



Accelerated Genomics on Xilinx Alveo™ Accelerator Cards

Genome sequencing on hybrid cloud

- 10x Speed up over original GATK pipelines
- Single acceleration platform across clouds
- Support for multiple GATK versions for germline & somatic pipelines

INTRODUCTION

By 2025, Genomics is projected to produce 40 exabytes of data annually as the latest genomic sequencers generate 2 or more terabytes of data daily. More data opens the opportunity to improved patient care with a caveat. Bioinformaticians and researchers face the dual challenge of having to accelerate workloads to achieve quick and timely results while performing cost-effective analysis. Falcon Computing's solution leverages Xilinx FPGAs to provide a fast, automated and cost-effective framework to accelerate genomics pipelines for secondary sequencing analysis.

Speed and accuracy are key requirements in genomics secondary sequencing. For a consistent and accurate analysis, the Broad Institute developed a set of pipelines that have been established as the industry standard in the field of genomics. Their pipeline takes over 50hrs to run a whole genome and hence the need for acceleration. With Falcon computing, bioinformaticians can run 10x faster while getting the exact same results as they would without acceleration as there are no source code changes.

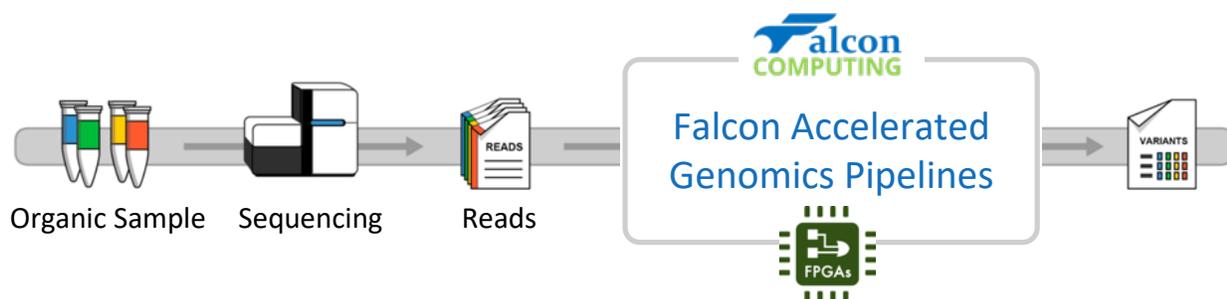
PRODUCT OVERVIEW

Working in partnership with Xilinx, the Falcon Accelerated Genomics Pipelines (FAGP) work seamlessly on local appliances with the Alveo™ U200 Acceleration Card or on the public cloud. On the cloud, the platform is currently available on instances for AWS, Huawei, and Alibaba Cloud.

SOLUTION OVERVIEW

Genomics Accelerated Platform for Secondary Sequencing on Hybrid Cloud

Falcon Computing accelerates GATK Best Practices Pipelines for secondary sequencing. Secondary sequencing is the process of cleaning up data coming out of the sequencer and aligning it to a reference genome to identify variants that can be linked to diseases or other unique characteristics. Using the industry standard GATK pipelines could take over 50hrs to finish this process. Falcon solution leverages Xilinx FPGAs to do it a fraction of the time.



Up to 10x improved performance using Xilinx FPGAs

Falcon provides a software solution with consistent performance across available platforms with Xilinx FPGAs. This provides users with the flexibility to continue enablement of their hybrid cloud strategy.

Platform	BWA (sec)	MarkDup (sec)	BQSR (sec)	HTC (sec)	Total (hrs)	Speed up
Alveo™ U200	7444	2154	4198	3557	4.82	10x
AWS: f1.16xlarge	8768	2077	4477	4023	5.37	10x
Huawei: Fp1c.8xlarge.8.I	11129	4761	2379	7178	7.18	7x

Speed ups compared to Original GATK runtime of +50hrs on 16-core CPU for NA12878 WGS

Merlin Compiler for proprietary pipelines

While bioinformaticians are heavy users of the GATK pipelines, there are instances where they've developed proprietary pipelines for specific projects or research opportunities. As part of the Falcon Computing product portfolio, users can also leverage FPGA acceleration on other pipelines using the Merlin Compiler which effortlessly transforms C/C++ into FPGA without requiring hardware development knowledge. With the Merlin Compiler you can achieve up to 35x algorithm performance improvement compared to CPU and you can achieve 10x productivity compared to manual FPGA optimization

CONCLUSION

Falcon Accelerated Genomics Pipelines and the Merlin compiler are the answer to a cost and performance efficient path to acceleration. Whether it is to accelerate GATK pipelines, data compression algorithms or any other proprietary pipeline, Falcon Computing products can make it easy.

TAKE THE NEXT STEP

Try it on the AWS cloud: <https://aws.amazon.com/marketplace/pp/B07C3NV88G>

Explore Falcon Genomics Acceleration Solution with Xilinx Alveo™: www.xilinx.com/alveo

Xilinx Alliance Program page: <https://www.xilinx.com/alliance/memberlocator/1-dfp9rh.html>

For more information and developer tools: www.falconcomputing.com



Falcon Computing HQ
3979 Freedom Circle Ste 530
Santa Clara, CA 95054
USA
www.falconcomputing.com

Falcon Computing Los Angeles
10990 Wilshire Blvd #920
Los Angeles, CA 90024
USA
www.falconcomputing.com

Falcon Computing China
64 N 4th Ring Rd W
Haidian Qu, Beijing Shi
China
www.falconcomputing.com

Falcon Computing enables data scientists and software developers working in Genomics, Machine Learning, Finance and Computer Vision to simplify their path to heterogeneous accelerated computing. With a simple push-button C/C++ to a highly-optimized FPGA implementation for both public and private clouds, users can achieve performance, energy-efficiency and productivity gains. Falcon Computing was founded in 2014 by Dr. Jason Cong, serial entrepreneur and Director of the CDSC at UCLA.

