Cloud Computing with SAS

Sy Truong, Meta-Xceed, Inc., San Jose, CA

ABSTRACT

Cloud Computing is not a new computer architecture construct but there are several forces that are going to make it an alternative, if not the main way SAS software will be delivered in the future. This paper will detail some of the forces that are pushing many software solutions into Cloud Computing while describing how SAS is a microcosm of this transformation. Some of the topics being presented in this paper include:

- Cloud Computing Introduction

- Driving Forces for Change

- Software Components of Cloud Computing

- Requirements for Success

The development of Cloud Computing has been in the making for some time. There are still some challenges and SAS needs to prepare and make a shift to this computing environment if it is to remain a dominant force in business intelligence software.

CLOUD COMPUTING INTRODUCTION

As software becomes more sophisticated to handle larger and more complex sets of data, the computer systems that run these applications also become more challenging to manage. New system architecture such as Grid computing for parallel processing and multi-tiered computer architecture are designed for scalability to optimize performance. This environment enables hardware configurations to deliver and match user demand for data mining and powerful analytics. IT managers in organizations that are responsible for the installation and management of these systems are faced with daunting challenges that go beyond their resources. This is because each vendor and software system has its own unique computing configuration requirements and thus demands a full time member to validate each installation and perform the administration and maintenance of that system. This poses challenges but also opportunities which give rise to Cloud Computing. Cloud Computing embraces the complexities and unique requirements of complex software solutions, yet delivers them to users in much less time and startup cost. Cloud Computing is therefore going to alter the way computer systems will be implemented by large organizations and small organizations alike.

Software as we know it is going though a fundamental change in the way it is delivered to users. Historically, it has gone through several stages in its evolution and it is now going to make another transformative leap as it adapts to Cloud Computing. Early in software development, the main processing of software resided on a mainframe computer. Users connected to it through a dumb terminal which only displayed text on monochrome monitors. All the computing processes and related files were centralized on the mainframe. With the advent of the personal computer, the processing power on the user's desktop became more powerful so software shifted and was installed on desktop computers. Many software packages were delivered in a box, shrink wrapped and stored on CDs. The SAS system reflects this evolution with earlier versions of SAS on mainframes and version 8/9 migrating to desktops. The complexities from multiple offerings from SAS pushed the boundaries of software in the box model and is transitioning it into a new environment of Cloud Computing. Elements of this are beginning to happen as the SAS software is now delivered as a download.

The mass market acceptance of services such as Google and Amazon has helped push the evolution of Cloud Computing. The complexity of the software configuration is moving to a separate set of servers managed externally to the internal IT resources or individual software user. The hardware and software for these new systems are placed behind a "cloud" on data centers which are managed separately and accessed through the internet. It is then delivered to users through a web browser so there is no longer the need for lengthy installation, validation, version upgrades and other related software maintenance. The portability of being able to access the software and related data combined with the ability to outsource the management of sophisticated software systems is proving to be more efficient and necessary for software such as SAS.

DRIVING FORCES FOR CHANGE

There are many forces at play that are creating a ground swell for the change that is enabling the growth of Cloud Computing. Some forces are technology advancements which have been evolving over the past decade, while others include market and economic forces that are more recent and can be more potent in specific vertical markets. The technology developments by themselves are not significant but the culmination of all them has profound effects across industries. The following list describes some of the reasons for the change and also provides insights on how you can adjust to and thrive in this dynamic environment.

1. **Telecommunications –** With the dot com boom of the late 90s, there were fiber optics and high bandwidth data infrastructure put in place for global connections. This allowed access to fast speed voice and data connectivity. The dot com bust during the start of the millennium affected many of the companies that established this infrastructure by forcing them out of business. This offered the fiber optics and equipment for whole sale or at a market discount. This created cost effective opportunities for remote computing in many environments and in particular for the development of Cloud Computing.

2. **Commodity Hardware –** Computing has evolved from expensive centralized mainframes to personal desktop computers (PCs) that have become ubiquitous. The mass production and personal consumption of PCs has made it a mainstream appliance which has driven down their prices. Although these PCs were originally designed for individual use, they can also be networked and clustered together into grid computing to form powerful super computers which provide the back bone of Cloud Computing. The power of Cloud Computing is no longer restricted to large institutions with large budgets but can be implemented at a fraction of the budget due to the commoditization of hardware accompanied with new clustering software.

