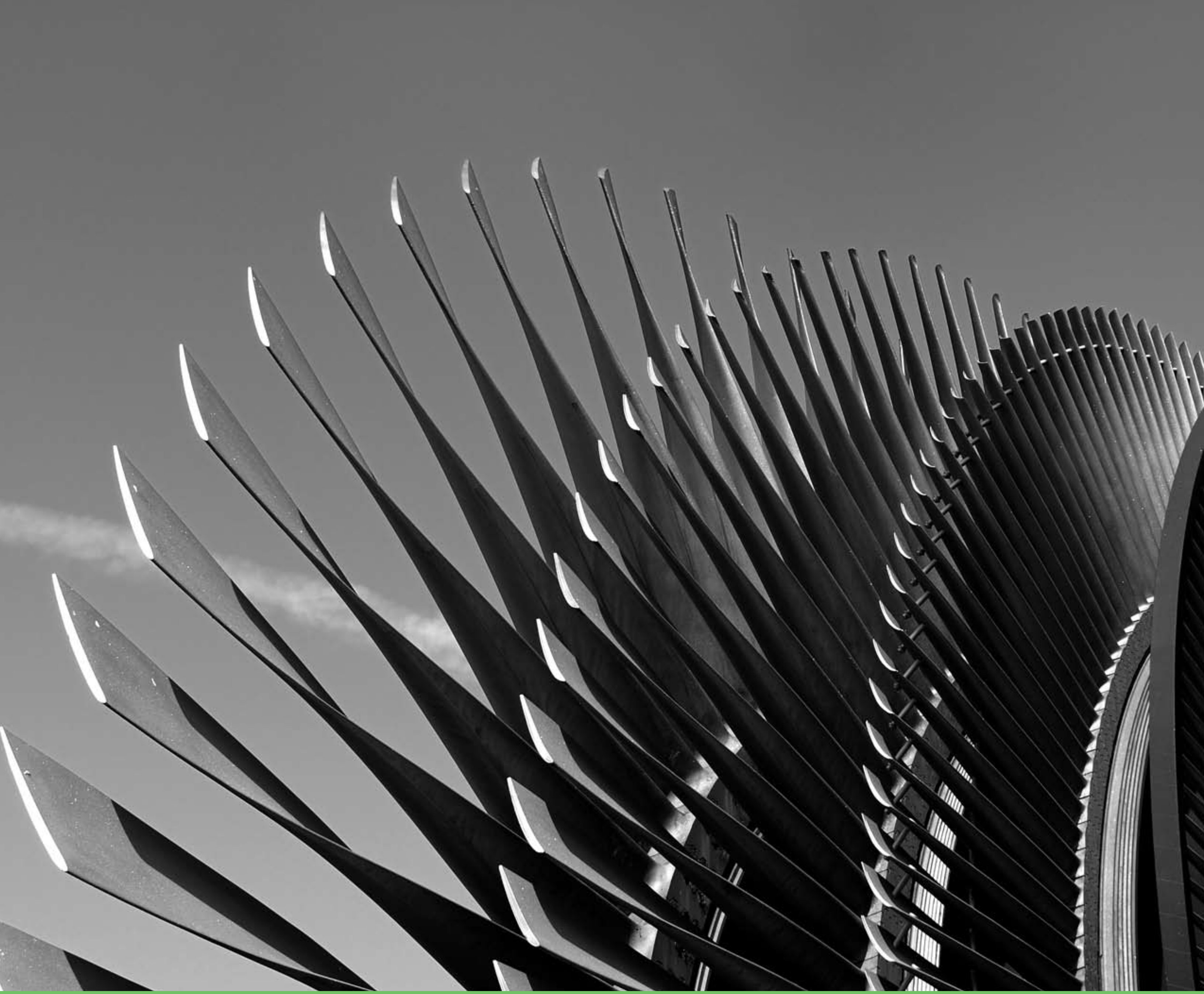

The importance of HPC management software for the enterprise deployment of simulation lifecycle management tools

A Platform Computing white paper

Author:
Joseph Cieslak, P.E.
Platform Computing
jcieslak@platform.com

Platform Computing

Contents

1. Introduction

Simulation lifecycle management (SLM) is best defined as the capture and management of engineering knowledge acquired from computer-aided-engineering or simulation tools performing product design validations. SLM addresses the need for data management, authentication, historical reference, process automation, collaboration, and IP protection – this is similar to the advantages product lifecycle management (PLM) has provided for the users of computer-aided-design tools over the past decade.

