Application Layer vs. TCP Layer WAN Optimization
Executive Summary

When planning the capacity of a WAN optimization solution, it’s best to base calculations on a WAN optimization appliance’s ability to perform application-layer functions, such as managing user counts and protocol optimization throughput, rather than on low-level metrics such as TCP counts. WAN optimization appliances function as WAN proxies, changing network traffic to accelerate it and secure it. These changes can affect TCP counts in unpredictable ways, making TCP counts a poor choice of a metric for capacity planning. A better approach is to determine how many application sessions a WAN needs to support and then to design the WAN solution accordingly. An effective WAN acceleration solution will make optimal use of lower-level TCP functions while also scaling its higher-level application and session-based functions to meet the growing demands of users.
Choosing the Right Metrics

Consider this scenario. A financial services firm has grown rapidly over the past decade, and now it’s time to upgrade the network. The goal is to increase the network’s capacity to accommodate new employees and branch offices. The upgrade will include a new WAN optimization solution with WAN concentrators deployed at core data centers and at branch offices.

To plan the capacity of the WAN solution, the company’s Network Operations Center (NOC) engineers determine how many TCP connections their employees are using in their daily work. To give themselves a margin of error, the engineers increase this number by 20%, then multiply it by the projected number of employees the company will have in 3 years. Then they buy and deploy their new WAN optimization appliances.

But there’s trouble right away. Network performance is sluggish everywhere. Even though it was designed to accommodate a much larger organization, the WAN optimization solution is already maxed out and not delivering the promised performance; in fact, it’s slowing down applications. What went wrong?

The NOC team made a critical error in scoping their WAN optimization solution. While it’s perfectly reasonable to use TCP connection counts as a capacity planning metric for routers and other internal network devices, it’s inappropriate for scoping the capacity and design of WAN optimization appliances. Why? Because WAN optimization appliances are essentially proxies, intercepting LAN application traffic and changing it – for the better – on the WAN. And the effect of proxy services on TCP count is unpredictable. Improving applications can increase, or decrease, the TCP connection load in ways that defy simple back-of-the-envelope calculations.
A Closer Look at WAN Optimization

Effective WAN optimization solutions manipulate traffic in a variety of ways in order to deliver applications and data quickly and securely to users throughout the enterprise. Authentication, byte-caching, compression, protocol optimization, policy enforcement, and other proxy activities all have the potential to change the number of TCP connections active on a WAN. Exactly how these activities change TCP counts varies, and it varies so much that TCP counts turn out to be a poor metric for planning WAN optimization capacity.

For example, consider the effect of protocol optimization on TCP connections. A powerful technology, protocol optimization aligns high-level protocol behavior with low-level network realities. However, the "optimal" way to transmit data could involve more or fewer TCP connections, depending on the protocol and the circumstances. For example, a user Web request going across an optimized WAN might have two to five TCP connections on the LAN side, one connection on the WAN between appliances, then balloon to 50 or more connections on the far end, as the appliance leverages parallelization to improve performance. But protocol optimization can also do the reverse, as when a proxy aggregates CIFS file service requests, thereby reducing overall TCP count.

TCP counts are a misleading metric for capacity planning. They're distracting, too. Protocol optimization, authentication, and other proxy-like activities tax any WAN appliance far more than holding open TCP connections do. How well a WAN optimization appliance performs these acceleration and security functions for users will ultimately determine the capacity of the WAN solution. By overlooking these higher level functions and instead focusing on TCP counts, NOC engineers risk deploying an appliance without the horsepower to do its job well, however many TCP connections it can hold open.
The Best Practice for WAN Optimization Capacity Planning

When capacity planning a WAN optimization solution, it’s best to base calculations and scoping requirements on user sessions and application throughput. WAN optimization solutions need to optimize the requisite number of user sessions, regardless of how many TCP connections result through parallelization and other optimization techniques. Whether a single session, such as a user running a Web application, results in 50 TCP connections or only 5, the appliance will still have plenty of capacity for managing TCP connections, as long as it has enough horsepower to perform these other higher-level optimizations.

The goal, ultimately, is to deliver low-latency, network efficient and policy-compliant, applications to users everywhere. To avoid network bottlenecks and other unpleasant surprises, focus on user sessions and upper layer metrics, rather than lower-layer metrics that ultimately serve only to implement the higher-layer, optimized WAN solution.

Running a pilot project and monitoring how well a WAN optimization appliance manages its higher-layer functions is a good way to ensure the broad WAN optimization solution will have adequate capacity when it is deployed.

The Blue Coat Solution for WAN Optimization

Blue Coat is a leading provider of WAN optimization solutions that accelerate business applications across the distributed enterprise. Blue Coat’s family of appliances and client-based solutions – deployed in branch offices, Internet gateways, end points, and data centers – provide intelligent points of policy-based control enabling IT organizations to optimize security and accelerate performance for all users and applications.

Blue Coat appliances use a proxy/cache architecture that is user- and application-aware. By analyzing higher-level application functions, such as authenticated user sessions, and making use of caching and session- and application-layer techniques, Blue Coat ProxySG appliances can optimize application delivery far more effectively than WAN solutions whose designs treat WAN optimization as a networking problem confined to the packet delivery layers of the OSI model (layers 2-4). As a result, in real-world deployments, Blue Coat WAN optimization appliances consistently deliver faster performance and greater scalability than other optimization products.
Conclusion

When comparing WAN optimization architectures, it's important not to get sidetracked counting packets and TCP connections. Focus instead on delivering applications quickly. That means looking where users, applications and their data live: at layer 7. Measure WAN solutions by their ability to deliver accelerated applications, not packets, and you’ll achieve your WAN optimization goals.