CATO
NETWORKS

The Future of the SLA

How to Build the Perfect Network Without MPLS

The Anxiety of Today's Networks

Anyone considering SD-WAN or SASE for their global WAN transformation project must be a bit anxious about transitioning to Internet last-mile access. Compared to carefully controlled MPLS services, Internet access seems positively chaotic. Multiple ISPs need to be coordinated with no single throat to choke; no Service Level Agreements (SLAs) are there to guarantee service commitments.

Yes, enterprises gain in many ways — lower cost per bit, faster deployment of last-mile connectivity to new sites, quicker response to network changes, more innovation, and far greater network reach, to name but a few benefits.

Yet, there's no denying it. If you're not getting end-to-end SLAs, how do you ensure your providers deliver on what they're supposed to deliver?

The short answer? First, build the systems and technologies that guarantee your network won't fail and only then SLA the rest.

The Dirty Little Secret of SLAs

Ensuring network uptime is a far more effective approach than banking on an SLA. Collecting the penalties for missing an SLA has always been challenging. The contractual language limits the SLA scope. And, if penalties can be gathered, they'll never compensate the enterprise for brand reputation damage or revenue loss.

But what was an organization to do? There wasn't much of an alternative. Sure, the largest and most critical sites could justify fully redundant deployments, but such last-mile redundancy wasn't cost-effective for many locations. Running active/active last-mile connections with automatic failover (and failback) was science fiction. In short, SLAs were the best possible approach to a bad situation.

Use the Right Service for the Right Problem

SD-WAN changed everything and that was the first step in disrupting the legacy approach to designing reliable networks. SD-WANs separated the underlay – the underlying transport (Internet or MPLS) — from the traffic engineering and routing intelligence (the overlay), determining the use of that capacity.

With the overlay run by their SD-WAN devices, enterprises benefit from the competition that reduced Internet access costs and increased service diversity. Each location can be connected by the best available Internet service. And by connecting the site with SD-WAN devices to two, and ideally, three, active last-mile connections, enterprises improve network performance, reducing packet loss and increasing uptime, for even the smallest office.

CATO NETWORKS

National Technology Hub

Packet Loss

Across those connections, packet loss compensation technologies help loss-sensitive applications, such as VoIP. Where loss does occur, as we'll see, Cato further improves loss recovery by segmenting the connection into three parts – two last miles and a middle mile. With the TCP connection terminating at a Cato PoP, not a distant location, packet recovery operates far faster. In addition, the SD-WAN automatically manages path selection, instantly selecting the optimum link in the event of brownout or blackout.

Uptime

Running active/active connections with automatic failover/failback improves last-mile uptime. When coupled with diverse routing, where connections run over separate, physical last-mile infrastructure, SD-WAN achieves extremely high uptime targets, meeting and exceeding MPLS targets, even when using less-reliable Internet connections. This can be shown by the following formula for calculating total network availability using multiple links:

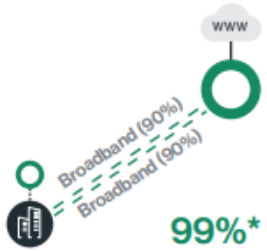
$$1-(1-A1)*(1-A2)$$

where "Ax" is the availability of a given last-mile connection.

CATO NETWORKS

National Technology Hub

How much uptime you want will depend on the site's importance. You can think of this in stages of increasing availability.

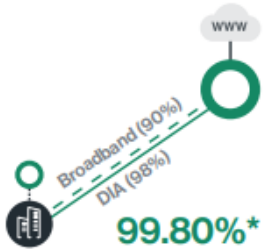


99%*

STAGE 1

Two Broadband Links

Broadband links typically have 90% uptime, but when aggregated, they achieve 99% uptime or no more than 7 hours and 16 minutes per average month.

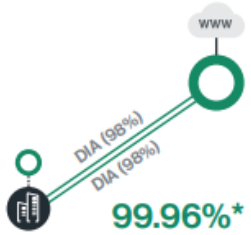


99.80%*

STAGE 2

Broadband and DIA Links

Improve the quality of one link, and uptime jumps to 99.8% of downtime of no more than 1 hour 27 minutes per average month.



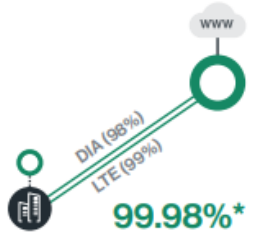
99.96%*

STAGE 3

Two DIA Links

Improve the quality of the second link, and uptime jumps to 99.96%.

