

# WaveRider

## for Sentieon Genomic Tools

### Two Waves of Acceleration & Cloud Cost Reduction for Secondary Analytics

Next Generation Sequencing (NGS) machines process short fragments of genetic material using massively parallel technology to produce data files in industry-standard text format (FASTA or FASTQ). In secondary analysis (the bioinformatics phase), fragment sequences are assembled into contiguous sequences and then into targeted regions of DNA or RNA, or even a complete genome. Secondary analysis is computationally and resource intensive and can take much longer than fragment sequencing.

#### Wave 1: 5-10x shorter runs and 5-10x less cloud costs

Sentieon® provides a commercial package of bioinformatics analysis tools for processing genomics data. The Sentieon Genomics tools are new implementations (complete rewrites) — stressing computational efficiency — of the algorithms underlying popular open-source tools such as GATK, Picard, and BWA-MEM. Sentieon Genomics tools show a five- to ten-fold performance improvement over the equivalent open-source tools.

#### Wave 2: Another 40% faster and 40% less cost

Memory Machine Cloud (MMCloud) streamlines the deployment of containerized applications on cloud compute instances. MMCloud includes AppCapsule, a feature that automatically migrates running processes, from one compute instance to another, without losing execution state. This feature dynamically rightsizes compute instances and protects against spot instance reclaims.



#### How it works

In variant calling - the analysis that identifies variants from genome sequence data - resource demands on CPU cores and memory vary as the pipeline executes. When Sentieon tools are used for variant calling, multiple data processing steps execute on a single cloud compute instance. The compute instance must be sized for the peak resource demands, or the pipeline will take too long or fail because of OOM (out-of-memory) errors. But it adds cost. When Sentieon tools are used with MMCloud, MMCloud migrates the running processes from one compute instance to another as the resource demands change. This ensures that the resources available to the job are not over-or under-provisioned — a capability not available from other cloud automation solutions, which provide horizontal scaling only.

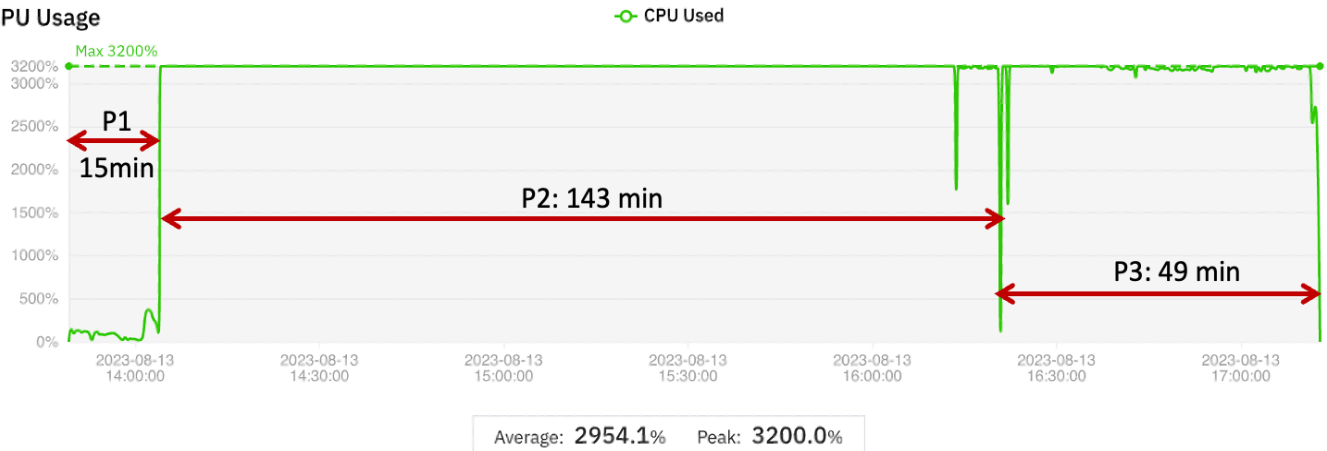
Learn More



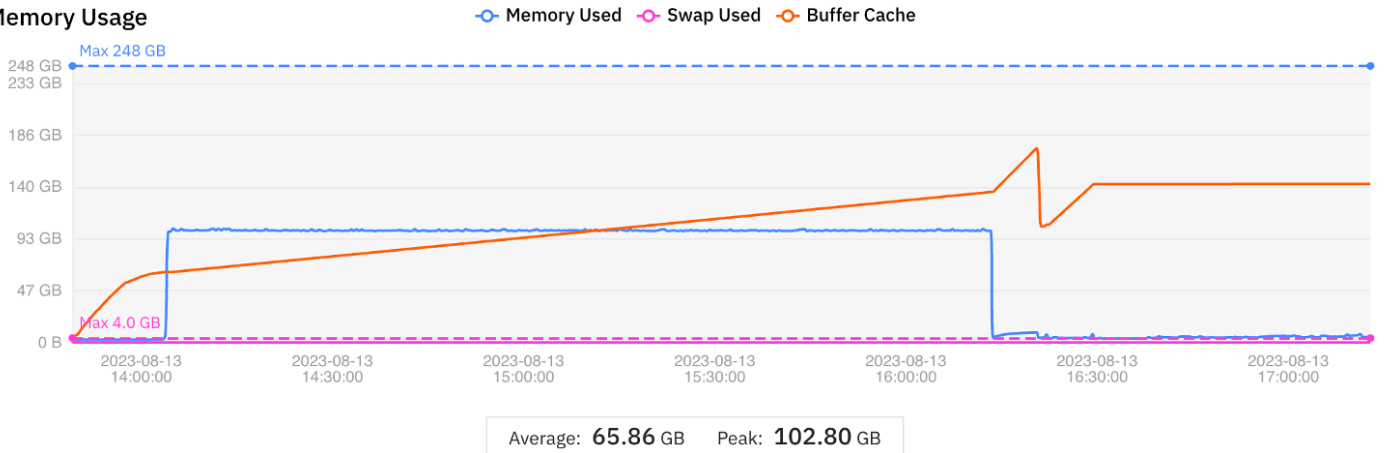
## Test Results

Benchmark tests compare the performance of a whole genome sequencing (WGS) pipeline analyzed with Sentieon's DNAscope product, executed with and without MMCloud, running on AWS EC2 instances. With MMCloud, wall clock time and cost are both reduced (by 40% and 34%, respectively)

### CPU Usage



### Memory Usage



Learn More

