
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

LFS305-S

SPONSORED BY INTEL

Optimized software for improved performance in cryo-EM and genomics

Michael J. McManus, PhD
Director of Precision Medicine & Principal Engineer
Intel Corporation

Kshitij Kumar
Founder and CEO
Clovertex

Dr. Toshio Moriya
Associate Professor
High Energy Accelerator Research Organization (KEK)

Brendan Gallagher
Head of Business Development
Sentieon, Inc.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cryo-EM discovery path

FROM SAMPLE TO 3D MOLECULAR STRUCTURES

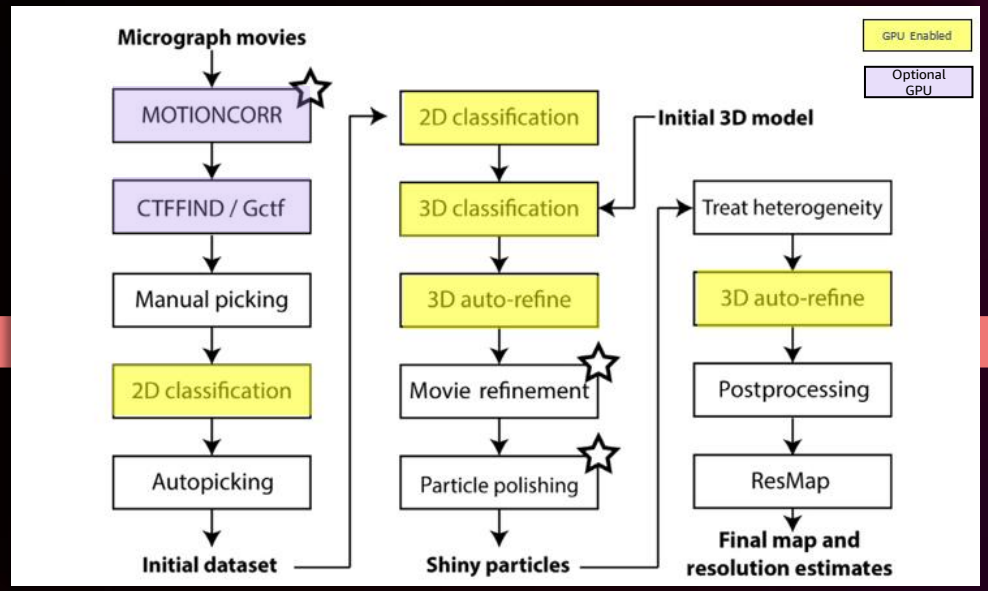
Sample



Experimental
(instrumentation)

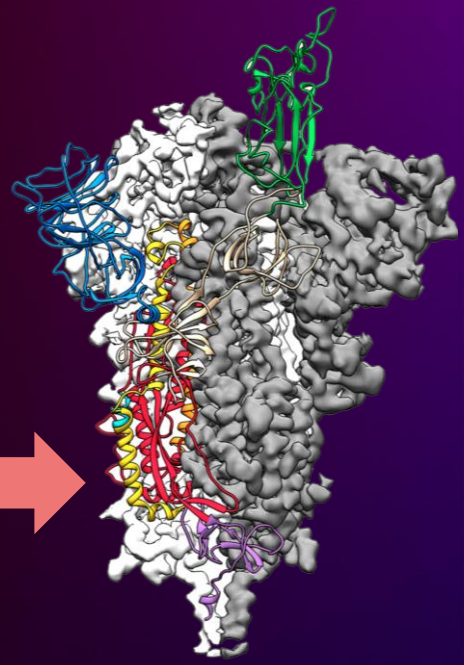


Computational analyses
(HPC resources + software)



HPC: high-performance computing

3D structure



Antibody binding
SARS-CoV-2 spike

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Clovertex
modernization made simple

Kshitij Kumar

Founder and CEO

Clovertex



KEK

Professor Toshio Moriya

KEK: High Energy Accelerator Research Organization

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Technology enablement across bio-pharma

