ITU-D Regional Development Forums 2010 on NGN and Broadband for the Arab Region
"NGN and Broadband, Opportunities and Challenges"

Traffic Trends and QoS in NGN

Cairo, Egypt, 13 to 15 December 2010

Oscar González Soto
ITU Consultant Expert
Spain
oscar.gonzalez-soto@ties.itu.int

Agenda

- Services traffic trends on NGN
- Traffic modeling for NGN services
- QoS for NGN service types

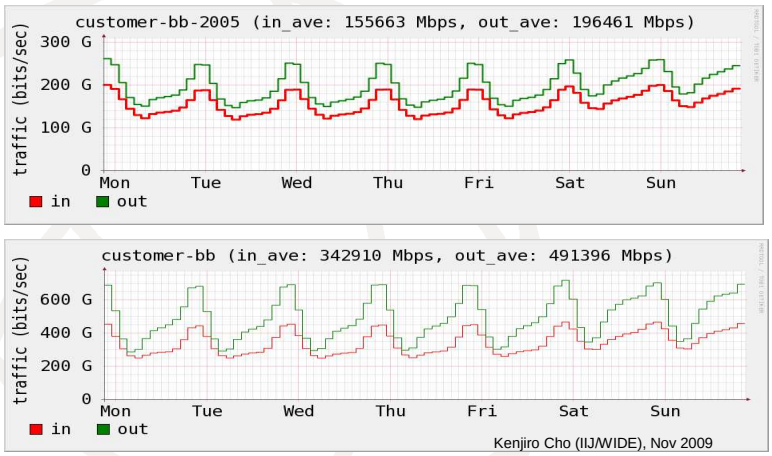
Traffic related questions

- Is traffic important in NGN and IP flows?
- What units to consider for dimensioning and engineering?
- Which traffic activities are needed in operation?
- Which units to consider for interconnection and SLA?
- Is QoS well defined and what parameters?
- Others……?

NGN Service traffic demand: Recent trends

- Major increase of data traffic both in fixed and mobile networks (mainly video, browsing and social networks)
- More heterogeneous behavior (up to 20:1 in volume and 10:1 in signaling for new terminals) and in session composition due to the multiplicity of terminals and applications
- Different proportion of busy hour traffic to the overall daily traffic as compared with traditional (> 25%)
- Several Origin/Destination patterns/matrix and flow modeling
- Need for a continuous process in traffic measurement, projection and dimensioning

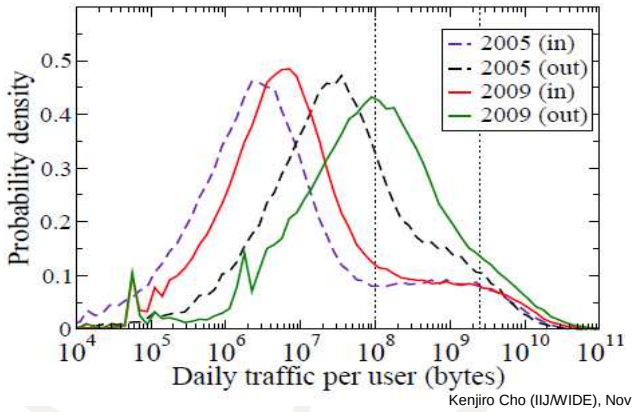
NGN Service traffic demand: Recent trends



Kenjiro Cho (IIJ/WIDE), Nov 2009

• Evolution of traffic daily profile in Japan 2005-2009: (daily periodicity maintained, increase of peakdness for busy period and asymmetry D/U)

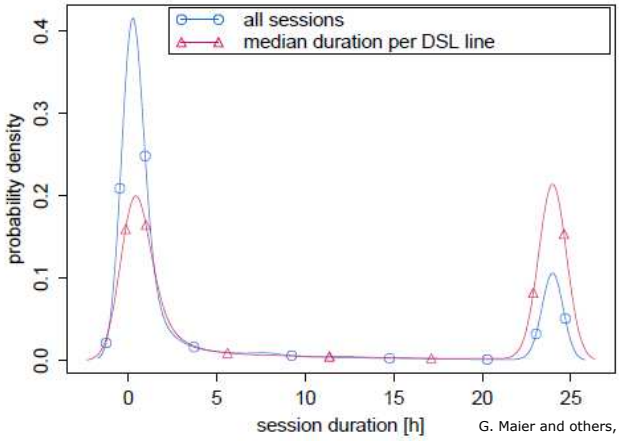
NGN Service traffic demand: Recent trends



