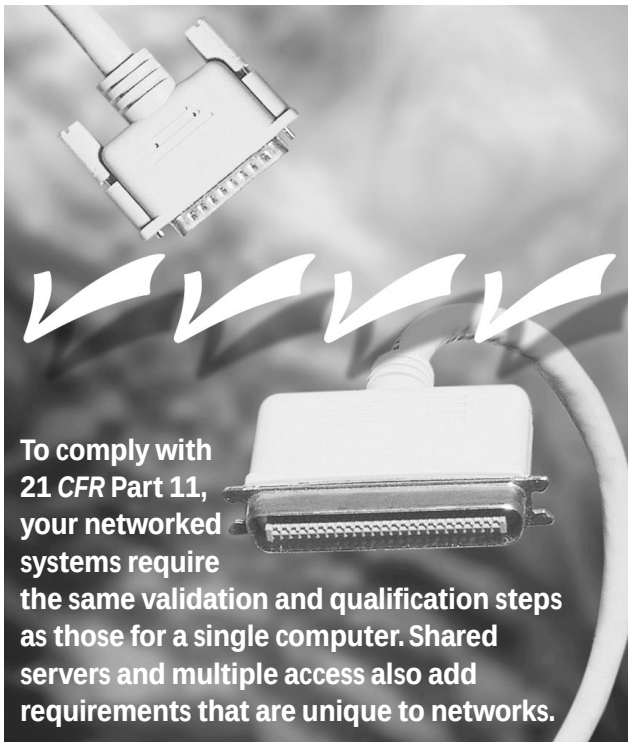
Qualification of Network Components and Validation of Networked Systems, Part I

Ludwig Huber* and Rory Budihandojo



ARTVILLE

Ludwig Huber is worldwide product manager for pharmaceutical solutions at Agilent Technologies GmbH, PO Box 1280 D-76337, Waldbronn, Germany, +49 7243 602 209, fax +49 7243 602 501, ludwig_huber@agilent.com, www.agilent.com.
Rory Budihandojo is head of R&D for IT quality and testing at the Centre of Excellence, GlaxoSmithKline, Collegeville, PA 19426.

*To whom all correspondence should be addressed.

In Part I of this article, the authors discuss FDA expectations, specific requirements, and validation plans of networked systems and include a glossary of related terms. Part II will address the types of specifications, the installation, testing, and implementation of required validation for networks.

Networked systems with integrated or distributed databases increasingly are used in the pharmaceutical industry, and like all computerized systems, they must be qualified and validated to demonstrate their suitability for their intended use. Although the validation of stand-alone computer systems is well described (1), many companies still are uncertain about how to qualify networks and networked systems. But FDA is increasingly looking into such systems, as evidenced by recent warning letters and inspection reports. Networked systems and their applications must be validated for compliance reasons, but that validation also is important for business reasons. Missing data in an electronic batch record or laboratory information system or lost data from a research project can be disastrous for a company and its employees. Production delays caused by network failures also are costly.

We will highlight the qualification of network components (switches, hubs, routers, software) and the validation of networked systems, assuming that readers already are familiar with the principles of computer system validation and with network technology. (Reference 1 is a good source for beginners on the subject of computer validation; network technology and terminology can be found in Reference 2 and similar books.)

Network quality assurance has been addressed by Crosson, Campbell, and Noonan who recommend that a network be qualified (because it is a piece of equipment) and then managed through documented control (3). Olthof discussed information technology (IT) quality in a paper at the ECA conference (4). A special interest group at the Good Automated Manufacturing Practice (GAMP) forum (5) emphasized that quality assurance principles are critical to the management of the IT infrastructure. That group recommended bringing IT infrastructure into initial compliance with established standards through planned qualification processes. Once in compliance, the infrastructure should be maintained using documented standard processes and quality assurance activities. The effectiveness of the program should be audited periodically.