Clovertex
modernization made simple

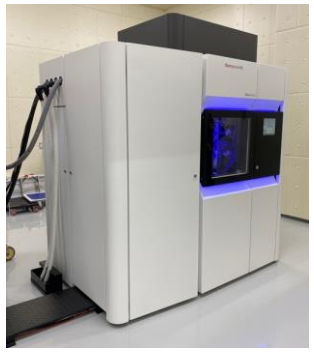
- Cloud-managed services
- HPC
- Database-managed services and modernization
- Application security and vulnerability testing
- Data science
- Robotic process automation (RPA)



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

KEK: High Energy Accelerator Research Organization

KEK Cryo-EM facilities



Talos Arctica
200 kV
Cryo-EM
Falcon 4

Coming soon

Titan Krios G4
300 kV Cryo-EM
Falcon 4i
Selectris X Filter



KEK

- Institute of Particle and Nuclear Studies
- **Institute of Materials Structure Science**
  - Synchrotron Radiation Science Div.1
  - Synchrotron Radiation Science Div.2
  - Neutron Science Division
  - Muon Science Laboratory
  - **Photon Factory**
  - Slow Positron Facility
  - **Cryo-EM Facility**
- Accelerator Laboratory
- Applied Research Laboratory
  - **Computing Research Center**

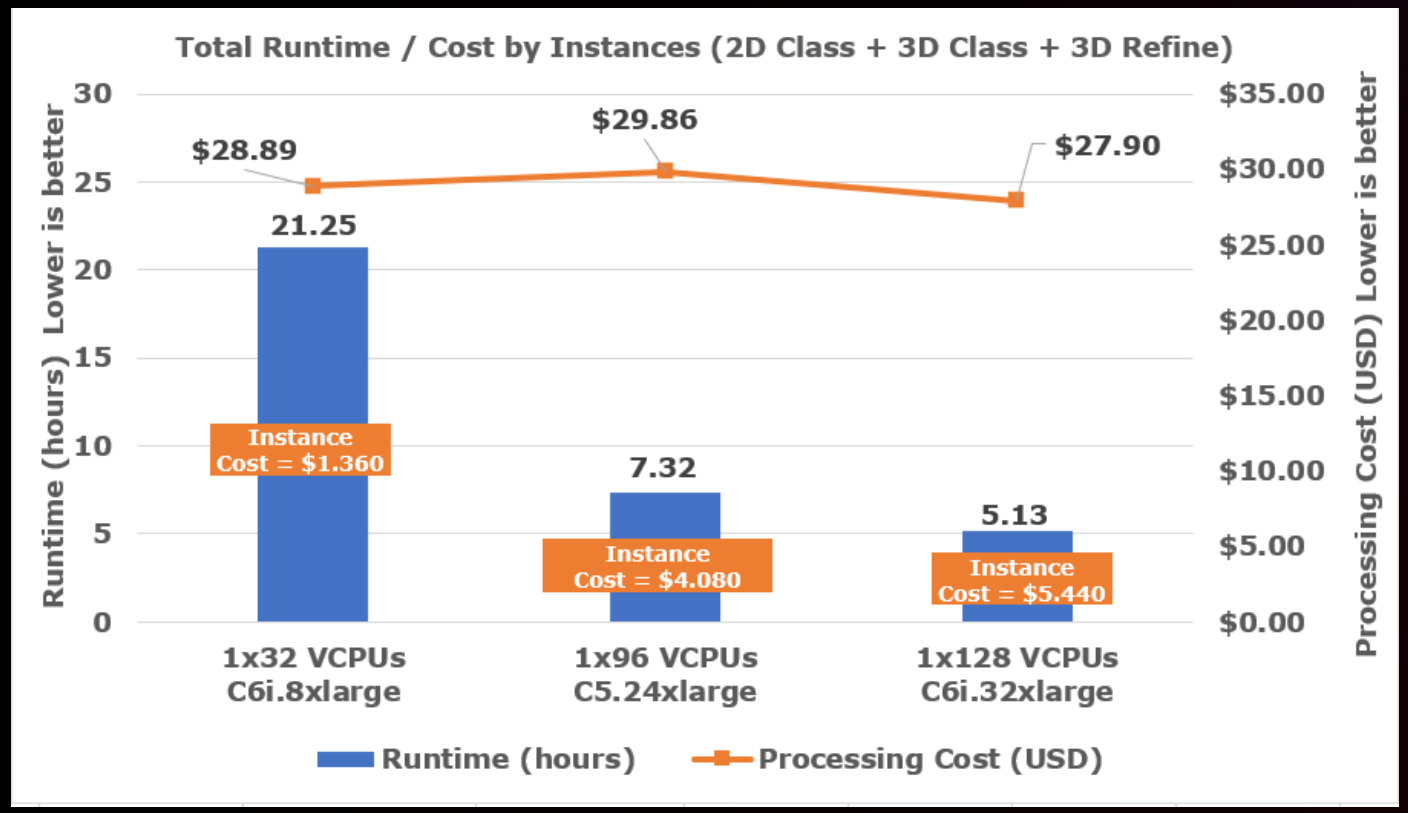
Center for Integrative Quantum Beam Science

