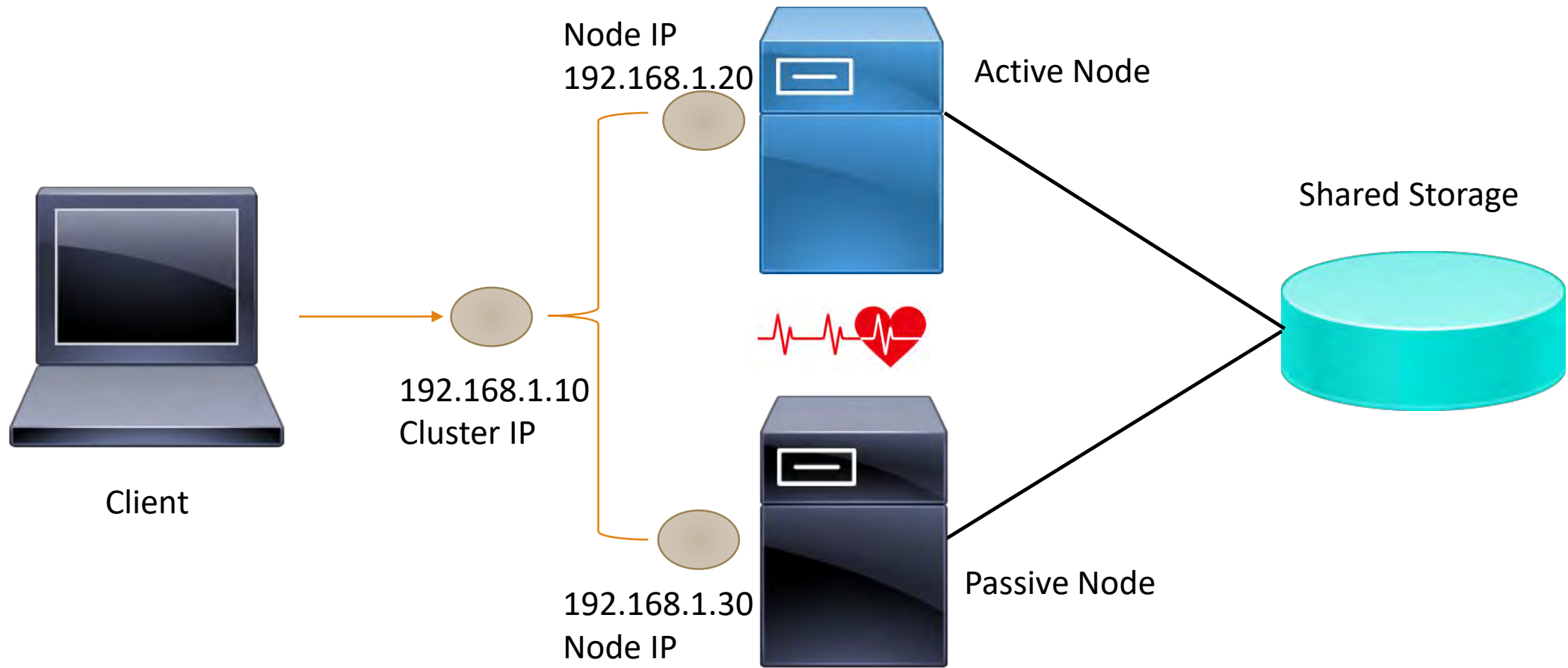
Active/Passive:

- One system is active
- The other system is in standby (passive) mode
- Passive system listens to the “heartbeat” of the active system
- If it stops hearing the heartbeat, the passive system takes control of the data/database/service

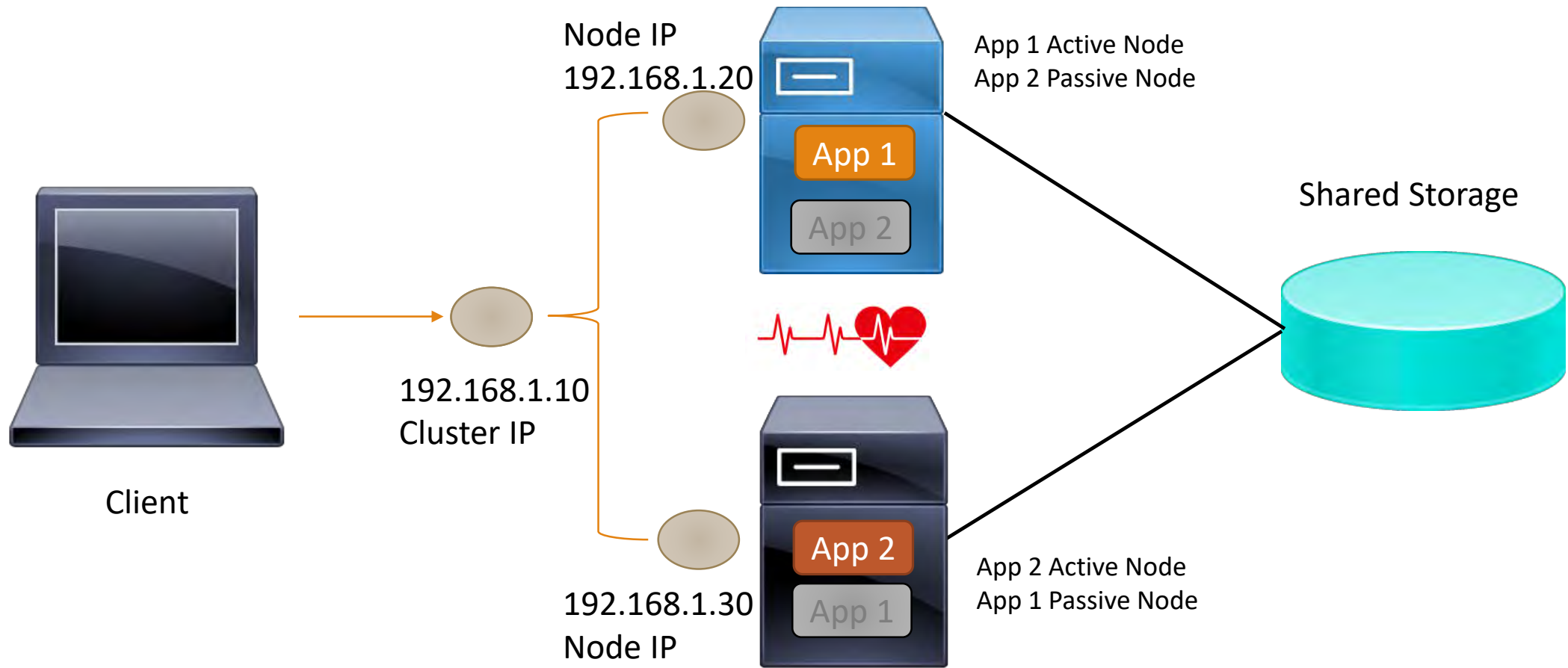
Active/Active:

- Both systems are active
- Each system is the primary provider for a different service (e.g. SQL and Exchange)
- Each system acts as backup for the other
- Either system can take over both services