3. **Open Source –** The decentralized approach to software development challenged the institutionalized form of software development. It allowed for individuals to contribute program code updates to form sophisticated operating systems such as Linux that runs on most web servers and many computer electronic devices. Open source software also played a significant role in web server technologies that form the core components of Cloud Computing. In addition to the Linux operating system that was first installed on commodity hardware, apache web servers provided web services. The server side software on these web servers along with middle ware including client side XML based components are all open source forming the foundation for many Cloud Computing applications.

4. **Web 2.0 –** The combination of the hardware and software technologies mentioned above have created a new computing platform that is very distinct and different from its predecessor. This computing platform is referred to as Web 2.0 and drives many ecommerce and social network services. Web 2.0 has increased in adoption because it has made web applications more interactive with features from mouse over popup to drag and drop features that were once only available on desktop applications. The interactivity and sophistication shine in ecommerce since it has made the user experience much more intuitive and user friendly. These are the same forces that are helping to bring Cloud Computing to the mainstream.

5. **Social Network –** What was once viewed as a social experiment that college kids did with Facebook or a group of eccentric academics tinkering with their Wikipedia has matured into a powerful business tool. The same principle that drove the social networks to connect, share and learn has become a powerful tool for collaboration, marketing and other business applications. Organizations develop new features to their products, get ratings, and connect to their customers with greater efficiencies compared to the traditional print and broadcast media. The success of Cloud Computing leverages on the implementation of social networks. Cloud Computing shares some of the same core technologies and therefore works symbiotically with social networking.

6. **Cloud Computing Services –** As the technologies and processes mature and large companies such as Amazon, Google and IBM operate large operations using Cloud Computing, they are starting to put together offerings to external organizations and individuals. These pioneering organizations have worked out many of the challenges of implementing Cloud Computing from the IT hardware infrastructure to the software and development tools needed. More and more services are beginning to be offered to businesses and individuals to implement new Cloud Computing services based on the existing infrastructure so you do not have to re-invent the wheel. This can be seen at all levels of the traditional software market including consumer based individuals extending to large enterprise solutions.

7. **Economic Down Turn –** The economic down turn has put a strain on IT budgets and refrained capital from implementing large systems. Companies still need to deliver services to their customers and IT departments' demand solutions for their customers. However, the credit crunch has halted many projects that require a huge setup cost. Rather than purchasing or committing to large systems, businesses can only spend a small amount on critical tasks to operate their business. This includes many different aspects of the system such as hardware cost, software cost and support. Cloud Computing can modularize specific

business functions and deliver to clients in an a la carte fashion with a complete service including hardware, software and support. The IT budget constraint has made Cloud Computing a viable and cost effective alternative.

8. **Outsourcing IT –** With the economic environment, companies turn to outsourcing as a way of cutting cost to maintain competitiveness. Various different parts of IT are outsourced. Support call centers can provide cost savings when supporting their customers. In order to cut cost within their own organization, the ability to outsource the implementation of software systems is another layer of services which can be outsourced. This frees the resources of an organization to focus on core business. Cloud Computing will be the enabler and the fuel to this type of software as a service (SaaS).

9. **Mobile Computing –** The need for users to access some or all their information to conduct business when they are away from their desk is becoming more of a necessity. The ubiquity of laptop computers is replacing desktops and the proliferation of BlackBerry and iPhones is creating a new platform in computing. The dimension of mobility along with the increasing power of hand held devices provides new opportunities for Cloud Computing. Currently, the web browser is the dominant method for delivering software to clients across the cloud. Mobile computing represents the next stage that will enhance the delivery of applications.

The ecology of software in the Business Intelligence (BI) sector is multi-facetted and complex. There are many components that are required for delivery of large computer systems within businesses large and small. There are many factors that go into changing the way software is delivered. Some of the agents of change are evolving technologies that provide better solutions to existing business requirements. This has been proven within the telecommunication and web technologies. Other factors that contribute to the transition to Cloud Computing have to do with business processes and people, such as the case of access to capital in an economic environment or social networking. When all the factors work in concert, it propels Cloud Computing beyond just an alternative solution to traditional software.