(<17 minutes of downtime per average month).



99.98%*

STAGE 4

One DIA and LTE

Now, add a wireless connection (99% uptime) to a DIA link, and last-mile availability jumps to 99.98% uptime or no more than ~8 minutes of downtime per month.

(With dual DIA and LTE connections, availability goes to better than five nines).

CATO NETWORKS

National Technology Hub

Latency: A Story of the Middle Mile

Better Internet access services, though, do nothing for the performance of Internet core, the middle mile. The global Internet is notoriously unpredictable, and with Internet routing optimized for cost, latency is often far higher than with managed networks.

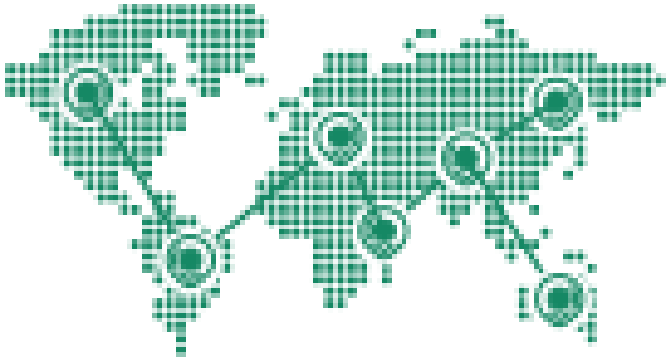
And it's not just the longer latencies that matter. It's also consistency. The Internet's latency fluctuations wreak havoc with delivering a predictable voice, video, and real-time application user experience across the Internet.

So, yes, the global Internet may suffice for backup or applications more tolerant of latency and predictability, but mission-critical or loss-sensitive applications require another solution – a privately managed middle mile. With a private middle mile, enterprises eliminate the delay and unpredictability of the Internet core. And, when done right, end-to-end performance can exceed MPLS throughput without compromising on cost, complexity, uptime, or insight.

There are two cloud alternatives to answering the middle mile question: the private cloud backbone or the Cato Global Private Backbone. In both cases, sites are equipped, ideally, with redundant SD-WAN devices in HA that establish encrypted tunnels across local Internet connections to an on-ramp to the nearest provider PoP. But beyond those basic details, offerings are very different.

Global Private Backbones

AWS and Azure offer private cloud backbones for connecting third-party SD-WAN devices. Connecting to private backbones involves a complex provisioning process. SD-WAN features may also be unavailable. Limited bandwidth, routing limits, increased operational burden, and higher costs are some of those problems. Deployment, monitoring, and securing traffic routing are further complicated by requiring third-party solutions for next-generation firewall (NGFW), threat prevention, SDP/ZTNA, and remote access.



Claims of network reach also need to be scrutinized. While cloud providers may claim many PoPs, only a fraction can act as SD-WAN on-ramps. At last check, for example, only 39 of Azure's PoPs out of some 65 PoPs supported Azure Virtual WAN.

Finally, there's the question of availability. Uptime SLAs offered by cloud providers run 99.95% or ~264 minutes of downtime per year. By contrast, traditional telco service availability typically runs at four nines, 99.99% uptime for ~52 minutes of downtime per year. The SLAs also only govern the backbone transit, not the access to the PoP. Should there be a brownout or blackout event, automatic failover to a secondary or tertiary PoPs is not part of the service and will depend on the SD-WAN configuration.

CATO NETWORKS

National Technology Hub

The Cato Global Private Backbone

As part of the Cato SASE Cloud, Cato edge SD-WAN devices, the Cato Sockets, automatically connect to the nearest Cato PoP into the Cato Global Private Backbone. The Cato backbone is a geographically distributed, SLA-backed network of 70+ PoPs, interconnected by multiple tier-1 carriers that commit to SLAs around long-haul latency, jitter, and packet loss. Cato backs its network with 99.999% uptime SLA (~5m of downtime per year).

As such, enterprises do not need to do any HA planning with Cato, which, as we alluded to above, would be required with private cloud backbones or when telcos deliver SD-WAN services. In those cases, enterprises need to ensure that the telco or provider designs in the necessary redundancy. Are there redundant appliances? What happens when there is a lockup, not just an outage, will the system failover properly? What about the underlying memory, storage, and server system underpinning what are often virtual appliances? Are they redundant? What happens if the PoP becomes inaccessible? The list goes on.

