CLOUD-BASED RESEARCH INFORMATICS
IMPROVE COLLABORATION, AGILITY AND TCO ACROSS DIFFERENT INDUSTRIES

White Paper

SELECTING A CLOUD-BASED SOLUTION TO SUPPORT COLLABORATIVE RESEARCH, DELIVER OPERATIONAL AGILITY AND REDUCE TCO

To address rising cost and risk pressures, improve innovation and focus on core competencies, many science-based organizations are moving collaborative relationships beyond traditional boundaries and creating flexible networks of researchers. Some are in-house; others are with industry and academic partners, research institutes, consortia and contract research organizations (CROs). Over time, these externalized networks are increasing in size and complexity. Many combine numerous partners with diverse objectives involving single or multiple research projects that, in some cases, can tie up more than 50% of a commissioning organization's IT budget. Internet-based collaboration solutions such as email, SharePoint, VPN, Citrix and other data exchange mechanisms often introduce security challenges, incompatible data formats and the need to prepare and curate files manually. These difficulties can reduce productivity, decrease data quality, lengthen project timelines and increase failures.

With these challenges in mind, combined with an enormous pressure to reduce their informatics footprint, organizations in life sciences; chemical; consumer packaged goods; energy, process & utilities; industrial equipment and other industries are turning to cloud-based solutions as a scalable, secure, state-of-the-art environment for research collaboration. With cloud adoption significantly enhancing collaborative projects, increasing operational agility and lowering total cost of ownership (TCO), cloud computing has become a valuable and viable solution today; however, organizations are often uncertain about the best way to evaluate, select and implement a cloud collaboration platform.

This paper will discuss the challenges of internal and external collaboration and how scientific research organizations can manage the transition to a more highly performant, flexible and agile partner workspace in the cloud with the help of a qualified cloud provider. The paper will also consider the benefits of cloud computing and the questions to ask when investigating and selecting a cloud-based collaboration solution.

LIFE SCIENCES **CONSUMER PACKAGED GOODS** **ENERGY, PROCESS & UTILITIES** **INDUSTRIAL EQUIPMENT**

Figure 1. Cloud-based collaboration can improve the efficiency of internal and externalized research projects across diverse industry sectors, while also securing IP, enhancing agility and minimizing IT costs and effort.

SCIENTIFIC, BUSINESS AND IT CHALLENGES

In today's fast-changing project landscape, creating and managing a complex high performance computing (HPC) and IT infrastructure to handle, analyze and share data presents significant scientific, business and IT challenges. From the scientist's point of view, it can be difficult to communicate effectively with partners down the hall, let alone in different time zones, cultures, languages and geographies. Data exchange with partners is often manual and therefore error-prone; and every collaborative project brings its own set of data representations. Highly-paid scientists can spend up to 50% of their time manually processing and checking collaborator data prior to analysis. Errors may go unnoticed for weeks or months, resulting in significant project delays and IP risk. Most importantly, the level of effort required to understand different types of data originating from many different sources can be formidable. External collaboration can require scientific staff to radically change the way they access, manage and interpret research data, disrupting established workflows and even requiring significant retraining. As this is often not feasible, access to collaboration data is sometimes restricted to a few "gatekeepers," further extending project timelines and reducing efficiency.

Externalized collaboration comes with numerous business challenges. Managing multiple partners, establishing key performance indicators, assessing results and protecting intellectual property (IP) across multifaceted partnering networks are costly, time-consuming, labor-intensive activities. With so many puzzles to deal with, there is little wonder that externalized collaborations typically do not run as smoothly as internal projects, especially as the number and complexity of partnerships increase. As a result of these challenges, external projects often do not meet their original expectations, or worse, they fail completely. This is a huge risk factor as organizations rely more than ever on external partners to advance their discovery initiatives.

The challenges for IT departments tasked with building and managing partner workspaces are especially daunting. The IT difficulties spring from in-house data management systems that are typically not designed to support externalized research. An organization's internal data management infrastructure tends to be structured on the assumption that scientists working for the same company can, for the most part, freely access the data they need to make informed decisions—and that the company owns all of its IP. Externalized collaboration networks are different. External partners are likely authorized to see only the data that applies to them, and the ownership of IP resulting from collaboration sometimes rests with the commissioning company, or it may be spread among the partners. Likewise, collaboration and CRO networks are not static but evolve over time with different partners involved in different stages at different times. As a result, externalized collaborations often need to be spun up and down quickly with partner data securely partitioned in the system so that it can be managed and shared with selected internal and external researchers as projects progress. Additional IT challenges include synching incoming partner data with in-house legacy and on-premises systems and making sure that all data is secure. Reliable, automated procedures for tech transfer, data standardization/harmonization and data transfer to legacy databases can be expensive to implement and maintain and often require highly skilled software developers.