Kenjiro Cho (IIJ/WIDE), Nov 2009

• Evolution of traffic volume distribution per day in Japan: (mode multiplied by ~ 50 in 4 years)

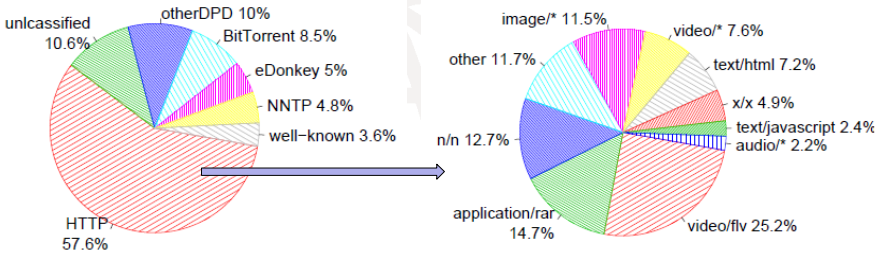
NGN Service traffic demand: Recent trends



G. Maier and others, T-Lab, Nov 2009

- Diversity of customers behaviour with peaks at short and long durations

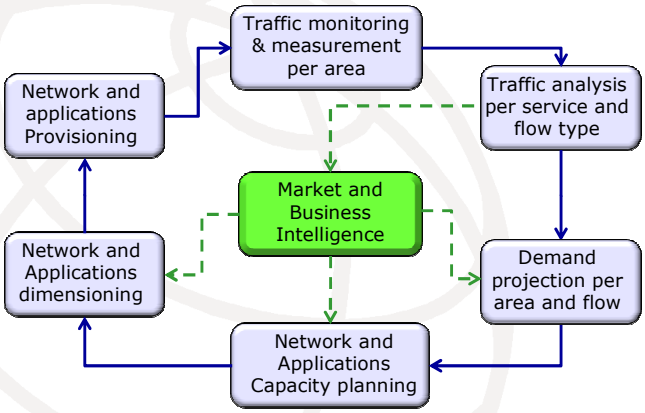
NGN Service traffic demand: Recent trends



G. Maier and others, T-Lab, Nov 2009

- Mix of applications with HTTP dominant overall and video dominant within HTTP

NGN Service traffic demand: Required Traffic Activities for NGN



Traffic/Planning activities and continuous cycle to engineer network according to the very dynamic evolution of applications (specially on mobile)

NGN Service traffic demand: Traffic activities for NGN

Demand projection per area and flow

- Customer segmentation per lifestyle and profile
- Customer location
- Customer/terminal usage pattern for media and signaling
- Customer purchasing criteria for services and bundles
- Service and application success rate

NGN Service traffic demand:
Traffic matrix characterization criteria

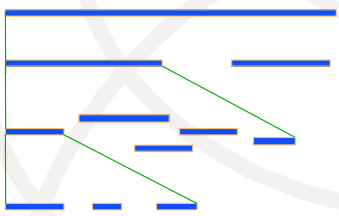
- Per O/D network points (end to end user, user to service providers location and multiple to multiple O/D)
- Per dimensioning criteria (constant, guaranteed streaming, best effort)
- Per application type (Video, Web, Bulk, P2P, Social networking, Gaming, etc.)
- Per customer category (Wholesale, LAN, business, residential)

Agenda

- **Services traffic trends on NGN**
- **Traffic modeling for NGN services**
- **QoS for NGN service types**

Traffic Characterization

• Hierarchical modelling for call driven communications generating traffic flows in NGN



Level 1: Customer Service time at "Call" level

Level 2: Activity/Communication times at Session/Application level

Level 3: Communication times at Flow/Burst leve

Level 4: Transmission times at Packet level

• Aggregated average traffic per level as a weighted average of the services categories (i) and customer classes (j) at that level.

Traffic Characterization for NGN

- **Which** units used to predict traffic demand ?

