RESONATE | APPLICATION NOTE

Scaling Application Performance on Servers Using Port Load Balancing

Resonate Central Dispatch Application Note

RESONATE®

At companies with heavily used Internet or intranet sites, system administrators have the ongoing task of determining how to get the most out of server capacity—without sacrificing end-user service.

Resonate's Central Dispatch™, a software-based server management solution, can help. Using Central Dispatch's ability to load balance traffic across multiple server processes running on the same machine, administrators can maximize the utilization of system resources, and at the same time, maintain a high level of service.

KEY BENEFITS OF PORT LOAD BALANCING

- **Increases application scalability**
- **Enhances server availability**
- **Reduces capital expenditure requirements**
- **Reduces site administration costs**
- **Straight forward to configure and modify**

Increasing Application Scalability

Often, in heavily trafficked sites, the Web or application server process can reach maximal performance while the server hardware still has available operating cycles. The site can benefit from port load balancing. Targeted for high-end servers, port load balancing maps a single incoming port (or service) to a range of ports on a physical server, with each port running a separate instance of the Web/application server process. This more efficiently utilizes both the application processes and server hardware resources, effectively increasing total system capacity.

Enhancing Availability and Serviceability

Central Dispatch also delivers the fastest client responses possible. Incoming requests are intelligently forwarded to the most suitable server based on predefined scheduling policies. Using Resonate's triangular data flow, the server then responds directly back to the client, avoiding potential bottlenecks and maximizing data throughput. When port load balancing is configured, port selection at the server is performed in a round-robin manner; if one port does not respond, another port on the same machine is selected. This ensures that end-user connections go through even if a process hangs, improving server availability. In addition, port load balancing minimizes the potential for performance degradation associated with a single highly utilized server process environment.

Reducing Costs

Central Dispatch enables system administrators to control overall site administration costs and minimize the need for costly system upgrades. Since Web/application server processes typically show a decline in performance before the server hardware does, the former is usually the limiting factor for system capacity. Because port load balancing increases application scalability, administrators of growing sites can maximize the utilization of existing systems, and postpone or eliminate the need for new systems.

Easily Configure and Modify

With Central Dispatch, users can flexibly allocate system resources to accommodate ongoing traffic changes. After creating the additional server processes in the actual server system(s), Central Dispatch users simply specify the range of ports on the server(s) to which incoming requests should be mapped. As traffic demands change, additional servers and/or server processes can be configured to use port load balancing, or resources can be reallocated to different applications or to different sites.

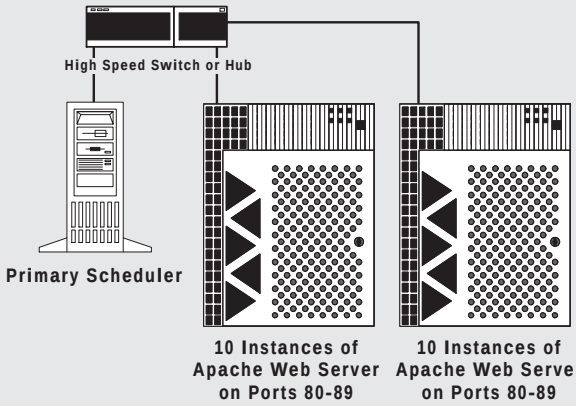

Figure 1—Port Load Balancing to Support Multiple Instances of Web Server Application on Servers

Considerations When Evaluating Alternative Solutions

- Can the system perform port load balancing? e.g., can incoming ports be mapped to a range of ports on a physical server, each running the Web/application process?
- In heavily used sites, how does port load balancing impact: (a) server availability, (b) end-user response times, and (c) system administration costs?
- What is the traffic flow within a site when port load balancing is active (from client request to client response)?
- How does this enable full utilization of server capacity and eliminate potential bottlenecks?

Conditions for Considering Port Load Balancing

- Are there Enterprise class servers where the Web or application server process is at maximal performance but the computer is still operating below full capacity?
- Are there systems with a single highly utilized server process?
- Is there only one server providing content for particular applications?
- Are there any issues with response times for particular applications?

Summary

Central Dispatch's port load balancing functionality offers tremendous application scalability, helping to control capital equipment costs, while ensuring the highest application performance. This solution provides flexibility for the system administrator and network architect, eliminating the common difficulties associated with constrained application scalability.

Resonate, Inc.
385 Moffett Park Drive, Sunnyvale, CA 94089
tel: 408.548.5500 fax: 408.548.5679
www.resonate.com