How Teridion Kumo-X Works

A Technical Overview of the Teridion Kumo-X

Teridion Engineering

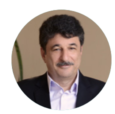
March 2017

Introduction

Sub-second application response time is an important but often unrealized goal for SaaS providers. Everyone knows that application responsiveness is directly linked to user productivity. However, inconsistent Internet performance makes response times of over 5 seconds typical, particularly for end users who are located far from the application's data center.

With Teridion Kumo-X, SaaS providers are now able to take back control of their application's performance to deliver the best possible user experience. This paper describes how Teridion Kumo-X works and how it solves the "Internet Backbone Problem" to deliver better performance and ROI.

> "Teridion Kumo-X accelerates the Thru platform's data transfers on global networks via intelligent routing over the inefficient segments of the Internet."



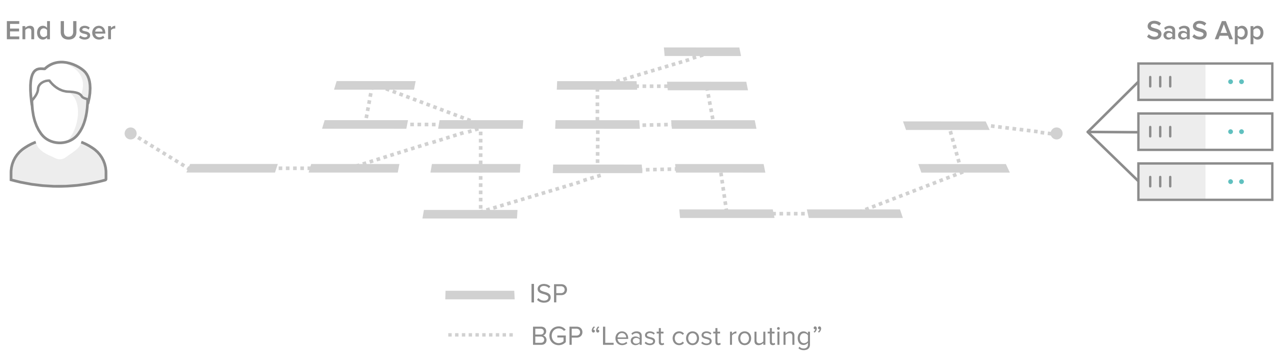
SERGEY ARUTIUNOV
CTO, *Thru*

The Internet Backbone Problem

What is the Internet Backbone?

Nearly all web content flows through the Internet Backbone. The Internet Backbone is made up of many large Network Service Providers that interconnect with each other. These large networks charge ISPs to transport data packets long distances.

Internet Backbone



How Least Cost Routing Hurts Performance

Within the Internet Backbone, there is no mechanism for SaaS providers to pay ISPs more to deliver traffic faster. Thus the only way for an ISP to maximize profitability is to minimize the cost of sending traffic. Least cost routing is the process of selecting the path traffic will take along the Internet Backbone based on the lowest cost, not on performance.

Least cost routing happens as a result of rules that are baked into Border Gateway Protocol (BGP), the routing protocol of the Internet Backbone. These BGP rules allow providers to prioritize traffic using cost-based weighting factors. This puts the customer experience for SaaS providers at the mercy of the network providers' cost-cutting routing tables.

How BGP Hurts Performance

BGP is what network providers use to route data from their own machines to others, and vice versa. When you visit a website, that data traverses networks all over the world, through machines belonging to many different organizations. In order to ensure that data transmissions eventually get to their intended locations, Internet Backbone routers keep a table of known, trusted routes.

The negative performance impact to SaaS providers of the Internet relying on BGP is that the protocol's rules for moving traffic between networks (called EBGP) dictate that traffic between two points will always take the same path regardless of network congestion. The combination of least cost routing and lack of congestion detection explain much of the performance problems with the Internet Backbone.

