Cloud Connectivity Management Made Simple: Your Cloud Services Are As Good As Your Network

Bruno Giguère, Advisor – CTO Office, EXFO

INTRODUCTION

Cloud services are the Holy Grail for service providers. As enterprise customers look to reduce IT expenses, cloud services are looking more promising than ever in terms of achieving these savings. Although multiple stakeholders are increasingly involved in the delivery of cloud services, none have as much impact as cloud carriers.

But first, let's start with a simple question: "What exactly is cloud computing?" According to the definition recently published by the National Institute of Standards and Technology (NIST), "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." The document further elaborates on its definition by providing a list of essential characteristics, in addition to service and deployment models. The essential characteristics are: on-demand self service, broad network access, resource pooling, rapid elasticity and measured service. A definition for each characteristic is provided in the following **NIST publication**. By applying a network view of these characteristics, it becomes immediately clear that a static and classic network is not able to deliver on these promises. An intelligent and dynamic network is mandatory to delivering on these essential characteristics.

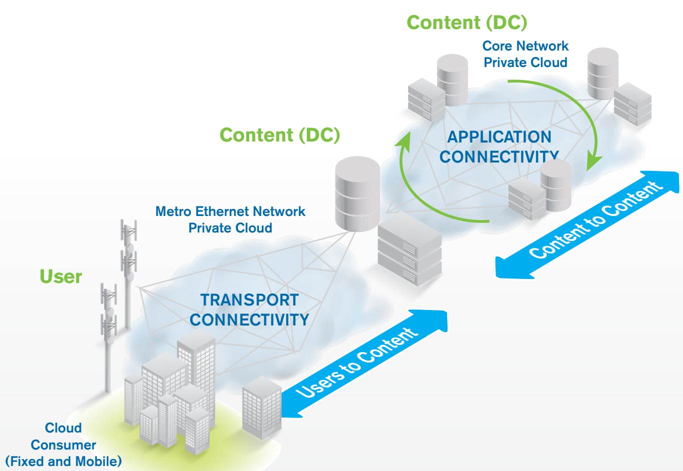

Figure 1. Cloud connectivity.

Figure 1 shows the data flow for delivering cloud computing services and the two types of connectivity required, as follows:

1. Cloud transport connectivity
2. Cloud application connectivity

Cloud transport connectivity can be described as the connectivity between cloud consumers and cloud content. This connectivity makes delivery of cloud computing services to the cloud consumer possible. In order to deliver this type of connectivity, cloud carriers need high-performance services with multiple classes of service and high availability.

Cloud application connectivity can be described as the connectivity between cloud content. This connectivity provides the ability to move data between the computing resources via different data centers. For cloud application connectivity, services must be high performance and be available at all times.

This application note discusses the technologies being used to deliver the different type of connectivity, and the management steps (e.g., activation, performance monitoring and troubleshooting) required to successfully deploy cloud computing services.

GOT CONNECTIVITY?

To deliver cloud computing services, cloud providers must possess the computing infrastructure necessary to process and store very large amounts of data, and rely on cloud carriers to connect all the pieces together. This section covers the different types of connectivity used to deliver cloud computing services in more detail, and also covers the technology used to achieve this connectivity.

Cloud Transport Connectivity

The original connectivity strategy for cloud computing was the Internet. What better a solution than the most ubiquitous WAN technology to access data centers? Although the Internet is great for consumer applications, it lacks the attributes mandatory for enterprises: security, network performance, data governance and regulatory compliance [**MEF CS**].

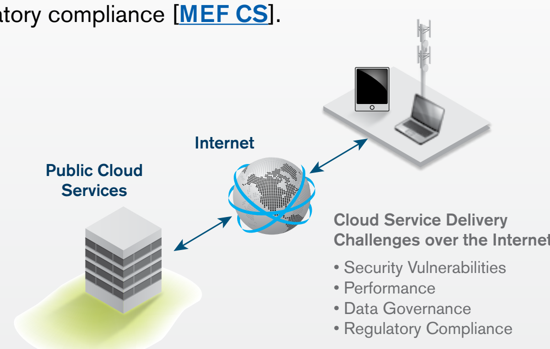

Figure 2. Current challenges with cloud service delivery. (Source: Metro Ethernet Forum)