Structural Biology Research Center
SBRC

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cryo-EM benchmarks – RELION (on AWS)

PLASMODIUM RIBOSOME – CRYO-EM PIPELINE – INTEL OPTIMIZATION FOCUS



KEY POINT:

The instance with the lowest hourly cost often leads to a higher processing cost than a more expensive instance

EXAMPLE:

- Instance cost for C6i.32xlarge is ~33% **higher** than C5.24xlarge
- However, the processing cost of the C6i.32xlarge is 7% **lower** than the C5.24xlarge instance

RESULT:

- Your job is a little lower cost and ~2 hours faster using the more expensive instance

Instance Type	Instance Cost (/hr)	Compute Hours	Cost Ratio	Compute Ratio	Compute Cost
C6i.32xlarge	$5.440	5.13	1.333	1.427	$27.901
C5.24xlarge:	$4.080	7.32			$28.89

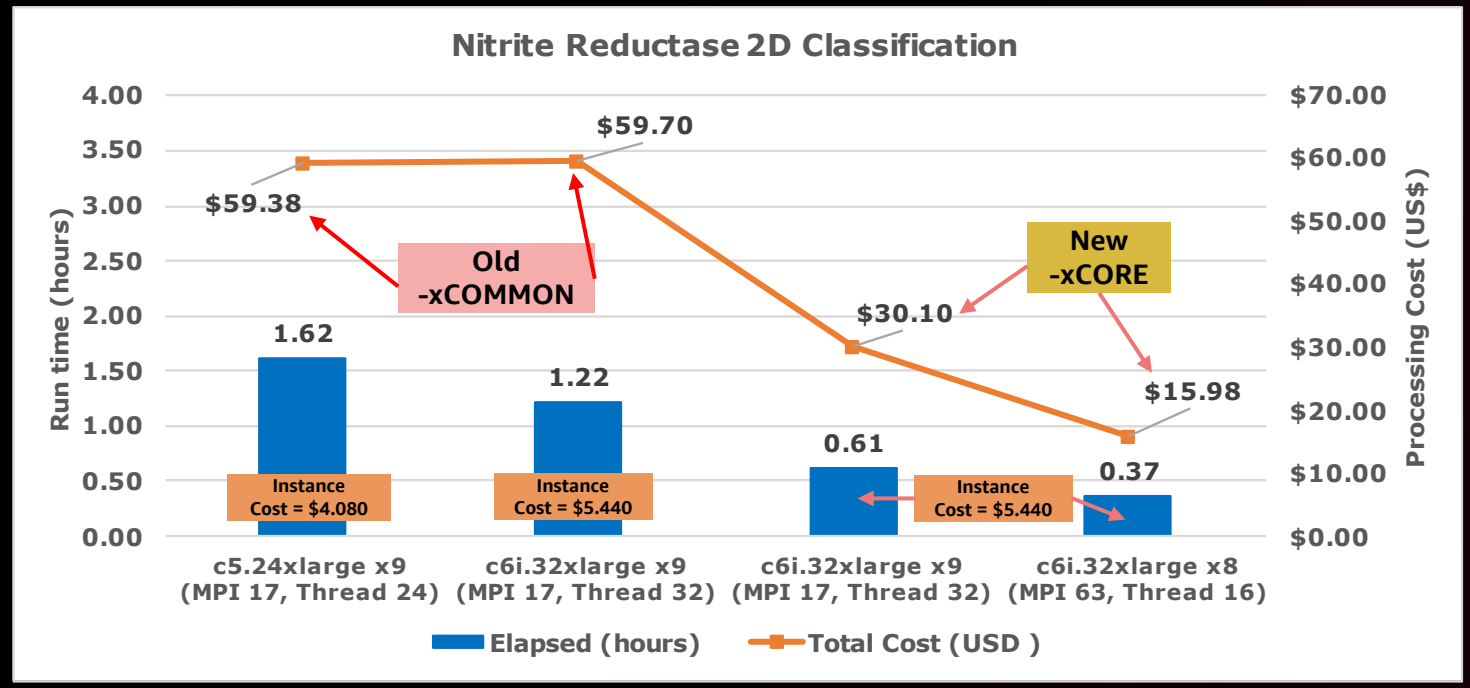
Notices and Disclaimers: Performance varies by use, configuration, and other factors. Learn more at www.intel.com/PerformanceIndex. Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. No product or component can be absolutely secure. Your costs and results may vary. Intel technologies may require enabled hardware, software, or service activation. © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Clovertex modernization made simple

Cryo-EM benchmarks – Life science – RELION on AWS

NITRITE REDUCTASE: 2D CLASSIFICATION | VARYING INSTANCE COUNTS AND COMPILER SETTINGS



KEY POINT:

Using the right compiler flags can shorten the runtime and reduce the processing costs

RESULTS:

1. Using –xCORE instead of –xCOMMON provides a ~3.4x **shorter** runtime and 3.7x **lower** cost
2. Using 8 instances vs. 9 instances with –xCORE is ~1.6x **faster** and a ~1.9x **lower** cost
3. Using more MPI threads and **all** virtual cores vs. less MPI threads and **mostly** physical cores yields the shortest time and lowest cost

Notices and Disclaimers: Performance varies by use, configuration, and other factors. Learn more at www.intel.com/PerformanceIndex. Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. No product or component can be absolutely secure. Your costs and results may vary. Intel technologies may require enabled hardware, software, or service activation. © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

KEK



Sentieon

Brendan Gallagher

Head of Business Development

Sentieon Inc.

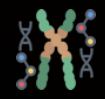
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sentieon: Precision data for precision medicine

>400
customers

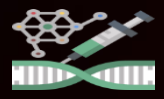
1. Sequencer agnostic → what does this mean?
2. Support all major sequencers (5+)
3. Same great speed and cost performance for each sequencer
4. Custom machine learning to account for each sequencers data type and error profile
5. Robust, high accuracy, germline, and somatic results for widely available Intel CPUs
6. Joint calling on short and long reads
7. On AWS and on premises



