BNL Box

Hironori Ito

Brookhaven National Laboratory

Concept

- All of us need the convenient method to transfer or access data in different systems
  - Users might need to copy their analysis scripts and the data between their workstations and central analysis farm separated by different network and firewalls
  - System administrators might need to transfer custom software packages to their systems for installations.
- In BNL RACF, AFS has been the storage of choice for moving small amount of data in/out of various systems.
- AFS limitation
  - Not really universally accessible.
  - Not easiest one to use in various platform.

- Commercial cloud storage seems to be popular among some of users and sys-admins.
  - Dropbox, Box, Amazon Cloud Drive, Google Drive, MS OneDrive, etc…
  - Advantages of commercial cloud storage
    - Already available for use
    - Easy to use. All of then provide https-based access.
    - Free (up to some level)
    - Available in various platforms.
  - Limitations
    - Size/Cost/Performance.
    - Archive
    - Not really meant to stream data

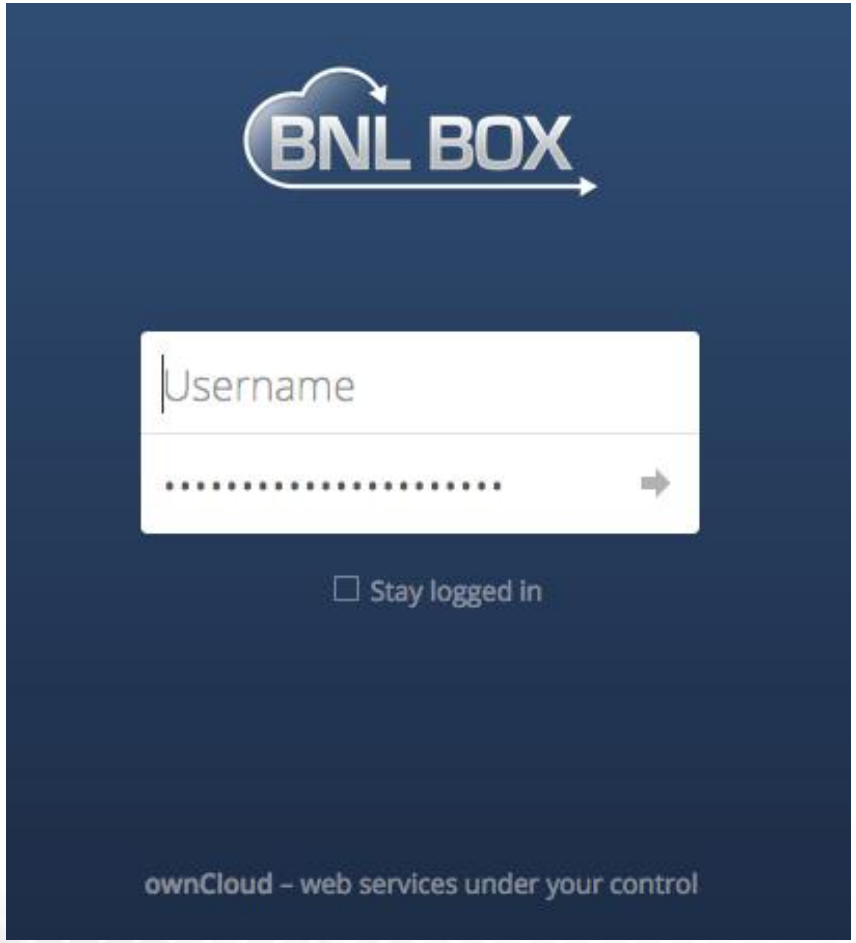
Target users

- All scientific users of BNL
  - HEP/Nuclear physics communities
  - RACF Staff
  - Users from different science domains than HEP
    - CFN
    - NSLS-II
    - Chemistry, Biology, etc…

Target usage

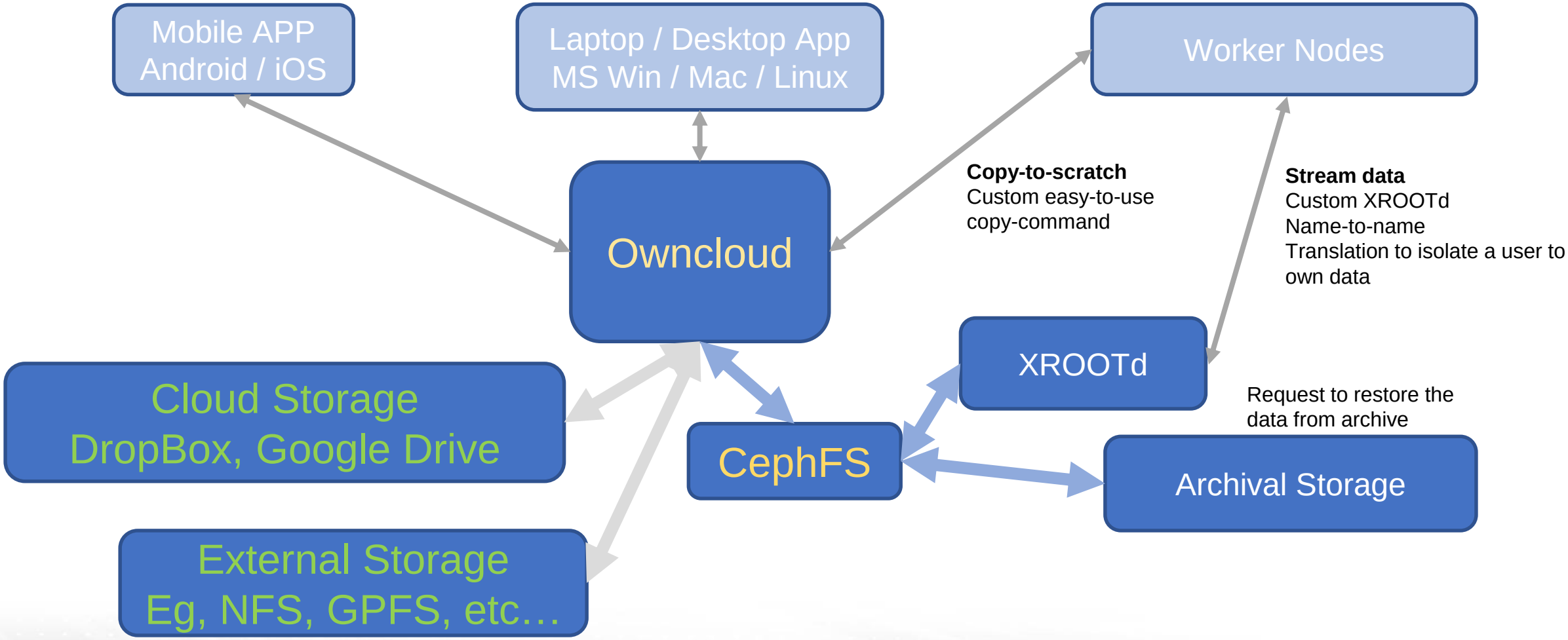
- Transfer small data, documents in & out of BNL between central interactive computing nodes, desktop workstations, laptops, and tablets/smart phones.
- Transfer large data in & out of BNL between detector data stores, central storage, remote storage of users.
- Access data to/from analysis computing farm
- Archive data

