

## Premise-Based Optimization: How It Works

By multi-homing with a premise-based optimization product, the ubiquity and diversity of the Internet can be leveraged to its fullest potential as a business communications platform. This white paper explores the concept of a premise-based optimization solution, its functional and operational requirements, integration and deployment scenarios, and the Internap approach to delivering such a solution.

An Internap White Paper

©2008 Internap Network Services Corporation. All rights reserved.



## Table of Contents

|                                      |   |
|--------------------------------------|---|
| Executive Summary                    | 3 |
| Introduction                         | 4 |
| Premise-Based Optimization Solutions | 5 |
| Route Optimization Components        | 5 |
| Measurement Modules                  | 5 |
| Control Module                       | 6 |
| Reporting Modules                    | 6 |
| Flow Control Platform                | 6 |
| Determining Prefixes                 | 7 |
| Measuring Quality                    | 7 |
| Policies                             | 8 |
| Construction of Policies             | 8 |
| Route Updates                        | 9 |
| Configuration                        | 9 |
| Management                           | 9 |
| Reporting                            | 9 |
| Summary                              | 9 |



## Executive Summary

Even though business use of the Internet has grown tremendously over the last decade, little of that growth has come from the use of advanced applications like VoIP, office-to-office transactions, supply chain management or customer relationship management. The lack of adoption of these advanced applications stems mainly from the technical and operational shortcomings of the current Internet architecture.

These limitations primarily affect cost control, quality assurance and resiliency requirements. Though standards groups have attempted to address these problems in different ways, most approaches suffer from interoperability limitations or prohibitively complex operational requirements.

Therefore, early adopters have chosen to achieve quality in their business infrastructure either by deploying significant private infrastructures or by purchasing Internet bandwidth from multiple providers (multi-homing). However, both approaches force companies to accept compromises in the cost, quality and resiliency of their communication infrastructures.

While private infrastructures provide many of the desired aspects of a business communications system, they are expensive to deploy and maintain. Private infrastructures also lack the ubiquity and the associated efficiencies of the Internet. Multi-homing to the Internet, on the other hand, is a useful strategy to achieve resiliency and some degree of quality. Unfortunately, multi-homed networks are difficult to manage in terms of cost control, and quality assurances are delivered only on a “best-effort” basis.

These advanced applications are now becoming a requirement for enterprises and the companies that service them. It is possible to meet the needs of these advanced applications through multi-homing by deploying a premise-based optimization solution capable of maintaining business and application specific policies whether it be VoIP, CRM, ERP, etc. Whatever the application network requirements are, it can be improved with the Internap® premise-based optimization solution, Flow Control Platform™ (FCP) solution. The Internap FCP solution enables enterprises to control costs and optimize quality by taking advantage of the ubiquity of the Internet and the path diversity provided by multiple Internet connections. These measurement capabilities allow companies to better manage its network operations and look at defined parameters for enhancing operations.

This white paper explores the concept of a premise-based optimization solution, its functional and operational requirements, integration and deployment scenarios, and the Internap approach to delivering such a solution.



## Introduction

The Internet is defined as the ultimate network of networks. Indeed, the number, variety and size of the networks are as diverse as the applications and content that the Internet carries. It is, therefore, understandable that the Internet and its associated protocols have not always evolved in a manner best suited for even the majority of the networks connected to it.

The Border Gateway Protocol (BGP), for example, was designed to provide Network Layer Reachability Information (NLRI) across administrative (inter-AS) boundaries. The algorithms behind BGP ensure only reachability and a loop-free path to a given destination prefix. BGP is unaware of network performance issues, often maintaining routes even under poor network conditions such as high latency or packet loss.

Furthermore, BGP was primarily designed, as are most Internet protocols, to accommodate the most complex of networks in a simple yet effective manner. It was not designed with edge networks in mind. However, BGPv4 is the standard dynamic inter-AS routing protocol, making it the protocol of choice for leaf networks with multiple Internet entrance points. Networks with a single connection to the Internet typically utilize a default static route as dynamic routing is not necessary.

