



**ADVANTECH**

*Enabling an Intelligent Planet*

**intel**®



**Accelerating Personalized Precision Medicine for Rare Diseases:  
Enhancing Genomics Workflow and Computational Efficiency with  
WASAI and Advantech's High-Speed Analytics Platform**

intel.  
partner  
Titanium

## AGENDA

- **Introduction**
- **Challenges**
- **Solution Overview**
- **Case Study:**
  1. South Korea Yonsei University Mirae Campus
  2. National Taiwan University College of Medicine Use for Large Population Cohort and Disease Study
- **WASAI Lightning Bio-IT Platform**
- **Advantech High-Performance, High-Density Servers**
- **WASAI Lightning Performance Validation**
- **Conclusion**
- **Why Advantech**
- **About WASAI**
- **Additional Information**



## INTRODUCTION

The emergence of next-generation sequencing (NGS) has transformed DNA sequencing capabilities, generating vast amounts of genomic data that require effective processing, management, and analysis. However, the challenge for NGS is maintaining accuracy when handling the extensive raw data generated during the secondary analysis phase. Despite their significance, the development of specialized in-house bioinformatics tools has proven to be costly. Consequently, laboratories prefer solutions based on open-source tools with hardware acceleration that can be seamlessly integrated into their existing infrastructure and flexibly customized. The demand for time and cost efficiency in handling complex tasks requires that servers, racks, storage, and network infrastructure be upgraded to meet the evolving needs of genomic research and diagnostics.

To address these challenges, WASAI collaborated with Advantech to leverage the capabilities of its SKY-820V3, SKY-620V3, and SKY-8234D platforms for genomic analytics. These innovative platforms combine 5th Gen Intel® Xeon® Scalable Processors with Altera's Agilex® field-programmable gate array (FPGA) cards, harnessing WASAI acceleration for genomic secondary analysis using the BWA-MEM and GATK HaplotypeCaller algorithms.

This groundbreaking approach leverages the latest Altera® Agilex™-based FPGA cards to concurrently process billions of DNA base pairs. The result is a cost-effective, precise, and rapid genome processing solution rooted in the Intel Select Solutions of Genomics Analytics. This not only harnesses exceptional computing power but also provides users with the freedom to use open-source software platform tools for building custom analysis pipelines. This dual advantage significantly reduces turnaround time and total cost of ownership (TCO).



## CHALLENGES

The exponential growth in Life Sciences, fueled by massive developments in next-generation sequencing (NGS), has created urgent demand for high-performance computing (HPC). The proliferation of NGS technologies, resulting in the rapid accumulation of massive genomic data, underscores the need for efficient data processing solutions. While NGS technologies are cost-effective, challenges arise in managing and analyzing the vast data generated. HPC is crucial for effective big data management in the field of Life Sciences, particularly during the secondary analysis of the human genome.

The increasing availability of millions of genomes nowadays further intensifies the need for robust computational solutions. Each genome, comprising billions of nucleotides stored in gigabytes, poses unprecedented challenges in data handling. Whole genome sequencing (WGS) involves deciphering approximately 3 billion bases, with data for a single human genome sequence ranging from 1 to 200 gigabytes. The sheer volume of genomic data highlights the critical role of high-performance computing in overcoming the challenges posed during secondary analysis. As demand for high-performing appliances grows with the surge in genomic data, the associated high financial costs are a barrier to user access. This underscores the pressing issue faced in the current landscape of genomic analysis, where a balance between cutting-edge technology and affordability is vital for advancing Life Sciences.

