BNL FAX Setup

Hironori Ito
Brookhaven National Laboratory

List of FAX related services

- dCache Storage
- dCache xrootd services in pnfs namespace
- FAX Reverse proxy redirector/servers
- FAX (Forward) proxy redirector/servers

dCache Storagte and xRootd door

- dCache provides all T1 data from disks and tapes
- dCache storage servers are behind the firewalls. They are not accessible from the outside BNL computing facilities.
- dCache provides own xRootd interface (door) to access the data from xRootd clients.
- Files are accessible via pnfs name space path

  NOTE: It is possible to change via N2N. But, it is not used in BNL.
- Clients must have direct access to the storage servers hosting requested data.
  - Data is transferred directly from the storage server to the client host.
- Example.
  - Path
    - root://dcdoorXY.usatlas.bnl.gov:1096/pnfs/usatlas.bnl.gov/...
  - Use
    - xrdcp root://dcdoorXY.usatlas.bnl.gov:1096/pnfs/usatlas.bnl.gov/...
  - Storage servers
    - dcABC.usatlas.bnl.gov
      - Clients must be able to ping these hosts.

FAX Reverse proxy

- FAX Reverse proxy service allows the access to BNL data behind the firewall by clients located outside BNL. (It also allows clients within BNL to access the data. But, see the caution for the next section.)
  - NOTE: It will also allow clients to access none BNL FAX data.
- FAX Reverse proxy understand ATLAS global namespace via its N2N plugin.
- It is consisted of xRootd reverse proxy manager (redirector) and several xRootd reverse proxy servers (aka, data servers).
  - They are being run on the dCache's gridftp servers since they are used for same function. They proxies the data-transfer into/out-of the dCAche storage servers without them going through site firewalls since ATLAS traffic can easily overwhelm the site-wide firewall, resulting in possible, complete shutdown of site network traffic.
  - Reverse proxy servers have dual network interfaces to access both external and internal networks.

Configuration detail in FAX reverse proxy

/etc/xrootd/xrootd.cfg

```
all.manager proxy dcdoor11.usatlas.bnl.gov:1213
all.manager meta atlas-xrd-us.usatlas.org+ 1095
all.export /atlas r/o

if dcdoor11.usatlas.bnl.gov
   all.role proxy manager
   cms.dfs lookup central
   cms.delay servers 0
else
   all.role proxy server
   xrootd.redirect atlas-xrd-us.usatlas.org:1094 ? /atlas
   pss.setopt ReadCacheSize 0
   xrd.report atl-prod05.slac.stanford.edu:9931 every 60s all -buff sync
   xrootd.monitor all auth flush 30s  window 5s dest files info user redir atl-prod05.slac.stanford.edu:9930
fi

xrootd.trace emsg login stall redirect
xrd.trace all
cms.trace all
all.adminpath /var/run/xrootd
all.pidpath /var/run/xrootd

pss.origin dcdoor05.usatlas.bnl.gov:1096
pss.namelib /usr/lib64/XrdOucName2NameLFC.so root=/pnfs match=usatlas.bnl.gov

sec.protparm gsi -vomsfun:/usr/lib64/libXrdSecgsiVOMS.so -vomsfunparms:certfmt=raw|vos=atlas|grps=/atlas
xrootd.seclib /usr/lib64/libXrdSec.so
sec.protocol /usr/lib64 gsi -ca:1 -crl:3 -gmapopt:10
acc.authdb /etc/xrootd/auth_file
acc.authrefresh 60
ofs.authorize

all.sitename BNL-ATLAS
```

