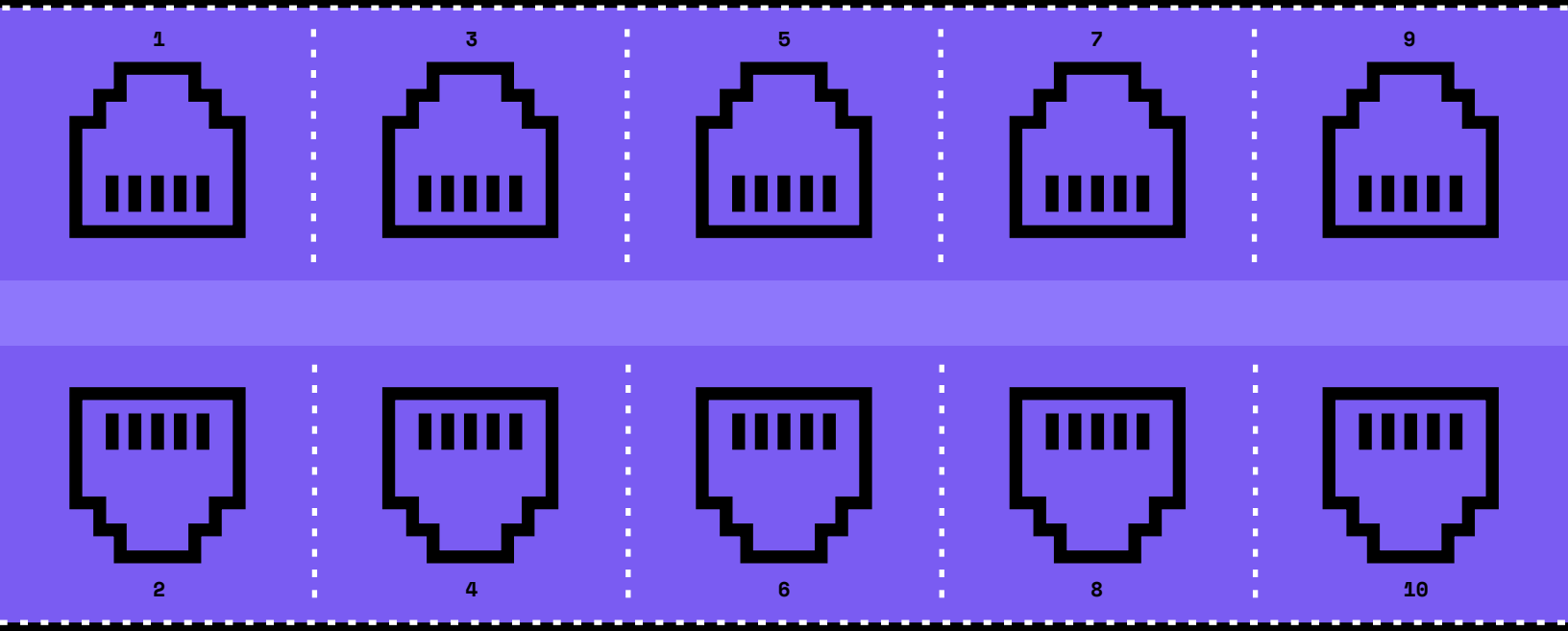
Auvik



The IT pro's guide to network switch physical design configuration

Ethan Banks

Ebook

About Ethan Banks

Ethan Banks is a network architect, independent IT writer, frequent conference speaker, and co-host of the Packet Pushers podcast. Follow him on X @ecbanks.

About Auvik

Auvik is a cloud-based IT management platform that helps IT departments proactively manage their networks, endpoints, and SaaS applications.

The key is absolute simplicity: seamless deployment, an intuitive interface, and effortless automation.

The result is less friction for IT departments, so that everyone can work however and wherever they want.

Contents

Introduction

An on-site switch plays a critical role in a network: that of connecting users to the rest of the IT infrastructure.

Don't think of that jack in the wall or under the cubicle as a simple Ethernet port. Rather, think of it as the mission-critical gateway to IT services that it is.

