

FAX setup at BNL

Hironori Ito
Brookhaven National Laboratory



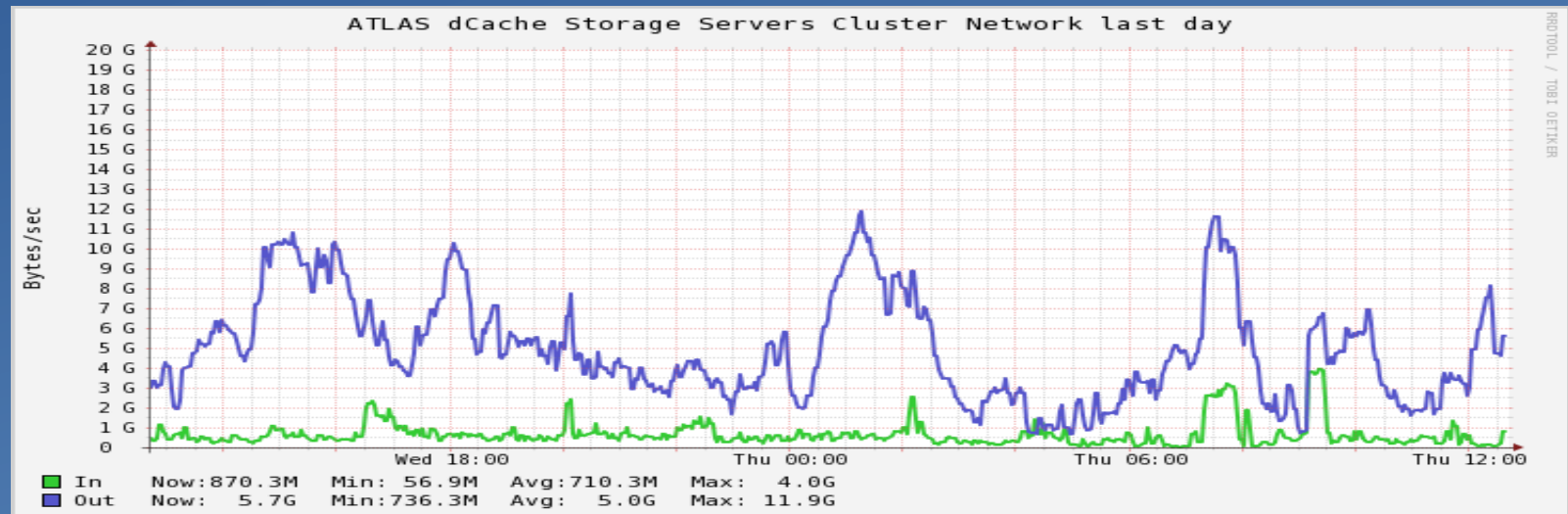
Requirement

- Storage are behind firewall
 - No direct access from WAN
- WAN throughput
 - Limited to 10-20 Gb depending on sites
 - Expect upgrade to 100 Gb in the future
- LAN
 - No firewall
 - Direct connections are allowed
 - Larger available total throughput >300 Gb
- dCache storage (2.6.x)
 - Various types of storage with total of about 12PB
 - Nexan, thor, DDN, JBOD.
 - Various performances depending on the types