How TCP/IP Hurts Performance

Among routing protocols, BGP is unique in using Transmission Control Protocol (TCP) as its transport protocol. TCP provides the Internet with reliable, ordered and error-checked delivery of data between two systems. By design, TCP is optimized for accurate delivery rather than timely delivery. Thus TCP sessions can incur long delays (often for seconds) while waiting for out-of-order messages or retransmissions of lost messages. For this reason, real-time applications like VOIP opt to use different protocols.

For SaaS providers this means that the TCP data transfer algorithms are not efficient, particularly for larger files which require many round trips to transmit. TCP/IP requires each chunk of data to be acknowledged by the receiver before the sender sends the next batch of data. Since these data chunks are typically small, typically a thousand bytes, it means that transferring even 1 MB of data can require hundreds of separate trips through the Internet Backbone.

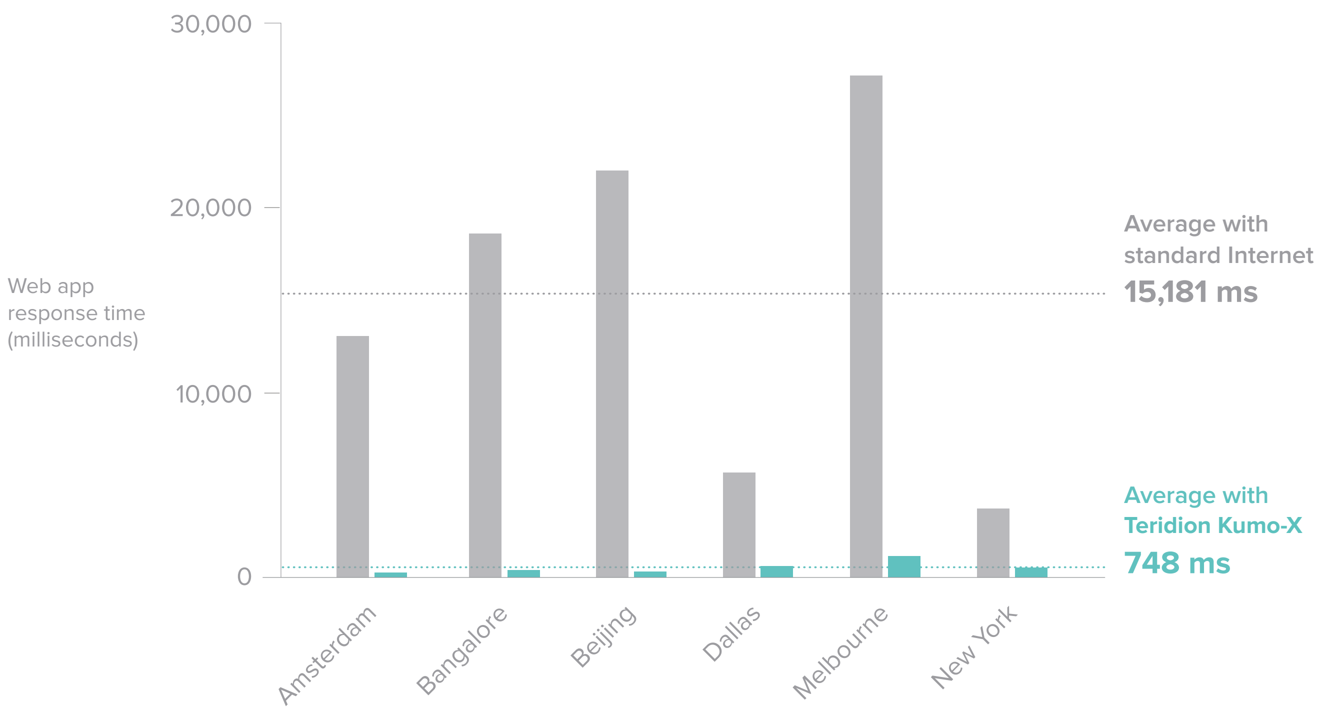
Teridion Kumo-X solves the Internet Backbone problem

Teridion Kumo-X opens up a fast lane through the Internet Backbone that gives the end-users of SaaS providers sub-second application performance around the globe. Teridion Kumo-X routes traffic around the congested networks in the Internet Backbone, eliminating the need for workarounds. This in-turn reduces the cost and complexity of application, storage and networking infrastructure.