COMPONENTS OF CLOUD COMPUTING

Cloud Computing is a compilation of several components.   This section will elaborate on these components and describe some examples to help you understand the technologies and processes so that you can implement your own Cloud Computing service. Cloud Computing is a larger framework of technologies and processes that incorporate other existing technologies. A couple of key components that are used in Cloud Computing include Web 2.0 and SaaS or Software as a Service.

1. **Web 2.0 –** Web 2.0 is a set of technologies that is used on the World Wide Web to enable more collaborative and interactive applications. The term was originally used by O'Reilly media at the conference held in 2004. The number "2.0" is a borrowed construct from software releases which implies that this is a new version of the web. For example, when SAS comes out with a new version, it is labeled as SAS 9.2 which has newer technologies and functional specifications compared to the previous release of SAS 9.1.3. However, the moniker "2.0" in Web 2.0 refers to the second coming or revival of web technologies after the dot com bust at the beginning of the millennium. It is therefore more similar to a movement comprised of a collection of technologies and tools rather than a specific set of technical upgrades. If you had to quantify some of the technologies involved, it could include tools such as wikis, blogs, asp (application service providers), and the social networking websites.

   Many concepts and accompanying technologies of Web 2.0 are implemented in Cloud Computing. In the early "Web 1.0", the World Wide Web was more like the Wild Wild West in that there were a myriad of sites and technologies with little coordination and cohesion. This may have helped in democratizing some aspects of computing but it was challenging for business applications. The set of sites and tools for Web 2.0 has organized this chaotic environment into a computing platform that is utilized to form Cloud Computing.

   Some examples of well known Web 2.0 applications include eBay, Craigslist and Wikipedia. These applications take the full advantage of the users on the internet to work together and even contribute to the content which makes the platform unique compared to other software environments. Another aspect that is unique to Web 2.0 is its ability to be accessed from anywhere on the internet rather than having to be installed on a particular desktop or internal intranet. The content and interface of Web 2.0 applications are more interactive and rich in media compared to its predecessors. This is analogous to the early days of SAS/AF where the GUI was developed as character text displayed on monochrome text monitors. This evolved into a windows user interface and then into a web interface that is more media rich and interactive. Web 2.0 is implemented using underlying technologies such as Ajax, Flex and Siverlight. These technologies are transparent to the users since they are not aware of or concerned about how it was developed but rather that the result is a more user friendly, dynamic and media rich experience.

2. **SaaS –** It is by coincidence that software as a service is pronounced like SAS. It is not however an analytical BI system but rather a construct for software delivery that is distinct from the traditional software in a box. This means that the software is no longer installed on the user's desktop or on a server local to the user's organization. This is available across the Internet so it alleviates the need to install, validate and support all the components of the software. In the example of SAS, rather than having the SAS executable and core catalog located locally on your SASHELP path, the application service provider (ASP) delivers the same software function across the internet. This is very suitable for sophisticated systems where the computing resources and expertise needed to maintain the system goes beyond the capabilities of the user and its IT support group. Some examples of software developed in this way include CRM, human resources, web analytics, and clinical information management systems.   Some attributes that enable software to be delivered as a service in this manner include the reliance on the network which delivers the key information stored centrally. It is usually installed in a single or few instances with many users accessing it. In the model of SaaS, SAS is not installed on each desktop but rather it is centrally installed on a server and many users can access its core logic and related data across the network. This is different from having a desktop based display manager tool, but instead the user interface is deployed via a web browser or a remote desktop. Cloud Computing fits very well into the model of SaaS since its logic can be centralized and then delivered across the network. This methodology takes advantage of the infrastructure of Cloud Computing.

Cloud Computing may appear at first to be a disorder mash up or patch work of a myriad of open source technologies, but rather it is evolving into a cohesive computing platform. As the technologies mature and all the various underpinning processes function together more seamlessly, Cloud Computing functions to form a new platform. The key to its success is shown more in the next section which demonstrates that it is not in any particular technologies employed that plays a pivotal role, but rather that the user's experience in the end with global access via a thin client browser.

