

Building a Hosted Statistical Computing Environment: Is it Possible?

John Leveille, d-Wise Technologies, Raleigh, NC, USA

ABSTRACT

Small pharmaceutical companies face a large challenge of implementing a compliant and validated environment with limited resources, no QA, and an IT staff who thinks SAS[®] is a European airline. They go searching for new fangled technology to solve their problems, only to find high priced commercial solutions that are more of a sledgehammer on a pin. So small companies are stuck: they can't afford commercial solutions and don't always have the knowledge to implement their own. This paper will present a case study of how a small pharmaceutical company implemented a hosted computing environment including development and production workflow, version control, syncing of clinical data and a set of robust processes. And the best part is that they did it at a price tag that was right for their budget allowing them to grow in the future.

INTRODUCTION

Despite their size, small pharmaceutical companies must adhere to the same regulations that govern large ones. Production statistical computing environments must be validated and compliant; no exceptions. With limited resources, minimal QA, and an IT staff having little to no understanding of SAS[®], it is difficult to know how to proceed. Small pharmaceutical companies can become paralyzed by the risk of various failure scenarios such as:

- creating a system that they can't afford to validate
- creating a system that doesn't grow as the company grows
- creating a system that is too costly to support.

They fear the big bad audit from the FDA; however, they don't have the time to implement technology and processes to alleviate that fear. Some companies buy a costly package labeled 'statistical computing environment' only to find that they cannot envision doing enough trials to recoup the capital expenditure. Buyer's remorse really sets in when they realize that they still have to spend time and money to validate the system and define work processes.

So what is a small band of statisticians and programmers holding the next blockbuster compound to do? Build it. But before setting out to build it, the team needs a plan, a template, if you will, to keep the project scope focused and ensure that the project achieves its end goal: a validated computing environment with a controlled, documented work process.

The case study presented here describes such a project. Technical consultants from d-Wise teamed with a cutting edge biotechnology company to develop a project approach and system design. We compressed our initial costs by keeping the design simple and using free, open source technologies, where possible, for several system components. We employed virtualization technology to enable a return on investment and minimize validation overhead..

ASSESSMENT

The first step in the project was to sit down and evaluate the company's current environment and how the users work on a day to day basis. We conducted a series of team meetings to map the current process and identify critical elements as well as gaps or deficiencies. Key personnel, departments, roles and responsibilities were identified within the process mapping. The following high level requirements were defined for the new environment:

- A hosted server easily accessible by users with performance similar to the existing local environment
- Ability to develop code locally and promote to a production environment
- Version control for clinical programs
- Automatic synchronization between data management systems and the production environment
- Secure, encrypted communication channels between the user's local machine and the hosted server
- Ongoing hardware, software, and process support within the hosted environment
- Standard operating procedures, work process instructions and training for the new environment
- System validation

The project team was formed from a combination of d-Wise and customer resources. d-Wise provided the roles of system architect, system administrator, and project manager. The customer provided various roles to represent the users of the system. A liaison from the customer IT department was assigned to the project to confirm that the

design was in compliance with IT policies and security. A resource from the customer QA department was assigned to the project to oversee validation of the system and vendor/facility audit. Together, the team assessed the scope of the project and formulated four project goals:

Project Goal 1: Keep it simple, but don't compromise any requirement.

Project Goal 2: Shorten the validation effort.

Project Goal 3: Allow the system to scale up without validation overhead.

Project Goal 4: Minimize cost and impact to current working processes.

SYSTEM DESIGN

Based on the requirements, an infrastructure was designed with the goal of minimal impact to the user community. The design focused on the workflow required by users as well as the software and hardware components needed to support the workflow.

Appropriate server class hardware was selected along with Microsoft remote desktop services to satisfy the hosting and performance requirement. A free, open source version control software package named Subversion satisfied the requirements for version control of clinical programs and code promotion. Since Subversion software has a client-server architecture, it is accessible from a local machine development environment and from the hosted production environment.

A custom data synchronization mechanism was designed to model the customer workflow. Free, open source software was utilized to perform the data synchronization and secure the channels for transmission of data across the Internet. Figure 1 provides the high level design of the hosted environment.

