

Performance,
Scalability,
&
Integration

Valencia Systems Aruba Suite™ for NetFlow Management

Valencia Systems
1040 Woodcock Road,
Suite 100
Orlando, FL 32803

Phone: (1) 407.228.4417
Fax: (1) 407.228.4419

www.valenciasystems.com

Table of Contents

Section 1: Introduction

Section 2: NetFlow backgrounder

Section 3: Aruba Suite Overview

Section 4: Summary

Section 1: Introduction

IT organizations are challenged with maintaining the cost and quality of applications delivered over the network. It has become a corporate mandated to maintain or improve quality of mission critical applications while reducing the cost associated with the service. NetFlow provides the data necessary to determine the who, what, where, when and why of Network IP traffic.

Section 2: NetFlow backgrounder

NetFlow gives network managers a comprehensive view of application flows on the network. Cisco IOS® NetFlow technology collects and measures data as it enters specific routers or switch interfaces. NetFlow provides a set of services for IP applications including network traffic accounting, usage-based network billing, network planning, security, Denial of Service monitoring capabilities, capacity planning, troubleshooting and network monitoring. NetFlow provides valuable information about network users and applications, peak usage times, and traffic routing.

A major advantage of NetFlow is as a data source for measuring the quality and cost of applications delivered over the network. NetFlow provides a key component in deterring how your services are performing and what needs to be done to optimize service delivery for standard applications and emerging applications like VoIP or Video. Prior to the advent of NetFlow, the only way to accomplish this measurement was to deploy, costly, network-RMON2 probes. This became extremely costly since most applications were crossing multiple segments throughout the enterprise and to obtain these measurements, probes were required on every traffic segment.

NetFlow is data is accumulated by Flow caching in the routers and switch interfaces. NetFlow analyzes and collects IP data flows entering the router and prepares data for export for analysis. It enables the accumulation of data on flows with unique characteristics, such as IP addresses, application, and CoS. NetFlow supports key technologies, including IPv4, IPv6, Multicast, and Multiprotocol Label Switching (MPLS).

NetFlow captures seven data keys in each captured record:

- Source IP Address
- Destination IP Address
- Source Port Number
- Destination Port Number
- Layer 3 Protocol Type
- Type of Service-ToS-bits
- Input logical interface

NetFlow captures flow records and bundles into export packets. Each packet contains approximately 5 KBytes, is typically 20-50 flow records. These export packets are broadcasted over UDP and the frequency of transfer is dependent on the amount of traffic on the NetFlow enabled interfaces. As traffic increases, the frequency of broadcast goes up. In a busy, NetFlow enabled, network the amount of data accumulated is extensive, potentially in the millions of flows per hour. This is the main reason for requiring a tiered, scalable, performance management solution for NetFlow analysis. Most industry solutions are designed to poll data on regular intervals and not designed to handle the vast amounts of UDP traffic collection.

Section 3: Aruba Suite Architecture & Features



The Aruba Suite leverages a Highly Scalable data collection capability and database to provide a "carrier-class" consolidated IT performance management reporting solution. Aruba collects extensive raw SNMP data from hundreds of different networked devices to isolate "hot spots" and report on real-time conditions for troubleshooting, Capacity Planning, Service Level Management (SLM), Business Service

Management (BMS), Quality of Service (QoS), financial-bill back, and operational needs. Raw data is rolled-up into historical reports for current time periods or multi-year analysis. Aruba is easily configured and installed. Aruba ships with hundreds of preconfigured reports and can be easily customized (using XML) for individual requirements for specific users (IT Managers, Operations, Engineering, etc). Aruba's powerful base lining capability allows enterprises and service providers to determine typical vs. atypical behavior, trending and predicting behavior and identifying security or policy breakdowns. Aruba is designed to be integrated with existing IT applications to provide a comprehensive, scalable, IT performance management solution.