BNL BOX



- Owncloud Software
  - Clients are available in many popular platforms; Linux, Mac OS, MS Win, Android and IOS
  - Extremely easy to use.
    - Synchronize data automatically
    - NOTE: **Requires the same amount of storage in local and remote storage.**
  - Quota for each users
  - Users can share data with the members and non-members
  - Has been used with large number of users at different sites
- CephFS Storage

Diagram


Mobile APP
Android / iOS
Laptop / Desktop App
MS Win / Mac / Linux
Worker Nodes
Owncloud
Copy-to-scratch
Custom easy-to-use
copy-command
Stream data
Custom XROOTd
Name-to-name
Translation to isolate a user to
own data
XROOTd
Request to restore the
data from archive
Cloud Storage
DropBox, Google Drive
CephFS
Archival Storage
External Storage
Eg, NFS, GPFS, etc…

WebDAV access and Sync

- Default sync app seems to synchronize data at the top rate of about 100MB/s per client. (100MB/s = 360GB/Hr = 8.6TB/day)
  - Sufficient rate for small data ~ less than TB.
    - For the purpose of using under the desktop/laptop/smartphone, the performance of the sync program is sufficient because
      - Spinning DiskIO on desktop(~100MB/s).
      - Wifi N (max 300Mbps~40MB/s)
      - LAN (1Gbps=120MB/s)
      - A single disk is not much larger (currently max at about ~10TB)
- High volume users will require higher throughput.
  - 10TB or more.
  - Owncloud supports standard WebDAV protocol
    - Easy to write a custom copy tool. (copy_bnl_box.rb)
      - Easily achieve 150MB/s per single file transfer.
      - Concurrent multiple transfer of files will results in obtaining desired throughputs.
      - NOTE: Different SSL library seem to impact the observed throughput of WebDAV command. For an example, “curl” in RHEL 7 is compiled with NSS. This version of “curl” produces 1/5 of throughput of “curl” using OpenSSL.

Stream Access

- XROOTd and WebDAV can stream data
- Would like to separate the data-sync operations from the data-read access as much as reasonably possible.
- XROOTd can cleverly map user data in BNL Box in a very simple way.
  - Owncloud web URL maps a user data by https://host/owncloud/index.php/apps/files/MYDATA
  - This is different from how Owncloud physically stores user data in its storage as /base-directory/**username**/files/MYDATA
  - XROOTd can cleverly hide “username” of physical files by providing access by root://host/files/MYDATA
    - Courtesy of Andrew Hanushevsky from XROOTd

Archive data

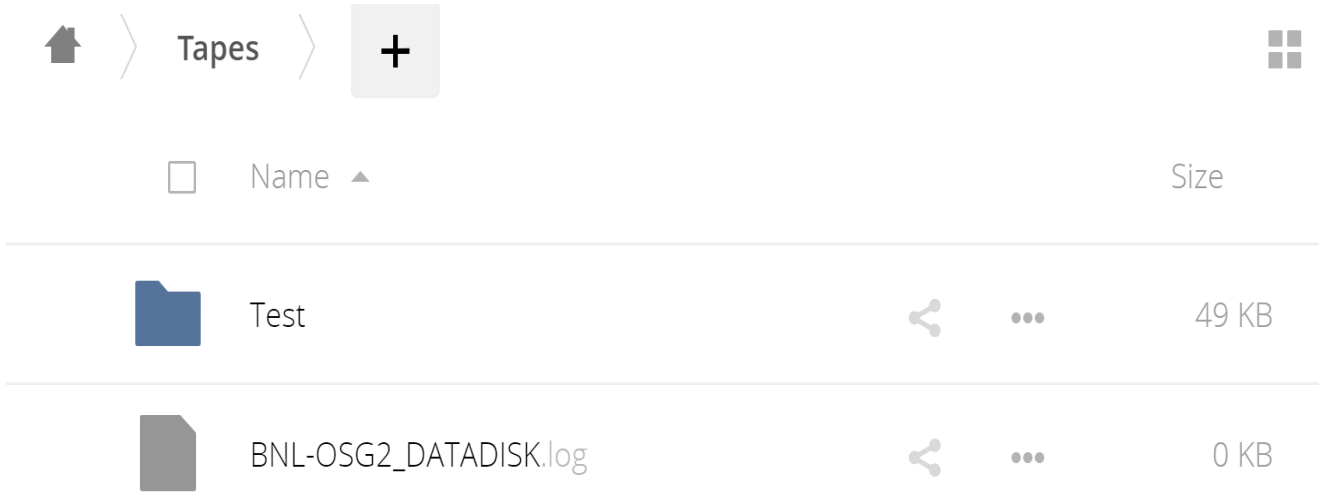
- Some users would like to archive rarely or never read data to archive
  - Will the data be read again?
- Difficulties
  - Tape system is relatively slow in
    - Seek
    - Mounting a tape
  - Number of Tape drives are limited.
    - Limited concurrency
- Must write in "right way" to produce the good or even decent read-IO for everyone.

