FTP server for receiving vendor data

Overview

Two commercial vendors competing for LSST business have been awarded contracts for a set of preliminary devices. Before shipment, each device will undergo a series of tests by the vendor, producing ~25 GB of data. These data must be transferred to SLAC, analyzed, archived and distributed to other LSST laboratories prior to a "pre-ship review". Only after successfully passing this review will a device be authorized for shipment. It is expected that SLAC will receive multiple data deliveries per month although they are not precisely scheduled and deliveries may be bundled.

Requirements

- Vendors must be able to control the timing of the data delivery.
  - The intranet within the company as well as their connection with the internet can be overwhelmed by these deliveries
  - Vendors insist on controlling the exact timing for the electronic data transfers
- Vendors must be able to *restart* an aborted data delivery.
  - The internet connection between SLAC and the two vendors is subject to intermittent instability.
  - The internet connection to at least one of the vendors is quite slow (100 Mbps)
  - The need to restart a large and time-consuming delivery from scratch would cause an unacceptable delay
- Vendors must be able to create, modify, or delete files in their FTP areas
- A simple solution: vendors have very limited IT expertise and are unwilling or unable to perform software installations or complex configuration changes to their systems
- The transfer buffer must be able to hold multiple data deliveries per vendor, so at least 200 GB
- LSST must do its best to prevent data from Vendor A from being visible to Vendor B, and vice versa

Proposed Solution

- LSST operated advanced FTP service
  - vsftp server software: very secure; high performance; restartable transfers; virtual ftp-only accounts
  - installed and running on LSST service VM (VM is "SCS Standard")
  - access to /nfs/farm/g/lsst/u2
- New lsst-ftp account to have ownership privs on a single NFS partition: /nfs/farm/g/lsst/u2 (which will be a short-term buffer from which a permanent archive will be made)
- Individual virtual vsftp accounts for Vendors A and B.
- This FTP area would be considered a "vendor playpen" from which copies would be archived to permanent LSST storage

Potential Security Issues and Mitigations (not complete!!)

1. Hacking into a vendor account
   a. Possible consequences
      i. loss or corruption of vendor data
      ii. use of storage for illicit purposes
      iii. interruption of vendor data deliveries
      iv. load on "u2" server (currently wain006)
   b. Possible mitigations
      i. configure vsftpd to recognize only certain IP addresses to log in
      ii. vendors must agree with the level of security and the risk
      iii. monitor disk usage with ganglia and look for abnormalities
      iv. configure vsftpd for secure userid/pwd transfer, e.g., tls
2. Hacking into the vsftp server
   a. Is this likely? This server is generally considered "very secure" as its name suggests. No hard data on this claim.
3. Hacking into the lsstlnx VM
   a. Independent of vsftp and, therefore, no different from other VMs at SLAC with externally visible ports. Server restricts login to a small set of authorized SLAC users.

Suggestions from the Cyber Security Group

1. Ask vendors to send MD5 checksum via a separate channel than FTP. Response: ask them to send it via email in their announcement message
2. Employ and IP filter (or virtual IP addresses), preferred would be to add this filter to the perimeter router. Response: request sent to net-admin

Why Existing FTP Service is Unacceptable

1. Non-anonymous (s)FTP requires a SLAC unix account and that has been deemed unacceptable by LSST project team
2. Anonymous FTP server suffers from several shortcomings:
   a. The server software cannot restart an interrupted data transfer
   b. The AFS-backed store is possibly not scalable to the hundreds of GB needed
   c. The 3-day dwell period is too risky for the data
   d. The AFS permissions combined with the 3-day dwell do not allow for the type of permissions that would allow for a convincing separation between the two vendor's data
   e. The dropbox paradigm does not allow for vendors to manage its data once at SLAC, i.e., to replace faulty data files.

Installation details

We are using vsftpd daemon running on a dedicated virtual machine. The machine is running a standard SLAC RHEL6 installation, with taylor and NFS access. Login is restricted to members of the sca-admin group.

Modifications to standard installation:

```
sudo yum install vsftpd
cd /etc/vsftpd
create file virtual_users.txt:
ITL
password1
e2v
password2
sudo db_load -T -t hash -f /etc/vsftpd/virtual_users.txt /etc/vsftpd/virtual_users.db
```

Modify standard /etc/vsftpd/vsftpd.conf as follows

```
12c12
< anonymous_enable=NO
---
> anonymous_enable=YES
96c96
< chroot_local_user=YES
---
> #chroot_local_user=YES
116a117
> pam_service_name=vsftpd
119,127d119
<
< # Virtual user setup
< guest_enable=YES
< virtual_use_local_privs=YES
< pam_service_name=vsftpd_virtual
< user_sub_token=$USER
< local_root=/nfs/farm/g/lsst/u2/$USER
< hide_ids=YES
< guest_username=lsst-ftp
```

Add a new file, /etc/pam.d/vsftpd_virtual

```
#%PAM-1.0
auth required pam_userdb.so db=/etc/vsftpd/virtual_users
account required pam_userdb.so db=/etc/vsftpd/virtual_users
session required pam_loginuid.so
```

Start vsftpd

sudo /etc/init.d/vsftpd restart