Considerations for Planning and Implementing ChemStation Plus Networked Data Systems in Laboratories Regulated by FDA 21 CFR Part 11

Application

Introduction

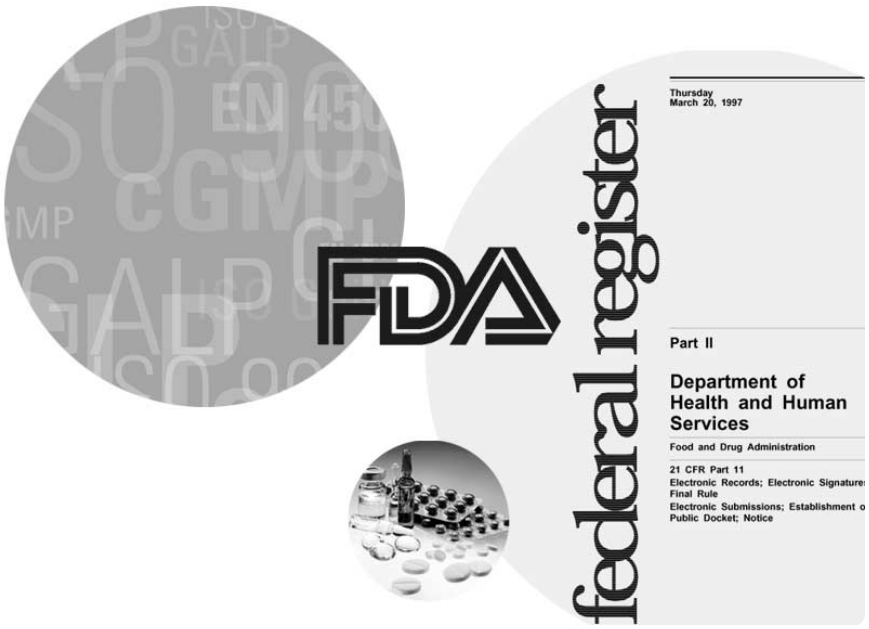
As analytical laboratories implement data systems to meet the Food and Drug Administration's requirements for electronic signatures and electronic records, 21 CFR Part 11, they find their vendors proposing different products, combinations of products, or complete system configurations. The selection of a data system reaches beyond meeting the requirements of the regulation because the usability, manageability, reliability and operating costs of the laboratory must be considered.

Networked chromatographic data systems have an impact throughout the enterprise, involving not only laboratory management but also IT, quality assurance and senior management in the decision-making process.

This Application Note discusses both general considerations for data systems operating in a 21 CFR Part 11 compliant environment and specific considerations that affect configuration decisions for Agilent networked data systems.

The information here is not intended to provide an exhaustive analysis but to cover four key areas

- data security
- data integrity and traceability
- file management, and
- analytical results management.



Agilent Technologies

Data Security

The trade press brings a constant stream of stories about computer security threats and tools for combating unauthorized system access. Computer security experts have built a variety of tools to create secure systems and operating systems are constantly improving security and security management.

21 CFR Part 11 and analytical instruments bring some special considerations to software security. Microsoft® Windows® NT® security is built around the concept of a single user owning a computing session. When another user wants to use the same PC, they either continue using the same identity or the first user must close all applications and log off so that the next user can log on.

21 CFR Part 11 does not allow users to operate using each other's identity. However, laboratory space and cost constraints make it desirable for several users to operate more than one instrument from the same PC, see figure 1.

Automated sequences of analyses must continue uninterrupted in order to generate valid results. In shift work, several analysts may need to interact with the ongoing sequence of analyses to add samples to a sequence in progress.

These requirements prohibit the use of the standard NT session security and mean that the analytical application must have its own built-in security functions.