- Intended.
  - "/Tapes/" directory will be used to indicate data to be stored to the tape system.
  - Files smaller than 1GB(?) will be copies to the archival "disk" storage until a better solutions are found.
    - Tar/zip option?
  - Files larger than 1GB(?) will be copied to the tape storage
    - File family for user???
      - Group sets of tapes to a particular user.
      - Probably to large data users.
    - Once copied, the size of the file will be changed to zero and the dummy file (with the same name) will be kept in same path.
    - Also, the copy is recorded in the external DB.
  - Restore requests will be made through Web interface.
    - Data will show up in the system place.

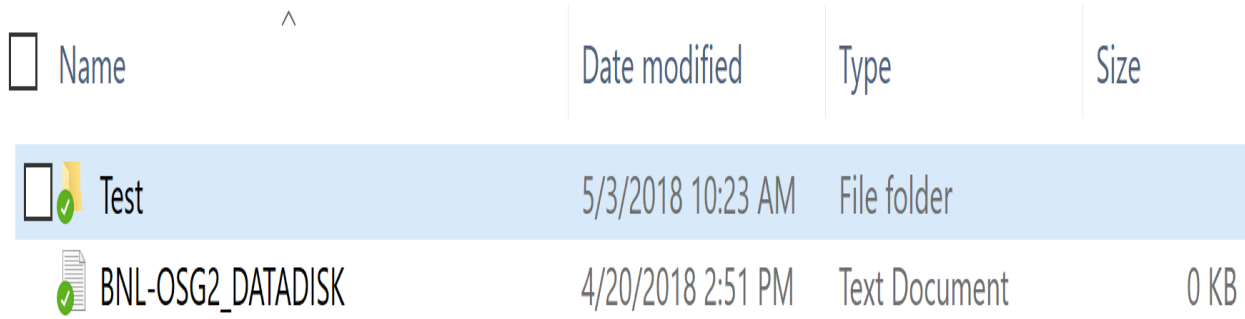
Technical details of Archive

- All newly written files under “/Tapes” will be monitored by Inotify
- Once a file is written to “/Tapes”
  - Copied to external storage and resized to zero.
  - The file will be rescan by BNLBox to properly register “0” size file.
  - The client now will sync “0” size file (if “sync” is used).