A solution: To **boost innovation** and **improve** R&D ROI, more companies are turning to **external collaborations.**

Who's collaborating?

25%
External research

23%
Academic institutions

17%
Consortium collaboration

Everyone

24%
Internal, multi-site collaboration

11%
CRO/CMO providing collaboration tools to client

And how often?

75%

of respondents had **10 or more** collaborations.

These collaborations took an average of

SIX WEEKS
to get started

(That is a long time to wait!)
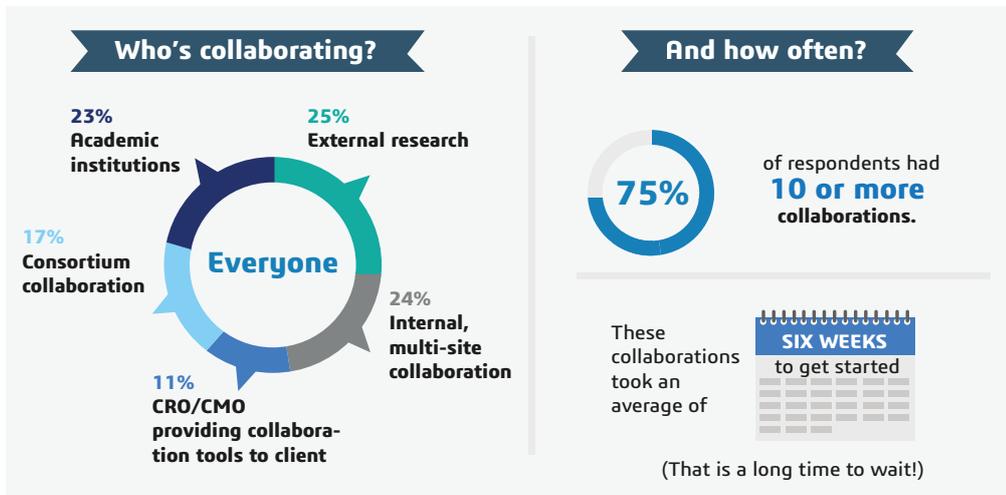
INCREASE OPERATIONAL AGILITY

Many companies are looking to lower costs and become more productive in operational environments where IT infrastructure (hardware and personnel) is expensive and not viewed as a core competency. They need to get projects up and running quickly, while research goals and software needs are continuously evolving. As a result of these critical business issues, many companies are taking a cloud-first approach to moving their informatics capabilities to the next level. Not only does the cloud accelerate the provisioning of computer resources, it also helps companies roll out new offerings more quickly—accelerating their business responses to changing conditions or opportunities and improving their competitive position. This improved operational agility can be even more compelling than shortening the provisioning timeframe.

REDUCE TOTAL COST OF OWNERSHIP

Some companies are requiring up to 80% reductions in capital and operating budgets in response to economic pressures. As a result, they are exploring cloud computing as a way to lower costs by eliminating up-front capital investments and ongoing maintenance costs associated with on-premises solutions. While point solutions are no longer sustainable, there is also low tolerance for software complexity and increased pressure to consolidate software vendors. Because it eliminates the need for companies to buy, deploy and maintain IT infrastructure and application software individually, the cloud is a feasible and attractive strategy for lowering TCO today.

EXPLORE THE COLLABORATIVE CLOUD LANDSCAPE

Faced with these data management/analytics, partnering and IT challenges, organizations are looking more and more to the cloud as a solution to relieve administrative burdens, speed analytical workflows and improve collaboration. A 2016 survey of IT and business decision makers conducted by IDG and published by Informatica found that 68% of respondents intend to investigate or deploy cloud analytics solutions over the coming year.[1] In addition, 74% said they expect to adopt a hybrid or cloud-only approach to analytics over the next three years. Respondents with currently deployed cloud analytics solutions cited the advantages of lower up-front costs (60%) over on-premises solutions, as well as greater agility and faster time to market (61%), more rapid and cost-effective scaling for larger data sets (60%) and self-service capabilities for non-technical users (51%). It is noteworthy that IT executives are assuming a leadership role in this emerging migration to the cloud, as confirmed by another 2016 survey of 1,850 executives and managers conducted by ServiceNow. A majority, 52%, report having "cloud-first" policies when making new technology purchases, an adoption stance that will increase to 77% within the next two years.[2]

