www.pipelinepub.com
Volume 20, Issue 3

Cloud Outages and the Flattening of the Internet

By: Khalid Raza

Since the public internet was first introduced in 1995, two metrics have grown exponentially: The importance of internet ***reliability*** and the ***complexity*** of the internet. Of course, complexity works against reliability. As a result, we've been engaged in an arms race since 1995. Every time the internet became more complex, we had to work harder to keep the internet reliable.

Here is some context. In 1995, there were 16 million users on the internet. (Remember, the public internet grew out of the private Arpanet, so there was an existing set of users from day one). Today, we have more than 5.6 billion users. That's an annual growth rate of 23% per year for 28 years. Traffic has grown from 0.1 exabytes per month in 2000 to 403.3 exabytes per month today (annual 43% growth).

This exponential growth of users coincides with other complexity drivers, such as hybrid cloud, edge networks, IoT, and an explosion in remote workers to create a vastly more complex internet.

Internet reliability is falling. According to Uptime Institute's 2022 Outage Analysis, the number of outages is slowly growing due to increasing network complexity.

The reality is that we've outgrown the way the internet works. We need to relook at how we build enterprise networks.

Where Outages Come From

Recent network outages that caused loss of connectivity to cloud providers have highlighted the need to change how we implement peering. What made sense in a simpler era now creates an increasingly fragile internet.

Figure 1

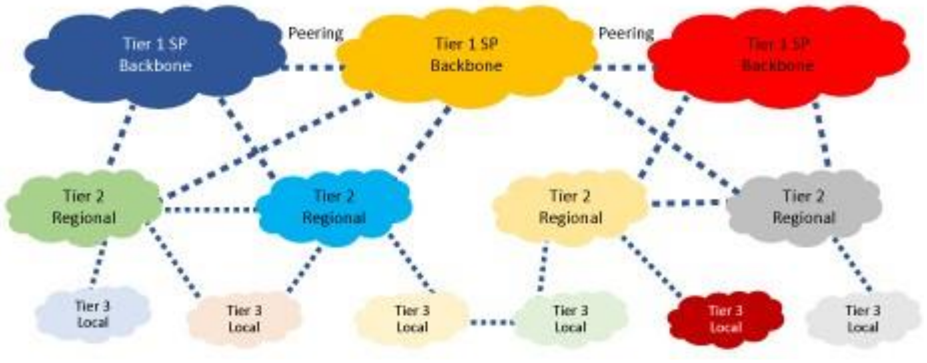


Figure 1: Tier 3 providers offer customers on-ramps through cable, fiber, and wireless line access.
click to enlarge

To understand why we need to change peering, let's explore how the internet routes traffic today. Most service providers connect with each other, and there is an established hierarchy:

Tier I Providers

These are usually the national or global carriers. They own extensive backbone and peer with prominent national or global carriers. These providers don't buy transit from any regional carriers but provide them with transit. These carriers maintain full internet routing tables that provide transits to Tier 2 or regional providers.

Examples of Tier 1 service providers in the US are AT&T, Verizon, Sprint, Century Link, etc.

Tier 2 Providers

These are regional providers. They peer with Tier 1 providers, buy transit from them, and provide transit to downstream providers. Essentially, they connect the hierarchy of local providers to the global provider by purchasing and selling transit. Examples of Tier 2 service providers include Hurricane Electric, Comcast, and Vodaphone.

Tier 3 Providers

They are local loop connections, or are what are called "Stub" networks. They don't sell any transit but buy transit from Tier 2 providers or sometimes directly from Tier 1 providers. Tier 3 providers offer customers on-ramps through cable, fiber, and wireless line access. Examples include Comcast, Deutsche Telekom, and Verizon Communications (see figure 1 above). Generally speaking, traffic flows up from Tier 3 to Tier 1, across Tier 1, and then down the tiers to the provider. But that's changing. Most user-facing network providers like Google, Facebook, and Amazon have increased their relationships with Tier 3 or local providers, and direct peering with regional transit providers. Cloud providers are also doing the same. This dramatically changes the dynamics of the Tier 1 provider. Tier 1 is being replaced by the cloud provider backbone. Why? Because large providers (like Facebook, Google, or cloud providers like AWS or Microsoft Azure) don't trust the current state of routing traffic across the digital wilderness of the public internet (see figure 2 below).

