Granular Metadata Distribution With rsync

Granular Metadata Distribution With rsync

It has been suggested that one potential alternative to the current "hosts file"-like approach to sharing metadata would be to deploy a real time solution that's much like the current DNS system, allowing required metadata to be retrieved (and presumably cached) on a site-by-site basis in real time, much as DNS resource records are currently maintained in a distributed model, only getting retrieved as needed. While that model is intriguing, we know from the DNS itself that it also potentially introduces a raft of complications and potential vulnerabilities unless architected for redundancy and high reliability.

An alternative model to consider might be a periodically synchronized granular metadata distribution model, perhaps leveraging rsync (see http://en.wikipedia.org/wiki/Rsync). Rsync is a terrific tool routinely used to mirror web sites, transferring only those files that have changed since last invoked. For example, the Oregon Routeviews Project (see http://www.routeviews.org/) uses it to distribute BGP data to researchers.

If you'd like to see how it works, you can try retrieving some Routeviews data files with:

```
% rsync -av --skip-compress=bz2 archive2.routeviews.org::routeviews/bgpdata/2002.11/RIBS .
```

In the metadata distribution case, instead of distributing routing snapshots, each site might have its own individual signed metadata file, and rsync would make it possible for sites consuming that metadata to just download any newly-changed site metadata files.