Post-Quantum

Cryptography Conference

Accelerated Quantum Supercomputing and Post-Quantum Cryptography

Accelerated computing is revolutionizing numerous fields, including quantum computing (QC) and artificial intelligence (AI), and is also set to accelerate the development of robust post-quantum cryptographic solutions. This talk explores how cutting-edge AI techniques are addressing challenges within QC across the hardware and software stack to create more optimized circuits, bringing practical quantum computers one step closer. Additionally, this talk will cover how GPU-based acceleration serves as a safeguard against emerging quantum cryptographic threats. We will reveal how GPU-based algorithms are accelerating cryptographic research by examining technical challenges in parallelizing cryptographic workloads across GPUs, managing memory bandwidth, optimizing performance, and overcoming hardware limitations. We will also highlight how these technologies are accelerating QC research. Real-world applications in sectors such as finance, healthcare, and data privacy will be showcased, demonstrating the practical benefits of AI, QC, and PQC.

Yarkin Doroz
Product Manager at NVIDIA

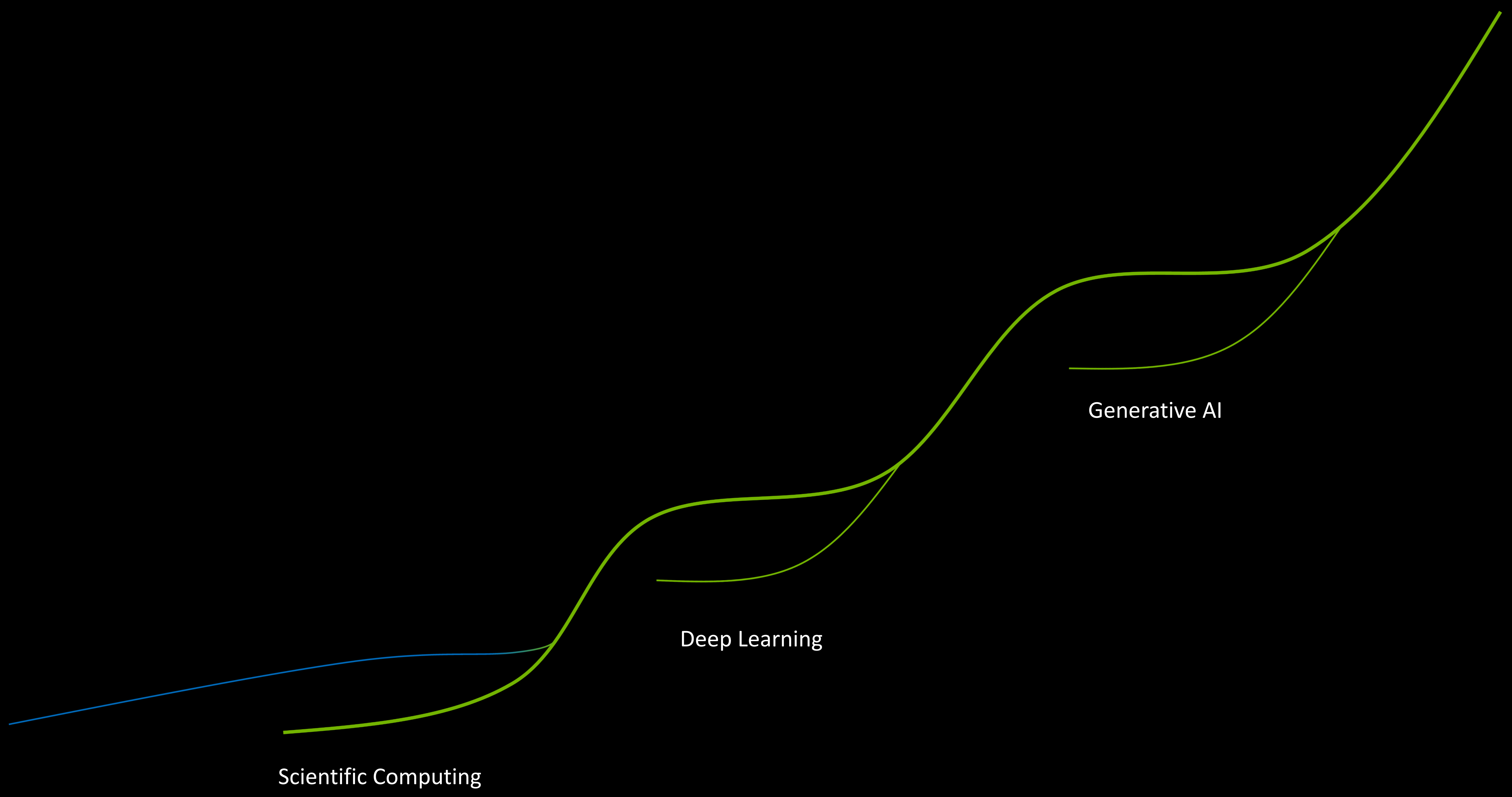
January 15 and 16, 2025 - Austin, TX (US) | Online

