Dynamic Workload Utility: A Key Component in Grid and Enterprise Cloud IT Infrastructure

Integrating Moab with HP iLO, HP SA, and Applications

Dynamic Workload Utility: A Key Component in Grid and Enterprise Cloud IT Infrastructure

1 Introduction

The Dynamic Workload Utility (DWU) is an automated scale-out infrastructure for Enterprise Clouds and Grids that is designed to:

- Optimize utilization of compute resources to 90% or better.
- Guarantee and enforce SLAs.
- Manage a mix of workloads including service and batch jobs in a shared environment.
- Dynamically respond to resource failures and workload surges.

These capabilities accelerate delivery of new services and products, decrease new hardware purchases, and reduce energy costs and administrative overhead.

Within a Cloud environment, the DWU is a joint solution between HP and Cluster Resources[1] that:

- Performs SLA-driven dynamic allocation of resources to workloads.
- Flexes resources based on dynamic workload changes.
- Manages workload scheduling and load-balancing of batch jobs and service workloads on physical and virtual servers in the same environment.
- Integrates with provisioning, monitoring, event management, virtual machine (VM) management, billing and reporting tools.

The DWU optimizes resource allocation based upon workload and SLAs in a very dynamic, ever-changing environment. The DWU can manage servers, networks, and storage—whether physical or virtual. The DWU goal is to make the environment as self-managed as possible.

DWU includes Cluster Resources' intelligent meta-scheduler and workload management suite Moab Utility/Hosting Suite[1], HP management tools, for example, HP Server Automation (SA)[2], HP Operation Orchestration[2], and a number of best-in-breed open source packages. The solution is SOA-based, allowing the components to be integrated using standards-based interfaces. The solution, therefore, enables easy integration of home-grown or third-party tools. Another capability of the DWU is that it supports non-HP hardware as well as a variety of operating systems.

Customers are looking for solutions that can help them turn their data centers into private Clouds or automated Grids. Gartner describes solutions such as the DWU as IT Workload Automation Broker-based solutions[3] and real-time infrastructures[4], which Gartner claims will be invaluable to the Enterprise to fulfill a requirement to manage mixed workloads across a heterogeneous computing landscape. This type of solution is not only capable of end-to-end automation with minimal human intervention, but this resource-aware automation and workload management is driven by business policies[3].

2 The DWU solution

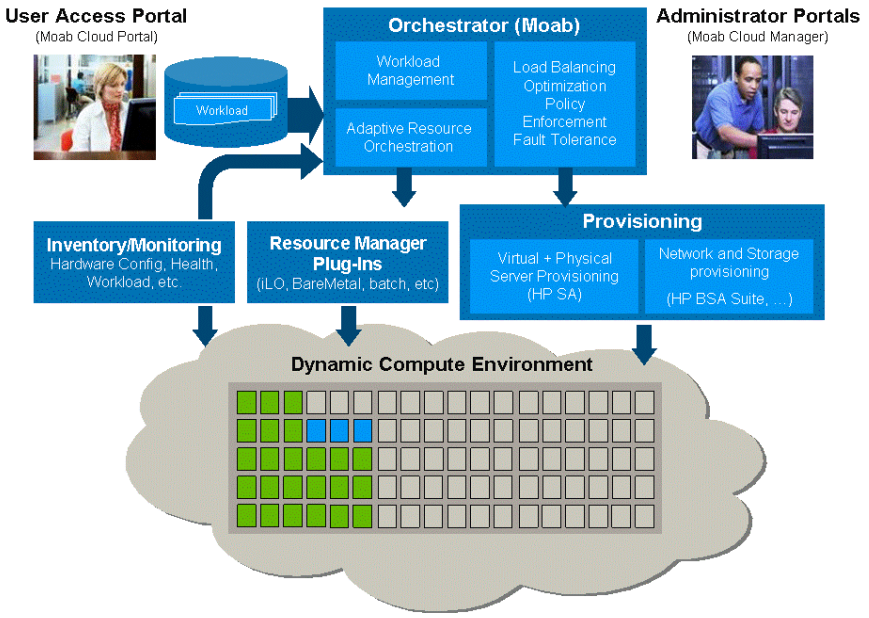
HP created the DWU to respond to customers' need to manage dynamic environments such as private Clouds and Grids where automatic adaptation of the resources to the workload is required in a changing cross-enterprise IT environment. These customers want to optimize utilization of compute resources by using an automated environment to manage a mix of workloads (batch, real time, online, service jobs, and more) on heterogeneous and geographically distributed servers. Service jobs are workloads that are running and acquire or release capacity in response to external triggers such as user demand or scheduled activities.