Ongoing updates about recent developments in network validation and compliance can be found on the Web sites of FDA (www.fda.gov), the GAMP forum (www.gamp.org), and PDA (www.pda.org) as well as some private sites such as www.labcompliance.com and www.computervalidation.com. Our objective is to provide practical recommendations for qualifying individual network components and validating networked systems (as part of the validation of the applications that are supported by the network).

FDA expectations

FDA is inspecting networked systems and has issued related warning letters and 483s or inspectional observation reports. Studying such letters and reports is instructive because they highlight what inspectors are looking for and what mistakes others are making. The following excerpts are from warning letters available publicly on the FDA Web site (Two network-oriented warning letters issued this year contain information for laboratory information management systems and stability test programs. The second letter also contains information on databases):

> The network program lacked adequate validation and/or documentation controls. For example:
>
> - The system design documentation has not been maintained or updated throughout the life of the . . . software dating back to 1985 despite significant changes and modification[s] that have taken place. These include program code, functional/structural design, diagrams, specifications, and text descriptions of other programs that interfere with [this program].
> - The software validation documentation failed to adequately define, update, and control significant elements customized to configure the system for the specific needs of the operations.
> - Validation documentation failed to include complete and updated design documentation, and complete wiring/network diagrams to identify all computers and devices connected to the . . . system.
> - The QCU [Quality Control Unit] failed to ensure that adequate procedures were put in place to define and control computerized production operations, equipment qualifications, documentation review, and laboratory operations (6).
>
> The . . . computer system that is accessed by personnel from various departments to include manufacturing, testing laboratory, and Quality Assurance lacked the following:
>
> - Audit trail function of the database to ensure against possible deletion and loss of records.
> - Absence of documentation defining the database, operating system, location of files, and security access to database.
>
> Your response fails to discuss extending the retrospective evaluation to other elements of the system needing to be defined and controlled as part of the overall configuration management (7).

An FDA 483 from July 2000 cited a company for insufficiently documented training records:

> There are no records to document that the Information Technology (IT) service provider staff personnel have received training that includes current good manufacturing practice regulations and written procedures referred to by the regulations (8).

FDA's warnings and inspection reports repeatedly emphasize controlled updates, a focus of this article. Additional examples and updates of extracts from warning letters related to computer and network compliance can be found at the Labcompliance Web site (9).

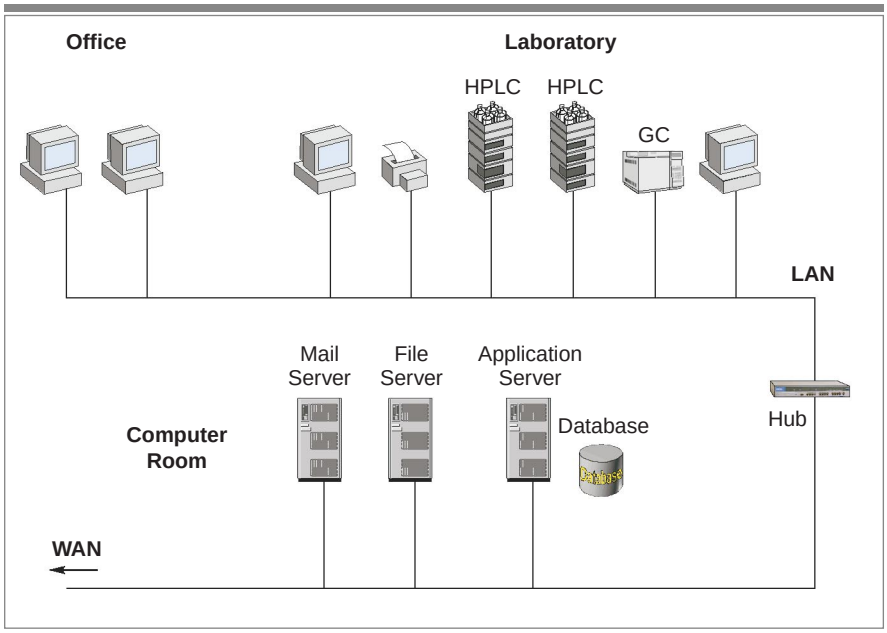


Figure 1: Example of client/server networked system (4).

