Virtualizing IT Service Management
Table of Contents:

2 . . . . IT Service Management for Federated Cloud Infrastructure as a Service

2 . . . . Virtualizing IT

3. . . . . IT Service Management

6 . . . . Novell Cloud Manager

9. . . . . Conclusion
Business information technology (IT) is constantly evolving. In recent years, it has moved from the proprietary world of mainframes and minicomputers to the more distributed world of client/server computing, and now to a networked world of Web services based on open protocols and standards. Standardization, componentization and mass market commodity hardware and software infrastructure continue to underpin the evolution of IT. Compared to earlier market segmentation due to proprietary hardware architectures plus associated operating system and ISV software stacks, today’s commodity compute and storage servers, plus Ethernet-based network convergence, are unifying IT architectures.

This paper describes infrastructure software for this new IT landscape. It will discuss the service-oriented distributed management necessary to create an IT platform. This platform provides many of the traditional mainframe application hosting qualities of virtualized compute and storage, usage accounting, high availability and security for federated cloud infrastructure as a service (IaaS).

Virtualizing IT

Computers have CPUs, memory and disk(s) to hold data when the power is turned off. Virtual memory presents the illusion of more main memory than is physically available. Virtual disks create the illusion of a disk larger or more fault tolerant compared to the many physical disks they comprise. Virtual machines present the illusion of a whole computer that is actually contained by a real computer sharing its physical resources among competing virtual machines. Clusters present the illusion of a single reliable computer by coupling together and masking the failures of physical computers. Grids provide dynamic scalability, and clouds are an Internet-scale multi-computer made from grids of clusters of virtual machines.

Today, data center servers are connected to network storage. By removing and relocating storage from individual servers to a central network location, server form factors have shrunk. Compute servers are granted access to virtual disks (named storage containers) located inside specialized storage servers. When a server fails, processing fails over to some other server with access to shared storage. When a service (running on a server) runs out of storage, more space can be allocated using standard management APIs. When services themselves are virtualized, such as when they are hosted inside a virtual machine, they gain the flexibility to migrate from one physical server to another.

Virtualization eliminates physically imposed static boundaries: CPU, memory and disk are allocated dynamically. Services and data gain mobility—the freedom to optimally consume physical resources and the agility to rapidly switch to alternate physical resources while adapting to workload demands. High availability is a natural consequence of virtualized systems.

Legacy line-of-business applications are also being virtualized. Static, monolithic client-server software is being augmented or replaced with Web services. Web-based services replace earlier distributed object systems. There are new WS-* protocols for anything that wasn’t XML based before. And line-of-business applications now comprise a number of cooperating services. Infrastructure services provide naming and discovery, and via XML, a data integration and exchange.
Virtualizing IT Service Management

format. Line-of-business components execute in virtual machines and communicate using Web services protocols. Virtualization, services and WS-* protocols are creating a new platform for distributed computing.

Finally, with so many distributed moving parts, identity management creates the fabric necessary to securely name, associate, authenticate and authorize service consumers with providers, regardless of service type. Identity is the context that binds a flow of service requests all the way from the end user, through multiple processing tiers, to data on disks. Users are granted rights to services, and services are granted rights to other services. And if we haven’t experienced enough virtualization yet, identity itself has been virtualized by tokens that support assertions about role.

Back to the Future

An industry inflection point is occurring as a result of commodity server, storage and network hardware convergence—plus open source software, that, in isolation, is practically worthless. The value is due to “sophisticated integration of otherwise isolated components.” This integration is comprised of open source plus standards-based distributed systems software that transforms racks of network attached compute and storage servers into an enterprise IT platform.

IT Service Management

Delivering IT services requires successful coordination of numerous management disciplines and processes. IT must understand authorizations, roles and related security policies; ensure configuration and patch levels of multiple interacting resources; provide non-disruptive change and release management processes; track issues or problems with service continuity management; and scale by dynamically provisioning new service-oriented components, requiring licensing and approval. Manual processes violate service-level objectives; automation is critical. All of the traditional IT silos must work together to achieve the CIO’s goal of an automated service delivery infrastructure.

