The Quest® Cloud Automation Platform

Private Cloud Automation for Enterprise and the Government

Closing the IT Service Delivery Gap with the Private Cloud

Despite tightly constrained IT budgets, the demand for complex services and infrastructure to support business and government initiatives continues to increase, creating a critical service delivery gap between available IT resources and the ever-growing demand. This gap materially impacts the enterprise and government agencies with lengthy and missed delivery SLAs, growing request backlogs, and expensive resource sprawl.

While the transition to virtualized infrastructure has largely helped to consolidate servers, centralize resources, and reduce some operational costs, it has also introduced new challenges. Administrators are finding it challenging to effectively manage the virtualized infrastructure to provide complex, user-focused services with their limited staff and resources. They're in need of a way to delegate administrative responsibilities and speed delivery while maintaining total control over the infrastructure and enforcing compliance.

The clear emerging path for both enterprise and government agencies is the adoption of private cloud automation technologies. ***This paper describes the fundamental technologies necessary to implement and grow a successful private cloud.***

Cloud Computing Defined

To begin, it's helpful to level-set on the definition of cloud computing. The National Institute of Standards and Technology (NIST) defines cloud computing as follows:

"Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." [NIST, v15]

Based on this definition it should be pointed out that cloud computing is not a product, but a computing model (or paradigm) that is enabled by six essential descriptive characteristics:

- **On-Demand Self-Service** – A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed, automatically, and without requiring human interaction with the service provider.
- **Self-Management** – Provisioning requests by users must not involve human interaction to fulfill these requests; therefore, workload provisioning, placement, movement, and tear-down must all be fully automated and self managed via IT policies.

- **Shared Resource Pools** – The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location-independence in that the consumer generally has no control over or knowledge of the exact location of the resources, but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

- **Rapid Elasticity** – Services can be rapidly and elastically provisioned (in many cases automatically) to quickly scale out and then rapidly released to scale back. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

- **Measured Service** – Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

- **Broad Network Access** – Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g. mobile phones, laptops, and PDAs).

Given these essential characteristics, cloud computing differs from traditional computing paradigms significantly and therefore requires purpose-built technologies that leverage IT's investment in physical and virtual infrastructure.

The Quest® Cloud Automation Platform

The Quest Cloud Automation Platform is the premier solution for creating, automating, and managing successful enterprise-class private clouds. With Quest, IT and business leadership create robust, secure infrastructure-as-a-service (IaaS) clouds to efficiently manage and deliver complex IT services across a wide range of use cases.

The patented, comprehensive Quest Platform is a purpose-built cloud automation platform that seamlessly integrates with existing data center infrastructure and automation investments, providing organizations with:

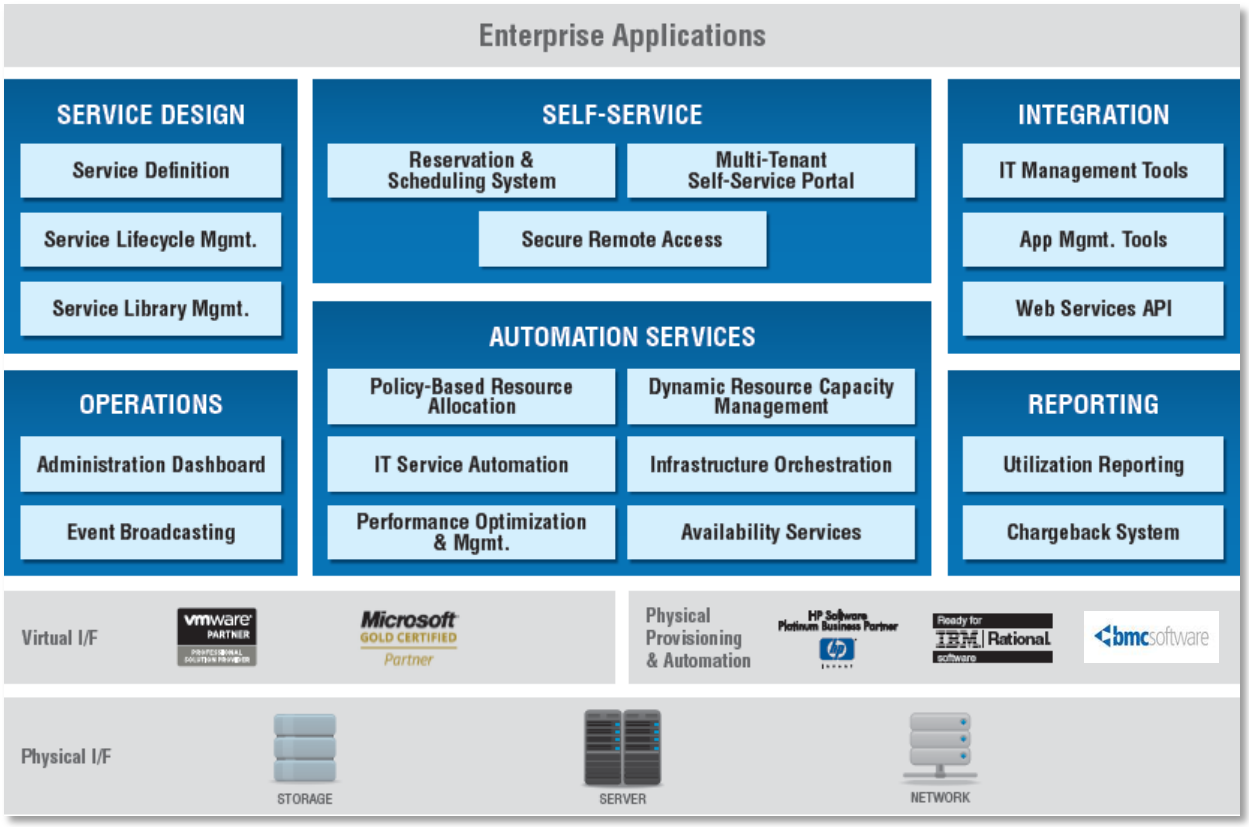
- Fully automated delivery & reclamation of complex IT services
- Policy-based self-service infrastructure delivery with guaranteed reservations
- Dynamic infrastructure capacity management & resource allocation (resource pooling)
- Scalable, elastic services
- VM sprawl prevention and elimination
- Proven enterprise scale with 160+ deployed enterprise clouds
- Rapid implementation in less than 30 days