The DWU extends the usual NGDC/AI concept by having a scheduler or workload manager drive the automation and act as the orchestrator or Service Governor[5] between the SLAs, the resource capacity, and other components within the environment. This section describes the solution as well as the components which are part of the Utility.

The DWU is a policy-based automated solution, which manages the workloads and resources (storage, network, compute resources, and more) in a heterogeneous environment, thereby ensuring that the SLA objectives for users, projects, and applications, as well as for availability and performance are met. The DWU schedules and dynamically assigns resources to workloads while monitoring the environment and triggering actions to be executed when key thresholds are crossed, scheduled times are reached, or specified events are detected. This enables compute resources, both virtual and physical, to be customized on-the-fly based on job requirements rather than having statically provisioned resource pools. With the DWU, the automation is done by creating and managing rules which define thresholds and trigger events. Each workload is treated differently so that the environment only triggers the allowed actions for the specific workload (migration versus no-migration, virtual versus physical server, minimum and maximum number of servers available for a given workload, flexing versus no-flexing, and more.)

The Moab, a meta-scheduler and workload manager, is the orchestrator in the DWU as illustrated in Figure 1. Moab makes intelligent decisions about what needs to be done in the environment to enforce and optimize existing and future SLAs. It acts as the glue between the various components in the Cloud infrastructure such as monitoring and provisioning. The user portal (MCP) gives users the ability to manage jobs and reservations from a single portal. The administration portal (MCM) enables administrators to manage users, workloads, reservations, event triggers, and more. from a single "dashboard" as well as to set policies, rules, and priorities for the environment. Reporting and billing are available through the MCM portal.

Figure 1. Dynamic Workload Utility Architecture



Moab has the capabilities to fulfill the role of orchestrator. It enables resources in the resource pool to grow and shrink based upon factors such as workload spikes, backlog, QoS (quality of service) guarantees, performance metrics, resources failures, and more. Moab ensures that no one exceeds intended usage with soft and hard limits for users, groups, projects and organizations. Moab is built on an advanced reservation infrastructure that enables it to track and optimally schedule both current and future workload. With this infrastructure, Moab can plan for future application requirements, anticipated one-time and periodic surges, and usage workload learning to utilize historical information to better plan future allocation decisions. Using the advanced reservation facility, Moab

can identify the impact of resource failures or the addition of future workload before they become issues and either notify administrators immediately or automatically begin taking steps to mitigate or eliminate the issues.

In its Enterprise Cloud version, the DWU integrates with HP provisioning tool HP SA to support a variety of operating systems and vendor platforms. Depending on the use case, the DWU can also take advantage of the IT process automation tool HP Operation Orchestration. The DWU includes support for IPMI for basic system management functionality such as start/stop in a multi-vendor environment. Additionally, in an HP environment, the DWU integrates with HP iLO, enabling intelligent scheduling based on system temperature, power consumption, and other available metrics, as well as provide remote system management capabilities.

Capabilities

The DWU capabilities are summarized below in the following six areas: general, workloads, provisioning, capacity management, monitoring, and reporting and billing.