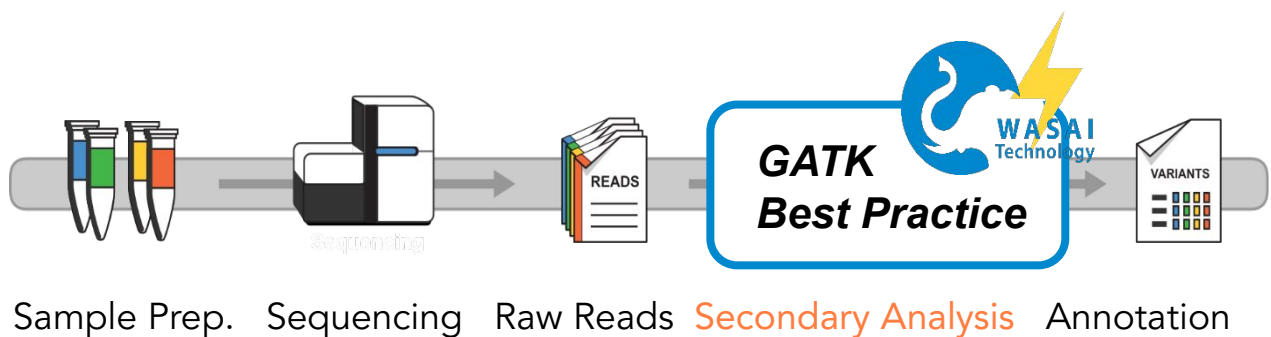




## SOLUTION OVERVIEW

### WASAI-Lightning Bio-IT Platform

The WASAI-Lightning Bio-IT platform is meticulously crafted to elevate the performance of the latest Agilex-based FPGA cards aimed at genomic secondary analysis. The platform employs a specialized memory hierarchy and pipeline design to target the BWA-MEM, Sorting, MarkDuplicates, and HaplotypeCaller algorithms. This enhances the efficiency of frequent memory access and provides a precise and high-speed analysis solution. Analyzing FASTQ datasets through the WASAI-Lightning mapping and variant-calling pipeline provides a highly accurate solution with significantly reduced execution time. Moreover, the reconfigurable nature of the platform ensures its continued suitability by offering a versatile and adaptive solution for the ever-evolving demands of bioinformatics.





## CASE STUDY

The Usage and Goal of South Korea Yonsei University Mirae Campus and Taiwan National Health Research Institutes for Large Population Disease Study.

Leveraging the performance, scalability, and cost-effectiveness of Advantech's genome-optimized server series, the WASAI Lightning solution assists research teams in South Korea and Taiwan with establishing unique population reference biobanks, eliminating the need to rely solely on Western population data.

The Yonsei University Mirae Campus in South Korea has long nurtured the medical ICT industry. Strategically located in the original state of the Digital Healthcare and Precision Medicine Regulation-Free Zone, the campus is equipped with the optimal infrastructure for medical AI research and drug discovery research. By integrating large-scale data such as public health data and genetic data, Yonsei University has established a state-of-the-art data center to support numerous internal and hospital-based disease research projects. Dr. Sungwook Ha, Professor in the Department of AI Semiconductor and CEO of NUMP, a startup founded by professors, remarked, "For various research projects, we need a highly scalable, flexible, and efficient hardware infrastructure. Integrating the WASAI Lightning Bio-IT platform with Intel Select Solutions for Genomics Analytics provides substantial benefits in terms of both performance and cost."



연세대학교  
YONSEI UNIVERSITY



## CASE STUDY

In Taiwan, WASAI has partnered with Taiwan National Health Research Institutes (NHRI) and collaborates with Professor Shih-Feng Tsai's team at the Institute of Molecular and Genomic Medicine. Their research explores the relationship between phenotype and genotype in Developmental Delay Intellectual Disability (DDID) using Whole-Genome Sequencing (WGS). They compare the findings with healthy samples from the Taiwan Biobank to gain a broader understanding of the genetic factors involved and hope to help the treatment and prognosis of the children patients with genetic diagnoses. Dr. Yung Feng Lin, a Postdoctoral Research Fellow in Professor Tsai's team, highlights that "the impact of the WASAI solution, noting its scalability, flexibility, and cost-optimization, significantly enhances the efficiency of the WGS analysis. This, in turn, aids in the early diagnosis of genetic disorders."