REQUIREMENTS FOR SUCCESS

This section will discuss key requirements for a successful Cloud Computing implementation. Examples include the use of Websas and ClinicalSMS which are applications used to manage clinical safety information in a Cloud Computing environment. The following will list a series of steps that can be taken detailing the requirements needed and provides insights in helping you implement your Cloud Computing implementation successfully.
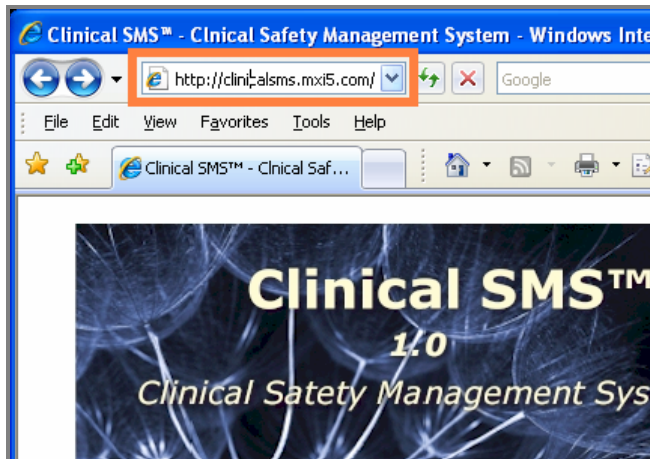
STEP 1 - USER ACCESS

figure 1 – Access via Internet URL

One of the unique aspects of Cloud Computing is that the user can access it from any computer that has access to the internet. The requirements on the user machine is therefore much lower so it is important that the provider of the software put efforts into making that single URL work universally for all users. Some of the configuration considerations include the following:

1. **Browser Compatibility** – The main browsers that are commonly supported include Microsoft Internet Explorer and Mozilla Firefox. Some applications require support for Google Chrome and Safari. HTML sometimes renders differently on different browsers and JavaScript also functions differently or not at all

depending on the browser. You can have your script identify the current browser being used and deliver the suitable HTML and JavaScript that matches the browser that the user is currently using.
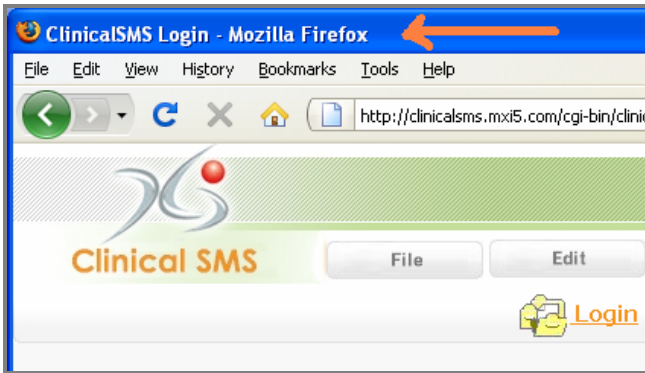

figure 2 – Access can be through Mozilla Firefox or Microsoft IE

2. **Connection Speed –** The user access may vary in speed depending on whether they are connecting from an office with high speed internet access or are traveling with a slow dialup connection. If the application can be optimized to support the slow connections, then it will accommodate all users. You can optimize all images by having optimal JPEG or GIF compression. If there are files that are commonly used throughout the application such as images, XML or script files, then have these files download during the opening splash screen. This will therefore be cached in memory and speed up response during usage.

3. **Secure Socket –** It is important to secure the communication between the server machine and the client browser. There are different protocols that can be used such as SSL but sometimes non-standard ports are used. So rather than the default: http://myserver:80, a non-standard port number other than "80" may be used. For intranet implementations, this works well since you have control over your internal router. However, for an external Cloud Computing system, this requires only standard port numbers. You have to design the system with secured communication all within the servers and limit the number of access from the cloud.

4. **Mobile Devices –** This may be a small part of your user base but it is growing. It is important to clarify with your users if they have a mobile device such as an iPhone or other smart phone that they plan to access the tools with. In this case, adjust the delivery similar to how you would handle a different browser. For true optimized implementation, an application developed for that mobile device can be made available to replace the web browser for specific functions that require the performance and optimal interface.

STEP 2 - SECURITY

The most widely used security on most ecommerce websites is the use of SSL. This secured socket layer encodes the information entered on your browser as it travels from the client to the server machine. This does ensure the protection for certain key data values such as passwords. In addition to this, there are additional measures that you need to implement to optimize the secure implementation of your Cloud Computing service.