Traditional

- Customers for given project (operator, country, region, worldwide)
- Ports associated to customers per class
- Calls generated at user interface
- Erlangs originated/terminated at user interface

New

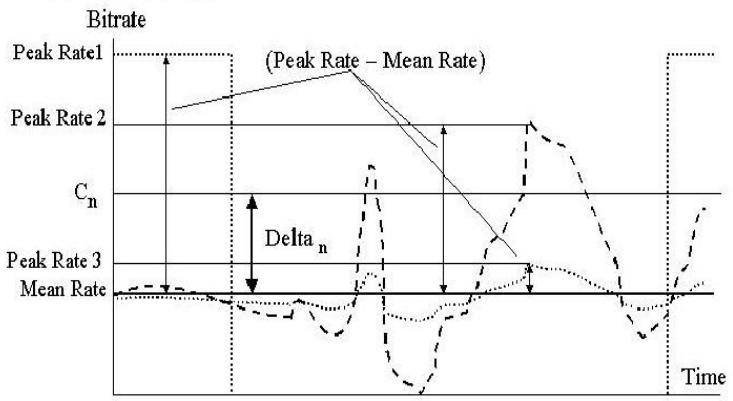
- Sessions/Information/requests generated at user interface
- Packets handled at a given resource through the network
- Mbits transported through a given network link/path

Traffic flow types for Quality of Service based dimensioning

- **constant stream**: bandwidth transmission at a constant speed with a specified delivery and jitter (ie: video distribution)
- **variable stream** : bandwidth transmission at a variable speed derived from a user information and coding algorithm which requires guaranteed quality and specified jitter (ie: VoIP, Video streaming, audio streaming, etc.)
- **elastic:** bandwidth transmission at a variable speed without jitter restrictions and asynchronous delivery (ie: browsing, file transfer, mail, UMS, etc.)

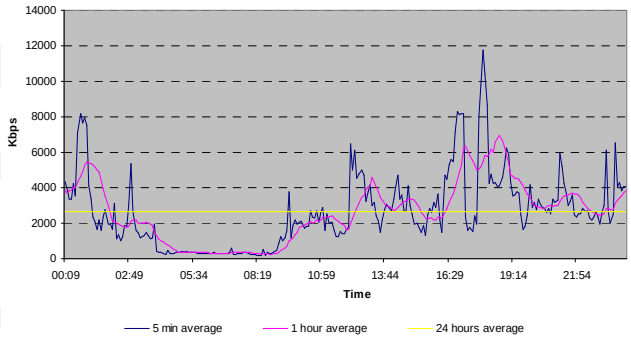
Traffic Characterization for NGN

- Different relation between peak traffic and average traffic per service classes: CBR (1), VBR(2), VBR(3)



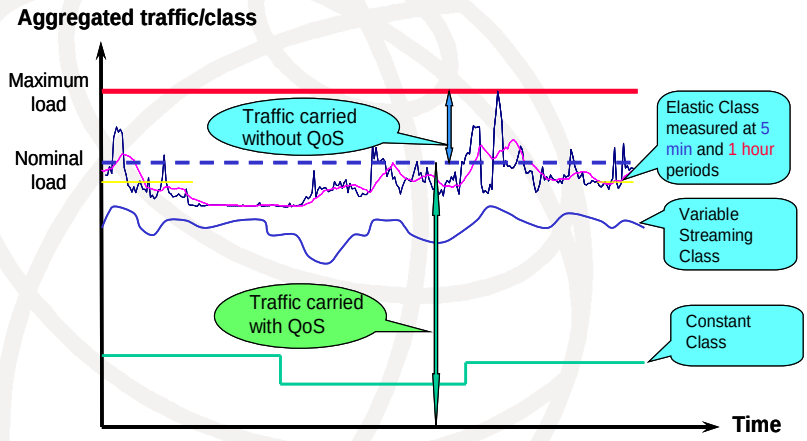
Example of time-scale influence on measurements

Variation per measurement averaging period at ENST campus measurements in 2001 for advanced internet applications



- Impact of averaging period
  - 1,8:1 ratio between "5 min" and "1 hour"
  - 2,3:1 ratio between "1 hour" and "24 hours"

