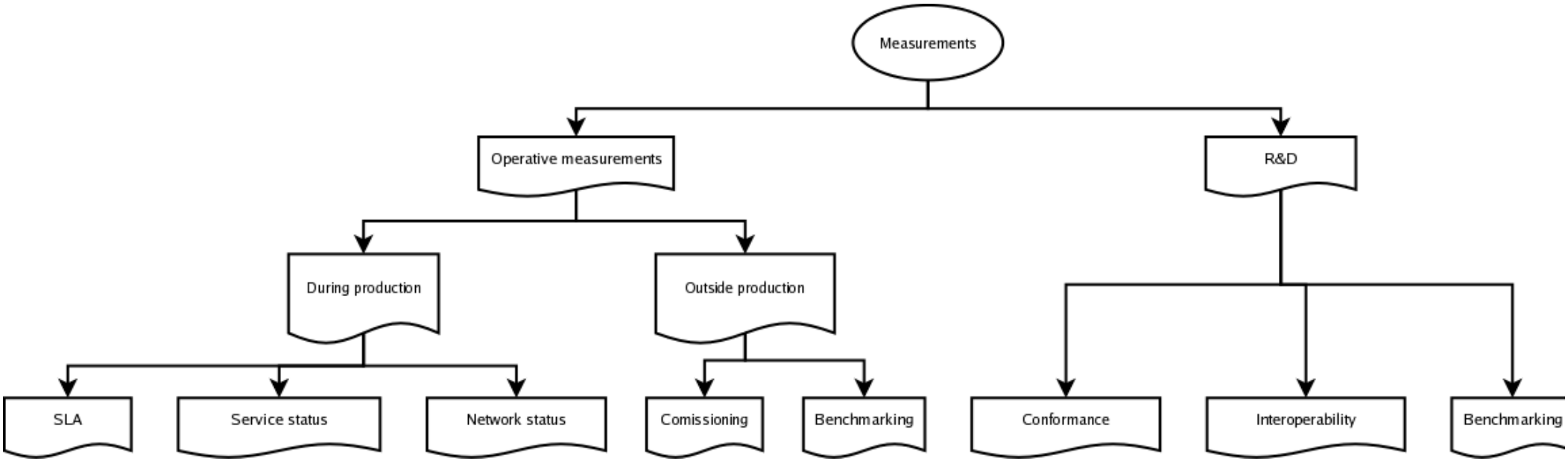


# S-38.3183: Internet Traffic Measurements and Measurement Analysis

## Lecture 11: Benchmarking routers and networking subsystems

# Different Measurements

- Measurements are performed for different reasons and goals

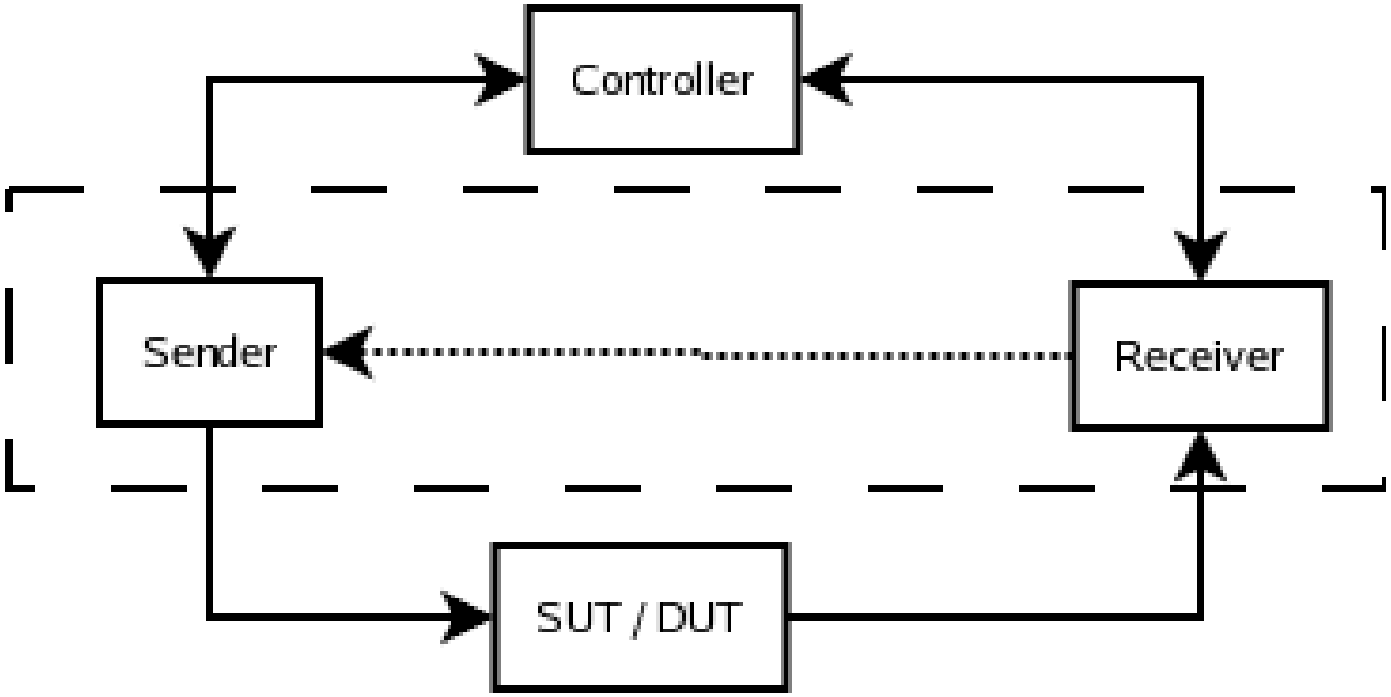


# R&D Measurements

- Measurement target varies depending on the stage of development
  - Algorithm
    - Scheduling, address lookup, route calculation
  - Protocol
    - OSPF state-machine, TCP state-machine
  - Device
    - Ethernet switch, IP router
  - Network
    - MPLS, IP, WiFi

# Generalization

- Generalization of R&D test flow with test-systems



# Approaches

- **White-Box**
  - System and its state can be observed during test procedure
  - System is well known and test flow can be designed to meet system characteristics
- **Black-Box**
  - System and its state cannot be observed during test procedure
  - System and its transfer function is not known
    - System state is approximated based on the input/output relationship

# Black-Box

- Common approach to determine performance of network subsystems
  - Controlled injection of test vectors
    - Packets, requests, etc
  - Operation of network subsystem is determined based on the response to the injected test vectors
  - Benchmarking of different systems

# White-Box

- Common approach in development of new protocols and algorithms
  - Tested algorithm contains suitable debug code to inject its internal state to external device
    - Causes extra processing which influences the performance
      - Discrepancies in real-time execution
  - External test vectors are injected to the system
    - Systems state and response are analyzed together

# White or Black

- In general knowledge of internal state helps in analyze of results
  - Large systems provide overwhelming amount of internal data
    - Important information is easily missed
  - In operative systems internal state analysis is restricted to non-real-time operations
    - Debug flags cause extra processing
- Black-Box analysis requires carefull inspection of measured system
  - Causality of injected test vectors

# Benchmarking

- Common measurement is benchmarking
  - Each system has operative boundaries
  - What is the performance of tested system in
    - Normal operative conditions
      - System is injected with vectors that conform protocol specifications
    - Abnormal operative conditions
      - System is in transient state caused by
        - » External influence
        - » Test vectors that cause transients

# Good Testing Practices

- Enable all protocols to be tested and conduct all testing with no further configuration of protocols.
  - This gives baseline performance.
  - A lot of protocols have variable timers which can be tweaked to tune in the performance.
    - This may lead optimized performance in test conditions
      - In real network these timer values may cause unstable network ,-)
  - Fine tune parameters after baseline testing
    - Apply only values which YOU know to work in real network.