What to validate

Networks are systems connecting several computers and peripheral devices. The main purpose of a network is to transport and control data traffic. Whether data are stored on the network servers or elsewhere, a network must provide assurance that data integrity and security have been maintained during the transport and traffic control functions. Data integrity and security are ensured by controlled and properly managed network access and by appropriate security for data stored within the network.

Computer systems used in regulated environments must be qualified and validated to demonstrate suitability for their intended use. That means all systems (including networks) used for work regulated by good laboratory practices, good clinical practices, and good manufacturing practices (all can be referenced as GxPs)

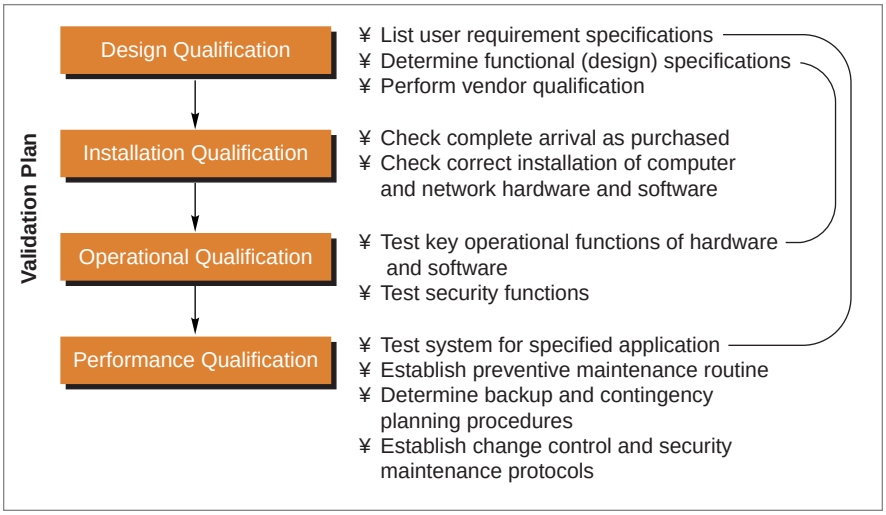


Figure 2: The 4Q model qualification phases of networked systems.

must be validated. The electronic record regulations in 21 *CFR* Part 11 give more detail about which computer systems are regulated: all systems that create, modify, maintain, archive, retrieve, or distribute electronic records (10). To ensure compliance, we recommend that you analyze all information that FDA may request during an inspection. If documents or data ever went through a computer or other device with the possibility of being modified, then that computer or device should be validated.

Validation is therefore necessary for computers that acquire and evaluate critical data from measurement systems in plant control and also for office computers that generate reports submitted to FDA. Word processing systems that generate standard operating procedures (SOPs) also must be validated; and frequently, the computers running them are connected through networks. In addition to meeting regulatory requirements, it is simply good business practice to validate all computer systems used for generating and evaluating critical data.

Networked systems

The diagram in Figure 1 shows a typical client/server networked system connecting client computers in a laboratory and offices to a server located in a computer room. The computer room also hosts mail servers. Laboratory computers with data system applications software acquire data using TCP/IP protocols and control equipment with built-in local area network (LAN) cards. Application software on the client computers also is used for data evaluation. Computers are connected to servers through a hub. Each server uses a relational database (such as that from Oracle, www.oracle.com) with customized applications for data management; for control charting and other statistical evaluation; for review, backup, archiving, and retrieval of data; and for generating electronic signatures compliant with 21 *CFR* Part 11.