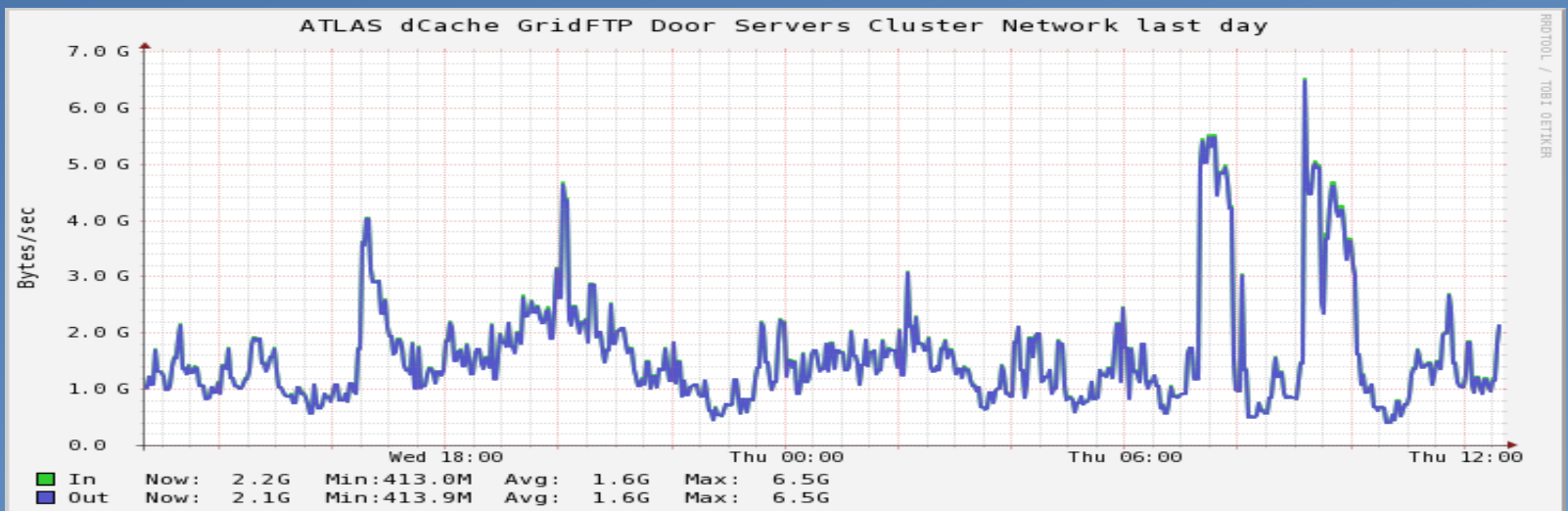
Normal load in the storage

Large variations in requested loads

LAN+WAN



WAN



Xrootd Configuration

```
all.manager proxy PROXYMASTER_HOST:1213
all.manager meta REGIONAL_GLOBAL_REDIRECTOR_HOST:1095
all.export /atlas r/o

if PROXYMASTER_HOST
  all.role proxy manager
  cms.dfs lookup central
  cms.delay servers 0
else
  all.role proxy server
  xrootd.redirect REGIONAL_GLOBAL_REDIRECTOR_HOST: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