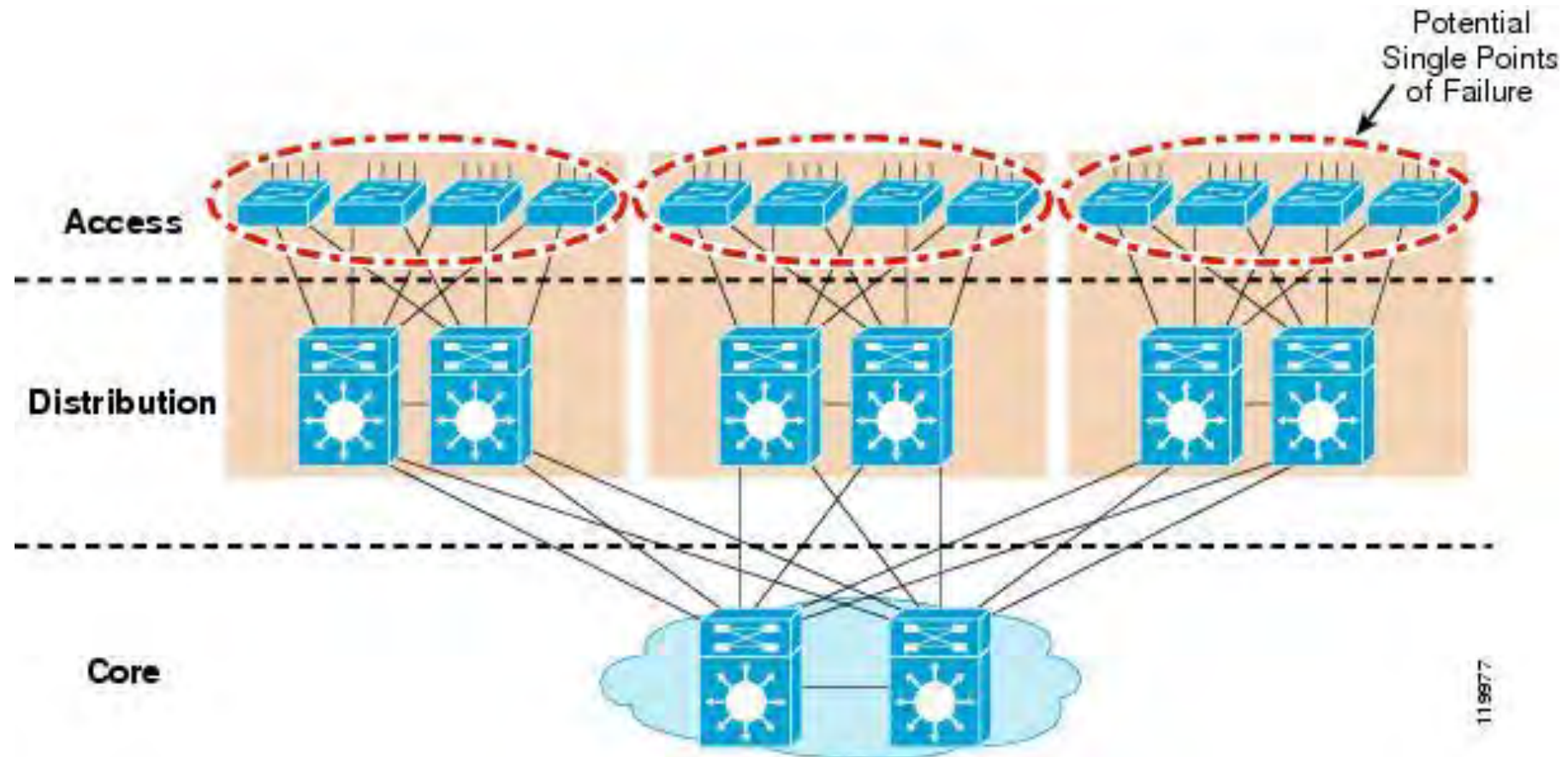
Active / Passive Clustering Example



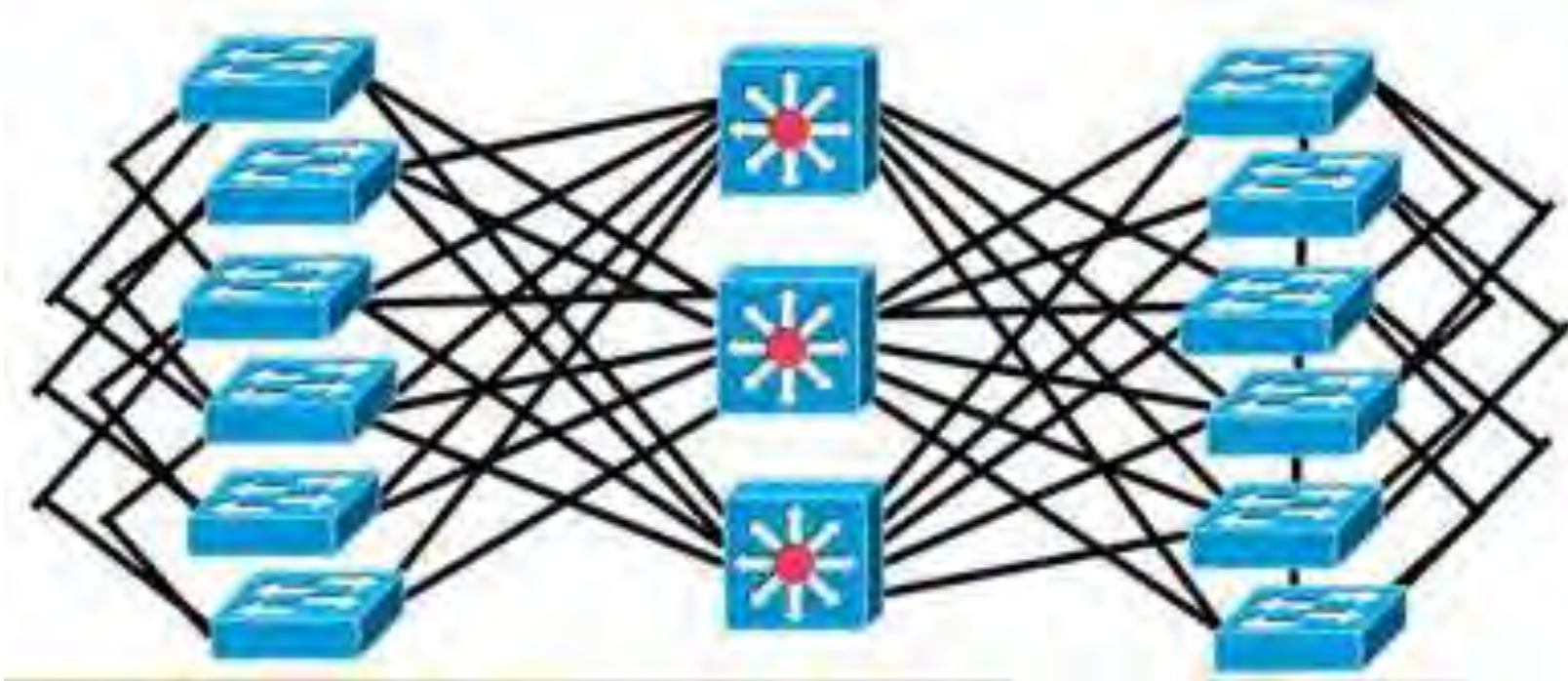
Active / Active Clustering Example



Redundant Switches Example



Too Much Redundancy!



Port Aggregation

Logical aggregation of Ethernet switch ports

Used to increase the bandwidth of a “single” link

Commonly used in uplinks / trunk links

Also referred to as EtherChannel

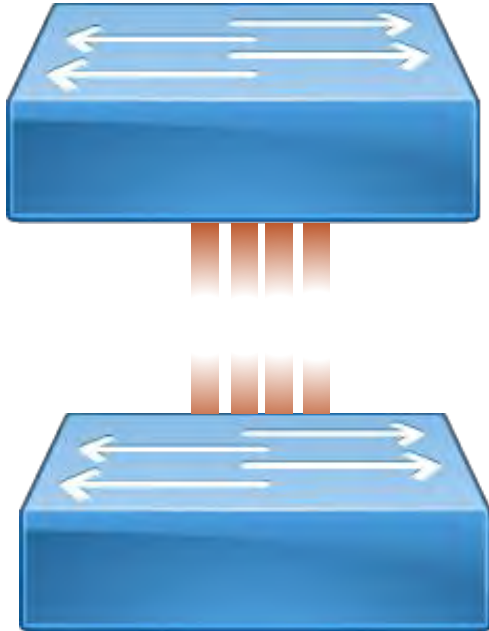
Two common methods:

- Cisco proprietary PAgP
- Vendor-neutral LACP (IEEE 802.3ad / 802.1ax)

Port Aggregation Example

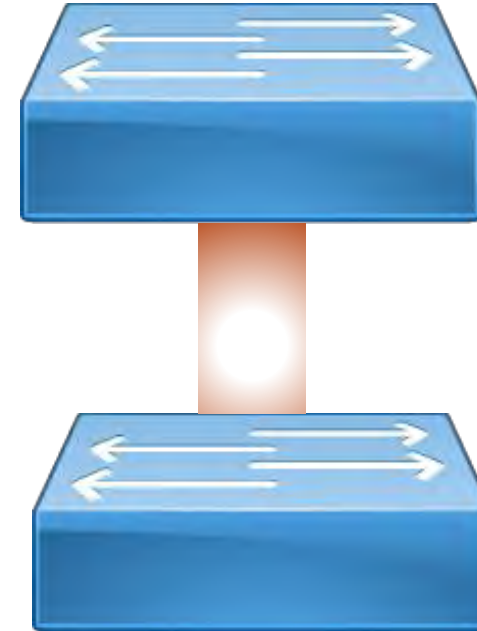
Physical View

Multiple ports defined as part of an EtherChannel Group



Logical View

Different subsystems running on the switch see only one large link



Redundant Routers

You can cluster routers

First Hop Redundancy Protocols (FHRP)

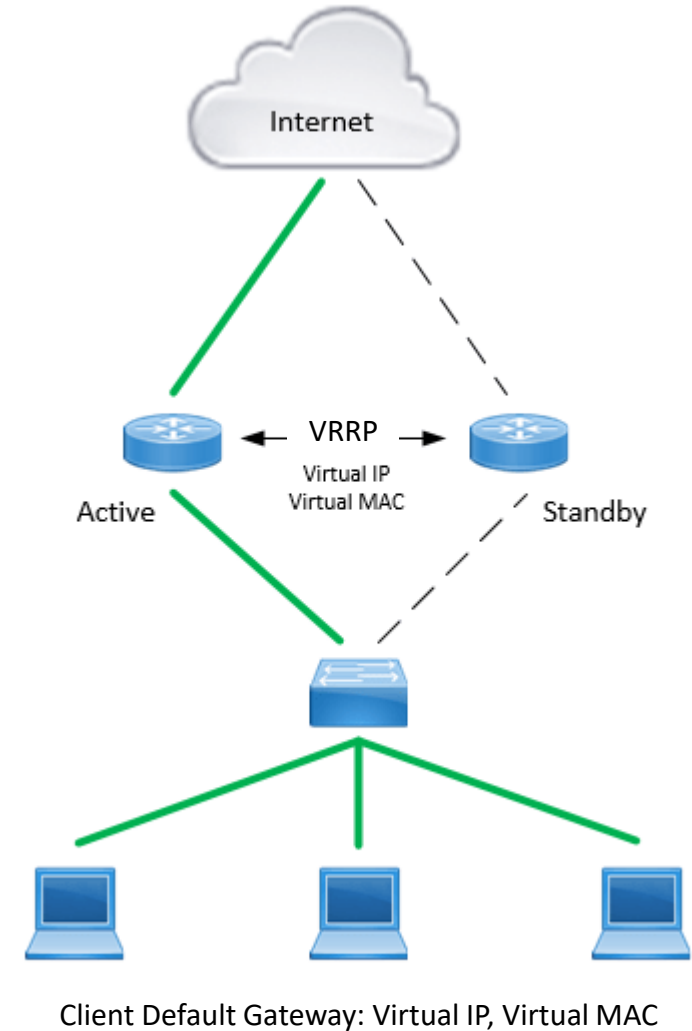
- A class of mechanisms that allow default gateway redundancy