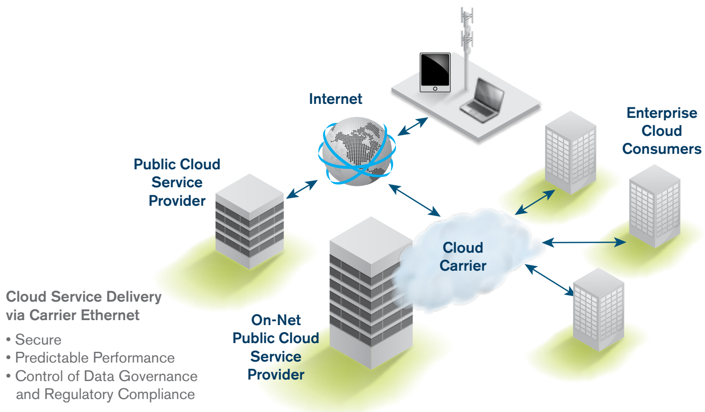

Figure 3. Ethernet services to public/private cloud service delivery. (Source: Metro Ethernet Forum)

One subject of particular interest involves the infrastructure inside the cloud carrier's network. Unfortunately, cloud carriers cannot always deliver services directly to cloud consumers. Not having a footprint in all markets will force carriers to rely on partners in order to connect remote cloud consumers. This creates a new breed of services referred to as cloud connectivity backhaul services (CCBS). The main issue with the use of this type of product is the expectation that it is the cloud carriers who must deliver to contractual SLAs, regardless of whether or not they are using CCBS to connect cloud consumers.

As with other services based on Carrier Ethernet 2.0, cloud transport connectivity needs to be high performance and use multiple classes of service with high availability. The table below provides an overview of the different service attributes that relate to cloud transport connectivity services.

Service	Priority	Committed Information Rate (CIR)	Excess Information Rate (EIR)	Frame Delay	Frame Delay Variation	Frame Loss Ratio	Availability
VoIP calls	0	10 mbit/s	0	5 ms	< 1 ms	0.1%	≥99.99%
Telepresence	1	50 mbit/s	0	25 ms	< 10 ms	0.1%	N/A
Mission Critical data	2	25 mbit/s	0	5 ms	< 1 ms	0.01%	≥99.995%
Streamed Live content	3	40 mbit/s	0	5 ms	< 1 ms	0.01%	≥99.99%
Non real-time content	4	15 mbit/s	500 mbit/s	25 ms	10 ms	1%	≥99%

Table 1. Cloud transport connectivity service attributes and examples of related SLAs.

As with other types of backhaul services, cloud connectivity backhaul services have their own service attribute requirements. In this case, the SLA value for the service attributes for CCBS will need to be more stringent than the end-to-end cloud transport connectivity.

One last characteristics of cloud transport connectivity is interactive provisioning of the bandwidth. Because cloud computing is becoming dynamic in nature and the services can be requested on-demand, the cloud transport connectivity needs to be as flexible as the services running over it. Cloud carriers have created a new breed of services that can be configured dynamically. By using web portals, cloud consumers can modify their bandwidth requirement and apply the new configurations very rapidly. The goal is to apply the new service configurations in seconds. To achieve this, cloud carriers will need to provision enough bandwidth in advance to double or triple the cloud transport connectivity without any truck rolls. This capability needs to be engineered and tested upfront, which can add additional burden to cloud carriers. We will address these issues in the cloud connectivity management section.

Now that we have tackled cloud transport connectivity, let's have a look at cloud application connectivity.

Cloud Application Connectivity

The other type of connectivity applicable to cloud computing is cloud application connectivity. As previously mentioned in this application note, cloud application connectivity provides the infrastructure to move large amounts of data between computing resources via the different data centers. For cloud application connectivity, services must be of high performance and high availability. The reason that cloud application connectivity has such requirements resides in the necessity of having all services available at all times, and ensuring that the response time for the applications running in the cloud is similar to that which would apply if the applications were residing in the enterprise data center.

An example of this would be the migration of a virtual machine (VM) from one server to another. When the system administrator initiates migration of the VM, transfer of data begins. This migration translates into a transfer of 1 GB to 10 GB of data, which must be executed in less than two minutes. This requires a transmission control protocol (TCP) connection capable of completing the transfer with a throughput of up to 3 Gbit/s. When viewed from a cloud application connectivity perspective, this means that the infrastructure is capable of achieving virtually zero packet loss with a very small round-trip frame delay (this value is vendor-specific, and is usually in the 10 ms range).

