COMPAQ

White Paper

April 2002
155E-1101D-WWEN
Prepared by:
Industry Standard Server Group

Compaq Computer Corporation

Contents

Compaq TaskSmart C-Series Servers Performance Guide

Abstract: Compaq TaskSmart C-Series servers are based on industry leading Compaq *ProLiant*™ server platforms and the leading content acceleration software suite–Inktomi Traffic Server. The TaskSmart C-Series servers are optimized to accelerate content serving and accesses. This white paper is to provide the reader with a brief summary of the performance results of the TaskSmart C-Series servers. The performance parameters used are also discussed to provide the reader with guidelines in deploying TaskSmart C-Series servers in forward proxy, reverse proxy, and streaming applications.

Introduction

A proxy cache can be deployed as either a client (forward) accelerator or a server (reverse) accelerator.

- Because a forward proxy is located closer to the clients than to the Web or stream server, where the server caches commonly accessed (web and stream) objects, it can serve these objects much faster to the clients than the origin Web or stream server. The upstream network bandwidth utilization is significantly reduced, when most clients are served by a forward proxy instead of a Web or stream server.
- A reverse proxy is located closer to the origin Web or stream server than to the clients. It intercepts and responds to clients' requests, off-loading the Web or stream server TCP connections as well as serving already cached objects.

Regardless of the deployment choice, a proxy cache is located somewhere between clients and server to effectively improve the client access time (also commonly known as latency).

The effectiveness of a cache to improve the client access time depends on several factors:

- Hardware selection and software parameter configurations, which are primarily controlled by the IT department.
- Client browser applications using more HTTP/1.1 persistent connections, which increase server performance by reducing TCP connection processing.
- Fewer no-cache directives (in HTTP headers), which contribute more objects served by the proxy cache.
- Origin Web servers with fewer no-cache HTTP header directives, which allow a proxy cache to cache more objects. Servers with a greater percentage of objects in cache can increase both the effectiveness and the performance of a proxy by increasing the cache hit ratio.
- Problems associated with network architecture and traffic conditions, including cache server placement in relation to the client.
- Access patterns of multiple clients, which can have significant impact on the cache server effective latency.

Because the actual latency depends on these factors, throughput is commonly used as a performance indicator of a cache. Throughput is the amount of traffic that can pass through a proxy cache server, and is commonly represented in rps (requests per second)[1] and Mbps (Million bits per second). Higher throughput can provide lower latency, which in turn translates to faster access times.

Related to TCP performance is the robustness of a TCP implementation inside a proxy cache server. Some vendors optimize the TCP performance to the point that it begins to downgrade their system reliability. Those vendors will have cache that can show attractive performance numbers, but eventually, the system user starts experiencing instability of the system in a heterogeneous networks environment. Compaq uses proven networking stack that complies with all necessary standards in its proxy cache servers to enhance reliability.

[1] Some also use operations per sec (ops), transactions per sec (tps) or objects per sec (obj/sec).

Benchmarking Cache-Performance

The "Performance Summary" section in this white paper presents performance benchmark numbers for several models of the Compaq *TaskSmart*™ C-Series appliance servers. Each benchmark test was run in a condition that best matched a deployment scenario, such as forward/reverse proxy and number of network connections.

Benchmark programs are used to measure the performance of a cache. Commonly used benchmark programs are WebBench[2] and Web Polygraph[3], as well as proprietary programs. Compaq uses the WebBench test suite to establish performance benchmarks for reverse proxy, and Web Polygraph to establish the forward proxy benchmarks. The WebBench data set size is only a few hundreds of million bytes and it will fit within most Web cache proxy servers' memory. Consequently, WebBench test stresses a server for its CPU, memory, and network subsystems. The type and number of disks is irrelevant for a WebBench test. Web Polygraph and its data set have been designed to stress a server for its CPU, network, and disk subsystems. The higher the number of disks, the higher the system throughput can be for Web Polygraph tests, provided that the other subsystems can keep up with the workload.

Product vendors use benchmarked performance numbers in sales tools such as advertising and marketing. They also use benchmark performance numbers to provide useful information for customer use when making purchasing and deployment decisions. Most vendors publish one or two performance numbers to give the customer an idea of a product's position in the market regarding its performance. When making purchase decisions, customers should not rely only on limited benchmarks published by different vendors but also consider benchmarking done in a neutral and consistent environment. An example of a neutral environment is the Cache-Off events[4] organized occasionally by Measurement Factory. When compared to simulated conditions provided by a benchmark program, most proxy cache server traffic patterns present completely different workload conditions.

After design and implementation, performance evaluation can serve to alert support personal to a potential problem that could be caught early. For example, if the MRTG (Multi Router Traffic Grapher) graphs of the cache miss latency show a degrading trend, it can be concluded that bandwidth limits are being reached in the upstream Internet transit connections.