Example of aggregated flows per category



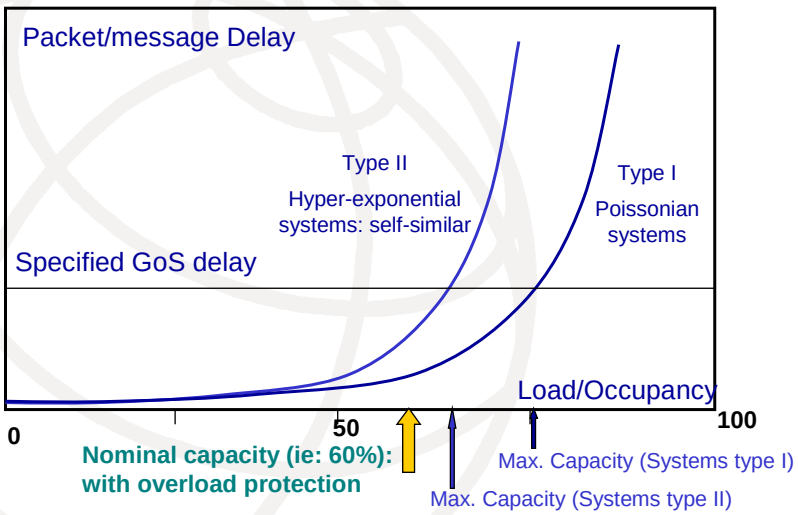
Traffic units for aggregated flows

- Traffic Units definition for network dimensioning
  - Equivalent Sustained Bit Rate (ESBR) or aggregated equivalent rates for same QoS category flows in a common reference busy period (ie. 5 minutes)
  - Computed as weighted average of the services at QoS category (i) and customer classes (j) at each network element: $\sum_i \sum_j ESBR_{ij}$

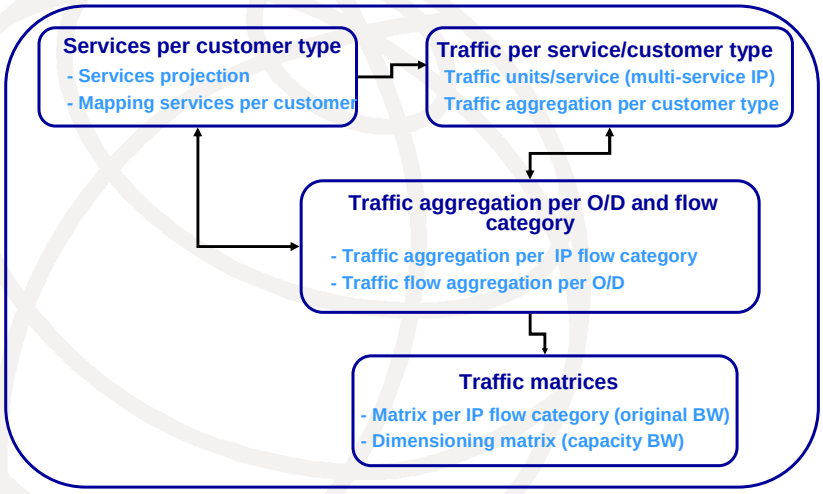
Dimensioning Criteria

- **Stream traffics** need reserve capacity procedures like MPLS and Call Acceptance Control (CAC) in the access and may be modeled with equivalent bandwidth methods.
  - ⇒ Available "multi-rate formulas" with different peakdness factors for a given quality.
- **Elastic traffics** may be modeled with resource shared models.
  - ⇒ Available "processor-sharing" models that provide a minimum capacity and a delivery speed as a function of simultaneous users
- **Constant rate** traffics need to be aggregated and reserved on top of the others with a given protection factor
- Overall dimensioning will be a combination of the previous procedures with different degrees of detail as a function of the model granularity

Typical dimensioning curves for delay based systems



NGN Service demand evaluation process



Agenda

- Services traffic trends on NGN
- Traffic modeling for NGN services
- QoS for NGN service types

Global QoS and Security

• Quality of Service (QoS): Characterization of the service accessibility and quality both with quantitative and qualitative (user perception) parameters and values

• Domains for QoS evaluation:

- **Grade of Service**
  - Service accessibility: capability to access a service
  - Connection establishment: Capability to get connection
  - Information transfer: Quality of information delivery
- Reliability: Failure probability
- Availability: Probability of system being active
- Survivability: Capability to provide service in abnormal conditions
- **Security:** Information and systems protection level
- Qualitative: Intelligibility, audibility, visualization, etc. of information content as derived from user perception