General

The DWU is vendor neutral and supports a variety of operating systems and hypervisors. The DWU solution enables integration with homegrown, open source, third-party and additional HP tools and solutions. The solution scales to tens of thousands of servers.

Workloads

Customers have a mix of workloads, such as service, batch, online, real-time, and workflow jobs, running across heterogeneous computing environments.

In the DWU, jobs are automatically given attributes when launched, which inform the orchestrator about a preferred and required allocation. These attributes are also satisfied when adding resources to an existing service job. Examples of attributes are: CPU architecture, virtual or physical server, specific application or hardware accelerator required, license availability, and more.

Each workload can be configured to meet its specific requirements so that the environment only triggers the allowed actions for the specific workload (migration versus no-migration, virtual versus physical server, minimum and maximum number of servers available for a given workload, flexing versus no-flexing, and so on). For example, real-time jobs must only run on physical servers and within a dedicated locked-in low-latency environment, and service jobs, which are allowed to grow or shrink, must be guaranteed a minimum set of resources so they do not disappear.

In the DWU default configuration, Moab uses Torque, an open-source resource manager, to launch batch jobs as well as to capture the requirements put on these batch jobs. Moab manages the total workload across the servers which include both batch and service jobs. A scheduler needs to drive this environment so that it receives information of past, existing, and future workloads and can compare it with existing and future resources availability. Moab can orchestrate the provisioning of any generic resources object: compute, network, storage, security, services. HP is currently investigating the addition of network and storage virtualization to the DWU with the already existing server virtualization and provisioning to provide a complete solution.

The DWU supports sharing of the resources by a mix of workloads and by many different applications. The DWU orchestrates mixed workloads in an environment composed of dedicated application servers and pools of physical and virtual servers.

Provisioning

The DWU solution enables the provisioning of bare-metal servers as well as virtual machines through its integration with HP SA. The DWU can also create VMs and provision both OS and application stacks.

Some of the DWU basic provisioning capabilities include:

- Deploying to multiple target systems at the same time.
- Updating inventory and discovery of resources, to install applications and service packs as required along with the ability to make configuration changes across multiple systems.
- Installing patches and updates as necessary; to deploy a target operating system with the hardware specific drivers and configure the server e.g. Network, SAN, NAS, and more.
- Updating systems after deployment, for example, configurations (ip, ports, and more), mounts external storage, and changes to application executables and third-party applications.

The DWU enables system administrators to do the above in their choice of an automated or manual procedure. HP SA also enables audit node configurations for compliance against organizational standards and best practices. This includes monitoring and reporting of configuration drift of provisioned systems. The audit can incorporate out of the box checks for governance standards such as SOX, HIPPA, PCI, or CIS.

Additionally, the DWU solution integrates with ticketing and change systems (via Operation Orchestrator[2]) to make sure any change done within the Cloud follows best practices Change and Configuration management processes. The ability to have changes for the Cloud automatically routed through Change Management systems/CABs for approval before implementation has great value in minimizing risk and for overall enterprise security and stability.

Capacity management

The DWU enforces business-driven SLAs by managing the resource capacity, job priorities, and other components within the environment. The DWU enables service jobs to add or remove capacity. Adding capacity means to be able to add or provision new servers based on automated procedures and automatically have them join the Cloud or Grid. Remove capacity means to take away unused and underutilized servers to allow repurposing of hardware. Adding and removing capacity have to be carried out while managing resource availability to support high transaction throughput and maximum server utilization. VM placement across multiple hosts and data centers, as well as live migration and stateless VM management is also supported.

Along with managing resources such as sockets, storage, and licenses for workloads, the DWU can be extended to become network-aware by integrating with the Voltaire Unified Fabric Manager[6], thereby enabling optimization of SLAs based on dynamic management of network congestion. HP is currently investigating adding this feature to the DWU since it enables the scheduling of fabric resources as well.

