

HP Network Node Manager iSPI Performance for Traffic Software 10.00



Collecting traffic data using sampling

Table of contents

Capturing Network Traffic Statistics	2
Introduction - Sampling	2
Types of Sampling	2
NetFlow Sampling.....	2
Collector (Leaf) Sampling.....	3



Capturing Network Traffic Statistics

With rapid growth of IP-based networks, a simultaneous need for measurement technology has emerged that can provide the pieces of information that are necessary for recording network and application resource utilization. Cisco's IOS NetFlow provides a solution for this requirement.

NetFlow is a functionality of Cisco routers and switches that captures a rich set of traffic statistics. These statistics contain user, protocol, port, and type of service details that you can use for a wide variety of purposes such as network analysis and planning, accounting, and billing. You can configure the NetFlow feature on a per-interface basis on Cisco routers and switches.

Introduction - Sampling

NetFlow is designed to be able to process all the packets on an interface, which can be expensive in environment like Internet backbones because of the extra processing power required for capturing all the information from each packet.

The NetFlow sampling technology enables you to collect NetFlow statistics from a subset of incoming (ingress) IPv4 traffic on the interface. You can select only one packet out of N sequential packets (you can configure "N" for an interface). By choosing to process only sampled packets, you can substantially reduce the CPU utilization needed to process the complete set of packets. Sampling reduces volume of data to be processed, but still provides a view of most IP flows passing through the network.

Types of Sampling

Cisco provides three types of sampling:

- Deterministic sampling – This type of sampling lets you select every Nth packet. This type of sampling can give incorrect statistics when traffic flows arrive in fixed patterns.
- Random sampling – This type of sampling selects, on an average, one out of each N sequential packets. It is statistically more accurate than the deterministic sampling.
- Time based sampling – This type of sampling selects a sampled packet every N milli-seconds.

Sampling Support in the NNM iSPI Performance for Traffic

NNMi SPI Performance for Traffic 10.00 supports following types of sampling –

- NetFlow sampling
- Collector(Leaf) sampling

NetFlow Sampling

As outlined above, NetFlow sampling is configured on per-interface basis on a router to export NetFlow statistics to a collector (Leaf Collector).

NNM iSPI Performance for Traffic supports random sampling mode for version 9 export format.

Version 9 is a flexible and extensible export format and is on the IETF standards track in the IP Information export (IPFIX) working group. The NetFlow version 9 export format can export a rich set of details like Layer 2 information, new security detection and identification information, MPLS, Multicast, BGP information, IPv6, and so on.

NetFlow version 9 export format includes two types of templates:

- Template Record – This template is used to define the format of subsequent flow records exported by the router. It includes information on template id, number of fields in the template record, type of fields and length of these fields, in bytes. Template records are periodically resent by the router.

NNM iSPI Performance for Traffic would cache any template record received and then parse any flow records received by locating the appropriate template record within the cache.



- Options Template – It is a special type of template record; this type is used to export the format of the information related to the NetFlow process (for example, sampling information). The sampling data is exported only when sampling is configured at the router. Options template, too, are periodically resent by the router.

The NNM iSPI Performance for Traffic can cache any options template received, and then parse options data records containing NetFlow process (for example, sampling information). Based on this sampling information, the NNM iSPI Performance for Traffic extrapolates bytes, packets, and flows details in sampled flow records to estimate the actual traffic in network.

Areas of Application

Sampling is very useful for traffic engineering, network planning and capacity planning where every flow may not be needed to understand the network behavior.

Advantages

The Sampling feature addresses the performance bottleneck experienced while turning on full NetFlow. Statistical traffic sampling significantly reduces the utilization of router resources while exporting useful NetFlow data. Packets sampling results in less flow records being exported by the router, which reduces the processing load on the collector.

Disadvantages

Sampling, at the router side, works by sampling IP packets flowing from a source to a destination. If there are significant changes in IP packets volume (byte value) then it becomes difficult to calculate near estimate of the actual traffic flowing in the network. Some of the IP flows may not be captured at all.