Other examples of networked systems frequently used in the biopharmaceutical industry include enterprise asset management (EAM) systems, manufacturing resource planning (MRP) systems, manufacturing execution systems (MES) with electronic batch record functionality, and electronic document management systems (EDMS). The arrangement can be the same for these systems as shown in Figure 1; validation requirements can be the same as well.

Validating a networked system requires qualifying its individual components (such as the applications running on each computer) and authorized access to the system as well as qualifying data transfer between the related computers (that is, the interfaces of the components at both sites). The whole system, including the network itself, is validated by running typical daily applications under normal and worst-case conditions and then verifying that the system and its functions are meeting previously specified criteria.

For qualifying the components and for validating the complete system, it is important to define a validation box. The goal of a validation box is achieved by subdividing the network into subnetworks (or sub-LANs) containing network components that are used by each application. A validation box helps define which parts of the complete network must be qualified and which are unaffected. The validation box for the laboratory data system in Figure 1 would include the lab computers, the file server, the applications server, and the database. Limiting the network qualification tasks to those components used by the network applications saves time.

The 4Q model

Validation of networked systems should, in principle, follow the validation practices of all other computer systems. Everything that is important in validating a single computer also is important in validating a network. Network validation activities should follow a validation plan. If such a plan already exists for other components, then the validation of network-specific tasks should be added to the validation plan. A networked component should be treated like any other piece of equipment that is installed and qualified.

A network component should be treated like a piece of equipment that is installed and qualified (for example, chromatography software functions such as peak integration and quantitation). Typical network functions such as limited access and network transactions should be qualified. Because of the complex nature of a network, a cross-functional team should control validation activities. For the validation of the network, any structured approach (such as a life-cycle model) should be followed (see Figure 2). It involves design, installation, operational, and performance qualification.

Design qualification (DQ). DQ is the first step, ensuring that the design of a network meets the user's requirements. In this phase, the user requirements for each function are specified. For example, a user requirement could state, "There should be limited access to the networked system." The required function to ensure that requirement could be stated, "There should be user ID and password entry fields when entering the system." The computer system vendor should be qualified during the DQ phase.

Installation qualification (IQ). IQ is the second phase. An indi-

Glossary

Bus. An electronic pathway along which signals are sent from one part of a computer to another. A PC contains several buses, each used for a different purpose. The address bus allocates memory addresses. A data bus carries data between the processor and the memory. The control bus carries signals from the control unit.

Checksum. A record of the number of bits transmitted and included with a transmission so that the receiving program can determine whether the same number of bits arrived. If the counts match, it's assumed that the complete transmission was received.

Client/server. A network architecture in which each computer or process on the network is either a client or a server. Servers are powerful computers or processors dedicated to managing disk drives (file servers), printers (print servers), or network traffic (network servers). Clients (PCs or workstations on which users run applications) rely on servers for resources such as files, devices, and even processing power.

Data flow. Movement of information between clients and servers that is tracked to ensure accuracy and security.

Data system applications software. The software that controls equipment, such as chromatographs, and acquires, evaluates, prints, and stores data.

Distributed databases. Computing is said to be "distributed" when the programming and the data that computers work on are spread among more than one computer, usually over a network.

Electronic document management system (EDMS). A system for tracking and locating electronic documents and for managing them throughout their life cycle.

Enterprise asset management (EAM). Knowledge within a company exists in many forms: in databases, knowledge bases, filing cabinets, and peoples' heads. All too often one part of an enterprise repeats the work of another part simply because that knowledge is poorly tracked. EAMs allow companies to manage legacy and object components, inventorying assets (what they are, where they are located, and how they are used).

Fault tolerance. The ability of a system to respond gracefully to unexpected hardware or software failures. The lowest level of fault tolerance is an ability to continue operation in the event of a power failure. Many fault-tolerant computer systems mirror all operations — that is, perform each on two or more duplicate systems — so that if one fails the other can take over.