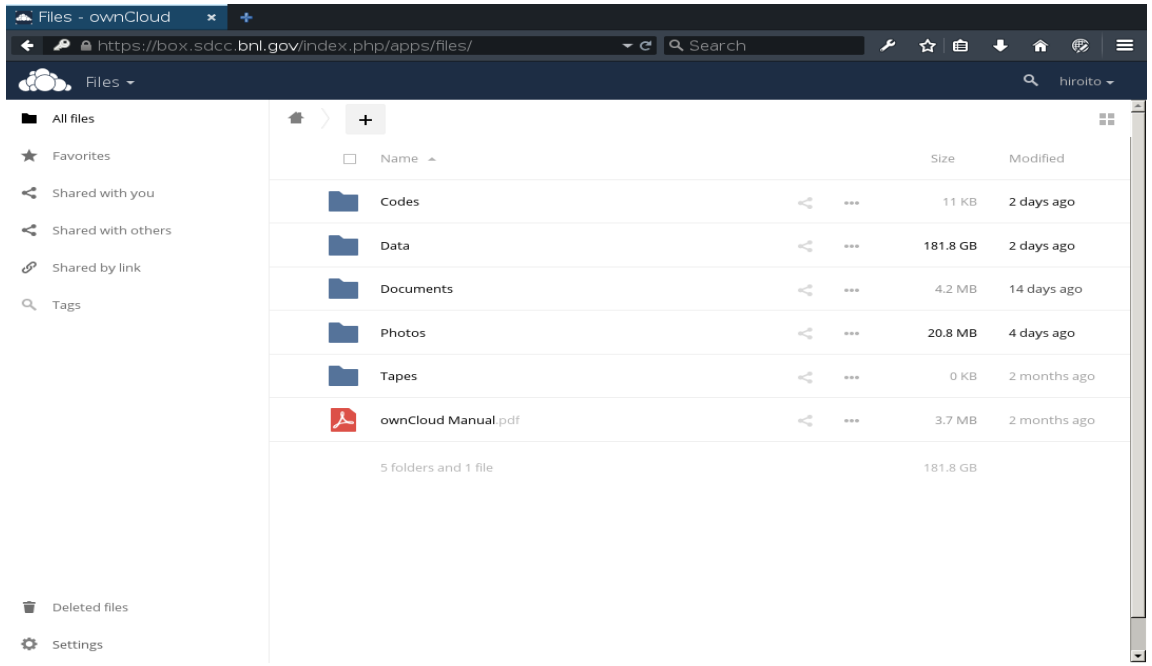
In BNLBox WebDAV

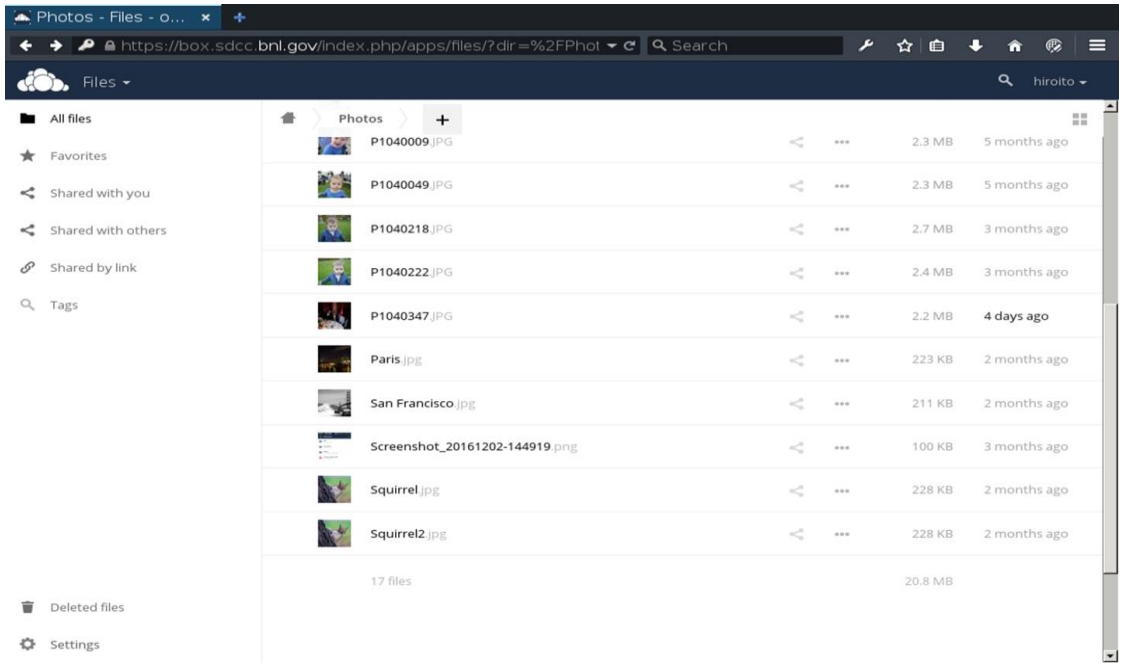


Owncloud Client (Windows 10)