Accelerated Quantum Supercomputing and Post-Quantum Cryptography

Post-Quantum Cryptography Conference PKI Consortium

Yarkin Doroz – Product Manager

Computing Revolutions



Computing Revolutions

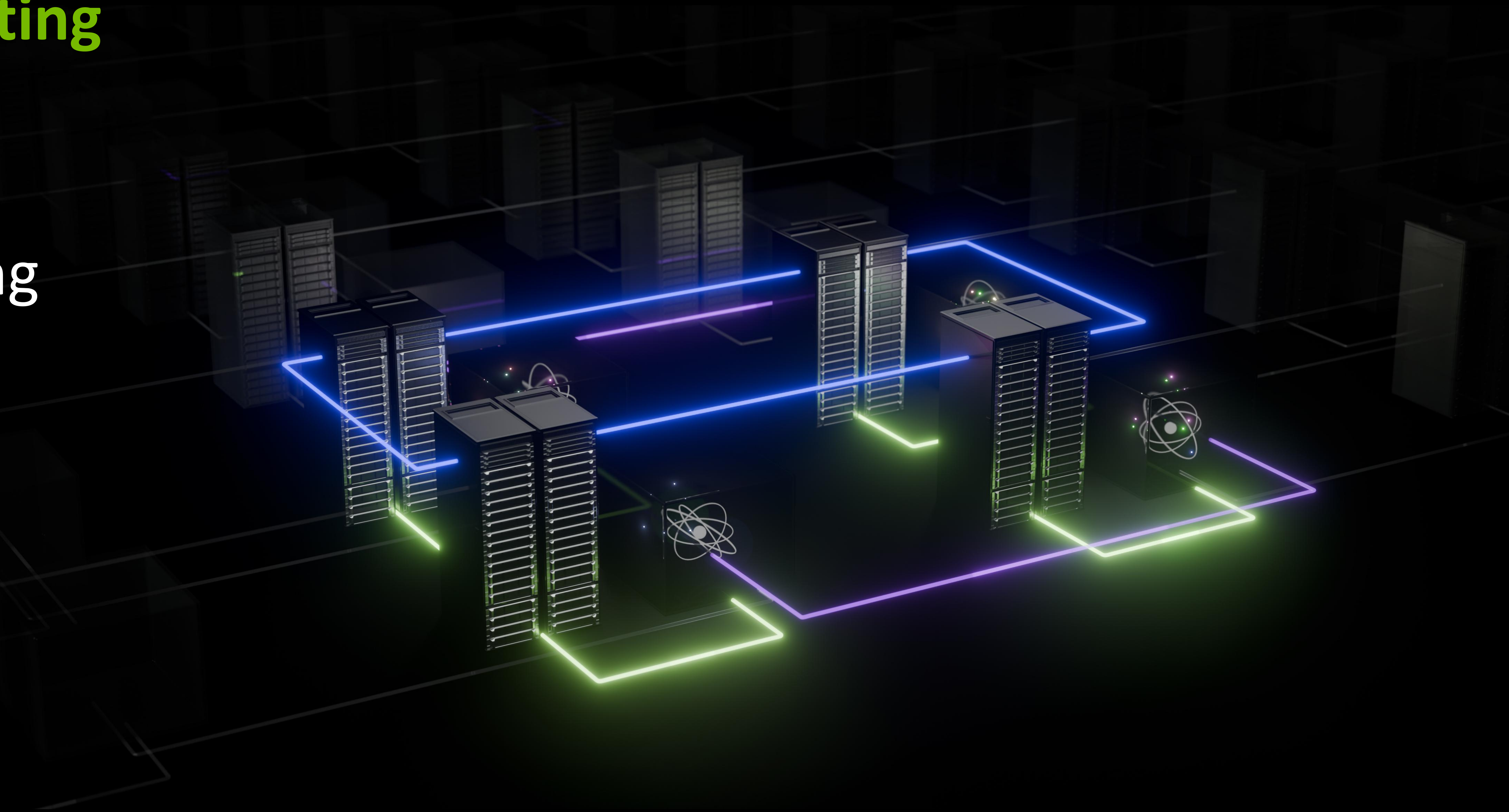
NVIDIA is not building

Qubits

NVIDIA is building all
Accelerated Quantum Supercomputers

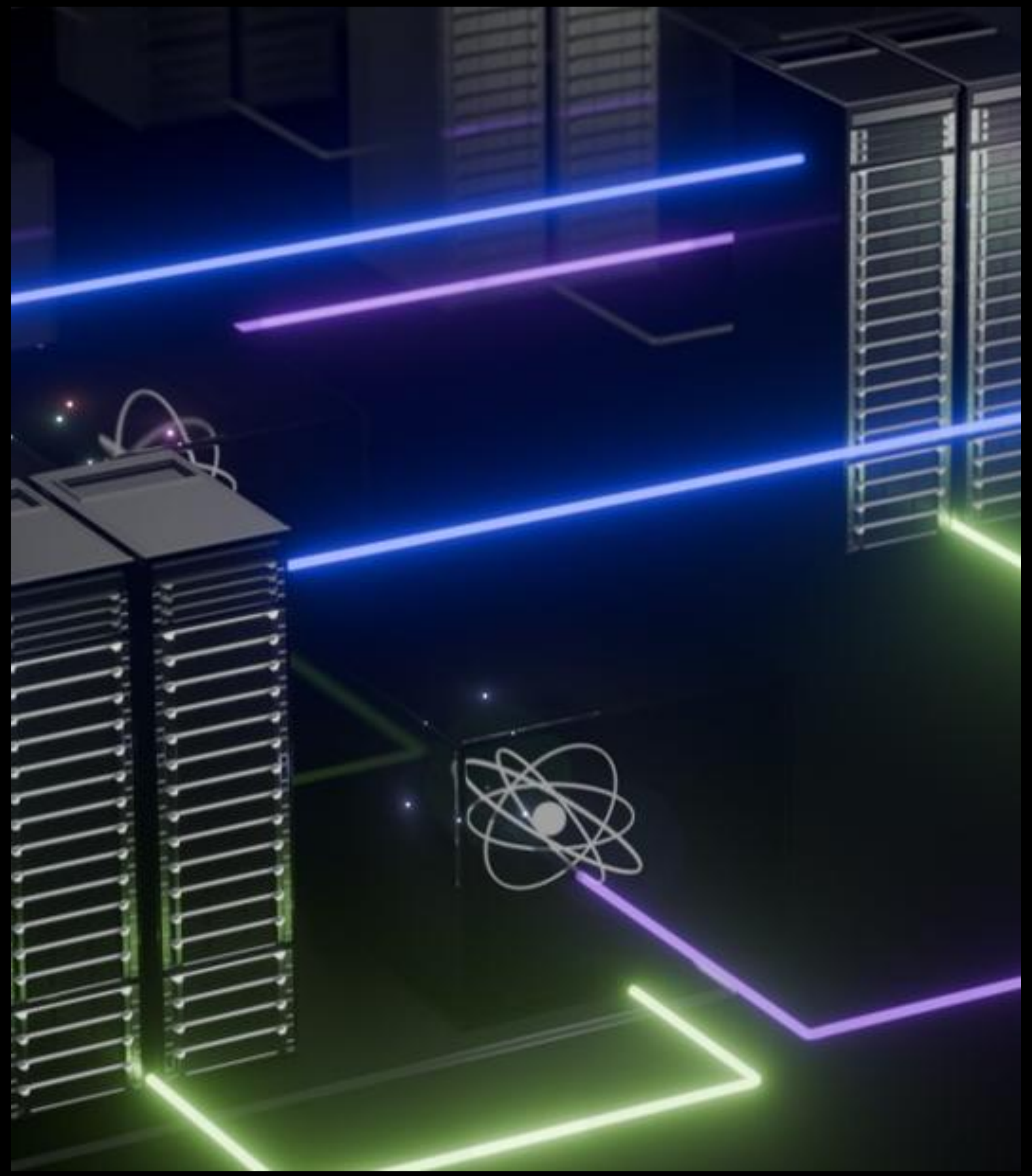
The Accelerated Quantum Supercomputer

- Supercomputing architecture **connecting quantum hardware**
- Ability to run **hybrid algorithms** - using GPUs and QPUs
- A **software platform** that seamlessly connects hybrid applications
- The ability to perform **qubit-agnostic** development of control and error correction