Since most enterprise networks are leaves that do demand resiliency and quality, BGP is not an ideal protocol for dynamic inter-AS routing between these enterprises and its associated transit providers. Unfortunately, there is no other dynamic routing alternative for inter-AS connectivity. As a result, enterprise network engineers managing multi-homed networks must attempt to craft traffic engineering policies within the confines of the BGP protocol in order to achieve its unique quality, bandwidth, cost and resiliency requirements. There have been attempts to address the quality assurance issues associated with this structure, however, most remain un-ratified as Internet standards, and generally suffer from a lack of interoperability. Most approaches also require complex configuration and management and are not capable of providing quality assurances across network boundaries. The most prominent of the current approaches are MPLS, Diffserv and Intserv, all of which are constrained to administrative boundaries and all require complex configurations and operational maintenance.

A pragmatic and very common approach to achieving quality on the Internet has been to trade cost for quality by over-provisioning bandwidth and by multi-homing. While multi-homing is not necessarily a costly proposition, it is usually accompanied by significant bandwidth over-provisioning. In fact, given the constraints of BGP routing, it is often necessary to over-provision bandwidth due to resiliency requirements and also due to the fact that BGP's distribution of traffic can be unpredictable and difficult to control.

However, with newer pricing structures that are based on bandwidth usage, over-provisioning of bandwidth can be cost-effective for enterprises, if they have the ability to control peak distributions of traffic. This control, and hence cost-effective multi-homing, can be achieved while providing private infrastructure quality on the Internet – by implementing a premise-based route-control solution.



## Premise-Based Optimization Solutions

Typically, an enterprise is considered to be multi-homed to the Internet when it purchases multiple access circuits to a variety of ISPs. Once multi-homed, the enterprise establishes a BGP connection to each service provider and receives default routing information, a full routing table or a subset of the ISP's routing tables (commonly customer and internal routes sometimes accompanied by a default route). Policies are then applied to the various routes received in order to assert traffic engineering preferences and to ensure fault tolerance in case of a failure of one or more of the ISP connections or related routing equipment.

Typically, initial setup is not the last time a network engineer will have to deal with BGP routing policies. For most companies, traffic engineering and BGP route management are ongoing tasks, as outages, increases in traffic, changes in traffic distribution and customer complaints cause BGP configurations to require substantial care and feeding.

Premise-based optimization solutions are designed to automate the 'care and feeding' of BGP management and route selection based on:

- Real-time traffic and quality characteristics;
- Sophisticated business semantics that allow the network engineer to configure the network as a system rather than as individual pieces each effecting individual concerns.

Premise-based optimization solutions accomplish the task of managing network resources in the context of policies. Once policies have been established, the solutions manage network resources by taking over the 'grunt' work of measurement and modification of the routing table.

## Route Optimization Components

The three basic components of a premise-based optimization solution are measurement modules, a control module and visualization/reporting modules.

### Measurement Modules

In order to characterize the network connections and the associated metrics that are relevant to the enterprise, measurement modules must gather a variety of data including circuit utilization, utilization by prefix, frequency distribution by prefix and loss and latency by prefix.

Circuit utilization is important to the solution in order to manage an overall context of traffic across each circuit, especially as it relates to cost. Utilization by destination prefix is a crucial metric in order to understand the effects of shifting traffic from one provider to another.

Frequency and utilization by destination prefix is the key to identifying only those prefixes that need to be monitored out of the thousands that may exist in Internet routing tables. Frequency is a particularly important metric in terms of scaling a premise-based optimization solution. The current routing table contains approximately 250,000 destination prefixes accounting for about one billion individual hosts. By considering frequency, only the



destinations that are relevant to the enterprise need be accounted for; hence easing the operational burden of the solution.

Loss and latency are the primary metrics necessary to make quality assessments of a given network path. It is necessary to measure both of these metrics in order to satisfy the individual and unique requirements of an enterprise and the applications in use.

