# Route Control – Optimize Multi-homed Connections for Performance, Load and Cost

By John Bartlett
January 2002

The Internet is coming of age, in large part because of its ability to open up markets and to make them more transparent. If I want to buy a new microwave, I can click on MySimon (www.mysimon.com) and instantly check dozens of Internet vendors, and get a list of microwaves organized by price. The natural result of this process is that most Internet based offerings of the same product come in at virtually the exact same price. Internet service providers (ISPs) are about to encounter similar pressures. Service providers like InterNAP constantly monitor ISP performance and availability. They report back to their enterprise customers on a regular basis, and can switch traffic to the best available service. A new breed of route managers has also appeared, whose job is to help the enterprise with multiple Internet connections (multi-homed) optimize the use of their Internet connections.

## So what is the problem?

Actually, there is not so much a problem as there is an opportunity. Large enterprise customers, especially those who need constant presence on the web, are connecting to multiple carriers at each of their sites.

The most immediate benefit this provides is high availability. With a redundant Internet connection, the temporary failure of one service provider does not cut the Web connection. But like many other redundancy situations we don't want the second link standing idle while it waits for the occasional ISP failure. Instead we want to use that bandwidth, and experience 'graceful degradation' when one link goes down. In the mean time and, importantly, most of the time, we can put that bandwidth to work.

Bandwidth Sharing is the first logical step. If I can manage my traffic so that I balance the load between my links, no link becomes over utilized. This can improve overall performance by minimizing peak loading, but more importantly it may help reduce cost by keeping traffic within contracted traffic agreements. With many contracts using "average" and "burst rate" charges, shifting traffic among ISPs to stay below surcharges is important. Management of traffic against bandwidth contracts based on traffic types, time of day and/or demand (load) provides real value, and can demonstrate a positive ROI for many businesses.

***What about Performance?***

Does it make sense to route traffic through one carrier or another to obtain better overall network performance? Users would certainly say yes.
Performance directly affects the user's perceived experience, which can positively affect a host of marketing issues – customer retention, return rate and conversion rate (see NF Report 5055 "Understanding Web Performance"). A pack of new product vendors, discussed below, believe they can find the better performing route, and deliver it dynamically.

Performance problems can be broken down into first mile, middle mile and last mile issues as shown in Figure 1. Most folks are familiar with the last mile, where bandwidth is often a constraint. The last mile includes the access link and the ISP serving the client,

**NetForecast Report 5057**

This report was first published as an article in *Business Communications Review*, January 2002. See www.bcr.com.

and this part of the connection has only one path choice – e.g. one ISP and one access link.

The middle mile consists of one or more carrier networks that move traffic from the access ISP to the "final" ISP. This is where path diversity takes place; indeed traffic typically takes two different paths through this middle mile – traffic flowing from the client to the content provider takes one route, while traffic from the content provider back to the client takes a different path. This is due to the carrier's policies (e.g. hot-potato routing) or to the decisions the Border Gateway Protocol (BGP) makes from each perspective. The middle mile is the section of the link where path diversity is possible.