# Good Testing Practices

- Take into account architectural differences of devices
  - Pizza-Box vs Modular system
    - Pizza-Box devices are usually brought out by integrating interface module and management module into closed chassis.
- Performance of these two is different when packets cross the backplane.
- In general:
  - The best results come from packets destined to same interface module
  - The worst results come from packets destined to other modules.
- Again depends on internal flow of packets ;-)

# Good Testing Practices

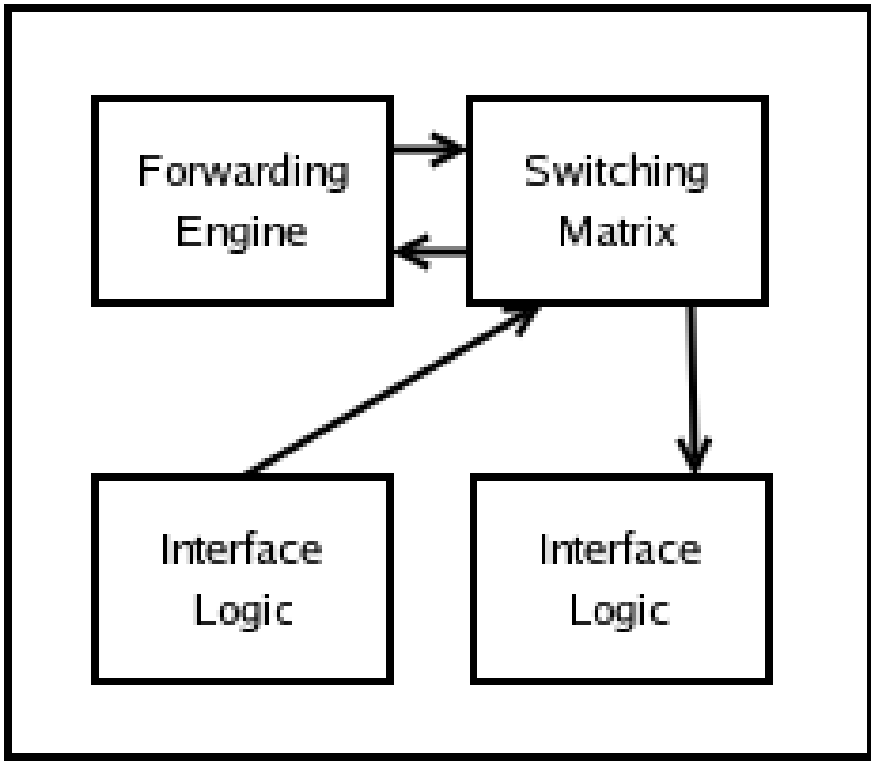
- Depending on the device type and architecture
  - Make sure that
    - System is reachable over the management interface throughout the testing procedure
      - Processor based systems are eager to stop communicating to outside world during the heavy load
    - There are valid control protocol adjacencies
      - It is common situation that routing protocol stops functioning but FIB contains entries that were there before crash

# Good Testing Practices

- Pick up the devices for testing from the shelves of a shop
  - Usually this can not be done with HC networking gear
  - At least update the software to a version that you know to be commonly used
    - Eliminates 'special price only for you versions' that are boosted to provide good results on your test cases
- Do not use IMIX address distributions as such
  - Device may be optimized for common test pattern

# Benchmarking Ethernet Switch

- Forwarding process
  - Frame is received from the interface
  - Frame header is forwarded to forwarding engine
  - Forwarding engine makes delivery header
  - Frame is delivered through switching matrix to outgoing interface
  - Frame is served to the link



# Constraints

- Processing power of interface logic
  - How many packets per second (pps) interface logic is able to handle
- Processing power of forwarding engine
  - How many pps FE is able to process
    - Lookup delay from the address database
- Capacity of switching matrix
  - How many bps switching matrix is able to handle