Virtual Router Redundancy Protocol (VRRP)

- Standards-based FHRP
- Active and Standby routers are organized into a Standby group
- They share a virtual IP and virtual MAC
- Active router is configured with a higher priority so it is preferred
- The standby router has a lower priority, but can take over for the active at any time

Cisco has proprietary FHRPs as well:

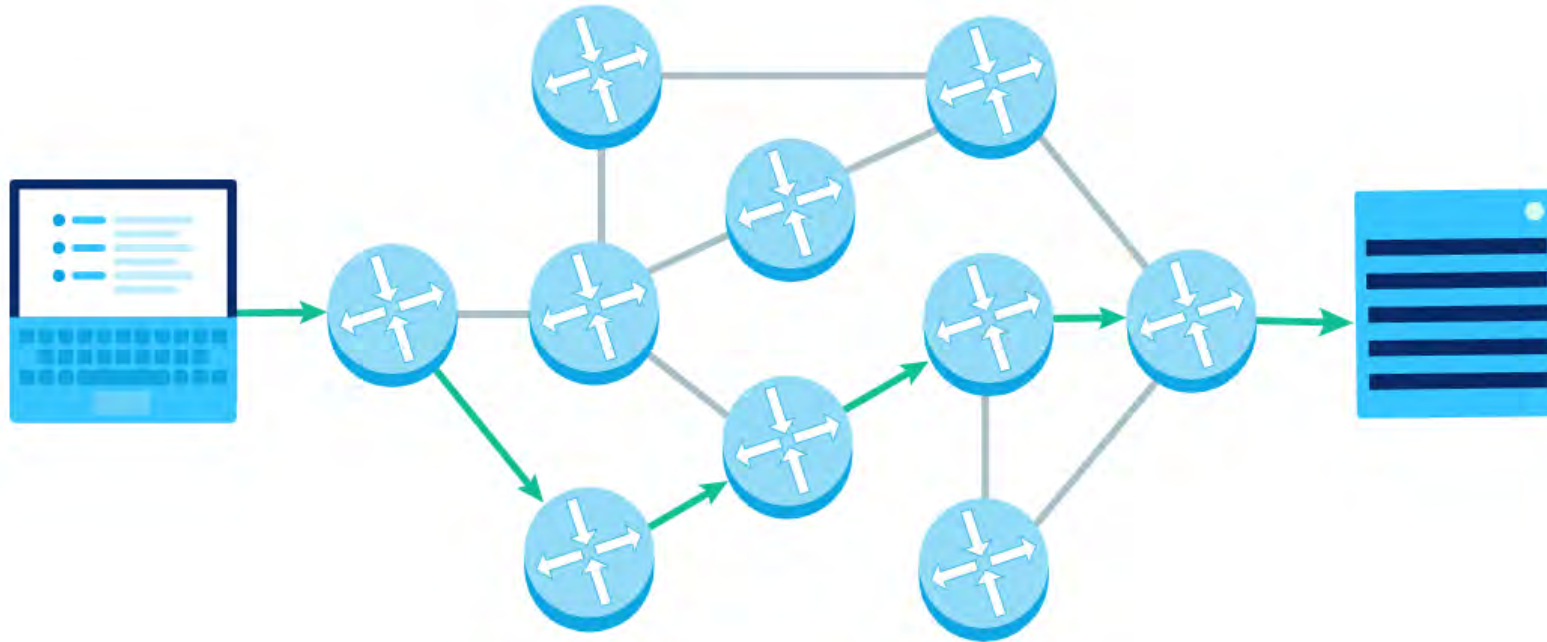
- HSRP – very similar to VRRP
- GLBP – like HSRP but also allows Active – Active load balancing



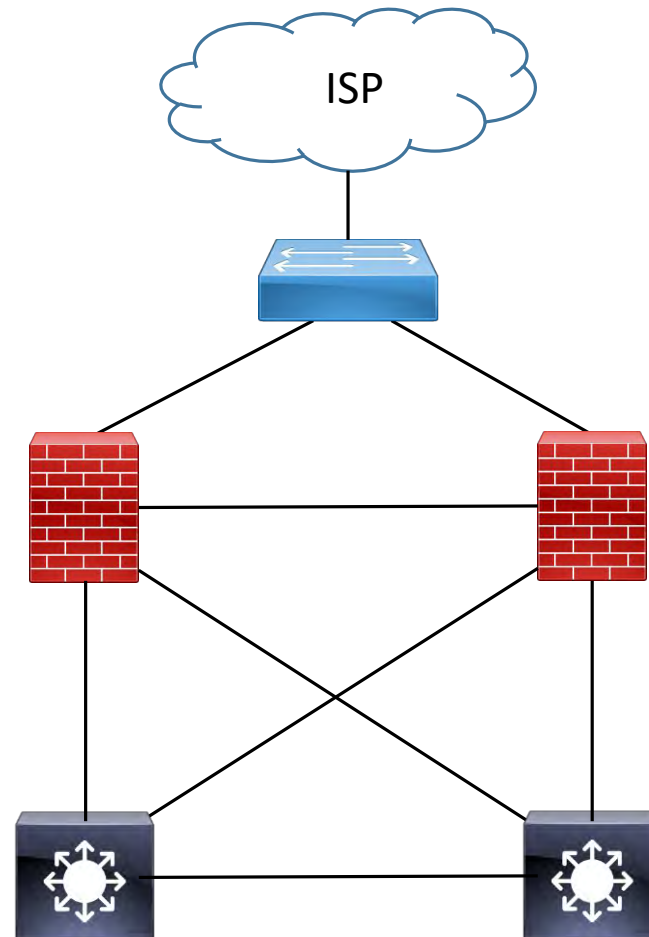
Redundant Routes

The routing protocol must decide the best path

More redundancy = more fault tolerant = more expensive

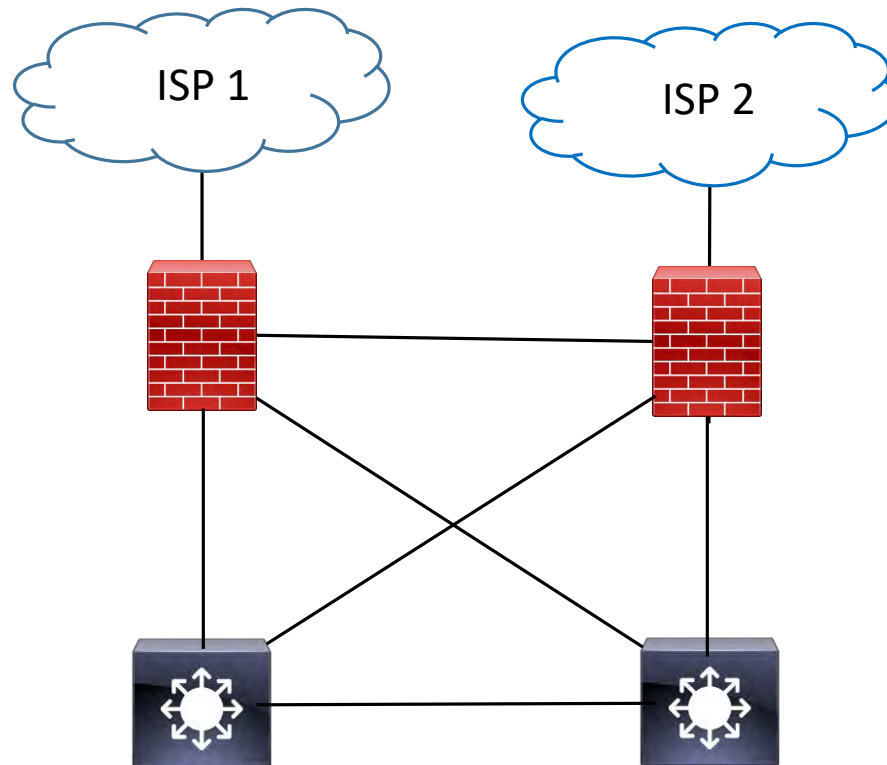


Redundant Firewalls



Multiple ISPs / Diverse Paths

Configure ISP2 as a standby link
Or load balance between the two





Disaster Recovery Mechanisms



Backups

Protect against hardware failures, data loss, corruption, disasters (manmade or natural) etc.