How Teridion Kumo-X works

Teridion Kumo-X intelligently analyzes the Internet Backbone to find the fastest routes between any two endpoints, avoiding congestion and overcoming the performance problems caused by least cost routing. In the process it delivers better data transfer performance. Teridion Kumo-X is deployed across public clouds and delivered "as-a-Service," so there is no hardware or caching to configure.

Teridion Kumo-X is faster than the Standard Internet



Source: time to transfer 5MB from San Jose, CA

How Teridion Kumo-X complements CDNs

Content Delivery Networks (CDNs) are distributed servers that serve cached content to users based on their geographic location. CDNs are excellent for serving static content to users near the CDN Point of Presence (PoP), but do not handle dynamic data that must created "on the fly," nor can they accelerate uploads or bi-directional data transfers.

CDN caches help end users avoid having to request content from the origin server. Teridion Virtual Networks accelerate content requests from the origin server. Both approaches improve the end user experience.

Teridion excels at accelerating uploads and dynamic content, making Teridion Kumo-X a natural complement to CDNs within a SaaS application architecture. Teridion can complement a CDN in several ways:

- **Accelerate dynamic or uncacheable content**: Teridion Kumo-X does not require the replication or synchronization of data at the edge to PoPs. Instead, Teridion Kumo-X accelerates the data to the end-user directly from the origin server.
- **Meet stringent regulatory requirements**: Replicating content across a global network of CDN caches may conflict with financial regulations, EU data residency laws and Chinese Internet Content Provider licenses.
- **Support security requirements**: CDNs cache unencrypted static content in many locations, breaking the chain of custody for security conscious SaaS providers and opening up potential security vulnerabilities.

How Teridion Kumo-X complements SD-WANs

Software-Defined Wide Area Networks (SD-WANs) optimize internal point-to-point business communications between central and branch offices. Teridion Kumo-X is able to deliver private network quality of service over the public Internet at a fraction the price of traditional leased line solutions, like Multiprotocol Label Switching (MPLS dedicated line.) Teridion Kumo-X can add value to a traditional SD-WAN deployment in several ways:

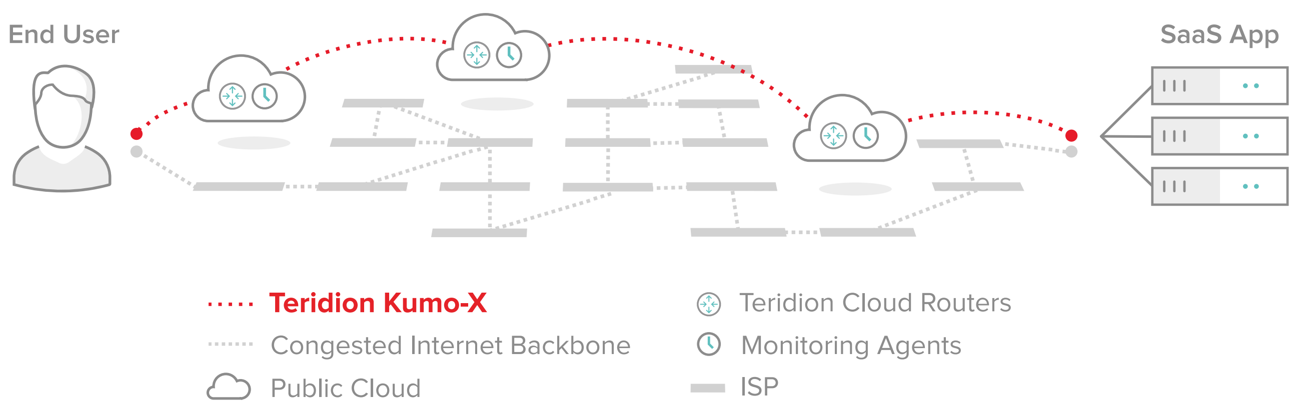
- **Remote Office Workers**: Teridion Kumo-X improves networking to remote workers in small regional offices where an SD-WAN may not be cost effective.
- **Mobile Workforce**: Teridion Kumo-X can improve network performance for corporate mobile users anywhere in the world.

