InMon
internet monitoring

Experiences with Widespread Packet Sampling

Sonia Panchen

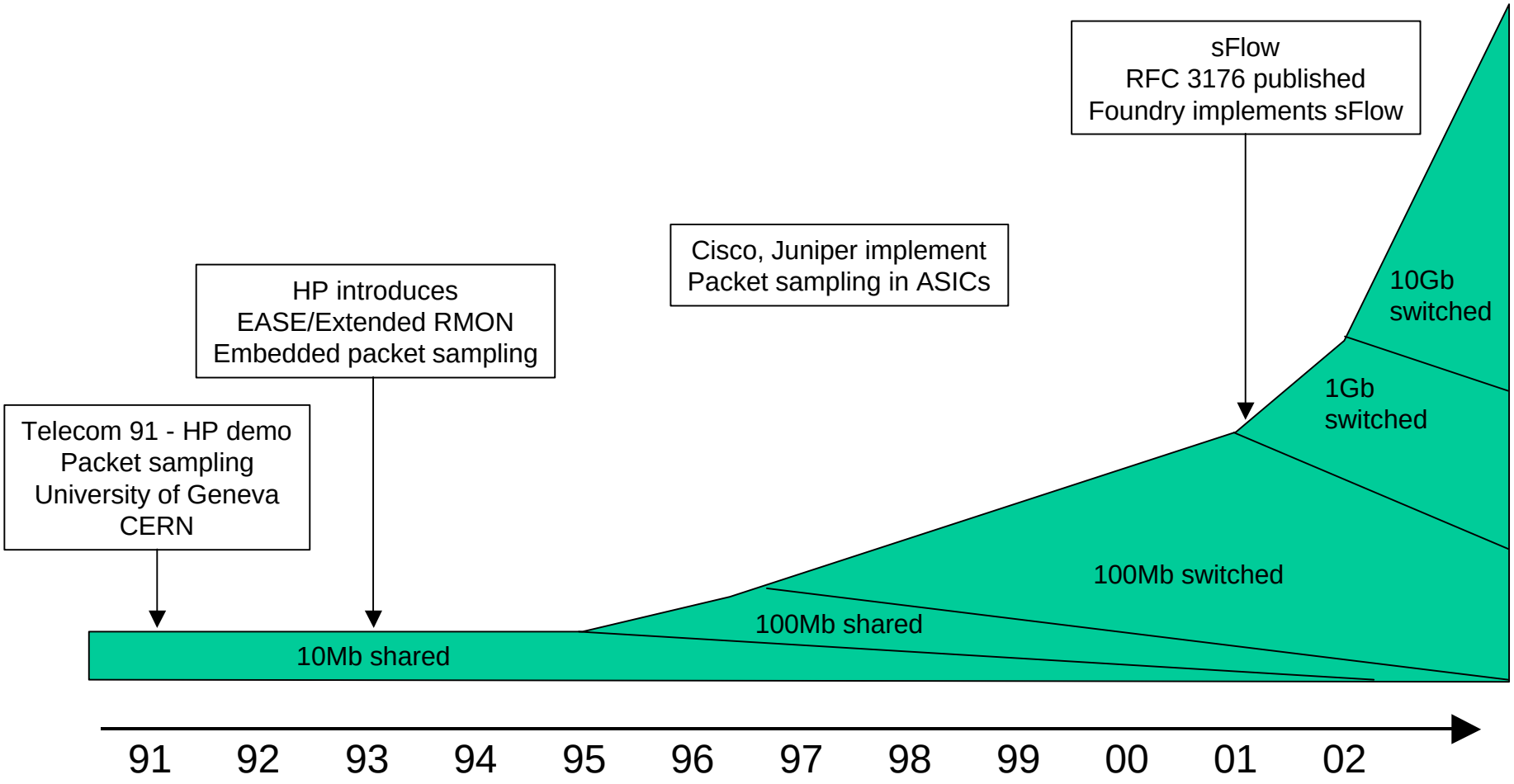
InMon Corp

sonia_panchen@inmon.com

Topics

- Packet sampling history
- sFlow (RFC 3176) – a packet sampling system
  - sFlow measurement system design goals
  - Packet sampling algorithm
  - sFlow datagram
  - Statistical model for packet sampling
- Using sampled traffic data
  - Example usage
  - Sampling system configuration
  - When not to use sampled traffic data
- Discussion
- More information