What is behind their distrust? We've recently seen significant outages in major cloud providers due to configuration and policy errors from Tier 1, 2, and 3 service providers. Network reachability failure occurs due to a few reasons.

- Router configuration errors. Someone in one of the tiers misconfigures their router, and traffic gets misrouted as a result.

- Link problems.

- Delays in network convergence. This is where the routing rules to get to a specific provider change in one of the tiers, and it takes a while for the change to propagate to the other tiers. During this delay, traffic to the providers fails to reach the provider.

Figure 2

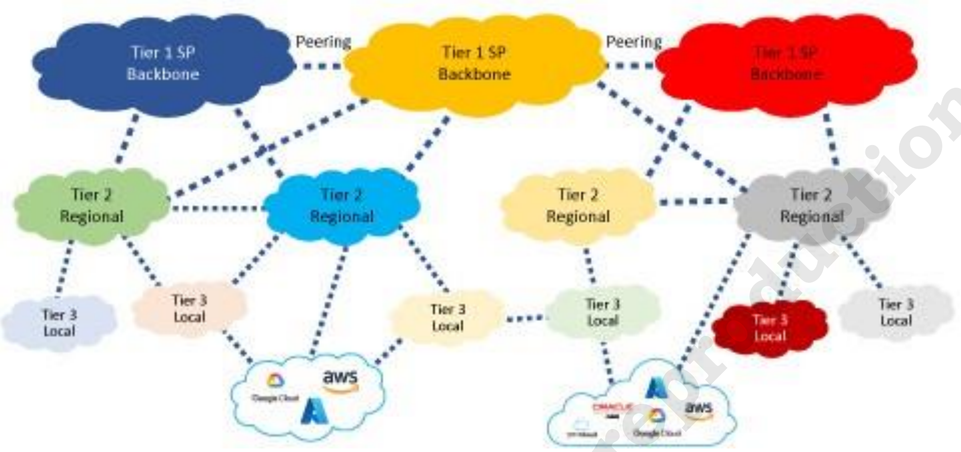
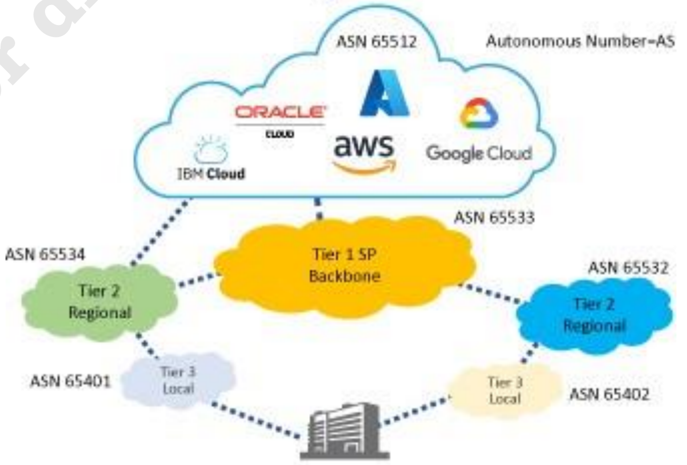


Figure 2: Tier 1 is being replaced by the cloud provider backbone.
click to enlarge

Larger providers understand these problems very well, especially with small carriers. Local issues sometimes cause global outages. For example, if your local ISP (Tier 3) has one of the aforementioned routing problems, your connection to the large provider (e.g., Facebook or AWS) will fail. Essentially, the ISP promises a destination or service that it can't deliver. They don't have end-to-end control or visibility. They have to trust that routing rules propagated from higher tiers are accurate, but as we've seen, that is not always true.

Figure 3



Fixing the Internet

We can't afford the current model's internet-based connectivity for mission-critical services. On the other hand, if we change a few things, which in my opinion and experience is critical for nextgeneration business class internet, we can establish direct control plane peering between the end customer and the cloud provider.