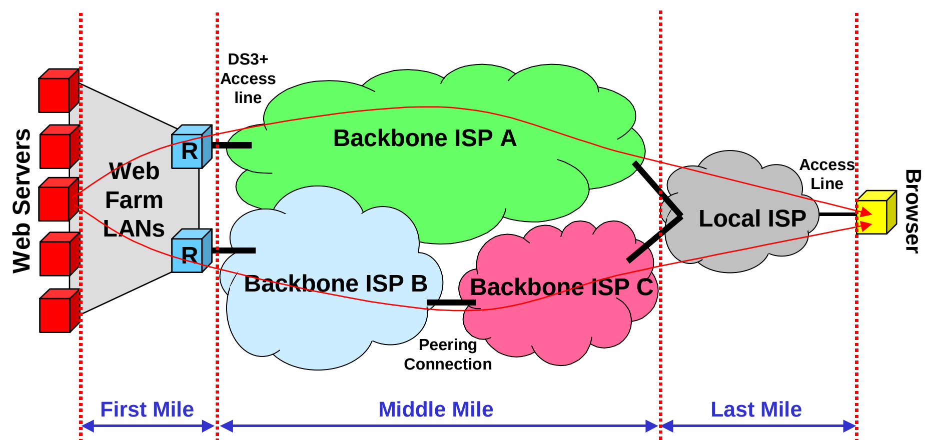


**Figure 1 - Multihomed Network Diagram**

The first mile consists of the internal network of a content provider, and the access ISP that connects the content provider to its carrier. With multihoming, the content provider "shortens" up the first mile and connects directly into the core diversity. The route manager can now help make routing decisions into this core that make business sense for the content provider.

What does Internet performance look like? Well, it is problematic. Remember, there is no master plan for the Internet; instead it is an aggregation of separate carriers and ISPs, connected in whatever way they choose. Network providers have a great deal of incentive to make their own networks work well, but no one is worrying about making the combination of networks perform as a system.

Congestion occurs in this network of networks due to natural peaks of traffic, system perturbations such as large scale demand (e.g. news events) and failures of nodes or links. A consistently fragile area is the peering points, the places where Internet providers connect to each other. The current economic model does not encourage carriers to worry about keeping peering points uncongested, even through most Internet packets traverse at least one peering point to reach their destination.

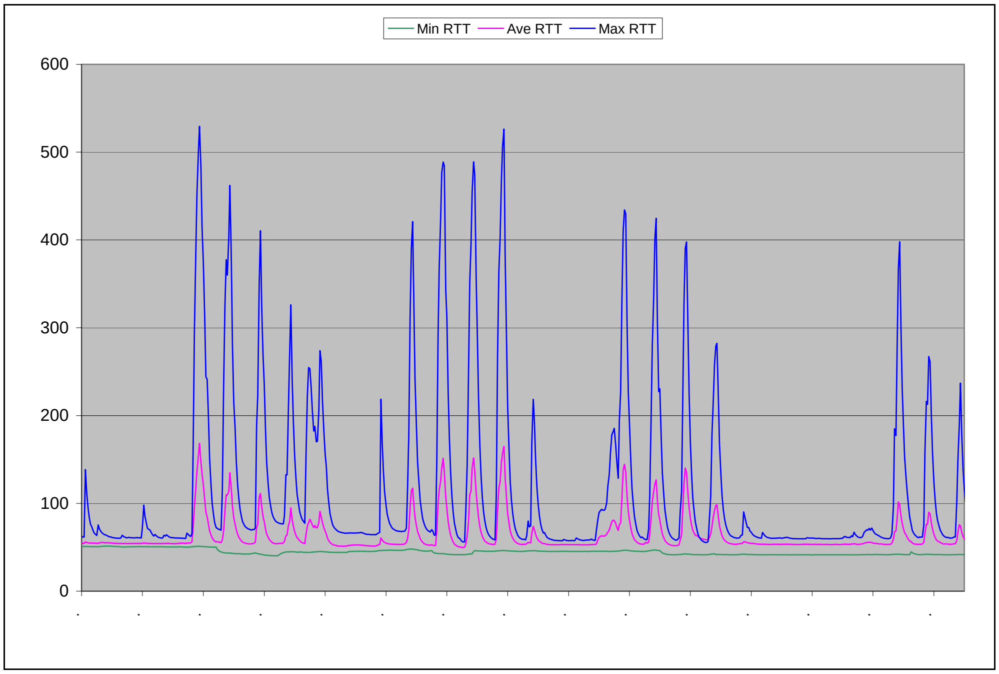


**Figure 2 - Latency to General Motors (data courtesy of NetVmg)**

Figure 2 shows the backbone performance from a Network Access Point (NAP) on the west coast to the General Motors web site. The data in the figure comes from tests run simultaneously with 7 different carriers over a 1-month period. Test probes similar to a traceroute were sent over each carrier towards the destination site at five-minute intervals.