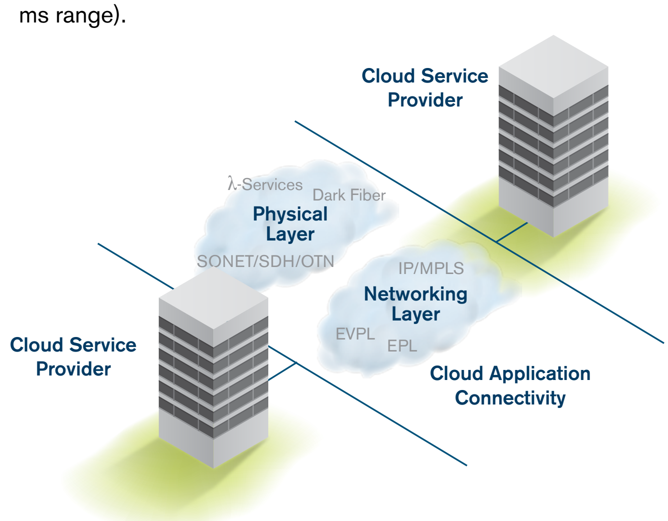

Figure 4. Technologies used to deliver cloud application services.

Historically, the technology of choice for cloud application connectivity was dark fiber, wavelength services or SONET/SDH/OTN. As is true for other services, a migration toward a packet-based architecture is also occurring in cloud application connectivity. Because this connectivity also needs to be dynamically configurable, Ethernet Private Line (EPL) and Ethernet Virtual Private Line (EVPL) and IP/MPLS services are now being deployed to deliver this connectivity.

CLOUD CONNECTIVITY MANAGEMENT: THE BASIS FOR DELIVERING QUALITY CLOUD COMPUTING SERVICES

Cloud connectivity management comprises the steps used to deliver cloud computing services. From cloud connectivity service activation to monitoring and troubleshooting service performance, cloud carriers need to ensure that each step is performed so that cloud consumers are delivered an optimal quality of experience.

Cloud Transport Connectivity Service Activation

Cloud transport connectivity services are dynamic in nature and will evolve over time. Cloud carriers need to ensure that the infrastructure used to deliver these services is correctly evaluated at service turn-up. The elasticity of bandwidth, although a very promising concept, cannot deliver more bandwidth than the underlying network infrastructure. For example, a 100 Mbit/s service on a 100 Mbit/s layer 1 connectivity will not be “dynamically” reconfigured to 1 Gbit/s. If the cloud carrier wants to provide “elastic bandwidth” functionality, it will have to be enabled from the outset. In terms of cloud transport connectivity, this means that testing for the original service configuration will not ensure that QoE will be optimal once the cloud consumer reconfigures its connectivity to a higher rate.

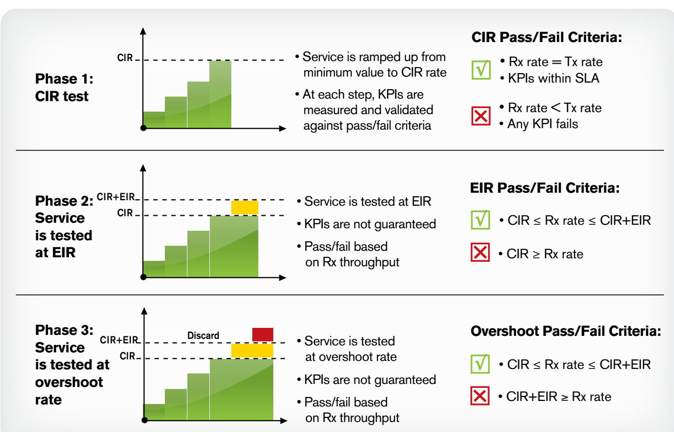

Figure 5. Y.1564 Phase 1: service configuration test.

Fortunately, there are test methodologies that can be leveraged to deliver these connectivity services. In section 2, Carrier Ethernet 2.0 was presented as the technology of choice for cloud transport connectivity. With its multiple class of service and manageability, CE 2.0 has all of the attributes for such connectivity. In order to create methods and procedures to activate these services, cloud carriers are leveraging a recommendation from the International Telecommunication Union (ITU) called Ethernet service activation test methodology, numbered Y.1564. This recommendation provides a dual-phase methodology that confirms the configuration of the different class of services and validates its performance. Figure 5 provides a visual description of the service configuration test phase of the methodology.

The first phase is referred to as the service configuration test. For each service, a ramp test is used to gradually reach and exceed the committed information rate (CIR). During this part of the test, all key performance indicators (KPIs) are measured against a threshold. These KPIs are frame delay, frame delay variation and frame loss ratio. Once the CIR is confirmed, the excess information rate (EIR) is tested. Because the KPIs are not usually guaranteed between CIR and CIR + EIR, there is no pass/fail criterion for this part of the test apart from the confirmation of EIR.

The last step for the service configuration test is the policing test. Test frames are sent at a rate that is greater than CIR + EIR, and the received information rate is measured to confirm that the services do not have more bandwidth available to them than CIR + EIR.