Packet Sampling History



sFlow Measurement System Design Goals

Accurate

- Quantitative traffic measurements even at Gb speeds
- Forwarding information

Timely

- Up-to-date statistics on traffic flows

Scalable

- monitor 1000s agents from a single point

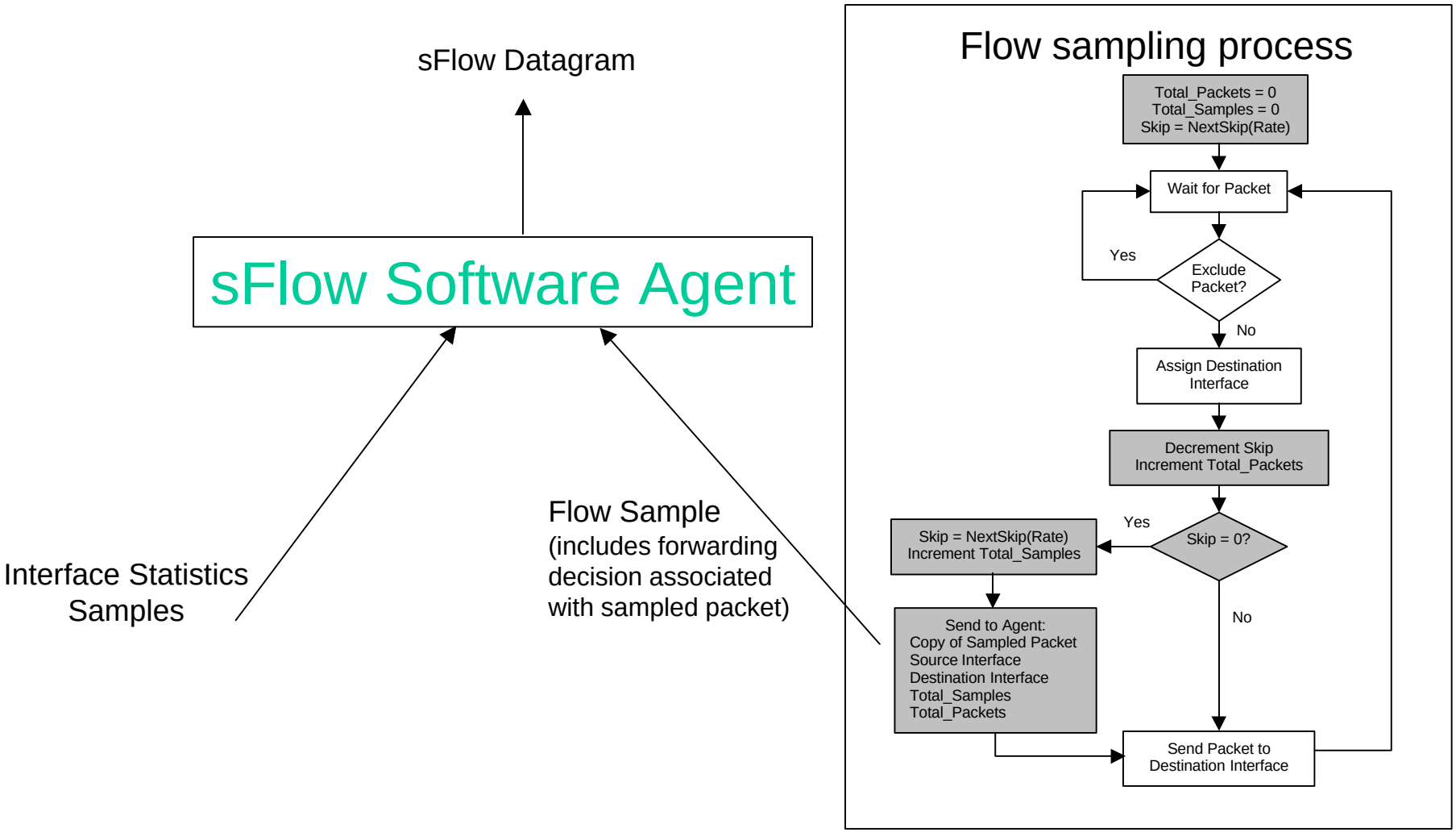
No impact on performance

- Switch or router
- Network overhead

Low cost embedded implementation

- Encourage pervasive deployment

sFlow Packet Sampling Algorithm



sFlow Datagram

- Interface statistics samples
- Flow sample
  - Packet header (MAC,IP,IPX,AppleTalk,HTTP,FTP,DNS…)
  - Sample process parameters (rate, pool etc.)
  - Switch
    - Input/output ports
    - Priority
    - VLAN
  - Router
    - Source/destination prefix
    - Next hop address
  - Gateway
    - Source AS, Source Peer AS
    - Destination AS Path
    - Communities, local preference
  - User
    - User IDs (TACACS/RADIUS) for source/destination
  - URL
    - URL associated with source/destination