ITU framework for QoS support to operation

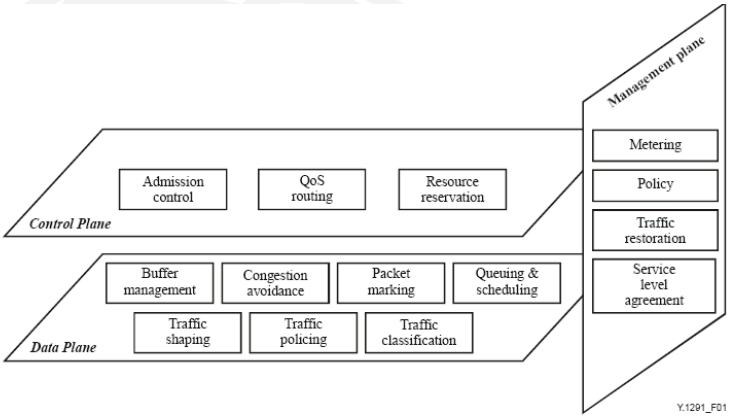


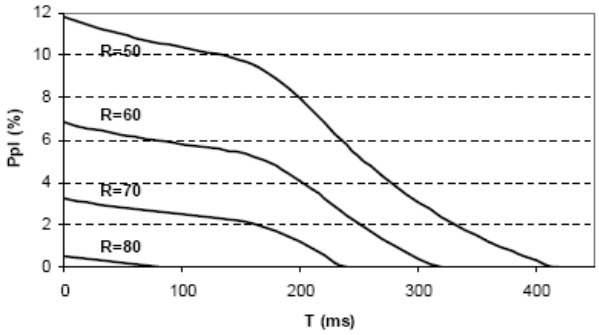
Figure 1/Y.1291 – Architectural framework for QoS support

QoS application phases and views

- - Phases of the service life cycle to analyze like:
  - service provision, service enhancement, service support, service connection, service billing, service management, etc.
- - Criteria for the quality observation like:
  - Availability, accuracy, speed, security, reliability, etc.
- - Customer view: QoS requirements and perception
- - Service provider view: QoS offering and achievement

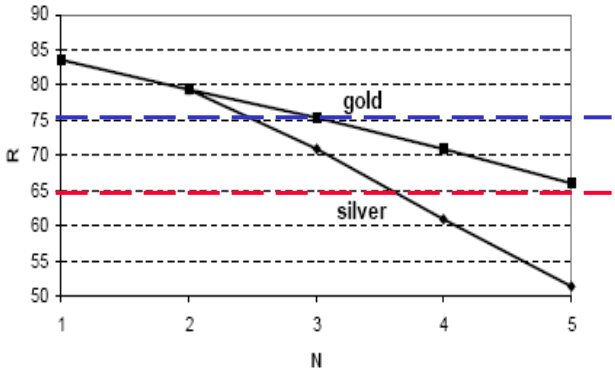
QoS rating for VoIP as function of packet loss and delay

Iso-quality curves as a function of packet delay and packet loss probability for G.729 (19th International Teletraffic Congress September 2005, Beijing)



QoS rating for VoIP as function of crossed hops

Perceived Quality of Service as a function of the number of crossed domains for the G.711+PLC coding with ppp = .01 and gold /silver SLA (19th International Teletraffic Congress September 2005, Beijing)



- **"Gold"** implies a T< 100ms, ppp< 1% and jitter <10ms (business customers)

- **"Silver"** implies a T< 150 ms. ppp< 2% and jitter < 30ms (residential customers)

QoS parameters and values for network design and planning

QoS	Priority	Bit loss Probability	Packet loss probability	Packet delay	Jitter	Availability
Stream constant	High	<10e-9	<10e-3	<150 ms	<10 ms	>99.999%
Stream Variable	High and medium	<10e-9 <10e-5	<10e-2 <5x10e-2	<150 ms < 400ms	<10 ms <30 ms	>99.999% >99.99%
Elastic	Low	<10e-3	Without guarantee	Without guarantee	Without guarantee	Without guarantee

Reference guide for QoS classes in IP operation by ITU

Table 2/Y.1541 – Guidance for IP QoS classes

QoS class	Applications (examples)	Node mechanisms	Network techniques
0	Real-time, jitter sensitive, high interaction (VoIP, VTC)	Separate queue with preferential servicing, traffic grooming	Constrained routing and distance
1	Real-time, jitter sensitive, interactive (VoIP, VTC).		Less constrained routing and distances
2	Transaction data, highly interactive (Signalling)	Separate queue, drop priority	Constrained routing and distance
3	Transaction data, interactive		Less constrained routing and distances
4	Low loss only (short transactions, bulk data, video streaming)	Long queue, drop priority	Any route/path
5	Traditional applications of default IP networks	Separate queue (lowest priority)	Any route/path
NOTE – Any example application listed in Table 2 could also be used in Class 5 with unspecified performance objectives, as long as the users are willing to accept the level of performance prevalent during their session.			