Quantum Computing Needs Accelerated Computing

AI SC for QC Deployments



Quantum Error Correction

Hybrid algorithms and applications

AI for
- Calibration
- Control
- Readout

AI SC for QC Development

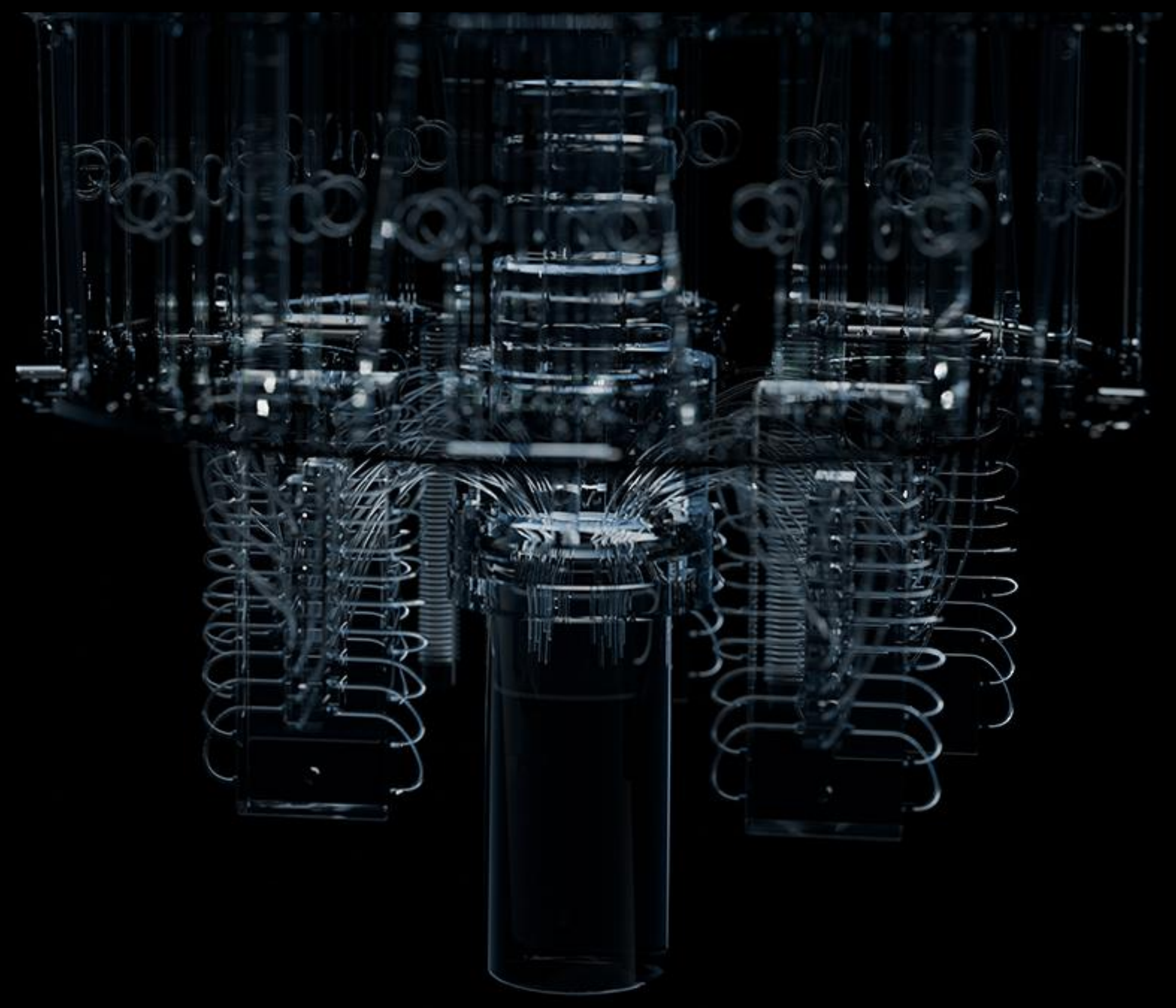
Accelerated application development

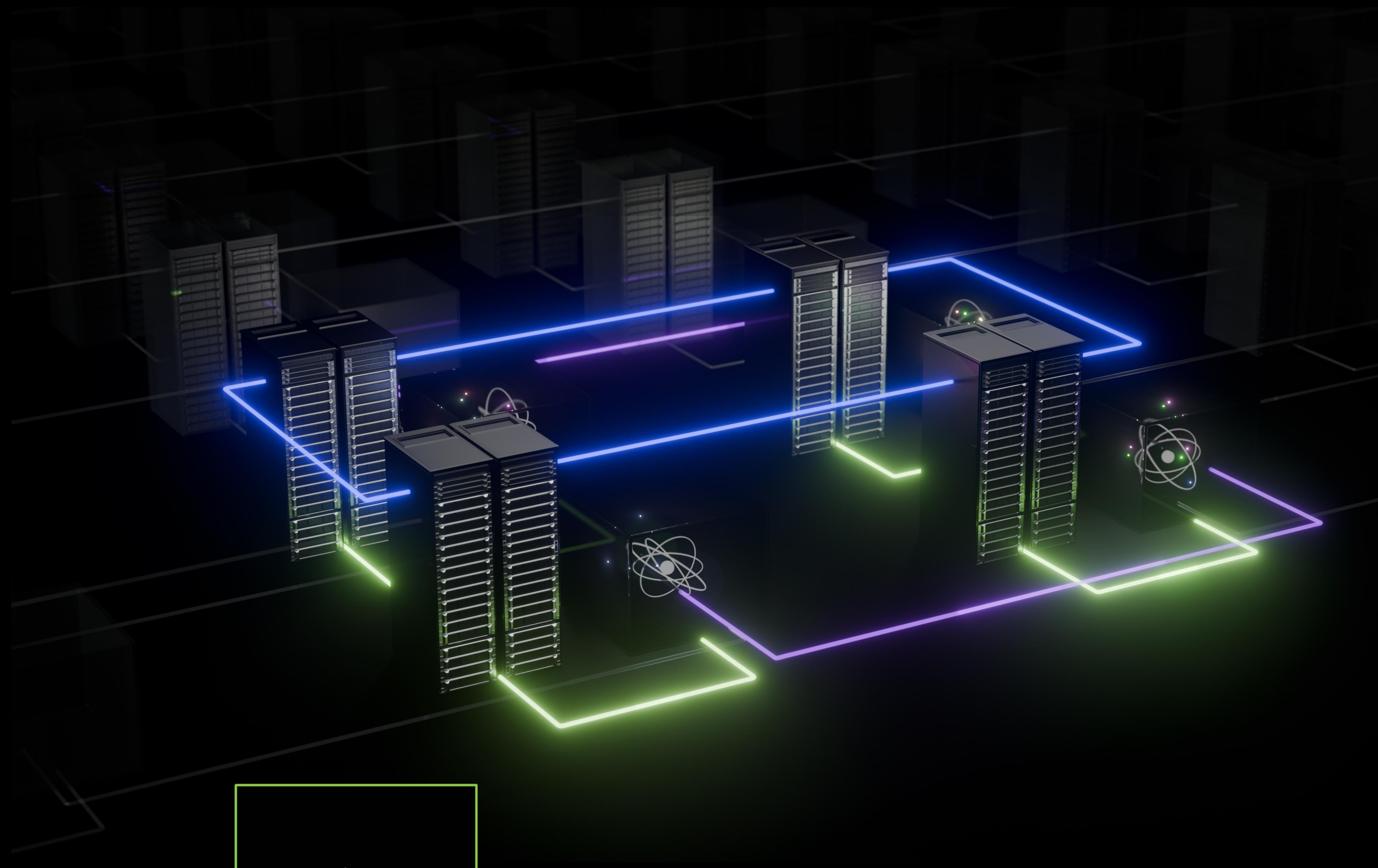
AI assisted circuit design

Dynamical simulations

Noise modeling

Accelerating the Journey From Qubits to Supercomputers





Qubits

Better quantum hardware

GPU

Hybrid algorithms

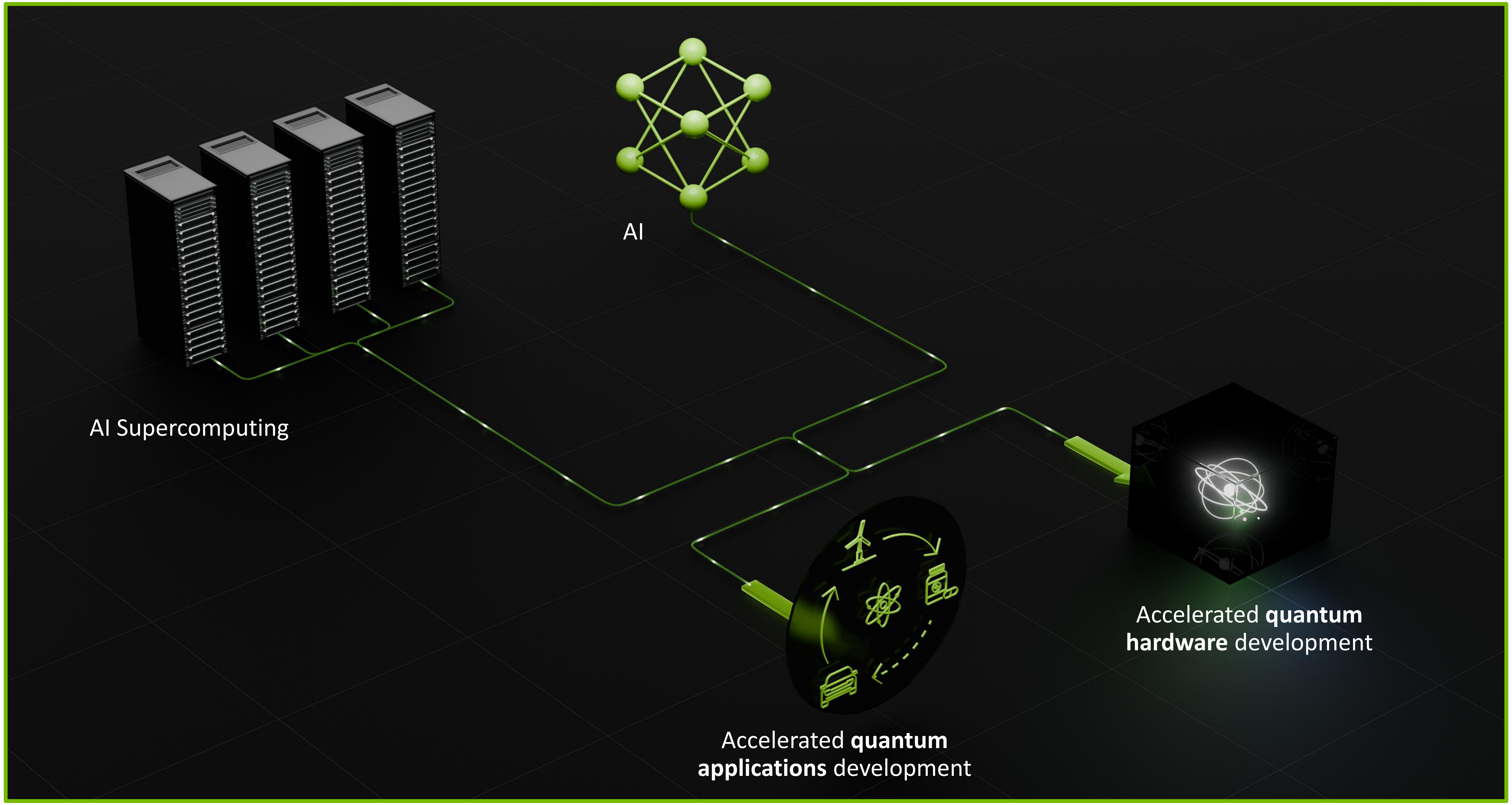
Implementing QEC

Infrastructure challenges

Accelerated Quantum Supercomputers

AI supercomputing

Bringing AI to quantum computing



Artificial Intelligence for Quantum Computing

Yuri Alexeev[†1], Marwa H. Farag[†1], Taylor L. Patti[†1], Mark E. Wolf[†1*], Natalia Ares[2], Alán Aspuru-Guzik[3,4], Simon C. Benjamin[5,6], Zhenyu Cai[5,6], Zohim Chandani[1], Federico Fedele[2], Nicholas Harrigan[1], Jin-Sung Kim[1], Elica Kyoseva[1], Justin G. Lietz[1], Tom Lubowe[1], Alexander McCaskey[1], Roger G. Melko[7,8], Kouhei Nakaji[1], Alberto Peruzzo[9], Sam Stanwyck[1], Norm M. Tubman[10], Hanrui Wang[11] and Timothy Costa[1]

[1]NVIDIA Corporation, 2788 San Tomas Expressway, Santa Clara, 95051, CA, USA.

[2]Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom.

[3]Department of Chemistry, Computer Science, Materials Science and Engineering, and Chemical Engineering and Applied Science,University of Toronto, 80 St George St, Toronto, M5S 3H6, ON, Canada.

[4]Vector Institute for Artificial Intelligence, 661 University Ave Suite 710, Toronto, M5G 1M1, ON, Canada.