Collector (Leaf) Sampling

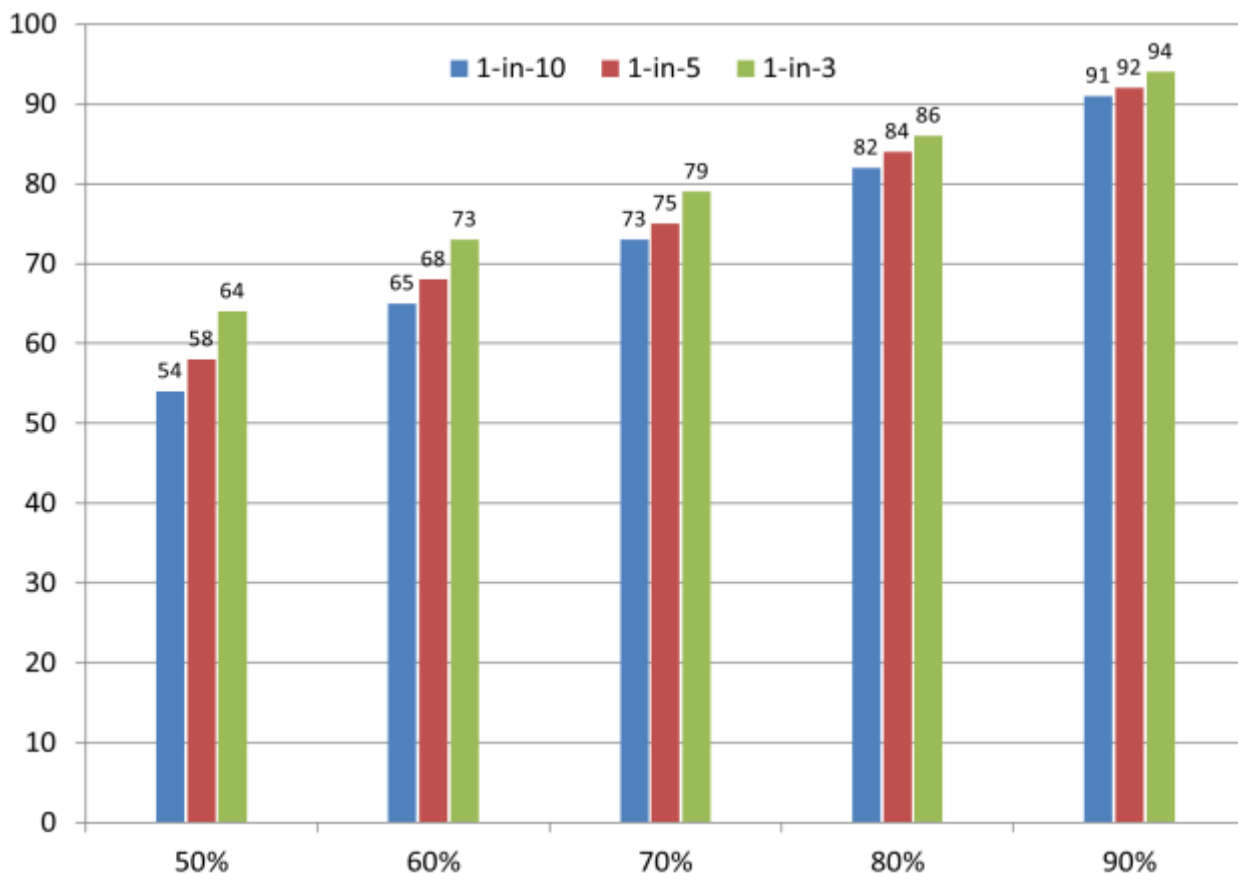
It is a new type of sampling incorporated in the NNM iSPI Performance for Traffic. This type of sampling is applied on the collection engine (Leaf in the NNM iSPI Performance for Traffic) and works by sampling flow records from exported datagrams (exported packets) received from the router. An export datagram packet may contain up to 30 flow records for version 5 or 9 flow export.

The NNM iSPI Performance for Traffic has implemented an algorithm that enables it to report high level of accuracy in terms of bytes, packets, and flows information for sampled IP flows. The algorithm is such that once the NNM iSPI Performance for Traffic samples a flow record, it processes all the subsequent flow records corresponding to the same IP flow. Thus, it produces better estimates of network traffic without involving any extrapolation.

The NNM iSPI Performance for Traffic allows operator to configure Leaf Sampling through a configuration page that accepts sampling rate and sampling mode inputs among others. For better accuracy, it is recommended that sampling rate is entered between 2 to 10. As sampling rate decreases (or value of N increases), system samples less and less flow records for estimating network traffic and thus it lowers reporting accuracy. Our lab results show that it is small but gradual decline in accuracy with decrease in sampling rate. The reporting accuracy is directly proportional to flow records uniqueness ratio



(70% unique IP Flows/min will produce more accurate results than 60% unique IP Flows/min).



Above figure shows sampling rate (Legends) versus accuracy % (Y-axis) for various uniqueness ratio (X-axis)

Leaf sampling can be configured in two sampling modes –

- Random Sampling – it samples 1 out of every N sequential flow records.
- Top Flow Record Sampling – it samples a flow record (from N sequential flow records) that has maximum value of bytes in it. This mode caters to specific use cases where operator is interested in capturing flow records with high volume (more bytes). Such high volume flow records may remain non-sampled with Random sampling as it randomly picks a flow record irrespective of its volume (byte value).

Areas of Application

- Leaf sampling should be used when the NNM iSPI Performance for Traffic flow records processing rate is not able to cope up with flow records arrival rate because of high scale or low hardware or something else then enabling Leaf sampling helps overcome such situations easily.
- Leaf sampling should also be used to gauge traffic trends in network where knowing exact traffic statistics is not required.

Advantages

Leaf sampling allows the NNM iSPI Performance for Traffic to process 1-out-of-N flow records only. Thus, it helps in reducing processing load, almost, by a factor of N. It provides a fairly accurate idea of traffic statistics and data trends in network. It enables NNM iSPI Performance for Traffic to tune the sampling rate depending upon its current processing load and hardware setup.



Disadvantages

Leaf sampling works by sampling (or picking) a subset of flow records. Rest (non-sampled) of the flow records are dropped and not processed further. In such a mechanism, some IP flows will not be captured at all and therefore will not be reported by the NNM iSPI Performance for Traffic.

