Integrate runs Azure Batch job using 150,000 cores

INTEGRATE
BUILT BY **Milliman**

Thomas Peplow, UK Development Lead

Overview

MG-ALFA® Compute for Azure™ was developed by Milliman as one of the first cloud-based hyper-scale computation platforms. Milliman first launched Compute for Azure in 2011, providing clients with access to the Microsoft Azure cloud for virtually unlimited computing capacity on a variable cost basis.

The transition of workloads from on-premises grids to the cloud has been significant.

Integrate™, a revolutionary cloud-based financial projection system for the life insurance industry, was built by Milliman and also launched in 2011. Compute for Azure is a key module in the end-to-end Integrate platform, providing clients a cost-efficient way to meet the computational demands of ever-increasing workloads, driven in part by regulatory requirements such as Solvency II and principle-based reserves and capital. By the end of 2016, clients using Compute used a total of 138 million central processing unit (CPU) hours, which is the equivalent of running one CPU for just under 16,000 years.

Milliman now leverages Azure Batch for more efficient provisioning and management of compute resources, leading to significant performance improvements and cost reductions. This new release of Compute for Azure provides clients the many benefits described below.

Faster start time

From an all-idle state, typical 1,000-task compute jobs now start running after two minutes, with 90% of calculations starting in just four minutes, and the balance of calculations starting inside of 11 minutes. This is around *85% faster* than the previous release of Compute.

Reduced cost

For workloads where the required number of CPUs varies over time, the system automatically shuts down the underutilized capacity. The graph demonstrates an example from an early adopter of Compute for Azure with Azure Batch. In this case, the client *saved 50%* of its total run-time cost.

FIGURE 1: COMPUTE FOR AZURE WITH AZURE BATCH
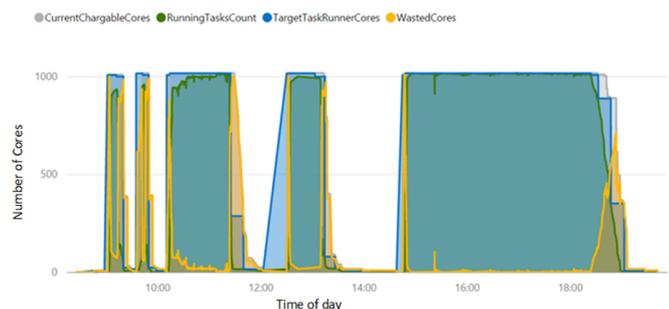
Reduced run time

Granular control over resources allows clients to increase the maximum number of CPUs that can be used in a given time window. For example, an Integrate *client reduced its run time by 10 hours*, using 30,000 CPUs for nine hours, instead of 8,000 CPUs for 19 hours. This used an additional 10,000 CPU hours, but roughly doubled throughput at a cost increase of only $1,500.

Increased scale

Compute for Azure has now been successfully tested with a single job that used *150,000 CPUs*. This scale test simulated a nested stochastic projection with 100 outer loops and 1,000 inner loops.

Nested stochastic modeling

A big push for Integrate this year is to enable nested stochastic modeling. Stochastic modeling is a complex mathematical process that uses a combination of probability and random variables to forecast financial performance or financial requirements. Nested stochastic models, as the name implies, are stochastic models inside of other stochastic models.

For example, a 1,000-scenario model with reserves and capital based on 1,000 paths at each valuation point for a 30-year monthly projection requires the cash flows for each policy to be projected 360 million times. And it can get even more complicated than that.

Compute for Azure utilizing Batch is making this goal a reality. The new functionality in the Compute platform supports this complexity for both Integrate and MG-ALFA clients. Compute enables much larger and more complex jobs containing many millions of interrelated tasks. As mentioned above, this approach has proven to scale to 150,000 cores, which makes a nested stochastic projection of 150 outer loops with 1,000 inner loops a reality, requiring minimal model changes.

With a simple desktop plug-in, MG-ALFA desktop customers can be up and running in minutes, routing nested stochastic workloads to thousands of cloud-based CPUs metered by the minute.

In the context of auto-scaling, this means the increase in cost that results from requiring 150,000 cores is managed by automatically adjusting the required computational resource to minimize overhead. In addition, Compute can ensure that all inner loops occurring in the same time period are 100% parallelizable, which will have a significant impact on run time.

Partnerships

Milliman has a long-standing working relationship with Microsoft. This partnership has seen many fruitful endeavors, which not only benefit Integrate clients, but the entire cloud ecosystem. The Integrate team is bolstered with industry-leading consultants from Arapsys and Endjin, who provide ongoing strategy and technology consulting to keep Integrate at the forefront of technology.

Additional Compute for Azure benefits

Beyond using Azure Batch, other enhancements include:

- **More variation in compute resources, including multicore virtual machines.** This is of great value to clients who have large memory requirements, because the memory-intense virtual machines open the opportunity to run these jobs faster.

- **Reduced data distribution overheads.** Using an innovative data distribution approach, Compute for Azure is able to reduce the cost of input and output data distribution by around 85%. That produces faster results and reduced data movement cost.

- **Increased number of concurrent tasks per job from 1,000 to 5,000.** For large scenario distribution jobs that can run fully parallel, this enhancement can yield results up to five times faster.

- **New map/reduce approach to cell-distributed jobs, enabling thousands of tasks with thousands of scenarios.** This capability has a huge run-time improvement for large cell-distributed workloads.

Milliman

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com   integrate-solutions.com

FOR FURTHER INFORMATION, CONTACT:
Thomas Peplow
thomas.peplow@milliman.com