# Device dependent constraints

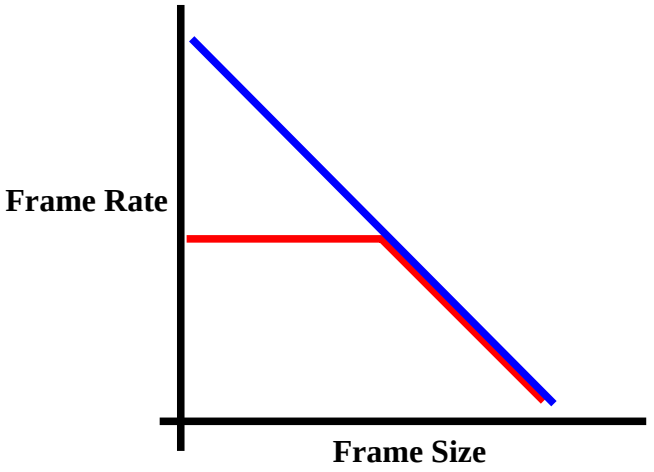
- Construction of device provides additional constraints
  - Pizza-box approach
    - All the logics are integrated into single blade
  - Modular approach
    - Functionalities are divided to different modules
      - Management module
      - Interface module
      - Processing module
    - Internal interface capacity vs sum(external interface capacities)
    - Packet flow within chassis (hot-spots?)

# Frame Sizes

- Minimum size from UDP Echo request frame
  - IP header (20 octets)
  - UDP header (8 octets)
  - MAC level header is required by the media in use
- Maximum frame size is determined by the limitations of the MAC
- Generally at least five different sizes containing minimum and maximum
  - Frame sizes to be used on Ethernet
    - 64, 128, 256, 512, 1024, 1280, 1518

# Analyzing Results

- Constraining element can be analyzed from the results
  - Blue line: Switching matrix or sum of interface capacities is restricting the operation
  - Red line: FE or interface logic is restricting the operation to certain extend after which Blue restrictions are valid again.