1. **Access Control List –** When the user logs onto a machine, it relies on the user account established on the operating system to control its access. In a similar way, you can add a layer of control to the access of your application. This will allow for controls to the specific functions and areas of your application that have greater granularity than the ones provided by the operating system. This can then be applied to functional roles that each user plays when using the system, such as an administrator would perform different tasks than a user. When accessing systems over the cloud, the user is removed from the operating system on the server. They are focused on getting the job done and not concerned with how their files are stored. It is therefore important for you to manage the user access separately from the operating system.

figure 3 – User access is controlled by an account established for the application

2. **User and Data Access –** The access control list allows you to identify and authenticate each user. The roles and user privileges granted to these individual users limit what access they have. This is important to control different functional areas of the system such as the distinction between administrators versus a casual user. Another important distinction is knowing which users will be granted access to what data. The databases that users access usually contain the business rules and consist of the primary intellectual property (IP) of your organization. It is therefore important to add an additional layer controlling users' access to the specific databases of the system.

3. **Password Rules and Expiration –** It is familiar to most users in a secured environment that their password expires over time and that they have to enter a new one periodically to maintain a secured system. The passwords they come up with need to contain special characters and numbers to ensure a level of protection against hackers. These are rules that the system can enforce to help users function in a more secure environment. One aspect of Cloud Computing that is very different from desktop software is the users' session. If a user is logged on and then decides to step away for a time, that session is considered a security threat. If another user is able to access that session and enter, security would be breached. It is therefore important to have the ability to control and expire a session after a specified time such as 10 minutes.   In that case, if the user tries to access after the 10 minute period, they would need to log in again. This measure may appear to be an inconvenience to the user but the little effort goes a long way towards closing security holes to the system.
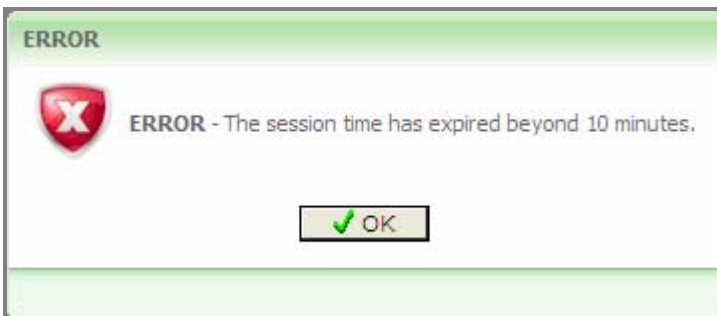
figure 4 – A standard message is presented to users upon expiration forcing a log in.

STEP 3 – INTERACTIVE NAVIGATION

The advent of Google and other developments of the web have raised the level of expectation of users to expecting that they can get to any piece of information they need with only a couple of clicks. This goes beyond the level of expectation of most desktop applications. The goal for a successful Cloud Computing application is to get users quickly to their desired location with as few clicks as possible, while also maintaining a user friendly interface. The use of popup navigational menus that are common on websites can be implemented into applications to enable navigation of an application as friendly and efficient as a good website.
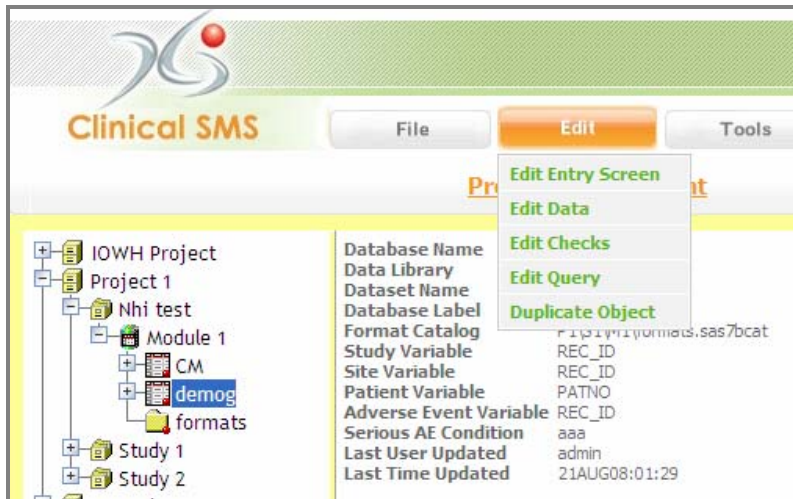
figure 5 – A mouse over popup navigational menu makes navigation more efficient

