Improving Performance for Mission-Critical Applications with QoSWorks

Solving Performance Problems at the Network Edge for ERP

QoSWORKS™

Sitara QoSWorks. Peformance. Control. Precision. Power.

Using QoSWorks™ to Manage ERP Applications

Business requirement:	An enterprise has deployed an ERP solution that consists of a mix of mission-critical transaction-oriented applications and back-end bulk transfer applications. The ERP applications are accessed by 70 branch offices around the country. The network also carries file transfers, e-mail, and Web traffic.
Network topology:	Each branch is linked to headquarters via a Frame Relay network with port speeds of 64Kbps. The permanent virtual circuit committed information rate (CIR) for each location is 50% of the port speed, or 32Kbps.
Application requirements:	The ERP application must contend for network resources with the file transfers, e-mail and Web access. If the network is congested and a record in transit is not fully received, the ERP application will hold the record open in a blocked state until the situation is resolved; an operator must manually unblock the record, which consumes both time and resources. Delays can also lock up an ERP database and cause time-outs and frustrating messages such as: "User disconnected — unknown error."

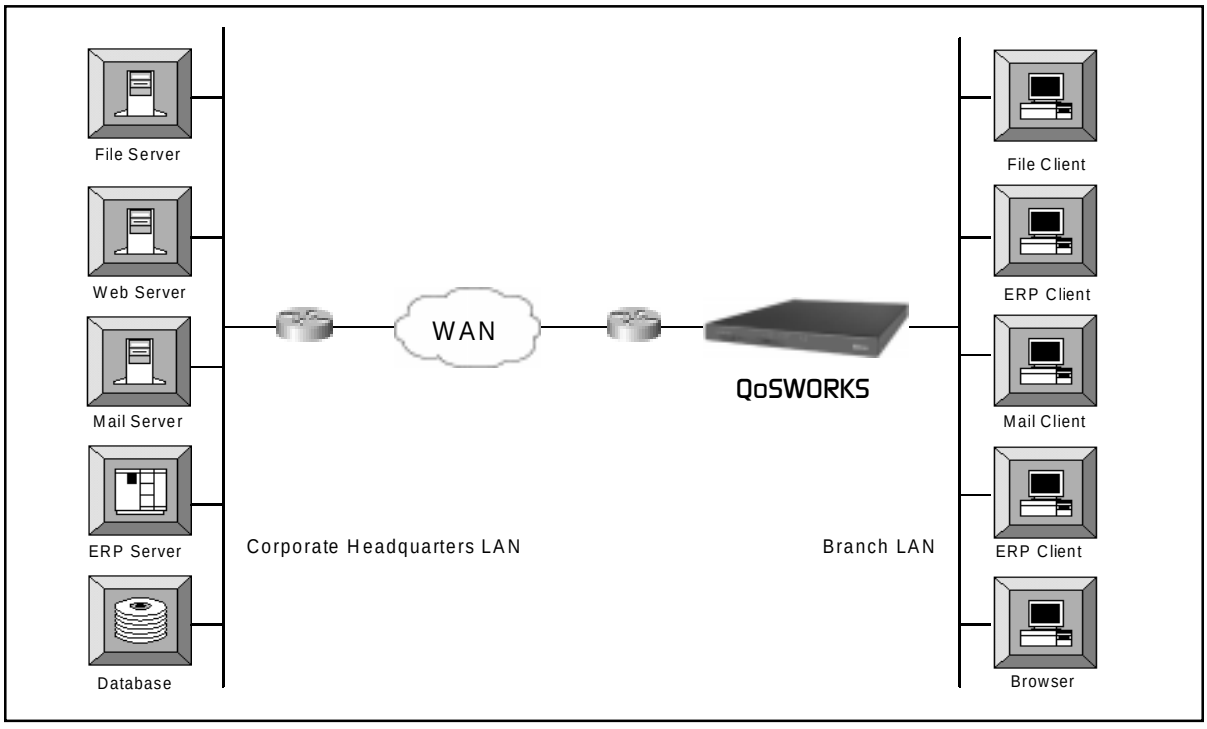

Figure 1. Deploy QoSWorks behind branch office routers to ensure improved, consistent response times for ERP applications.

