SOLUTION BRIEF

The Seven Keys to Successfully Managing Compute Workloads at Financial Institutions

Adaptive COMPUTING

Executive Summary

It is no small fact that compute workloads at financial organizations are substantially increasing. This is driven by the explosive growth in the amount of data, an increase in the level of analysis required on that data and new data centric risk/stress test requirements. Faced with this increase, organizations have the choice of either dramatically increasing their compute hardware spend or finding ways to improve efficiency and throughput.

Exacerbating these challenges further, a growing number of financial organizations face increasing fees for workload management solutions from IBM; while also needing to move away from IBM x86 hardware (which has been sold off to Lenovo).

In this solution brief, Adaptive Computing will lay out Seven Keys to success for financial institutions to meet these challenges: Consolidation, Convergence, Automation, Optimization, SLA Delivery, Agility and Modularity. It will articulate the role Adaptive's Moab Cloud for the HPC Suite and Nitro workload scheduling software can play in solving the Seven Keys to success while showing a low risk migration path from competitive products.

Adaptive Computing has a vision to accelerate customer's delivery of products and services by unifying and optimizing their data processing into end-to-end workflows.

Challenges Driven by Market and Industry Trends

Data Growth

In financial markets, the sheer amount of data that needs to be processed for quantitative, risk, and actuarial analysis is generally growing at least 2x faster than processor performance improvements. According to the Wall Street Journal, in 2014 "Data Analytics" became the number one priority for CIOs. Data processing is certainly important, but it is also increasingly difficult and costly to match.

Data Center Complexity

Compounding on top of this trend is the fact that organizations have expanded their compute resources reactively to meet the needs of individual departments and projects over time. This has caused resources to be fractured into siloed environments where there is no sharing of excess resources, a fact that can drive utilization down as low as 20 percent efficiency, and which requires many time consuming manual processes such as data migration and transformation steps to bring the disjointed process back together. Large portions of time, resources, and money are thus wasted.

Further, analysis and simulation work is demanding increasingly exhaustive examination of data in order to maintain an edge in today's financial and insurance markets. Holding the advantage requires the ability to analyze multiple sources of data and use multiple applications in a complex analytical workflow.

Politics further complicate the picture, bringing inconsistency in processes, ownership disputes, and complicated approval battles.

These attributes are just a few examples of the costs and inefficiencies that are introduced by data center complexity.

New Financial Market Pressures

The pressures of new governmental regulations and the competitive need to achieve greater operational efficiency add to the challenge. Due to the recent financial crisis and the ongoing trend of increased government oversight, the financial industry is being laden with new regulations and best practices recommendations that can significantly increase the amount of simulation and analysis work an organization is required to complete. Every region has these requirements, whether they are from the Dodd-Frank legislation in the US, Prudential Regulatory Authority in the UK, the European Banking Authority requirements across Europe, or the International Monetary Fund for any region in the world.

These regulations and recommendations are not a matter of calculating a simple ratio, but rather involve detailed analysis and simulation. One Top 5 US investment bank estimated that the current regulations, if not altered, could require up to four times more computing work than the institution currently processes.

The Seven Keys to Success

While the above trends and pressures bring challenges to financial and insurance organizations, there are seven keys that are fundamental in transforming chaos into order which can aid in setting any financial institution on the path to success.

1. Consolidation – Break down the siloed and fractured environments through consolidating resources and enabling an environment that is sharable between users and groups, creating a foundation for utilization efficiency.

2. Convergence – Facilitate the convergence of workloads from diverse disjointed disciplines (High Performance Computing, High Throughput Computing/Grid, Data Center, Big Data, etc.) into efficient end-to-end process.

3. Automation – Automate data migration, processing steps, re-purposing of resources, validation steps etc. to accelerate and reduce the costs of these end-to-end processes.

4. Optimization – Dynamically optimize the workload processing on these environments in order to reduce costs and improve speeds of producing end results.