## CASE STUDY

Calvin Hung, CEO and Co-founder of WASAI Technology emphasized, “WASAI Lightning supports researchers in disease and population cohort research by adopting the latest advanced hardware from our ecosystem partners — Intel and Advantech. Research institutes can establish their own data centers cost-effectively while ensuring data accuracy and reducing turnaround time. This facilitates faster genomic discoveries for the benefit of omics research.” Fast genome analytics represents a transformative shift in the landscape of medical diagnostics and treatment, emphasizing precision medicine's potential to revolutionize patient care. By enabling rapid analysis of an individual's genetic makeup, this technology facilitates the identification of genetic predispositions to certain diseases, including cancer, rare diseases, and genetic disorders, at an early stage when they are often more treatable. The integration of fast genome analytics into medical practice not only enhances the precision and efficiency of disease diagnosis and treatment but also marks a significant step forward in the journey toward truly personalized healthcare. As WASAI’s technology advances and becomes more accessible, its role in improving long-term health outcomes and transforming the healthcare landscape will continue to grow.”







## Advantech High-Performance, High-Density Servers

To complement the breakthrough capabilities of WASAI's Bio-IT platform, Advantech provided three advanced server solutions: SKY-820V3, SKY-620V3, and SKY-8234D. SKY-820V3 is a 5G BBU edge server with a 2U form factor powered by dual 5th Gen Intel® Xeon® Scalable Processors. Equipped with features like 16 DDR5 ECC-REG RDIMM/LRDIMM slots, four PCIe 5.0 x16 slots compatible with dual-width GPU cards, and versatile connectivity options, it offers a comprehensive solution for demanding computational tasks.

SKY-620V3 is a proprietary server board featuring 5th/4th Gen Intel® Xeon® Scalable processors and up to 2 TB of DDR5 RDIMM (4400 MHz) for efficient memory performance. Enhanced storage options with five SATA 3 ports and two M.2 connectors (SATA/PCIe compatible) and the provision of two PCIe x16 slots and four PCIe x8 slots ensure flexibility for expansion. Together, these two server products provide the advanced computing hardware needed for computationally demanding genomic analysis.



### **SKY-820V3**

2U Short Depth Server with  
5<sup>th</sup> Intel Xeon Scalable Processors



### **SKY-620V3**

2U GPU Server with  
5<sup>th</sup> Intel Xeon Scalable Processors



## Advantech High-Performance, High-Density Servers

The SKY-8234D is designed and optimized to meet the high availability and for mission-critical use cases. It is based on the 5th Gen Intel® Xeon® Scalable Processors, which not only upgrades the total core counts per each consumed power watt, but also provides a high-performance and highly reliable platform. By pairing with multiple FPGA Cards, SKY-8234D offers a stable business-critical server with networks and computing accelerations. Powered by 5th Gen Intel® Xeon® Scalable Processors and the highly flexible PCIE and IO expansion capability, SKY-8234D can also be applied in enterprise private network infrastructure, scientific computation, network edge computing, and specific mission-critical application platforms. SKY-8234D is also designed for carrier-grade environments and to accommodate a broader range of temperature, dust, and humidity to deal with the application system running in harsh environmental conditions. Its redundant and field-replaceable PSU, fan modules, and management firmware minimize costly downtime, service interruptions, and onsite interventions.



### **SKY-8234D**

2U Edge Server with  
5<sup>th</sup> Intel Xeon Scalable Processors



## WASAI LIGHTNING PERFORMANCE VALIDATION

### Pipeline Comparison: Altera Agilex™ FPGA + Intel Xeon SP CPU Versus CPU Only

When working with vast datasets, more significant amounts of system memory can accelerate analysis speeds, while increasing the number of CPUs can reduce the time required for read mapping. WASAI's Lightning Bio-IT platform provides an all-in-one accelerated solution that integrates hardware and open-source genomics analysis software tools. For customers, this drastically reduces the initial investment and maintenance costs of deploying onsite infrastructure.

- **Hardware Configuration**

Built for demanding applications, Advantech's servers feature dual sockets for 5th Gen Intel® Xeon® Scalable Processors, 24 x memory DIMMs, 9 x PCI Express expansion slots, 8 x hot-swappable SATA bays, and 4 x LAN ports, as well as industrial-grade vibration tolerance to increase operational stability. The integrated fiber cards enable reliable high-speed networking, while the FPGA PCI Express cards allow accelerated deep learning and data analytics.