FAX Regional Redirector

Proxy Manager/director

Proxy Servers

dcache xRootd service

N2N

ATLAS VOMS

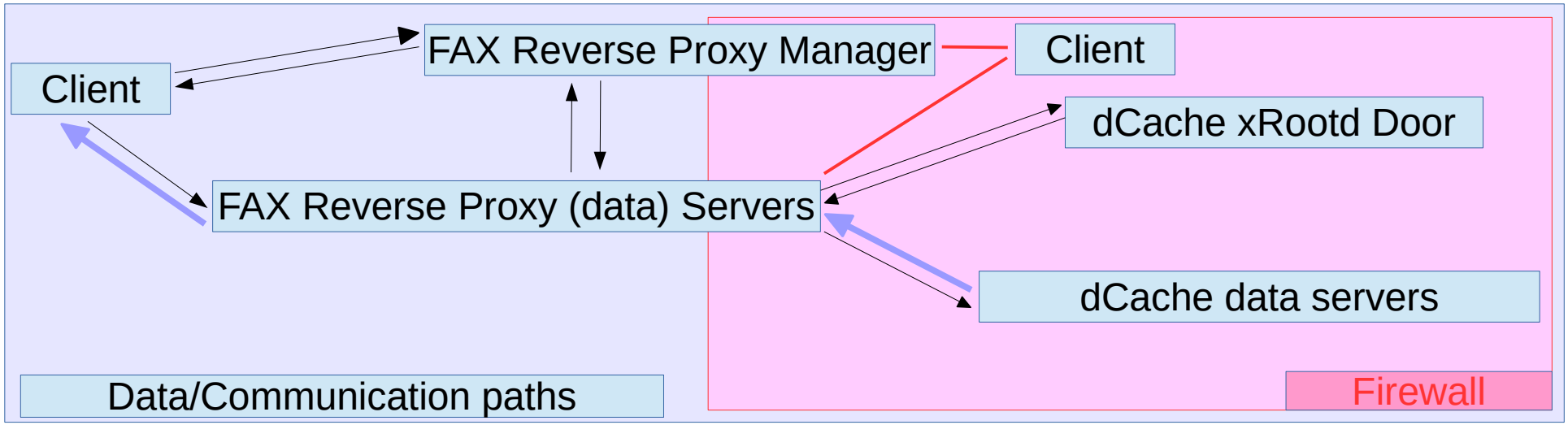
GSI options

/etc/sysconfig/xrootd

```
XROOTD_USER=xrootd
XROOTD_GROUP=xrootd
export X509_USER_PROXY=/tmp/x509up_proxy
export X509_VOMS_DIR=/etc/grid-security/vomsdir
XROOTD_DEFAULT_OPTIONS="-l /var/log/xrootd/xrootd.log -c /etc/xrootd/xrootd.cfg -k 7"
CMSD_DEFAULT_OPTIONS="-l /var/log/xrootd/cmsd.log -c /etc/xrootd/xrootd.cfg -k 7"
XROOTD_INSTANCES="default"
CMSD_INSTANCES="default"
```

/etc/xrootd/auth_file

```
g /atlas /atlas rl
```



Need for FAX (Forward) Proxy

- The FAX reverse proxy shown earlier will allow clients located within BNL and outside BNL to access all FAX data.
  - The data path is always from a data server or a reverse proxy server to a client.
- BNL's worker nodes are located behind the firewalls just like dCache data servers.
  - Accessing BNL data from BNL worker nodes will not cause the network path to cross the firewall. But, the data will go through the proxy host.
  - Accessing non-BNL data from BNL worker nodes will cause the network path to cross the firewall since the clients will directory communicate with remote xrootd services.
    - The simultaneous requests of remote data from thousands of local worker nodes can generate large network traffic, possibly over-whelming the capability of the site firewall.
- FAX proxy is necessary if worker nodes are behind the firewall!

Configuration detail in FAX (Forward) Proxy

/etc/xrootd/xrootd.cfg in Forward Proxy Server

```
all.export /atlas r/o
all.export /root:/
all.export /xroot:/
pss.origin = atlas-xrd-us.usatlas.org:1094
ofs.osslib /usr/lib64/libXrdPss.so
ofs.trace most
```

Forwarding proxy directives

Regional FAX director

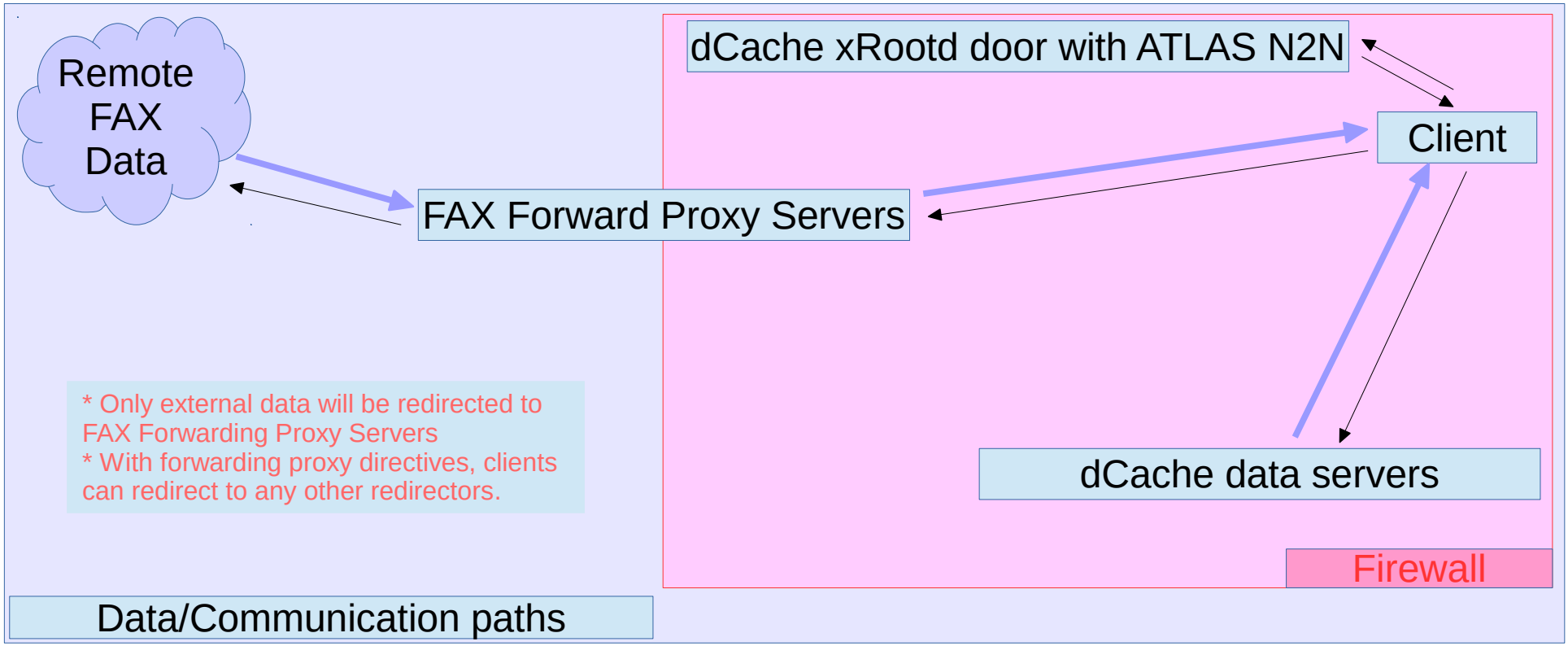
/etc/sysconfig/xrootd

```
XROOTD_USER=xrootd
XROOTD_GROUP=xrootd
export MALLOC_ARENA_MAX=4
ulimit -c unlimited
export X509_USER_PROXY=/tmp/x509up_proxy
export X509_VOMS_DIR=/etc/grid-security/vomsdir
XROOTD_DEFAULT_OPTIONS="-l /var/log/xrootd/xrootd.log -c /etc/xrootd/xrootd.cfg -k 7"
XROOTD_INSTANCES="default"
```