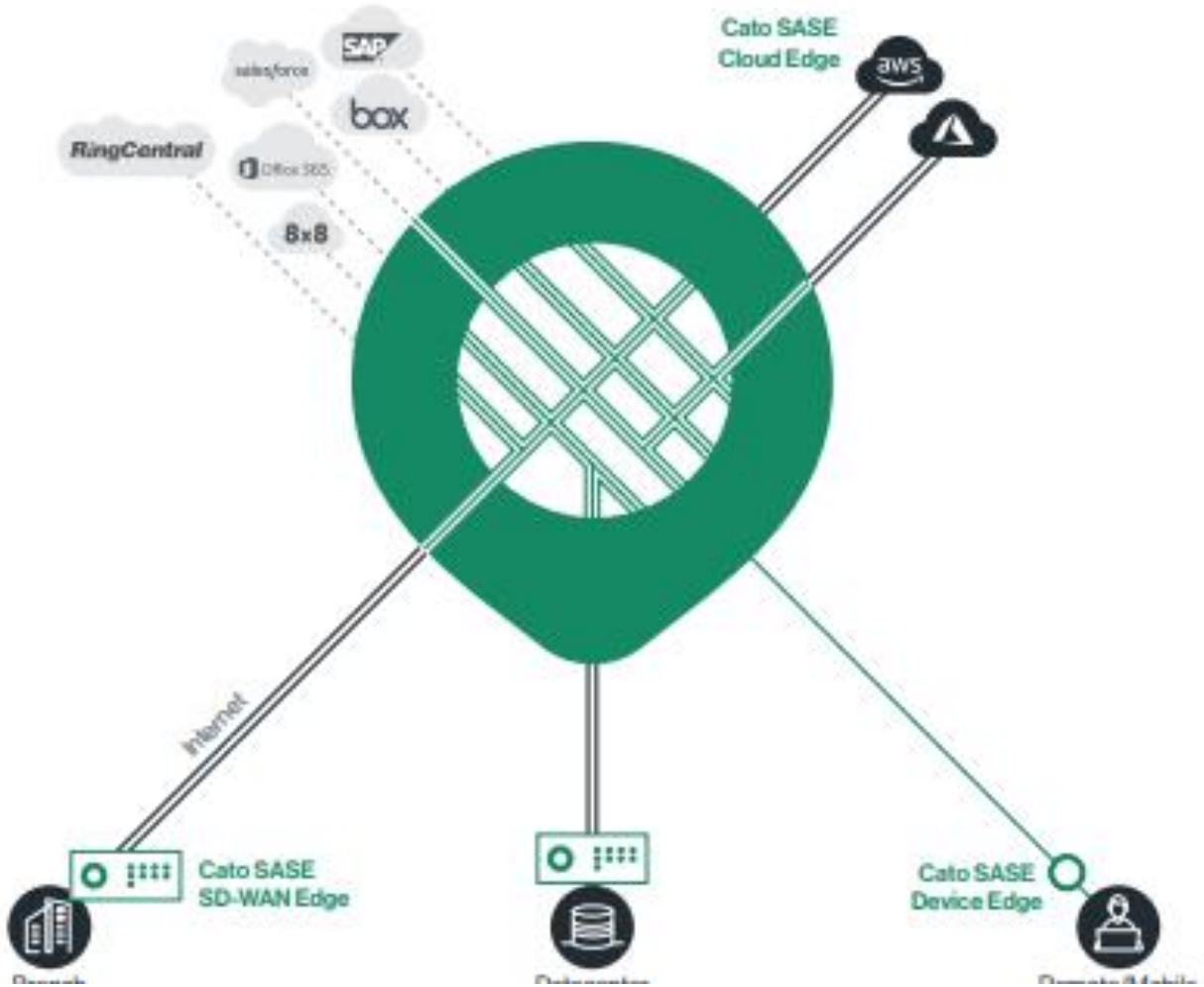
CATO NETWORKS

National Technology Hub

The Cato Global Private Backbone

By contrast, as a fully distributed, self-healing service, Cato includes many tiers of redundancies. Every Cato PoP consists of multiple compute nodes, and every compute node consists of multiple multicore servers. If the server core processing a flow fails, the flow will be handled by one of the other cores in the compute node. Should a compute node fail, other compute nodes in the Cato PoP assume the operation. Should the PoP become inaccessible, the connected users will automatically reconnect to the next best PoPs.