Monitoring

The automated DWU environment is governed using policy management for defining rules (thresholds) and integrated with an event management component that is based on alerts and predefined policies or procedures. The DWU solution provides a remotely accessible management console and a dashboard to monitor resource utilization.

For monitoring hardware, the DWU integrates out-of-the-box with Supermon[7], Ganglia[8], Nagios[9], and other common monitoring packages. HP is investigating the integration with OpenView Operation Manager[10] and HP SiteScope[10] for application monitoring, enabling the DWU to manage resources to support high transaction throughput.

Reporting and billing

The MCM dashboard enables administrators to:

- Identify resource consumption at the group or individual server level.
- Get reports on the status of individual servers within the grid close to real time.
- Record and display availability and activity levels at various levels.
- Choose the granularity of metrics on each server and for the whole grid.

The DWU also offers a wide variety of reporting and charge-back capabilities out-of-the-box and has been integrated with accounting packages such as Gold (open source).

Integrating Moab with HP iLO, HP SA, Applications

HP has integrated Moab with the iLO (integrated Lights Out processor) on HP blades and servers, enabling Moab to gather system temperature, power consumption, and other metrics. This gives Moab the opportunity to do more intelligent scheduling. For example, when a system gets warmer than the set threshold, Moab stops scheduling jobs on that server. This integration also enables the enhancement of the scheduling policies to include scheduling jobs on the coldest blades and more.

The DWU provisions a bare metal node through the integration of HP iLO, HP SA, and Moab. Moab orchestrates the process through two Moab native interfaces, constructed using the iLO and HP SA Web services, RIBCL and OSAPI, respectively. The iLO native interface provides power management and the SA native interface provides the current status of all nodes under SA's control. To provision a bare metal node, Moab powers on a node through iLO. SA then oversees the node provisioning, reporting the provisioning status to Moab. When provisioning and software installation completes, SA informs Moab the node is ready to be federated into the active pool where it begins receiving workload. This approach provides Moab with continuous provisioning status information.

The DWU comes out-of-the-box with many plug-ins designed to integrate the DWU with common service applications such as Apache, JBoss, and common management tools. These scripts present job service information to Moab, enabling Moab to handle all jobs in the same abstracted manner. Therefore, batch and service jobs coexist on the same system, managed by the same policy set and scheduler. The DWU flexible architecture makes it straightforward and easy to build scripts for homegrown or proprietary applications using the Moab native interface to control these applications.

Use cases examples

This section describes four scenarios in which the DWU is central to making the infrastructure dynamic and adaptive. Scenario 1 describes the basic business continuity use case and illustrates how the DWU enables better utilization of a existing resources. Scenarios 2 and 3 describe business continuity and capacity on demand. The DWU responds to unplanned and planned surges by managing the workloads and resources in the internal Cloud. The DWU also enables workloads to spill-over into an external Cloud such as Amazon EC2 or HP Flexible Computing Service. Scenario 4 describes how the DWU is a critical component in creating Cloud solutions such as Infrastructure as a Service solution.

Scenario 1: Business continuity

The DWU manages workloads and resources within a datacenter allowing for on-demand provisioning or customization of compute resources to fit the workloads. Using the DWU, companies avoid having unused statically preconfigured computing resources sitting idle. The DWU also packages existing workloads to get better utilization and turn off unused servers to save power. The DWU takes advantage of its advanced placement engine to place workloads and triggers live migration of virtual machines if required by the SLAs.

