Notes/Research/References on Performance Metrics

* In order to develop metrics for network performance in the cloud, we must first analyze the links that make up a general network.

http://www.cc.gatech.edu/~dovrolis/Papers/bwest\_survey.pdf

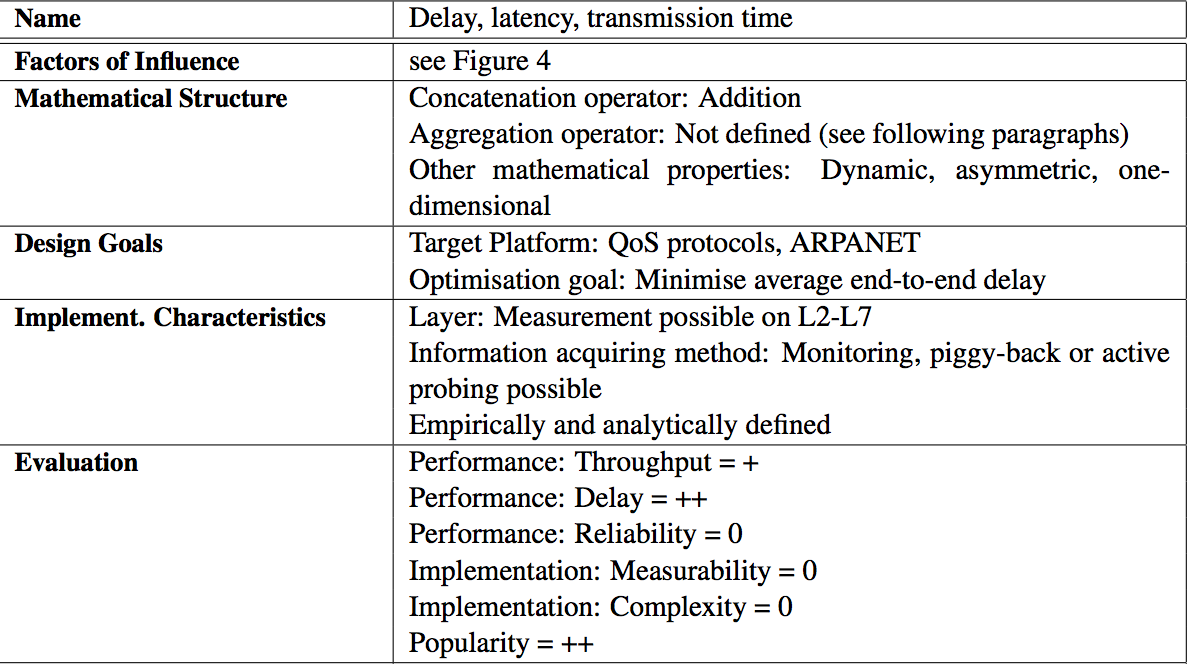
* In data networks, we look to quantify the data rate at which a network link or a network path can transfer.
* Bandwidth and throughput estimation is of interest to users wishing to optimize end-to-end performance.
* For many data-intensive application such as file transfers or multimedia streaming, the throughput available to the applications, directly impacts the application performance.
* Service level agreements (SLAs) between providers and customers often define service in terms of available bandwidth at key interconnection (network boundary) points.

<http://www.cisco.com/c/en/us/about/security-center/network-performance-metrics.html>