/etc/dcache/layout/myhost.conf in dCache xRootd Door with ATLAS N2N

```
xrootd.n2n.site=BNL-ATLAS
xrootd/xrootdPlugins=gplazma:none,authz:atlas-name-to-name-plugin,redirector
xrootd.redirector.host = XROOTD-FORWARD-PROXY-SERVER-NAME-ABOVE
xrootd.redirector.port = PORT # for FORWARD-PROXY
```

Must use xrootd version > 4.1.1
Currently on Git repo



* Only external data will be redirected to FAX Forwarding Proxy Servers
* With forwarding proxy directives, clients can redirect to any other redirectors.

Real examples of accessing files From BNL worker nodes (only!!!)

- Test files:
  - BNL file:
    - /atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M
    -
  - MWT2 file:
    - /atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M
- Test BNL dCache door with PNFS
  - Xrdcp -f root://dcxrd.usatlas.bnl.gov:1096//pnfs/usatlas.bnl.gov/BNLT0D1/rucio/user/ivukotic/f0/3c/user.ivukotic.xrootd.bnl-atlas-1M /dev/null
- Test BNL dCache door with N2N
  - Xrdcp -f root://dcxrd.usatlas.bnl.gov:1096//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M /dev/null
    - NOTE: Clients communicate directly with dCache data server.
- Test using BNL dCache door with redirection
  - Xrdcp -f root://dcxrd.usatlas.bnl.gov:1096//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null
    - NOTE: Clients will be automatically redirected to BNL xRootd proxy (shown below) when files are not found in BNL dCache.
- Test using BNL xRootd proxy without going though BNL dCache xRootd.
  - Xrdcp -f root://xrootdproxy.usatlas.bnl.gov:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null
    - NOTE: Proxy will contact North American Regional redirector for a file. Data will be proxy-ed through the server.
- Test using BNL xRootd proxy's explicit forwarding capability
  - Xrdcp -f root://xrootdproxy.usatlas.bnl.gov:1094//root://fax.mwt2.org//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null
    - NOTE: Proxy will contact the user defined (eg., fax.mwt2.org) remote redirector instead of fix one (eg. atlas-xrd-us.usatlas.org).
    - NOTE: By using the explicit forwarding, it will avoid doing stat-of-file over entire FAX storage.

Real examples of accessing files From Outside of BNL firewall

- Files
  - BNL file:
    - /atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M
  - MWT2 file:
    - /atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M
- Reading from BNL's proxy server
  - Xrdcp -f root://dcdoor11.usatlas.bnl.gov:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M /dev/null
  - Xrdcp -f root://dcdoor11.usatlas.bnl.gov:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null
- Reading from North Amercial Redirector
  - Xrdcp -f root://atlas-xrd-us.usatlas.org:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M /dev/null
  - Xrdcp -f root://atlas-xrd-us.usatlas.org:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null
- Reading from any FAX redirector (eg. North American Central Regional)
  - Xrdcp -f root://atlas-xrd-central.usatlas.org:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.bnl-atlas-1M /dev/null
  - Xrdcp -f root://atlas-xrd-central.usatlas.org:1094//atlas/rucio/user/ivukotic:user.ivukotic.xrootd.mwt2-1M /dev/null