Backup Types

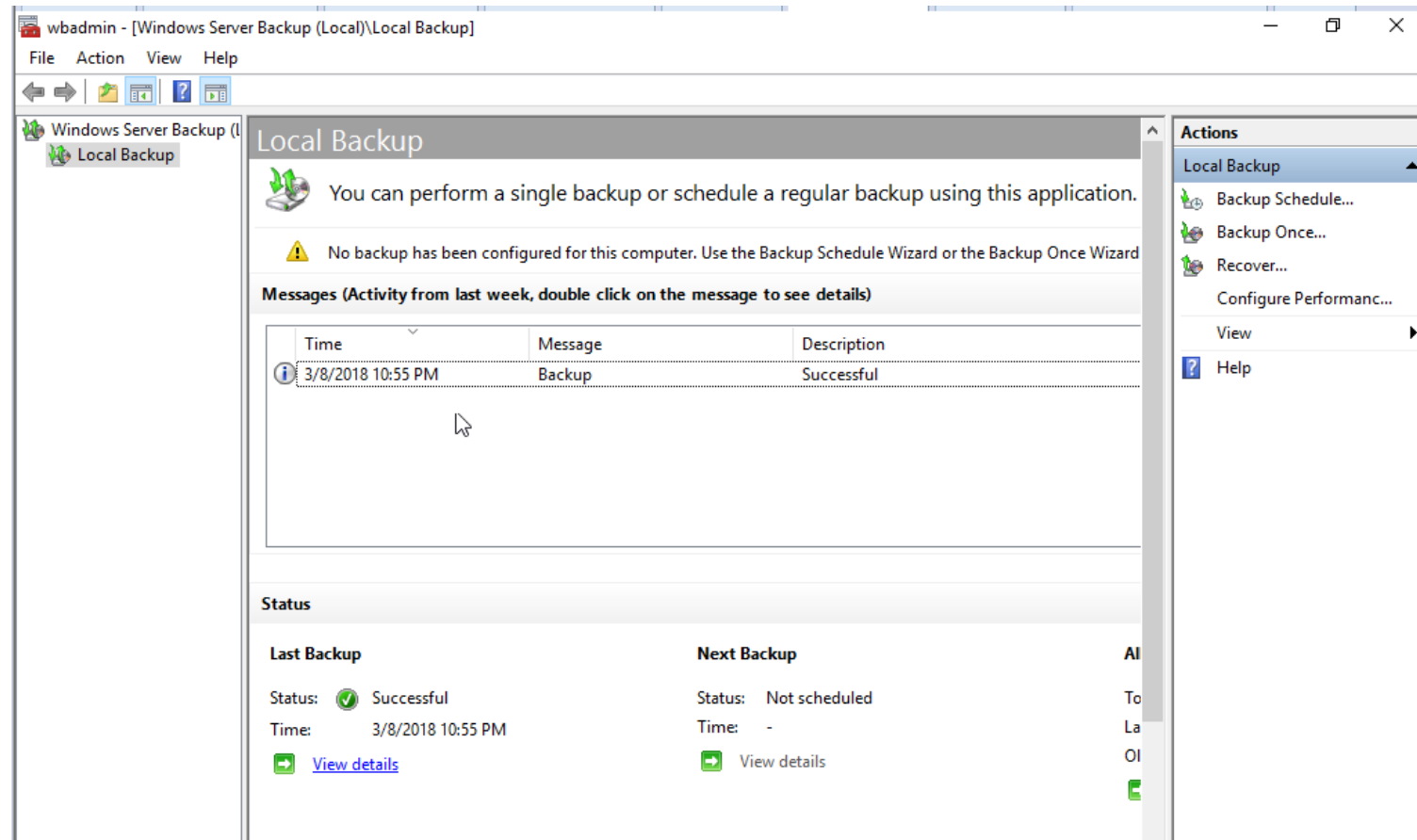
- Full, differential, and incremental
- Snapshots
- OS and configurations (network devices)

Backup Destinations:

- Local disks
- Network shares
- Cloud



Windows Server Backup Tool



Full Backup

The most basic and complete type of backup

Backs up all selected data to another set of media

- cloud, network share, local disk, tape

Provides a foundation for the other backup types

Changes the file's archive bit

Longest backup time

Shortest restore time



Full backup on Sunday

Differential Backup

Copies all data changed since last full backup

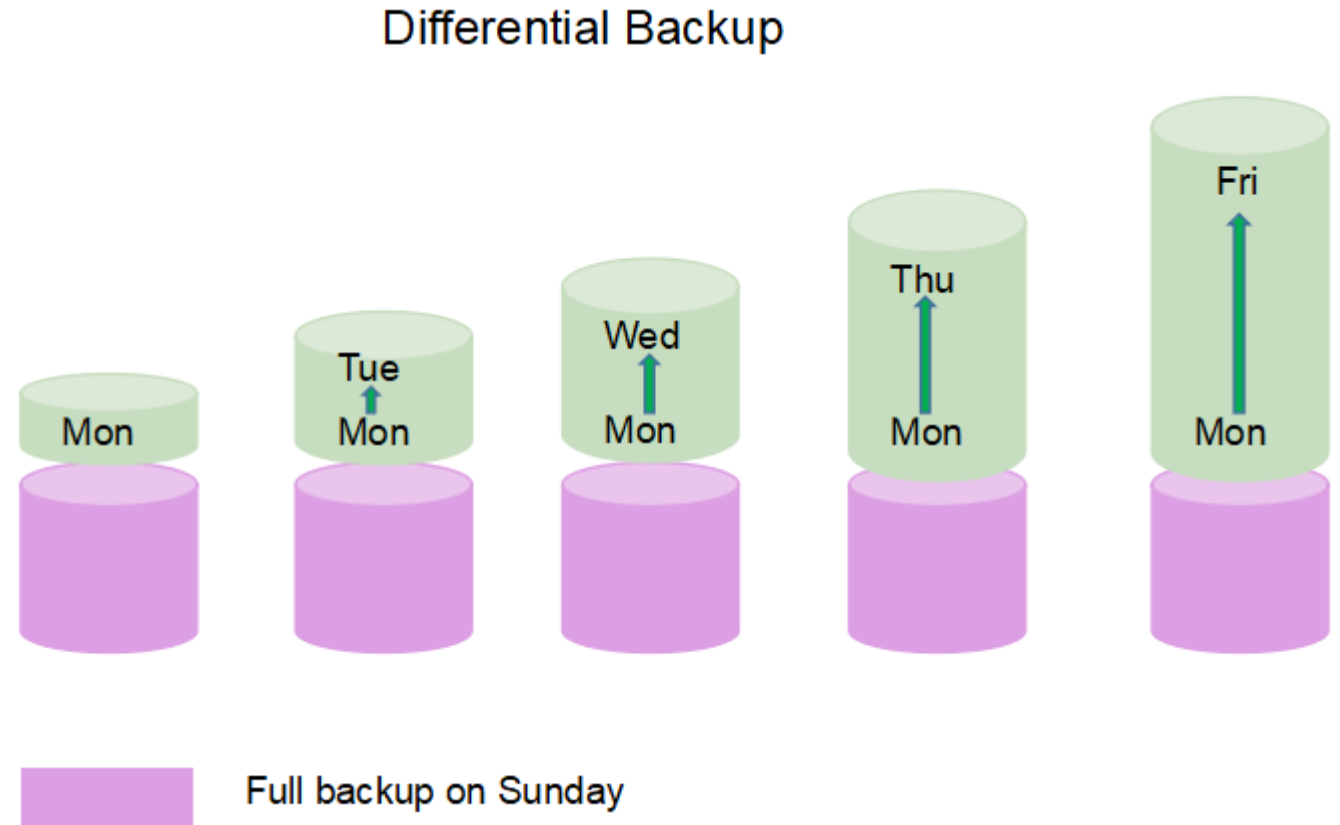
Does not change the archive bit

Can be thought of as a “running backup”