SLM benefits members of the entire engineering enterprise. This ranges from simulation experts who want to spend less time manipulating files and more time focusing on true design optimization, to designers, managers, and manufacturing experts who can now become involved at the critical conceptual stage of product design where simulation has the greatest impact.

SLM provides benefits to the entire engineering enterprise including process monitoring, error handling and a configurable notification and data management framework. But to take full advantage of SLM, several IT infrastructure challenges need to be addressed. This paper explains the necessity for integrations between SLM products and workload management (WLM) frameworks to make them robust enough for enterprise-wide deployment. It focuses on major areas where Platform Computing's technology can add immediate value to SLM functionality, including:

- Optimizing the use of distributed, heterogeneous IT resources
- Data-aware scheduling for dispersed data sets
- Integrating enterprise environments using multiple resource managers
- Intersection of SLM with cloud computing

The white paper draws on Platform Computing's extensive experience in high performance computing (HPC) management solutions that dynamically connect distributed computing resources and data to workload demand according to business policies. Since 1992, users of simulation applications from a variety of markets have benefited by leveraging integrations with Platform LSF® because of its advanced distributed resource management functionality for enterprise-wide computing. Platform Computing software helps manage nearly five million CPUs around the world, driving faster performance, delivering more flexible sharing, and ultimately, enabling the best possible use of valuable IT resources.

2. Optimizing use of distributed heterogeneous IT resources

Leading manufacturing organizations around the world recognize that the earlier engineering simulation can be introduced into a product's lifecycle, the more effective it is in reducing manufacturing cost and accelerating time to market. Due to recent advances in computational processing power, simulations that just a few years ago were performed as discrete, decoupled events are now being replaced by multi-disciplinary simulations capable of evaluating several performance variables at one time. Multi-disciplinary simulations use process managers to couple various applications used to evaluate structural, fluid flow, electromechanics and impact properties, into a single, virtual prototype. This approach more accurately represents actual physical behavior.

Many organizations have successfully used SLM techniques on individual projects or at a department level. However, extending SLM concepts to a broader enterprise user base is a major challenge. IT organizations faced with complex environments have typically designed and managed their infrastructure as application- or line-of-business-based silos. Historically these silos were deployed and optimized to run a single process or set of applications. A thick-client application handled all data access and business logic. Today's SLM products use web-based or services-oriented architectures that allow for network-wide access and data sharing among users. Unfortunately these products still view computing assets and third-party application software as dedicated resources made available without rules for prioritization.

SLM tools used in segregated workgroups are simple and manageable in this context because there are only a handful of infrastructure variables to manage – e.g. users, workstations, compute servers, applications licenses, and data servers (Figure 1).