The QoSWorks Solution

In order for the ERP solution to function effectively, the network manager must control the latency experienced by each ERP application at different times of the day and ensure that small packets from the order-entry application do not get crowded out by large e-mail, Web, and FTP packets. Only QoSWorks integrates all of the different QoS mechanisms necessary to enable the network manager to allocate bandwidth based on the relative priorities of the mission-critical and back-end ERP applications, while providing sufficient resources for the mail, Web, and FTP traffic and increasing overall network efficiency.

Wire-speed classification:

ERP flows are identified based on the combination of destination IP address and port. In this case, the IP address of the ERP application server would be used to identify all ERP traffic. The destination port number would indicate a specific application running on the application server (e.g., Bill-of-Material processing vs. Order Entry). The e-mail and Web traffic are classified by the destination port number (e.g., 25 for SMTP, and 80 for the Web). FTP file transfers can be classified by the IP address of the server or the server's port number (21).

Class-based queuing:

Class-based queuing (CBQ) provides the means of implementing traffic priority and bandwidth sharing. Using CBQ, the network manager can separate all traffic into classes based on a combination of the source and destination IP addresses and source and destination ports and then assign bandwidth to each class. This ensures that bandwidth is allocated fairly among the classes. The network manager can use QoSWorks to monitor each application over time and adjust bandwidth allocation as needed. CBQ also ensures that when excess bandwidth is available, it is distributed fairly and according to the priority assigned to each traffic class.

Transaction-oriented ERP applications like order entry can be assigned a higher priority than batch processing during business hours, and both can be assigned a higher priority than Web and mail traffic. FTP traffic is assigned the lowest priority. After hours, the policy can be reversed and the batch will get priority over order entry. Setting the priority higher for the transaction-oriented application will ensure that it is given bandwidth in the network first and will not be squeezed out by bulky batch jobs or bursty e-mail, FTP and Web traffic.

CBQ is an excellent method of controlling application latency. However, CBQ alone cannot guarantee that small packets from the order-entry application will never be queued behind large e-mail, FTP or Web packets, leading to delays. In addition, CBQ cannot guarantee that each connection within an application receives a fair share of bandwidth. These require two unique QoSWorks features developed by Sitara: packet-size optimization and fair allocation of bandwidth by connection.

Packet-size optimization: Minimizes delays for the order-entry traffic by reducing the size of e-mail and Web packets and file transfers. For example, it takes an average of 187ms to transmit a 1500-byte packet over a 64Kbps channel. This means that a high-priority packet, such as order-entry packets, would have to wait 187ms before it is sent. Packet-size optimization solves this problem by forcing TCP connections with large packet sizes, like e-mail and Web, to reduce the maximum packet size to about 512 bytes. As a result, a DLSw packet only has to wait 64ms before being transmitted, for a round-trip time of 128ms, an acceptable level of delay.

Fair allocation of bandwidth by connection: With fair allocation of bandwidth by connection, the network manager can further refine the allocation of bandwidth by fairly allocating bandwidth among connections within a class (e.g. multiple ERP order-entry streams).

TCP rate shaping: Rate shaping provides a mechanism to control the volume and rate of traffic being sent into the network by SMTP, FTP and Web servers. By manipulating the window sizes of the SMTP, FTP and Web connections, the TCP rate shaping feature helps manage the burstiness of this traffic to ensure that it does not interfere with the time-sensitive ERP traffic.

Clearly, TCP shaping should be used in concert with other QoS mechanisms such as queuing, as it controls only TCP-based traffic. Also, it cannot ensure a consistent rate for applications where the amount of data sent by the server is small and the rate is low (e.g. transactional applications such as order entry). This is where QoSWorks' CBQ and packet-size optimization features are particularly beneficial.

Caching:

In order for critical applications like ERP to function effectively, network managers must use every byte of available bandwidth as efficiently as possible, especially over low-speed links. This means eliminating redundant requests for Web objects. QoSWorks' caching feature stores frequently accessed Web pages to reduce traffic load on the WAN and free up more bandwidth for the ERP traffic. If an ERP application has a Web-enabled user interface, the static parts of the user screens (e.g., logos, images, text) can be cached as well. Typically, 30-40% of Web content can be cached locally, resulting in significant savings in bandwidth utilization.

For further information contact:
Sitara Networks, Inc.
60 Hickory Dr.
Waltham, MA 02451
www.sitaranetworks.com
(781) 487-5900

For more information on the problem solved by Sitara QoSWorks, please read the Sitara white paper "Taming the Impact of the Internet on Enterprise Networks: Making QoS a Practical Reality."

991202-012-01