The second phase of the Ethernet service activation test methodology is the service performance test. This phase of the test was created to validate that all services are capable of being delivered to the agreed service level agreement (SLA).

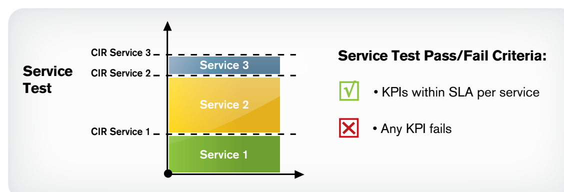

Figure 6. Y.1564 Phase 2: service performance test.

The methodology for this phase of Y.1564 is as follows: all services are generated simultaneously at their configured CIR. While these services are generated, all KPIs are monitored simultaneously and in a bidirectional manner for all services. These KPIs are then compared to the pass/fail criteria (called service activation criteria in Y.1564). If all KPIs pass, this phase of the test can be considered to be complete.

Although KPIs are great for characterizing and defining SLAs, they only cover network performance up to the IP layer of a network. These parameters provide the cloud carrier, cloud provider and cloud consumer with a measure of whether the network is capable of transporting frames. However, they will still not know what level of performance they should expect from their mission-critical applications. So, how can the cloud carrier ensure that the end user’s most important application is able to make use of the full bandwidth of the newly configured cloud transport connectivity services? The only way the cloud carrier can demonstrate that its services are able to perform as expected is by performing a TCP throughput test. By sending stateful TCP traffic to a remote test instrument, a cloud carrier can demonstrate that it is possible to fill the complete bandwidth with TCP-based test traffic and validate the service. Because there are a lot of issues related to the performance of the application running over TCP, it is recommended to test with Y.1564 to obtain a measurement of the service frame delay, and to then perform a TCP throughput test demonstrating the required window size for applications running over the cloud transport connectivity services. More information on **Understanding Carrier Ethernet Throughput [MEF TCP]** is available on the Metro Ethernet website.

Cloud Application Connectivity Service Activation

As discussed in section 2.2, the technologies used to deliver cloud application connectivity vary according to the requirement of the cloud provider. To better describe the methodologies used to activate each type of technology, let's group them in two categories: physical layer (dark fiber, wavelength services and SONET/SDH/OTN) and networking layer (EPL/EVPL and IP/MPLS).

Physical layer service uses the classic service activation test methodologies to demonstrate the capability of the link: bit-error-rate test (BERT). Complementary to BERT, other methodologies can be leveraged to provide additional information during service activation. Because cloud application connectivity services are delivered over fiber, optical time-domain reflectometers (OTDRs), and chromatic dispersion (CD) and polarization mode dispersion (PMD) can provide great insight into the quality of the fiber. These measurements can be used later on in the lifecycle of the service should any issues with the quality of the service arise.

In the case of wavelength services, the use of an optical spectrum analyzer (OSA) can be useful to ensure that the wavelength of the optical transmitter is centered properly. The optical signal-to-noise ratio (OSNR) is also an important measurement for wavelength services.

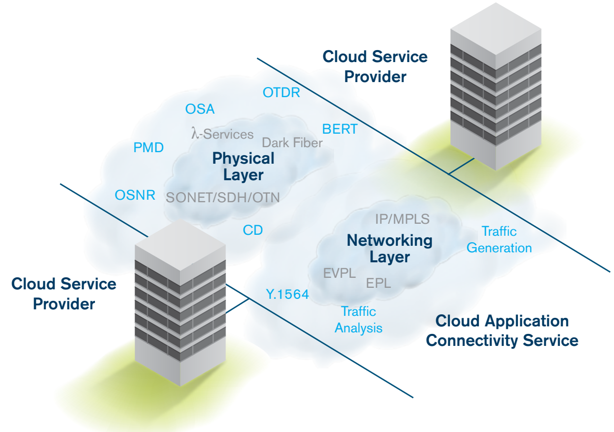

Figure 7. Cloud aplication connectivity – service activation test methodologies.

Service activation in a packet-based environment such as the networking layer requires that the test information be presented as frames or packets. Having the capability to measure the frame loss ratio of such a service with its frame delay and frame delay variation is quite important, especially in a cloud computing environment. Because the cloud application connectivity services are point-to-point and have very high bandwidth with no predefined class of service (from a cloud carrier perspective), the use of a test methodology such as Y.1564 could be recommended, but is not mandatory.