When we speak of the cloud, we are really referring to cloud services that are available in three basic flavors: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). The SaaS model for centrally hosting, licensing and delivering software on a subscription basis (sometimes referred to as "on-demand software") manages everything for the subscriber including applications, security, databases, operating systems, virtualization, servers, storage, networking and data centers. For the purposes of collaborative research, this can be preferable to other cloud service options that require subscribers to manage their own applications (in the case of PaaS) or their own applications, security and databases (in the case of IaaS). A secure, flexible and scalable SaaS-based platform that manages the entire environment can better support the end-to-end workflows associated with collaborative projects whether they are undertaken in a public or private cloud computing workspace.

Selection of a public or private cloud computing environment is an essential component of an overall cloud strategy. The public cloud offers the greatest level of efficiency in shared resources by providing scientific applications securely over the public Internet via a system operated by a public provider (e.g., Amazon Web Services or Microsoft). Public cloud services in a multi-tenant environment generate economies of scale and enable sharing of resources that can reduce costs and improve technology availability while keeping each client's data separate and secure through logical partitioning. IDC states that the worldwide revenue from public IT cloud services exceeded $69.6B in 2015 and estimate that it will reach $141B in 2019, a compound annual growth rate of 19.4%[3]. IDC also expects 65% of companies' IT assets to be located offsite in colocation, hosting and cloud datacenters by 2018.[4]

Figure 2. Cloud services in a multi-tenant data center generate economies of scale and enable sharing of resources that can reduce costs and improve technology availability while keeping each client's data separate and secure.

[1]Forbes, Joe McKendrick, "Cloud Computing Becomes a Home for Data Analytics," February 6, 2016
[2]Forbes, Joe McKendrick, "Cloud Computing Is Moving Technology Executives Closer to the Boardroom, Survey Shows" October 29, 2016
[3]"The Compelling Case for Cloud Computing in Life Science Research," Alan Louie, Ph.D., Research Director, IDC Health Insights, Dassault Systèmes Presentation; November 16, 2016
[4] IDC FutureScape: "Worldwide Cloud 2016 Predictions—Mastering the Raw Material of Digital Transformation," November 2015, pg.2

Private cloud computing enables customers to access applications via their own intranet which is connected through a Virtual Private Network (VPN) to a datacenter hosted by a cloud operator (e.g., IBM or HP). The private cloud provides separate virtual data storage for every customer while requiring them to manage their own private cloud infrastructure—offering customers a high level of security and control but also increased capital expenditure and management responsibility.

Depending upon where an organization operates along the research-development-manufacturing continuum, it will likely gravitate towards public cloud hosting if they are in a primarily non-regulated space or to private cloud hosting if they operate in a highly regulated space like development or manufacturing. Public cloud offerings for non-regulated environments need to support the full breadth of scientific informatics and collaboration capabilities that are essential in today's collaborative upstream research space (e.g., Electronic Lab Notebooks (ELNs), data registration, data pipelining, etc.). Private cloud offerings supporting regulated operations should provide the same capabilities in downstream development and manufacturing where organizations can manage their own private cloud, or select a third-party hosting service to oversee their private cloud infrastructure (e.g., ELNs, Laboratory Information Management Systems, manufacturing quality/compliance capabilities, etc.).

Selecting an optimal SaaS-based, public and/or private cloud hosting solution is only the first step toward embracing innovation in the cloud. There are, of course, additional challenges and steps when it comes to actually deploying a cloud solution, especially for organizations with little or no cloud experience. Utilizing the cloud can radically alter existing data access expectations within an organization, posing governance challenges as you move from planning to deployment. The best way to ensure compliance from stakeholders (and deal with cloud skeptics) is to create a strong cloud vision and build strategies that align with organizational needs. Simplify governance issues by transparently assessing impactful use cases, current workflows/processes, security risks and the highest priority functions to be moved to the cloud. Also carefully consider the cultural changes and new policies that will be needed to streamline governance and align employees with the new cloud environment. Provisioning and managing users and groups across the smorgasbord of cloud apps the organization needs can quickly become complicated. Your SaaS provider should be able to guide you through the process of creating and managing groups, controlling who has access to apps, enabling self-signup, managing password requirements and the many other business process changes that come with a move to the cloud.