NGS service providers



Academics



Bio-pharmaceutical



Clinical diagnostics



Hospitals

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key results

TOP CURRENT SPEEDS

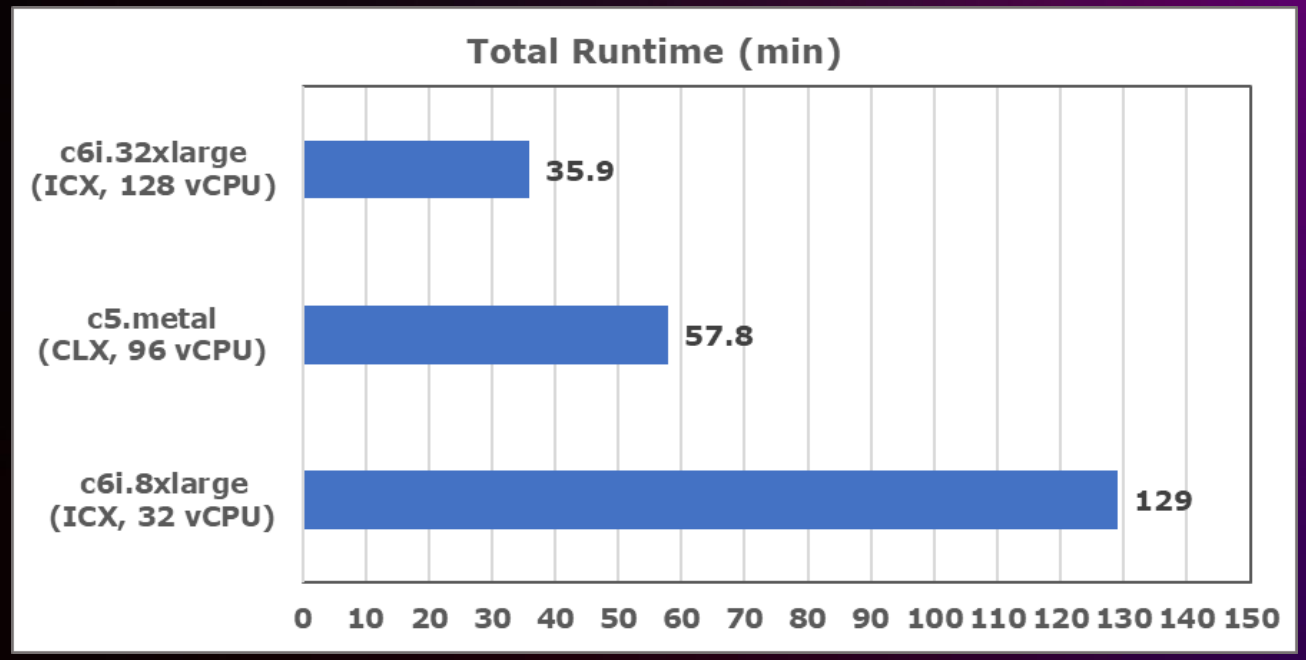
30x WGS from FASTQ → VCF

128-thread (AWS c6i.32xlarge) –

64 vCPU hours – 31 minutes

On-Demand cost – ~$3

Spot cost – $0.80



aws | intel

Notices and Disclaimers: Performance varies by use, configuration, and other factors. Learn more at www.intel.com/PerformanceIndex. Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. No product or component can be absolutely secure. Your costs and results may vary. Intel technologies may require enabled hardware, software, or service activation. © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

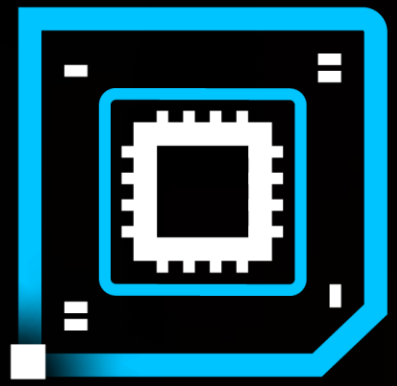
Key takeaways

1. CPUs continue to get better – just as performant as GPUs and FPGAs
2. Widely available on AWS, high Spot availability
3. Sentieon is sequencer-agnostic
4. Support all major sequencers (5+)
5. Same great speed and cost performance for each sequencer
6. Robust, high accuracy, germline, and somatic results for widely available Intel CPUs