### **Control Module**

The control module of a premise-based optimization solution does not reside in the actual forwarding paths of the managed network. This is important in that integration with current forwarding hardware and the state maintenance necessary in complex network environments would require the invention and/or ratification of an entirely new 'enterprise edge' or premise-based route control routing protocol. Therefore, it is desirable that the control element operates as a signaling mechanism communicating via standard BGP formatted messages. In addition, it is important that the BGP speaker interoperates with standard BGP implementations, and adheres to all ratified Internet Engineering Task Force (IETF) BGP attributes, extensions and message types. The control element must also take into account operational considerations such as dampening mechanisms, prefix length controls and appropriate debug outputs.

In addition, the policy and configuration semantics on which the control decisions of a premise-based optimization solution are based must be robust. They must allow the integration of network characteristics such as loss, latency and utilization, along with business criteria such as circuit cost. The policy grammar must also be presented in a simple manner to encourage ease of operation and maintenance.

### **Reporting Modules**

A premise-based optimization solution also needs the means to validate and visualize network effects. Utilization reports, prefix distributions, and performance graphics and data can be used not only to validate the functionality of the solution, but also as a means of SLA enforcement, bandwidth billing validation, capacity planning and further policy refinement.

### **Flow Control Platform**

The Internap premise-based optimization solution is called the Flow Control Platform (FCP) solution. The FCP solution is based on flexible software modules that may be integrated to fit an individual enterprise's needs. The basic building blocks of the platform are the FlowCollector™ and FlowDirector™ processors and the FlowView™ reporting tool. The FlowCollector processor measures the distribution, volume and quality levels of traffic managed by the FCP solution. The FlowDirector processor optimizes routing of traffic and the application subsets based on user-defined policies, and is comprised of configurable policy inputs, interfaces to the FlowCollector processor, algorithms for performance and cost constrained optimizations, and a BGP daemon (BGPd) that conveys intelligent route control decisions to the routing system being optimized. The FlowView reporting tool is the representation component of the solution, and provides a graphical reporting interface for displaying the performance, usage and route change activity information generated by the various FCP components.



## Determining Prefixes

The FCP solution's first job is to determine which destination prefixes are relevant to the installation. This is accomplished automatically by software modules or by manual configuration.

The Passive Flow Analyzer™ (PFA) tool, one of the most powerful and innovative components of the FCP solution, can be used to identify important destination prefixes. This tool is a passive tapping technology that is used to analyze IP and TCP header information from an Ethernet segment in order to analyze flow distributions. The PFA tool is also able to measure quality metrics and bandwidth usage associated with individual and aggregate flows (functionality covered later in this white paper), which can be an important criteria used to select specific destination prefixes for optimization.

Somewhat less powerful, the Netflow Analyzer (NFA) requires Cisco Netflow or sflow (compatible) exports information to analyze traffic distributions. Like the PFA tool, the NFA also dynamically identifies prefixes in the context of a policy or policies. NFA stats provide prefix and volume information but do not provide any passive quality metrics.

Another method of identifying relevant prefixes in situations where destinations are known is to manually configure static prefix lists. Each prefix list may be referenced by policies that allow for multiple quality or cost-based parameters, which can exist in tandem.

The FCP solution learns about prefixes in any combination of these three methods. PFA is the preferred method, because it is real-time nature and it can provide prefix quality metrics. Per prefix bandwidth utilization measurements are coupled with aggregate link utilization measurements gathered by the FlowDirector processor's usage collector in order to rationalize bandwidth usage and distribution across ISP uplinks. Per prefix and per ISP volume information is used for cost-based policy decisions.

## Measuring Quality

To measure quality, the Flow Control Platform solution employs both active and passive network measurement technologies. Passive measurement is performed by the FlowCollector processor (using the Passive Flow Analyzer tool). Because the FlowCollector processor captures Layer 3 and Layer 4 header information for all flows, it is able to reconstruct the flows and record performance metrics based on the traffic's TCP information.

While the FlowCollector processor's PFA tool is extremely accurate, flexible and scalable, the solution must also gather information periodically as to the viability of alternate paths. Since routing decisions (specifically with BGP) are constrained to a single next hop per prefix, the PFA only has visibility into the single, active path for any given prefix. To discover the network quality metrics of alternative uplinks, the FlowDirector processor can send active probes across the alternative uplinks when the PFA tool detects a quality-based policy violation on the current active path for a given prefix.