[5]Quantum Motion ,* 9 Sterling Way, London, N7 9HJ, United Kingdom.

[6]Department of Materials, University of Oxford, Parks Road, Oxford, OX1 3PH, United Kingdom.

[7]Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West., Waterloo, N2L 3G1, ON, Canada.

[8]Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, N2L 2Y5, ON, Canada.

[9]Qubit Pharmaceuticals, 29, rue du Faubourg Saint Jacques, Paris, 75014, France.

[10]NASA Ames Research Center, Moffett Field, California, 94035-1000, USA.

*†These authors contributed equally to this work.

AI for QC

- System characterization
- Platform design
- Gate and Pulse optimization

**Section 2:
QC Hardware
Development
and Design**

AI for QC

- System characterization
- Platform design
- Gate and Pulse optimization

Section 2: QC Hardware Development and Design

Section 3: Preprocessing

- Quantum circuit synthesis
- Circuit parameter learning
- State preparation

AI for QC

- System characterization
- Platform design
- Gate and Pulse optimization

- Device tuning
- Qubit control
- Device characterization

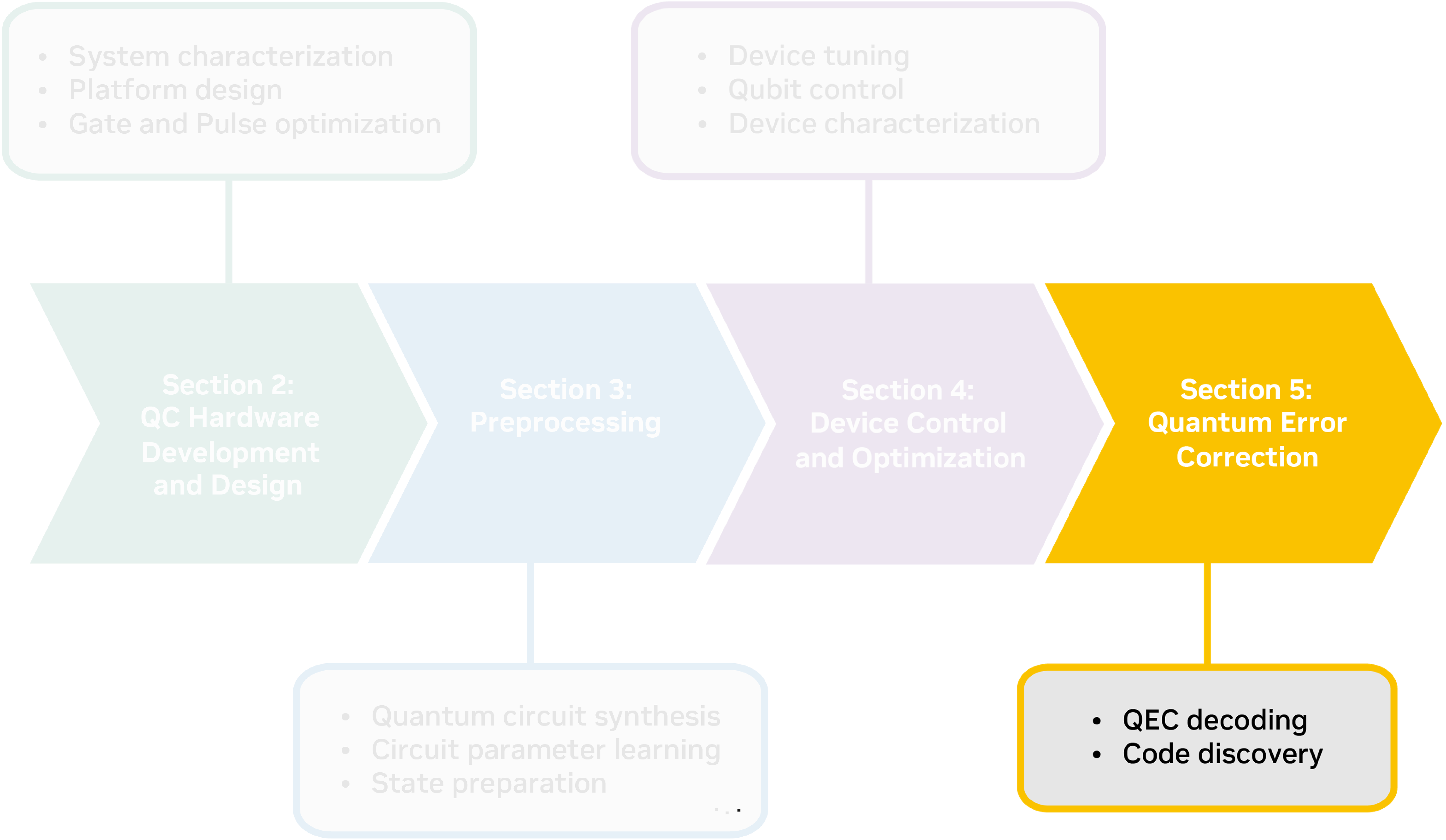
Section 2:
QC Hardware Development and Design

Section 3:
Preprocessing

Section 4:
Device Control and Optimization

- Quantum circuit synthesis
- Circuit parameter learning
- State preparation

AI for QC



System characterization
Platform design
Gate and Pulse optimization
Device tuning
Qubit control
Device characterization
Section 2: QC Hardware Development and Design
Section 3: Preprocessing
Section 4: Device Control and Optimization
Section 5: Quantum Error Correction
Quantum circuit synthesis
Circuit parameter learning
State preparation
QEC decoding
Code discovery