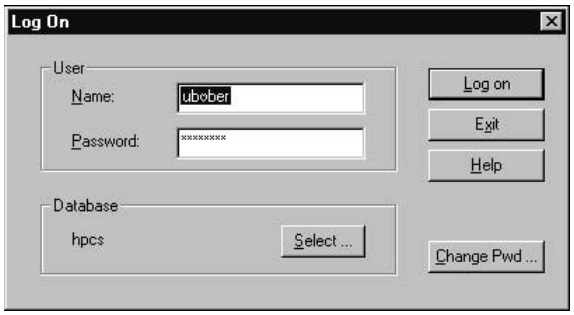


Figure 1
Agilent ChemStation has a password-protected logon for security according to 21 CFR Part 11

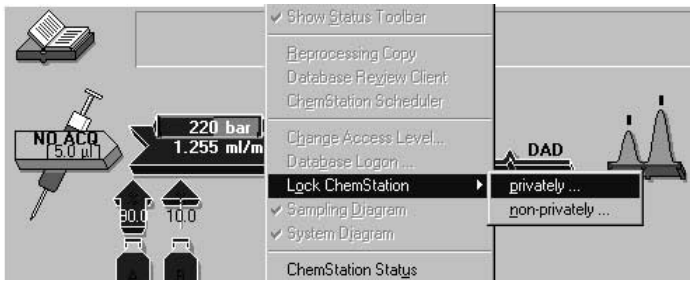


Figure 2
The security pack add-on module for Agilent ChemStation includes a Windows-like private lock that allows only administrators and the person who locked the application to release it, as well as a non-private lock that allows each authorized user to release the application. The non-private lock is highly important in production areas where instruments work on a 24-hours-a-day, 7-days-a-week basis, allowing the authorized change of users between shifts without interrupting the analytical instruments.

An additional reason why security must be managed within the analytical application is the need to control the level of access, in other words, to restrict the functionality accessible to specific users. 21 CFR Part 11 requires users to be properly trained in the use of the system capabilities that they are allowed to access. For example, a user might be capable of retrieving and printing results, but not trained to modify the query used to retrieve the data or the template used to generate the report.

The security mechanism must have access to control of specific functionality within the application to be able to control system access with sufficient precision.

The Agilent ChemStation Plus family of networked data systems includes the required security such as session-by-session timeouts and locks, see figure 2.

ChemStation Plus can be configured as a standalone system on a single PC or as a client-server system in which an Oracle® database is used to authenticate users. Both solutions provide sufficient security under Windows NT and Windows 2000 with appropriate operating system security settings.

For up to eight instruments and five users, multiple standalone systems are recommended, whereby the final choice between multiple standalone systems or a single client-server system should be based on the number of instruments and users as well as on the level of administration required to manage the system.

Choosing multiple standalone systems increases the amount of system administration because the user configuration must be repeated on each ChemStation individually. New employees must be given new accounts and when they leave their account must be disabled. If they receive additional training and are prepared for a more responsible level of system access, their account must be modified. This user setup and modification must be done on each ChemStation individually.

This also increases the likelihood that account configuration errors may occur, either granting access or permissions which the user should not have on certain ChemStations, or blocking access to important functions until an administrator can correct the account profile.