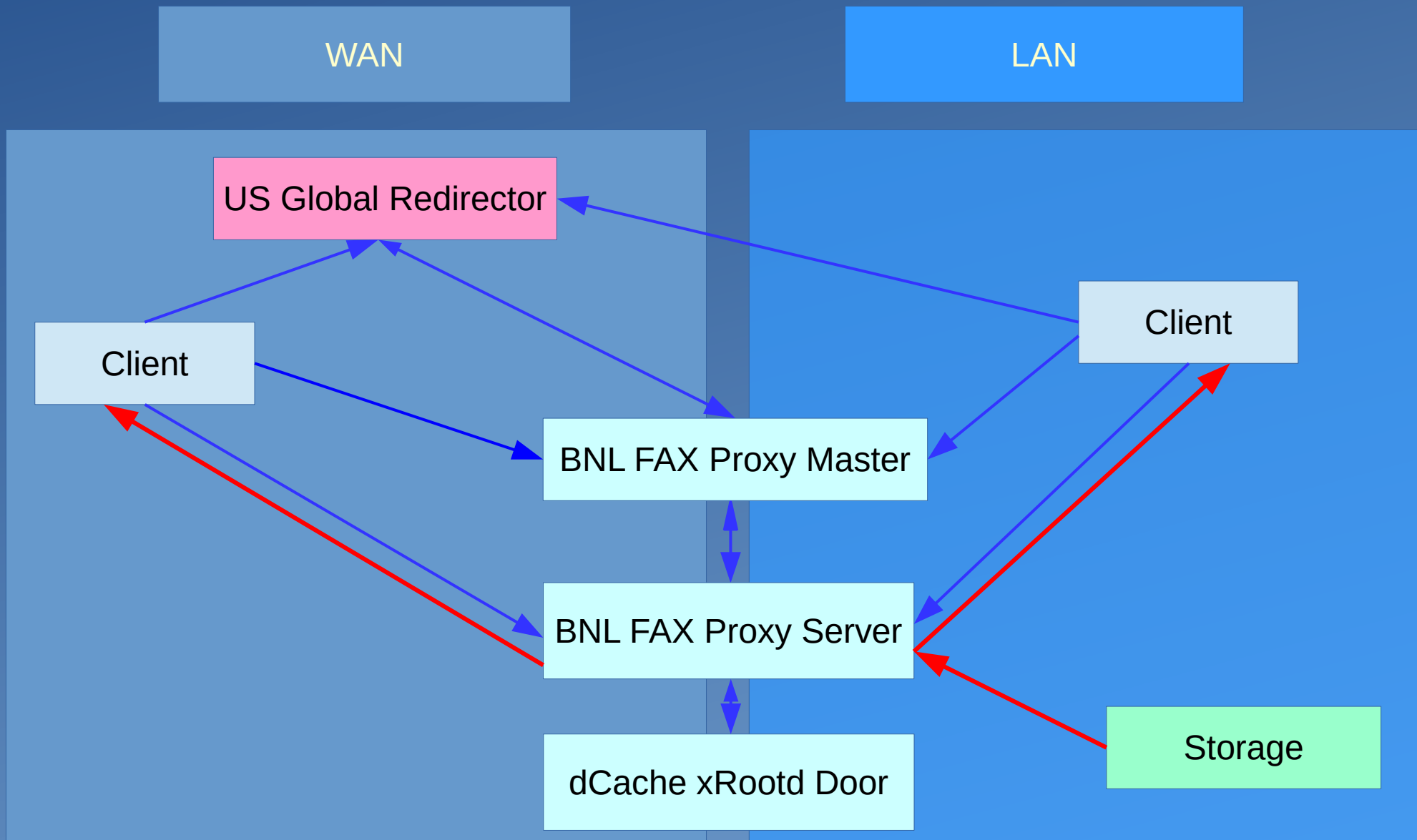
pss.origin NATIVE_DCACHE_XROOTD_HOST:1096
pss.namelib /usr/lib64/XrdOucName2NameLFC.so
sec.protparm gsi -vomfun:/usr/lib64/libXrdSecgsiVOMS.so -vomfunparms:certfmt=raw|vos=atlas|grps=/atlas
xrootd.seclib /usr/lib64/libXrdSec.so
sec.protocol /usr/lib64 gsi -ca:1 -crl:3 -gmapopt:10
acc.authdb /etc/xrootd/auth_file
acc.authrefresh 60
ofs.authorize

all.sitename MY_SITE_NAME
```