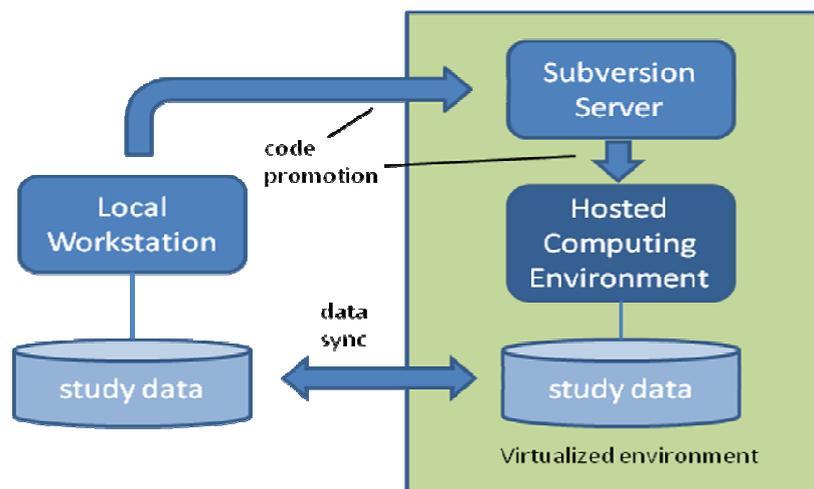


Figure 1. Hosted Platform Design

IMPLEMENTATION

Once the assessment and design were complete, the team began the implementation of the platform. The design diagram presented above shows a single production environment, but the implementation actually involved three separate, identical environments known as DEV, VAL, and PROD. These three environments are necessary for compliance with 21 CFR Part 11. If you have only one environment, "production", then you have no ability to stage and test changes prior to implementation in PROD. The three independent environments are essential for system validation and change control processes. This was the point in the project where virtualization technology began to provide some return on investment. Virtualization technology allows you to create something called a virtual machine (VM). A VM is essentially a file residing on your computer hard drive that includes the operating system, software applications, and data storage for a computer. This means that a single hardware device -- a single server in the data center -- can host multiple VMs and provide what appears to the user as multiple remote computer systems. This technology is sweeping the computing world because of obvious savings on physical space and data center power consumption.

Now back to how virtualization provided a return on investment within the project. In a traditional validated system build, the team must purchase three separate, identical sets of hardware and perform three separate install and configuration exercises. Figure 2 shows the very separate traditional environments.

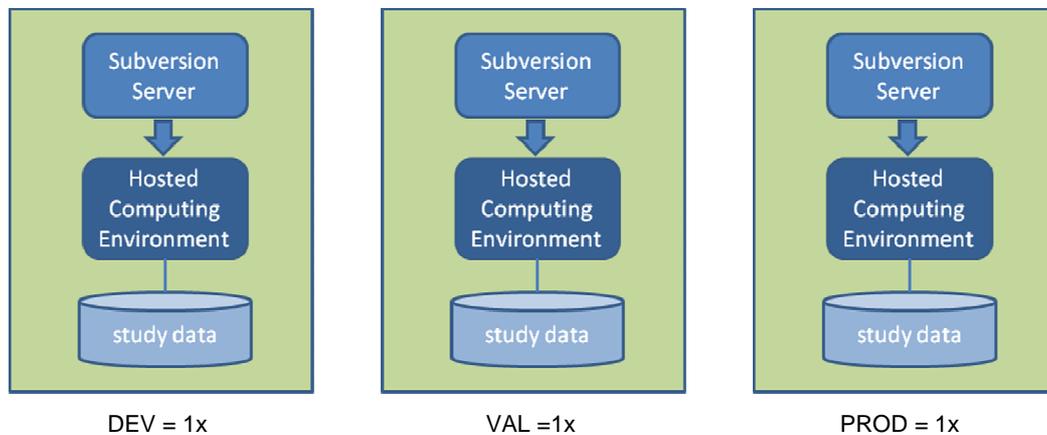


Figure 2. Traditional System Implementation Model

This approach yields a 3x hardware cost. Using virtualization technology, we were able to purchase a single set of slightly larger hardware capable of hosting the three environments. Since hardware vendors often give you more for your money we were able to procure an appropriate machine in the range of 1.5x to 2x. In addition, this single machine takes up approximately 1/3 of the space and power of the traditional setup, saving on hosting costs, shipping, insurance, physical installation time, etc. Figure 3 shows the virtualization model deployment.

