GANESHA, a multi-usage with large cache NFSv4 server.

Work In Progress Submission. Contact: Philippe Deniel (philippe.deniel@cea.fr)

The NFSv4 protocol is quite different from former versions of the NFS protocol: it provides enhanced caching capability and it is generic enough in its semantic to handle many different file systems. GANESHA is a NFS server which runs in user space. Its design is oriented to take advantage of these two aspects: aggressive caching and capability to manage as many different sets of data as possible.

Caching is done for data and metadata. Performance is a key point, it is addressed by building the complete metadata cache within the memory. This cache is based on a associative addressing via hash tables based on Red-Black Trees based which scale like O(log(n)). The GANESHA server runs in user space, this makes it possible to allocate huge pieces of memory (several Gigabytes). An internal mechanism (based upon Buddy Block paradigm) optimizes memory use. One of the goal of the product is to keep several days of "active inodes" in memory.

GANESHA makes use of backend modules called "File System Abstraction Layers" or FSAL. A FSAL provides all the necessary calls to access a file system. NFSv4 requires only a few "mandatory" attributes for files and directories, which makes it possible to address very basic file systems. Therefore, a FSAL can be built on top of sets of data organized as trees, where each entry has a name. For example, SNMP and LDAP meet the requirement for building a FSAL on top of them. They could be "plugged" into GANESHA if the right FSAL is developed.

Currently, GANESHA supports the two following FSAL:

● "HPSS FSAL": with this module, GANESHA will provide NFS support to HPSS (HPSS is an HSM developed by the DOE and IBM Government Systems).

● "POSIX FSAL": with this module, GANESHA will provide NFS support to any file systems with no "handle based" APIs, through the use of the standard POSIX calls.

The following FSAL are under development:

● "NFSv4 Client FSAL": if you have a FSAL that is a NFSv4 client with GANESHA's aggressive caching, it will turn GANESHA into a NFSv4 proxy (by using NFSv4 file delegation and future NFSv4.1 directory delegation);

● "SNMP FSAL": with this FSAL, GANESHA will provide NFS access to SNMP. MIBS related information will be browsed like the /proc file system.

● LDAP FSAL: with this module, GANESHA will provide access to LDAP trees. Accessing them will be similar to browsing the /proc filesystem.

GANESHA has been used in production on our site for one year. A website and a submission to Freshmeat and Sourceforge will be available in a few weeks.