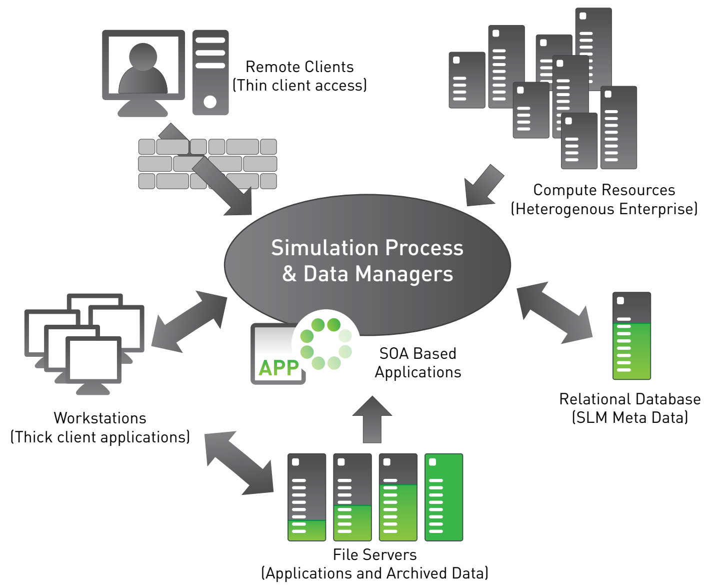

Figure1. SLM infrastructure building blocks

However, to make SLM a comprehensive solution for an entire enterprise several IT infrastructure challenges need to be addressed. They include:

- How to integrate SLM's service-oriented architecture (SOA) applications within existing engineering compute infrastructure – SLM process managers may need to call out several applications for a single multi-disciplinary simulation including:
  - o Parallelized impact and computational fluid dynamics (CFD) software optimized for efficiency on Linux clusters
  - o High end visualization tools and proprietary engineering tools on Windows-based servers
  - o Computer–aided-design software for geometry creation and spreadsheets that run exclusively on Windows workstations
- Providing SLM process managers with access to globally dispersed compute servers, application software and data sets using heterogeneous hardware and operating systems without disrupting locally-tuned resource prioritization
- Security and tracking of simulation processes and data

- Application software license prioritization and management
- Integration with workstations and cluster-based servers that use new multi-core and GPU chipsets
- Integration with high speed interconnects for applications that use parallel processing
- Prioritization of simulation events based on process interdependencies
- Job monitoring and administration to assure effective resource utilization

To meet these challenges, simulation applications have been re-written, removing the barriers created by their previous client-server structure. Now operating within a service-oriented architecture (SOA), both the application and its associated data can be accessed by users located across the entire engineering enterprise.

But just as SOA reconstitutes simulation software as a service to make SLM a reality, the computing resources used for SLM need to be virtualized as well in a service-oriented infrastructure (SOI) – this allows the sharing of infrastructure between multiple applications and user groups.

Platform ISF provides a SOI that manages resource supply across the enterprise, decoupling it from resource demand. This allows SOA-based SLM software to send requests through Platform ISF, ensuring that IT resources are made available when and where they are needed to provide all consumers with appropriate service levels (Figure 2).

Platform ISF has been designed to provide an additional layer of resource management that is operating system agnostic. In the case of a SLM process manager requiring applications in a predetermined sequence, Platform ISF marshals the required resources as they are called out in the workflow. During this process, Platform ISF allows discrete instances of Platform LSF or other third-party grid and cluster products to orchestrate their current jobs, tasks and sessions as usual.