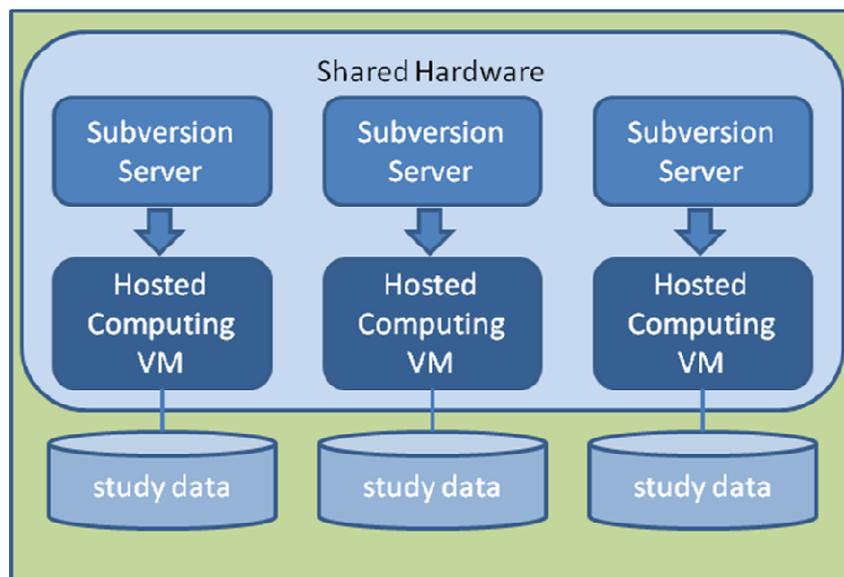


Figure 3 Virtualized System Implementation Model

The next cost saving came from reducing the validation effort by cloning VM images. Again, the traditional approach is to perform separate installation and configuration of the environments by laborious, repeated execution of an Installation Qualification protocol. Typically, this is done for all three environments, but only documented for VAL and PROD. The traditional validation process is shown in Figure 4.

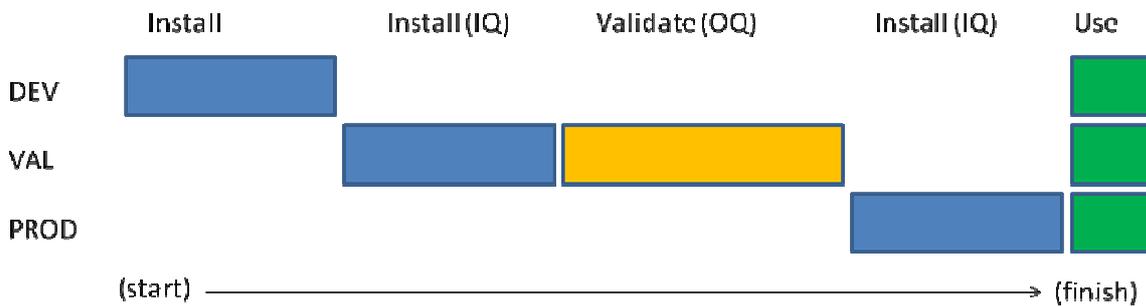


Figure 4. Traditional Validation Process

With virtualization, we were able to adopt a different approach to the deployment and validation of the environment. This new approach is shown in Figure 5.

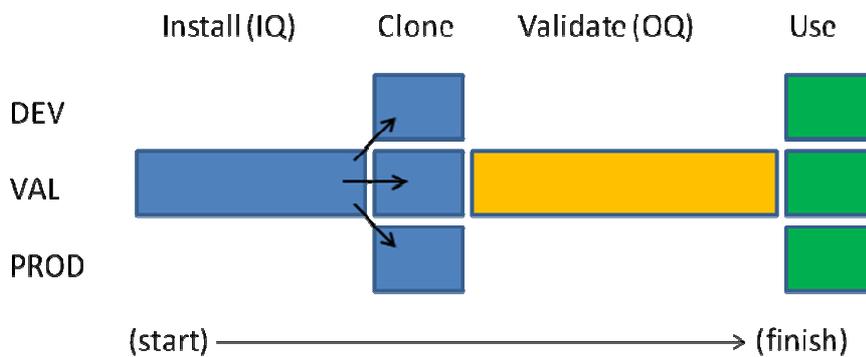


Figure 5 Validation Process within a Virtualized Deployment

Not only did this approach save time because most of the installation and configuration was done once and then copied, it actually results in a better equivalence between VAL and PROD. The principle of equivalence between VAL and PROD environments is essential for effective validation. In the traditional project, this equivalence is maintained through reliance on a human being to perform steps exactly as documented and to accurately record every step performed. Human error is a very real and significant risk factor and is typically combated by adding in more documentation and control procedures, heaping on the cost in both time and money. In contrast, the use of virtualization to literally copy the baseline VM into a VAL and a PROD environment removes nearly all of the human error by allowing the computer to do what it does best -- make a digital copy. As if this savings and risk reduction was not enough, the baseline VM image is archived and can be used to stamp out additional PROD VMs, allowing the customer to easily scale up their computing resources as their needs grow

In addition to having more bang for the buck in our hardware purchase, virtualization allows us to turn off the DEV and VAL environments and reclaim those computing resources for other purposes. For example, suppose that the system validation cycle is complete and the PROD environment is in active use. The customer has a drug approaching submission and wants to add 4 subcontractors to handle the additional workload. It is a simple matter to shut down the DEV and VAL virtual machines and enable another virtual machine within PROD. Since PROD VM2 is a clone of PROD VM1, it is validated and ready for use within a matter of hours! This level of agility is rarely possible in a validated system. Figure 6 shows the ability to define one virtual machine image and deploy it multiple times. The figure depicts the VAL environment being shut down making hardware resources available to the additional PROD VM.