When a user plugs their workstation into the network, it's their single connection to their email, instant messaging, financial system, sales engine, and other company resources. Even voice communications are likely provided through that jack. If the service is poor, that user can't do their job effectively. If enough users have a bad experience, the entire business is affected.

Sadly, in business networks, the on-site switch is often neglected. Read on to learn the basics of physically setting up and configuring a reliable switch that will serve users well.

Part 1: physical design

Physical location

Stuffed into a hot, poorly ventilated corner to provide connectivity to an office area, a closet switch is all too frequently treated as a cheap commodity item that simply doesn't matter much.

I've seen switches in bathrooms, stuffed inside drop ceilings, hanging by a single screw on the wall, underneath cubicles, and on filthy shelves in a tiny closet with no airflow or climate control.

The attitude seems to be, Does it work? Are the users getting connected? Well then, good enough.

It's possible to get away with this sort of approach if all you're looking for is rudimentary connectivity, but "cheaper is better" is a bad idea for any organization whose employees rely on the network to get their jobs done. The network is too critical an element to try to save money in the network closet.

I recognize that retrofitting aging buildings with wiring infrastructure is a challenge, and sometimes the solutions we're stuck with aren't ideal. I've been involved in many such installations.

That said, try your best to get the switch into a place with, at the least, airflow. Study the switch you're installing, figure out where the air intake and exhaust locations are, and keep them clear.

Even for fanless switch designs with external power supplies, it's usually a bad idea to put a switch right up against a wall. There's still heat generated that the chassis needs to radiate, which is part of the reason most switches come with rubber feet that can be applied to the bottom.

Dust and dirt are also concerns. Switches get clogged with crud when they're installed in filthy locations, and this can shorten their lifespan due to overheating. Switches in dirty places should be cleaned periodically to reduce this risk, even industrial models that are built for difficult environments.

"Cheaper is better" is a bad idea for any organization whose employees rely on the network to get their jobs done.

Uplink design

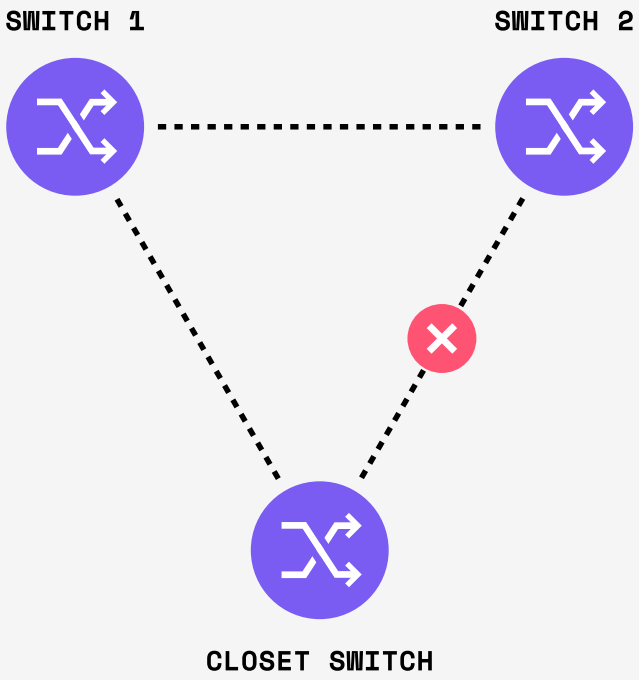
Consider the uplink design between the closet switch and the rest of the network carefully. While the simplest thing to do is to connect a closet switch via a single cable back to the rest of the network, a dual uplink is preferable for redundancy and possibly capacity. There are several ways to safely achieve a dual uplink.

A backup link

A backup link is created when a second line is connected in parallel to the primary line. Topologically, this makes a loop between the closet switch and the uplink switch. Spanning tree will detect the loop and block the backup link. If the primary link fails, the backup link becomes active.

A better alternative to this design is to connect to multiple switches rather than having both links running to the same switch. Spanning tree will behave the same in either case—a loop will be detected and one link blocked until the primary path fails.

> A dual uplink is preferable for redundancy and possibly capacity.