Aruba Product Features

The Aruba Report/Server

- Carrier-class scalability
- Historical trending & predictive analysis
- Hardware & network capacity planning
- Comprehensive System Management reporting
- Resource utilization & optimization reports
- Efficient provisioning of carrier capacity reports
- Proactive problem area detection
- Verify/Audit QoS goals
- Verify/Audit SLA goals
- Automatic Baseline comparisons
- Threshold based exceptions and alarms
- Secure Access
- ODBC compliant database
- XML- customization

The Aruba Distributed Flow Collector

- Unprecedented scalability, millions of concurrent flows

- Automatic Baseline comparisons
- Service Level (SLA) assurance
- Measure service quality (QoS)
- Utilization
- Congestion
- Errors
- Real-Time Analysis
- Historical trending & predictive analysis
- XML-Web based reporting
- Comprehensive picture of Network Usage, raw & aggregate data
 - Top N Talkers, Hosts, Listeners
 - Top N Conversations
 - Top N Ports
 - Top N Applications

The Aruba Distributed Poller

- Distributed remotely or centralized
- Add modularly as requirements grow n-tier architecture
- Multithreaded, high speed, polling of SMNP data cache
- Standards-based SNMP polling

Aruba is a highly scaleable architecture utilizing a client, server, collection layers and an Oracle™ database

The general Aruba Architecture is a scalable architecture which supports the remote/distributed collection of data from routers, hubs, switches, computer systems and other SNMP capable devices (i.e. firewall, or traffic shaping appliances). By distributing the polling function the data may be consolidated from the distributed collectors in a central data warehouse. Data is analyzed against preconfigured or customized report formats to produce actionable, information-based, web reports. Reports are easily categorized and grouped by function, location, priority and detail allowing Network Operations-to-Management utilize performance information for IT troubleshooting or financial and business planning.

Aruba was specifically architected to support very large-scale SNMP/NetFlow data collection and reporting. The architecture is both multi-tiered and distributed. Aruba can be divided into three logically separate layers: **Client layer**, **Server layer**, and **Collection layer**.

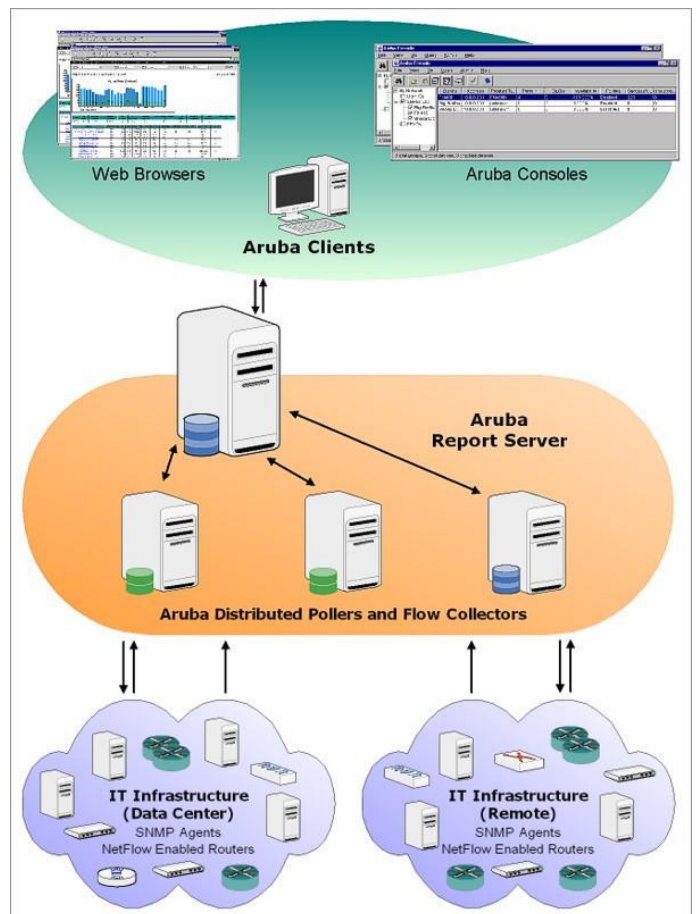
The Client layer provides the user interface to the system. The Aruba Console allows the administrator to configure and maintain the system, as well as perform real-time and custom reports.

The Server layer (*Aruba Report Server*) is responsible for generating and publishing reports, storing configuration information, and managing the underlying Pollers and Flow Collectors. There is a single Report Server in an Aruba system.