The active probes, based on User Datagram Protocol (UDP), are designed to be extremely lightweight and accurate in its measurements. In some cases, active probes are a desirable augmentation to passive technologies, as they eliminate lags and loss caused by DNS, client



software, last-mile bottlenecks or zero-diversity path segments. The administrator does have the option of excluding active probes to a given prefix list, in cases where the far end has administrative or acceptable use policies that prohibit such measurement.

## Policies

Policies are organized into three categories: detection, network quality and business semantics.

The detection category of a policy is used to decide the method and the metrics to be used in determining the prefixes affected by a policy. As noted previously, a configured list may be referenced by the policy or dynamically populated lists may be referenced.

The network quality category allows the administrator to define quality metrics for a given policy, referencing a given prefix list. The quality metrics supported by the FCP solution are **loss and latency**. Recent enhancements to the FCP solution include the **jitter** matrix. This is important for any Real-Time Protocol (RTP) flow such as VoIP or Video. Quality metrics may be defined as absolute values, but for policies in which it would be unreasonable to set absolute values, a dynamic metric may be employed. Dynamically computed metrics use historic measurements to assert a baseline value for a prefix across a path. These quality metrics can also be determined on an application basis. With these options, the administrator can define policies based on the type of traffic generated.

The business semantics policy applies business logic to the solution. Business semantics include billing rounding method (P-95) per uplink, cost per provider (flat rate, tiered or per megabit), and minimum bandwidth committed per provider.

## Construction of Policies

To construct a policy used by the FlowDirector processor, the administrator must perform four steps:

1. Name the policy
2. Configure policy clauses that match prefixes
3. Set policy parameters for the matched prefixes
4. Apply the policy instance

Prefixes can be matched with static prefix-lists or dynamic filter-lists. Filter-lists are similar to extended access-lists that will match against the prefix information generated by PFA and NFA. You can define a filter using any number of lines specifying packet characteristics such as protocol, source, destination, ports, etc.

All policies are named configuration elements and must be applied as a policy instance in order to take effect. The application of a policy instance ensures that policy actions do not take place until the policy configuration is complete.



The FCP solution uses billing methods as well as cost and bandwidth commitments to reconstruct actual and hypothetical bills to minimize cost or to meet the administrator's cost objectives.

### **Route Updates**

The algorithms employed by the FCP solution accept policies and metrics from the FlowDirector and the FlowCollector processor. The FlowDirector processor then performs the policy-constrained optimizations necessary to deliver routing updates that best meet the policies applied.

Routing updates are communicated to the routing solution via BGP update messages. The BGP update messages are only sent to internally configured peers and contain third-party next hops that direct outbound traffic to the appropriate upstream ISP. The optimized NLRI uses a user-defined high local-preference value to ensure route installation. The BGPd component also periodically withdraws updated routes when they are no longer necessary, ensuring that the optimized routing table stays as small as possible.

### **Configuration**

The configuration component of the FCP solution provides access to all configurable elements of the solution including prefix lists, policy templates and applications traffic breakdown and management configuration options. Two interfaces are provided: a GUI (called the FCP Manager) and a sophisticated Command Line Interface (CLI). The full configuration can be displayed and copied from the CLI for offline archiving.

### **Management**

The management component of the FCP Manager is tightly integrated with the FlowDirector and the FlowCollector processors, allowing it to provide the system health and operational details of the FCP solution via an easily accessible and secure GUI. System health parameters include CPU, memory and component process details. Operational details gathered by the FCP Manager include all processes system messages, as well as logs of actions taken by each component. In addition, information on system health and operational details can be accessed via SNMP and SYSLOG for use by third-party management tools.

### **Reporting**

The FlowView reporting tool provides graphical and archived data for performance, usage and bandwidth cost.

### **Summary**

The Internap Flow Control Platform (FCP) solution can make the 'business grade' Internet a reality by combining state-of-the-art measurement technology (the FlowCollector processor), customer defined policy-driven route optimization algorithms (the FlowDirector processor) and advanced reporting, planning and data storage modules (the FlowView tool). By multi-homing with a premise-based optimization product, the ubiquity and diversity of the Internet can be leveraged to its fullest potential as a business communications platform.