Other high-level considerations when assessing and selecting a SaaS-based information management and collaboration environment are:

Should a cloud system integrate with existing on-premises solutions?
In many organizations, cloud adoption will be a staged process in which scientists accomplish their work using both existing in-house (on-premises) systems as well as cloud collaboration. A hosted informatics/collaboration system needs to support this 'hybrid-cloud' environment in which data flows between on-premises and cloud applications. For example, a design team can start an experiment in an on-premises ELN and then transfer the information to the cloud for execution at a contract research organization; or molecules registered into a cloud database can be automatically synchronized into a company's on-premises corporate registration system. This bi-directional integration can facilitate an organization's staged movement from existing legacy infrastructure to an emerging cloud-hosted environment.

How can scientists focus on research instead of cleaning up data from different sources?

Some organizations claim that it can take up to two days per week for scientists to clean up scientific data they receive from collaborators. The cloud system should minimize time lost on data transformation and interpretation by ensuring that collaborative data complies with proprietary 'business rules' established by the sponsoring organization. Organizations will have a difficult time storing and disseminating collaborative data that has not been standardized using these rules, and they will face the possibility of costly errors that can bring projects to a halt. The cloud system should also support data annotation, which can provide critical context for other project team members using the data.

How can IT specialists and developers extend the system to support specific tasks?

The system should be highly configurable, allowing for the easy definition of sites, projects and user roles with appropriate data access and upload permissions for all. The system should also be extendable to support the creation and management of scientific services and the implementation of standard business rules enabling integration between cloud and on-premises systems with harmonized data flowing automatically from one environment to the other.

Figure 3. Poor communication with remote partners means it can take up to two days per week for scientists to clean up data they receive from collaborators. A cloud system can minimize time lost on data transformation and interpretation by ensuring that data complies with proprietary business rules.

Is it possible for managers and team members to receive continuous updates on project status?

The system should support data and project dashboards accessible from mobile devices to provide instant snapshots on collaboration progress. When managers have full visibility into projects in real time, they can plan and schedule tasks in the most efficient manner possible.

Can decision making be handled as a group including external project members?

With the cloud workspace connecting all project team members and enabling real-time data synchronization, data should flow seamlessly across the network and project lifecycle. Permitting all scientists, internal and external, to see in-house and networked data in a single place will streamline decision making for the entire team.

How can the system securely capture and manage IP collected during and after project completion?

The cloud-based collaboration workspace should protect each organization's IP, controlling and auditing access and ensuring that each participant can see only the information they need and are authorized to access. The system should protect IP by holding it in a secure neutral zone until ownership is clear. This is important because there are many contractual models for collaboration and some do not resolve IP ownership until final results are known and the project ends.

Can collaborations be started and stopped quickly?

Given the short-lived nature of many research collaborations, the cloud workspace should provide the agility to spin up and spin down new projects quickly, as well as the ability to add or remove collaborators from existing projects with minimal effort. User credentials and a user name will be all that is needed to get a new partner up and running and integrated with the project lifecycle.

WHAT DOES A SUCCESSFUL CLOUD SOLUTION LOOK LIKE?

Science-based organizations investigating the feasibility of cloud-based data management and collaboration should look for a hosted system that provides the following essential characteristics of a successful cloud solution.

Ensure a secure SaaS infrastructure

Data security is top of mind for organizations moving critical collaboration infrastructure to the cloud. You should align your organization with a SaaS provider that offers an infrastructure and set of applications with strong data safeguards. ISO 27001 is one of the most widely recognized and internationally accepted best-practice standards for information security management. Ideally, the SaaS provider you select will have earned ISO accreditation for the systems, technology, processes and data centers supporting their cloud environment based upon a formal audit confirming their compliance with the standard. You need complete assurance that your confidential and proprietary information remains secure in the cloud. The systems in place to safeguard your data should be certified to an accepted industry standard now and in the future.

Provide secure data and document storage and sharing

A successful hosted collaboration system will allow multiple partners to capture the structured and unstructured information that they are responsible for generating. They can then quickly and securely share this real-time information with the rest of the team and immediately access all the data from other collaborators that they need to advance their parts of the project. If your organization is commissioning research, there is a benefit to keeping this data in the cloud, because there is a clear delineation between your hosted collaboration system and your in-house server-based systems and data. Data automatically flows and resides where it needs to be depending on who owns it, who needs it and what it is for.