The graph shows three composites; the red line represents the worst time of any provider, the blue line shows the mean time, and the green line shows the best time. Note that none of these lines represents one carrier (yes, we all want the green one!), but instead represents, for each point in time, the worst, best or mean time.

The graph shows a couple of things: First, there are major slow-downs, with latencies moving from 60 ms up to 400 ms or 500 ms for an extended period of time. The green line, however, shows that if we could *dynamically* choose between these seven providers, we could always find a well performing path.

Figure 3 gives us a different look at the test data. This chart shows the best carrier choice for a specific destination, for the duration of the test period, as a percentage of the test duration.

There are two interesting results in this picture. First, the best path to one destination is not the same as the best path to other destinations; clearly routing must be done on a per destination basis. Second, no carrier comes out as consistently the best for all destinations, or consistently as the best carrier even for one location. Specific carriers dominate the best choice for individual destinations, but even these cannot keep it up all the time.

***Routing Decisions Today – BGP4***

Don’t the routing protocols determine how traffic will be routed through the Internet? Well, of course, we have BGP4, which is responsible for determining to which next carrier each packet will be delivered. BGP4 propagates reachability information between carriers to indicate how well it can reach each destination. The protocol then consolidates this information, and makes a forwarding decision based on the ‘best route’ to each destination as viewed from each BGP router.