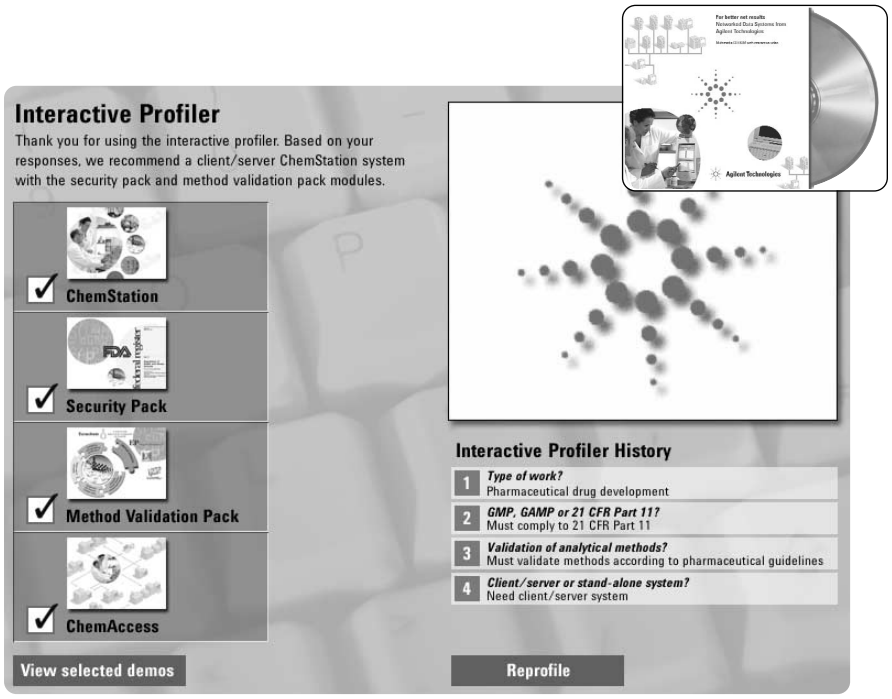


Figure 3
A new multimedia CD-ROM from Agilent includes an interactive profiler. Based on the answers to five simple questions, the profiler helps users to select appropriate data system components and decide whether a standalone or client/server configuration is best suited. (publication number 5988-4666EN)

However, lower purchase costs and novice-level database expertice – for example, Microsoft Access – may sway the decision in favor of standalone systems, if the number of instruments and users falls within the cited limits.

With a client-server system there is a single point of user management, and a user's profile can be updated in a single operation. There is a single system logbook that can generate a history of account management activities, allowing straightforward audits of user account management. The system logbook can also be used to easily relate training records to changes in users' system access permissions.

User configuration and user management is one example of the impact that the configuration of the data system has on the daily working practices of the laboratory. The choice between standalone and client-server-based security is not only a matter of meeting 21 CFR Part 11 requirements, nor simply the costs or effort to manage an Oracle server, but also impacts the quality, resources and costs associated with security management and record keeping.

Agilent provides an interactive tool to help laboratories decide whether a standalone or client-server system fits best to their needs, see figure 3.

Data Traceability

Traceability requires a clear audit trail of both application-level activities such as individual data processing steps, and system-level activities such as logins, failed logins or changes in user permissions. When tracing the integrity of specific results or of the system itself it quickly becomes clear that the two audit trails are not independent of each other.

Each instrument must be qualified and its performance verified. This means that the instrument identity (see figure 4), records all of the automated operations it performs. The status reported for each operation are also important for assuring data integrity. The Application Note "Levels of Instrument Control with Chromatography Data Systems and the Implications for Compliance", Agilent publication number 5988-2907EN, discusses this aspect in detail.

For each material or lot analyzed there will be many analytical results that together build the entire picture of the quality testing performed and the integrity of the analytical data. These results are often generated on more than one instrument, by more than one user, for more than one lot of a material, over a longer period of time.

A final element in the complexity surrounding integrity and traceability is what is known in the database world as *referential integrity*.

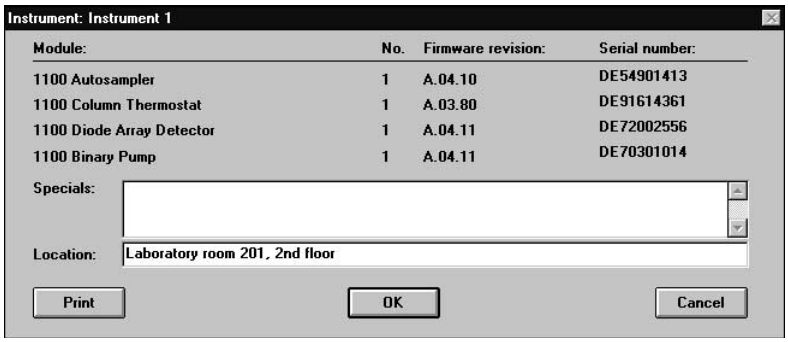


Figure 4
Agilent offers level-4 instrument control—the highest level available today—for its latest generation of instruments such as the 1100 Series LC or the 6890 Series GC. Key features are bidirectional communication and handshake, sophisticated error reporting and handling and advanced control and maintenance features, such as EMF, diagnostics and other service functions. A further, even more important aspect of instrument control at this level is the tracking of instrument or module serial numbers and firmware revisions. This information is not only important for inventory tracking but also fulfills some of the device check requirements specified by 21 CFR Part 11.