The Quest Platform provides the most comprehensive cloud automation solution in the industry, with rich capabilities for:

- **Self Service**: End users provision IT infrastructure, services, and applications via self-service
- **Service Design**: IT engineers and architects design and create a catalog of pre-approved IT services

- **Operations**: IT administrators monitor and manage pools of dynamic infrastructure
- **Reporting**: IT managers capture, analyze, and report on user resource utilization for planning, billing, and chargeback
- **Integration**: Systems are easily integrated into IT service management systems and processes

The illustration below provides an overview of the Quest Platform capabilities. The remainder of this paper describes each functional area in more detail: self-service, automation services, service design, operations, integration, and reporting.



» Self-Service

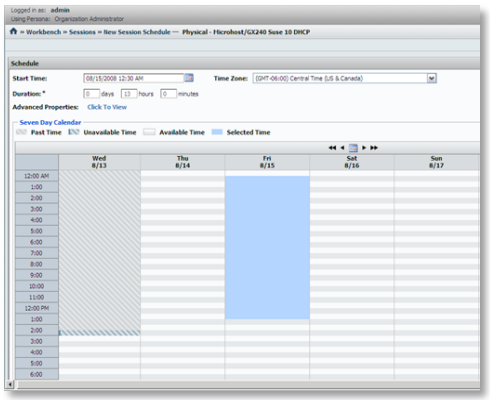
End users and consumers provision IT infrastructure, services, and applications via self-service.

Multi-Tenant Self Service Portal

The self service portal of the Quest Platform provides the user interface for end users to easily provision, reserve, and access their IT services. It provides a catalog or library of services that are made available to them based on their access credentials, including browse and search capabilities. The Quest self-service portal is multi-tenant, meaning a single instance can support multiple organizations with different user interfaces and

capabilities. The self-service portal is highly customizable and can be modified to meet the varying needs of end users. In addition, Quest provides multiple "out-of-the-box" user interfaces to support the workflows of a number of common user types, including development, test, training, demonstrations, and software evaluations.

Reservation and Scheduling System

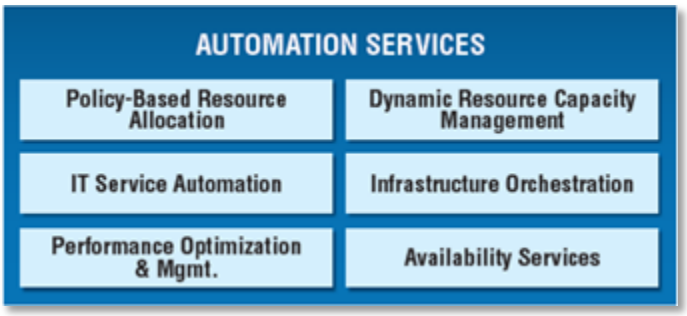


Users can provision IT infrastructure and services on-demand from within the Quest self-service portal. In cases where users want to plan ahead or when resources are not immediately available, Quest enables users to schedule and *reserve* IT services for a future time. When a user makes a reservation, the Platform reserves all of the resources – compute, network, storage – that the service requires from the start time for the full duration of the reservation. Therefore, the user is guaranteed in advance to have the service provisioned and available during the time specified. Once the resources have been provisioned, the user is notified and provided access to their environment. In addition, users have the ability to view a calendar of reserved capacity to determine future resource availability.