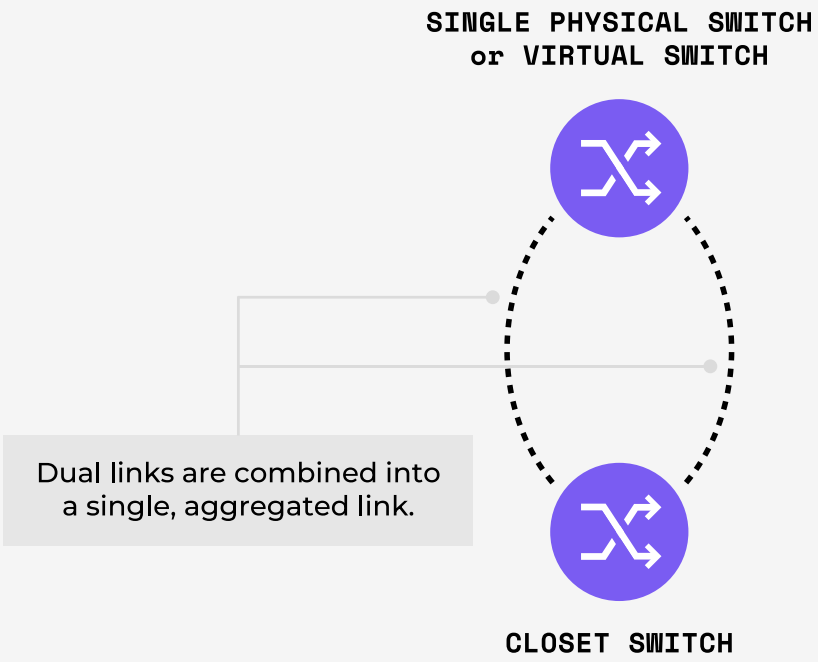


A parallel layer 2 link

A parallel link between switches can be achieved with link aggregation protocols such as Cisco's Port Aggregation Protocol (PAgP) or industry-standard Link Aggregation Control Protocol (LACP). This scheme allows for both links to be active and carrying traffic.

The link aggregation protocol makes the two links appear as a single link as far as spanning tree is concerned, while still maintaining a loop-free topology. Link aggregation can scale parallel links beyond two. Four- and eight-way link aggregation bundles are common.

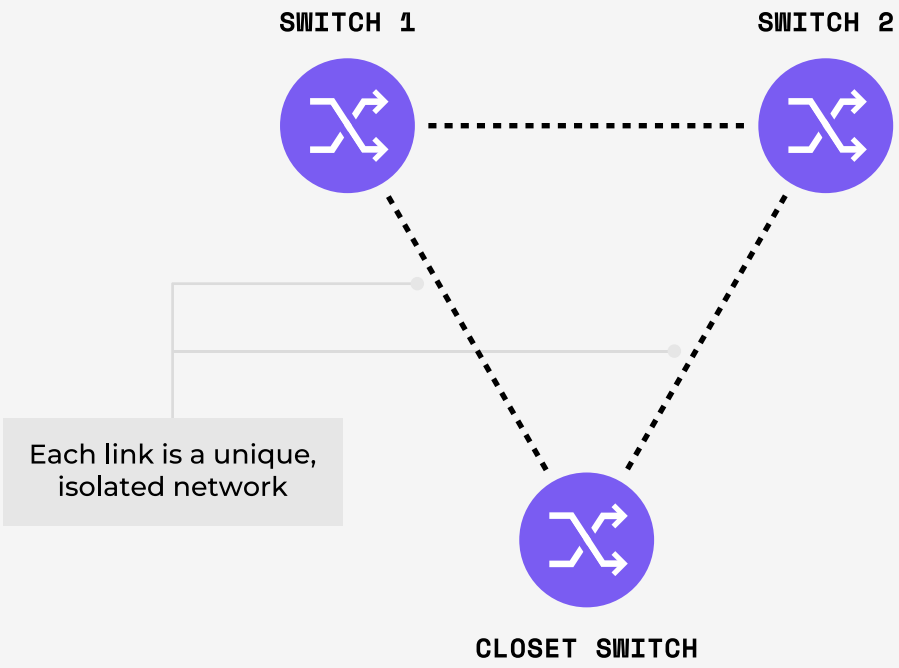
For diversity, link aggregation bundles can be split across two physical switches that act as one virtual switch, such as the stackable versions of Cisco's Catalyst switches or Cisco chassis switches running Virtual Switching System.