Relational databases store different elements of a data record in different areas and structures within the database. Referential integrity provides the assurance that each element within a complete record is maintained and that the proper relationships between the different elements are accurate.

Data Integrity

Based on the SOP different sets of standards and controls are required for a *good quality result* according to GLP/GMP and 21 CFR Part 11. As an example for the requirements of SOPs, most governmental regulations require particular calibration procedures. Modern calibration is dynamic and automated, especially with the now common use of bracketing standards.

Data systems for compliant laboratories must reflect these requirements and all related modifications as part of their default functionality. Other requirements may include system suitability standards, calibration standards, control samples and check standards.

Each result may have more than one revision due to modifications in processing. For each revision there are parameters and actions the operator used to create the particular result. All intermediate results such as integration values and the related meta data (the applied integration parameters) must be maintained and stored for each result revision. To ensure data integrity raw data, meta data and results need to be linked unbreakably.

To determine the integrity of a given result, the revisions that preceded the result must be known as well as whether newer revisions have superceded the result. All of the dependent elements for each revision must be carefully maintained. This maintenance requires the revisioning to be an integral part of the application software. Complete revisioning with all dependent elements is not possible using an external application-independent result management software for the following reasons.

- External applications are not able to poll fast enough to capture changes to files that may occur seconds apart, particularly during automated reprocessing of analytical sequences.

- In contrast to external applications, analytical applications may use intermediate results (such as every individual integration iteration) that were previously considered transient and are currently captured only by specific compliance tools.

- Predicate rules may require comments to be captured and electronic signatures applied to specific actions taken within the instrument's application environment.

This limitations make it necessary to have instrument and application-specific tools for data integrity and traceability.

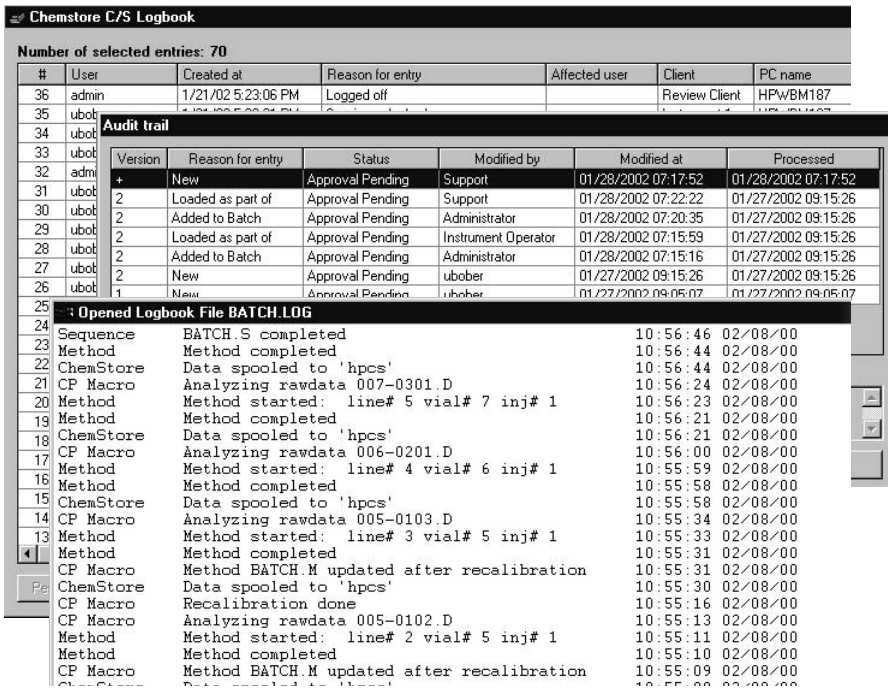


Figure 5
The security pack add-on module for Agilent ChemStation uses three different levels of audit trails to ensure data traceability. The first level refers to all acquisition-related data in the run and method logbook. The second level relates to the data analysis changes and result versions, which are stored with the run in a separate audit-trail. And the third level relates to all application-specific events. These events include logon failures, data archiving and user administration. All changes are stored in the database logbook. Audit trails are system-generated without any user interaction and are stamped with local time and user ID, or with the command stamp that performed the change, together with the actual change setting.

