

TGen improves the visibility, efficiency, & performance of Nextflow workflows on AWS

By deploying the Memory Machine Cloud executor for Nextflow

Translational Genomics Research Institute (TGen) is a leading nonprofit based in Phoenix and focused on translating genomic research into life changing results.

Cloud Cost Optimization is Not Easy at Scale

TGen struggled to run its large Nextflow workflows cost-efficiently on AWS. Using on-demand EC2 instances was too expensive, especially for larger jobs requiring thousands of nodes. Using cheap Spot EC2 instances sounded good on paper, but in practice failure rates could be as high as 80%, resulting in longer completion times as failed jobs had to be restarted multiple times and less predictable cost savings.

More Visibility, Efficiency, and Performance = Better Research

TGen leverages MemVerge Memory Machine Cloud to execute its Nextflow workflows on AWS. Memory Machine Cloud features such as FLOAT, SpotSurfer, WaveRider, and WaveWatcher seamlessly integrate with Nextflow as a supported cloud executor.

- Automated cloud resource management (FLOAT)
- Deep insights into cloud resource utilization (WaveWatcher)
- Cost and performance optimization (SpotSurfer + WaveRider)

Improving the Nextflow Runtime Experience in the Cloud

Since deploying Memory Machine Cloud, TGen is now able to run its workflows easily and cost-efficiently on AWS without the high costs of on-demand EC2 and without the high failure rates that can come from running large workflows entirely on Spot EC2.

- Organization-wide visibility into job-level resource utilization reports and insights
- Enhanced workflow execution with Spot EC2 (from 80% failure rates → below 1%)
- Automated rightsizing of EC2 instances at runtime



"WaveRider picks the most appropriate VM type for each of the thousands of jobs launched by NextFlow. That is giving me more efficiency than ever. Very cool."

– Vince Pagano

Senior Scientific Programmer, TGen

Before

Memory Machine Cloud

~80% Job failure due to Spot reclaim

After

Memory Machine Cloud

~0% Job failure after Spot reclaim