Typically get larger over time until the next full backup

Backup takes longer each day as the week goes by

Restore takes less time as you only need the full plus the latest differential



Incremental Backup

Copies only the data that has changed since the last full or incremental backup

Restore requires the full backup plus all subsequent incremental backups

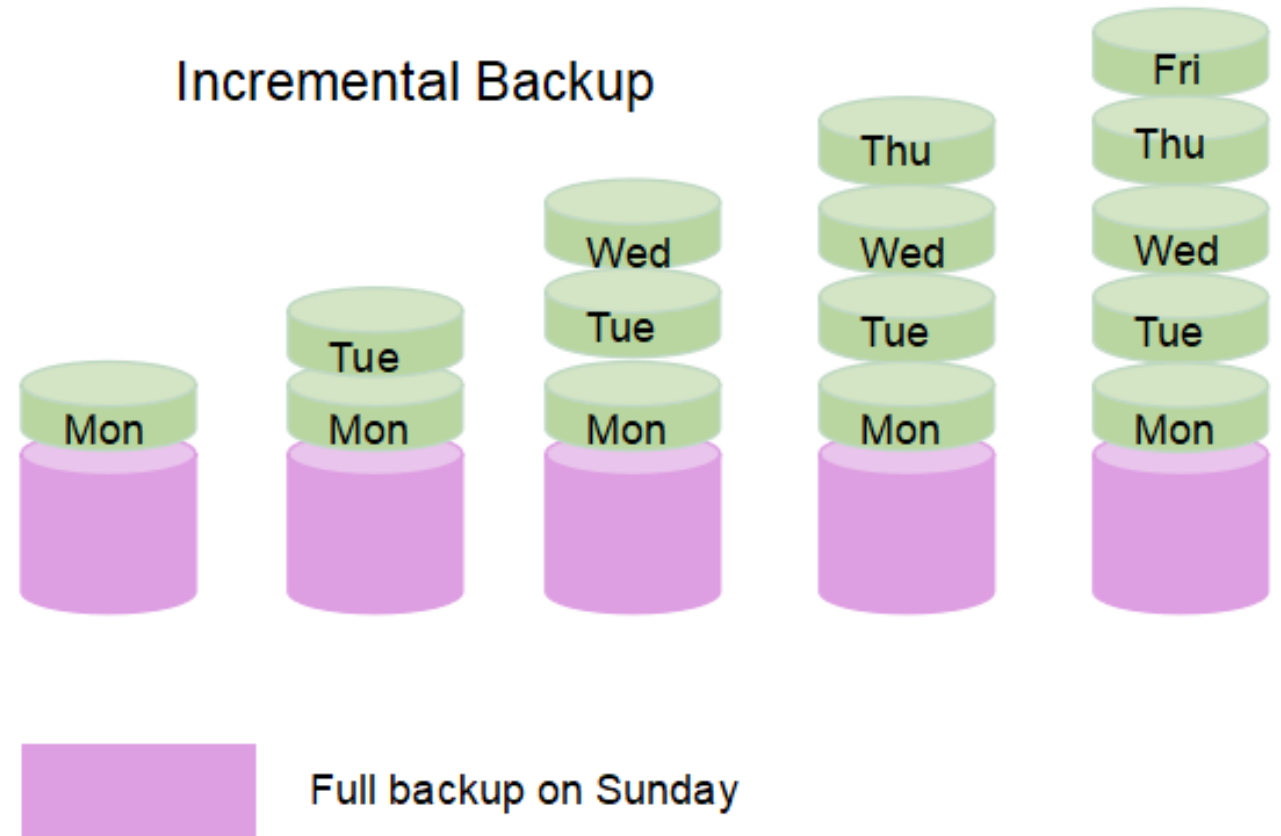
Changes the archive bit

Fastest type of backup

Longest restore

- You will have to restore the full
- Plus every differential in order

Takes up less storage space



Snapshot

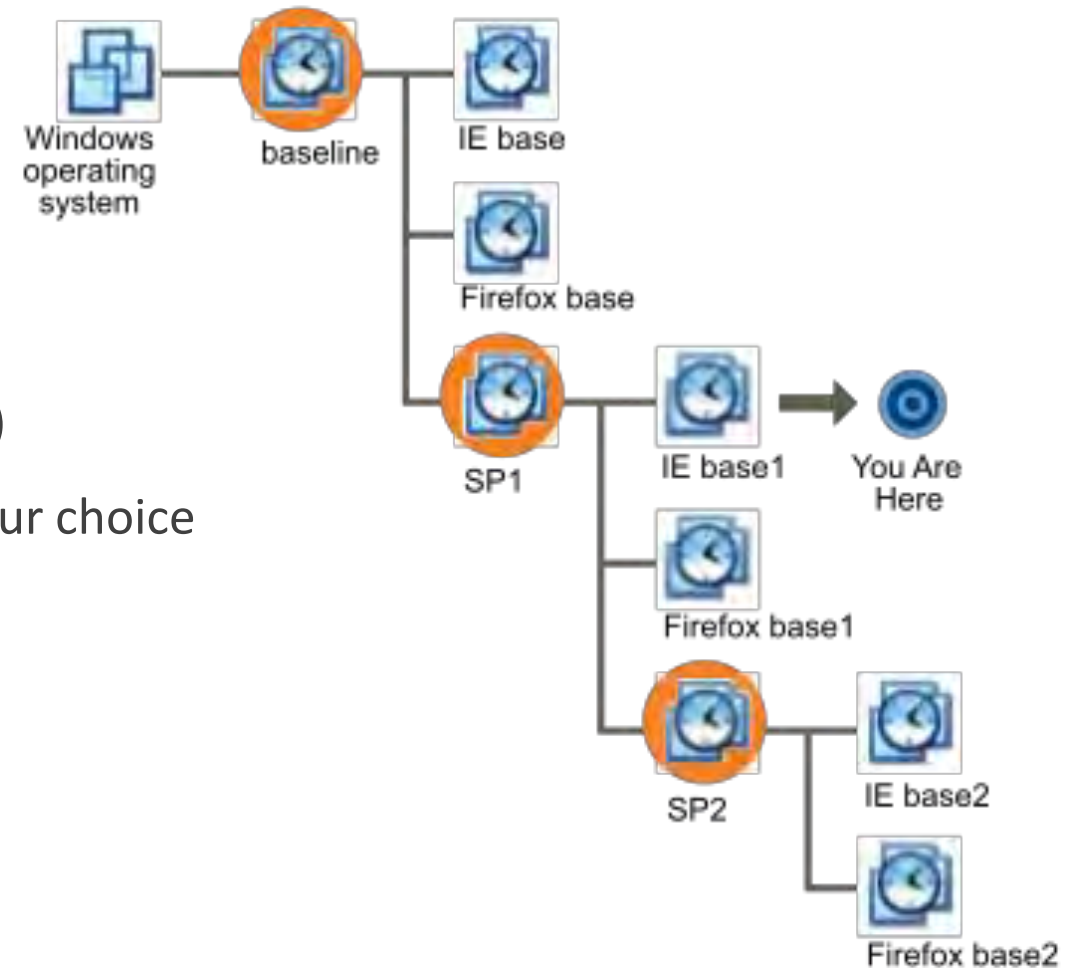
Used to restore a virtual machine to a previous state

An image of the state of a VM at a point in time

Typically requires the base image plus the snapshot(s)

As with backups, you can revert to the snapshot of your choice

Should not be your only backup solution



Network Device Backup/Restore

Device Configuration

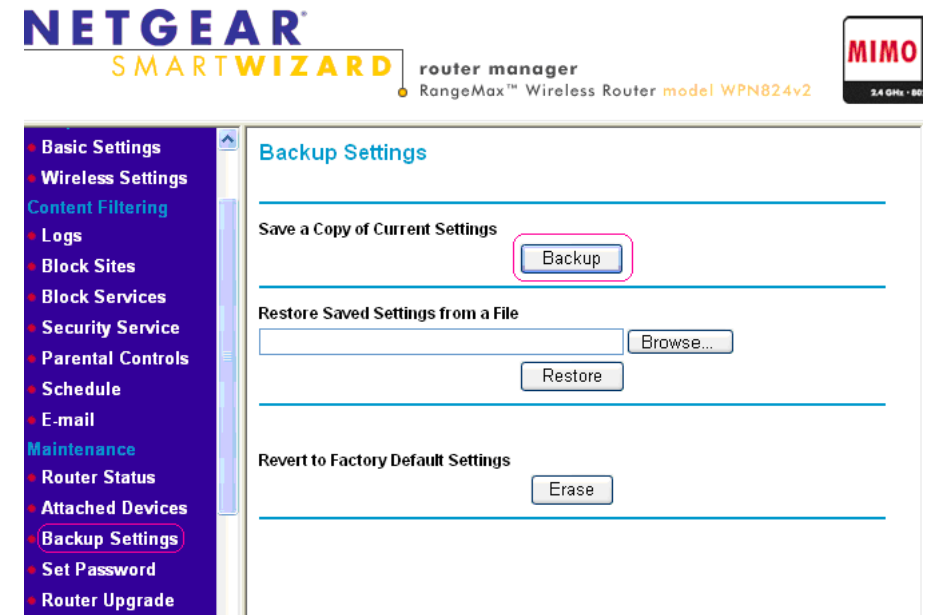
- AKA startup configuration
- Loaded when the device reboots
- Starts the device in a “clean” state

Device State

- Back up the startup configuration
- Also back up the current dynamic information that would be deleted if the device rebooted
 - E.g. route table, dynamic firewall rules, ARP cache, MAC table, NAT table, etc.

Most devices have their own backup utilities

- GUI (such as a browser) to download the backup to your PC
- Command line, typically using TFTP to backup to a TFTP server on the same segment



Backup Site

A location where the organization can relocate to following a disaster

Often hosted by a company that specializes in disaster recovery services

Site can be “cold”, “warm” or “hot”

Site can also be in the cloud

Some companies engage in a “reciprocal agreement” with a competitor

- Personnel from one company use facilities and equipment at the competitor site
- Strategy often used in the print media business
- The guest group operates at minimum productivity to not overburden the host group

Some specialty companies will even bring a container or trailer to you

- Has enough equipment and connectivity to act as an emergency office or comm center

Cold Site

Empty building

No equipment

Power

Security

Telecommunications (dark fiber)

