XROOTd for cloud storage

Hironori Ito
Brookhaven National Laboratory

Cloud Storage

- Remote storage popularized by commercial vendors like Dropbox or Google drive
  - Many commercial free storage; Dropbox, Google drive, MS Onedrive, Amazon Drive,etc...
  - Free for small size. Some of them offer unlimited space with small costs.
- They are used like backup of your laptop/desktop or archival storage of your pictures and videos.
  - Synchronize local contents with those at cloud storage
    - Think of it like “rsync”
- Open source alternatives for cloud storage.
  - Owncloud, Nextcloud, pydio, etc...
    - CERN box, etc...

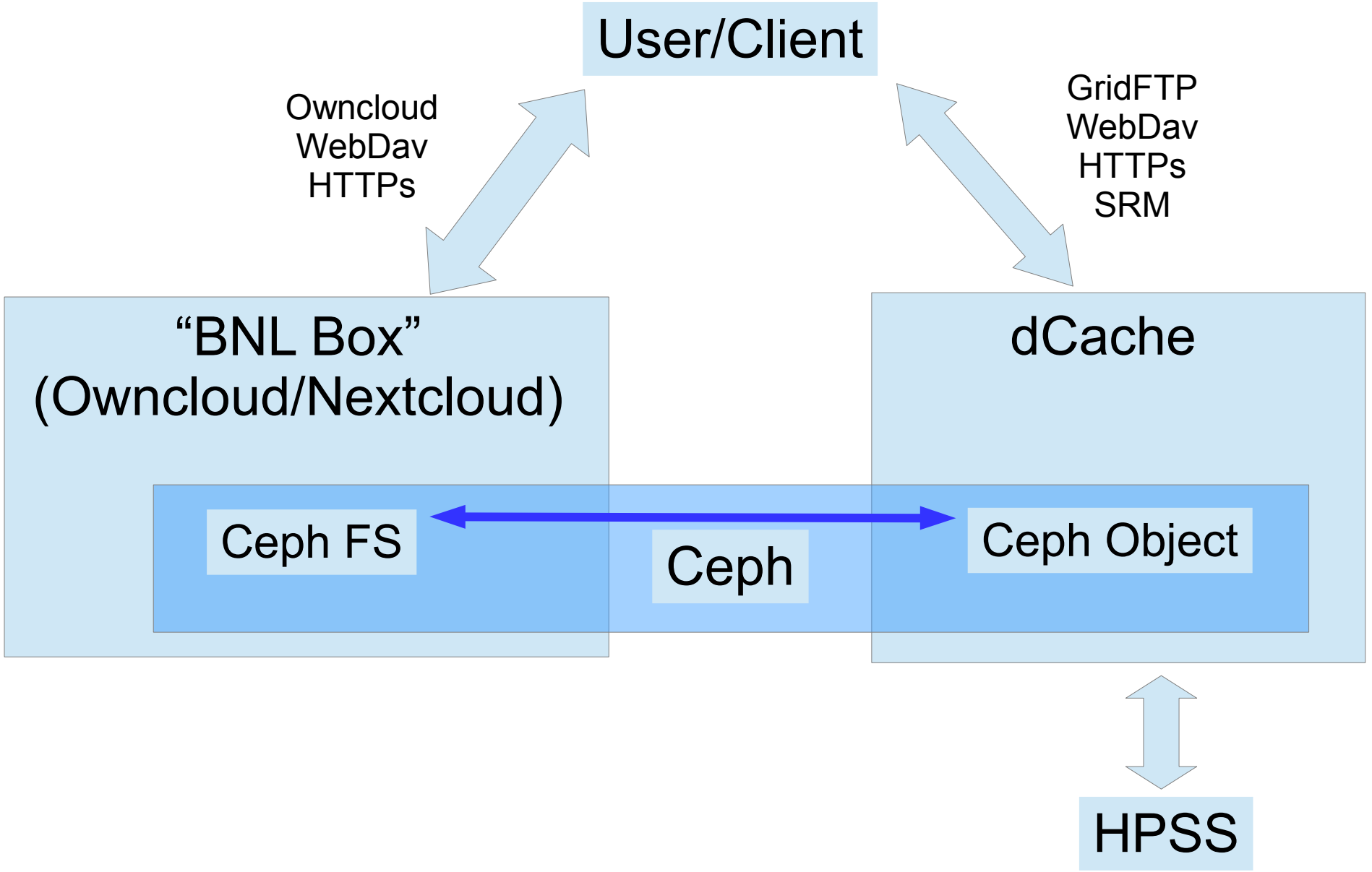
Cloud Storage

- Remote storage popularized by commercial vendors like Dropbox or Google drive
  - Many commercial free storage; Dropbox, Google drive, MS Onedrive, Amazon Drive,etc...
  - Free for small size. Some of them offer unlimited space with small costs.
- They are used like backup of your laptop/desktop or archival storage of your pictures and videos.
  - Synchronize local contents with those at cloud storage
    - Think of it like “rsync”
- Open source alternatives for cloud storage.
  - Owncloud, Nextcloud, pydio, etc...
    - CERN box, etc...