[2] From Ziff Davis. See www.WebBench.com for the details.
[3] From Measurement Factory. See www.Web-Polygraph.org for the details.
[4] Compaq hosted the second and the third Cache-Off events in Houston, Texas in January 17, 2000 and in September 11, 2001. See these Cache-Off events reports at the Measurement Factory website at www.measurement-factory.com.

Product Summary

The Compaq TaskSmart C4000 Model 30 server is best suited for a forward proxy deployment to improve web browsers' access time and to reduce upstream network bandwidth requirements. A forward proxy accesses origin web servers in lieu of the individual client browsers.

The Compaq TaskSmart C4000 Model 40 server is best suited for a reverse proxy deployment to accelerate the origin web servers. Because a reverse proxy is memory-intensive and is not disk-intensive, 1 Gigabyte of memory and two ATA drives are sufficient for most deployments.

The Compaq TaskSmart C4000 Model 50, 60, and 70 servers are best suited for a forward proxy deployment to accelerate web browsers' access time and streaming media applications. Model 50 is suitable when multiples of them are used in a load-balanced environment. In-the-box redundancy features are less important in this case. In addition to a higher number of drives, Models 60 and 70 offer hot-plug drives as well as hot-plug redundant power supplies. Model 70 offers the highest capacity of memory and drives.

By deploying one or more of these TaskSmart C-Series appliance servers, website users will experience faster and better quality access to the Internet.

Table 1. System configuration of C4000 models

	Model 30	Model 40	Model 50	Model 60	Model 70
Usage	Client-side acceleration (forward proxy)	Host-side acceleration (reverse proxy)	Client-side acceleration (streaming)	Client-side acceleration (forward proxy)	Client-side acceleration (streaming)
Software	Inktomi Traffic Server	Inktomi Traffic Server	Inktomi Traffic Server + Media IXT	Inktomi Traffic Server	Inktomi Traffic Server + Media IXT
Processor	(1) 933 MHz Intel PIII	(1) 933 MHz Intel PIII	(1) 933 MHz Intel PIII	(1) 1.13 GHz Intel PIII	(1) 1.13 GHz Intel PIII
RAM	384-MB	1-GB	1-GB	1-GB (Upgradeable to 4-GB)	2-GB (Upgradeable to 4-GB)
Storage	(2) 1-in, 20-GB ATA/100, 7200 rpm	(2) 1-in, 20-GB ATA/100, 7200 rpm	(2) 1-in 18-GB WU3 SCSI, 10K rpm, non-hot-plug	(4) 1-in 18-GB [+(2) 18-GB optional] WU3 SCSI, 10K rpm, hot-plug	(4) 1-in 36-GB [+(2) 36-GB optional] WU3 SCSI, 10K rpm, hot-plug
Disk Cache	~30-GB	~30-GB	~26-GB	~62-GB (4 drives) ~98-GB (6 drives)	~134-GB (4 drives) ~206-GB (6 drives)
Network I/F (See Note)	(2) 10/100 Fast Ethernet	(2) 10/100 Fast Ethernet	(2) 10/100 Fast Ethernet	(2) 10/100 Fast Ethernet (1) 1000SX Gigabit Ethernet	(2) 10/100 Fast Ethernet (1) 1000SX Gigabit Ethernet
Chassis	1U rack-ready	1U rack-ready	1U rack-ready	2U rack-ready	2U rack-ready
Power Supply	(1) 180W	(1) 180W	(1) 180W	(2) 400W Redundant hot-plug	(2) 400W Redundant hot-plug

Note: A gigabit network interface card or the Compaq Remote Insight Lights-Out Edition board (157866-001 NA) may be added as an option in the PCI slot for Models 30, 40 and 50. The Compaq Remote Insight Lights-Out Edition board (157866-001 NA) may be added as an option in one of the PCI slots in Model 60 and 70.

Note: The performance numbers shown here are PRELIMINARY. Compaq is still in the process of tuning various system parameters to improve these numbers.

Performance Summary

Table 2. Web performance parameters achieved by the TaskSmart C-Series appliance servers

	Model 30	Model 40	Model 50	Model 60/70 4 drives	Model 60/70 6 drives
Forward Proxy **(Web Polygraph)**	150 rps (15 Mbps)	150 rps (15 Mbps)	300 rps (30 Mbps)	450 rps (45 Mbps)	700 rps (70 Mbps)
Reverse Proxy **(Web Bench)**		1,825 rps (70 Mbps)	1,825 rps (70 Mbps)	>2,240 rps (>110 Mbps)	>2,240 rps (>110 Mbps)

Notes

All Web Polygraph results reported in this table are independently run and are not associated with The Measurement Factory or Cache-Off events. These performance results are based on PolyMix-3. All Web Polygraph tests reported an average response time under two seconds and document-hit ratio of 55 percent.