Cheapest

Takes longest to bring online



Warm Site

Contains most of the hardware needed to recreate your current data center

- Might be a smaller version of the production site

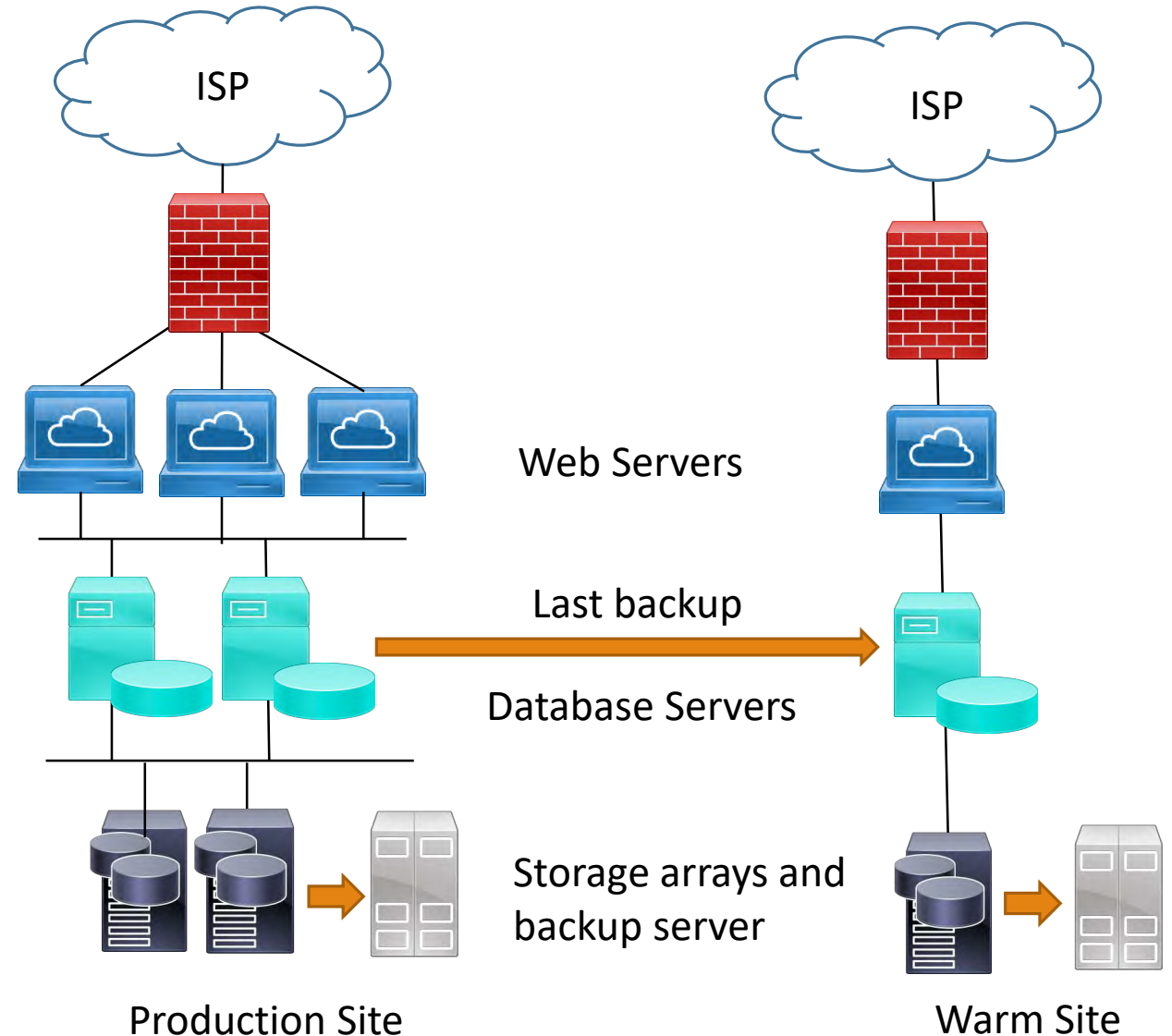
Restoring service requires the last backups

- From offsite storage
- Carried with admin during evacuation

Can be restored more quickly than a cold site

Could end up with outdated equipment

More expensive than cold site



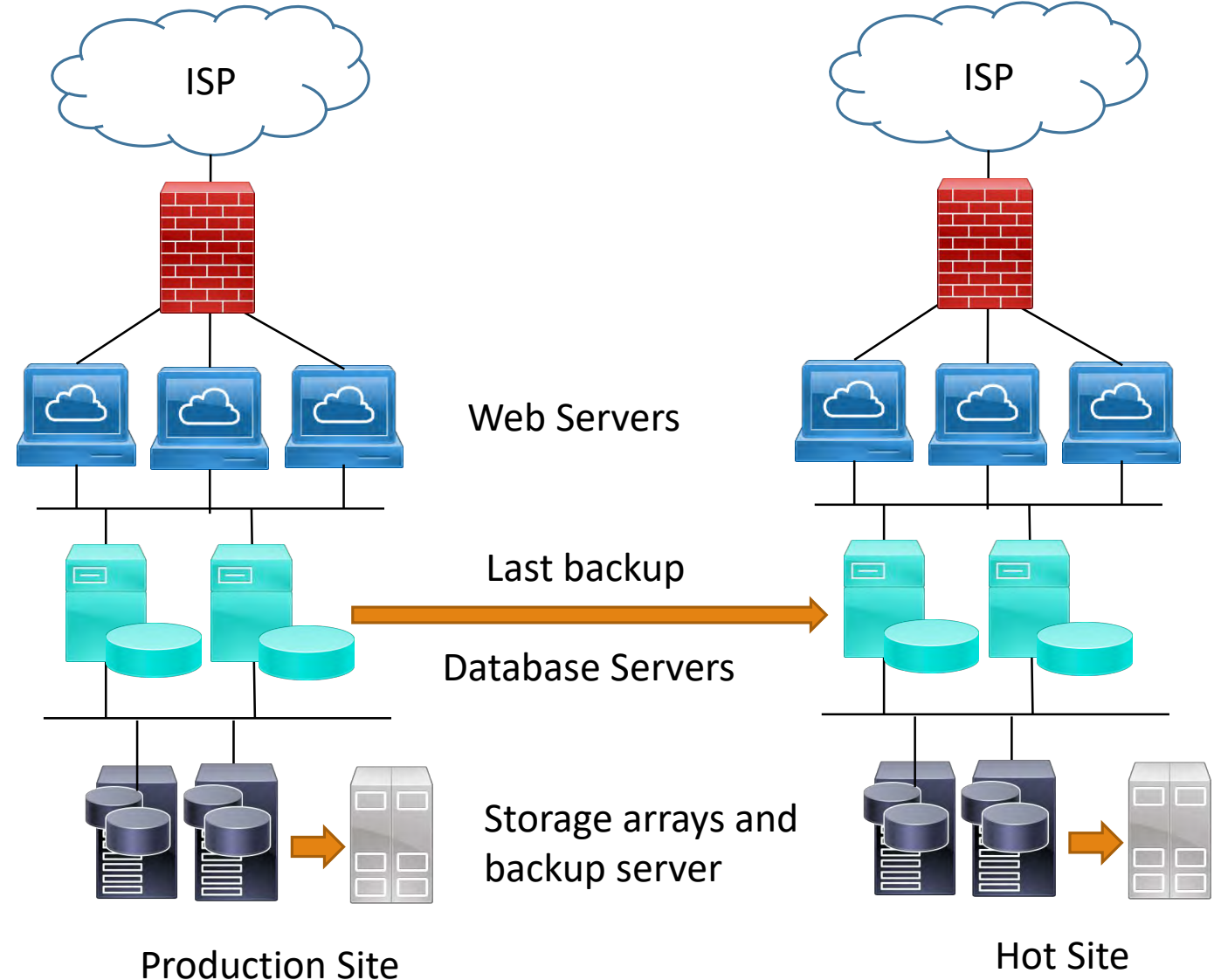
Hot Site

A practical mirror image of your current data center

All systems configured and waiting only for the last backups of your user data from your off-site storage facility

Up to full production in a few hours

A hot backup site is the most expensive approach to disaster recovery



Mirror Site

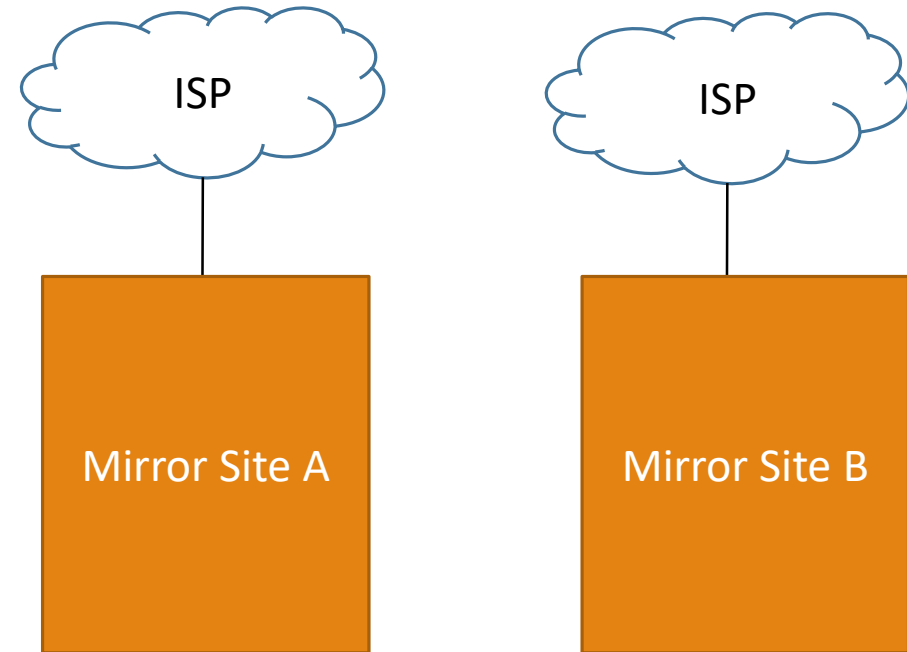
An always-ready hot site that exactly mirrors the production site