STEP 4 – DYNAMIC RICH USER INTERFACE

The delivery of context help or tutorial is one example of how Cloud Computing can take advantage of providing media rich content to users in an instructive way. There are many examples including motion video and tutorials. Rather than having the user sit through a lengthy video however, the specific piece of information that the user needs is usually on the screen that they are currently viewing. A mouse over popup bubble illustrates how elegantly context information can be provided on any object on the screen within the context of the users' current screen.
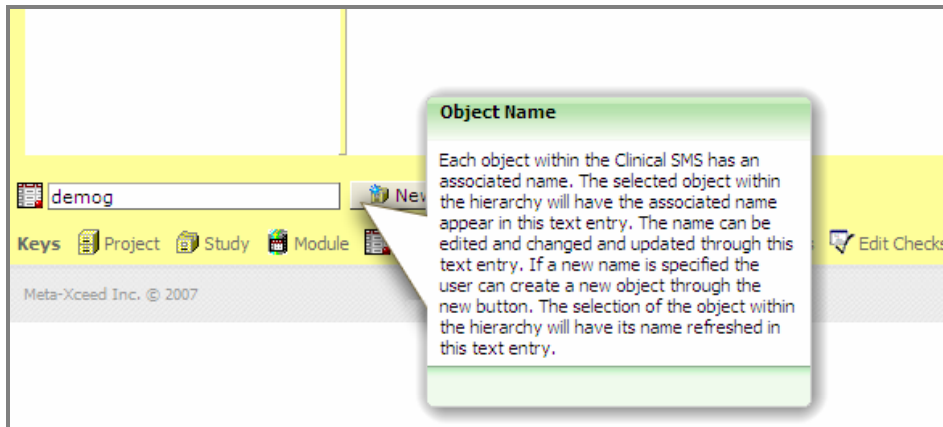
figure 6 – A mouse over popup can display additional information on any object on the screen.

The mouse over accompanied with a popup can replace the need for a tutorial or user manual in some cases and provides a more interactive experience. Another common behavior on desktop applications that is not commonly seen on web applications on the cloud is the ability to drag and drop objects. This is particularly useful for copying files from one location to another.
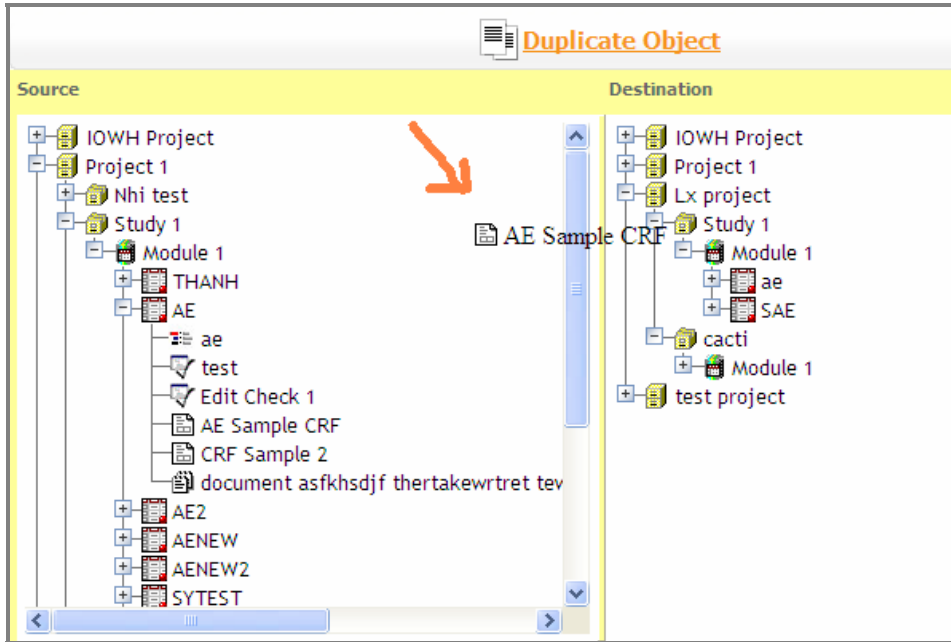
figure 7 – Drag and drop effectively issues a copy command