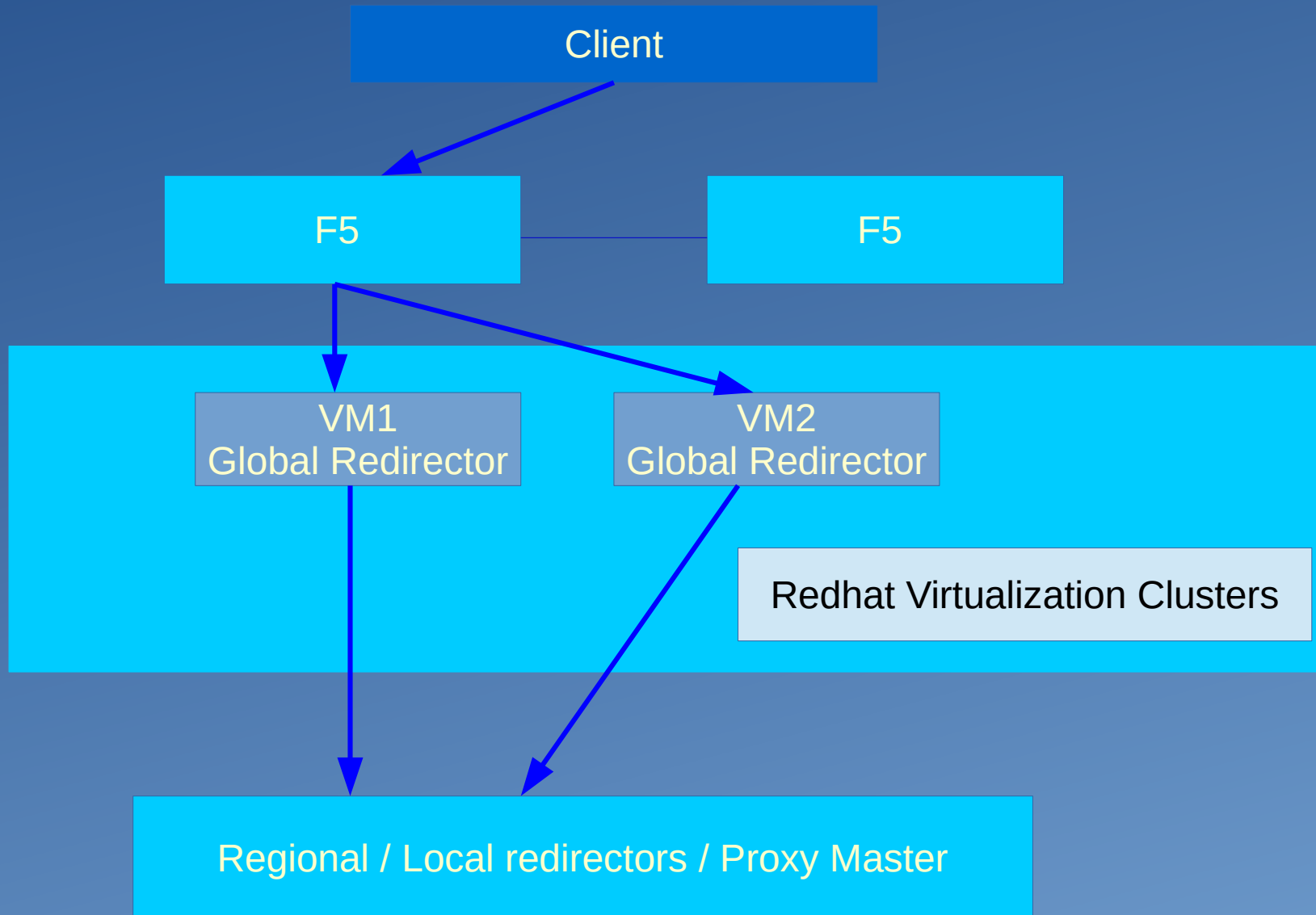
Redundancy in BNL's FAX setup

- Redhat Virtualization clusters
 - Multiple VMs for redirectors
 - Auto-restarts VMs
- Redundant F5 switch
 - Automatic disabling/enabling on dead VM hosts
 - Examine ports and services
- Multiple xRootd Proxy servers
- Multiple dCache xRood Doors