The Cato Global Private Backbone



CATO NETWORKS

National Technology Hub

The Cato Global Private Backbone spans more than 70 locations worldwide servicing over 140 countries.

CATO NETWORKS

National Technology Hub

The Cato Global Private Backbone

Cato delivers not only a highly reliable global backbone but also an optimized one. Cato accelerates WAN traffic by maximizing bandwidth for activities such as file downloads. Real-time path selection and packet loss correction means sensitive applications, like voice and video, get the predictable transport they require. And Cato optimizes cloud access by routing application-specific traffic to the PoP closest to the cloud destination.

Overall, Cato customers have seen 10x to 20x improved throughput when compared against running across MPLS or an all Internet connection, at a significantly lower cost than MPLS. And as part of Cato's SASE solution, the Cato global private backbone is easily activated and instantly deployed. (Cato also offers an optional last-mile management service; see the Appendix for further details.)

Telco Services: Same Legacy Problem with A New Look and Feel

And what about telco backbones as part of their managed SD-WAN services? With the convenience of a fully managed telco service also come the many limitations that led enterprises to leave the telcos in the first place.

While telco networks might have substantial reach within their operating regions, they lack global coverage. Once again, enterprises are left with themselves or the telcos establishing relationships with third parties to connect locations outside their operating area. And with those deployments comes a loss of control and visibility.

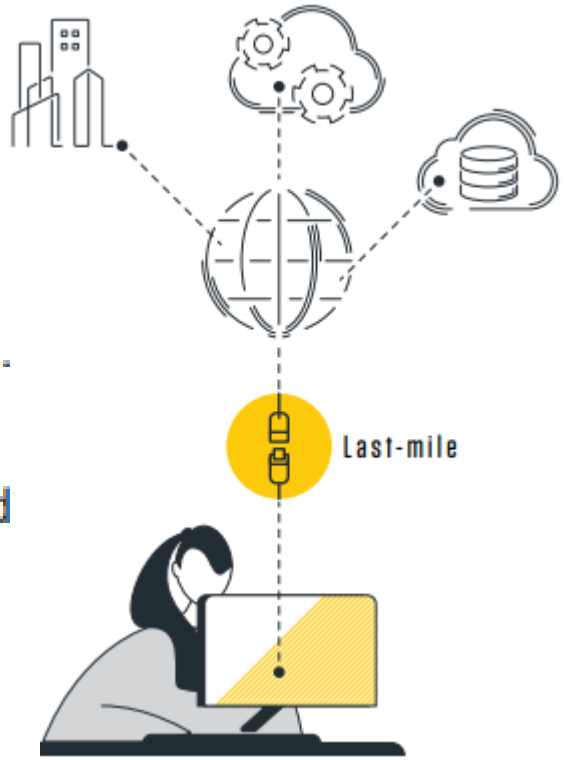
Part of the power of SD-WAN was empowering enterprises to change their WAN configuration themselves. With telco networks, those capabilities are typically limited to monitoring the network. Any changes must be implemented through the telco.

Telco Services: Same Legacy Problem with A New Look and Feel

Costs also remain high. Telcos use legacy, dedicated infrastructure for each customer. The cost of acquiring and integrating appliances is necessarily passed onto the customer at a premium.

The service itself is rigid. For one, organizations must have at least some of their offices on the telco's network, limiting their provider section. Enterprises also need to wait for the telco to deploy the last-mile and appliance stacks, unable to switch providers if necessary. What's more, product expertise sits in the technology supplier, not the telco. The telco is limited in its ability to accommodate support requests and lacks the roadmap control of the underlying technology platform.

SLAs: Sometimes Being Apart is Better Than Together



Building a global network on the backs of SLAs is an artifact of antiquated thinking. Yes, SLAs have a role, but better to rely on technology to ensure uptime than an SLA. By separating the underlay from the overlay and the last mile from the middle mile, enterprises can implement a strategy for building a global network that is reliable and optimized without the cost or lock-in of legacy MPLS services.

With separate underlay (last mile) and overlay (SD-WAN) providers, IT benefits from the market competition that has driven down the cost and increased the range of Internet last-mile access options. IT can now match the best SD-WAN vendor with the last-mile provider in each region. With one provider delivering both the underlay and overlay, enterprises lose that flexibility.

CATO NETWORKS

National Technology Hub

How the $^&# Am I Supposed to Manage so Many Internet Circuits?!