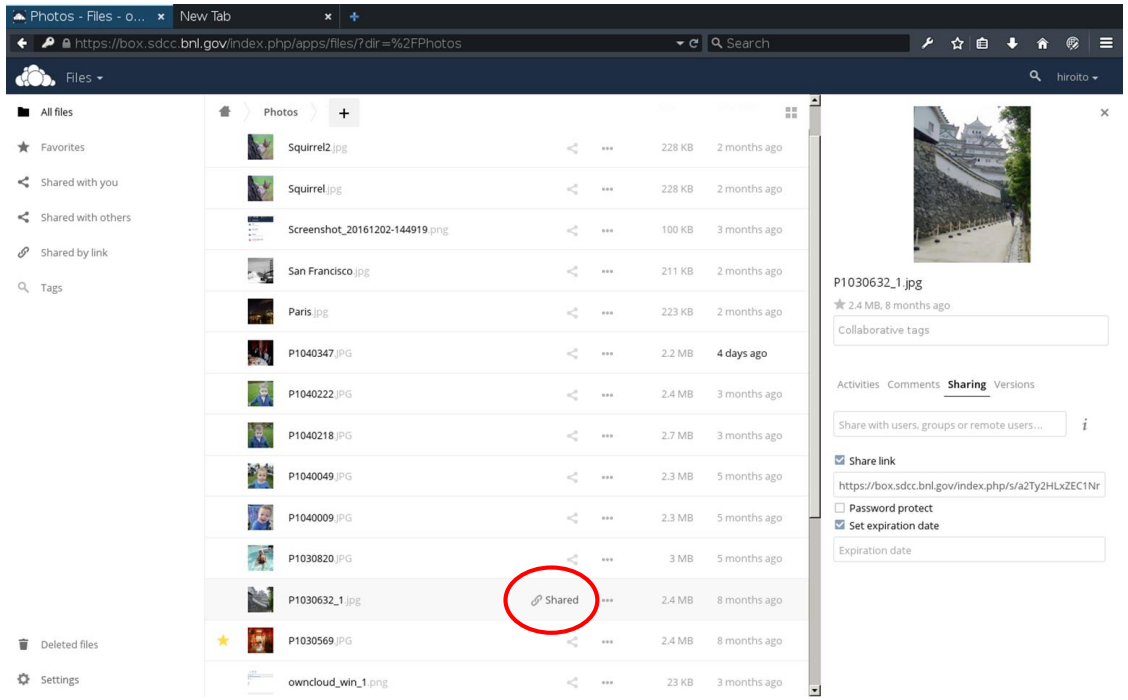
Sample images

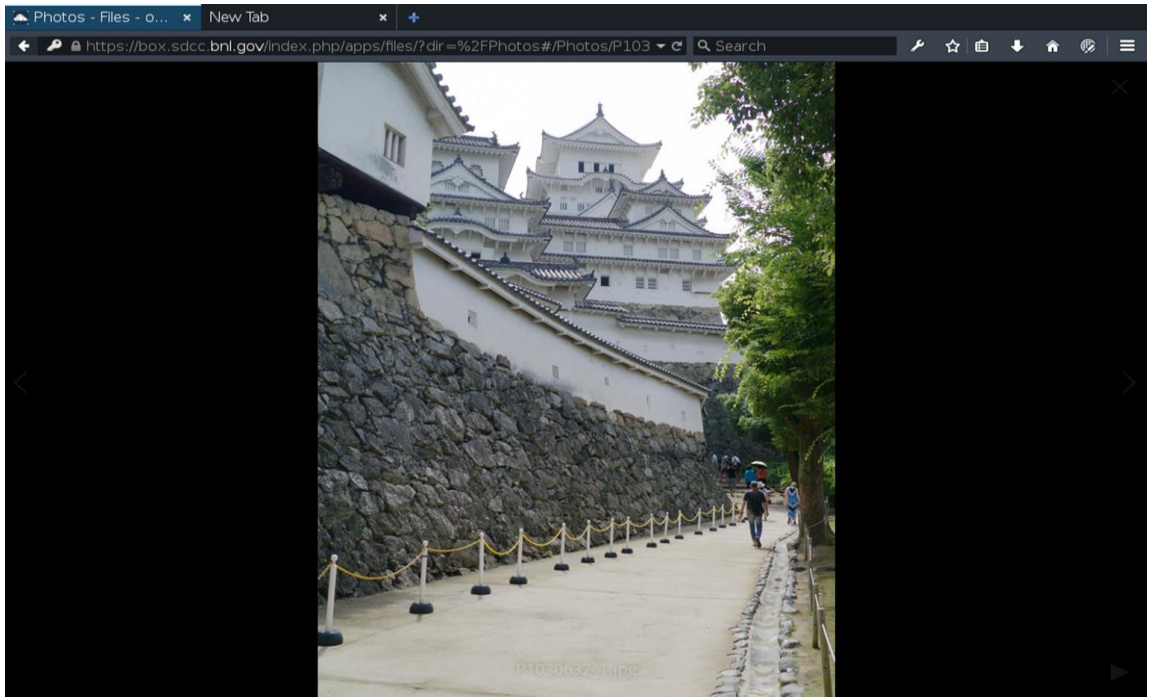




- Users only see their own directory.

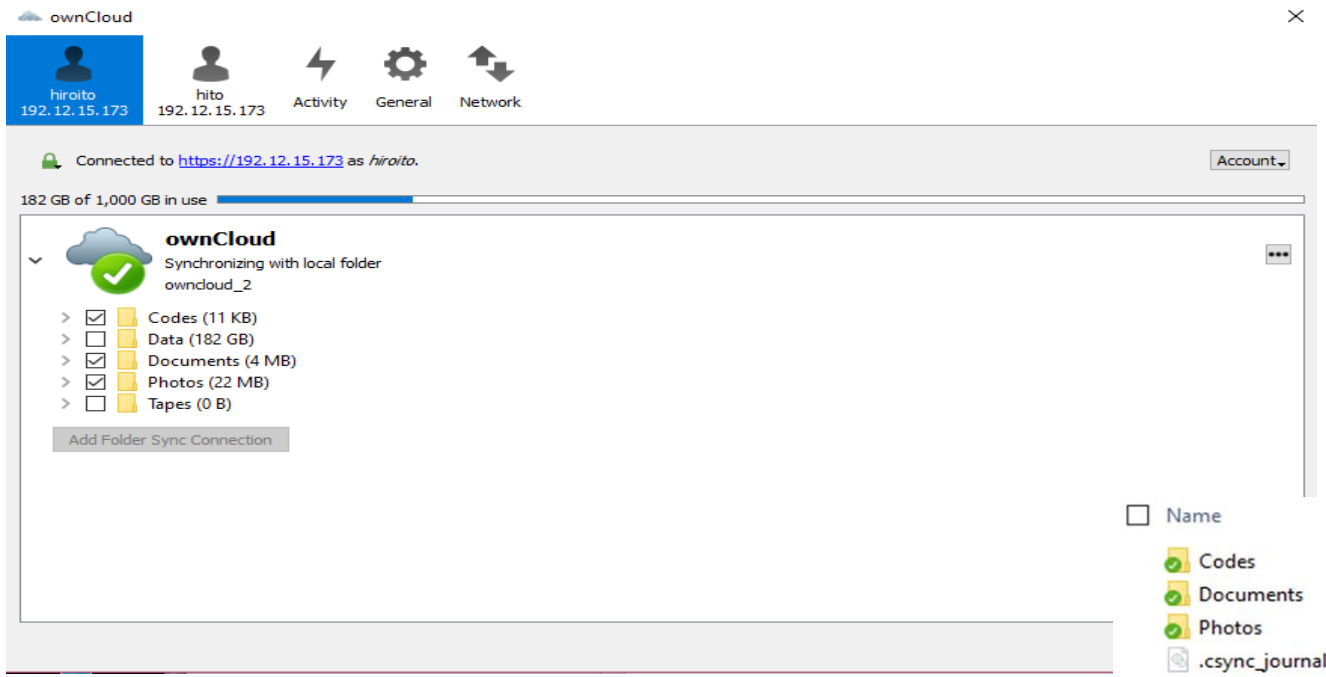
Share data

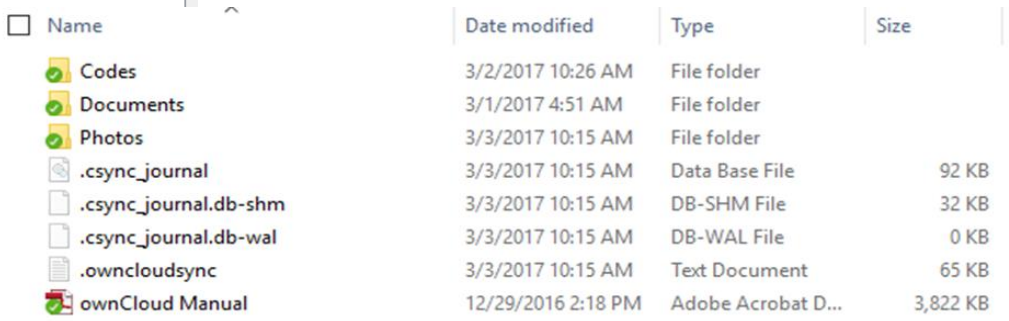




- User can share their data publicly or privately with password.
- User can set the expiration dates.
- User can share users in the system or anyone (with/without passwd).

Users decide what to sync





Using the provided app, users can decides what to sync automatically.
For an example

- Data and Tapes directories are not synchronized.
- Codes, Documents, Photos directories are synchronized automatically.

