Accelerating Web Applications with Cisco Enterprise Content Delivery Networks

Background

Because of the managed and distributed nature of content-delivery networks (CDNs), it is necessary to understand how this infrastructure can tie into application development and deployment. By using the appropriate methods in conjunction with a Cisco Systems Enterprise CDN, current Hypertext Transfer Protocol (HTTP)-based applications can now be dramatically accelerated. For enterprises, this means that many applications, which were previously not deployable because of certain network restrictions, are now feasible.

The combination of standard transparent or reverse proxy caching with managed content distribution provides a system that offers a great deal of control over when content is distributed and which edge nodes, or content engines, should receive content. In many instances, most of the objects that are requested in a Web-based application are fairly small, and can easily be handled by standard transparent cache methods. With little additional overhead, it is even possible to pre-populate those objects onto Cisco Content Engines in advance of users’ requests, providing an additional level of acceleration for the first user’s request for Web content from sites such as Yahoo.com or for any number of Web-based applications.

In addition to the caching acceleration, it is also possible to pre-distribute large objects such as Java applets, Flash animations, Shockwave programs, and any other file format that can be retrieved via a HTTP request. These items can be distributed using a Cisco Content Distribution Manager (CDM), which provides a system for managing when and where content should be distributed, and even allows the administrator to scope how much bandwidth is available and different times of the day or week for distribution. The system also provides built-in streaming media server technology for streaming RealNetworks, Windows Media, QuickTime, and Motion-Picture-Experts-Group (MPEG) media to a client’s system.

By combining caching and content-distribution management, an even more powerful system is created, giving enterprises the best of both worlds: content that does not exist on the system can be retrieved via on-demand pull, and can be integrated with content that has been pre-populated onto the same content engine via the CDM management parameters.

How It Works

Caching

Most of the content from a particular application or Web site is distributed using a simple scheduled pre-population of the content-engine cache. This is done by configuring the content engine to create a cache request for all the content located at the origin Web server, which stores the primary copy of the application content. Every night, the content engine(s) verifies that the content is still current, having the added effect of increasing its
age in the cache, thereby reducing its likelihood of being removed from the cache. There is still a possibility that a small number of files may be flushed from the cache because of normal cache operations, but because most Web objects are small, a small number of cache misses is not usually considered significant. The following cache commands distribute that content listed in the preload.txt file that is retrieved between the hours of 8 p.m. and 4 a.m.. Because the preload feature actually generates requests that are directed to the cache, content that already exists is checked against the origin for validity, but the object is retrieved only if it is not up-to-date:

```
pre-load enable
pre-load url-list-file /local1/preload.txt
pre-load schedule every-day start-time 20:00 end-time 04:00
```

where /local1/preload.txt is something like:

```
http://appserver.foo.com/WebApp1/
```

### Managed Distribution

For the larger files such as the Java client applet, the network administrator pre-positions these elements at the edge, so they are in no danger of being removed from the cache. This pre-positioning of content is efficient primarily for smaller numbers of large files, and is not currently an efficient solution for delivering the large number of very small objects that make up the bulk of most Web applications, because of the current architecture of the system. In order to provide a transparent environment for the client requests, a content rule is enabled on the content engine, which redirects requests for the applet and any other large elements to a separate fixed cache on the content engine. The rule would be something like:

```
rule rewrite url-regsub appserver.foo.com/WebApp1/applet/ 127.0.0.1:8002/Cisco/0004eadbeadcb/AppServer/WebApp/applet/
```

Note that the internal redirection is to port 8002, which needs to be specified in the Cisco CDM for each content engine under Device Editor->TCP/IP. This keeps the system from going into a loop.

In conjunction with Web Cache Control Protocol (WCCP) content routing at the edge locations, the application can be seamlessly integrated into a production network. In addition, the same content engine can be used for standard transparent HTTP proxy caching, RealNetworks® proxy streaming, VoD (Video on Demand) caching and splitting, WMT VoD delivery, and even MPEG digital-to-analog VoD playback on a standard NTSC or PAL display with a built-in analog converter card.

### Conclusion

The combined caching and ECDN software from Cisco enables enterprises to dramatically increase network efficiency and efficiently enable the deployment of Web-based applications and other rich media services at the edge of a network. By taking advantage of WCCP transparent redirection in conjunction with Cisco IOS Software-based routers at the edge of the network, a Cisco Enterprise Content Delivery Network solution can accelerate HTTP-based content in a fashion that is transparent to the end user, and provides network redundancy in the rare case of a content engine failure. In addition, the Cisco ECDN solution can provide streaming-media services without impacting WAN bandwidth, while at the same time providing HTTP content acceleration. By incorporating caching technology in conjunction with content pre-positioning technology, and taking advantage of Cisco’s leadership in the switching and routing infrastructure, Cisco ECDNs can help a network administrator roll out the services that a corporation needs in order to reliably accelerate the growing number of HTTP-enabled applications that are now mission-critical to the enterprise.
Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the Cisco.com Web site at www.cisco.com/go/offices.