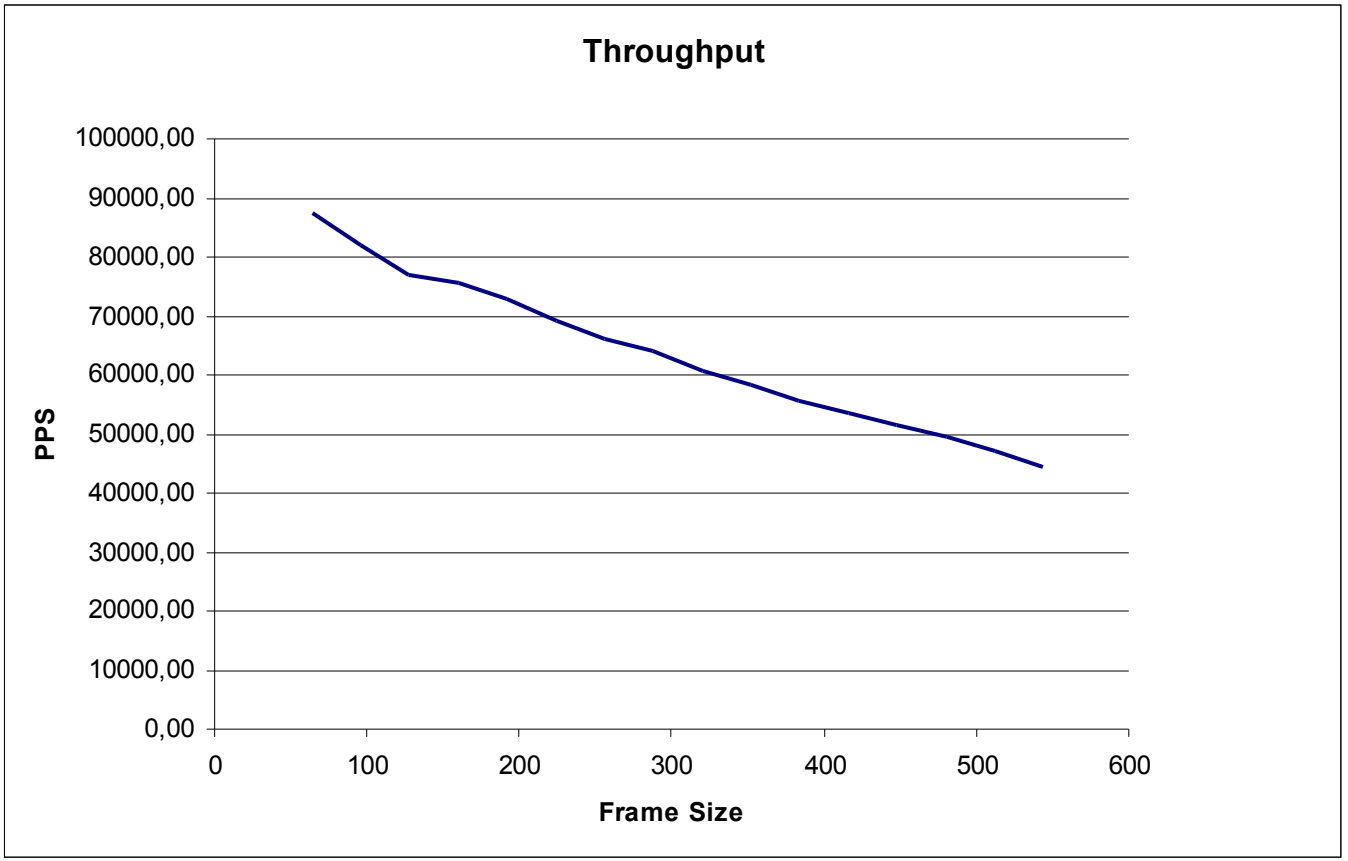




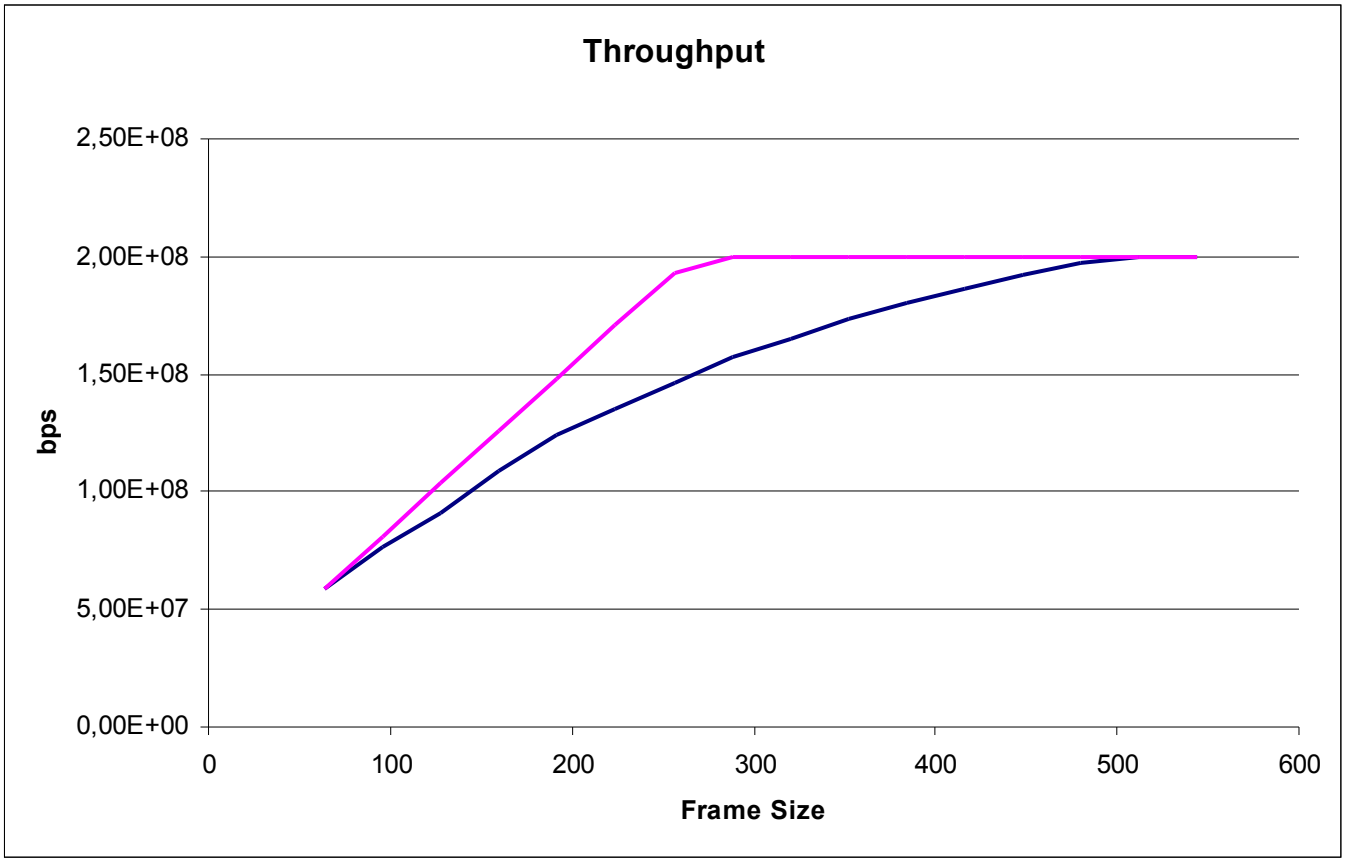
# Puzzle

- These are benchmarking results from a router
  - PC hardware
  - Linux Debian OS
  - 3 x 10/100
  - 2 x 1GbE
  - Connected:
    - 10/100 management
    - 2 x 10/100 test device
- Test methodology
  - Zero loss frame rate
    - Binary search with 1% accuracy
  - Frame size 64B +N*32B
  - Accelerated testing
    - 15s test period
      - too short for comprehensive testing
  - L2/L3 address discovery between iterations

# Puzzle


Throughput
100000,00
90000,00
80000,00
70000,00
60000,00
50000,00
40000,00
30000,00
20000,00
10000,00
0,00
PPS
0
100
200
300
400
500
600
Frame Size

# Puzzle


Throughput
2,50E+08
2,00E+08
1,50E+08
1,00E+08
5,00E+07
0,00E+00
bps
0
100
200
300
400
500
600
Frame Size

# Conformance measurements

- Conformance measurement is based on the analyzes of the DUT against formal definition of operation
  - Measurement is done with test system that fully conforms the protocol specifications
- Measurement reveals whether the system conforms the abstract protocol specification
  - Majority of protocol specifications allow some room in implementation
    - Mandatory part (conformance requirement)
    - Optional part



# Interoperability measurement

- Interoperability measurements are done against other manufacturers devices
  - Conformance measurement is only abstract verification of device operation.
  - Interoperability measurement reveals differences in different vendor implementations
    - What optional elements are implemented
    - Does DUT operate in multi-vendor environment
  - Only internal diagnostics available
    - What does the device tell about the network status

# Operative Measurements

- These measurements are done on real network
  - Cold network
    - No customer traffic is delivered in the network (real or virtual)
    - Comissioning the network
    - Benchmarking the network
  - Hot network
    - Customer traffic is delivered during the measurement
    - SLA measurement
    - Network status measurement

# Comissioning

- Comissioning is probably the least formalized measurement type
  - New network is build up and delivered to the customer
  - Before actual delivery, network is measured for the performance
    - Packet delivery
    - Routing stability
    - Service operation
      - VPN route propagation
      - Multicast delivery

# SLA Measurements

- Service Level Monitoring is important aspect for service provider
  - SLA contains measurable quantities for the offered service
  - These quantities are measured either with passive or active method
    - Active measurements are based on the probes which are used to determine delays and capacities
    - Passive measurements are based on
      - Sniffing on the traffic from different locations with time synchronized
      - Adding trailers to packets through the network