They either share off-site storage, or continuously replicate data to each other

You regularly “flip” between the two sites to ensure total readiness

Cutover should be seamless and practically instantaneous

In some implementations, the two mirror sites are active / active, with either ready to take on the load of both



Cloud Site

A cloud site can be your backup destination

- Slower, but highly secure and cheaper than maintaining backup servers and storage

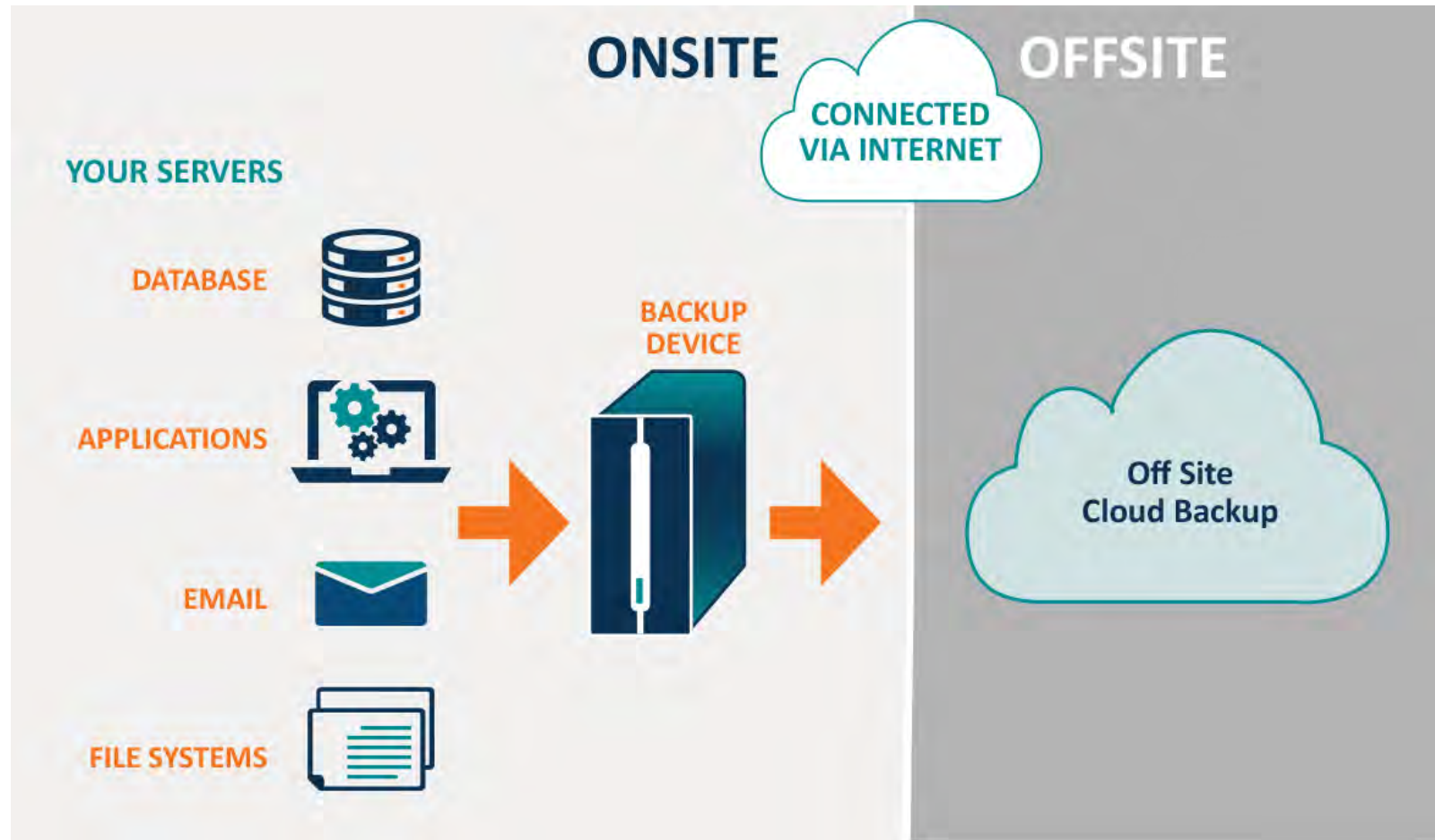
You can also extend your on-premises network into the cloud

- Replicate data to cloud storage
- Deploy additional authentication, email, database, etc. servers in the cloud

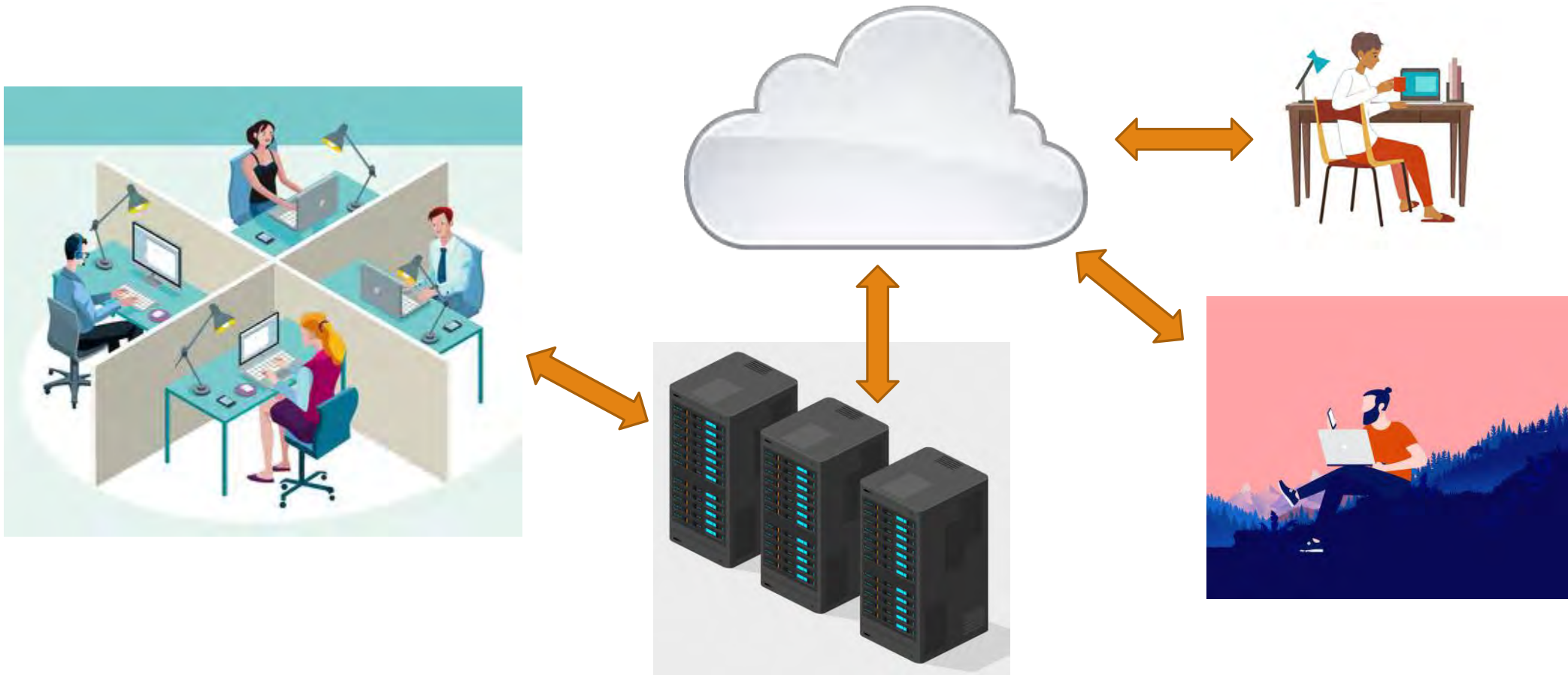
Relieves the organization from having to physically relocate to another site

- Users can telework and still access their data and network services
- You need not have a disaster for users to benefit from a cloud site

Cloud Backup Site Example



Extending On-Premises into the Cloud





Facility and Infrastructure Support



Power Management

Planning for fault tolerance and disaster recovery must include power management

What will you do if the power goes off in your area?

