dCache NFSv4.1(pNFS*)

included into 1.9.3

*) pNFS (Parallel NFS) - is a set of OPTIONAL features of NFSv4.1 which allow direct client access to the storage devices containing the file data.

> "Separation of control and data flows provides a straightforward framework to accomplish this, by allowing transfers of data to proceed in parallel from many clients to many data storage endpoints. Control and file management operations, inherently more difficult to parallelize, can remain the province of a single NFS server, inheriting the simple management of today's NFS file service, while offloading data transfer operations allows bandwidth scalability."
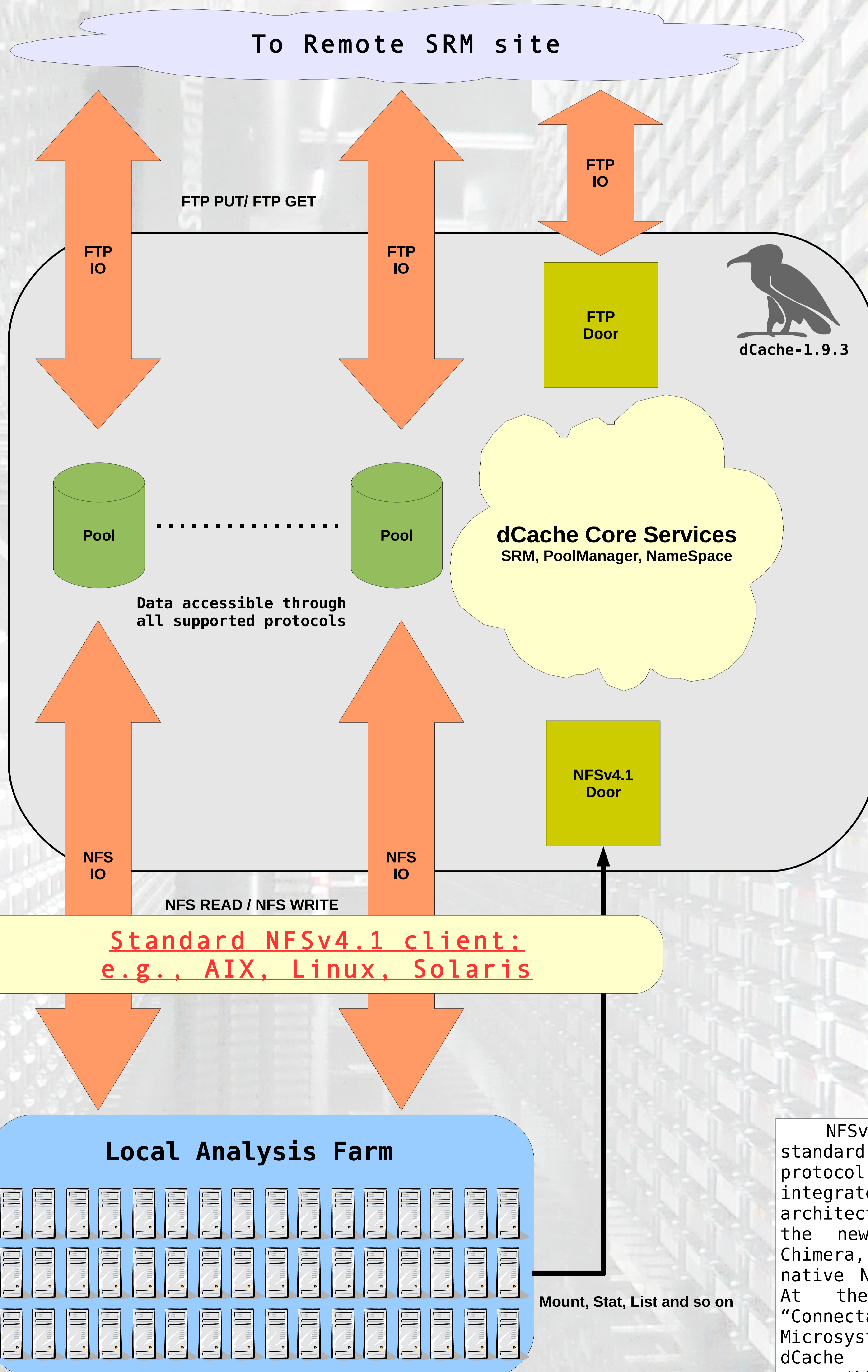>
> Internet-Draft - pNFS Problem Statement
> Garth Gibson , Peter Corbett

The NFSv4.1 protocol, defined by IETF and implemented by various Operating System and Storage Box vendors, e.g. EMC, IBM, Linux, NetApp , Panasas and SUN, provides : security mechanism negotiation (GSS-API, GSI, X509, UNIX), data access protocol negotiation (NFSv4 mandatory), clear distinction between metadata (namespace) and data access, support of multiple dataservers, ACLs, client and server crash recovery. The client modules are being developed for AIX, Linux, and the Solaris kernels.

On December 19 2008, draft-29 approved for proposed standard.

pNFS command set:

LAYOUTGET:
   used by a client to get a
   layout segment for a file.
LAYOUTRETURN:
   used to return a layout
   segment or all layouts
   belong to a file system
   to a metadata server.
LAYOUTCOMMIT:
   used to inform the metadata
   server that the client wants
   to commit data it wrote to
   the storage device
GETDEVICEINFO:
   returns the mapping of
   device ID to storage
   device address.
GETDEVICELIST:
   allows clients to fetch the
   all the device ID to storage
   device address mappings of
   particular file system
CB_LAYOUTRECALL:
   used by server to recall
   a layout segment or
   all layouts belonging to
   a file system, or all
   layouts belong to a
   client ID.

To Remote SRM site

FTP PUT/ FTP GET

FTP IO

FTP IO

FTP IO

FTP Door

dCache-1.9.3

Pool . . . . Pool

dCache Core Services
SRM, PoolManager, NameSpace

Data accessible through all supported protocols

NFS IO

NFS IO

NFSv4.1 Door

NFS READ / NFS WRITE

Standard NFSv4.1 client:
e.g., AIX, Linux, Solaris

Provided by OS Vendor.

From the client point of view it's a regular NFSv4.1 server, while in reality it's dCache.

No special software required. The same client can talk to different storage systems in parallel (Federated FS).

Local Analysis Farm

Mount, Stat, List and so on

NFSv4.1 is an open standard, industry backed protocol which easily integrates into the dCache architecture. Together with the new namespace provider, Chimera, dCache provides a native NFSv4.1 implementation. At the most recent NFS "Connectathon" at SUN Microsystems February 2009, dCache has proven to be compatible to all existing clients. The NFSv4.1 support included into upcoming 1.9.3 release of dCache.