Secure Remote Access

The self service portal provides secure remote access to deployed IT services through a standard web browser. Browser controls enable non-VPN secure remote access to each server within a deployed service session through RDP, VNC, ICA, and VM console protocols. Services can be pre-configured with the appropriate remote access protocol auto login and authentication for each server to provide a seamless and easy access experience for users. In addition, while using a deployed service, users can snapshot, save, extend or cancel the deployed session, as well as stop, start, restart, rollback, or suspend one or all servers within the deployed service.

» Automation Services



The automation services are the core enabling technologies for all of the Quest Platform capabilities – self-service, service design, operations, reporting, and integration. These core services provide the orchestration of the physical and virtual infrastructure, IT service automation, and policy based resource allocation and management to ensure highly available and optimized services operating on shared dynamic pools of infrastructure.

Infrastructure Orchestration

The Quest Platform orchestrates infrastructure across physical and virtual servers, networks, and storage devices to support the provisioning and management of IT services and applications. Quest takes an infrastructure-agnostic approach to virtualization management and leverages multiple industry-leading hypervisors – VMware vSphere, Microsoft Hyper-V, and Red Hat KVM (in progress). In addition, the platform

leverages leading data center automation solutions from HP, BMC, and Symantec for bare metal provisioning and management of physical infrastructure. Quest utilizes an "agent-less" approach to managing infrastructure systems through remote management capabilities.

Policy-Based Resource Allocation

The Quest Platform provides capabilities to create multiple shared pools of infrastructure resources and allocate them to users, groups, and organizations. Quest provides a mechanism to populate resource pools with servers (CPU, RAM), VMs, storage, network, and software licenses that can be monitored and managed in aggregate or by specific types of resource. In addition to full server resource capacity, portions of servers can be allocated across multiple pools or just part of a server included within a pool. The Quest access control model provides for users to have certain resource access rights, as well as administrative roles to perform tasks within the system. A number of common administrative personas (access rights and roles) are included "out-of-the-box," as is the ability to develop custom personas. This provides for multiple levels and granularity of administrative privilege and task delegation. Access controls and administrative policies can be configured at the individual user, group, or organization level. These policies govern access to resources, control user consumption of resources based on pre-set quotas, and functional privileges.

IT Service Automation

IT service automation enables the on-demand or scheduled provisioning of complex multi-tiered IT services. These services can support a wide range of server, network, and storage configurations. A single service can be heterogeneous – containing servers that are provisioned as VMs on multiple hypervisors as well as on physical bare metal servers. Quest supports both image- and task-based server provisioning methodologies. At the end of a session, or when a reservation expires, Quest automates the tear-down of the deployed service, returning resources back to the pool to be used by others. Policies can be set that allow users to save the state of one or more servers within the service back to the library for re-deployment in the future. Quest supports the concept of "elastic services," which provides the ability to add or subtract one or more servers on-demand to a running service to support fluctuations of infrastructure resource capacity based on load or demand. Quest provides mechanisms to migrate an entire service (or individual VMs) from one virtualization host server to another within a pool while maintaining the integrity of the service. Finally, the Platform provides a mechanism to both isolate and clone services across multiple physical servers known as cross-host fencing. This enables rapid deployment of multiple cloned instances of a service while provide each instance with isolation and a unique addressable identity.