IT governance, risk and compliance (GRC) requires provable protection against security threats, and adherence to regulations. To assure compliance, many organizations are adopting IT best practices prescribed by the IT Infrastructure Library (ITIL)3 and Control Objectives for Information and Related Technology (COBIT)4 and requirements set forth in new standards (ISO 200005, the first international standard for IT service management). The goal of ITIL is to align IT to the business:

“IT needs to find clearer, more relevant metrics for showing business alignment and relevance, service quality supportive of that alignment, and cost efficiency in delivering its service products to the business. Moreover, IT is also being asked to support compliance and security initiatives that are adding extra relevance and context to this drive towards accountability.” —Enterprise Management Associates6

The following are considered fundamental foundations of deterministic IT services management:

Software (Delivered) as a Service

In the physical world, enterprise Linux* distributions are certified for multiple IHVs’ physical servers and multiple ISVs’ applications. This creates tension between the conflicting schedules of hardware and software partners; enterprise Linux distro vendors act as brokers that balance both groups of partners on behalf of mutual customers. One practical response has been validated configuration programs that comprise a select combination of IHV servers, OS and applications that are bundled together, certified and supported for a particular use case. Some vendors offer certified and supported application stacks. A popular format is the Amazon Machine Image (AMI).

“One of the overall design goals is to create a computing system which is capable of meeting all of the requirements of a large computer utility. Such systems must run continuously and reliably 7 days a week, 24 hours a day […] and must be capable of meeting wide service demands. Because the system must ultimately be comprehensive and able to adapt to unknown future requirements, its framework must be general, and capable of evolving over time.”

Corbato and Vyssotsky on Multics
Circa 19657

2  www.multicians.org/fjcc1.html
3  www.itil-officialsite.com/home/home.asp
4  www.isaca.org/cobit
5  www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41332
6  www.enterprisemanagement.com/research/asset.php?id=851
At the OS layer, some OSVs offer pattern deployed or role-based servers—software packaging that optimizes deployment of OS and services for a pre-determined customer use case. Another approach is to consider each validated configuration as a special-purpose software appliance.

We define a pDistro as a “just enough” Linux distribution intended for IHVs. Server applications are hosted by a dedicated virtual machine that executes on virtual hardware abstracted and managed by a hypervisor. Package selections available when installing a pDistro need only consider the kernel, hypervisor, appropriate drivers and management agents necessary to support an IHV’s brand (or class) of server.

A vDistro on the other hand, is a “just enough” Linux distribution intended to offer ISVs an enterprise application runtime environment. Apart from processor architecture, vDistros are hypervisor and IHV-neutral; there is no longer a dependency between pDistro physical drivers versus applications hosted by the vDistro’s OS kernel. This allows for independent upgrades of pDistro kernels, physical device drivers, vDistro kernels (security patches) and applications.

**IT Workloads**

IT workloads are physical or virtual machines installed with a just enough operating system (e.g., Linux or Windows*) plus applications and/or services that were selected during automated installation. Virtual machines execute on physical hardware managed by a pDistro. SUSE® Studio™ generates p and vDistros—enterprise software appliances that are manufactured and distributed with service-level declarations due to DMTF’s Open Virtualization Format (OVF) standard for distribution and deployment.8

**Workload Portability**

Workload portability is the ability to relocate a workload from one environment to another. Portability results from standards like OVF and tools to convert physical to virtual machines via automated image manipulation and build recipes. Offline image manipulation provides the method for injecting just enough management, lightweight agents, monitor probes and other instrumentation necessary to securely manage workloads. Virtual to virtual machine conversion enables workload portability across heterogeneous environments.

**Workload Profiling**

Profiling is required for initial consolidation planning; afterward, profiling provides the business intelligence necessary to continuously optimize workload execution.

**Service Declaration**

The execution requirements and qualitative service levels of workloads are also based on OVF standards, plus references to a data center scale service model that provides the context necessary to automate workload provisioning and optimization. Service declarations are part of the metadata necessary to publish a workload based service catalog.

**Discovery**

The corollary of declaration is discovery; it is required to correlate actual versus declared (or expected) infrastructure state. Drift is the result of differing these two models. Discovery is necessary also for populating an analytics model in support of various use cases including workload consolidation planning and risk assessment (performance, availability or compliance).

**Data Center Automation**

Gartner describes two kinds of IT automation: passive and active.9 Passive automation is applied to IT administration processes such as requesting, approving and applying a configuration change. Active automation is what operates the (real-time) data center infrastructure, such as failing over virtual machine hosted server applications in response to...
server failure, or live migrating virtual machines under changing load or other environmental conditions. In comparison with ITIL prescribed best practices, change management is typically passive in nature (until released) while service-level management is generally proactive.

**Process Workflow**

Workflow is a stepwise processing pattern, and workflow engines provide an execution environment for automating different kinds of process. Processes are defined using modeling tools, and executed by a workflow engine. Workflow processes invoke a sequence of well-defined activities.

