

# “BIONANOSPECTO” Biotech INC.

The Database and Network  
Architecture 5 Year Plan

# Problem

- Biotech company is moving into an empty office space where it will set up lab and administrative spaces. It has plenty of in-house expertise for choosing appropriate scientific equipment but none implementing a database solution and very little implementing the supporting IT framework across which data will travel. Whatever architecture is put into place now will have to last 5 years without significant overhaul.

# IT infrastructure needs

- 30 – 50 computers
- Database application and server
- Printers, fax machines, scientific equipment
- Network to connect all this together
- High data throughput resulting from data feeds from scientific equipment and data analysis.

# Proposed Solution

- A multi-mode fiber optic backbone for distance < 3000 feet from computer to server. “Multimode fiber gives you high bandwidth at high speeds (10 to 100MBS - Gigabit to 275m to 2km) over medium distances” (source: <http://www.arcelect.com/fibercable.htm>)
- 30 computers outfitted with IEEE 802.3z standard network interface cards. These will be installed on every computer on the company’s network and a fiber optic cable run between each card and a switch.
- A switch will act like a relay, passing information between computers using the MAC addresses that are assigned by the network cards.
- A router will connect to the switch and will enable the company’s LAN to communicate with external networks (aka ‘the internet’).
- Router behind a firewall and computers behind security software.
- SQL Server 2008, database management software, hosted on on RAID 10 or RAID 5 Server running on a Storage Area Network (SAN).

# Mission Alignment

To have a product ready in 5 years.

Response:

The network and servers will need to support today's and high speed, high volume data throughput needs for the whole of 5 years without interruption.

Multi-mode fiber optic lines can carry 1 Gb of information/second. UCLA uses a 1 Gb network....we are a smaller operation.

# Organizational Fit

Go with what is proven to work and be good stewards of venture capital.

Response:

Focus on proven technology.

The equipment used for communications over multi-mode optical fiber is proven technology. “Because of its high capacity and reliability, multi-mode optical fiber generally is used for backbone application in buildings” (Source: Wikipedia: Multi-mode optical fiber)

# User Access

100% high availability needed during typical work day defined as 07:00 – 19:00.

Response:

- Best practice for SQL Server to be 100% available is to have it running on fault tolerant servers. RAID 5 and RAID 1+0 disc arrays ensure that data is written twice. In the event of a hard disc failure the faulty disc can be replaced with no loss of data.
- Properly configured data will ensure faster reads and writes and quick recoverability as well. If OS is saved to a separate hard drive from data, transaction, and back-up logs, then concurrent reads and writes can occur with the least amount of disc contention.
- Automated backup procedures will ensure that corrupt data does not destroy more than 5-10 minutes of data.

# Users to be able to do their Jobs

Research assistants (RAs) and scientists cannot be waiting around for scientific equipment to finish running assays on time sensitive cultures and chemical reactions. The ability to run a set of tests must not be slowed down because of a network that is trying to catch up with the data output. RAs and scientists should not have to learn SQL Query language to issues queries against database.

Response: fiber optic network can handle 1Gb of through put per second. If needed we will dedicate one network to the scientific equipment and use a secondary (copper network for communication and admin work) to eliminate competition with other data traffic. DBA will need to get specifications for each scientific instrument that will be acquired in the labs. Specs to include maximum data output it is capable of when working at max capacity.

Third party software can be purchased for scientific data analysis. Knowledge of SQL Server is not needed to issues queries against database.

# Security of Organization Resources

Offsite backups of intellectual property,  
marketing data, human resources information,  
etc.

Response:

Backups to tape which are then removed off-  
site.

# Privacy Issues

Company needs to protect the identities of clinical sample's donors and company's unpublished work.

Response:

Database wide and column specific encryption is available with SQL Server 2008.

Database will not have any open ports on it's servers that are not specifically listening for data from internal sources of data.

Routers will be behind a firewall.

Computers will be running anti-virus and security software.

# ROI

Installation must support data needs 5 years out so that no reinvestment has to be made during this time into the IT infrastructure. This would be an unwelcome distraction.

Response: The need for precise sizing of database for current and future needs can be mitigated with the installation of a SAN where stored data does not reside directly on any of a network's servers. Storage can be sized and resized as needed for growth. Return on investment is thus realized immediately as no more storage is purchased than is needed.