The ability to graphically organize information such as files into a hierarchy tree view is a crucial function for file management. The use of plus and minus signs in this example interface allows users to go view and navigate through large sets of data files. When this interface is combined with the ability to drag and drop objects within the tree view, it adds an interactive experience that is common on desktop applications but has previously eluded web applications. This is no longer a limitation and can be deployed on Cloud Computing.
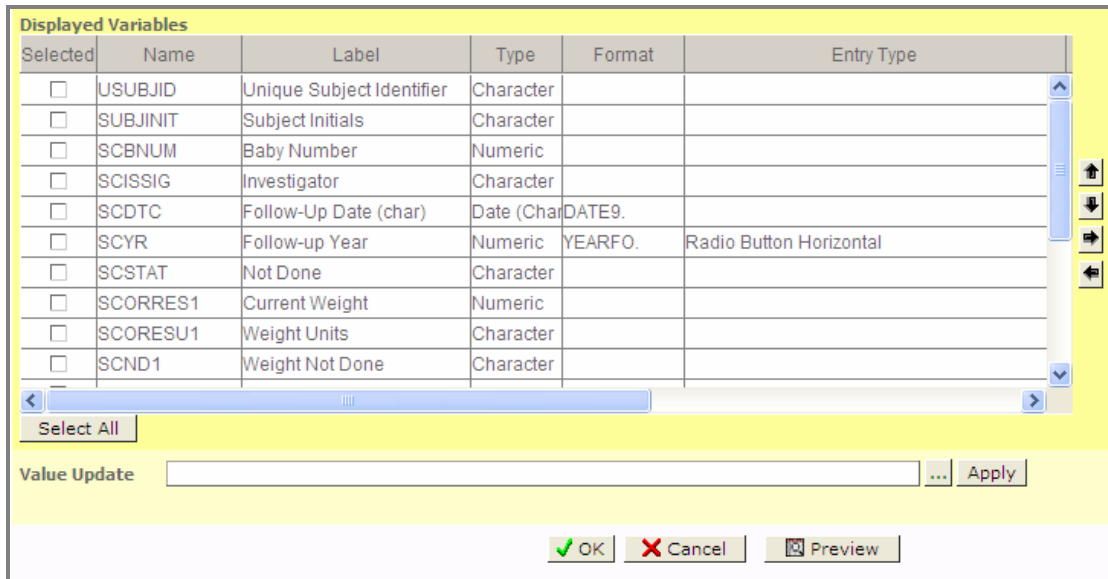
figure 8 – Interactive spreadsheet view of any data view

The spreadsheet is a "killer app" when it was first introduced on desktop graphical user interfaces. It allowed users to visually see and interact with their data while performing calculations among the cells interactively. These capabilities are not limited to desktop GUIs but have extended to Cloud Computing as the tools mature. This example illustrates how variables and attributes are listed in a spreadsheet view allowing users to scroll through, stretch and adjust cells, interactively modifying values and interactively working with individual or groups of cells.

The use of interactive interfaces such as the spreadsheet view, tree view and interactive popup can provide a very rich user experience. This is different than desktop applications however in that there is no additional software

installed on the desktop computer and the data file is still centrally managed on the cloud servers. This example illustrates how Cloud Computing can deliver user friendly enterprise solutions to a global set of users while maintaining central control with optimized servers.

STEP 5 – USER COMMUNICATION

The integration of communication methods with the user such as email plays a significant role in Cloud Computing compared to other platforms. Email is the universal method for messaging when on the internet so using this method to communicate detailed information or sending attachments with additional information is an effective way of communicating to users.

figure 9 – Integrating Email as a way of sending users information is essential

In addition to email, other methods can accompany it such as SMS or text messaging. This is not as universal as email but it can be a very useful alternative for specified mobile users.

STEP 6 – SYSTEM AND USER CONFIGURATION

Cloud Computing applications are centralized in their installation but the users can access them from any location. It is therefore important to have a central method of configuring the options and settings for the entire application while also storing user specific options for each user. Each user therefore would have a user profile stored on the server. This would contain selections on each screen so that when the user returns to the same screen, defaults would display specific to that user. In addition to user specific profile settings, a system wide series of configuration options can be set by an application administrator. This will have settings for the application behavior for a site or organization. An example could be the length of time for the expiration of passwords or the length of time before the user is timed out and needs to relogin.

figure 10 – System wide configuration can be listed as parameters that can be set by an administrator.