Teridion Kumo-X Architecture

Teridion Kumo-X is comprised of three components deployed globally across hundreds of public cloud data centers:

- **Teridion Kumo-X Management System (TMS)**: Orchestrates virtual agents that are constantly calculating the fastest path between every route
- **Teridion Kumo-X Cloud Routers (TCR)**: Virtual router VMs that direct traffic along the fastest paths (calculated by TMS) by using the most efficient algorithms
- **Teridion Kumo-X Portal**: A network administrator UI for monitoring traffic and troubleshooting connectivity issues

Teridion Kumo-X is unique in leveraging public cloud infrastructure to provide optimal routing across the Internet Backbone. Other Internet overlay providers are trapped in expensive and inflexible private clouds. Teridion Kumo-X takes advantage of the scale of public cloud investment in both location and in peering relationships with local ISPs. Teridion Kumo-X also has the advantage of neutrality over public cloud vendors in being able to pick the optimal cloud providers for each region.



Teridion Management System

The Teridion Management System (TMS) is a centralized orchestration service that automatically operates Teridion Kumo-X. The TMS is responsible for finding optimized paths and providing management APIs for configuring virtual networks.

The TMS gets performance data from Teridion Measurement Agents (TMA) which are installed in each of the hundreds of public clouds that Teridion Kumo-X operates in. These agents perform real time measurements of Internet performance between other data centers and send that data to the TMS.

These Teridion orchestration components operate in the control plane and have no access to the data flowing through the Teridion Kumo-X. The TMA is written in Scala to enable the real-time parallel processing of large numbers of data events.

TMS is designed according to a number of cloud best practices, including:

- **Multi-cloud**: TMS can deploy virtual networks across over a dozen cloud providers and hundreds of cloud data centers.
- **Elastic scaling**: TMS orchestration automatically scales up and down TCR VMs based on real time demand for each region.
- **Resilient**: TMS can detect and recover from a number of failure scenarios, including DNS, network and data center outages.

Teridion Cloud Routers

A TCR is a virtual router that runs on public cloud Linux VMs. TCRs take advantage of public cloud peering relationships with local ISPs to provide a fast “on-ramp” to the Internet Backbone for end users anywhere in the world.

The TMS orchestrates deployment of sets of TCRs in the public cloud to create one or more fast lanes. TCRs are dedicated to a particular SaaS provider’s application traffic and do not cache or store network content. DNS CNAME redirection is used to map a particular origin server target to its fast lane.

Each TCR VM performs network monitoring to detect and recover from network interruptions. TCRs support multiple protocols (TCP, UDP, SIP, FTP, SFTP) and can use service chaining to add additional protocols and services. For example, TCRs can be configured to capture Teridion SessionFlow data which provides end-to-end visibility into how content moves across the Internet.

Teridion’s approach to high availability

Teridion Kumo-X was designed as a self-healing network that is resilient to a wide range of potential outages, including problems affecting entire data centers or regions.

- **DNS provider failure**: Every Teridion Kumo-X has multiple Domain Name Server (DNS) providers. If one provider fails or slows significantly, Teridion Kumo-X will switch to a secondary DNS provider.
- **Network failure**: The TMS automatically reroutes traffic in response to any network congestion, including data center or network failures. This makes Teridion Kumo-X more reliable than dedicated MPLS dedicated line lines for large outages.
- **TCR edge failure**: TCRs are constantly monitored and unresponsive TCRs are automatically removed from Teridion Kumo-X.