Data Integrity and Tracebility: Standalone or Client-Server?

The difference between standalone and client-server configurations is that the server provides a single data repository from which any result, revision history or audit trail can be retrieved by submitting a query from any client in the laboratory network. Standalone configurations mean that the information is distributed on several PCs and the amount of data that can be maintained on-line on each PC is small.

With standalone configurations, building the history for a material may require retrieving and printing or exporting information from several individual PCs. It may also require retrieving more than one archived database file for each PC, if the information needed dates back more than a month or two. The material retrieved would then have to be assembled manually into a complete history.

The central data repository in a ChemStation Plus client-server configuration is an Oracle relational database. By design its configuration is flexible enough to address the online storage needs for each installation. A typical configuration allows capacity for about six months of data – and can be expanded easily according to individual needs – with all the associated raw data, results, meta data, audit trails and logbooks, see figure 5.

This allows to audit easily the system or the results and facilitates investigations into out-of-specification results. For longer term result comparison, data can be restored easily from their archive media and added to an existing data set, thereby maintaining full traceability and audit-trail documentation.

Integrity: An Additional Benefit of Integrated Data

Intimacy between the system application and the analytical instrumentation also allows automated documentation of elements of information that were previously unavailable. An example of broadening the scope of *all-in-one* data management is automated column documentation, see figure 6.

As a user sets up an instrument to perform a different analysis than was performed the previous

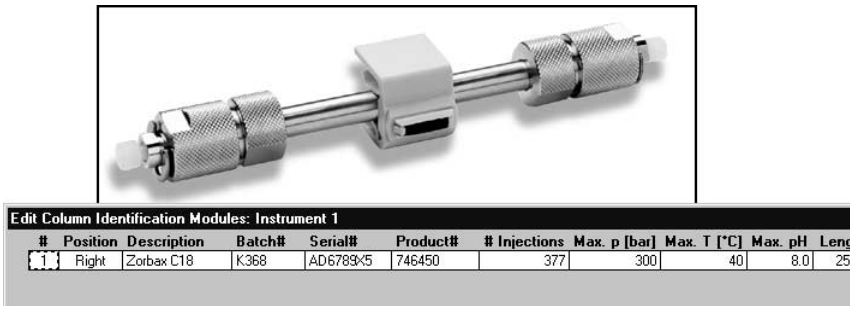


Figure 6
The Agilent column-identification module allows for maximum data traceability. A microchip contains the column's personal electronic ID number and all important column parameters, including a counter for the number of injections performed using the column. These parameters can be accessed through the ChemStation.

day, they might forget to change the column prior to starting the new analysis. If this happens, a result review would probably reveal strange results. Without the capability of the system to sense and record the column identity, the user might suspect that the wrong column was used but if the system has already been reconfigured at the time of the data review, there would be no way to confirm this suspicion.

If the data system automatically tracks all column information used for this analysis, just one mouse-click is needed to confirm the suspicion and save some hours of error investigation.

Similarly, as the column degrades during 24-hours-a-day, 7-days-a-week usage, automated tracking of the number of injections on a specific column helps to understand and predict performance decreases due to column ageing.

By bringing together in one application, all analytical results and their revisions, related meta data, result audit trails, system audit trails and instrument identity and performance records, it is easy to retrieve all the information related to a set of results and the system activities in that time frame. In most cases, direct evidence will be available of any system setup errors or instrument performance issues related to out-of-specification results. By scanning back it is possible to determine definitively the starting point and scope of a problem and clearly understand which data is questionable.