For streaming big data, a constant 400 Mbs input/output throughput between the server and storage is required. Accordingly, Advantech's servers are equipped with a 10 Gigabit network link to ensure reliable always-on network performance. This allows analyzed data to be uploaded to a centralized database for subsequent downstream analysis, variant annotation, data sharing, and genome referencing. Individual laboratories, each with its own fleet of sequencers, can temporarily store raw sequencer output data on cache storage proximal to their instruments. Data can then be easily transferred to centralized computing and storage infrastructure, including network-attached storage (NAS).



## WASAI LIGHTNING PERFORMANCE VALIDATION

The 5th Gen Intel® Xeon® Scalable Processors include more cores per processor with improved instructions per cycle, AVX-512 technology, and higher memory bandwidth, allowing for greater parallelism, more operations per clock cycle, and quicker access to large genomic datasets stored in memory. This reduces the overall time required to complete the alignment process and accelerates the computation of complex algorithms involved in sequence alignment and variant calling, where large datasets are divided into smaller data-intensive tasks, enabling faster genomic data processing with rapid access to vast amounts of information. With AVX-512, large blocks of genomic data can be processed in a single instruction cycle, reducing the computational overhead.

The improved performance-per-watt and high-speed transceivers in Altera Agilex 7 FPGAs provide complex computational tasks like sequence alignment and variant calling to be performed more efficiently and move large genomic datasets faster between different components of the processing pipeline, with less energy consumption. This is particularly important in large-scale genomic projects, where sustained high performance is required over long periods, making the process more cost-effective and environmentally friendly.



## WASAI LIGHTNING PERFORMANCE VALIDATION

- **Speed**

With WASAI's FPGA accelerated genomics solution, it only takes 3 hours to sequence a whole human genome, with the capacity to process eight whole genomes per day.

"The current combination of CPU, memory, SSD, and FPGA in a 2U server provides the optimum cost-performance fit. However, we will continue working to increase performance using multiple FPGA cards." Calvin Hung asserted.

- **Accuracy**

The WASAI Lightning solution with enhanced performance generates intermediate output files from the GATK4 best practice tools at a speed that rivals their CPU-generated counterparts, ensuring full equivalence. Additionally, the final variant call format (VCF) results from the GATK4 HaplotypeCaller tool exhibit an outstanding accuracy rate of 99.99% for both single nucleotide polymorphisms (SNPs) and insertion-deletions (INDELS) in terms of sensitivity and precision. Notably, the GATK4 baseline variant caller is nondeterministic, producing slightly varied results based on runtime parameters such as the number of threads, and aligning consistently with these variations during GATK4 execution.

In the comparison between the NA12878 and NA24385 VCF files and the NIST/GiaB v3.3.2 benchmark VCF file, constrained within the coordinates of the accompanying GIAiB v3.3.2 BED file, precisionFDA's NISTv3.3.2.vcf.gz (benchmark VCF) and NISTv3.3.2.bed (benchmark BED) are utilized. This comparative analysis assesses the accuracy of the pipeline within the "confident" regions of the Genome in a Bottle NA12878 and NA24385 truth datasets. Table 1 below compares precisionFDA truth data, NA12878 (HG001) and NA24385 (HG002), using GRCh38 and GRCh37 as reference genomes.



## WASAI LIGHTNING PERFORMANCE VALIDATION

The comparison results indicate that the WASAI Lightning Bio-IT platform solution not only impressively accelerates data processing but also maintains a remarkable accuracy of 99.99% for sensitivity and precision compared to standard GATK best practices, as evident in both SNP and INDEL procedures. Examining the genome samples of NA12878 and NA24385 with 50X coverage, the comparison table highlights virtually identical accuracies between standard GATK best practices and WASAI's Lightning solution. These results demonstrate that the WASAI solution acceleration cards not only uphold the accuracy standards set by GATK software tools, but also achieve a tenfold acceleration in genome processing during secondary analysis.