A parallel layer 3 link

One strategy for uplinking closet switches is to connect them to the rest of the network via routed (Layer 3) links as opposed to a switched (Layer 2) link. While the wiring looks the same, the end result offers better isolation from the rest of the network.

With this approach, the two L3 links don't create a loop because each link belongs to its own network segment isolated from the other. The challenge of this design is that user VLAN segments can't span to different closets, as an L2 network segment (a VLAN) will not extend beyond the L3 uplinks. Dual L3 uplinks should connect to separate switches for resiliency.



Does the network need 10Gbps Ethernet?

You might be tempted to upgrade the network to 10Gbps Ethernet because it's faster. And faster is better, right? But it's a good rule of thumb to never do anything with technology "just because." Think about whether users really need that extra bandwidth and whether the cabling and budget supports it.

When you're dealing with closet switches, it's crucial to appropriately size the uplink from the closet to the rest of the network.

While every network is different, in general, closet switches can tolerate an enormous amount of oversubscription. By oversubscription, we mean the ratio of user-facing ports to uplink ports.

For example, in a switch with 48 1Gbps user-facing ports and two link-aggregated 1Gbps uplink ports, the oversubscription ratio is 24:1. In other words, there are 24 user-facing ports for each uplink port.

Over the years, I've seen very high oversubscription ratios for closet switches—as high as 96:1—with no problems.

The traffic patterns from user workstations tend to be both bursty (a quick spurt of traffic) and unsynchronized (traffic bursts happening on one machine at a time, and not all at once). For these reasons, it's possible to get away with high oversubscription rates for a closet switch on-site at a business that wouldn't be acceptable in most data center designs.

The question then arises: Are 10Gbps Ethernet uplinks a requirement between the closet and the rest of the network? The answer is that it depends. Let's consider a few important facts that could tip the balance in favor of 10Gbps uplinks one way or the other.

> Never do anything with technology "just because."

Can you afford 10Gbps Ethernet?

Closet switches must very often traverse scores or hundreds of meters to uplink back to the main network. To go long distance, fiber-optic cabling is required. To go exceptionally long distances, fiber of a particular optical quality is required, and possibly a different sort of transceiver. So while it's not as expensive to deploy 10Gbps links as it once was, cost is definitely still an important consideration.

Not all fiber is created equal; not all 10G optical modules push light the same distance. I've fortunately been able to re-use existing multi-mode 62.5 micron cable for 10G uplinks, despite slightly exceeding the recommended maximum distance. Cisco-branded transceivers in particular have been good to me this way over the years, frequently outperforming their specifications. Your mileage may vary.

Do you have copper cabling?

10GBase-T is 10Gbps over copper cabling. Its primary purpose is not for closet uplinks. Cisco mentions, "The primary use case for 10GBASE-T is high-speed server connectivity. Other less common scenarios use 10GBASE-T for interconnecting distribution or core switches that reside within a 330-feet (100-meter) distance."

Not only is 10GBase-T intended mostly for in-rack connectivity between servers and access switches, but it also has stringent requirements for the copper cabling.

Over distances of less than 100m and with the correct type of copper cabling (Cat6 certified to 500MHz, shielded Cat6, Cat6A, or Cat7), it may be possible to use 10GBase-T as a closet switch uplink. But 10GBase-T is not an obvious, cost-effective answer. Ethernet's leap from 1Gbps to 10Gbps over copper is a higher one than the leap made years ago from 100Mbps to 1Gbps.

Do your users move lots of data around?

As we've discussed, high rates of oversubscription are often fine for a switch. But some user groups might be moving more data around than the average worker. Folks that work on their local workstations with large datasets fit that category. Think multimedia artists, developers working with large test databases, and similar scenarios.