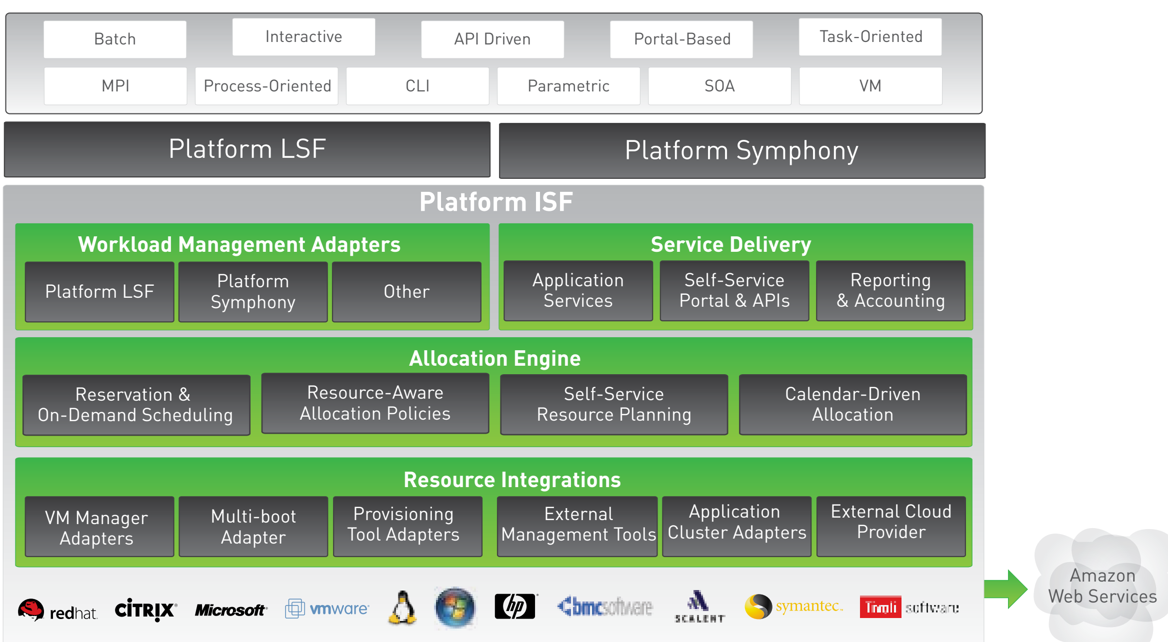

Figure 2. Platform ISF components

Platform ISF is the culmination of almost two decades of Platform's experience in managing enterprise wide complex production data centers which need to:

- Deliver IT resources faster with minimum TCO
- Enable higher utilization of resources
- Achieve service levels at lowest possible cost
- Reduce data center administration and infrastructure costs
- Optimize SOA and cloud computing infrastructure

Platform ISF is leveraged by Platform LSF and Platform Symphony, and is an essential component of SAS Grid Manager, an OEM component provided exclusively by Platform Computing (Figure 3).

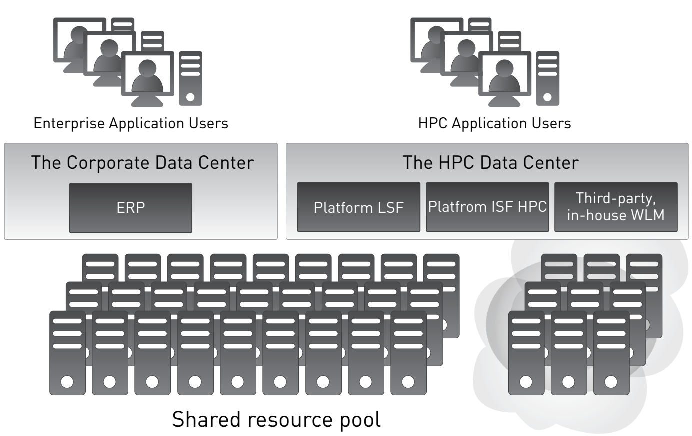

Figure 3. Platform ISF in a heterogeneous enterprise

3. Data-aware scheduling for dispersed data sets

SLM driven applications need access to data stored on either local or remote storage resources in order to place the data into an area where it can be used by one or several CPUs in the enterprise executing a simulation. If the resource management system does not explicitly account for the amount of data each job needs to access, some jobs will occupy the CPU resources for much longer than necessary, spending significant amounts of time waiting for file transfer. Since an HPC system can have as many as 100,000 multi-threaded processors and can cost many millions of dollars, it is imperative to operate these systems efficiently.

In a typical HPC system, while some jobs are waiting for data to come in, the CPU can work on other jobs. However, job-level schedulers embedded in current SLM process managers usually do not provide this level of optimization when sharing CPU resources with other consumers. This results in unacceptable throughput. Therefore, the addition of an intelligent scheduling algorithm is critical for the successful enterprise deployment of process managers or simulation lifecycle management systems. Commercially available workload management systems like Platform LSF solve the problem of staging multiple jobs on every CPU at the core level.