Table 3. Streaming Performance parameters achieved by TaskSmart C-Series appliance servers

	Model 50		Model 70, 4 drives		Model 70, 6 drives	
Bit Rate	**# Streams**	**Network Bandwith Utilization (Mbps)**	**# Streams**	**Network Bandwith Utilization (Mbps)**	**# Streams**	**Network Bandwith Utilization (Mbps)**
22 Kbps	4200	88	7000	140	7250	140
37 Kbps	2500	89	4900	170	5000	175
56 Kbps	1600	86	3550	190	3575	180
100 Kbps	900	88	2825	270	2875	265
200 Kbps	470	90	1350	240	1500	290
300 Kbps	310	90	1000	280	850	240
750 Kbps	125	89	240	170	200	130
1.5 Mbps	45	65	100	140	110	150
3.0 Mbps	25	60	55	125	60	130

Notes

- All the above streaming performance numbers were achieved using the Windows Media Load Simulator (WMLS) with zero packet loss reported by the WMLS clients.
- The system throughput on number of streams was higher with a few WMLS client packet losses without perceived video quality degradation, especially for the mid-range bit rates (100 Kbps, 200 Kbps, and 300 Kbps). The higher number of streams varies depending on the bit rate and the packet loss behavior, such as random versus bursts.

Discussions

Note in Table 2 that the forward and reverse proxy throughput numbers are not related, since they were generated using different benchmark programs and payloads. Although Model 30 is intended for forward proxy and Model 40 for reverse proxy, both models may be deployed as forward or reverse proxy.

Forward Proxy Test

A forward proxy serves a relatively small group of clients and these clients can access diverse websites. Therefore, a forward proxy normally caches a wide range of URLs and objects. It is not practical to provide large amount of RAM to have high hit ratio in the memory cache. Instead, providing a number of high performance disks is a more balanced approach in sizing a forward proxy. For this reason, Models 30 and 40 have two optimized ATA drives that are used to provide good economic models without sacrificing performance.

Reverse Proxy Test

A reverse proxy will cache only from one Web site. Knowing the access pattern of a given Web site will help tremendously in selecting an appropriate reverse proxy server. For example, knowing the total storage requirement of a Web site and taking a factor (10 percent) of the most commonly accessed objects, will result in approximate disk capacity in a reverse proxy cache. Taking another factor of the disk capacity will result in approximate memory capacity. For example, assuming the size of all the objects for a Web site is 200 GB and 10 percent of this Web site is accesses 90 percent of the time. Since 20 GB of this Web site is commonly accessed, and should be cached, a proxy cache server with disk cache size of at least 20 GB will be sufficient from a storage perspective. Models 40 and 50 both have 1 GB of memory and they have similar reverse proxy benchmark results because the benchmark load test is more memory-centric.

Streaming Test

For streaming throughput on the Model 50, Model 70 with four drives, and Model 70 with six drives, Table 3 shows a set of numbers as guidelines for the number streams for different bit rates. Here, a stream corresponds to a client user. The actual number of streams will depend on the application, and users tolerances on video and audio qualities. For example, telecasting of a CEO presentation will not require as stringent video quality (in terms of size, resolution, and frame rate) as a high quality movie. How a video is compressed at the source also has significant impact on the streaming throughput. For variable bit rate encoded videos, the overall network bandwidth utilization and the cache server CPU utilization will vary according to the instantaneous bit rate of the video traffic. When the peaks of the video traffic fluctuations are within the network and CPU utilization upper limits, there will be no packet loss. However, when the peaks of the video traffic fluctuations start getting over the network and CPU utilization, packet losses start occurring. The severity on the video and audio quality with respect to packet losses depend on the application, users tolerance, and the packet loss interval and burstiness. For example, users may tolerate some infrequent video quality degradation as long as the audio quality is acceptable (for example, when viewing a CEO presentation). Streaming players also contribute to video and audio qualities perceived by users. Upon packet loss events, some players reduce the video frame rate, and the audio sampling rate necessary in dealing with variable bit rate streams. Others freeze the video frame, until there are virtually no packet errors.

Conclusion

TaskSmart C4000 models provide reliable and well-priced performance choices for customers to accelerate content serving and access. The models can be deployed as forward or reverse proxies depending on customer requirements, with Media IXT streaming capability if needed.

Related Documents

Compaq TaskSmart C-Series Content Acceleration Servers Deployment Guide white paper

Compaq TaskSmart C-Series Streaming Servers Deployment Guide white paper

Compaq TaskSmart C-Series Servers Feature Procedures Guide white paper