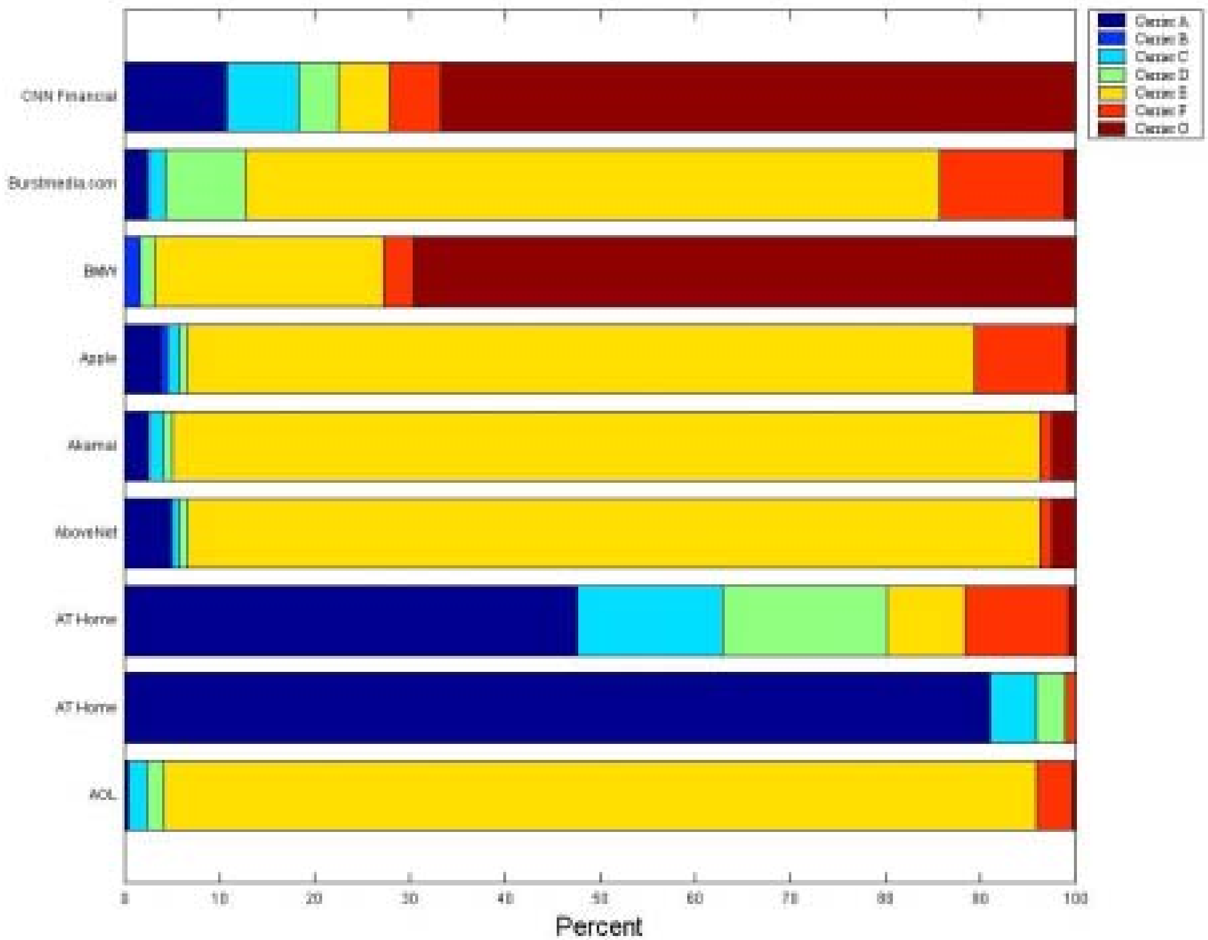


**Figure 3 - Best Carrier Percentages**

Unfortunately, BGP4’s definition of ‘best’ is usually limited to knowing how many Autonomous Systems (ASs) or carriers lie between the source and destination. A path with only 2 AS hops is preferred over one with three or more AS hops. In the case where more than one route is available with the same number of hops, BGP4 makes its decision based on which router has the lowest assigned BGP identifier number. The Internet used to be much more strung out, causing connections to cross three or four carriers on their path from source to destination. As the size of backbone carriers increases, and as their numbers decrease, the Internet is becoming denser. Fewer carrier hops are now required, because carriers cover a much broader expanse of the network.

The BGP4 decision yields a path that works, but not one that necessarily best suits the enterprise. No consideration of link loading, link cost, round trip delay or packet loss is included in this decision. These parameters may directly affect the ability of the business to maintain customers or to minimize operational costs.
So, given this complex and dynamic performance environment, how can I get the best connection for my business? This is the job of the new breed of route controllers.

## How Route Controllers Work

This new intelligent route management boxes (or services) monitor the performance from your site to other sites in the Internet. Then, these units have a BGP conversation with the enterprise routers that are connected to the service provider links, and indicate, either by destination AS number or by route prefix, which carrier to use. With their updated forwarding tables, the routers then direct traffic onto one link or another to obtain optimal results.

Optimal, of course, is in the eye of the beholder. So the vendors of these products provide lots of knobs to manage the behavior in ways that better benefit your business. Time of day is one knob, allowing certain configurations during one part of the day, and different ones later. Cost configurations are another knob. Tell the system the relative cost of your links, and the thresholds involved with changes in cost, and the route manager will deliver a minimum threshold of performance and will then favor your lower cost links, reducing overall cost of operations.

Frequency of updates also often can be managed. Most route control vendors recommend that the BGP tables in the edge router not be changed any more frequently than every five to 15 minutes. Longer periods are used by many of the solutions. Thresholds for the route change operation are another knob. Changing routes too often can cause instability, so options are available – for example, the ability to set a policy like, “Performance on an alternate carrier must be 10% better than on the current carrier before switching.”

Note that route controllers only affect the outgoing traffic link decision; they do not try to propagate information into the Internet BGP tables, only into the local table of the enterprise edge router. Traffic coming to the enterprise from a remote client continues to use the path chosen for it by routers running BGP on the reverse path. This solution applies best to content providers, where the preponderance of traffic is in the outward direction, and is subject to route control. This solution would not work as well for a disk-backup provider, for example, where most of the traffic is flowing inward, over uncontrolled routes.

### *Mapping Internet Performance*

How these solutions determine current Internet performance is an area in which there is much diversity among the vendors. Each has a unique approach for determining the current Internet performance, and deciding on an alternate route. That said, however, there are two basic ways of getting raw data: passive monitoring and active probing. There also are two approaches to assimilating the data: local knowledge and centralized knowledge.