AI for QC

System characterization
Platform design
Gate and Pulse optimization
Section 2: QC Hardware Development and Design
Section 3: Preprocessing
Quantum circuit synthesis
Circuit parameter learning
State preparation
Device tuning
Qubit control
Device characterization
Section 4: Device Control and Optimization
Section 5: Quantum Error Correction
QEC decoding
Code discovery
Observable estimation and tomography
Qubit Readout
Error mitigation
Section 6: Postprocessing

AI for QC

System characterization
Platform design
Gate and Pulse optimization
Device tuning
Qubit control
Device characterization
Observable estimation and tomography
Qubit Readout
Error mitigation
Section 2: QC Hardware Development and Design
Section 3: Preprocessing
Section 4: Device Control and Optimization
Section 5: Quantum Error Correction
Section 6: Postprocessing
Quantum circuit synthesis
Circuit parameter learning
State preparation
QEC decoding
Code discovery

The Generative Quantum Eigensolver

First demonstration of GPT-generated circuits

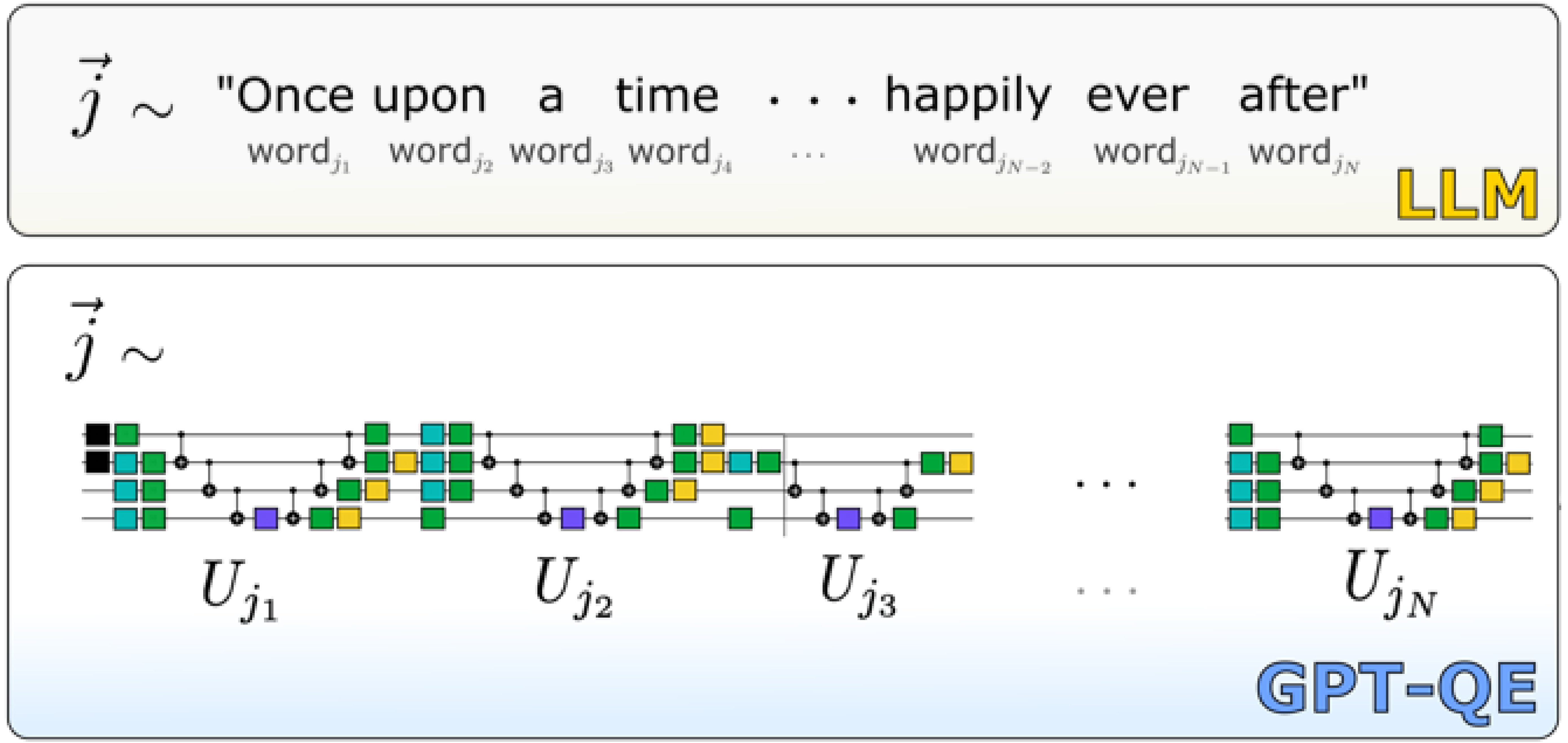
CUDA-Q + St. Jude Children's Research Hospital + University of Toronto

Challenge

- Variational quantum algorithms offered promise for running drug-discovery applications on small quantum devices - but **suffer from serious optimization issues**
- Many of these problems are tied to how circuits are parametrized.

Solution

- The generative quantum eigensolver acts like a Large Language Model – but generating quantum circuits from quantum operations, rather than sentences from words.
- Using a generative model like GPT to create quantum circuits **avoids the limitations of traditional variational quantum algorithms**



New approach in Using AI for building quantum applications

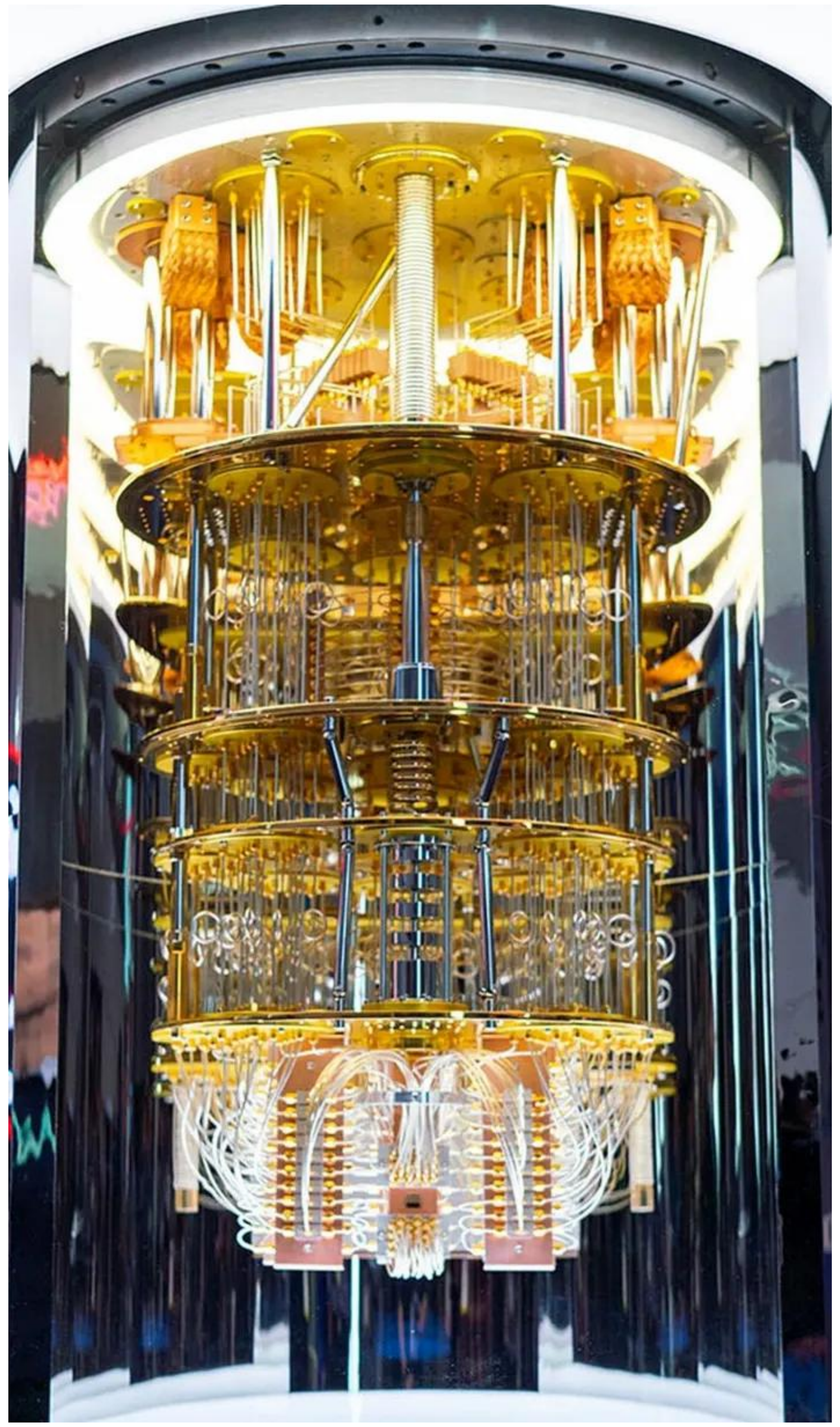
Can be extended to various application areas