Long-term Data Management: General Considerations

Considerable attention has been focused on the issue of retrieving and archiving data files from the many PCs used in today's compliant laboratories. It is critical that result data be properly and verifiably archived, both from an operational and regulatory point of view.

Larger operations have had to build expensive and sophisticated network infrastructures. Many have begun to layer these systems with configurable data archive applications that sweep files into central archives and provide audit trails for the archiving process.

Unfortunately, less attention has been paid to the issue of referential integrity. The previous section showed that not only must each element within a complete record be maintained, but also that all existing elements are understood clearly and that the relationships between these elements are verifiable.

External software applications may be configurable to capture an appropriate method file with the raw data file but will have limitations in their ability to know the exact relationship between a specific moment in the life of a method and an exact result.

Benefits of Integrated File Management Utilities

With the security pack add-on module for Agilent ChemStation the complete web of elements maintaining referential integrity is placed in a database.

Because the elements are created by the ChemStation and transferred to the database automatically, it is assured that the necessary elements are included and all of the relationships are maintained.

The startling number of network applications and databases that could be put in place to handle different sources of data for 21 CFR Part 11 compliance has lead many vendors to try to use standalone solutions at the PC level and tie them together with network archiving applications. This approach is more expensive and resource consuming than allowing one network and database application for the chromatographic data management and another network application that centrally archives all data from various sources throughout the enterprise.

File Management Requirements for Standalone Systems

With the standalone version of the ChemStation Plus database, the web of data and results resides in a single Microsoft Access .MDB file for each PC. There is a 1-gigabyte limit for these files and headroom of about 20% must be left for new revisions and audit entries if the data is reprocessed in the future. Depending on the type and number of raw data files acquired on the system, a new database file can be needed for each PC at any interval from a couple of weeks to a couple of months.

Backup and even contents indexing of these .MDB files can be accomplished with archiving utilities. But the standalone configuration was designed to satisfy the needs of small-scale implementations that create only small amounts of analytical data. It does not require any advanced database expertise but requires a series of manual maintenance procedures that can become a burden when done on a larger scale.

For example, each new database file must be created individually. They can be based on a previous database file to include the same users and assigned privileges as well as report templates. However, studies are not carried over to the new database and thus need to be recreated. This means that some significant reconfiguration may be needed on each PC when the database file is changed. Changes in the user administration—such as adding new users or password changes according to the implemented password policies—must be performed on each PC individually.

Perhaps more critical is the embedding of security settings in the database file. If the database file is archived without changes in the security settings and user profiles, the settings at the time of the archive are maintained.

This means that to access the database when it is retrieved from the archive, the users who were able to access the system must be known and the users must know their passwords. This can be problematic because personnel changes occur and users are required to change passwords regularly.

Data Management with Client-Server Systems

Oracle databases can handle data integrity and archiving much more smoothly. Standard Oracle utilities can be used for cold or hot backups of the central database. A set of archive and maintenance utilities to manage the data in the central Oracle database is provided with the client-server configuration of ChemStation Plus.

These utilities allow archiving of runs to a separate file on a durable storage medium. These archive units are based on queries and include all result versions. To create space in the database, runs can be deleted from the online dataset in the actual database instance after archiving is complete and verified. This allows data to be removed not only based on age but also, for example, based on a selected project or a specific user. The system maintains audit trails for these archives as well as electronic signatures on archive activities.

The archive units can be interrogated for their content without restoring the archive and specific runs or sets of runs to the online database. Throughout these archive processes the system maintains referential integrity for the archive process, making sure each element within the record for a specific result is kept as part of an integral unit.

Conclusions for Long-term Data Storage

Substantial effort is involved in properly maintaining multiple standalone systems. The likelihood that some errors will occur during ongoing maintenance procedures increases with the number of systems and the total effort required for maintenance. The critical nature of any potential for data loss, makes the client-server configuration a better choice for data management when more than five ChemStations need to be managed in a 21 CFR Part 11 compliant facility.

Analytical Results Management

The pressure to meet regulatory requirements often creates the assumption that the laboratory must simply accept the burden imposed by the regulations. There is a sense that regulations always make life more difficult.

