Week 5: Traffic measurements

1. Introduction
2. Packet capture
3. Interface counts
4. **Flow capture**
5. Traffic matrix
6. Anonymization of packet traces
7. Conclusion

IP Flows

- Set of packets with common properties
  - Definition can vary
    - ✓ Traditional 5-tuple: src IP, dst IP, src port, dst port, protocol
    - ✓ Packets from one ingress to an egress point
- Packets that are “close” together in time
  - Maximum spacing between packets (e.g., 15 sec, 30 sec)

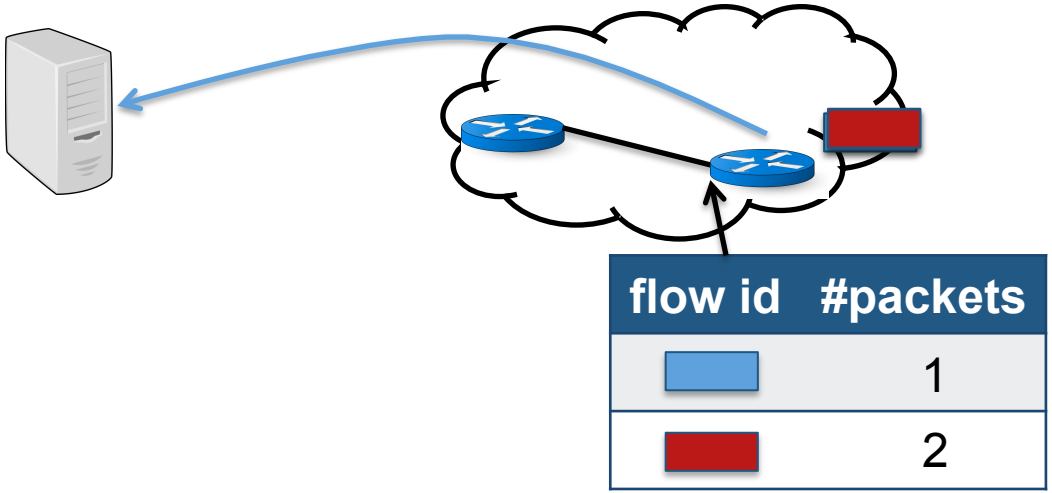


Flow ≠ application session

- Application session may be composed of multiple flows
- Packets in application session may not follow same links
- Hard to measure application session inside the network

Capturing flow statistics in routers

- Basic method
  - Specify set of properties that define a flow
  - Router log statistics per flow (flow records)
  - Push flow records to collecting process (IPFIX)



Packet Sampling

- Creation of flow records over the sampled packets
  - 1-out-of-m sampling of individual packets (e.g., m=100)
- Sampling reduces overhead
  - Avoid per-packet overhead on (m-1) out of m packets
  - Avoid creating records for a large number of small flows

Flow records: Flow identifier

- Packet header information
  - Source and destination IP addresses
  - Source and destination TCP/UDP port numbers
  - Other IP & TCP/UDP header fields (e.g., protocol, ToS bits)
- Routing information
  - Input and output interfaces
  - Source and destination IP prefix (mask length)
  - Source and destination autonomous system numbers

Flow records: Flow properties

- Aggregate traffic information
  - Start and finish time of the flow (time of first & last packet)
  - Total number of bytes and number of packets in the flow
  - TCP flags (e.g., logical OR over the sequence of packets)

Tools

- In-router capture
  - ✓ Cisco NetFlow
  - ✓ Juniper JFlow
- Collection and post-processing
  - ✓ Flow-tools
  - ✓ ntop

Flow monitoring: Pros and Cons

Pros

- More details about traffic compared to counters
- Lower measurement volume than full packet traces
- Available on high-end line cards (Netflow, Jflow)
- Control over overhead via aggregation and sampling

Cons

- Less details than packet capture
  - No individual packet arrival times
  - No information on packet content
- Not uniformly supported (getting better with IPFIX)
- Computation/memory requirements for the flow cache

Picture credits

- **Router clip art** by www.openclipart.org (public domain) (slide 4)
- **Server Linux box clip art** by OCAL, www.clker.com (public domain) (slide 4)