FAX Diagram

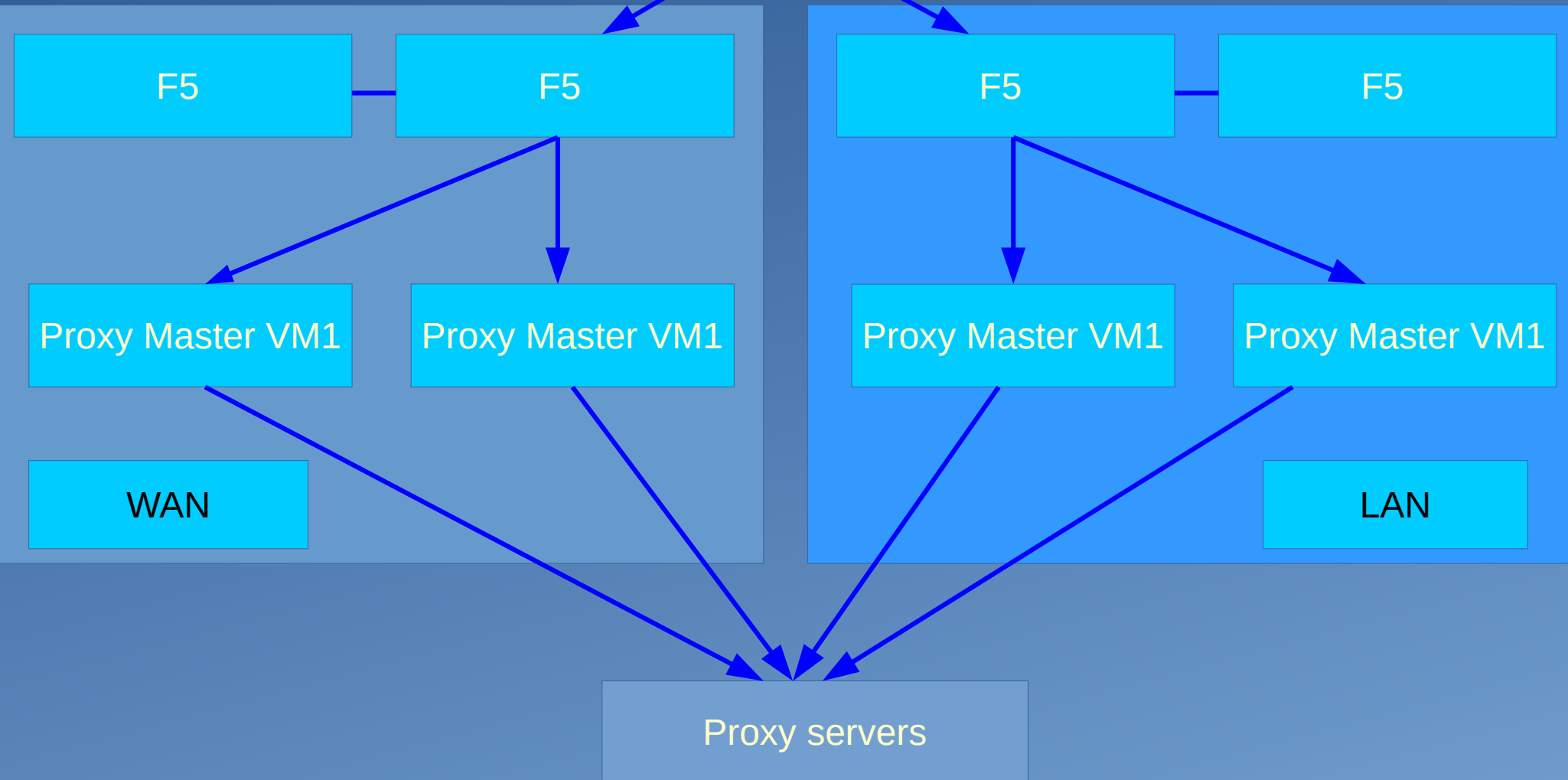


Redundancy in Global redirectors

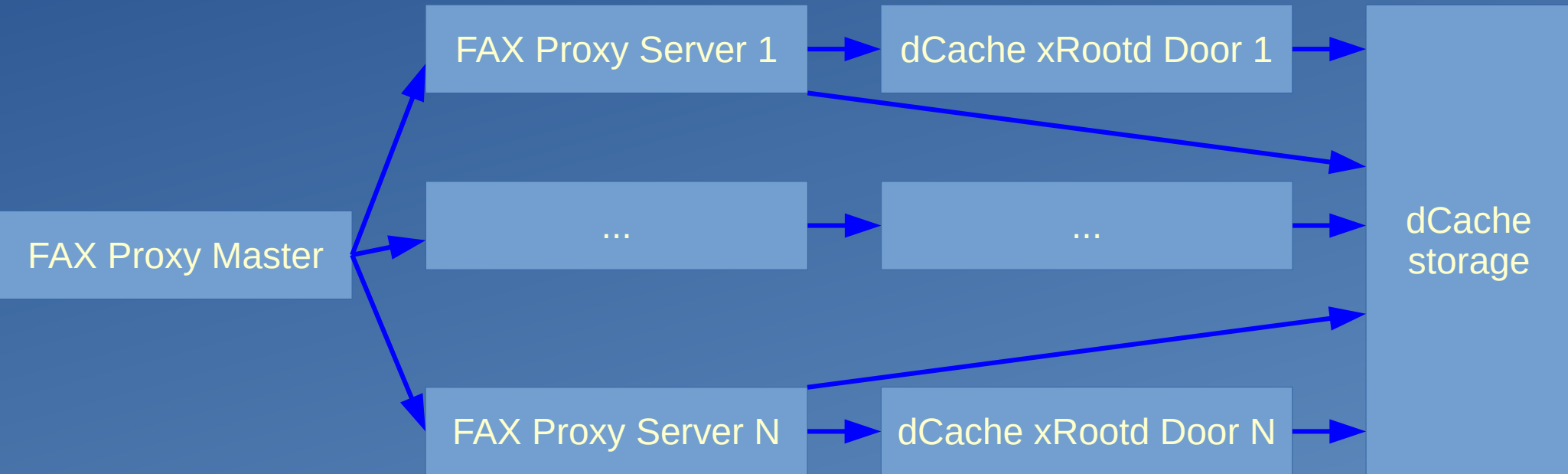


Redundancy in Proxy Master

Clients / Global Redirector



Redundancy in proxy servers and dCache xRootd doors



- Proxy server = Data Server
- Each proxy server is limited to 10 Gb
- Maximum simultaneous connections can be configured for each proxy server.
- dCache xRootd door = redirector

Configuration Management

- All configuration are managed by Puppets
 - No exceptions
- All puppet scripts are stored in Git and pulled automatically to hosts
 - BNL staff access only
 - Various checks to prevent bad commit
- Some hosts and services specific parameters are stored in GLPI
 - Eg, open specific ports in firewall, run dCache doors, etc...
- All hosts are built by Cobbler