QoS reference parameters per call processing class by ITU

Table IV.1/Y.1530 – Call processing QoS class definitions and performance objectives

Call processing QoS class	Connection set-up parameters			Connection disengagement parameters	
	Connection set-up delay	Connection set-up error probability	Connection set-up failure probability	Connection disconnect delay	CPDP CCFP
QoS Class E (High or Excellent priority level)	Mean < 7500 ms 95%ile < 8450 ms (FFS)	Default (FFS)	Mean < A (FFS)	Defaults Mean = 3500 ms 95%ile = 4250 ms [I.352]	Defaults (FFS)
QoS Class 1 (Ordinary telephone level)	Mean = 7500 ms 95%ile = 8450 ms [I.352]		Mean = A (Value A is FFS) [I.359]		
QoS Class 2 (IP telephone level)	Mean > 7500 ms 95%ile > 8450 ms (FFS)		Mean > A (FFS)		
QoS Class U (Best effort level)	U	U	U	U	U

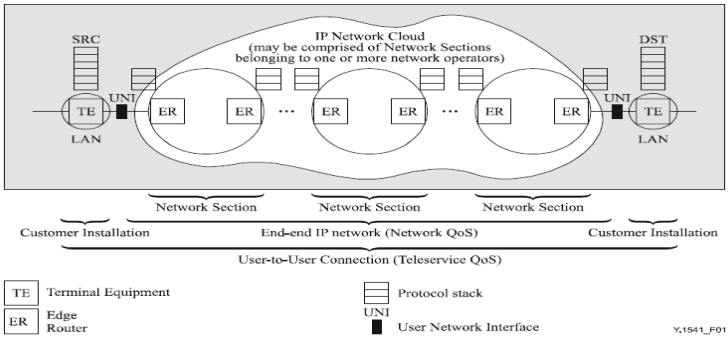
CPDP Connection Premature Disconnect Probability
CCFP Connection Clearing Failure Probability
U Unspecified or Unbound
FFS For further study

QoS reference values per call processing class by ITU

Table 1/Y.1541 – IP network QoS class definitions and network performance objectives

Network performance parameter	Nature of network performance objective	QoS Classes					
		Class 0	Class 1	Class 2	Class 3	Class 4	Class 5 Unspecified
IPTD	Upper bound on the mean IPTD (Note 1)	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	Upper bound on the $1 - 10^{-3}$ quantile of IPTD minus the minimum IPTD (Note 2)	50 ms (Note 3)	50 ms (Note 3)	U	U	U	U
IPLR	Upper bound on the packet loss probability	1×10^{-3} (Note 4)	1×10^{-3} (Note 4)	1×10^{-3}	1×10^{-3}	1×10^{-3}	U
IPER	Upper bound	1×10^{-4} (Note 5)					U

Reference path for end to end QoS application by ITU



NOTE – Customer Installation equipment (shaded area) is for illustrative purposes only.

Figure 1/Y.1541 – UNI-to-UNI reference path for network QoS objectives

Reference points for inter-domain performance measurement by ITU

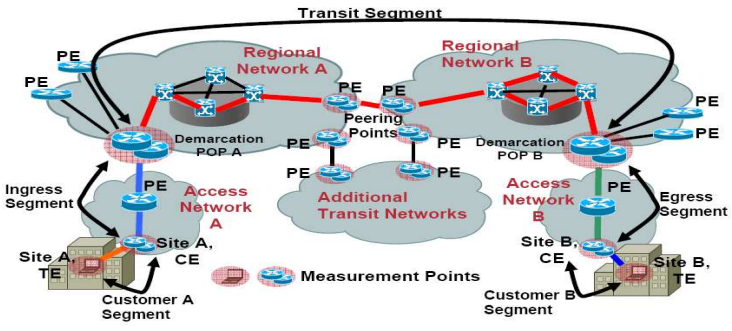


Figure 5/Y.1543 - TE-TE Model

Maturity of standards related to traffic and performance

- Standards well defined for NGN protocols and interfaces. Specific implementations may vary in additional features.
- Standards well defined for NGN intra-domain routing. Specific implementations may vary in additional features.
- Solutions defined for inter-domain routing but agreements on applicability and adoption at early stages. SLA negotiated on a bilateral basis
- Solutions for dimensioning and quality of service provisioning defined at intra-domain level and early definitions at end to end level.
- Traffic units and engineering rules available at scientific forums (i.e.: International Teletraffic Congress) but still not of extended applicability or standardization. Today case by case applications by experts.