The first order of business is to test the performance of each carrier link to a specific user, or group of users in the Internet. This is done by on-site hardware at the point where multiple carrier connections come together in the enterprise. Passive testing consists of watching the traffic that is initiated by remote clients coming in across the links. Measuring TCP open time, TCP ACK times, and TCP retries will yield latency and packet loss information for those connections. These measurements will yield performance values for one carrier, the carrier that BGP has chosen as the best route.

***Central Knowledge vs. The View From Here***

**The view from here:**
A route controller that determines network performance from one point of view on the network.
*Advantage*: Only needs to compute paths sourced at this location, only needs to compute paths of interest to this enterprise. Does not care which carriers are causing congestion, it makes decisions based on measured performance. It's a standalone solution, not dependent on a central resource.
*Disadvantage*: Routes of interest must be specified or learned dynamically. In a dynamic environment, newly active routes will not be optimal until new probes can be launched and network performance analyzed for this new link.

**Central Knowledge:**
A route controller service that collects information from many sources and builds a global view of Internet connectivity and performance.
*Advantage*: Global view provides routing for connections even before they are required by edge systems. Constant monitoring of congestion points allows predictive capability about when congestion will occur or clear up. Dynamic database of information provides a platform for other interesting service offerings. This may be a more accurate approach under some circumstances.
*Disadvantage*: Global view requires large capital investment, or subscription to a service. Central service is subject to single-point-of-failure due to technical or business reasons, and requires a monthly subscription fee.

Active testing involves sending ping-like probes from the local site towards the remote client. These probes can be sent over each carrier network to determine their relative performance. Vendors have improved on the basic ping to give them better results and to avoid looking like a denial of service (DOS) attack.

Once this data is obtained, there are two choices on how to deploy this information. The vendors who provide a standalone self-contained system assimilate this data, and make a routing decision based on the view from that site. If Carrier B has 20% better performance to client 1 than carrier A, the system can choose carrier B. This routing preference is then sent, via a BGP session, to the enterprise edge router, where forwarding tables get appropriately updated. Further traffic to that client is then passed out over Carrier B, until some future performance change updates this decision.

The second approach, used by vendors providing a service rather than a stand-alone box, is to send performance information to a central location where a composite map of traffic jams on the Internet is created. Think of the morning traffic report on TV, where they show the slow roads and intersections in pulsing red. This composite map can be used to determine the best routes from anywhere to anywhere else on the Internet. A subset of these 'best routes' is then sent to the client sites, and used by the local controller to help determine the optimal path.

The stand-alone approach has the advantage of not needing the centralized computing resource. Incoming traffic defines the routes that are of interest to this site. Active testing can be done on the alternate carriers to determine if their routes give better performance, and routing decisions are made.

Given that it can make routing decisions for all the traffic that site experiences, what advantage does the centralized resource provide? The primary difference is proactive vs. reactive routing.

The localized approach is typically only probing routes of current interest, e.g. routes being used by active users. If a new user arrives at the site from a corner of the Internet that is not being monitored, his or her path to the site may not be optimal until the local probing and performance algorithms have an opportunity to scope out this new set of paths. A central resource, on the other hand, can update routes preemptively, insuring that traffic to this new user is returned over the optimal path even on the first access. Because the central resource collects information from many client locations around the network simultaneously, it has seen the congestion point and routed around it even before the new user comes on line.

***Reporting Results***

Each of these route managers provides feedback on the performance they measured and the decisions they made. This information is a direct reflection of the quality of the service being provided to the enterprise by the contracted Internet carriers.

In Quality school we learned that shortening the feedback loop improves quality. Share this data with the carriers on a regular basis, especially when renewing contracts, and you can expect their quality to improve. This information may provide the financial incentives carriers need to improve peering point performance bottlenecks.