Business process management was the first domain to adopt workflow technology for coordinating people-oriented business processes. The IT Infrastructure Library (ITIL) prescribes best practices for IT administration processes that are also coordinated by people with roles: responsible, accountable, consulted and informed.

A runbook is documentation of an IT administration process. Runbook vendors use workflow and related business process management technologies to create IT process automation products. Runbook products capture IT knowledge gained by integration with a variety of IT infrastructure, by offering reusable libraries of workflows and activities.

Our archetypal use case is “patch virtual machine.” This is an approval-driven sequence of IT process automation steps required to apply change to an offline image of a virtual machine, and then release to production while ensuring minimal service-level disruption.

A key distinction in the popular use of the word orchestration is with regard to what is being orchestrated: IT administration processes or server workloads (running in virtual machines). The virtualization market generally uses the word “orchestration” to describe workflow-based automation of IT administration processes.

**Workload Orchestration**

Orchestration here means the automation of workload placement across many physical or virtual computers based on resource matching policy. Automation is usually specific to the resource type. For example, orchestration of virtual machine deployment means identifying a virtual machine host that satisfies the requirements of the virtual machine (constraints matched with resource capabilities). It also means executing a sequence of jobs as needed to position an OS image, and define and run the virtual machine on the chosen host. Policies, constraints and capabilities are expressed in XML based formats.

Orchestration of IT processes, when integrated with workload placement, provides the framework for executing ITIL best practices. Driven by service declarations, resources are provisioned, related to one another and managed as whole business services. Monitoring performance and availability relative to declared service levels enables closed-loop optimization.

Consider this workflow: it constructs a workload, and then releases it to production:

```plaintext
workload = new Workflow::CreateWorkload() {
  vm = VMMgr->CloneVM(qos)
  disk = StorageMgr->ProvisionDisk(qos)
  StorageMgr->ConnectDiskToVM(vm, disk, qos)
  app = AppMgr->DefineApp(packages)
  AppMgr->InstallApp(vm, app)
  DefinitiveSoftwareLibrary->CheckinVM(vm)
}
if (workload.ApproveRelease(admin) == TRUE) {
  DefinitiveSoftwareLibrary->Release(vm)
  Orchestrate->PlaceWorkload(workload)
}
```

**Federated CMDB**

A federated configuration management database (CMDB) is a data center scale repository of aggregated resource management configuration data. The CMDB collects
and records information in the context of the relationships that exist between multiple managed resources and the business services they compose. Managed resources are either compute servers, virtual machines, storage servers or networking devices. The CMDB’s meta-schema provides:

- Dynamic domain specific structure
- Instance relationships as types
- Extensible resource class and attributes
- Version control via consistent snapshots

The federated CMDB is the arbiter of change management; it uses a model of current configuration and dependencies that provides the context for comparison with future change requests and potential impact on business services. The CMDB can also function as the authority for how systems should be configured. This enables detection of unauthorized or unapproved change that might constitute a security or compliance breach when systems are actually configured differently.

Change Management
Change management is a request of desired future state, enacted via orchestration of resource reconfiguration. Before change can be activated though, its impact must first be analyzed, reviewed and approved. Querying the CMDB’s model of relationships is therefore a required activity within the change management process.

Policy
Policy provides control:

- Relative to time—frequency of drift
- Relative to events—continuous remediation

Change and configuration management is periodic:

- “Make it this way, next Sunday”

Governance, risk management and compliance is real time:

- “Ensure it is this way, all the time”

Version Control
Every change request is recorded and audited to enable version control:

- Roll back to any prior snapshot state
- Roll forward to any declared future state

Federated Identity Management
Every resource (modeled object and relationship) has an identity token. All management operations are executed with, and governed by context that federates resource, workload, process and administrative identities. Explicitly recognizing “customer” as an identity enables multi-tenancy. For example, virtual machine X was started on virtual machine host Y, as initiated by process (workflow) Z. Process Z was executed by admin A on behalf of customer B. Identity enables a correlation between events and processes. The governing CMDB model is the foundation for security, cost accounting and change commit use cases.

Novell® Cloud Manager
Novell Cloud Manager is a new product architected around many of the previous principles. It provides management of business services that comprise multiple cooperating workloads, with declared service levels, that are deployed across federated private and (soon) public virtual machine based hosting infrastructure. Novell Cloud Manager enables multi-tenant self-service ordering of IT workloads and is built on an enterprise service bus hosted business process execution language (BPEL) workflow engine for approval based IT service management:
A number of existing Novell products function as service components and are integrated by Novell Cloud Manager. They are deployed as virtual machine based software appliances and thereby gain high availability from the underlying virtualization technology.