In those situations, a high oversubscription rate will be less tolerable, as longer bursts of traffic will increase the likelihood of simultaneous user traffic streams hitting the uplink. In that case, 10Gbps uplinks reduce the oversubscription by a factor of 10 when compared to 1Gbps uplinks, and make contention much less likely. Lower contention equals faster network throughput for the user community.

How many supported devices do you have on the network?

Most organizations now support employee phones, tablets, and related wireless devices on their LANs, which means the sheer number of devices on the network has increased.

Today's average network user might represent two or three devices instead of a single workstation. As device count increases and access points proliferate, the likelihood of uplink contention also increases. 10Gbps can be helpful here.

Bottom line: Think about whether you really need the extra bandwidth of 10Gbps and whether the cabling and budget supports it. It's entirely possible that an existing 1Gbps connection suits the job just fine.

Physical path diversity

When possible, route cabling coming into the closet by alternate physical paths. The idea is that if one cable connecting the switch to the rest of the network is cut, flooded, burned, or otherwise damaged, the other cable won't share the same fate.

Separate paths are admittedly challenging in buildings where conduits are limited and where construction is just not amenable to such an approach. But when possible, keep the uplink cables separate.

When possible, route cabling coming into the closet by alternate physical paths.

Physical ports

A problem I've run into repeatedly is a 24-port switch that runs out of ports. Budget permitting, a 48-port switch—preferably with additional ports to serve as the uplinks—should always be installed to allow for additional growth.

Yes, 48-port switches cost more, and that cost is often the driver for the purchase of a 24-port switch to begin with. But in the long run, 48-port switches are almost always preferable in my experience.

That said, consider the long-term impact of wireless networking in the environment. If you're finding that 802.11ac provides wired-equivalent performance, then you're on a different track.

You're moving away from wired connections to wireless, where your wired port density concerns have shifted from user workstations to access points. But if "wires everywhere" is still the choice, then my comment about 24- versus 48-port switches is hopefully thought provoking.

Switch stacks

A popular choice in the network closet, stackable switches and chassis switches offer a great deal of port density with a single point of management. From a design perspective, these advantages have a few challenges worth keeping in mind.

Single point of failure

Stacks and chassis are single points of management, but can also be single points of failure. To decide if this is a concern, consider the impact to the organization if the entire closet was out of service for several hours or days due to a catastrophic system failure.

Chassis switches often offer dual supervisor engines and dual power supplies to mitigate this risk. Some might point out that the chassis switch itself is a single point of failure, but in nearly 20 years of networking, I've only ever run into a chassis failure once.

Chassis are all in

Stackables are generally better off than chassis, in that each physical switch in the stack can usually function as the "stack master," where a new stack master will be elected in the case of a failure.

In addition, power is distributed throughout the stack; a failure of the switch's power supply will only impact one switch, and not the rest of the stack. Interestingly, certain Cisco stackable switches offer StackPower, which can mitigate blown power supplies in a stack.

Stackable switches and chassis switches offer a great deal of port density with a single point of management.

Uplinks should be spread across the stack

Dual uplinks should be spread across the chassis or stack to maximize resiliency. The idea here is to make sure the uplinks from a chassis or stack do not come from the same switch or module.

Too often, I've seen the dual uplinks coming from the same supervisor engine, meaning that if the supervisor fails, the chassis might be disconnected from the rest of the network, even with dual supervisors.

Stackable switches have the same concern. Dual uplinks should be spread across different physical switches in the stack. My practice with closet switch stacks is to place one uplink at the top of the stack and the second at the bottom. Assuming a break in the stack between the top and bottom, this means that both parts of the fragmented stack will still uplink back to the main network

Software upgrades can be challenging

Software upgrades are sometimes an all-or-nothing affair. When upgrading chassis switches, you might need to reload the entire chassis to bring up the new software, meaning users aren't able to access the rest of the network until the chassis is back online.

Cisco and other vendors offer in-service software upgrades (ISSU) that can mitigate that issue, but there's usually a specific hardware requirement such as dual supervisor engines to have ISSU capability.

You might intuitively assume that switch stack upgrades can be completed one switch at a time, but reality is that switches in a stack often require very close software versions to be members of the same stack. Therefore, upgrading one switch might render it offline until all the other switches in the stack have been upgraded as well.