Sample	Reference	Compare	SNP-Recall	SNP-Precision	SNP-F-Score	INDEL-Recall	INDEL-Precision	INDEL-F-Score
NA12878 50x	GRCh38	Lightning	98.24%	99.62%	98.90%	97.67%	99.30%	98.50%
		Standard	98.24%	99.70%	99.00%	97.67%	99.34%	98.50%
NA12878 50x	GRCh37	Lightning	99.95%	99.09%	99.50%	99.51%	99.14%	99.30%
		Standard	99.96%	99.33%	99.60%	99.52%	99.26%	99.40%



- **Throughput and Minimized Time**

Sample	Data Coverage	Data Size	BWA-MEM + GATK HaplotypeCaller Execution Time		
			5th Gen Xeon Only	Agilex™ FPGA + 5th Gen Xeon	Speed Up%
NA12878 (NIST v3.3.2)	WGS, 30X	~90GB	24hrs4mins	3hrs10mins	88%
HG002 (NIST v3.3.2)	WGS, 30X	~60GB	26hrs20mins	3hrs50mins	84%

## KEY BENEFITS

- Ultra-efficient workflows combined with hardware acceleration boost overall performance by over 88% for a shortened runtime of nearly 3 hours.
- Innovative analysis software with industrial-grade server hardware provides a breakthrough secondary analysis platform for genomic data sequencing and pattern referencing.
- Highly reconfigurable FPGA technology dramatically accelerates mapping, alignment, and variant calling.
- Advanced smart fan control optimizes FPGA thermal management to reduce noise levels and maximize computing power.
- The entirety of the solution hardware, including multiple processors, memory, and network cards, fits into a compact two-unit rackmount case, saving installation space.



## CONCLUSION

The WASAI Lightning Bio-IT solution, dedicated to optimizing GATK based on the Intel Select Solutions architecture, provided a remarkable 9-fold improvement in overall performance, reducing runtime from 24 hours to just 3 hours—surpassing the capabilities of generic solutions. This enhanced GATK demonstrates excellent performance in terms of accuracy and throughput while offering flexible scalability. The WASAI Lightning Accelerated Solution leverages the latest Altera® Agilex™ FPGA card with GATK pipeline and system optimization to increase computational efficiency. Furthermore, it showcases the scalability of CPU+FPGA, enabling further increases in other processing workloads within the infrastructure, and allowing scientists to plan HPC resourcing effectively.

The improvements from 4th Gen to 5th Gen Intel® Xeon® Scalable Processors and from Altera Stratix 10 to Altera Agilex 7 FPGAs translate into substantial gains in processing speed, efficiency, and accuracy for sequence alignment and variant calling. These tasks, foundational to genomics research, benefit from the increased computational power, memory efficiency, AI acceleration, and customizable architecture offered by these next-generation Intel hardware platforms. Consequently, genomic researchers can achieve faster, more reliable insights, enabling quicker decision-making and advancing personalized medicine and other genomic applications.

In partnership with Advantech, WASAI is committed to developing and implementing cutting-edge technology to drive innovation that propels global sequencing projects. Such initiatives bring us closer to making precision medicine a reality. Looking ahead, the integration of hardware acceleration in genomic analytics holds immense promise. It's not just about faster processing; it's about revolutionizing personalized medicine by tailoring treatments based on individual genomic data.





## About WASAI Technology Inc.

WASAI Technology's mission is to deliver acceleration technologies of high-performance data analysis (HPDA) in future data centers for targeted vertical applications with massive volumes and high velocities of scientific data. To strengthen and advance scientific discovery and technological research via big data-intensive acceleration in high-performance computing, WASAI Technology aims to improve the commercialization and commoditization of scientific and technological applications.

## WHY ADVANTECH

Advantech provides an extensive portfolio of mature, industrial-grade products that can be easily customized to deliver breakthrough solutions. For the bioinformatics market, Advantech offers comprehensive IoT solution architecture, including edge sensing, computing, connectivity, and server storage devices with an extended product lifecycle of up to 5 years. From components to systems, Advantech solutions offer design flexibility, operational reliability, and extended longevity support.

## ADDITIONAL INFORMATION

To learn more about the information described in this document, reference the following documents and web resources:

- [WASAI Technology](#)
- [SKY-820V3 product documentation and video](#)
- [SKY-620V3 product video](#)
- [Advantech Server Series](#)