The notion of availability was discussed in section two, although not within the context of service activation. Because availability is a long-term measurement (generally over a monthly time period), its relevance in a service activation scenario (service activation test times for cloud connectivity services are in hours) is very limited. If we translate 99.99% availability over a one-week period, we would be looking at a downtime period of 6.05 seconds. To bring this value into context in order to declare a period of time as unavailable, the errors need to be present for 10 consecutive seconds.

Cloud Transport/Application Connectivity Performance Monitoring

Once the cloud connectivity services are in service, network performance monitoring is mandatory to change the operational behavior of a company from that of reactive to proactive. By monitoring the performance of services, cloud carriers have access to performance information that can be used for trending and reporting. By providing access to the service KPIs through a web portal, cloud carriers provide a network view to the cloud consumer. Because cloud consumers are just starting their migration to cloud computing, having a view of the KPIs provides them with the assurance that the network delivers as promised. By having a network view, cloud consumers can become interactive with their cloud carrier and cloud provider, and control the amount of bandwidth required to achieve a great cloud computing experience.

On the cloud carrier side, creating advanced thresholds enables cloud carriers to become proactive in the detection of possible issues, and allows them to deliver these services according to the contractual SLAs. Because cloud connectivity services require high availability, monitoring the network provides assurance that SLAs are being met.

Additional information about cloud connectivity performance monitoring is available in EXFO's "**Testing the Cloud**" white paper.

Cloud Transport/Application Connectivity Troubleshooting

In the event that a problem were to occur in the network, the cloud carrier would need to react rapidly because SLAs are very stringent in cloud computing. Having the capability to remotely diagnose the issue gives the cloud carrier an edge, because the penalties related to SLAs (and the possible loss of cloud consumers) can be very high and bring unwanted publicity to the outage. Fortunately, the same tools used to activate services can also be used to troubleshoot these issues.

By having access to the historical data of the cloud connectivity service (either from the birth certificate or performance monitoring trending), the cloud carrier will have additional data points that can be used to locate the problem and reduce the time needed to complete the repair.

CONCLUSION

This application note provides an overview of the current network technologies used to deliver cloud transport and application connectivity services. Because cloud carriers are leveraging their converged network architectures to deliver cloud connectivity services, the Metro Ethernet Forum Carrier Ethernet 2.0 framework will become the technology of choice for delivering these services. With its multiple classes of service, interconnectivity and manageability characteristics, MEF CE 2.0 provides the basis for the quality cloud connectivity service required for cloud computing.

Cloud connectivity services are based on advanced network technologies that are predictable and manageable. The tools used by cloud carriers to operate and manage other services can be leveraged to deliver quality cloud connectivity services, the only difference lies in the KPIs; which are more stringent than those for the SLAs of other services.

The cloud connectivity management step presented in this application note should be leveraged by the cloud carrier. These steps ensure that cloud connectivity services are activated properly and monitored throughout the lifespan of the service. This enables cloud carriers to deliver the quality services mandatory for use of cloud computing services by cloud consumers.

REFERENCES

[EXFO] M. Sylor. "Testing the Cloud," EXFO, November 2012

[MEF CS] "Carrier Ethernet for Delivery of Private Cloud Services," Metro Ethernet Forum, February 2012.

[MEF TCP] "Understanding Carrier Ethernet Throughput," Metro Ethernet Forum, July 2010.

[NIST] "The NIST Definition of Cloud Computing," Special Publication 800-145, National Institute of Standards and Technology U.S. Department of Commerce, Peter Mell and Timothy Grance, September 2011.

GLOSSARY AND ACRONYMS

Term	Definition
Cloud User	*A person or organization that uses and benefits from the cloud.*
Cloud Consumer	*An organization or person who buys services from the cloud provider.*
Cloud Computing	*Cloud computing is the practice of using a network of remote servers (usually hosted on the Internet) to store, manage and process data, and to deliver computing resources (hardware and software) as a service.*
Cloud Provider	*An organization or person who provides a cloud service.*
Cloud Carrier	*A communications service provider who provides connectivity and transport between users and the cloud, or within the cloud.*
Cloud Broker	*An organization that helps cloud consumers find providers and manages the relationship between the consumer and provider.*
Cloud Auditor	*An independent organization that assesses the operations, performance or security of cloud providers.*
NIST	*National Institute of Standards and Technology*
CCBS	*Cloud connectivity backhaul services*

EXFO Headquarters > Tel.: +1 418 683-0211 | Toll-free: +1 800 663-3936 (USA and Canada) | Fax: +1 418 683-2170 | info@EXFO.com | **www.EXFO.com**

EXFO serves over 2000 customers in more than 100 countries. To find your local office contact details, please go to **www.EXFO.com/contact**.

APNOTE278.1AN

Printed in Canada 12/12