Owncloud/Nextcloud

- Using the already existing, one-shot command to synchronize the contents of local machine to those of the cloud storage, one can achieve about 100MB/s.
  - The reason is that the client software limits the bandwidth.
    - The most of normal users don't have WAN network throughput of 100MB/s or even 10MB/s. Users normally don't want to use all available network bandwidth.
- Owncloud has WebDAV interface.
  - The custom client is easily possible.
  - Using the WebDAV, one can achieve much higher rate (>1GB/s)
  - Owncloud interface is made by PHP.
    - It uses 8KB block size for IO read/write. Problem???
- The cloud storage normally requires the same amount of storage in local storage and remote cloud storage.
  - 100GB in local machine = 100 GB in cloud storage
  - It is not usually used like central storage.

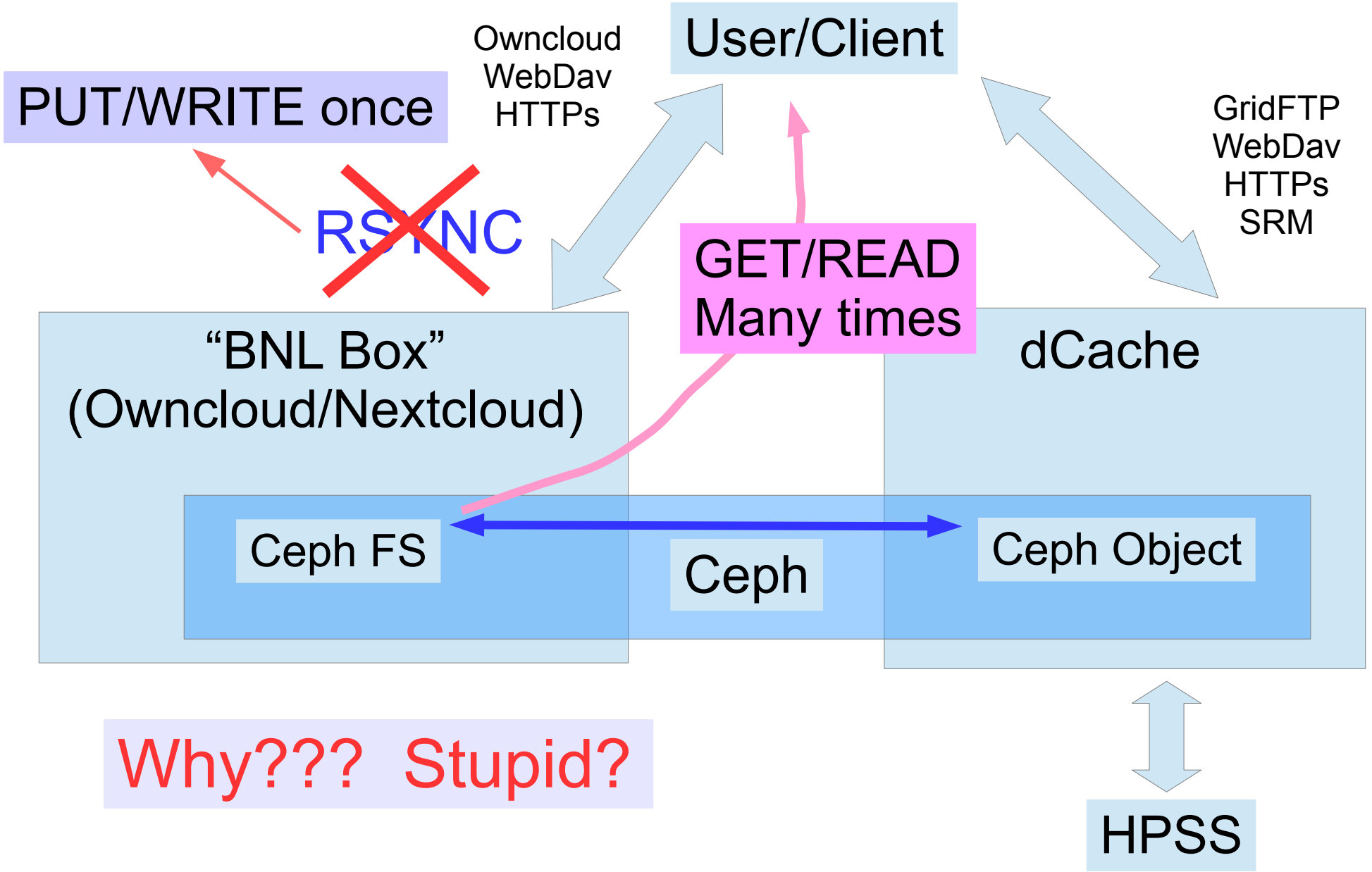
Proposed Scheme



Central Storage

- The cloud storage is great. But, I also want to use it as a central storage as well.
  - Reason
    - Data volume: The size of data won't fit in the local storage.
    - Still would like to use the easy-to-use Owncloud interface for the most of the operation.
      - Grid commands based on X509 are too complicated outside of HEP.
    - Data analysis: would like to use computing farms and not my local machine.