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Conclusion



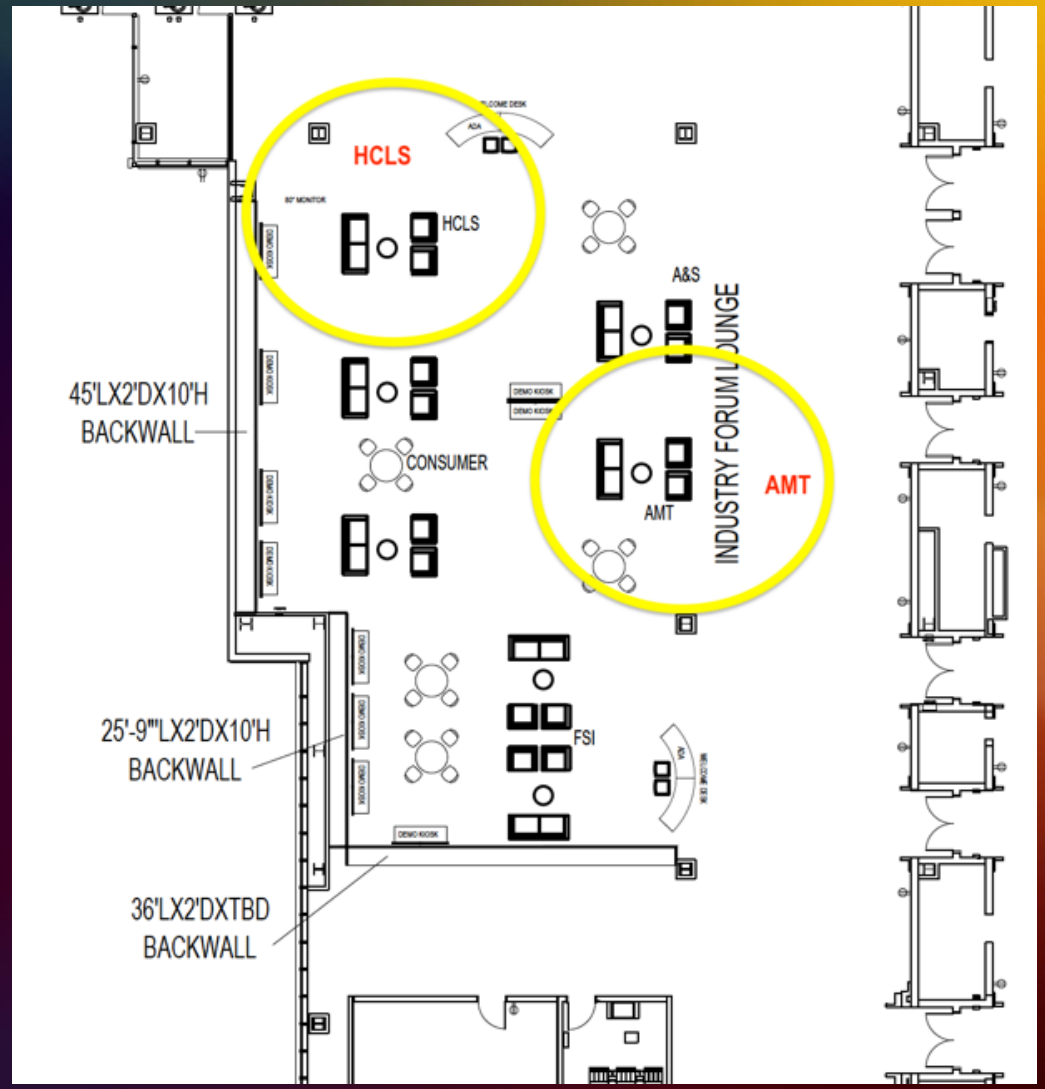
CPUs provide a performant and cost-effective experience across HCLS workloads, underpinning technologies that allow faster drug discovery and medical advancements

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Learn more at the HCLS industry kiosk takeover

Caesars Forum, Level 1,
AWS Industries Networking Lounge,
HCLS Industry Kiosk, 2:00–3:00 p.m.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.