Packet Sampling and Network Monitoring

CERN openlab Monthly Technical Meeting
13\textsuperscript{th} November, 2007

Milosz Marian Hulboj
milosz.marian.hulboj@cern.ch

Ryszard Erazm Jurga
ryszard.jurga@cern.ch
What is “Network Monitoring”?

- Network “Health” Inspection
- Observation and analysis of following objects:
  - Network devices
  - End systems
  - Network links
  - Network traffic
  - Network applications
Why Network Monitoring (1)

- Networks are getting more complex and harder to comprehend
- Networks are a business-critical element
- Occurrence of problems in any network is inevitable:
  - Increasing configuration and topology complexity
  - Increasing number and complexity of threats, attacks, viruses, etc.
  - Conclusion: It is just a matter of time
- Detect the problems as early as possible
- Reduce the unavailability time
Why Network Monitoring (2)

- **Network Statistics:**
  - Identification of performance characteristics:
    - For traffic engineering (pkt/s, bytes/s, connections/s, flows, traffic matrix)
    - QoS metrics, latency, bandwidth (SLA, billing)
    - Planning (busiest services, traffic distribution, throughput)

- **Network Inventory:**
  - Identification of equipment on the network

- **Troubleshooting:**
  - Failures of interface cards, power supplies
  - Connectivity problems
  - Service availability
Why Network Monitoring (3)

- Accounting
  - User activity

- Security
  - Policy violations:
    - Unauthorised services, machines
    - Unauthorised access
    - Unauthorised applications (e.g. p2p)
  - Intrusion detection
  - Compromised hosts detection
  - Protection against cyberattacks, worms, etc.
MOORE'S LAW

- Intel® 4004 Processor
- Intel® 8080 Processor
- Intel® 8086 Processor
- Intel® 286 Processor
- Intel® 386™ Processor
- Intel® Pentium® Processor
- Intel® Pentium® II Processor
- Intel® Pentium® III Processor
- Intel® Pentium® 4 Processor
- Intel® Itanium® Processor
- Intel® Itanium® 2 Processor
- Dual-Core Intel® Itanium® 2 Processor

transistors

- 10,000,000,000
- 1,000,000,000
- 100,000,000
- 10,000,000
- 1,000,000
- 100,000
- 10,000
- 1,000

Packet Analysis - old methods (1)

- Sniffing in the old times ("old shared Ethernet")
- Slow network speed
- Captures everything (all packets+payload)

- "Old shared Ethernet" is a history...
Port mirroring:
- Captures all the traffic (per port, group, VLAN, etc)
- Requires HW support
- Requires fast network interface
- Problematic determination of originating port

Network device-based data:
- Captures (partial) data from selected ports
- Sampled packet data
- Sampled flow data
- Requires HW support
Other Common Sources of Data (1)

- SNMP
  - Operations via simple variable manipulation
  - Standard mean for retrieving generic statistics, network status, etc:
    - Packet arrival and departure rates, packet top rates, error rates, system load, etc.
  - Used also for network configuration
  - Cannot customise monitored variables within agent
  - Different vendors use different proprietary MIBs for detailed information
Other Common Sources of Data (2)

- **RMON and RMON2**
  - Extension of the basic set of SNMP
  - Remote data collection and processing
  - RMON2 decodes packets at layers 3 – 7 and handles certain protocols
  - Collects aggregate statistics (volume, rate, Top Talkers, etc) about network and application traffic
  - Implementation of RMON agents is complex
  - Probes might be expensive and require administration
  - Cannot add new features to the existing MIB
Packet Sampling

- A mean of passive network monitoring
- Simple to implement
- Low CPU and memory overhead
- Sample only the packet header (~128 bytes)
- Traffic patterns estimated from the samples with certain error
Error Estimation

- Decreasing error = increasing the sampling rate
sFlow Packet Sampling

- RFC 3176
- Multi-vendor standard
- Complete packet header and switching/routing information
- Some SNMP counters information
- Low CPU/memory requirements – scalable
Usage of sFlow

- Profiling network traffic
- Building flow statistics
- Accounting and billing
- Route profiling (forwarding information)

- Security analysis / intrusion detection:
  - Packet headers analysis
  - Traffic pattern analysis
Current Research

- Influence of sampling on flow estimates
- Influence of sampling on anomaly detection:
  - Access only to packet headers
  - Unable to reconstruct the sessions from samples
- Traffic prediction:
  - Packet count prediction
  - Traffic volume prediction
- Adjusting of sampling rate:
  - Attempt to maintain the constant error
  - Attempt to fully utilise the hardware capabilities
• http://www.sflow.org
• And many more…