Dynamic Resource Capacity Management

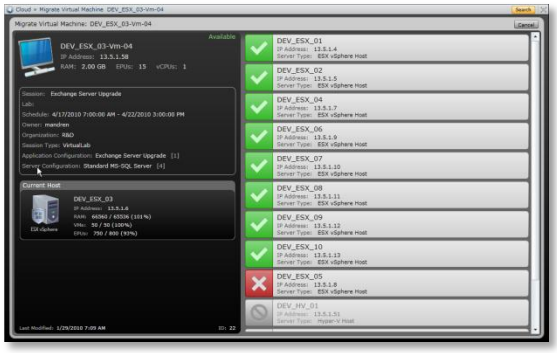
The Quest Platform provides an extensive dynamic resource capacity management (DRCM) capability that is tightly coupled with resource pooling, IT service automation, and the scheduling and reservation system. The DRCM system manages and tracks resources such as CPU, RAM, VMs, IP addresses, MAC addresses, VLAN IDs, storage resources, and software licenses within the resource pools in support of users' requests for IT services. The DRCM understands the

current utilization, current available, future reserved, and future available resources for all resource types with a given pool. Based on this pooled resource information and the resource requirements of a given user's request to deploy or reserve a service, the DRCM makes optimal decisions on where to deploy service resources on the physical infrastructure within a given pool.

Performance Optimization and Management

The Quest Platform leverages and integrates with industry leading performance management technologies to ensure IT service performance optimization across pools of infrastructure resources. Quest provides the ability to migrate a service or individual VMs within a service to improve VM workload performance when certain performance thresholds have been reached by taking action from events generated by performance management products and technologies. The Platform includes an optional integration with VMware vCenter technologies – vMotion, DRS, and DPM – to leverage virtual infrastructure performance optimization and management capabilities. In addition, Quest enables administrators to view DRS or DPM VM migration recommendations within the context of upcoming reserved VM deployments in order to make optimal performance-related decisions.

Availability Services



IT service availability is achieved through the Platform by with the following infrastructure automation capabilities – high availability (HA) resource pools, maintenance window management, and integration with VMware HA solution. The platform enables administrators to designate one or more HA pools. Within an HA pool, one or more servers can be reserved as fail-over capacity. If a potential server failure is detected or a failure occurs, the VMs are migrated or restarted on the designated reserved server, minimizing service downtime. The Quest Platform also provides a mechanism to designate one or more servers for maintenance. Servers in maintenance will not allow future deployments or reservations to be associated to them during their maintenance window. Administrators are provided a real-time and future impact analysis for servers requiring immediate maintenance with options to manage sessions, reservations, and migrate VM workloads in order to minimize downtime for users. In addition, Quest is fully compatible and interoperable with the VMware vCenter HA solution as another mechanism for providing business continuity.

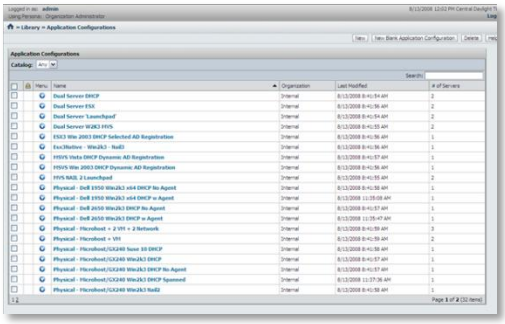
» Service Design

Engineers, architects, and administrators of IT infrastructure, services, and applications can design, create, and manage the lifecycle of IT service definitions, topologies, and configurations.

Service Definition

As described earlier, the Platform automates the provisioning of complex multi-tiered IT services. These IT services can support a wide range of server, network, and storage configurations. A single service can be heterogeneous – containing servers that are provisioned as

VMs on multiple hypervisors as well as on physical bare metal servers. The Quest Platform provides a user interface to assist engineers and architects to design, create, and manage service templates that are used as "recipes" for provisioning the service in real time. Service templates are composed of one or more server templates that can be auto-assembled, network and storage topology, hypervisor and physical server type, server boot order, server remote access methods, automation execution policies, and attributes that include resource capacity requirements and pointers to images, media, automation programs, and other "content."



Service Lifecycle Management

The Quest Platform provides the ability to model the stages of the IT service lifecycle from design, development, test, staging, and production through the configuration of organizational access controls and policies coupled with their associated resource pools. Quest provides management capabilities for service templates such creation, modification, duplication, save, promotion, and deletion in support of the IT service lifecycle. In addition, Quest provides tools to assist engineers and administrators in managing the configurations of servers and the optimization of server images to ensure security, availability, and performance.

Service Library Management

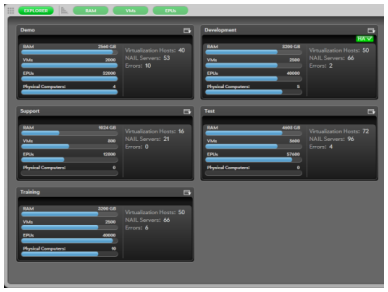
The Quest Platform provides a central library that can be physically distributed for storing and managing IT service configuration "content" in support of IT service delivery automation. Content can take the forms of virtual and physical image files, snapshot files, .ISO files, and automation programs such as software installation packages, patches, or customization scripts. The library provides a central place for architects, engineers, and administrators to manage the configuration components that comprise services. Capabilities include viewing, storing, dependency mapping to services, and clean-up of library content.