Most business equipment can be shut down gracefully or run for awhile on backup power

Be sure to always prioritize communications if you have limited backup batteries

Ensure that you do not sacrifice user or device safety if the power goes out

Battery Backup/UPS

Uninterruptible Power Supply

- Keeps the network running just long enough for graceful shutdown of servers
- Should have a USB, RS-232, or network link to the servers to send a shutdown signal
- Servers should have software to listen for shutdown signal

Protects hardware from damage

- Devices actually run off the battery
- Battery is constantly being charged
- Protects from blackouts, spikes, surges, sags, brownouts, and dirty power

Enterprise UPS will be directly wired to power on a dedicated line



Power Generators

Backup power supplies that kick on during a power outage

- Usually a short delay before generator starts and is running

Can run from a variety of sources:

- Gasoline /diesel
 - If power stays off too long, could run out of fuel
- Solar / wind



Dual Power Supplies and Redundant Circuits

One or more of following

- Building circuits include two separate dedicated circuits
 - “A” feed (120V, 20A)
 - “B” feed (120V, 20A)
- UPS Systems include two separate UPS systems
 - “A” UPS (120V, 2200VA)
 - “B” UPS (120V, 2200VA)
- Server Redundancy allows separate “A” and “B” inputs
 - “A” input connects to “A” UPS powered by circuit “A”
 - “B” input connects to “B” UPS powered by circuit “B”
- Redundant power supplies on a network device



Power Distribution Units (PDUs)

A long, specialized power strip

Meant to mount in a rack

- In a U-slot or along the side

May provide different outlet types

May have an on/off switch for the entire unit

May include power meters, surge protection, and line conditioners for network cables



HVAC

Datacenters require special HVAC design

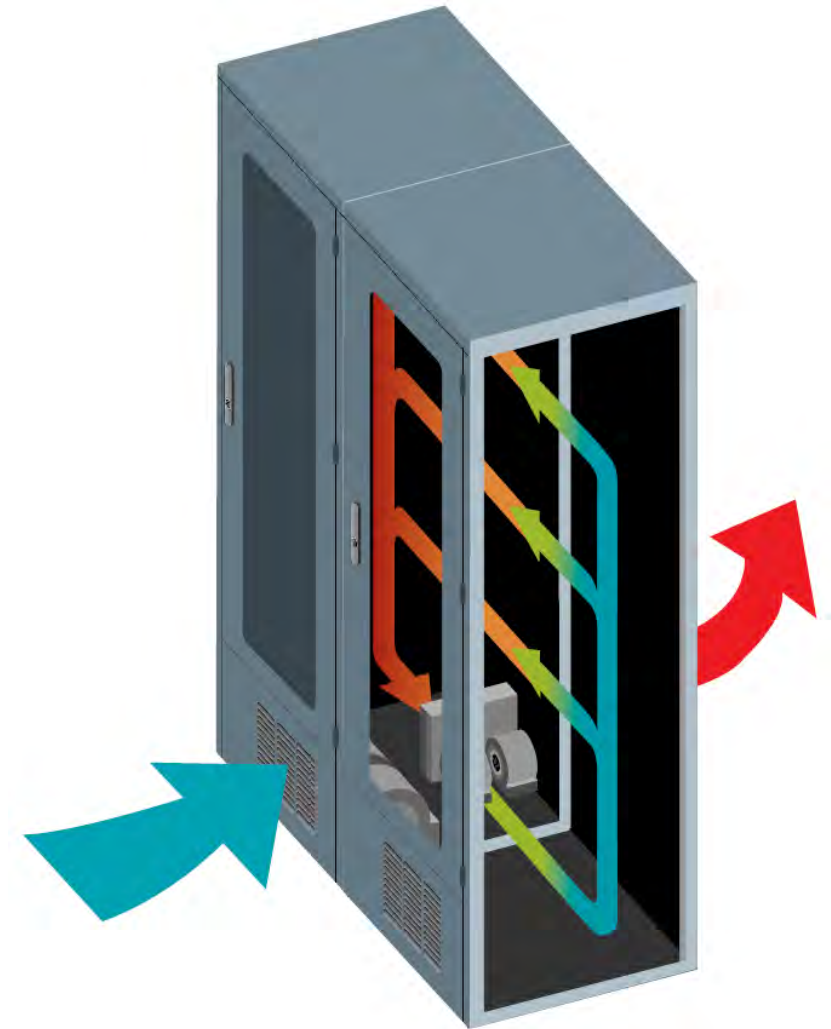
- Computers should be kept at 68 - 71 degrees F
- Recommended don't go below 50 F
- Or above 82 F
- 40 - 50 % humidity

Maximize efficiency of the cooling system

1. Draw in cold air
2. Blow through device intakes / output vents
3. Push out hot air

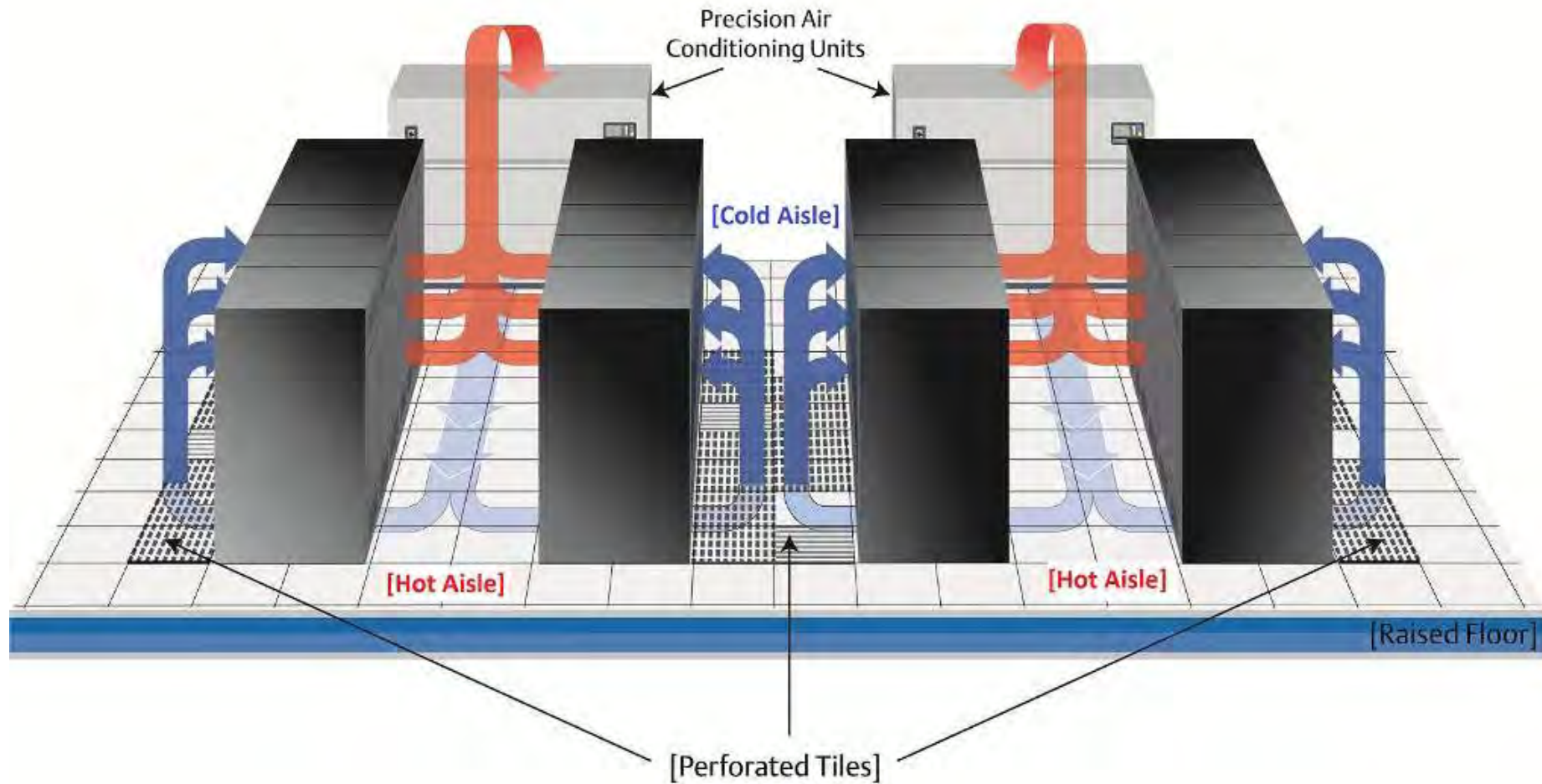
Try to avoid mass cooling the entire datacenter

Many datacenters use a hot aisle / cold aisle design



Cooling-efficient Rack

HVAC Hot Aisle / Cold Aisle Example



Fire Suppression

A fire needs the following:

- Ignition
- Fuel
- Oxygen
- Heat
- Chemical reaction

If you can disrupt any of these, you can put out the fire

There are three levels of fire protection for the datacenter:

- Building level
- Room level
- Rack level



Building Level Fire Suppression

Sprinkler and handheld fire extinguishers

There should be a handheld extinguisher for every 3000 square feet if class A combustible materials (cloth, wood, paper, rubber, and plastics) exist in the work area

The building can use firewalls and fire-rated floor assemblies to delay the spread of fire



Room Level Fire Suppression

You can use sprinklers, water mist, aerosols, or gas

Aerosol systems use strategically placed canisters that discharge when fire or smoke is detected

- Aerosol particles contain potassium nitrate that disrupt the fire's combustion process

Gas can be released at high velocity into the datacenter

- Preferred over water-based systems which can damage equipment

You can use inert gas (argon mixed with nitrogen) to displace oxygen

- Requires a large amount of gas to fill the room

You can also use a “clean agent” gas such as Novec 1230 (fluorinated ketone) or FM-200

- Absorbs heat and/or displaces oxygen to put out the fire
- Requires less gas to fill the room than inert gas

Note: Halon gas was banned in the 1990's due to its destruction of the ozone layer

Aerosol Fire Suppression Example



FM-200 Fire Suppression Example



Rack Level Fire Suppression

A direct-release system includes in-rack fire detection and tubing that delivers the chemical suppressant

One system can be set up to provide for multiple racks