40X

Speedup over CPU when running GQE on GPU

Quantum Computers and Quantum Algorithms

- Quantum Computers compute using qubits
- Quantum gates operate on qubits to run Quantum Algorithms
- Phenomena like superposition and entanglement allow qubits to accelerate some computations



Two main Quantum Algorithms Related to Cryptography:

- **Grover's Algorithm**
  - Quantum Search Algorithm
  - Search for an input in an unstructured database
  - Reduces complexity from $O(N) \to O(\sqrt{N})$

- **Shor's Algorithm**
  - Quantum Algorithm for Integer Factorization
  - Factorize a given integer into its prime factors
  - Reduces complexity from $O(N) \to O(\log_2 N)$

NIST PQC Standardization Timeline

Urgency for New Cryptographic Solutions

NIST PK-PQC Call
(Dec 2016)

Round 1 Submissions
KEM: 59
DSA: 23
(Nov 2017)

Round 2
KEMS: 17
DSA: 9
(Jan 2019)

Round 3
Main Candidates
KEM: 4
DSA: 3
Alternatives
KEM: 5
DSA: 3
(July 2020)

Finalists
KEM: 1
DSA: 3
(July 2022)

Round 4
KEM: 4
(July 2022)

Additional DSA – Round 1
DSA: 40
(June 2023)

Additional DSA – Round 2
DSA: 14
(Oct 2024)

NIST PQC Transition Timeline

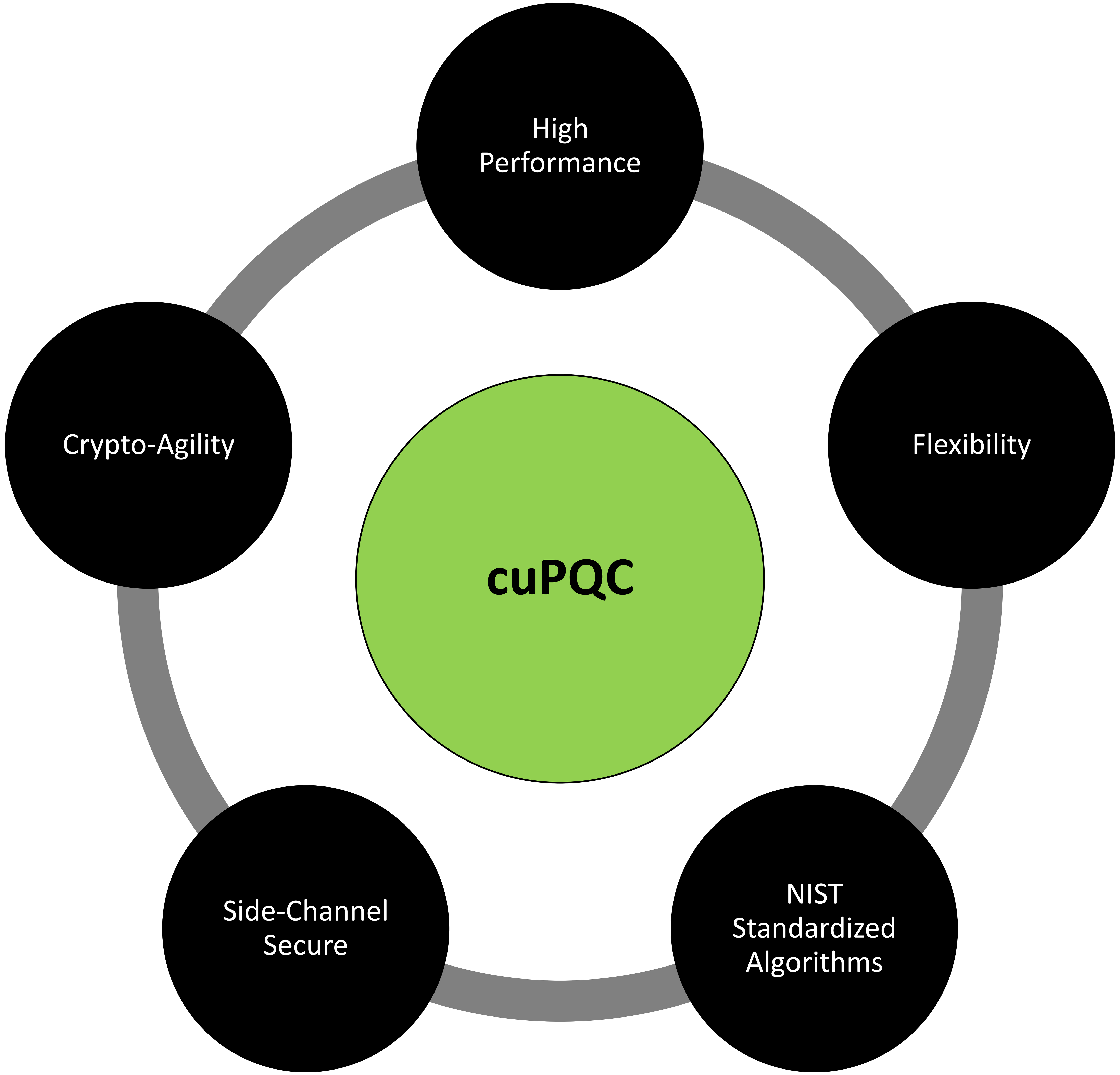
Algorithms	Parameters	Depreciate	Disallowed
ECDSA, EdDSA, RSA	112	2030	2035
ECDSA, EdDSA, RSA	128	-	2035
Finite Field/Elliptic Curve DH and MVQ, RSA	112	2030	2035
Finite Field/Elliptic Curve DH and MVQ, RSA	128	-	2035

Critical Applications Demanding High-Throughput Post-Quantum Cryptography

Key Sectors	Protocols
Cloud Service Providers	
Content Delivery Networks (CDNs)	
Financial Institutions	**TLS (Transport Layer Security)**
Telecommunications Companies	**IPsec (Internet Protocol Security)**
Military and Government Applications	**SSL (Secure Sockets Layer)**
Blockchain Companies	**SSH (Secure Shell)**
Healthcare Systems	**HTTPS (Hypertext Transfer Protocol Secure)**
Big Data Analytics Companies	**GSM/3G/4G/5G Encryption Protocols**
Autonomous Systems	**SNMPv3 (Simple Network Management Protocol version 3)**
Network Equipment Providers	
Internet of Things (IoT) Companies	