- **TCR core failure**: The TMS automatically sets edge TCRs to route directly to the origin server if core TCRs become unavailable.
- **TCR failure**: TMS failures will not affect Teridion Kumo-X, however automated routing adjustments to avoid congestion will stop until TMS is back online. Every element of TMS is replicated for high availability and mirrored for disaster recovery.
- **TCR Portal failure**: Teridion Kumo-X's network performance is unaffected by the availability of the Teridion Kumo-X Portal.

Teridion Kumo-X Portal

The Teridion Kumo-X Portal enables network administrators to view and analyze traffic flowing through the Teridion Virtual Network. There are two main tabs:

- **Dashboard**: Shows performance and status of your traffic flows
- **Analysis**: Provides in-depth analysis and comparisons of your traffic flows.

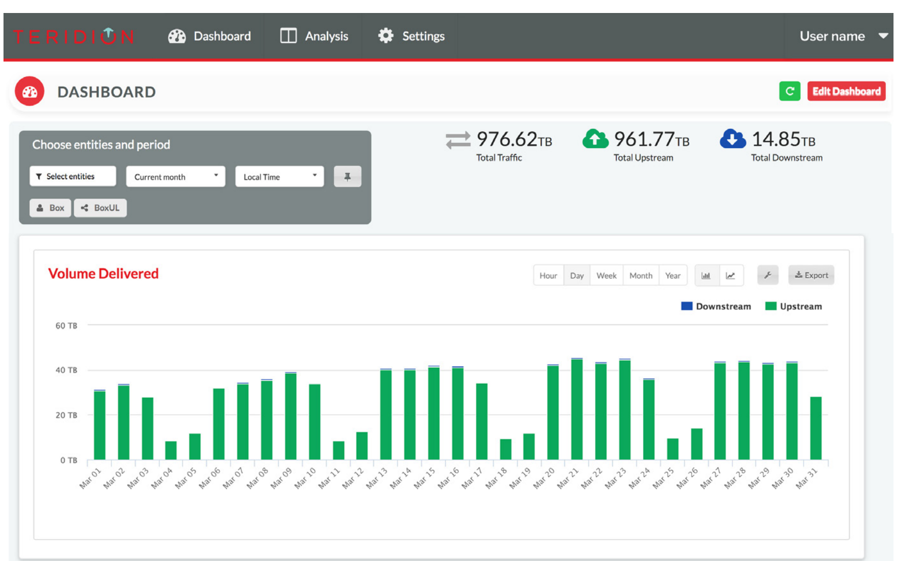
Dashboard Tab

The Teridion Kumo-X Portal Dashboard tab is customizable and includes programmable widgets that display upload and download traffic by time period, region and by end user domain. For each widget the dashboard can also show the performance improvement between the Teridion Kumo-X and the standard Internet.

To validate performance improvements, Teridion uses a third party web monitoring service to measure performance, latency and throughput for both the standard Internet and the Teridion Kumo-X This service compares end user performance from ISPs around the globe and then presents the results through the Teridion Kumo-X portal.

Analysis Tab

The Analysis tab shows traffic delivery and performance across time periods. For example, this can be used to compare traffic volumes delivered over several months.



Teridion SessionFlow Data

SessionFlow data is newly available in Teridion Version 2.0. This structured data output provides network administrators with end-to-end, real-time visibility into their Internet traffic.

SessionFlow provides a fast way to pinpoint network connectivity and throughput issues.

With SessionFlow data, network administrators can diagnose and troubleshoot traffic in realtime. SessionFlow data includes latency and throughput at each TCR location. The TCR data is sent asynchronously to a cloud storage location chosen by the SaaS provider (for example Amazon's S3 service).