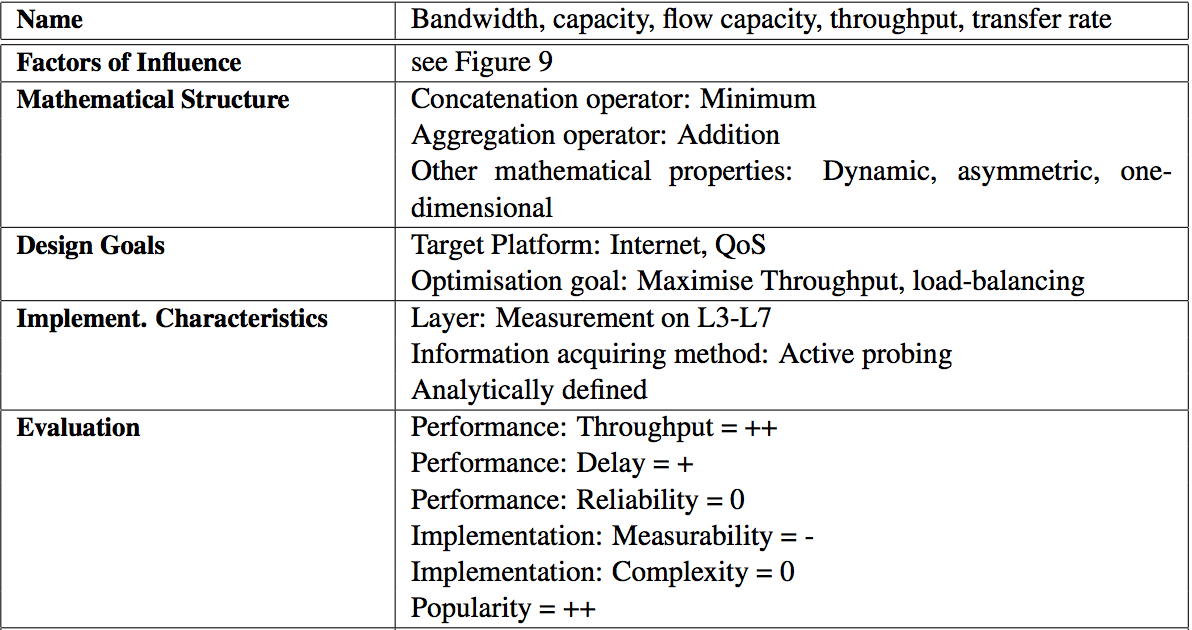
* Routers and switches are generally considered to be *stateless* devices because they forward each packet independently. Thus, metrics such as b/s and p/s may be sufficient to describe the performance of these devices.
* In addition to bandwidth (b/s), a metric such as p/s is also important in order to fully understand network performance. In literature, the interfaces of a network device are said to operate at line rate when the device is capable of forwarding packets, regardless of size. Thus, even for the smallest packets (highest packet rate), the network device will perform its functions.
* The the Cisco ASR 1000 Series Router, is capable of forwarding packets at up to 16 Mp/s with services enabled, it can support the processing of the equivalent of 10 Gb/s of traffic at line rate, with services, even for small packets.

<http://rainer.baumann.info/public/tik262.pdf>

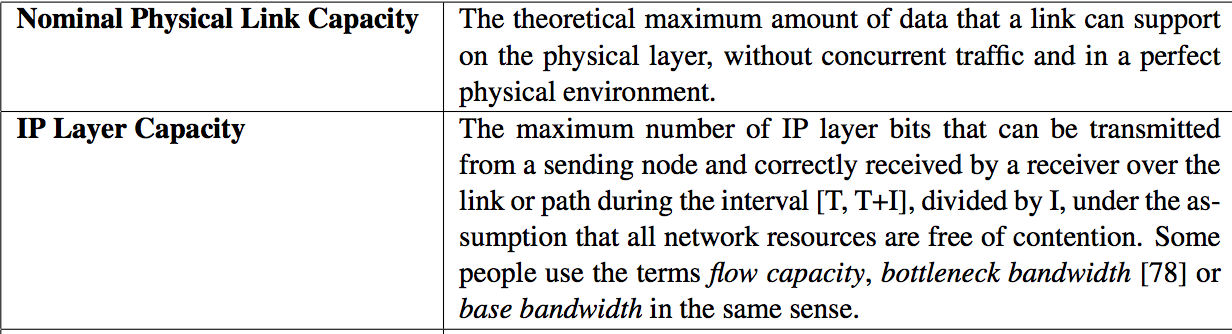
* The value of many metrics is governed by various different effects.
  + Network-Immanent Factors are defined as factors that depend directly or indirectly from the traffic within a network.
    - congestion,
    - internal interference (inter-flow and intra-flow interference)
    - the topology of the network