**PlateSpin Orchestrate**

PlateSpin Orchestrate is a grid-based distributed resource automation system for physical and virtual machines. It supports full virtual machine lifecycle management across networks of physical hosts for virtual machines. It completely automates the process of creating and managing virtual machines. It provides distributed storage repository management for virtual machine images and data, as well as constraint-based adaptive deployment of virtual machines to capability-matched physical servers.

PlateSpin Orchestrate schedules work to managed servers in the form of scripted jobs. Jobs are units of work that are assigned to servers by a real-time resource scheduler that continuously evaluates available resources and pending requests. Physical servers, like virtual machines, are considered resources that advertise their capabilities in the form of facts that describe the type and capacity of resource. For a physical server, example facts might include number and type of CPUs, memory and direct-attached storage capacity. Static facts are attributes of a resource that don’t change; an example might be a server with multi-core CPUs. Dynamic facts can change over time, perhaps due to physical hardware hot plug or memory ballooning of a virtual machine. Computed facts are calculated by the scheduler when referenced in a control policy. As an example, consider deploying a virtual machine into your data center production server pool. The virtual machine requires two CPUs, 512 MB of direct attached OS image storage, 1 GB of memory and 10 Gigabit Ethernet connectivity for access to a converged Fibre Channel SAN. These requirements are expressed as a set of deployment constraints, which are references to facts that are matched to available resources by Orchestrate when scheduling the virtual machine for deployment.
Sophisticated resource allocation is made possible when combining a number of constraints into policy statements that are applied to groups of resources, thus matching supply with demand. The deployment of a virtual machine to a physical server is therefore unified by a general purpose (grid-based) algorithm for assigning units of work to available resources. This is done in a manner that’s respectful of competing work and shared capacity.

**Novell Business Service Management**

Business Service Management solutions from Novell create a federated, real-time integration of a multitude of data sources, through a series of (often) bidirectional integration modules to underlying infrastructure performance and availability management systems. Performance data is also collected synthetically measuring the user experience. The IT service dashboard is a complete view of service performance, inclusive of business transactions.

After data is collected, it is normalized in the CMS, where automation is applied to manage multi-dimensional service levels for performance and availability. Thresholds are defined to manage to avoid service-impacting incidents; however, the service model views also provide speed to pinpointing root cause during an incident.

Critical to business service management is the ability to model complex state dependency rules that represent the physical and logical relationships between end-users, applications and the supporting IT infrastructure. Business Service Management solutions provide a pluggable state propagation and calculation engine to support the creation of a real model of the data center. Administrators can customize the state engine to produce reports of the overall performance of business services.

**Novell Sentinel™**

Novell Sentinel is a security information and event management solution that receives information from many sources throughout an enterprise. It then standardizes that information and prioritizes it for the purpose of making threat, risk and policy-related decisions.

Novell Sentinel automates log collection, analysis and reporting processes to ensure that IT controls are effectively supporting threat detection and audit requirements. Novell Sentinel replaces these labor-intensive manual processes with automated, continuous monitoring of security and compliance events and IT controls that are necessary to meet specific regulatory requirements, such as the Payment Card Industry Data Security Standard (PCI DSS).

Novell Sentinel gathers and correlates security information from networked infrastructure, third-party systems, devices and applications. It visualizes the correlated data to identify security or compliance issues, and tracks remediation workflows to streamline error-prone security management processes.

Automated incident response management enables formal documentation and tracking, escalation and response to incidents and policy violations, and provides two-way integration with trouble-ticketing systems.
PlateSpin Recon

PlateSpin Recon is a data center workload profiling, analysis and planning tool that combines consolidation planning with capacity management. PlateSpin Recon correlates workload inventory and utilization statistics to analyze infrastructure and applications running in the data center, including how and when resources are being used. PlateSpin Recon enables complex server consolidation and capacity planning initiatives and provides ongoing management of the resources and costs of an organization’s physical and virtual infrastructure.

Conclusion

The manner in which information technology provides solutions to business is radically changing: Virtualization frees applications from the static data center silos that restrict flexibility, but it also necessitates new approaches to retain enterprise control and compliance. This paper described a number of foundations and solutions for hosting business services as portable workloads via IT service management for federated cloud infrastructure.