STEP 7 – BUSINESS LOGIC

Each application has its own unique business logic. This can be to perform sophisticated analytics on a particular set of data specific to an industry. An example is to perform statistical analysis for clinical trials data to ensure a certain level of quality or statistically significant clinical safety or efficacy of a drug. In these cases, the computing speed and power needed to crunch the business rules are more effectively handled with a larger server multi processor with disk IO speeds optimized with high level of RAID. In these cases, you will be able to perform your business logic with greater efficiency in a Cloud Computing environment compared to executing things at the desktop computer.

figure 11 – An example business logic of handling sophisticated clinical data across the cloud

10

STEP 8 – SCALABLE SERVER

You may start out with just one or two team members using an application that is processing a small set of data. However, with dynamic business demands, this can grow into a larger group with a large set of data. The ability to scale to meet the computing demands in a dynamic environment is key to your success. Cloud Computing can scale with increasing number of application servers that can be added or removed depending on the need. This can be accomplished on the server by implementing VM (virtual machines) and/or the dynamic control of assigning the number of application servers used to serve the application.

In this example, an application is managed by Websas which dynamically assigns different numbers of application servers to a specific application. The number of application servers is described as a number in square brackets. Each application server handles the request on a different TCP socket port. This allows dynamic communication on different ports so there are no conflicts between the application servers. In this example, the application ClinicalSMS is described in abbreviated form as:

   clinicalsms [3] 5025-5030

This means that the application ClinicalSMS has been provided with three application servers which use TCP ports 5025 through 5030. When the users place a request from their browser, it will distribute the load upon three separate servers running the same application. Each application server is communicating through its own TCP socket to prevent communication collisions. This can easily be reconfigured to be: clinicalsms [30] which means that it will use 30 distinct application servers handling all the requests for this application for a large set of users.
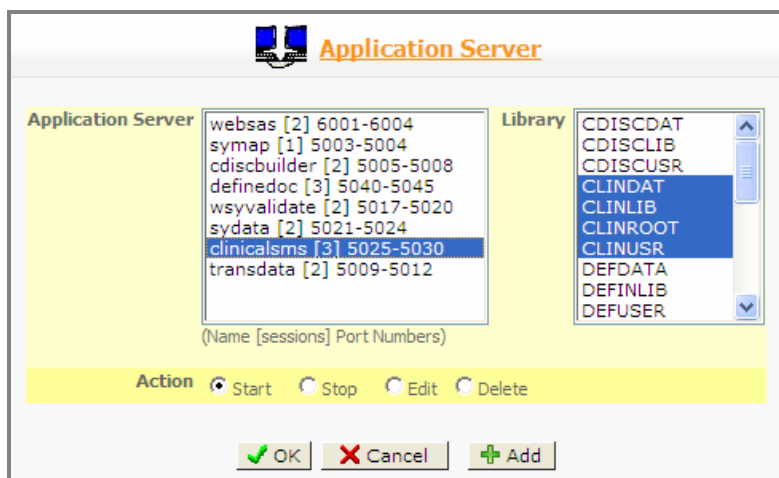

figure 12 – A dynamic set of application servers can be assigned to each application served to changing set of users.

The flexibility for an administrator to dynamically add or remove applications depending on demand is essential in optimizing a Cloud Computing application. This allows the application to scale to the changing user demands of the application.

CONCLUSION

The more complex software becomes, ironically, the more simple the user machines need to be to run them. This is because the goal is to simplify the usage and management requirements so that the user can access sophisticated software from any computer with an internet access. The growth of web technologies driven by ecommerce and social network technologies accompanied by economic constraints has acted as a catalyst for implementing solutions on Cloud Computing. The steps needed to implement a successful solution with this Cloud Computing environment requires changes in how software is implemented both on the client computer and also on the server. The foundation for changes on the client machine has already taken shape due to ecommerce and social technologies. The challenge remains for complex solutions such as SAS to adjust and adapt on the server side to function as a service within Cloud Computing.

CONTACT INFORMATION

Sy Truong is President of MXI (Meta-Xceed, Inc.) They may be contacted at:

Sy Truong
MXI, Meta-Xceed, Inc.
2185 Oakland Rd
San Jose, CA 95131
(408) 955-9333
sy.truong@meta-x.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.

ClinialSMS, Websas and other brand and product names are trademarks of Meta-Xceed, Inc.