GLPI Screen shot

Items Seen: contains dcdoor11

Display 15 items

Name	Domain	Status	Manufacturer	Model	Serial Number	Custom Fields - BNL Property Tag	Type	Operating system	Location	Last update	Technician in charge of the hardware	User
Proxy master	atlas.bnl.gov	production	IBM	IBM System x3550 -[7978B9U]-	99N1791	0	Main Server Chassis	Red Hat Enterprise Linux Workstation release 6.5 (Santiago)		2013-12-17 17:51	Jason Smith	

Uncheck All

Display 15 items

BNL Property Tag: 0

Puppet Classes:

```
base ganglia::node{cluster=atlas dcache gridftp_door_servers} ldap
auto_direct nagios::plugin{plugins-disk,nagios-plugins-
swap],script_name
tlxrootddcache atl version=2.6.18-1}
```

t1xrootddache

Puppet Parameters:

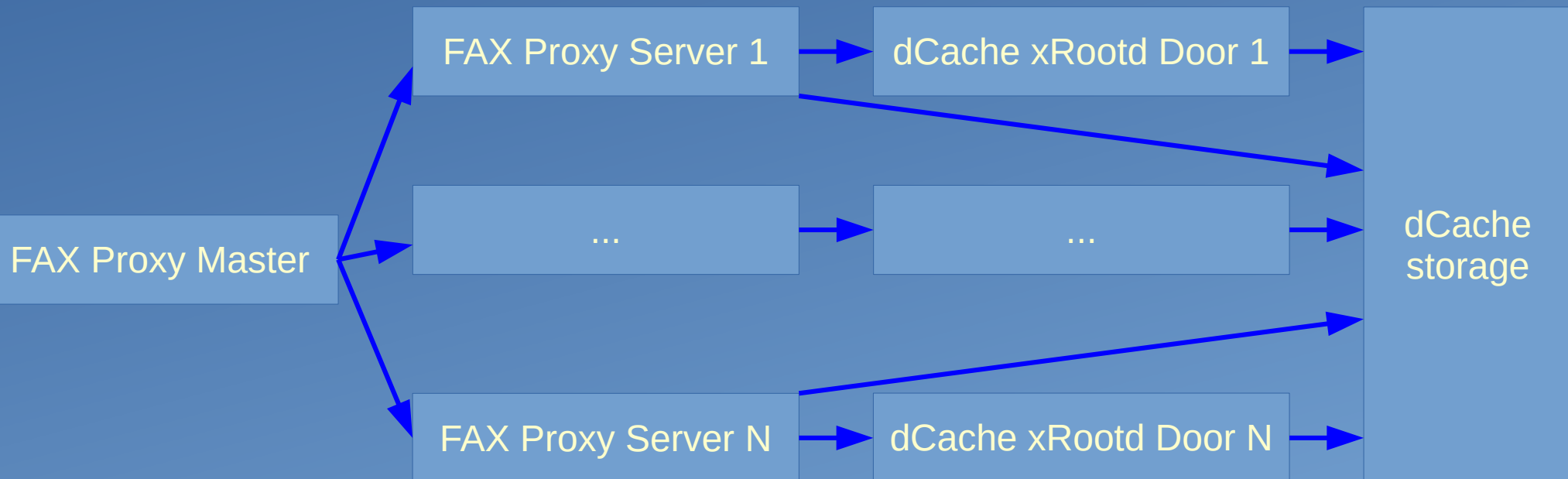
```
subnet=130.199.186
iptables_allow
iptables_no
ssh_allow_atlas
```

iptables_allow_tcp_ports=[xrood]

Update Custom Fields

Possible concern #1

- Proxy Server + dCache xRootd door = Data server for entire data of dCache
 - Different from regular data server which serves a fraction of data at a site
 - All proxy servers will answer the same.
 - Xstream copy might use all proxy servers for the same data.



Possible concern #2

- Cache turning.
 - Do you want to cache meta-data lookup?
 - If cached, proxy master might pick the dead proxy server.
 - If not cached, every request will result in name-space lookup for each proxy server. With multiple proxy servers (N), this will crease by the factor of N.
 - If every users request to the global redirector, the number of look-up will increase dramatically.
 - Might lead to an instability of name-space at a site which does not have a file.

Under Utilization of FRM with FAX

- FRM (File Resiliency Manager)
 - Utilize any file transfer tools to populate storage which understands xRootd
 - Protection against eager users executing many fetch commands
 - A single user has shut down a remote storage site
 - A single user has crashed own local storage
 - Data caching.
 - Fetch data according to the client requests
 - Populate data with any global file path
 - Delete data with threshold

Only need a few lines of scripts!