The Collection layer consists of one or more Pollers and/or Flow Collectors which automatically gather data from the network devices. These clean, normalize and store raw data, as well as aggregating and aging historical data.

The Client, Server and Collection may all be collapsed onto a single machine or they may be distributed on separate

computers. For small networks, a single Client, Report Server, and Poller running on the same machine will be appropriate. Larger networks may require multiple Clients



simultaneously connecting to the Report Server, that in turn, managing data from several distributed Pollers and/or Flow Collectors. Aruba's architecture will scale from small to very large distributed installations, without sacrificing simplicity and ease of use.

Why a multi-tiered, distributed architecture?

The optimal solution must minimize data movement by pre-processing the data at the collectors

A monolithic approach, a data collection engine and reporting engine both hosted on a single platform, will eventually be limited by processor performance, available bandwidth, and storage capacity. And long before these thresholds are reached, the user experience is impacted - continuous polling and database processing will consume much of the available machine cycles. Just adding more machines introduces a new problem - each machine now has its own collection and reporting domain. The user must know which machine is responsible for each device, and groups of devices can not span the domains.

An alternative approach is to separate the administrative and reporting functions from the data collection engines. This still gives the user a single machine to access for a holistic view of the entire network, for both reporting and administration. The CPU/bandwidth-intensive collection functionality is distributed across as many machines as needed. An additional benefit is that a collector can now be located close to the data sources, reducing SNMP polling across the WAN.

The downside of this approach, if the collectors rely on the server layer to actually process and store all the data that is collected, is the problem has not been solved. In fact, this only makes it worse, since the data is essentially moved twice. Many solutions which started with a monolithic architecture have been "scalability enhanced" by adding a "server layer" using this approach and must manipulate the data across platforms.

Valencia System feels the optimal solution must minimize data movement by pre-processing the data at the collectors - only summarized data is sent up to the server layer. Granular data (for example, raw polled data) is retained in the collector database, and is only moved to the server if user requested.

Reducing the ongoing maintenance burden is a major factor in a solution's scalability. To reduce maintenance burden the client layer must be distributed as well. The administrators console should tear away, and multiple administrators must be able to simultaneously make changes. In addition, grouping must be integrated into the administrative functions as well as the reports.

NetFlow

Cisco NetFlow IOS® tracks unicast IP Packets as they enter the router through a specific interface and tracks IP packets on a "per flow" basis. NetFlow answers questions regarding IP traffic: who, what, where, when and how. NetFlow provides

information for troubleshooting, capacity planning, policy management, security, and cost and quality used to evaluate financial decisions. This information can be collected to produce a baseline and audit information to support the foundation of Business Service Management (BSM).

System scalability is a critical requirement

Traditional performance management solutions are not designed to handle the demanding, real-time raw data, requirements of NetFlow management. The historical IT environment, "pre-NetFlow", for gathering data was not constrained to managing large amounts of real-time data but utilized off-hour polling techniques to collect the raw metrics into a datastore for analysis. With NetFlow, the paradigm has changed. When NetFlow is enabled on a Cisco router, a NetFlow cache is built by the router to track flows as they enter the router in a 64-byte record (one per flow) that details each respective flow. The size of the NetFlow cache (NetFlow data storage) is dependent on the router platform and/or the amount of memory in the router. Most routers are optimized to dump cache before the router performance will be impacted. This requires a NetFlow management solution to support a, scalable, multi-tiered architecture for real-time collection and processing. Further, most performance management vendors will tout their ability to monitor a large number of interfaces but the reality is that this number is misleading. With NetFlow the number of interfaces is irrelevant as *NetFlow needs to scale to support the number of concurrent Flows*. This analysis of flow traffic could easily reach into the MILLIONS of flows.

