

Web Page Caching

*WEB CACHING IS PART OF THE
INTERNET SELECT BECAUSE
CACHING WORKS FOR SCHOOLS!*

Packed into the TOSHIBA Internet Select are features developed specifically for the special needs of the educational community. There are many time- and bandwidth-saving innovations schools can take advantage of with TOSHIBA technology; Web page caching is perhaps the most useful in a classroom setting. Caching provides high-impact productivity. Let's see why, and how it works in the Internet Select.

As you may know, Web page caching is the ability for a server or computer to 'cache' or save Web pages onto its hard drive. When a Web surfer calls for the page, parameters will determine whether it's more efficient to ask for the page from the original Web server over the Internet (takes more time) or directly from the Internet Select (saves time). The benefit? Quick access. And

properly set parameters—using some very impressive algorithms—virtually eliminate serving old versions of Web pages when newer ones exist.

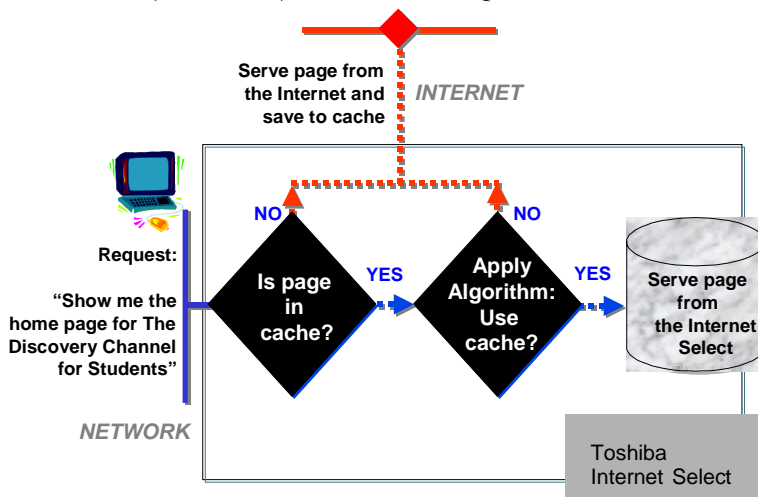
Consider the illustration. There's not much mystery to the logic in the caching scenario, except when it comes to the algorithm.



The Algorithm

We relied on a great high-performance web caching program called Squid to implement our caching technology. Squid—the result of the government-funded Harvest Program—is used world-wide by ISPs, and has gained great acceptance. At the risk of over-simplifying this very complex program, the logic works on levels of determination of whether the cached page is 'stale' or 'fresh.'

The logic is summarized as follows:



1. Is there a maximum age present in the page request? If yes, is the cached page older than this age? If yes, the cached page is stale, and the page is served from the source.
2. Is the maximum age less than the cached page? If so, the cached page is fresh and is served to the user.
3. Is there an EXPIRES time (expiry time from the server reply headers) present in the cached page? If so, has the cached page expired?
4. Is the object's age more than the MAX_AGE (maximum object age the client will accept)? If yes, the cached page is stale.
5. Is LM_FACTOR (the ratio of OBJ_AGE, how much the object has aged, to LM_AGE, how old the object was when it was retrieved) less than the configured value? If yes, the cached page is fresh.

The logic is quite advanced (and complicated!) but it all means that Web pages are always fresh, and are served from the cache whenever appropriate.

Why Schools Want Caching

Why is caching so important in a classroom setting? Unlike your typical surfing session at home, schools

use a focused, defined set of Web pages to support learning objectives in their curriculum. The same page (say, The Smithsonian Institute or The Discovery Channel for Students) may be accessed hundreds of times during the school day. Students will be constantly popping in and out of these sites to examine their contents.

Keeping a students' attention is a hard enough burden without waiting for Web pages to serve across the Internet. Thus, caching frequently-accessed pages makes the educational process that much smoother. To a teacher, having the information immediately available is crucial to easing the pressure of the process.

Caching speeds the educational process, reduces frustration, and makes Web access in the classroom more palatable to students and teachers.

The Bottom Line

Beyond the educational value, what really makes web page caching attractive to schools is *cost savings*. Since pages are served locally whenever appropriate, schools can get better performance out of less expensive connection scenarios, saving bandwidth and access costs. Complicated algorithms aside, *everyone can appreciate this concept!*