In addition, Platform Computing has designed a new set of data-aware scheduling policies that can leverage existing SLM metadata to further optimize the scheduling of jobs with data set dependencies. The SLM products use Platform LSF and a data-aware scheduler plug-in to reduce job wait times. The data-aware scheduler optimizes data staging versus job wait time in a distributed environment. Based on the estimated data staging time and job wait time in each location, it makes intelligent decisions to either move the data to where the job is or move the job to where the data is. The scheduler incorporates metrics such as network bandwidth and transfer latency to ensure accurate estimates.

Working in concert with Platform LSF, the data-aware scheduler can automate the process of staging input files from remote to local storage resources. It can also increase utilization by oversubscribing the computing resource's CPUs in order to overlay data staging jobs with compute intensive jobs. And, by integrating the resource management system with the SLM meta-data, users can focus on their work, as opposed to worrying about tedious data staging and data management issues (Figure 4).

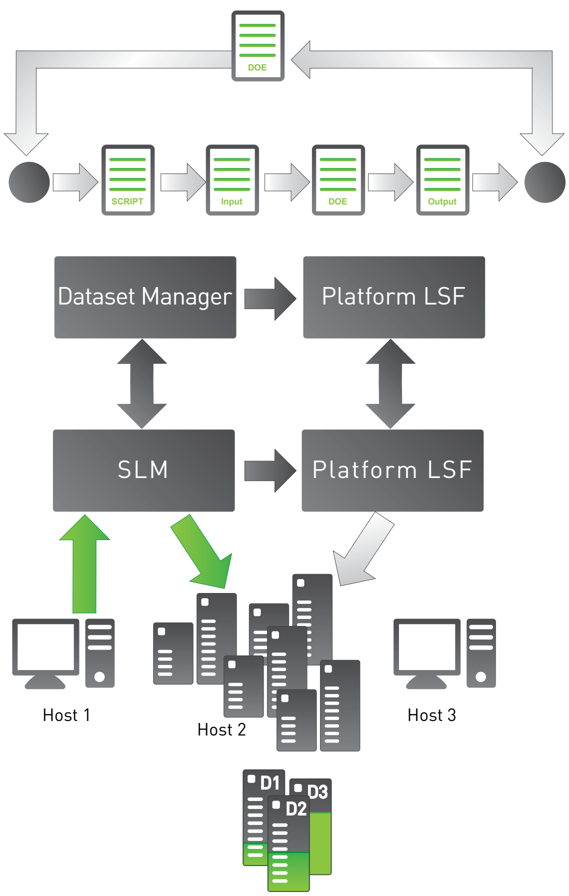

Figure 4. Data-aware scheduling architecture

4. Integrating enterprise environments using multiple workload managers

In a perfect world, a customer would use a single, resource management system with a universal workload management schema for all infrastructure assets in a SLM-driven enterprise. But as discussed above, engineering infrastructures are a mosaic of heterogeneous resources provisioned to accommodate the needs of individual departments, or maximize performance and utilization for a specific application. Even organizations standardized on a WLM, such as Platform LSF, may run several independent instances prioritized for the peak demand on siloed compute systems. As a result, SLM process managers have a difficult time integrating into existing infrastructures.

Platform ISF employs a practical approach designed to solve the issues of integrating multiple distributed resource manager instances. It allows organizations to plug-in process manager workloads, while providing individual workgroups and departments with the flexibility to deploy the WLM solutions best suited to their needs. Platform MetaScheduler also allows work to be submitted to disparate schedulers such as Sun Grid Engine™, OpenPBS, Torque, Altair PBSPro™, Microsoft Windows HPC Server 2008, and IBM Tivoli Workload Scheduler LoadLeveler™. This means that all jobs – including those running on workstations or remote clusters – can be controlled and monitored through a single management interface. From the perspective of both Platform Computing's job brokering utilities and web-based tools interacting with the Platform ISF API, jobs running on remote clusters appear similar in all respects (Figure 5).