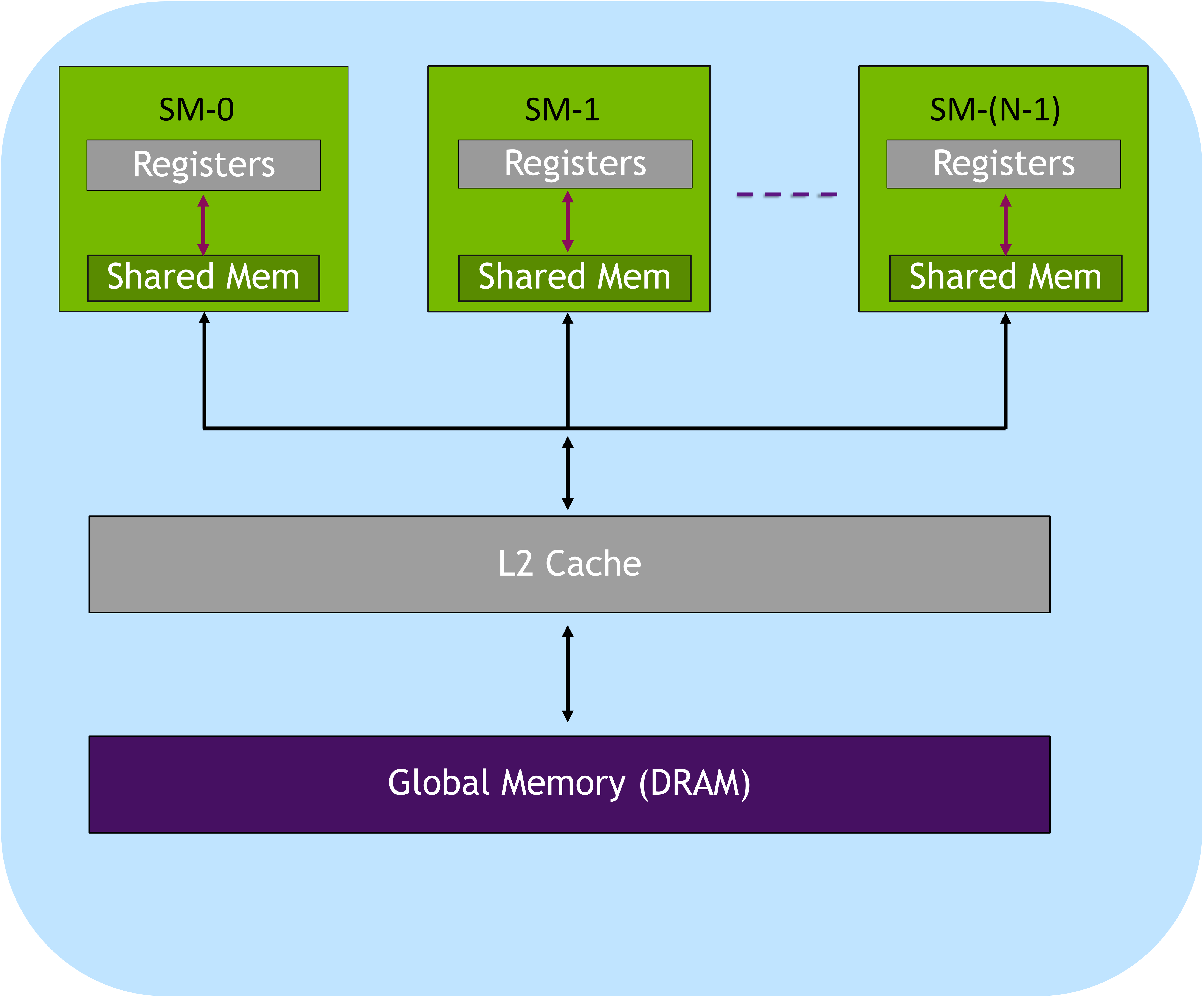
GPU based PQC Library Considerations

Our considerations when creating cuPQC



GPU Hardware Optimizations

- Certain mathematical subroutines are suitable for parallelization
  - Number Theoretic Transform
- Shared Memory used for collaboration of cores and accelerate mathematical subroutines
  - This space allows threads to work together to solve a larger problem.
- Optimization techniques to allow register-heavy components to be spread across multiple threads.
- Limiting data transfers between the CPU and GPU
  - CPU main memory transfers over PCIE take up valuable clock cycles.
- ML-KEM and ML-DSA are register-use heavy and can limit utilization.
  - Hashing algorithms utilizing Keccak, random sampling, and others.
  - Algorithms need to be reorganized with hardware considerations.

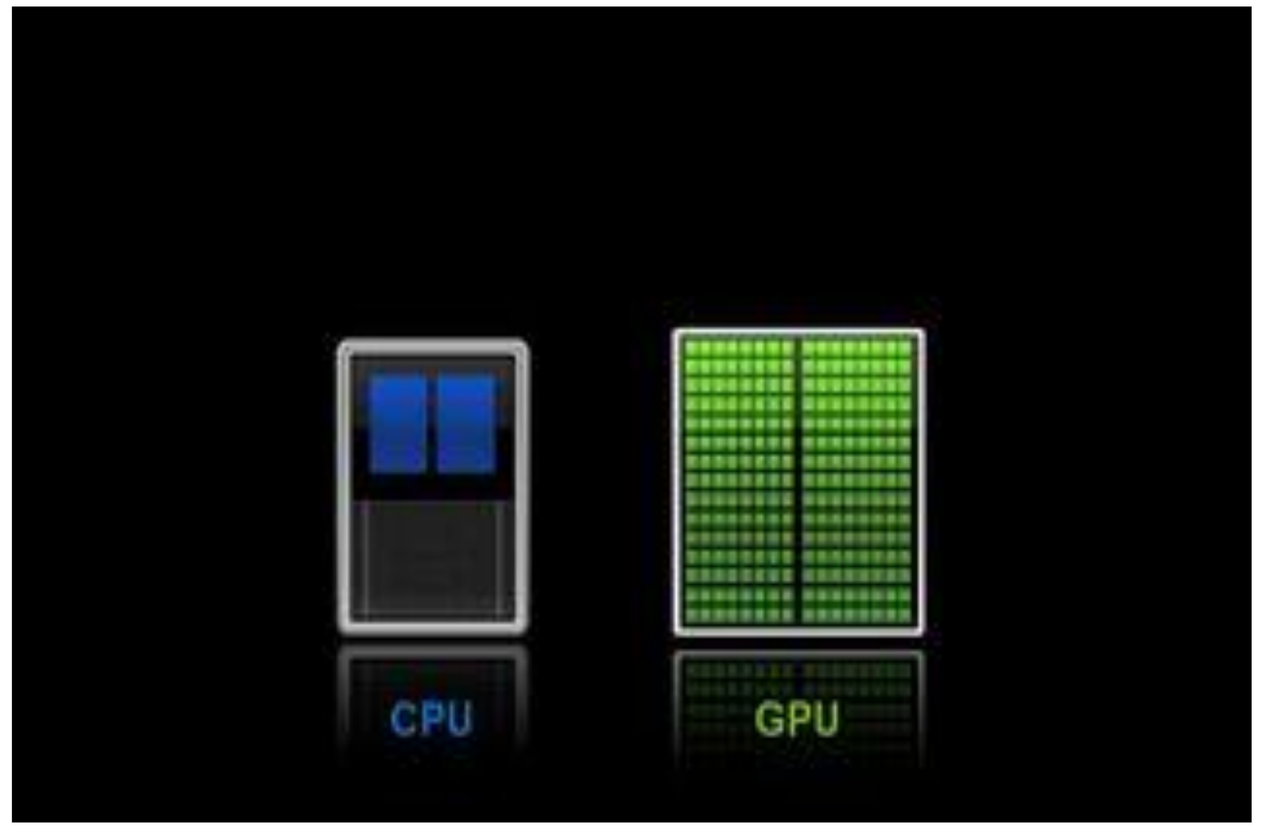


GPU Parallelism: Natural Noise Enhances Security Against Side-Channels

The thousands of processing cores in GPUs → Massive parallelism in arithmetic, cryptographic operations → Thousands of simultaneous