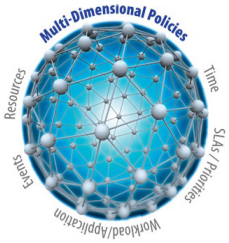
5. SLA Delivery – Enable service commitments to be met and business alignment to occur using policy-driven controls so that political barriers can be overcome.

6. Agility – Ensure the infrastructure is designed with agility in mind. It should be capable of handling changes quickly, being able to adjust to new priorities or projects, promptly respond to failures and be able to swiftly integrate with and take advantage of new technologies in a cost efficient manner.

7. Modularity – Implement modular technologies that are additive, not monolithic, in nature, and ensure that it is with technologies that are integration friendly and open.

About Adaptive Computing

Accelerate results with intelligent and predictive workload management



According to International Data Corporation (IDC), Adaptive Computing is the leading provider of HPC job/workload management software. Adaptive Computing's software runs on some of the largest systems in the world (systems with over 10,000 servers) as well as on systems of Fortune 500 companies such as ExxonMobil, Dow, Honeywell, and Boeing. In Finance, Adaptive Computing has its software running at a Top-5 US banking institution on thousands of nodes, where it has saved the bank hundreds of millions of dollars. Also among Adaptive Computing's customers are a Top-5 financial institution in Australia, a global European market maker, a global quantitative trading company, and others.

Adaptive Computing has achieved this leadership position through being a pioneer and innovator in scheduling, optimization, cloud, and other areas with over 70 patents issued or pending, and through global partnerships with HP, IBM, SGI, Intel, and Cray.

Vision -- Big Workflow

Adaptive Computing's vision is to accelerate our customer's delivery of products and services by unifying and optimizing their data processing into end-to-end workflows.

Products

There are two principle products from Adaptive Computing that meet these core needs in the financial and insurance industries.

Moab Cloud -- A traditional capability scheduler and policy engine which acts as the "brain" of a cluster or grid, ensuring that usage matches organizational objectives. Moab Cloud is able to connect to multiple monitoring and management tools to create a unified world view of the environment. It is then able to apply policy and scheduling to govern workload optimization, balancing and resource usage.

- Includes the core policy/scheduling engine, accounting/charge-back, workflow dependencies and workflow triggers, grid load balancing, a job/task submission portal and API, and an administration portal.

Nitro -- A high throughput execution engine which can launch millions of small single core jobs up to a rate of hundreds of tasks per second per core.

- Submits workloads up to 100x faster than traditional schedulers
- 40-60 percent less expensive than the leading alternative
- Works as a stand-alone product or in conjunction with a traditional capability scheduler like Moab Cloud or LSF.

Delivering on the Seven Keys to Success

1. Consolidation

- **Heterogeneous Resources** – Moab Cloud enables heterogeneous resources to be combined into a larger shared system and applies intelligent matching policies to place workloads where they will perform the best (based on speed, memory, GPUs, node-locked licenses, etc.)
- **Sharing Policies** – Moab Cloud helps apply rules that overcome key political barriers to sharing resources by taking into account all of the users, groups, projects, and accounts before applying prioritization, reservations, preemption, and other policies to achieve the highest service level possible.
- **Cluster/Grid Load Balancing** – Moab Cloud enables load balancing capabilities between projects on shared systems, and between multiple systems that are at a single site or geographically dispersed. With the needed data staging, credential mapping and policy management aspects of Moab Cloud, users are able to control which workloads must stay on local systems and which workloads can be allocated to remote systems as well as apply elastic computing to dynamically re-purpose resources to meet the needs of a new workload in order to facilitate "bursting."

2. Convergence

- **HTC and HPC Convergence** – Moab Cloud, when integrated with Nitro is able to manage both High Throughput Computing (HTC) workloads (many small tasks) and High Performance Computing (HPC) workloads (traditional batch and interactive workloads on both large parallel and serial workloads) on the same system.
- **Data Center Convergence** – Moab Cloud's workflow capabilities can allow it to link to non-cluster/grid applications and resources, thus enabling the importation of data from multiple data services in the data center. Moab Cloud can then transform the data into HPC/HTC tasks, interact with datacenter databases, and drive post-processing tasks.