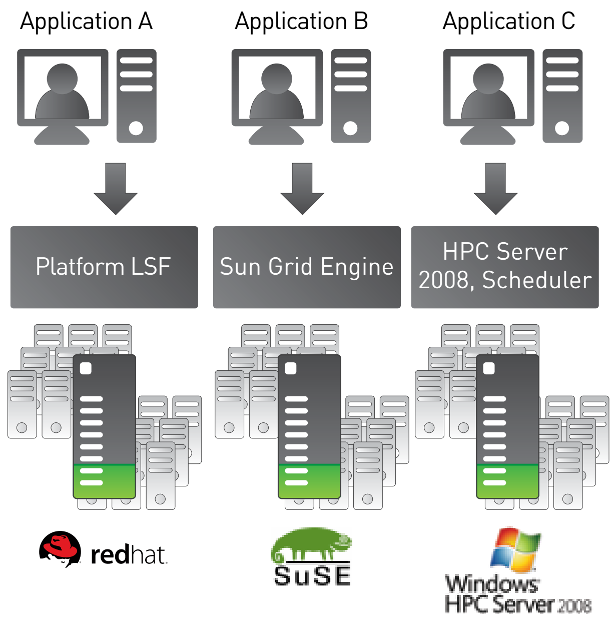

Figure 5. Platform ISF with multiple workload management systems

Platform ISF leverages the existing functionality in Platform LSF controls that handle resources processing, as well as application licenses and data staging. This allows Platform ISF to provide SLM products a way that enables users to easily unify cluster environments without having to manage the complexities of interacting with multiple workload management systems with different interfaces or scheduling methodologies.

5. The intersection of SLM and cloud computing

Cloud computing is one of the fastest growing segments of the information technology (IT) industry because, compared to traditional packaged applications, it offers a more cost-effective alternative for enterprises to achieve their business objectives. This is a new paradigm in which sharing benefits both end users and IT. Both private and public sector organizations are highly interested and moving toward potential adoption. For example, the US General Services Administration (GSA), which is responsible for $70 Million in government IT spending, now has a project known as the "federal cloud."[1]

The technology and best practices developed over the last decade in distributed computing architectures such as clusters and grids have contributed directly to the cloud computing model. Although it leverages recently developed technology, cloud computing is a business, not a technical trend, and because of its recent arrival on the scene, it often means different things to different people. But even though there are many different definitions of the term "cloud," they all have four key characteristics in common:

- **Infrastructure sharing** –Today's enterprise data centers are characterized by fluctuating resource demands from a variety of users. Cloud computing enables dynamic sharing of these resources so that demands can be met cost effectively.

- **Scalability** – To handle ever increasing workload demands and support the entire enterprise, cloud computing must have the flexibility to significantly scale IT resources. Scalability and flexibility allow the cloud provider to fulfill – or at least come close – to the promise of unlimited IT services on demand.

[1] Cohen, Reuven, "The US Federal Government defines Cloud Computing", , May 7, 2009 See: http://www.elasticvapor.com/2009/05/us-federal-government-defines-cloud.html

- **Self service** – Cloud computing provides customers with access to IT resources through service-based offerings. The details of IT resources and their setup are transparent to the users.

- **Pay-per-use** – Because cloud resources can be added and removed according to workload demand, users pay for only what they use and are not charged when their service demands decrease.

Cloud computing provides an ideal mechanism for collaboration and data sharing beyond the boundaries of an organization – a core construct for SLM. In the past, engineering infrastructures segregated their compute assets in workgroup or departmental structures with little resource sharing. Today's grid-enabled infrastructures do provide a mechanism to share resources between enterprise departments to maximize resource utilization. However, large engineering consortiums often include many individuals located around the globe, as well as companies and their supply chains that are outside the organization's infrastructure boundaries (Figure 6).

Cloud computing's promise is the ability to provision infrastructure quickly and flexibly based on actual demand and usage. For example, a virtual organization collaborating and interacting intermittently on a project-by-project basis, will find that cloud computing is the most efficient mechanism for sharing resources. Although internal or external SLM consumers do not have to own or manage the underlying infrastructure, they do have control over operating systems, computing capacity, storage, deployed applications, and networking components. And since users can be billed on a pay-per-use model, engineering organizations can benefit from cost savings, performance improvements, and enhanced operating flexibility.