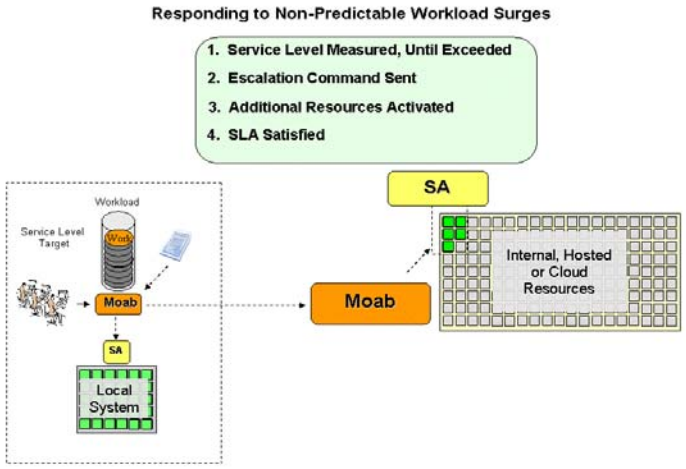
Scenario 2: SLA driven resource escalation

The DWU can respond to unplanned workloads surges by using available resources within the local data center as well as by accessing external Cloud resources. The SLA driven resource escalation is illustrated in Figure 2.

In this example the load on service Job A increases due to some unexpected external user activities, for example, a higher-than-expected result from a marketing campaign that begins to overload the web farm and ordering system.

When Moab detects that Job A exceeds its allowable service-level metrics, for example, CPU utilization, it reacts to the SLA violation. Based on the existing rules and policies, Moab performs a series of customizable escalation steps. In this example, Moab first searches for available stand-by resources, which can be brought into the active resource pool and enable Job A to meet its SLA. If there are not enough resources, Moab can preempt lower priority jobs. The preempted jobs are then requeued and relaunched when resources are available. If Moab has preempted all the jobs that were preemptable, used all the standby resources, and still needs more resources to satisfy the Job A SLA, it starts looking for bare-metal resources and externally accessible resources. In all cases, Moab can trigger HP SA to provision the right stack on the local or remote servers if it does not find resources with the correct stack. By going through these steps, Moab ensures that Job A as well as the rest of the jobs in the environment meet their respective SLAs. The additional hosted or Cloud resources utilized during the workload surges are released once they are no longer needed. The various steps in the Moab decision tree are customizable to fit the individual application need.

Figure 2. SLA driven resource escalation

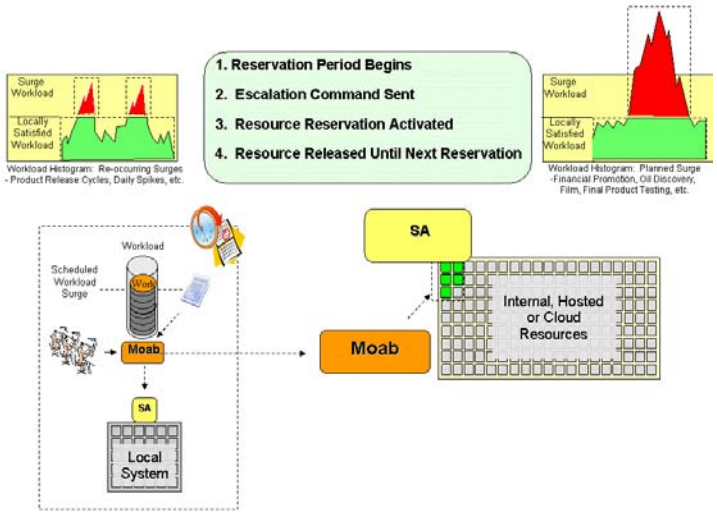


Scenario 3: Scheduled surges

The DWU can respond to planned workloads surges by using available resources within the local data center as well as by accessing external Cloud resources. The Scheduled surge scenario is illustrated in Figure 3.

In this example, Company A knows that during Christmas shopping the workload for Job A increases 100%. The administrator sets up this planned surge in the Moab advanced reservation system and thereby assures that the number of required systems for Job A to meet its SLA will be available. Moab then puts a reservation on a specified number of systems for a given time at a specified date and allocates them to Job A so it can fulfill its SLA. In the case where the Company A data center is already fully utilized, it can take advantage of remote underutilized/standby systems or of rented systems for that period of time as illustrated in Figure 3. The new systems are provisioned automatically by having Moab trigger HP SA. Moab thereby ensures that SLAs are met for Job A and for the rest of the jobs in the data center. The extra resources utilized by Job A are released when the scheduled surge is over.