The following is sample SessionFlow output:


```
"session-info": [
        {
            "domain-name": "dlnyc.teridion.com",
            "end-time": "2017-02-02T09:37:22Z.307ms",
            "hop-location": "IN",
            "hop-order": 1,
            "in": {
                "incoming-bytes": 205815,
                "incoming-throughput": 72435,
                "outgoing-bytes": 22646470,
                "outgoing-throughput": 7970250,
                "read-idle-time-ms": 22789,
                "retransmits": 0,
                "rtt": 60,
                "termination-type": "FIN Sent",
            },
            "ip": "5.102.254.127",
            "next-hop": "3e7bc2-digitalocean-ams2-roni2-n.teridion.work",
```

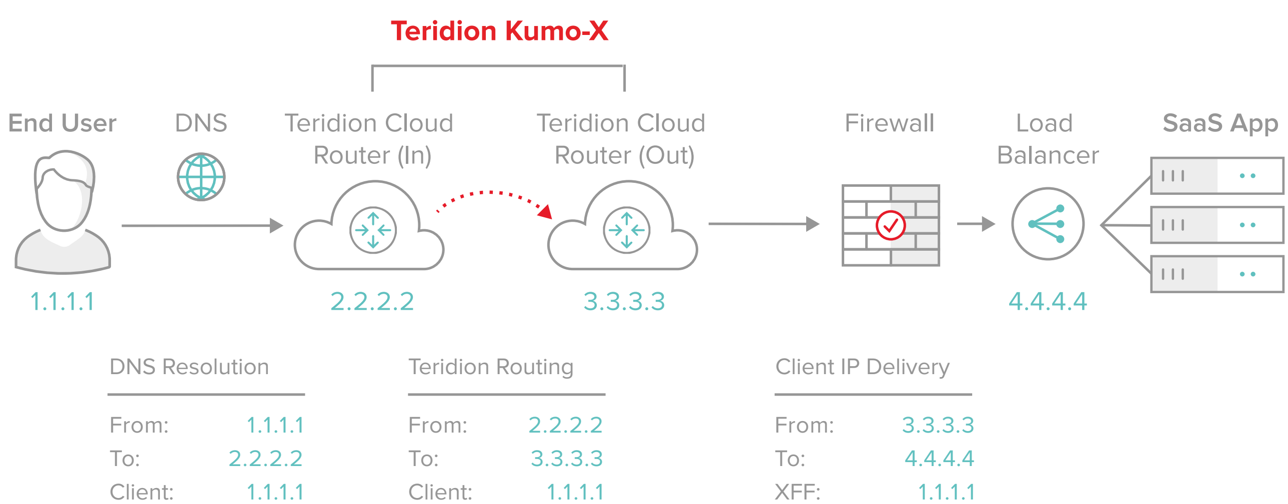
Teridion Kumo X Reference Architecture

Teridion Kumo-X carries content along a virtual network of TCRs, so Teridion Kumo-X looks to the rest of the network like any other network proxy - such as a load balancer, SSL offload or CDN cache. The end user connects to a TCR, which then routes that traffic along the virtual network to the origin server. Because TCRs act as network proxies, the connection request that reaches the origin server will have the IP address of a local TCR, referred to as a "TCR Out".

Teridion Kumo-X uses [Proxy Protocol] to give the load balancer at the origin connection information from the end user requesting a connection to the origin server. The load balancer then adds connection information to the `X-Forwarded-For` headers of each HTTP request such as the source IP address, destination IP address, and port numbers.

The following is an example of the Proxy Protocol format for IPv4. This format gives client IP, proxy IP, client port and proxy port, each separated by a single space):

```
PROXY TCP4 198.51.100.22
203.0.113.7 35646 80\r\n
```



Next steps

Interested in a FREE Trial?

Set up your app today, or call 1-844-TERIDION

Questions:

Contact Us

Case Studies:

Learn how SaaS providers are succeeding with Teridion

Datasheets:

Learn more about Teridion Kumo-X

Whitepaper:

"Internet Backbone Problem"