This issue varies from vendor to vendor and product to product. The point is to be sure the organization is able to cope with a software upgrade process that potentially takes hundreds of user-facing ports offline while the upgrade is going on.

Upgrading one switch might render it offline until all the other switches in the stack have been upgraded as well.

Part 2: configuration

Spanning tree configuration

If you chose the backup link (p. 6) or parallel Layer 2 design (p. 7), then you've extended your Layer 2 domain from the core network into the closet. That means your closet switch is participating in the global spanning tree domain and needs to be configured appropriately.

When dealing with a closet switch, the key is ensure that the closet switch doesn't become the root bridge.

If a closet switch does become the root bridge, then links in or around the physical core of the network could end up blocked. The result could be some odd (and unexpected) forwarding paths in the network that traverse the closet switch.

To help prevent this scenario, my recommendation is to set the root bridge priority to a very high number (the max of 65535 works), rather than leave it at the default value of 32768. If you've already set your core switches to some low value below 32768, changing them might seem unnecessary, but thinking ahead is critical. Setting the value to a higher number could help avoid an unexpected spanning tree result in future.

Also worth mentioning is that I'm assuming the use of rapid spanning tree. 802.1w is a significant rewrite of the original 802.1d spanning tree, and features a number of performance-related enhancements. Rapid spanning tree includes equivalents of several early Cisco spanning tree enhancements such as uplinkfast that will help your switch converge on a new topology more quickly if something changes.

The key is ensure the closet switch doesn't become the root bridge.

Practically speaking, let's say you've implemented the backup link scheme where one link is blocking and another forwarding. If the forwarding link goes down, your switch will notice and react.

Eventually, traffic will begin to flow across the remaining link. How long "eventually" is varies between original spanning tree and rapid spanning tree, with rapid spanning tree being faster.

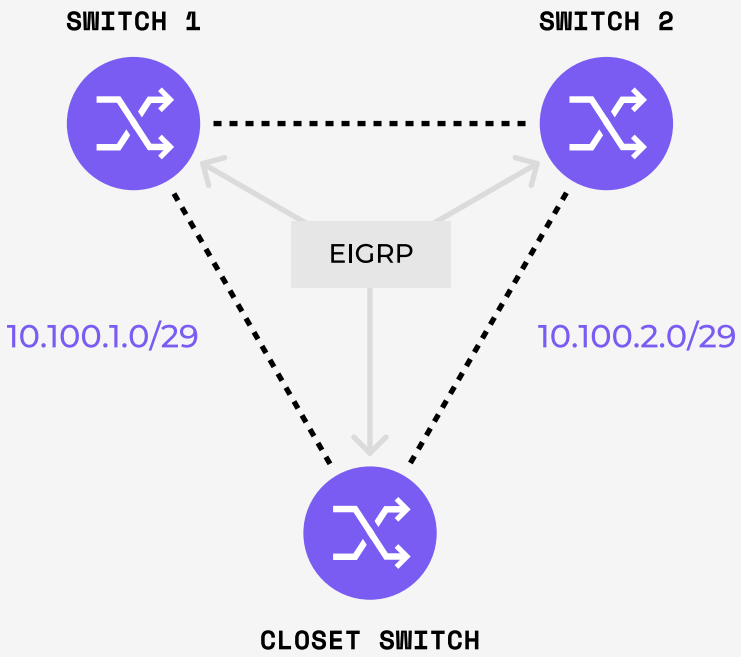
If you chose the parallel Layer 3 link design (p. 8), then you've created an isolated Layer 2 spanning tree domain on the switch itself. The question becomes whether or not the closet switch needs to be the root of the spanning tree domain.

In most cases, you'll want it to be the root because you don't want other switches that are plugged into the network, say, at a user's desk, to become the spanning tree root bridge of the network closet domain.

Routing configuration

Many closet switches are Layer 3 switches, meaning they have the ability to route traffic between different blocks of IP addresses resident in separate VLANs. In the Cisco Catalyst product line, 29xx series routers are usually Layer 2 only, while 37xx and higher model switches are Layer 3 capable.