Deployment

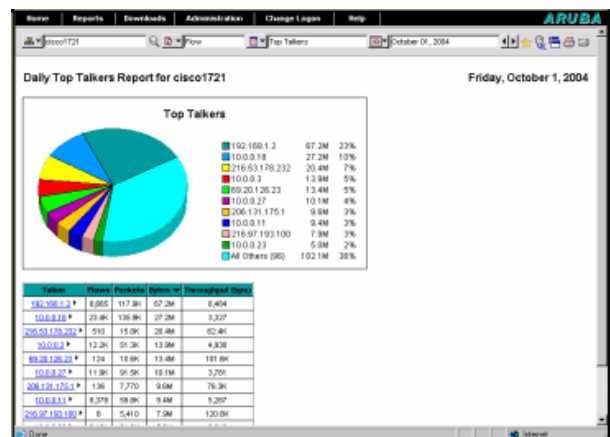
Because almost all routers support NetFlow (even non-Cisco devices), the capital investment required to use Aruba's NetFlow feature is very low. It is straightforward to configure, requiring only a few configuration commands to have the router automatically forward application traffic flow statistics.

Aruba's NetFlow reporting overlays the infrastructure already monitored by Aruba; it takes advantage of Aruba's flexible grouping capability, secure portioned access, and XML-based reporting architecture.

Aruba's NetFlow feature adds the next level of detail to Aruba's extensive SNMP reporting, leveraging Aruba's multi-tiered architecture to *support the collection of MILLIONS of application flows* and the analysis to support operational

"Number of interfaces managed" is not an accurate performance measurement when you need to manage millions of Flows

Aruba collects and reports on flow records from routers and switches that support Cisco's NetFlow, enabling extensive IP traffic analysis without the use of probes.



decisions. At the report level, a user may literally "drill down" from the SNMP based report and quickly troubleshoot the specific application traffic that has caused the link utilization to spike. At the highest level, Aruba shows the following information for any interface or device:

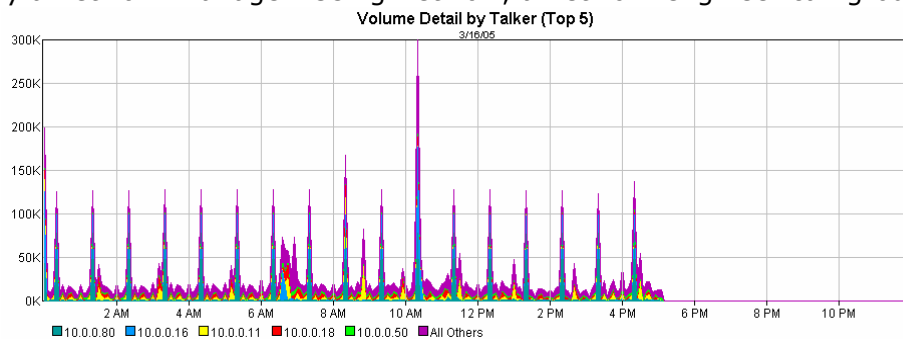
- Top N Talkers
- Top N Hosts and Listeners
- Top N Conversations (Host Pairs)
- Top N Network Protocols
- Top N Traffic Types
- Top N Applications

Aruba can automatically detect that a WAN link has suddenly exceeded its baseline, and will notify a network manager. Using NetFlow, a network engineer can group of interfaces for the selected time period.

Network Protocols are based on server ports.

Aruba comes with a standard list of Network Protocols that map IP ports to Network Protocol names. Users can also define IP ports (and layer 4 protocols) as Traffic Types, and multiple port numbers can be mapped to the same Traffic Type.

Business Application definitions are based on IP Address/Port pairs. Users can define an Application as one or more pairs. An application can be running on multiple machines and/or multiple ports. Specific hosts, pairs, ports, etc. are also searchable which allows the detection of specific traffic (i.e. virus on port X). Aruba also supports full customized reports, queries against flow data and integration with 3rd party systems.



Aruba collects NetFlow to support the measurement of cost and quality for Business Service Management

Section 5: Summary

NetFlow is a relatively new tool in the arsenal of IT managers. Augmenting and replacing expensive Network Probes to determine the cost and quality of services on the network has become of extreme importance when IT budgets are shrinking while the IP network is asked to do more. By effectively deploying a scalable performance management solution leveraging NetFlow data, IT can manage and lower costs, increase uptime, detect possible security threats, and improve the quality of service delivery.

To learn more about Valencia software solutions, please visit us on our web site at www.valenciasystems.com