Desktop/Laptop apps are available in MS Win, Mac and Linux.
The performance seems to be limited to the maximum of 100MB/s.

Encryption

BNLBox is not meant to store the sensitive data.

- Client **side**
  - **Advantages**
    - Encrypt at the point of file creation.
      - A user can only see encrypted and the decrypted content of the data
      - A system administer can't even see the decrypted content.
    - Various options to encrypt
      - **EncFS** (see the next slides)
      - GnuPG(GPG), Zip with password, etc…
    - As secure as you want to be.
      - Some might be more secure than the other.
  - **Disadvantage**
    - Integration with BNLBox is limited.
    - Users must chose your favorite encryption software.

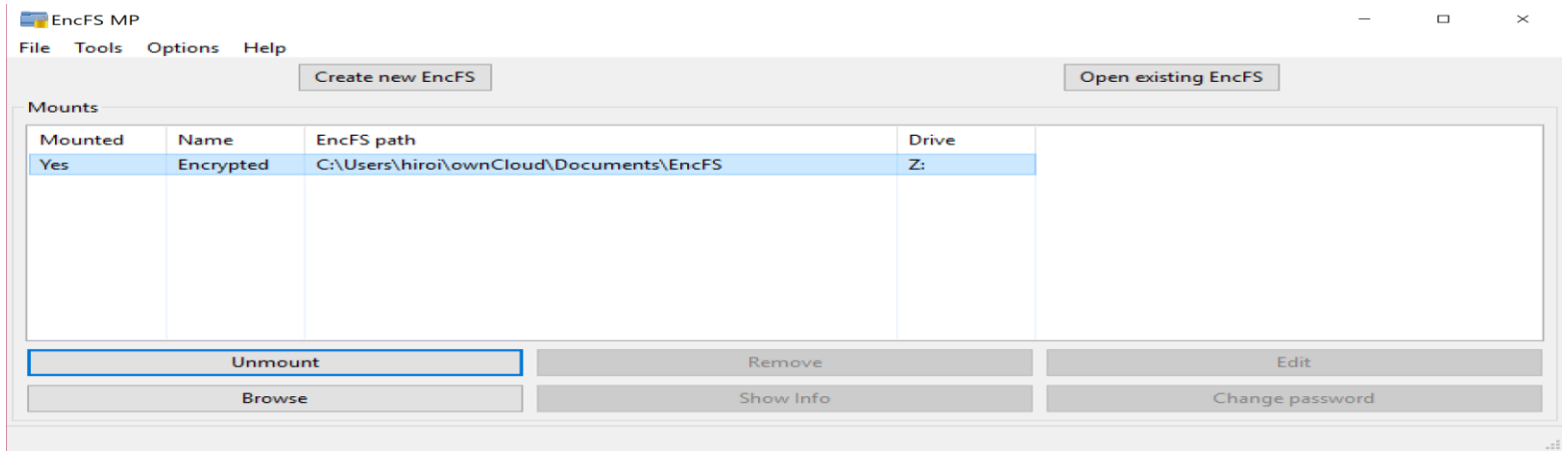
- Server **side**
  - **Advantages**
    - Tightly integrated with the software.
    - No need to learn the separate client software.
  - **Disadvantages**
    - All data are encrypted.
    - No choice of encryption software.
    - A system administrator might be able to see the content

- BNLBox will not encrypt your data.
- The access to the service is secure by the industry standard HTTPs.
- User are free to encrypt own data if desired.
- Encryption and decryption are the responsibilities of user.

EncFS Encryption

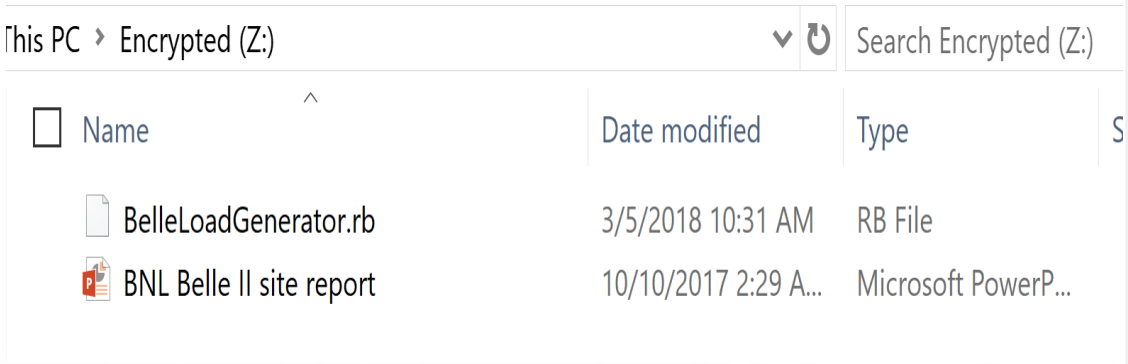
- Cross platform
  - Android, Linux, MacOS and MSWin
- Files are encrypted individually at a client host.
- The names of files are also encrypted, preventing peaking by anyone, including the administrator of BNLBox and the backend storage (CephFS).
- The files are encrypted at the client host.
  - The owner of the data is the only one who can decrypt the data.
  - NOTE: *BNLBox system administer can not recover your encryption key/password. Make sure to backup your key/password. But,* ***please don't put them on BNLBox or any other cloud storage***.
- Encryption and decryption happen automatically once it is setup.