Assuming an L3-capable switch running parallel links, there are a few different routing schemes you could choose. Rather than go through all the different possibilities, let's review one solid approach worth considering in a Cisco environment—that of dual EIGRP links.



In this design, each L3 link is assigned a /29 network block. A /29 is a block of eight addresses, six being actually usable. In our example, the link to switch 1 has a range 10.100.1.0-7, where IP addresses 1 through 6 can actually be used.

I recommend /29s over more customary /30 point-to-point links, which only offer two usable addresses. /29s allow for future flexibility without renumbering the link, such as when an additional device like a firewall, WAN optimizer, or replacement switch might need to be added.

IP address conservation is usually not an issue on private networks using RFC1918 address space, so a /29 is a useful addressing scheme for point-to-point links without going overboard.

In this scenario, our closet switch has two EIGRP links, one to each of a pair of switches back in the core network. Assuming both links have the same bandwidth and delay characteristics, the EIGRP routing process will see these links as equal costs, and load balance traffic between the two links. In the event that one of the links goes down, EIGRP will detect that there's no longer a connection and converge on the remaining link.

A couple of additional notes here.

1. The closet switch should very likely be configured as an EIGRP stub router. Unless you've made some unusual network design choices, there's no reason a closet switch should ever be a **transit router**.

   The only networks for which a network closet is the source are the connected VLANs accessed by users. So, there's no reason for the core network to query the closet switch about lost routes originating in other parts of the network.

2. Another way to think about it is to consider the access switch a dead end. If the edge of the network diagram is the closet switch, then that switch is indeed a dead end—nowhere further to go. And thus, configuring it as a stub is exactly the right thing to do.

   The closet switch doesn't need to know the entire routing table. All the closet really needs to know is a default route. Why? In our sample design, the only place a closet switch can send anything is into the core network. A granular routing table is only useful to a router (or Layer 3 switch) if traffic for some IP destinations are reachable via one link, and traffic for other IP destinations are reachable via other links.

   In our case, that's not true. Therefore, why clutter up the closet switch's routing table with a bunch of IP destinations that are all reachable via the exact same two links? Instead, the core switches can summarize traffic into a default route using an "ip summary-address eigrp" statement on the interfaces that uplink to the closet switch. The closet switch will only learn a default route instead of the entire core routing table.

The closet switch doesn't need to know the entire routing table.

Quality of service considerations

Another big topic that we can only give an introduction to here is Quality of Service (QoS). QoS is the big idea that some network traffic is more important than other network traffic.

Important traffic (like voice and video, for example, although it could be anything you define) must be delivered, while other traffic can tolerate some loss.

For this ebook, let's make a few high-level observations:

If you're not running IP phones or video devices through the closet switch, you can probably do okay without a QoS scheme.

1. If you're not running IP phones or video devices through the closet switch, you can probably do okay without a QoS scheme. Without voice and video traffic, what you're sending through the switch is all data traffic, most likely TCP traffic. In the rare case of congestion between a closet switch and the core network, the retransmission and sliding window mechanisms built into TCP can cope. In most situations, this is sufficient.

2. Exceptions to this general rule include interactive applications like SSH, where a delay due to congestion could make the application difficult to use. In such a case, a QoS scheme might be useful to prioritize that interactive traffic.

3. When running real-time voice and video through your switch, probably because you have IP phones rolled out to the organization, the idea is to send voice traffic—people talking on the phone—out a low-jitter "priority" queue, and video traffic out a queue that guarantees enough bandwidth for its needs. Jitter refers to the amount of time in between packet delivery. For voice traffic, delivering packets evenly is important for a good quality call. Video traffic is more tolerant of jitter than voice, but all the packets need to get where they're going.

   How, exactly, to achieve a correct QoS configuration for your switch varies, sometimes dramatically, by switch platform. Different switches have different sorts of switching silicon inside of them, and not all of the queueing structures and capabilities are the same. Therefore, the commands, while often similar, will vary from switch platform to switch platform. Cisco has been working to unify their QoS configuration tools across platforms, but there are still differences to be wary of.

Other configuration considerations

Once the physical connectivity and forwarding design have been nailed down, a smart next step is to ensure the design continues working as intended. There are a couple of major points to consider.