Access a broad range of extensible scientific applications

By offering integrated and extensible hosted applications, a cloud collaboration system can support a range of scientific workflows appropriate to your needs, enabling project team members to access the applications and data they need, wherever they are, at any time. With no software to manage and no hardware to maintain, your organization can improve operational excellence by conducting globally networked research in the cloud with significantly reduced IT cost and effort. Typical scientific research applications include:

- A flexible, multi-disciplinary electronic lab notebook that enables sponsor organizations and network partners to capture and share experimental methods, register molecules and biologics and manage assay data in a collaborative environment

- A scientific workflow authoring application supporting the creation and management of scientific services, data harmonization/synchronization, implementation of standard business rules for partners, application extensions and integration between hosted and on-premises systems, enabling data to flow from one to the other.

- A mobile application enabling scientists to access both cloud and server-based systems where they can execute tasks and share results with colleagues using mobile devices in a hybrid environment.

- Data analysis and visualization applications providing scientific analysis, data pipelining, visualization and charting capabilities for use with mobile devices, either on-premises or in the cloud. Scientists should be able to construct informative workflow protocols from pre-built services in chemistry, biology, imaging and other scientific domains and easily share workflows and results with colleagues in a browser or on a mobile device.

- An open portal where scientists can publish and share protocols and reports based on captured data, enhancing broad dissemination of essential research tools within the scientific community.

- Applications for managing inventory and tracking biologics are also critical in modern research labs.

Leverage social networking to accelerate project-centered science

Whether a project is internal or external, good communication is critical to staying on track and achieving a successful outcome. Today people are used to communicating via mobile devices on social platforms like Facebook and Twitter. Virtual teams can leverage social networking to advance collaborative project-centered science with real-time data access. For example, when a team member creates a new experiment and collects new data, he/she can immediately share this news with the entire team via social notifications, so they can drill in and view the data on their mobile phones. Cloud-based collaboration empowers researchers to capture and search user and application feeds, post from internal/external systems and speed project execution by building and sharing crowd knowledge.

Figure 4. Virtual teams can advance collaborative project-centered science in the cloud through social networking with real-time data access.

NOW IS THE TIME

Cloud-based collaborative discovery allows researchers to exchange scientific data and protocols securely and in real time—optimizing internal and external research projects, securing IP and accelerating innovation. As collaborative research drives secure cloud adoption across industries, cloud infrastructure will be integral to the adoption and integration of technological innovation and digital transformation. Cloud solutions also remain the best approach for dealing with the tidal wave of new data types and sources engulfing collaborative research today. Most importantly, the cloud delivers improved agility and lowers total cost of ownership for organizations tasked with responding more quickly to changing business needs while also lowering costs. Being in the cloud means you can set up a robust collaboration system—accessible anywhere, anytime—with minimal IT support quickly and easily, and you only pay for what you need and use.

By 2020, at least a third of all data will pass through the cloud.[5] More than 43% of organizations expect that the majority of their IT capability will be delivered through public cloud services by 2020, and they will access 78% of IT resources through some form of cloud—public, private, or hybrid—by 2018.[6]

The numbers speak for themselves. Now is the time to improve competitiveness by defining and beginning to implement the best cloud strategy for your organization. Contact Dassault Systèmes BIOVIA for more information on leveraging the cloud to improve agility, reduce cost and advance collaboration when working on new scientific discoveries and technological innovation.

[5] Forbes, Bernard Marr, "Big Data: 20 Mind-Boggling Facts Everyone Must Read," September 30, 2015

[6] IDC FutureScape: "Worldwide Cloud 2016 Predictions—Mastering the Raw Material of Digital Transformation," November 2015, pg.2

Our 3DEXPERIENCE® platform powers our brand applications, serving 12 industries, and provides a rich portfolio of industry solution experiences.

Dassault Systèmes, the **3D**EXPERIENCE® Company, provides business and people with virtual universes to imagine sustainable innovations. Its world-leading solutions transform the way products are designed, produced, and supported. Dassault Systèmes' collaborative solutions foster social innovation, expanding possibilities for the virtual world to improve the real world. The group brings value to over 190,000 customers of all sizes in all industries in more than 140 countries. For more information, visit **www.3ds.com**.

3DEXPERIENCE®

DASSAULT SYSTEMES | The **3D**EXPERIENCE® Company

WP-5505-0417