WCMS redirect issue post mortem

The WCMS has had an intermittent issue with pages reporting infinite redirect loops for some time (see RT 352675, 372084, and 376339 for some examples; note that 352675 was first submitted in June). Upon each investigation, the issue would either not be reproducable or would resolve itself without any intervention.

Without being able to reproduce the issue, testing was limited. We were able to confirm that user-created redirects were not the problem, nor were migration redirects. Nothing in code seemed to have potential to cause the problem. Jeff could not find any issues in the server logs.

In hindsight, these issues may have started with changes to the Varnish configuration that were implemented in early May (RT 337634), which changed the way HTTP/HTTPS was handled. The potential connection between the change and the issue was not made due to the infrequent nature of the issue.

On Tuesday, October 21, with the launch of the President's Town Hall Meeting registration form, we again saw this intermittent issue, and because of the high-profile nature of the event we received many reports. By Friday, October 24, it became important to find a resolution quickly despite the fact that we were still unable to reliably reproduce the issue.

Jeff looked through our code base and found only 2 examples of "301 redirects", and he conjectured that one of them might cause issues with the way the Varnish caching server handles HTTP vs HTTPS requests – specifically, that it doesn't know the difference.

Our web forms code checks to see if the user is on an HTTPS connection, and if not, it redirects them. Jeff's conjecture was that this redirect was getting cached by Varnish, and since it doesn't know the difference between HTTP and HTTPS, it would redirect both sets of users, causing an infitite redirect loop until the cached copy of the page expired. (The web forms themselves are deliberately never cached.)

Since we cannot in good conscience allow web forms to be accessed over HTTP, we agreed that we would simply move the WCMS to a HTTPS-only configuration, redirecting all HTTP traffic without touching Varnish.

We did some basic testing of this on our Pilots server on Friday morning without any issues, and Jeff implemented the necessary changes for the production servers early Friday afternoon.

Shortly after the changes were implemented on production, we received reports that some users were unable to access the sites at all through HTTP (which should have simply automatically redirected the user to the HTTPS site). With testing, we found very different results – some users were able to access everything without any problems, others could access http://uwaterloo.ca but not http://www.uwaterloo.ca, others couldn't access both. In addition, we received reports that CSS was not being properly implemented for some non-WCMS sites (it was actually not loading).

Upon investigation, Jeff was able to deduce that the problems were limited to users connected to the network using IPv4 instead of IPv6, and that the connection was failing at the load balancer. Jeff and I talked with Steven Bourque, the IT staff member in charge of the load balancer. He determined that the load balancer checked for the existence of a page over HTTP, and wouldn't pass on content if that page didn't load (and since we'd gone HTTPS-everywhere, that page didn't load). He adjusted the load balancer configuration to look for a 301 redirect from the HTTP page, and all issues with loading failing were resolved.

Communication to our clients was handled by Heather Wey, IST Communications Officer.

We believe we have resolved the infinite loop issue, but since we were never able to find a way to reproduce the issue for testing we cannot be absolutely certain unless we receive no further reports.