3. Automation

- **Workload Automation** – Moab Cloud's job/task dependency capabilities allow multiple steps in a workload to be automated. It can launch steps based on success, failure, or in conjunction with other steps. This allows results of one task to be used by a subsequent task for a refined analysis. An example of this is Moab's capability to run a market analysis simulation and spawn many subsequent pricing jobs/tasks based on the results.
- **Process Automation** – Moab Cloud's trigger capabilities allow it to build out workflow steps that are not directly part of a workload. Triggers can be attached to workloads, nodes, storage, network, reservations, etc. and allow actions based on combinations of data, object state, events, metrics, counters, service levels, or environmental conditions. With these qualities, Moab Cloud is able to stage data, turn on a node, run a health check, do a security scrub on a node after a workload is complete, and many other script or command driven tasks. These can be chained together with no direct workload involvement or added to a workload to automate a full end-to-end process or data pipeline.

4. Optimization

- **Run Time Optimization** – Moab Cloud's time-aware predictive scheduling takes the information from its world view of resources, current and reserved workloads, planned maintenance, deadlines, service commitments, wait time, etc., and models out into the future how to best pack workloads. This optimizes how early a workload can be run and which resources will run it the fastest. Through dynamic prioritization, Moab Cloud is able to preempt lower priority workloads while intelligently avoiding the preemption of workloads that are close to finishing or that have already been preempted a given number of times.
- **Power Optimization** – Moab Cloud is able to adjust power settings (on, suspend, hibernate, off, etc.) on nodes that are projected to stay idle for a period of time. Moab Cloud is also able to dynamically adjust clock frequency levels to match the needs of workloads and route workloads to resources that will give the greatest performance to power ratio.
- **License Optimization** – Moab Cloud is able to be configured to act licence-aware, knowing that a particular application is licensed per node and that workload from that application should have an affinity to run on nodes that already have the license in use. It can integrate with FlexLM for dynamic

license management, and know that there are a limited number of licenses for a given application and ensure work is not submitted to nodes when licenses are not available. These capabilities reduce the number of licenses needed and get more out of the hardware and licenses already owned.

- **Cost Optimization** – Moab Cloud can be configured to set a resource cost for different hardware, storage, GPUs etc. and have its scheduling algorithm achieve the lowest cost that will still meet SLA objectives. Further, simply by implementing show-back or charge back policies (showing costs or actually charging organizations based on usage), organizational behavior naturally adjusts to optimize according to costs.

5. SLA Delivery

- **Ownership, Efficiency and Fairness** – Moab Cloud is able to translate the equivalent of siloed resource rights that an organization possesses into policy-enforced guarantees. To do so, Moab Cloud uses resource reservations, allocations and other such policies.

- **Accounting** – Moab Cloud has accounting management capabilities that enable an organization to monitor and communicate successful delivery of guaranteed resources. It can track allocated resources and calculate the costs of usage for users, groups, accounts, etc. and then report it in "show-back" mode, or enforce it in "charge-back" mode by adjusting prioritization or blocking usage so users don't go past their allocation.

6. Agility

- **Changing Resource Allocations based on Business Priorities** – Since allocations of resources are policy-based and not physically siloed resources, organizations are able to make prioritization changes quickly and change allocations as fast as they can make the decision.
- **Dynamic Re-purposing** – If integrated with provisioning, virtualization or cloud management tools, Moab Cloud is able to dynamically drive the re-purposing of nodes to match the current demand of workload. This helps overcome the fracturing of one's compute resources because of OS or current application set.
- **Quickly Integrating with In-house Tools** – Moab Cloud's architecture is designed to monitor and manage different aspects of a compute environment or workflow by interfacing with the tools that are the domain experts for any given functional capability. In many cases, Moab Cloud is able to interface with in-house custom tools, thus retaining important historical investments.