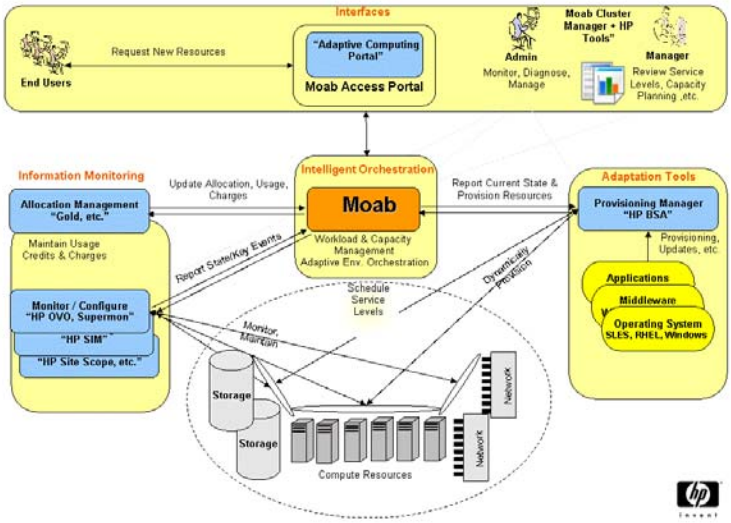
Figure 3. Scheduled surges



Scenario 4: Infrastructure as a service

The DWU can be a critical component in creating an Infrastructure as a Service solution as illustrated in Figure 4.

Figure 4. Infrastructure as a service



The DWU goal is to create project or test environments on demand. This creates custom, disposable infrastructures which only are available for a given period of time specified in the user request, and which can be recreated on-demand from templates stored in a database. The templates are created by IT Administrators and include hardware components and software and applications stacks. The templates are then selected by users to create a basic infrastructure for their jobs. The templates can be customized to fit the user's target environment.

In this environment Moab decides what resources are going to be assigned to a given user for the requested time. The DWU environment handles both physical and virtual servers and, though HP OO[2] and HP SA, automates workflows, and provisions with Linux, Windows, and Solaris. Users log in through a Secure Portal and, based on their credentials, are granted different views. The DWU-based Infrastructure as a Service environment integrates with accounting, monitoring, reporting, and charge-back.

3 Conclusion

The Dynamic Workload Utility is a key component in turning data centers into private Clouds or automated Grids. It is a unique solution that enables the Enterprise customer to use business policies to drive resource-aware automation and workload management with minimal human intervention, and to manage mixed workloads across a heterogeneous computing landscape. The DWU is currently being packaged as an SCI solution.

For more information

Send questions regarding the HP Dynamic Workload Utility to hpcdgrid@hp.com.

Send questions regarding installing Moab on HP Clusters to hpcdgrid@hp.com, or see the section about Grid and Adaptive Infrastructure on the HP website: http://www.hp.com/go/collaboration

---

1 Cluster Resources, Moab, http://www.clusterresources.com/solutions/adaptive-computing.php

2 HP BSA, http://www.hp.com/go/bsa

3 Gartner, IT Workload Automation Broker: Job Scheduler 2.0, ID G00157440, 28 April 2008

4 Gartner, Modernization and the Real-Time Infrastructure, ID G00167604, December 2008

5 Gartner, Hype Cycle for Real-Time Infrastructure, ID G00158969, 18 July 2008

6 http://www.voltaire.com/Products/Unified_Fabric_Manager

7 Supermon, http://supermon.sourceforge.net/

8 Ganglia, http://ganglia.info/

9 Nagios, http://www.nagios.org/

10 HP BTO, http://www.hp.com/go/bto

4AA2-6552ENW, May 2009

CLUSTER RESOURCES

hp invent