EncFS with BnlBox

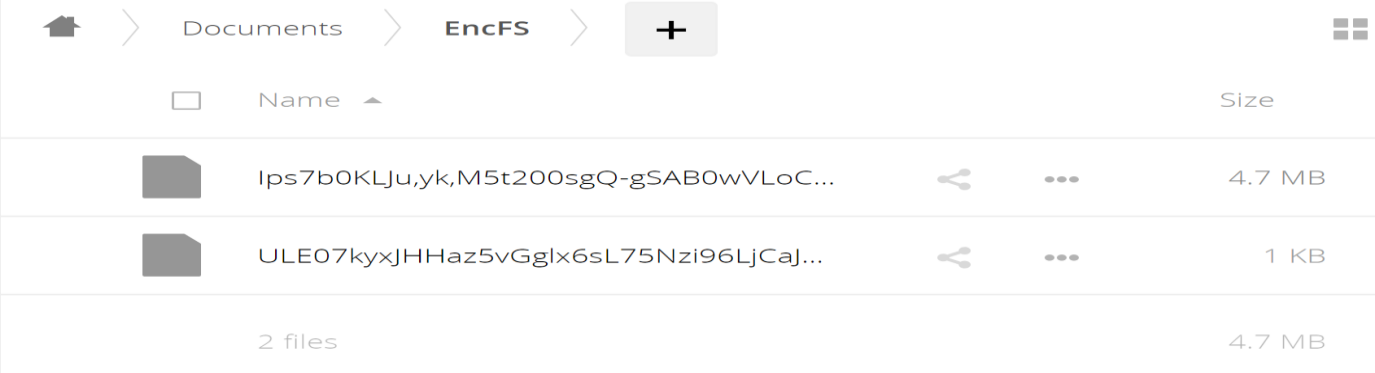


Specify EncFS to
(a) Original/non-encrypted directory (Eg, Z:)
(b) Encrypted directory (Eg, ownCloud/Documents/EncFS

Non-encrypted directory in the client host



Encrypted directory (in the client and BnlBox)



External Storage

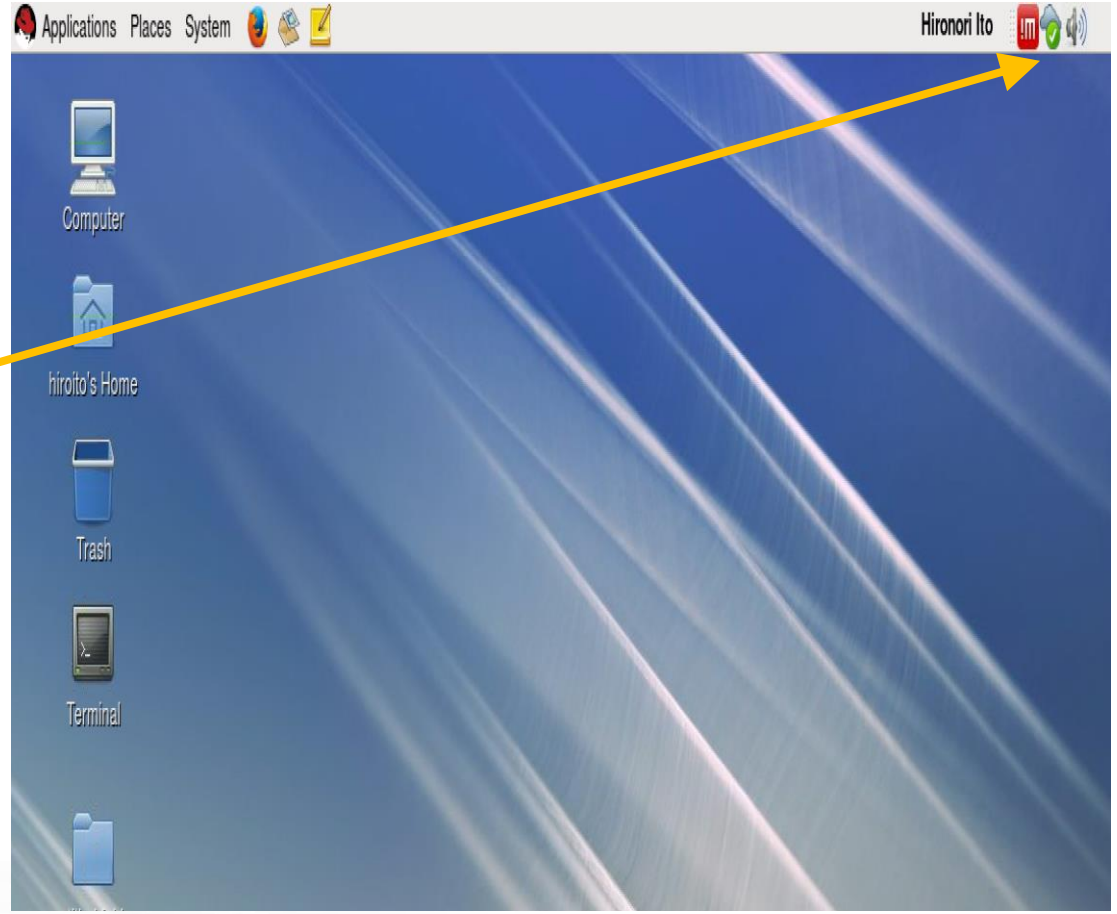
- Owncloud currently supports the following external storage.
  - FTP, SFTP, Owncloud(remote), Google Drive, Dropbox, Amazon S3, SAMBA/CIFS, WebDAV, OpenStack Object Storage.
    - Using SFTP option, BNLBox might be possible to serve the data stored in GPFS or NFS without copying to its own storage (CephFS).

To do lists.

- The deployment of the production CephFS.
  - Current CephFS is a test instance.
- The complete the development of archival system.
  - Deployment and configuration of the external archival storage for small files
  - Create the restore mechanism and integrate to the existing WebUI.
- Integration with single sign-on. Duo?
- Improve the “copy” client command
  - Recursive
  - Check the diff (rsync like)
- Creation of the custom preview for unsupported image types
- Test external clients. Particularly SFTP
- The performance tests
  - “Copy” command
  - Scalability of large number of sync files.
  - Write/Read with archival storage.