Platform Computing is leading the way in delivering cloud technology to help enterprises accelerate data center performance to match the speed of business demand. One such early cloud computing deployment is the Singapore National Grid Service developed in partnership with Platform Computing, Alatum (the cloud service provider arm of Singapore Telecom),

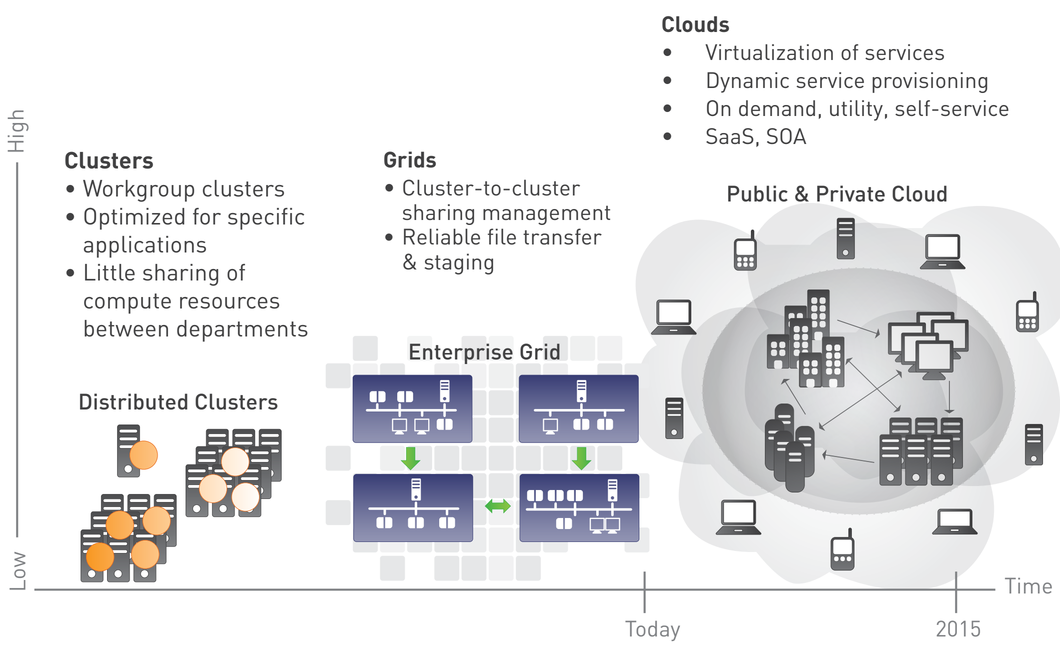

Figure 6. The evolution of engineering IT infrastructure

Hewlett-Packard, and others. The deployment is already helping enterprises and consumers benefit from affordable, secure, on-demand, pay-per-use access to high performance computing, software, and data storage services. Platform ISF enables delivery of Infrastructure-as-a-Service (IaaS), creating a shared computing infrastructure from physical and virtual resources to deliver application environments according to workload-aware and resource-aware policies (Figure 7).

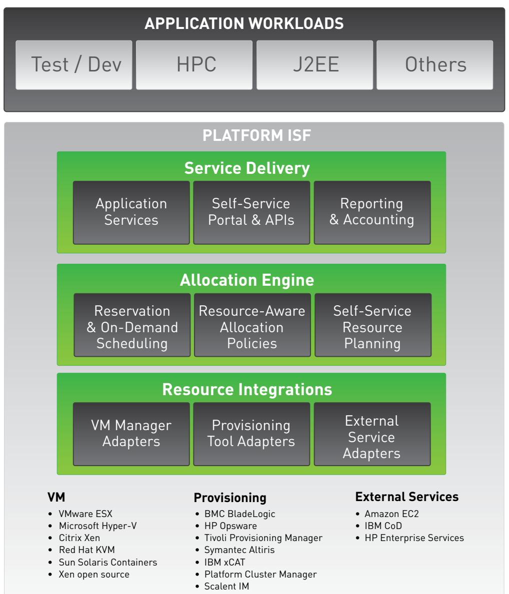

Figure 7. Three layers of infrastructure delivered by Platform ISF

The Alatum cloud allows consumers to not only access bi-directionally synchronized data, but to rent processing, storage, networks, and other fundamental computing resources and provision required software – including SLM applications and operating systems on demand[2].