Security

Security should be put in place to prevent switch configuration from being altered by an unauthorized party. Note that a switch configuration left to its defaults is not necessarily secure, so the wise administrator will take care to enter non-default usernames and passwords, and to disable unencrypted management protocols like Telnet, SNMPv2, and HTTP, instead using SSH, SNMPv3, and HTTPS.

Another common sense security step is to use an access list to limit the source IP addresses that are allowed to manage the switch. This prevents, say, a user at his desk discovering the switch and trying to log in, or malware from attempting to send a new configuration to the switch via SNMP.

If you run a highly secure environment, you might consider more aggressive tools in the network closet. For example, IP source guard, dynamic ARP inspection, DHCP snooping, and port security are features that can be deployed to ensure that systems using the switch are who they claim to be, aren't performing roles they shouldn't, and are behaving normally. Implementation of these features is complex and beyond our scope here, but are well-worth investigating and understanding whether you ultimately implement them or not.

Security should be put in place to prevent switch configuration from being altered by an unauthorized party.

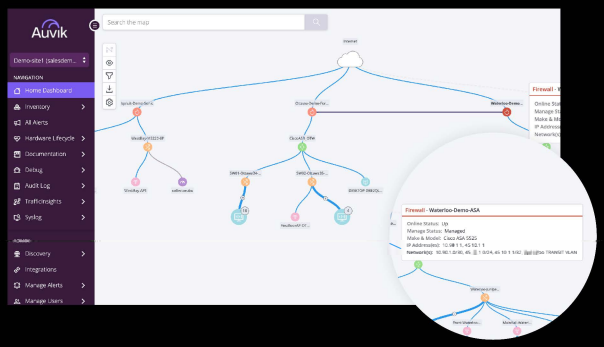
Monitoring

Another concern is monitoring of the switch to be sure it's up and running, that links (especially the uplinks to the core switch) are not being over-utilized, and that the switch ports are running error-free. Many network management products, including Auvik, can help with these tasks, handling monitoring along with reporting, troubleshooting, and configuration management.

The result of all this design and configuration work is a closet switch you can trust. You'll be able to rest a bit easier, confident in a switch that's quietly forwarding traffic from the closet in the very best way possible.

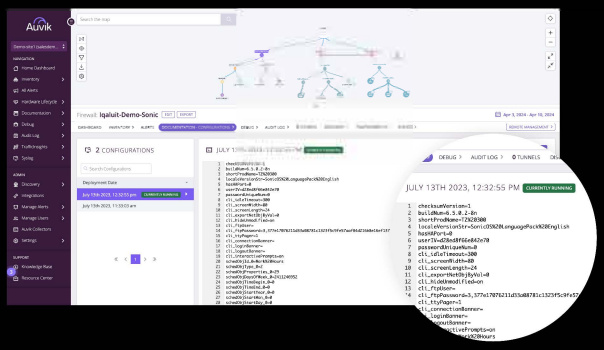
When networks run the world, network management is everything.

Gain true network visibility and control with Auvik.



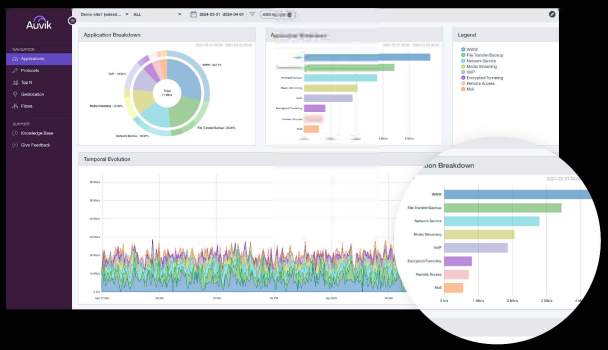
Real-time network mapping & inventory

Quickly discover & audit new networks. Then, stay in the loop—you'll always know exactly what's where, even as users & devices move.



Automated config backup on network devices

Mitigate network risk with no manual effort.



Deep insights into network traffic & flows

Quickly solve network bottle necks & spot potential security vulnerabilities.

Use free for 14 days

Auvik

auvik.com/try-now