***Who are the Players, and How do They Compare?***

I spoke with six vendors of route control solutions in the process of writing this article; they are listed in Table 1. Three of the vendors – NetVmg, Proficient and Route Science – are providing a stand-alone appliance that installs at the enterprise site, outside the firewall. These units need to have an active BGP session to each edge router, and access to link loading statistics, usually through SNMP reads of the router MIB. Some units also require a mirror port to passively monitor the traffic crossing the carrier links.

Two other vendors use the centralized approach – Sockeye and Opnix. These folks also provide an appliance, which collects information locally and converses with the routers. These units also converse with a central congestion monitoring facility, where the heavy computation is done on the whole Internet picture, and specific routes are then sent back to each box.

One vendor – Radware – does not play the BGP game, but instead focuses on the local links, their utilization and loss characteristics, and optimizes use of the links based on that information. Radware units sit in-line with the data so they are able to passively or actively monitor performance, and can add other benefits like routing based on traffic classification.

**Table 1 - Route Control Vendors**

| Vendor | Prod/ Serv. | Measurement Method | Route Determination | Controls |
|---|---|---|---|---|
| NetVmg | Product | Active & Passive | Local | Utilization, Latency, Loss, Cost, Time of day |
| Opnix | Both | Active | Central | Utilization, Latency, Loss, Cost, Time of day |
| Proficient | Product | Active | Local | Utilization, Latency, Loss, Cost, Time of day |
| Radware | Product | Active & Passive | Local | Utilization, Latency, Loss, Cost, Time of day, Traffic type |
| Route Science | Product | Active & Passive | Local | Utilization, Latency, Loss, Cost, Time of day |
| Sockeye | Service | Active & Passive | Central | Utilization, Latency, Loss, Cost |

Opnix, Sockeye and Proficient cause the least impact to the enterprise network, because they do not require passive monitoring of local traffic. Opnix and Sockeye have a global view of the Internet's performance, so they will deliver high performance routes for addresses that have not yet been used. Opnix creates a central database using the information from each enterprise-installed controller. Sockeye gets its data from Akamai, which is already creating this performance map for use with its content distribution network.

## The Route Management Forecast

Is there a place for route control in the network today? Yes, based on the graphs in this article. Managing for business results makes sense for any enterprise; so if your enterprise is large enough to support multi-homing, take a close look at route control.
Is there a future for route control? Again the answer is yes. Remember that the network is not centrally managed, so there will always be issues. Carriers will optimize their networks in different ways, and the aggregation of those methodologies will always create bumps in the road.
However, remember that this approach is not a definitive answer to Quality of Service (QoS). To use this approach to support services requiring high QoS (e.g. Voice and video) is playing the odds. This approach does not guarantee QoS, it just improves your chances.

The introduction of route control provides strong feedback to the business managers who negotiate with carriers for service. Information on monthly performance, with direct comparisons between vendors may make carriers more competitive on quality, and may give carriers a way to add value. Results would benefit both the Internet and the businesses that use it.

Will there be a place for these route control companies in the milieu of high-tech solutions? If the demand for these services is real, than the service and products will survive. Route control is a compute-intensive application, and it requires monitoring of the traffic for specific performance parameters. Neither of these jobs is easily taken over by the router.

Lastly, there may be significant value yet to be mined in the centralized Internet connectivity and congestion map. Look to the centralized service providers to introduce additional features or capabilities over time. ¤

## Companies Mentioned

| | |
|---|---|
| Internap | www.internap.com |
| NetVmg | www.netvmg.com |
| Opnix | www.opnix.com |
| Proficient | www.proficientnetworks.com |
| Radware | www.radware.com |
| Route Science | www.routescience.com |
| Sockeye | www.sockeye.com |

NetForecast develops customized analytic models to determine the business value of new technologies.

www.netforecast.com

**John Bartlett** is Vice President of NetForecast, and has 24 years of experience designing at the chip, board, system and network levels, with a focus on performance. John led the team that built the first VLAN implementation, one of the first ATM switches, and he is a leading authority on the behavior of real-time traffic on the Internet. He can be reached at john@netforecast.com.