Platform ISF is the leading end-to-end private cloud management software. It creates a shared computing infrastructure from physical and virtual resources to deliver application environments according to workload-aware and resource-aware policies. The software consists of three layers:

- **Resource integration** – This foundation layer integrates distributed and heterogeneous IT resources to form a shared system. All major industry standard hardware, operating systems (including Linux and Windows), and VM hypervisors (including VMware ESX, Xen, Citrix XenServer, Microsoft Hyper-V and Red Hat KVM) are supported. This layer can transparently integrate resources from external providers while maintaining its private cloud management environment.

- **Allocation engine** – Once a pool of shared resources is formed, a set of site-specific policies is configured in the allocation engine layer to ensure that applications receive the required resources. The policies also ensure that the organization's resource sharing priorities are applied, and that the quota constraints applicable to business groups sharing the cloud are reinforced. This private cloud "brain" is critical for IT agility.

- **Service delivery** – This top layer of Platform ISF provides interfaces to users and applications as well as supporting the cloud service management lifecycle. A self-service portal enables users to request and obtain physical servers and VMs in minutes instead of days or weeks. Platform ISF has a set of APIs that can be called by applications, middleware and workload managers to request and return resources without human intervention[3].

[2] Additional information about Alatum and other Platform ISF uses cases is available in "Enterprise Cloud Computing: Transforming IT", a Platform white paper (http://www.platform.com/resources/whitepapers), and in the Singapore National Grid customer Success Story (http://www.platform.com/resources/casestudies/Singapore-CS-web.pdf)

[3] To learn more about Platform ISF and cloud computing please download a copy of "Enterprise Cloud Computing: Transforming IT", a Platform Computing white paper (http://www.platform.com/resources/whitepapers)

6. Conclusion

As the industry pioneer and long-time leader in High Performance Computing (HPC) management software, Platform Computing has years of experience in dynamically connecting distributed computing resources and data to workload demand according to business policies for enterprise-wide computing. The company recognizes the need to work collaboratively with SLM solution providers and offer them integration points to Platform Computing's advanced distributed resource management functionality for cluster, grid, and cloud computing.

Platform Computing

Platform Computing is the leader in cluster, grid and cloud management software - serving more than 2,000 of the world's most demanding organizations for over 17 years. Our workload and resource management solutions deliver IT responsiveness and lower costs for enterprise and HPC applications. Platform has strategic relationships with Cray, Dell™, HP, IBM®, Intel®, Microsoft®, Red Hat® and SAS®.
Visit www.platform.com.

World Headquarters
Platform Computing Corporation
3760 14th Avenue
Markham, Ontario
Canada L3R 3T7
Tel: +1 905 948 8448
Fax: +1 905 948 9975
Toll-free Tel: 1 877 528 3676
info@platform.com

Sales - Headquarters
Toll-free Tel: 1 877 710 4477
Tel: +1 905 948 8448

North America
New York: +1 646 290 5070
San Jose: +1 408 392 4900

Europe
Bramley: +44 (0) 1256 883756
London: +44 (0) 20 3206 1470
Paris: +33 (0) 1 41 10 09 20
Düsseldorf: +49 2102 61039 0
info-europe@platform.com

Asia-Pacific
Beijing: +86 10 82276000
Xi'an: +86 029 87607400
asia@platform.com
Tokyo: +81(0)3 6302 2901
info-japan@platform.com
Singapore: +65 6307 6590
wliaw@platform.com