File transfer protocol (FTP). The TCP/IP Internet protocol used when transferring single or multiple files from one computer to another.

GPIB-IEEE. A general-purpose interface bus standard from the Institute of Electrical and Electronic Engineers, which develops standards for computers and the electronics industry. This standard allows as many as 15 intelligent devices to share a single bus, with the slowest device participating in the control and data transfer handshakes to drive the speed of the transaction.

GxP. All of the regulations that apply to good laboratory practices, good clinical practices, and good manufacturing practices, taken as a whole.

Handshake. Requires the recipient of a data record to acknowledge to the sender that the record has been received.

Hash algorithms (hash values). A hash value is an algorithmic method. Sometimes called the "digest" of a document in digital form, a number is generated from a string of text. The hash is substantially smaller than the text itself, generated by a formula that makes it extremely unlikely for some other text to produce the same value.

Hashes are used in security systems to ensure that transmitted messages have not been tampered with. The sender generates a hash of the message, encrypts it, and sends it with the message itself. The recipient then decrypts both the message and the hash, produces another hash from the received message, and compares the two. If they're the same, it is highly probable that the message was transmitted intact.

Hot site. A site designated to operate a network if the normal operation center fails (for example, in case of a natural disaster or fire).

Hub. A common connection point for devices in a network, such as a LAN. A hub contains multiple ports. When a packet of data arrives at one port, it is copied to the other ports so that all segments of the LAN can see all packets. A passive hub

vidual checks whether an instrument arrives as purchased, installs network components, and completes the necessary documentation.

Operational qualification (OQ). OQ is the third step, when critical key functions are tested. Testing always should follow a test plan and be compared against previously specified acceptance criteria.

Performance qualification (PQ). PQ is the last phase, which includes testing the entire system for specific application performance. PQ could involve a complete analysis by sample equipment for specific hardware, accessories, and software. It also includes preventive maintenance. For example, PQ would include both regular disk maintenance and change control.

Specific requirements

Network computer systems have some specific characteristics that differ from stand-alone equipment and must be addressed during validation. Unlike stand-alone computer systems, which consist of homogeneous hardware and software, networks are heterogeneous. They usually include a variety of hardware components, several software applications, and communication protocols. A change to one component can influence many other components and applications.

Cabling designs and specifications are as important as the hardware and software in a networked system — mainly because network components can be far away from each other. Many people and departments often access the network as a common resource, so security issues are quite important. Networks can include both components that must comply with regulations and those that aren't regulated. IT personnel have not always been trained in the GxPs.

Validation plans and teams

Validation master plans are not required by regulation, but FDA inspectors may ask for an explanation of your company's approach toward validation. The master plan is a good tool for demonstrating that approach, and plans should be available for both multisite and single-site companies. Validation master plans help ensure consistent and efficient implementation of validation throughout a site and throughout a company. If already available, such plans can be extended easily to include networks and networked systems. We recommend starting with a generic plan and adding network specifics, for example, include network terms in the glossary.

Generic network specifications (such as cabling, security, and vendor qualification) should be part of the master plan. It should

Glossary — continued

serves simply as a conduit for data, enabling it to go from one device (or segment) to another. So-called intelligent hubs (or manageable hubs) include additional features that enable an administrator to monitor traffic passing through the hub and configure each port. A third type of hub, called a switching hub, actually reads the destination address of each packet and forwards it to the correct port.

Information technology (IT). The broad area concerned with all aspects of managing and processing electronic and computerized information. Some companies refer to the department as information services (IS) or management information services (MIS).

Integrated databases. Databases that have two or more components merged together into a single system. Increasingly, the term "integrated" is reserved for software that combines word processing, database management, spreadsheet functions, and communications into a single package.

Local-area networks (LANs). Networks with computers geographically close together (in the same building). Wide-area networks (WANs) have computers farther apart, connected by telephone lines or radio waves.