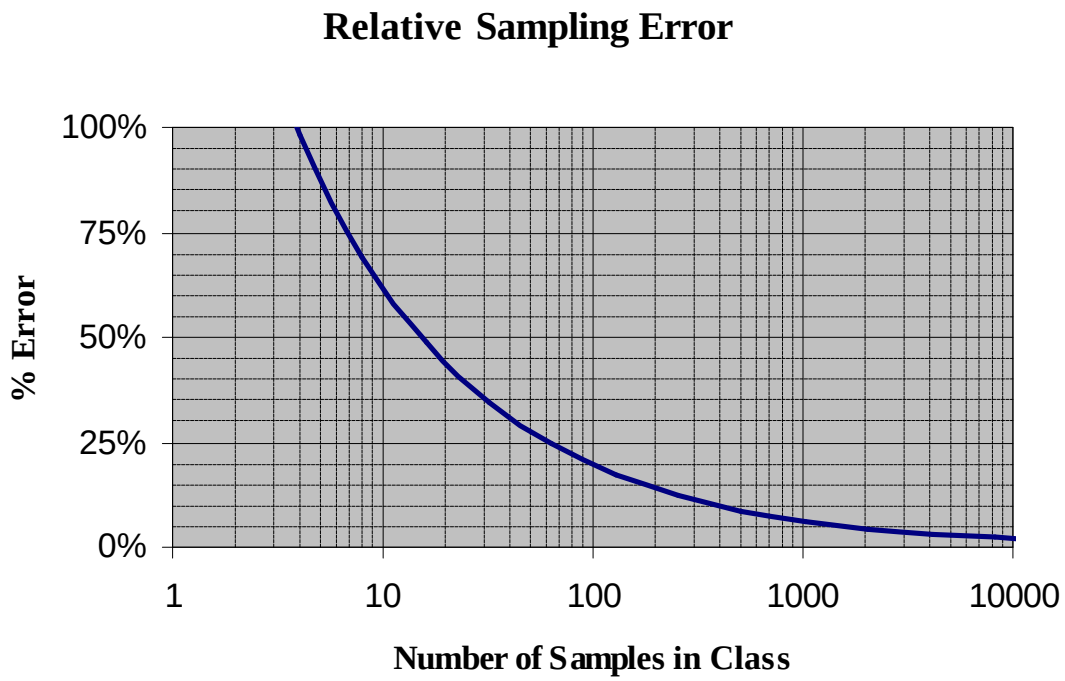
- Raw data
  - Delay processing
- Detailed
  - Traffic
  - Forwarding
  - I/F counters

Statistical Model for Packet Sampling

Estimating Traffic per Protocol

Total number of frames = N
Total number of samples = n
Number of samples in class = c
Number of frames in the class estimated by:

$$N_c = \frac{c}{n} \cdot N$$



$$\%error \leq 196 \cdot \sqrt{\frac{1}{c}}$$

Using Sampled Traffic Data

Characterizing Traffic

- Top Talkers for any protocol or group
  - Diagnose congestion
  - Characterize DOS
- Top AS Paths
  - Flows carried by AS paths
  - Understand overall traffic flow
  - Optimize peering relationships
- Traffic fan out analysis
  - Track worms
  - Network misuse

Using Sampled Traffic Data

Traffic Profiles

- TCP connection profiles
  - Packets/connection
  - Connections/second
  - Connection duration
- TOS distribution
- IP fragmentation analysis
  - Top sources/protocols associated with fragments
- Packet header pattern matching
  - Worm and infected host detection
- Packet size distribution

Using Sampled Traffic Data

Accounting

- Network-wide audit trail
  - Identify security threats
- Planning
  - Cost effective upgrades
  - Effective use of traffic shaping
- Billing
  - Allocate costs
  - Revenue generation

Sampling System Configuration

- Always-on
  - Invariant underlying packet sampling
  - Continuous real-time statistics
  - Historical audit trail
- Sampling rate
  - Static
  - Average of sample/second per sampling entity
    - Congestion characterization
      - lots of samples/interval (but still <1% of traffic)
    - Billing – sufficient accuracy over longer periods

When not to use Packet Sampling

- Understanding specific and unusual events
  - Use packet sampling for existence
  - Filtering and hashing for diagnosis
- Total address space usage

Discussion

Promote the use of sampling for reliable, detailed, timely traffic measurements

- Standard *measurement* primitives
  - Measurement
  - Data format
  - Configuration
- Simple and cost effective for pervasive implementation in ASICs
- Literature, experience, and models to encourage sampled data analysis
- Use existing export techniques appropriate to specific applications

IETF 53 Network Monitored with sFlow

- Foundry Networks switches with sFlow
- http://166.63.177.97
  - NTOP
  - Traffic Server

More Information

- **sFlow RFC 3176**
- **Statistical model & computing confidence intervals from sampled data**
  - www.hpl.hp.com/techreports/92/HPL-92-35.html
- **Forum for developers and users of products, services and tools based on the sFlow**
  - www.sflow.org
- **sFlow tools**
  - Open source
    - ntop, tcpdump, snort
    - sFlowTool (www.inmon.com/sflowTools.htm)
  - Commercial
    - InMon Traffic Server, Foundry IronView, HP-IUM
- **sFlow Agent source code**
    - www.inmon.com/sflow.htm