One of the significant challenges in shifting to an Internet-based network is managing the last-mile circuits. It's a big change from legacy services where the service provider managed last-mile circuits.

But while beneficial, tying the last mile to the service provider also restricts last-mile options. If a provider doesn't have a relationship with a local provider in-region or the last-mile provider's offering doesn't meet requirements, the enterprise is stuck.

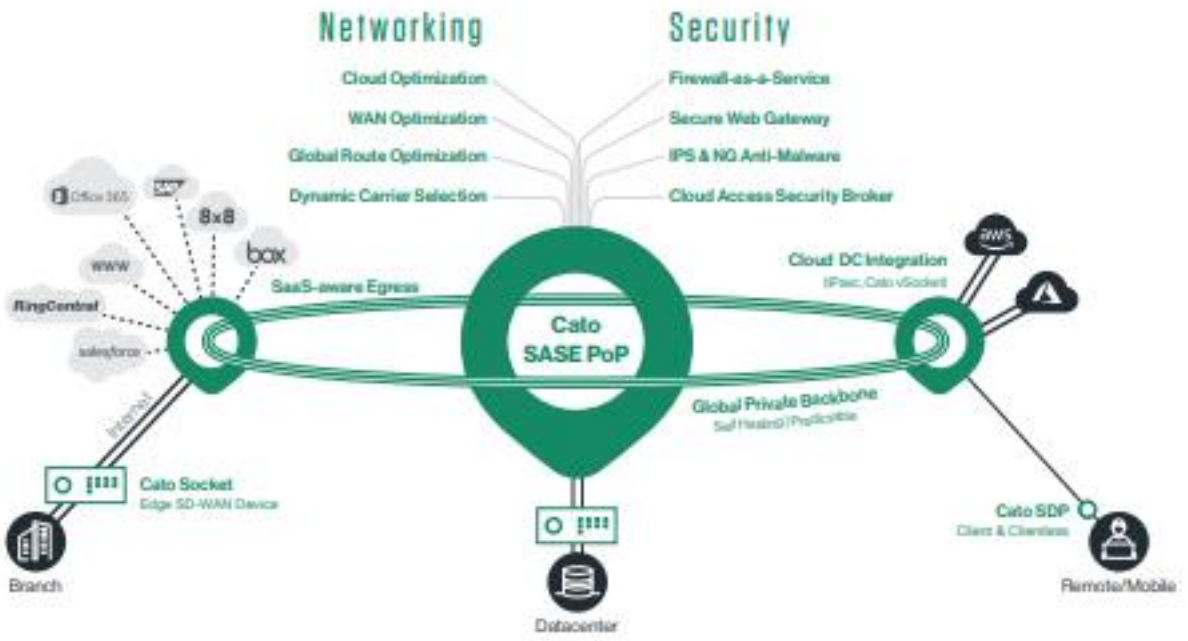
With Cato, enterprises retain flexibility and gain last-mile management. Cato offers Intelligent Last-Mile Management (ILMM) a-la-carte, which provides 24x7 monitoring of last-mile ISPs. We will detect link degradation or failures and work with the ISP until the issue is resolved. If needed, we can offer last-mile provisioning through an ISP aggregator. Unlike telcos and legacy service providers, enterprises maintain control and full self-service access to their network.

CATO NETWORKS

National Technology Hub

About Cato Networks

Cato is the world's first SASE platform, converging SD-WAN and network security into a global cloud-native service. Cato optimizes and secures application access for all users and locations. Using Cato SASE Cloud, customers easily migrate from MPLS to SD-WAN, improve connectivity to on-premises and cloud applications, enable secure branch Internet access everywhere, and seamlessly integrate cloud data centers and remote users into the network with a zero-trust architecture. With Cato, your network and business are ready for whatever's next.



CATO NETWORKS

National Technology Hub

Cato SASE. Ready for Whatever's Next

Cato SASE Cloud

- Global Private Backbone
- Edge SD-WAN
- Security as a Service
- Cloud Datacenter Integration
- Cloud Application Acceleration
- Secure Remote Access
- Unified Management Application

Use Cases

- MPLS migration to SD-WAN
- Optimized Global Connectivity
- Secure Branch Internet Access
- Cloud Acceleration and Control
- Remote Access Security and Optimization
- Flexible Management

SOC2 Approved

ISO 27001 Certified

GDPR Compliant

Gartner peerinsights. Customer first

CATO NETWORKS

National Technology Hub