» Operations

Administrators of IT infrastructure can create, manage, and monitor users, policies, and shared pools of infrastructure resources – compute, network, and storage.

Administrative Dashboard

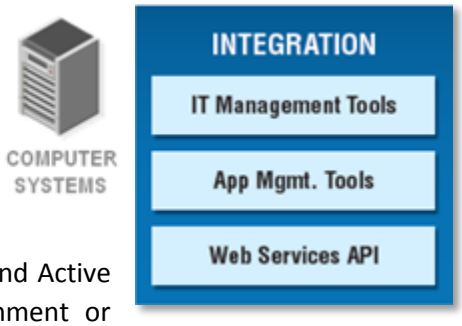
The Quest Platform provides an administrative dashboard for user and group management, infrastructure resource management, and resource capacity monitoring. The dashboard enables administrators to set up access control, resource allocation, and management policies through a two tier hierarchy – organizations and groups. Infrastructure resources – servers, network, storage, and software licenses – are pooled, managed, reallocated, or removed through the dashboard. In addition, the dashboard provides real-time infrastructure capacity monitoring views of the cloud, pools, and virtualization servers with search, drill down, and VM migration capabilities.



Event Broadcasting

The Quest Platform generates standard SNMP v3 events for common failures, error conditions, and warnings that occur within the operation of the system. Thresholds may be configured to generate additional SNMP v3 events based on performance characteristics of the cloud and underlying infrastructure. This provides standard data center monitoring tools the ability to monitor the operational health and performance of the Quest platform. The platform also provides a rich set of logs for auditing user access and troubleshooting problems that occur within a Quest-managed infrastructure.

» Integration



The Quest Platform provides a broad and rich web services API to support integration with 3rd-party systems, as well as productized integrations with industry leading management tools.

IT Management Tools

Quest provides productized integrations with directory services – LDAP and Active Directory to support Quest user authentication and automatic assignment or generation of machine names as part of the IT service provisioning process. Quest integrates well into standard ITIL v3 data centers through integrations with runbook automation, service desk, and service catalogs in support of standard ITSM processes.

Application Lifecycle Management Tools

Quest provides productized integrations with leading ALM tools – HP Quality Center, IBM-Rational Quality Manager, and IBM-Rational BuildForge. In addition, Quest has implemented numerous integrations with test automation, IDE, build and configuration management tools in support of the ALM process.

Web Services API

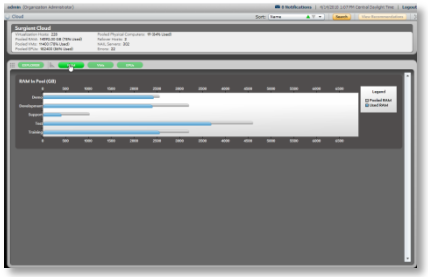
The Quest Platform provides a broad and rich web services API to support integrations with leading management tools. Web services APIs are provided for resource management, deployment management, scheduling automation, user management, and library management. In addition to web services APIs for integration, Quest provides a CLI for common commands, as well as an easily customizable service automation framework.

» Reporting

IT managers can run pre-defined reports, generate custom ad-hoc reports, and generate reports using their own business intelligence tools by leveraging the published Quest data dictionary.

Utilization Reporting

The Quest Platform provides pre-defined reports for resource capacity utilization, IT service deployments, account creation history, and user activity. Most reports provide daily, weekly, and monthly graphs for historical trend analysis. In addition, the platform provides a report generator that enables IT managers to run ad-hoc custom reports based on user, resource, and service data that is collected and stored in the reporting database.



Chargeback System

The Quest Platform collects detailed utilization information over time for user sessions, IT services, and pooled resources – CPU, RAM, VMs, network, storage, and software licenses. IT managers can combine this information with the published data dictionary and costing attributes to create custom billing and chargeback systems.

About Quest

Now more than ever, organizations need to work smart and improve efficiency. Quest Software creates and supports smart systems management products – helping our customers solve everyday IT challenges faster and easier. Our innovative solutions for application monitoring, database management, Windows management and virtualization management help customers achieve better results, faster.

Please visit **www.quest.com/cloudautomation** for more information.

8303 N Mopac Expwy | Suite C300 | Austin, TX 78759
info@surgient.com | 1–877–787-4436