Mail server. A mail server handles incoming and outgoing e-mail for Internet users. Most mainframes, minicomputers, and computer networks have an e-mail system. Some electronic-mail systems are confined to a single computer system or network, but most have gateways to others, enabling users to send electronic mail anywhere in the world.

Manufacturing execution system (MES). A system that delivers information on plant production activities. MES programs guide, initiate, respond to, and report on plant activities as they occur, resulting in rapid response to changing conditions.

Manufacturing resource planning systems (MRP). Production tracking systems used primarily in the 1980s, which are now being primarily replaced by EAMs.

MD5. A digital signature algorithm used to verify data integrity that is claimed to be as unique to that specific data as a fingerprint. Developed by Ronald L. Rivest of MIT, MD5 creates a digital signature, requiring that large files be compressed by a secure method before being encrypted with a secret key. MD5 is a one-way hash function, meaning that it takes a message and converts it into a fixed string of digits, also called a message digest. When using a one-way hash function, one can compare a calculated message digest with the message digest that is decrypted with a public key to verify that the message hasn't been tampered with. This comparison is called a hashcheck.

Networked systems (networks). A group of two or more computer systems (hardware, software, and peripherals) linked together. Computers on a network are sometimes called nodes. Computers and devices that allocate resources for a network are called servers.

Node. Any device attached to the network that is capable of communicating with other network devices. A node can be a computer or some other device, such as a printer. Every node has a unique network address.

RAID. Short for redundant array of independent (originally inexpensive) disks, RAID is a way of storing the same data in different places (thus, redundantly) on multiple hard disks. By placing data on multiple disks, input and output operations can overlap in balance, improving performance. Because multiple disks increase the mean time between failure (MTBF), storing data redundantly also increases fault tolerance.

Relational database. A type of database that stores data in related tables. A relational database is powerful because it doesn't assume how data are related or how they will be extracted. As a result, the same database can be viewed in many different ways.

Routers. Devices that connect any number of LANs and use headers and forwarding tables to determine where packets go. Routers use ICMP, an extension of IP, to communicate with each other and configure the best route between any two hosts. Very little data filtering is done through routers.

Submasking. Masks are filters that selectively include or exclude certain values. For example, when defining a database field, it is possible to assign a mask that indicates what sort of value the field should hold. Values that do not conform to that mask cannot be entered. Masks are hierarchical, and submasks are filters within filters.

Subnetworks (sub-LANS). Within a network, subnetworks are another name for nodes.

Switch. In networks, a device that filters and forwards a packet to its next destination. A switch also may include the function of a router in a generally simpler and faster mechanism.

TCP/IP. Transmission control protocol/Internet protocol, enables devices to exchange information over a network.

Validation box. Defines a set of network components that are required on the network — for example, a networked chromatography data system.

include recommendations for backup, contingency planning, disaster recovery, change control, validation reports, and archiving. The plan also should include naming conventions, which make it easier to identify components and track data flow within a network. Templates for daily operations should be included as appendices for consistent implementation, and reference should be made to existing SOPs. The master plan will be a good foundation for individual project validation plans.

Validation teams can coordinate validation activities for networked systems. The complexity of networks requires more than one expert for definitions, qualifications, and (most important) change control. The validation team should include expert IT professionals. They can best describe what might go wrong with a system and how individual network components can affect each other.

Laboratory personnel (or others who will use the network) should be part of the team because they should be aware of possible problems at the application level. End users also should be able to determine whether a network continues to operate effectively after the validation activities are complete.

QA personnel should be part of the validation team to ensure that documentation, control, and use are in compliance with regulations and company policies. The software engineering department should be involved if all or part of the software has been developed in house. Otherwise, vendor representatives can be included. Consultants can be brought into the team if necessary; they can be a great help with initial, big network validation projects.

Part II of this article, including references, will be published in *Pharmaceutical Technology* in November. **PT**