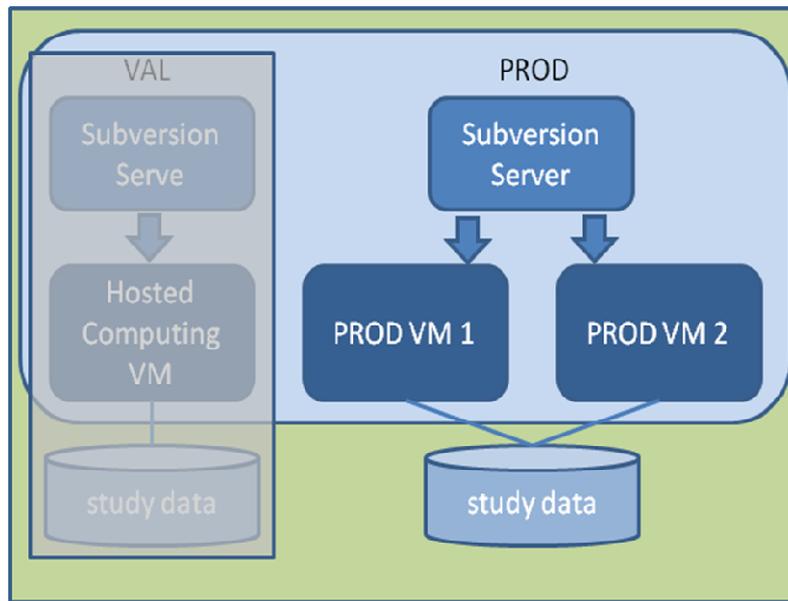


Figure 6. Scaling up a Virtualized System

PROCESS DEFINITION

This new system resulted in some necessary changes to the clinical programming Standard Operating Procedures and work processes. Every effort was made to minimize the extent of the changes for the users and to make the adoption of the new system as seamless as possible. SOP changes were limited to just a couple minor changes and we avoided writing any new SOPs. The emphasis was placed on the working process documents and the steps required for data management and clinical programming activities within the new validated, compliant environment.

The team kept the timeline tight by working through process definition in parallel with the system design and build. During the implementation phase, we prototyped each of these processes:

- Raw data synchronization from data management into the hosted environment
- Promotion of clinical programs from development to production
- QC of clinical programs
- Report and output data synchronization from the hosted environment back to the corporate network

Prototyping and analyzing the work flow of these processes allowed us to make sure the system would meet the needs of the users. The team developed a roll out plan, including user training and work process documentation to ensure users were ready to adopt the new platform and processes once validation was complete.

CONCLUSION

It is true that many IT departments, even at small companies, are using virtualization technology. You need only watch TV to see that “The Cloud” is everywhere. A small pharmaceutical company can easily find a hosting provider that will provide a virtualized server, and, in some cases, their own corporate IT department may be able to offer this service. The problem with these options is that the virtualized server is just a stand-in for a typical operating system installation. They have missed the advantage that comes out of validating the virtual machine as a complete package:

Validated Virtual Machine = [operating system + applications + IQ/OQ/PQ validation documentation]

The result of our project was a validated, hosted statistical computing environment. It seems, from the outside, like many other such systems, but at its heart is something quite unique: A virtualized configuration where each VM and each VM clone is validated and ready to go with the flip of a switch.

This paper has presented a case study of how an agile pharmaceutical company collaborated with a technology consulting company to build a full featured, validated, hosted computing system. The solution was implemented along with a set of robust processes and minimal impact to the users, positioning the company to work effectively

today and expand in the future. The best part is that they did at a price tag that was right for their budget. So the answer to the question “Is it Possible?” is a resounding yes!

ACKNOWLEDGMENTS

Special thanks to Chris Decker, d-Wise Technologies, for his guidance and editing.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: John Leveille
Enterprise: d-Wise Technologies
Address: 4020 Westchase Blvd Suite 527
City, State ZIP: Raleigh, NC 27607
Work Phone: 919-600-6236
E-mail: jleveille@d-wise.com
Web: www.d-wise.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.