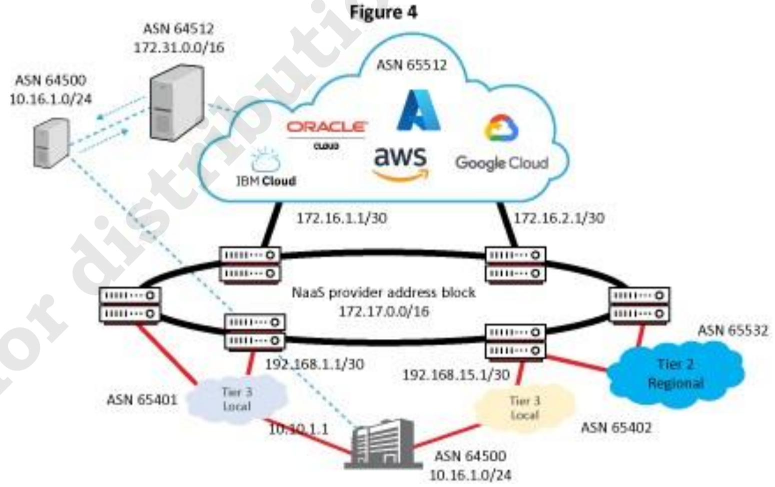
We start by introducing a Network-as-a-Service provider. This NaaS provider sits between the enterprise and the provider (Facebook, AWS, etc.). The NaaS provider has no transit information about customers and cloud provider routes. This means router errors within the NaaS are simply not possible. All POP routers are just forwarding stateless routes without routing, encryptions, VRF, etc. At the same time, they all understand how to route traffic to the cloud provider's next hops. The customer just needs to send traffic to the cloud provider network, while pointing it to the stateless core of the NaaS provider.

NaaS provider will only advertise a route to the next hop of the cloud provider it directly connects with. Now the underlay of the NaaS provider only sends the next hops it connects with, while the source and destination routes are only known to the customer and cloud provider. I learned this during my SDWAN days for business class internet. Unless we involve underlay through SDN principles, we will never provide service guarantees.

Referencing figure 4 below, the peering between the enterprise and the provider is at the off-path control layer. This peering will exchange routes and all the related security keys. For example, the cloud ASN 64512 will advertise 172.31.0.0/16 to the customer, and the customer will advertise the route of 10.16.1.0/24 to the cloud provider.

This is where the NaaS concepts become interesting. For the prefix 172.31.0.0/24, the next hop is the link between the NaaS provider and the customer. The customer builds an overlay to the closest POP of the NaaS provider, so the next hop for 172.31.0.0 would be 192.1681.1. The tunnel is built to the POP, and at the POP, the packet is routed to the exit of the cloud provider.

Figure 4



Traffic from the enterprise reaches the NaaS over the internet. The NaaS will provide details and visibility about how traffic traverses the NaaS network. The enterprise tells the NaaS to route traffic to the closest exit to the provider.

Now what we can do inside the NaaS provider is, instead of running hop-by-hop routing, we can use SDN tactics along with segment routing. Once the traffic enters the NaaS provider, a label stack is put in the header, and it routes the traffic to the peering interface of the NaaS and cloud provider. The traffic will use the prefix and adjacency labels until it reaches the exit interface. This helps eliminate transit issues even within the NaaS provider. Also, the NaaS provider only offers connectivity to the services it can directly connect to.

We can eliminate a lot of convergence, scale, and even security issues by deploying stateless networks within the NaaS and in the future with any transit providers. SDN came to solve networking issues, and it was about more than just network automation and configuration monitoring.

The practice of providers getting their users into their networks earlier and keeping them there for more of the journey is called “flattening the internet.”

Before we go into how the internet is flattening, we can flatten it further through state abstraction using SDN concepts.

In the past decade, we have seen the growth of FAANG (Facebook, Amazon, Apple, Netflix, Google), and now we should call them MAANA (Meta, Amazon, Apple, Netflix, Alphabet). These companies want to get traffic to their network as quickly as possible, and keep it in their networks. Similarly, the cloud providers have done the same.

Why? After all, large service providers want to get the traffic out of their network as quickly as possible to preserve capacity. Why would FAANG (and the larger cloud companies) be so interested in getting their users into their networks more quickly and keeping them there? The answer is that’s the only way they can improve reliability. There is a growing consensus that reliability isn’t likely in our digital wilderness of the public internet.