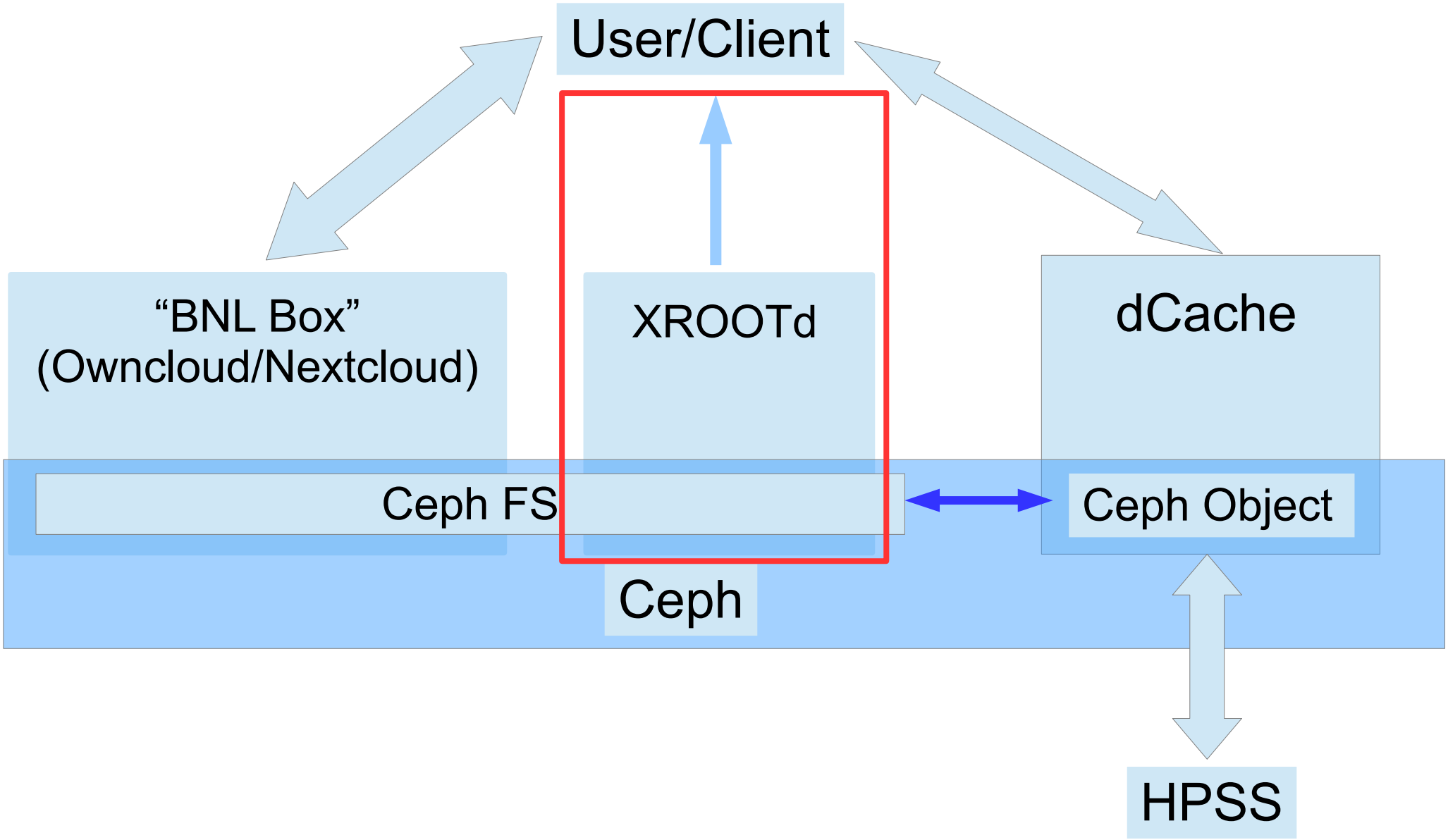
Use it like Central Storage



XROOTd as Read IO Service

- Owncloud/Nextcloud stores data with Apache owner.
  - XROOTd also usually run as “one” (xrootd) user.
- Owncloud/Nextcloud stores different user data in separate, isolated directories.
  - User A → /A/... while user B → /B/...
  - XROOTd's acc.authdb directive provides easy way to authorize a particular user to only his/her own data.
    - Can we not use **FLATFILE**
- Data is stored in Cephfs.
  - Multiple XROOTd services can be run on separate hosts to provide resilient read IO

Cloud storage as Central Storage



Sample Interfaces