More examples of its use

- Copying files from your laptop to interactive nodes at RACF?
  - Laptop with Sync software
  - NX servers at RACF with BNLBox client



Copying without sync clients

- copy_bnl_box.rb -h
  Usage: copy_bnl_box.rb [options] source destination
  NOTE: to specify the owncloud storage, you need to append 'box:' to your path seen in your bnl box service
  -h, --help Display help message
  -s, --https_proxy ProxyHost Set https proxy
  -r, --url URL Set none-default url
  -p, --password password Set User Password
  -u, --username username Set User Name

- Copy from local to BNLBox

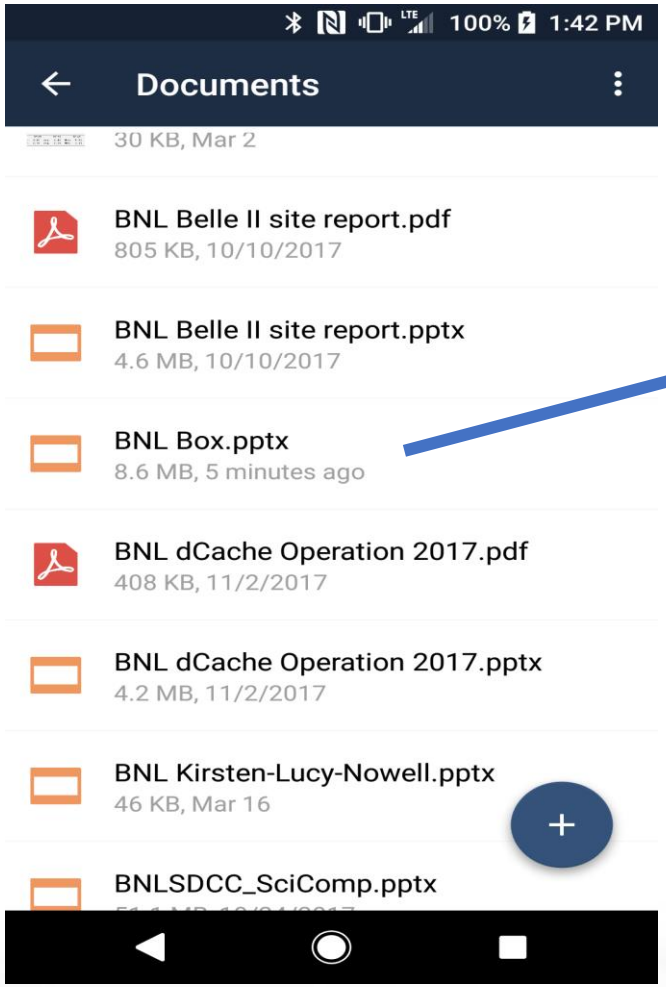
  ruby copy_bnl_box.rb vomses box:/Documents/testcopy.1
  Username:
  Password:
  Source: vomses
  Destination: /owncloud/remote.php/dav/files/hiroito/Documents/testcopy.1

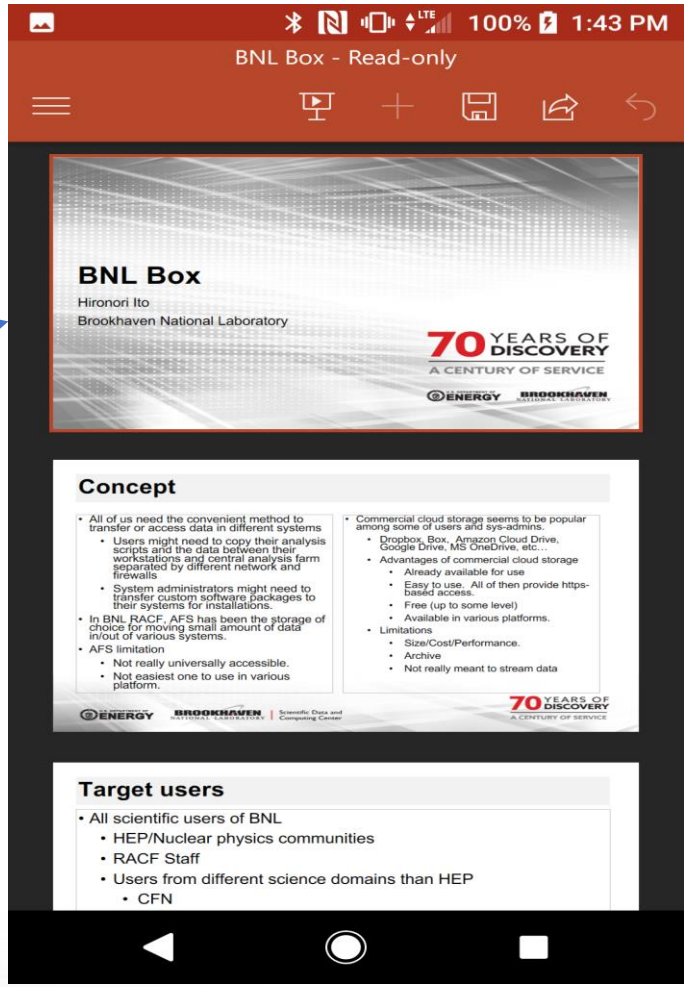
- Copy from BNLBox to local

  ruby copy_bnl_box.rb box:/Documents/testcopy.1 /home/hiroito/testcopy.1
  Username:
  Password:
  Source: /owncloud/remote.php/dav/files/hiroito/Documents/testcopy.1
  Destination: /home/hiroito/testcopy.1

Presentation on the go?

Open a file via Smart Phone or table





Conclusion

- Cloud storage could be useful for scientific communities.
- BNL Box will provide our users with ability to store and access their data anywhere by the easy-to-use applications on various platforms.
- BNL Box allows the owners of the data to share with anyone without involvement of the system administrator.