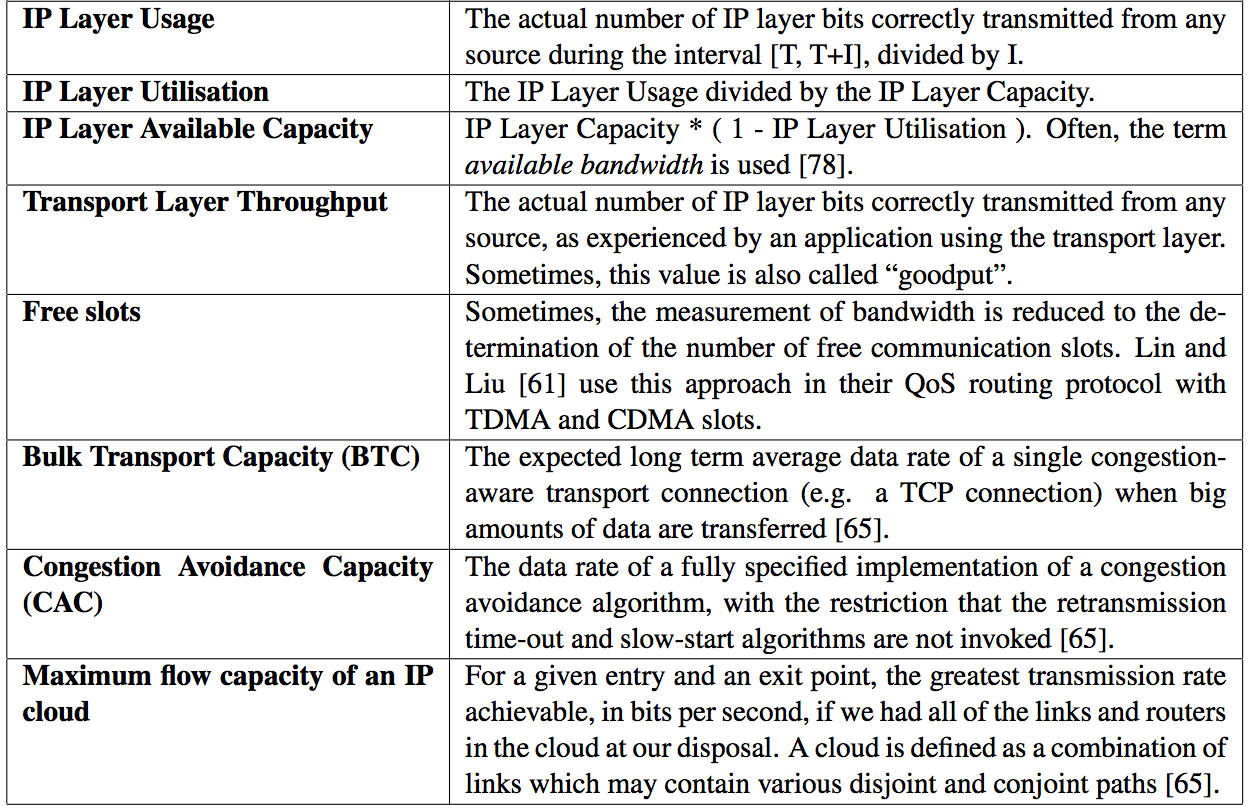


* The Round-Trip Time (RTT) quantifies the bidirectional delay of a link: A probe carrying a timestamp is sent to a neighbouring node. This node returns the probe immediately. Thus, the time is measured for a packet to travel to the neighbouring node and back again



Terminology:





<https://tools.ietf.org/html/rfc5166.html>

http://www.eurecom.fr/en/publication/4134/download/rs-publi-4134\_2.pdf

“Configuring and tuning OpenStack is complex: the large number of parameters that govern system behavior, the variety of hypervisor technologies (XEN, KVM or others), the different flavors of storage systems (various filesystem and logical volume management combinations), the number of alternatives to implement network switching (GRE tunnels, dynamic VLANs) all play a crucial role in determining the overall system performance.”

* VM-to-Host and Host-to-VM: these traffic patterns allow to measure the capacity available through the virtualization layer, that is between a network interface inside the VM and the corresponding TAP interface managed by the hypervisor;
* Distance 0: the physical network is not involved in this case, since all communications are established within the same physical host. As anticipated above, the bottleneck that determines the overall performance of this traffic pattern is related to VM-to-Host and Host-to-VM communications, which suffer from overheads due to network virtualization; • Distance 2: in this case, the physical network is the bottleneck. Virtualization overheads are low, as the loss in BTC for this traffic pattern is roughly 4% with respect to the physical upper-bound. Additionally, we remark that VM-to-VM communications at distance 2 achieve a BTC that is one order of magnitude lower than what can be obtained at distance 0;

• Host-to-Host through a GRE tunnel. GRE is a IP-over-IP tunneling technique used by OVS to instantiate a single switch spanning multiple physical hosts

<https://www.cs.rutgers.edu/~muthu/bquant.pdf>

* An important measure of performance perceived by the users is the round trip time (RTT) (which in turn affects dynamics of the network through mechanisms such as TCP flow control).
* the tails of the distribution of round trip times can become very stretched. Hence, to gauge the performance of the network in detail and the effect on all users (not just those experiencing the average performance), it is important to know not only the median RTT but also the 90%, 95% and 99% quantiles of TCP round trip timesto each destination.

https://github.com/tdunning/t-digest/blob/master/docs/t-digest-paper/histo.pdf

http://www.innervoice.in/blogs/2015/03/31/openstack-neutron-plugins-and-agents/