1. Current status. We currently have two SANs, EMC CX300 and AX4-5F. There are two IO servers (lynxpet and geminipet) which connect with the SANs. CX300 was out of warranty 5 years ago, and AX4-5F was out of warranty in end of January 2017. We want to move our valuable data to a system with warranty. Also current storage systems are pretty full, which makes our daily works inconvenient. And the server geminipet will become end-of-life in July 2017.

Current storage layout can be found here:

https://orionpet.med.yale.edu/IT/information/disk\_partition.htm

1. New SAN and server. We are purchasing a new SAN (Dell PowerVault SCv2080) and a new server (Dell PowerEdge R730). The order had been approved and sent out to Dell. Dell is producing them, and the estimated delivery date is February 20, 2017. The actual receiving date can be several dates later.
2. More details about SCv2080. This is 5U system, and we only use half of its space. It contains 3 drawers. Each drawer contains 14 disks. Each disk is 4T raw size. According Dell recommendation, we will use RAID 6 with hot spare. The initial plan is following.
   1. Each draw (14 disks) will be one group (volume, pool). One disk for hot spare, and RAID 6 uses 2 disks for parity on average. This gives us 11 disks for storage, which will become a large pool (volume). On this pool, several logic disks (LUNs) will be allocated. The logic is to allocate all logical disks in same pool for same IO server. In other words, all logical disks in same pool will be assigned to one server.
   2. The initial plan is to move all current partitions from CX300 and AX4-5F to new SCv2080, and size of all partitions will be at least doubled.
   3. Drawer 1 is about 44TB raw space. It will be divided into data1 (10TB), data8 (24TB), home1(10TB). All 3 disks will be served by lynxpet.
   4. Drawer 2 is about 44TB raw space. It will be divided into home2(10TB), data2(20TB), data3(14TB). All 3 disks will be served by geminipet.
   5. Drawer 3 is about 44TB raw space. It will be divided into data4 (22TB), data5(22TB). All 2 disks will be served by geminipet.
3. The new coming R730 server will become geminipet. It’s necessary for archive purpose because geminipet has been archive client for many years. The current geminipet will have a new name (Kathryn is thinking about it.)
4. CX300 will not be used because it’s too old and size is too small. AX4-5F will still be used, though it’s out of warranty. We will handle all problems and risks by ourselves, such as disk failure. It will be served by current geminipet, which also be out of warranty soon. My plan is not plug these two components into FC switches, which makes administration easier.
5. Known problems. EMC storage and Dell storage use different protocols. This makes them not able to co-exist in same server. EMC uses its own PowerPath, which we have been using since beginning. Dell uses Linux MultiPath, which is built-in Linux kernel. To use PowerPath in Linux, we must disable MultiPath and rebuild a new kernel without MultiPath, then install PowerPath and configure all storage. This is status currently in lynxpet and geminipet.
6. Note on user home. We will divide our users into two groups. One is for daily operation and production. This user group usually doesn’t need a lot of space and doesn’t run IO-intense works. But they need quick system response. This user group will go to home1. Another is for research users. This user group usually needs more space and often runs IO-intense works. They will go to home2. The purpose is to separate user IO so that they will not conflict with each other.
7. Pools in AX4-5F will be destroyed and re-created into 3 large pools. Each pool (RAID 5 with hot spare) will have one single logical disk. Their sizes will be 10TB, 20TB, 20TB. We have no plan on how to use them. Anyone is welcome to provide suggestions.
8. Data migration. This can become complicated because of our tight schedule and large data volume. It can only be done during weekend, and it will cause some disks offline. This can be done in multiple steps. The initial plan is following.
   1. Assume we can receive new server and storage at about February 24. Then I need some time to play around the new system and get familiar with them. The new R730 server will be given a new name (such as tttpet), and its own IPs. Linux system 6.5 will be installed, same as other servers. But PowerPath will not be installed. Instead, we will use MultiPath to communicate with SCv2080. The SCv2080 will be configured as above. All logical disks will be assigned to tttpet temporarily. The tttpet and SCv2080 will be connected via FC directly, not through FC switch. Only two of the four FC ports in SCv2080 will be used. At this moment, we are not worry about LoadBalance and Failover, because there is almost no IO at all.
   2. After one week, if step 1 goes well, we will transfer all data currently in CX300 and served by geminipet to SCv2080. This is scheduled on March 4-5. During this weekend, raw data and recon data will become unavailable.
   3. When step 2 is finished, we will switch geminipet and tttpet, i.e. switch their name and IPs. Then unplug CX300 and old geminipet from FC switch, and plug new geminpet and SCv2080 into FC switch. This step needs configuration in two FC switch. From this moment, new server and SAN are online. Users will not notice the difference because the server and all disks remain the same names. But I need more works on this new server, such as Samba setup, archive/backup setup, etc.
   4. In the next weekend (March 11-12), we will do another half. This will cause all system offline. We will move all disks currently served by lynxpet to new SAN. Then lynxpet needs some extra works. We will remove PowerPath from its kernel, and use MultiPath in its kernel. Without this step, lynxpet cannot talk with SCv2080.
   5. If step 9.4 works smoothly, we will go to SCv2080 and assign some logic disks to lynxpet. The remaining FC ports in SCv2080 will also be plugged into FC switch. At this moment, we should be all set. All disks and servers will become online again. All storage and servers have the ability of load-balance and failover. If any steps go wrong, we can rollback and still use AX4-5F. This is the most difficult step because lynxpet is unique and cannot be replaced by others. All NIS information is in lynxpet.
   6. The CX300, AX4-5F, old geminipet (at this moment has name of tttpet) will remain untouched for at least 1-2 weeks in case that any problem happens in above steps. If no problem, the CX300 will be discarded (recycled). And AX4-5F will be re-configured for other user. It will be served by tttpet.
9. Any suggestions or concerns are welcome. I maybe need one or two strong men/women to help me with these heavy boxes.