7. Modularity

- **Start Simple** – Moab Cloud's policies and capabilities can be turned on or off, enabling organizations to start with simply optimizing workload packing and enabling any of the above mentioned capabilities over time. Products are also sold in a modular way so that you only pay for the capabilities you need.
- **Additive and Open Foundation** – Adaptive Computing has built an open architecture that allows it to add value to pre-existing investments in tools and processes.

Low Risk Integration/Migrations -- Third-Party Tools

Moab Cloud and Nitro are architected to enable low risk integration with existing third-party products. They can act either as a supplemental service or as a means to migrate part or all workloads to Adaptive Computing products over time. This architecture also allows organizations to sample our products on a subset of workloads before deciding to either rollback to the third-party product or to increase the workloads that are managed by Moab Cloud or Nitro.

IBM Platform Symphony

Moab Cloud and Nitro can be integrated with Symphony to allow workloads submitted to Symphony to be routed to Moab Cloud/Nitro (typically Linux workloads that use simple Symphony API integration). The Symphony command line and UI are able to see both workloads in this case, and workloads running under Moab Cloud/Nitro appear to be running on a single large node. This allows Moab Cloud/Nitro to be evaluated and avoids disruption of current submission and management practices. Over time, workloads may be submitted directly to Moab Cloud/Nitro, thus picking up added capabilities and reducing licensing costs (up to 40-60 percent savings) as organizations chose to do so.

IBM Platform LSF

Nitro can be integrated to add high throughput task launching to LSF (up to 100x faster) with or without the use of Moab.

SGE/OGE/UGE

Moab Cloud and/or Nitro can be added to Grid Engine in order to add improved policy management or high throughput task launching (tens of times faster than grid engine and multiple times faster than grid engine short jobs). Moab Cloud is also able to connect clusters that use grid engine, TORQUE, SLURM, or other traditional HPC schedulers into a grid.

Other Custom and Commercial Schedulers

Moab Cloud's architecture has allowed it to connect to other schedulers in the market, as well as custom in-house schedulers and orchestration engines. This is commonly to add load balancing and policy controls to multi-system grids. In one case, Moab Cloud had the potential to save over one hundred million dollars just by applying its load-balancing capabilities on a custom tool that had low utilization.

Other Middleware

Moab Cloud has integrated with FlexLM, HP CMU, IBM Tivoli Provisioning Manager, BMC Bladelogic, xCAT, Nagios, Ganglia, VMware, and many other tools. These middleware integrations allow for health monitoring, dynamic re-provisioning, license scheduling, power management, etc.

Conclusion

Organizations are seeing that they may soon have a significant increase in new workloads that need processing. This comes from the explosive growth of data (2x faster than compute), the more in-depth level at which it must be analyzed, and the many new financial risk / stress tests that need to be run. As they face this increase, organizations can throw more hardware at the problem as they have in the past and thus substantially increase their hardware spend, or they can change how they are doing things to significantly improve efficiency and work within moderate to flat hardware budgets.

To further exacerbate the issue, a growing number of organizations have expressed that the high costs that they are paying for their historical workload management solutions from IBM are increasing yet again, and they need to change much of their hardware purchase from IBM System x to a new vendor due to IBM's sale of the server line to Lenovo.

Adaptive Computing's Moab Cloud and Nitro products, in conjunction with leading hardware partners, can be a key solution to enable the transformation of an organization's siloed, disjointed and under utilized systems (clusters/grids) into a consolidated, shared, and workflow-automated compute environment. This agile infrastructure delivers the needed utilization efficiency to offset what would have been significant cost increases, provides the optimization needed to get results faster, and the service-level guarantees needed to gain political buy-in from participating groups.

Contact a solutions advisor by phone or email, or visit our web site today.

North America, Headquarters +1 (239) 330-6083
Provo, UT, USA Office +1 (801) 717-3700

Corporate Headquarters
704 Goodlette Road North
Naples, FL 34102

Email: info@adaptivecomputing.com
www.adaptivecomputing.com