Agilent has turned a careful eye on the impact of 21 CFR Part 11 compliance on laboratory operations. It is definitely true that regulation creates overhead, but the same database used to maintain referential integrity for data and results can also be used to improve the efficiency and quality of information processing.

With the results in a modern relational database, the capability to query for results and then to filter and organize them in sophisticated ways, offers real improvements in results management.

In labs that have not completed their 21 CFR Part 11 compliance implementation, assembly of the final reports is still a relatively manual process frequently using spreadsheets to accumulate and calculate summary statistics or generate plots. This approach now breaks the chain of electronic record keeping and introduces the probability of transcription errors.

ChemStation Plus provides tools for creating sophisticated final reports. In cooperation with users, Agilent has developed complete sequence reporting templates that allow to create a report with system suitability, control and sample sections in a single operation. This includes statistical summaries and options such as individual values and means for replicate injections.

When working with results for a material, it is often necessary to combine results from more than one test method or more than one analytical technique. ChemStation Plus supports a wide range of analytical techniques and combines results in a single database. With carefully selected information fields, it is straightforward to query for the appropriate combination of results and combine them in reports.

Dissolution and stability studies generate large sets of results. For stability testing results are summarized across months or even years. Users may need to create charts of selected results such as the concentration of the com-

pound as it degrades. The ability to query for any data to assemble a report, eliminates the boundaries between data from different analysts or different instruments, or runs from different days or even different labs.

Before the use of databases in networked chromatographic data systems this gap was often addressed by a LIMS or other information management system. However, it was difficult to get results to flow smoothly and automatically into these systems. More importantly, LIMS were designed as repositories for final processed results. Although sophisticated fixed reports could be generated, LIMS were generally not designed with interactive tools for results investigation. If there were questions about the original results, such as integration results, data had to be reprocessed at the source and re-entered into the LIMS.

Conclusions

21 CFR Part 11 has forced vendors to expand the capabilities of their analytical software applications by using databases for result management, adding advanced result review and approval functionality as well as integrated tools for long-term data management and archiving.

Agilent ChemStation Plus now offers all the tools for 21 CFR Part 11 compliance, including result review and management capabilities that were formerly only available in LIMS applications. ChemStation Plus is thus a comprehensive solution for all Agilent instruments such as GC, LC, CE and MS. In laboratories with instruments from multiple vendors, ChemStation Plus can be easily interfaced to additional external LIMS applications for enterprise-wide result management.

With ChemStation Plus, data flows back smoothly to the ChemStation base software for reprocessing and the database is automatically updated with the new result revisions.

If the results are kept on each individual standalone PC and perhaps contained in archived .MDB database files, where the runs were deleted from the database, manual assembly of results is still required. When LIMS or other information management systems are in place, results assembly may take place within these applications but investigation and reprocessing at the raw data level is still be a largely manual process.

Using an Oracle database with a client-server ChemStation Plus system minimizes the manual assembly and interactions in result review and archiving. ChemStation Plus includes a single archiving utility for data from all clients along with a central and online repository for all user management and security actions where each change is easily traceable over the full database lifetime. For laboratories with the need for an integrated archiving that want to minimize manual user tracking and manual security actions on analytical data ChemStation Plus is the right tool to answer these needs.

Summary

This Application Note described the general requirements of 21 CFR Part 11 for data security, integrity and traceability, and long-term data and result management. The evaluation of ChemStation Plus has outlined the major differences and limitations of standalone systems compared to client-server systems. Agilent offers tools that help users decide whether a standalone or client-server configurations is the right choice. The major limitation of all standalone solutions — regardless of whether they use an external file management application or a LIMS solution as final data containers — lies in the area of cross-instrument result review, especially in conjuction with archived data.

Working with a centralized, often termed *in-process*, database for analytical results offers real improvements in efficiency and accuracy for analytical data processing. Once the results are finalized they can flow easily into enterprise-level information systems.

www.agilent.com/chem/nds

Printed in Germany, November 1, 2002
Publication Number 5988-4743EN

Agilent Technologies