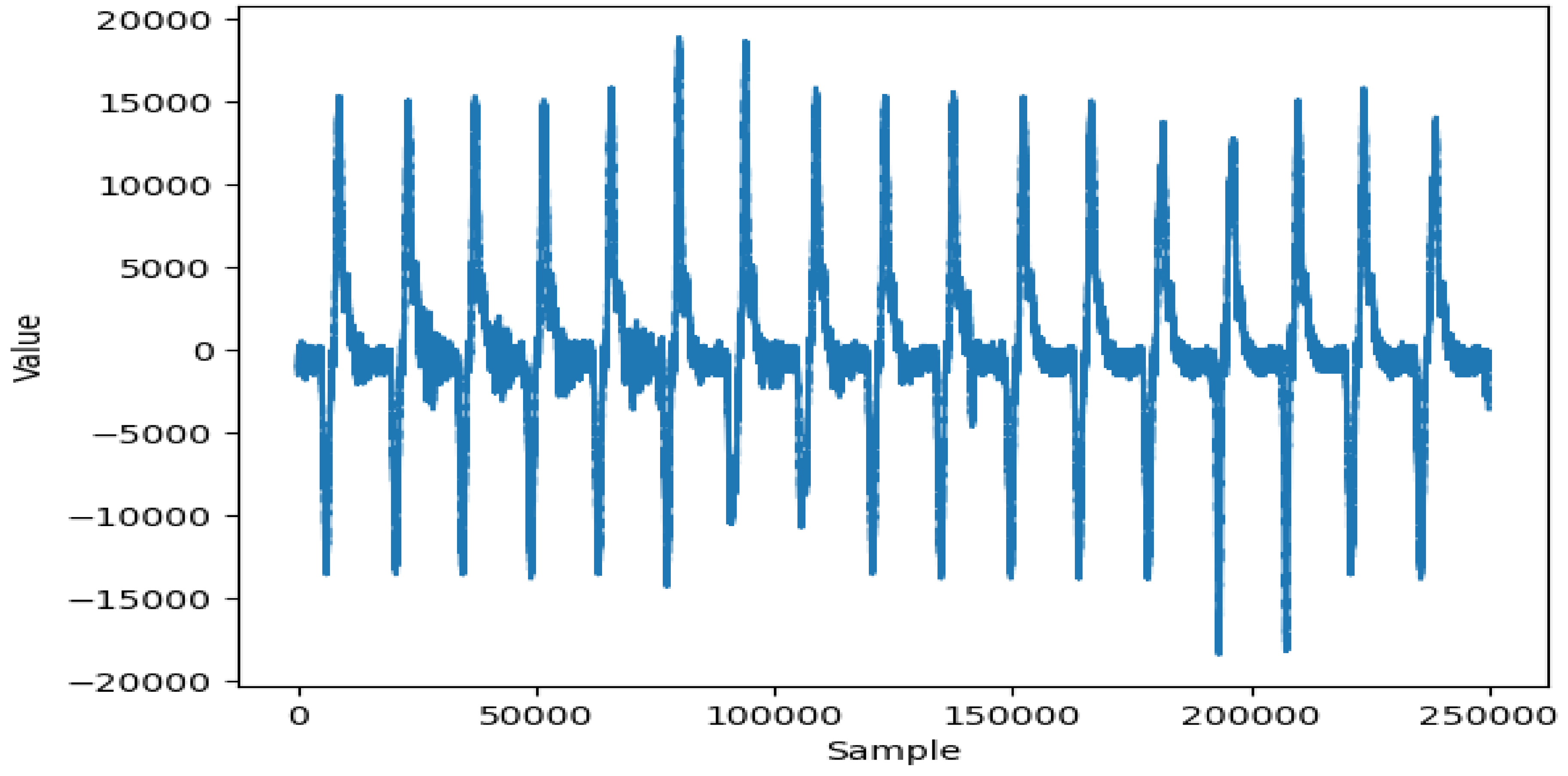
- Cache&Memory Accesses
- Arithmetic Operations
- Power Spikes

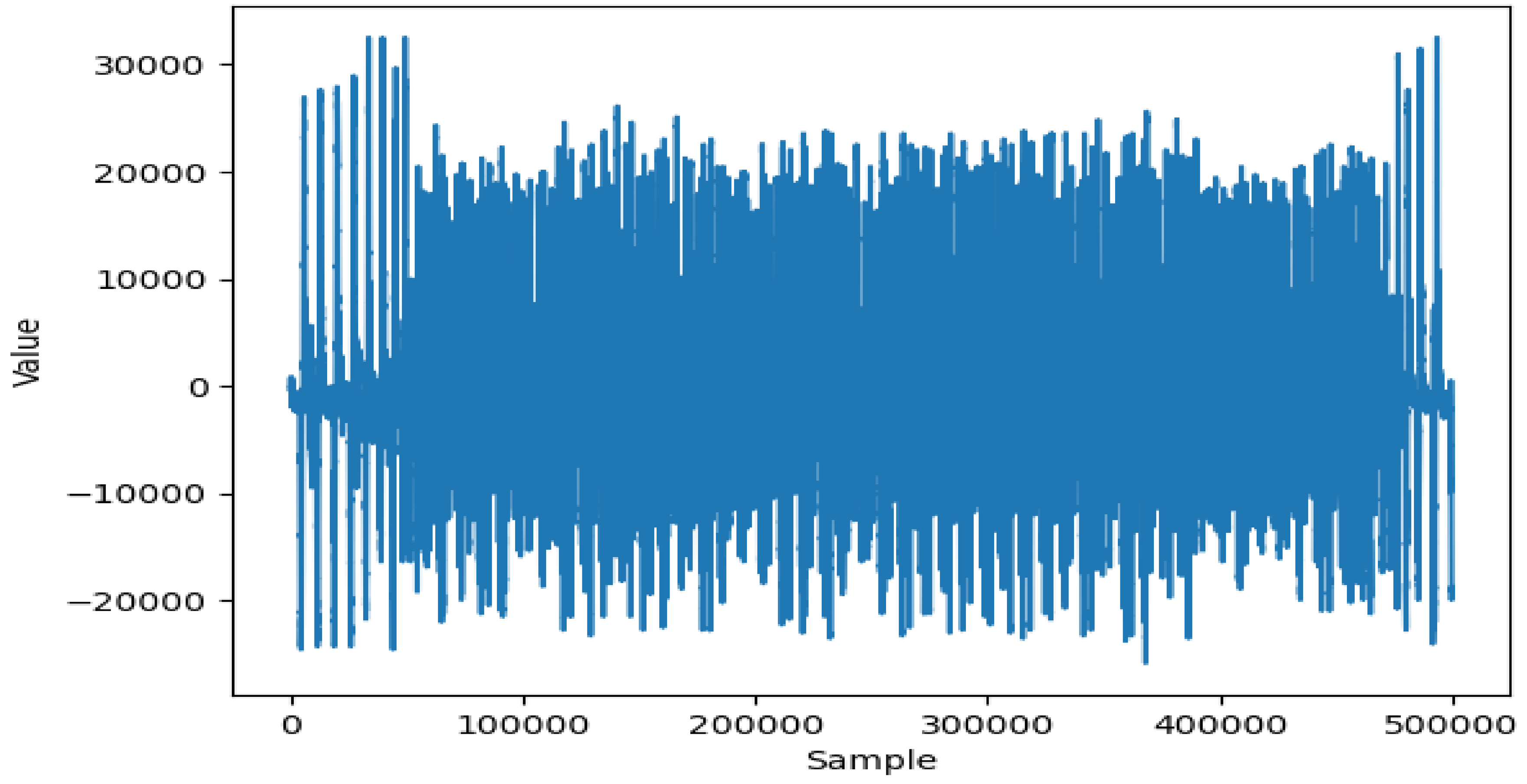
→ Causes garbled power, EM, memory traces → More resistant to side-channel attacks



Only a small subset of known attacks apply to GPU threat model

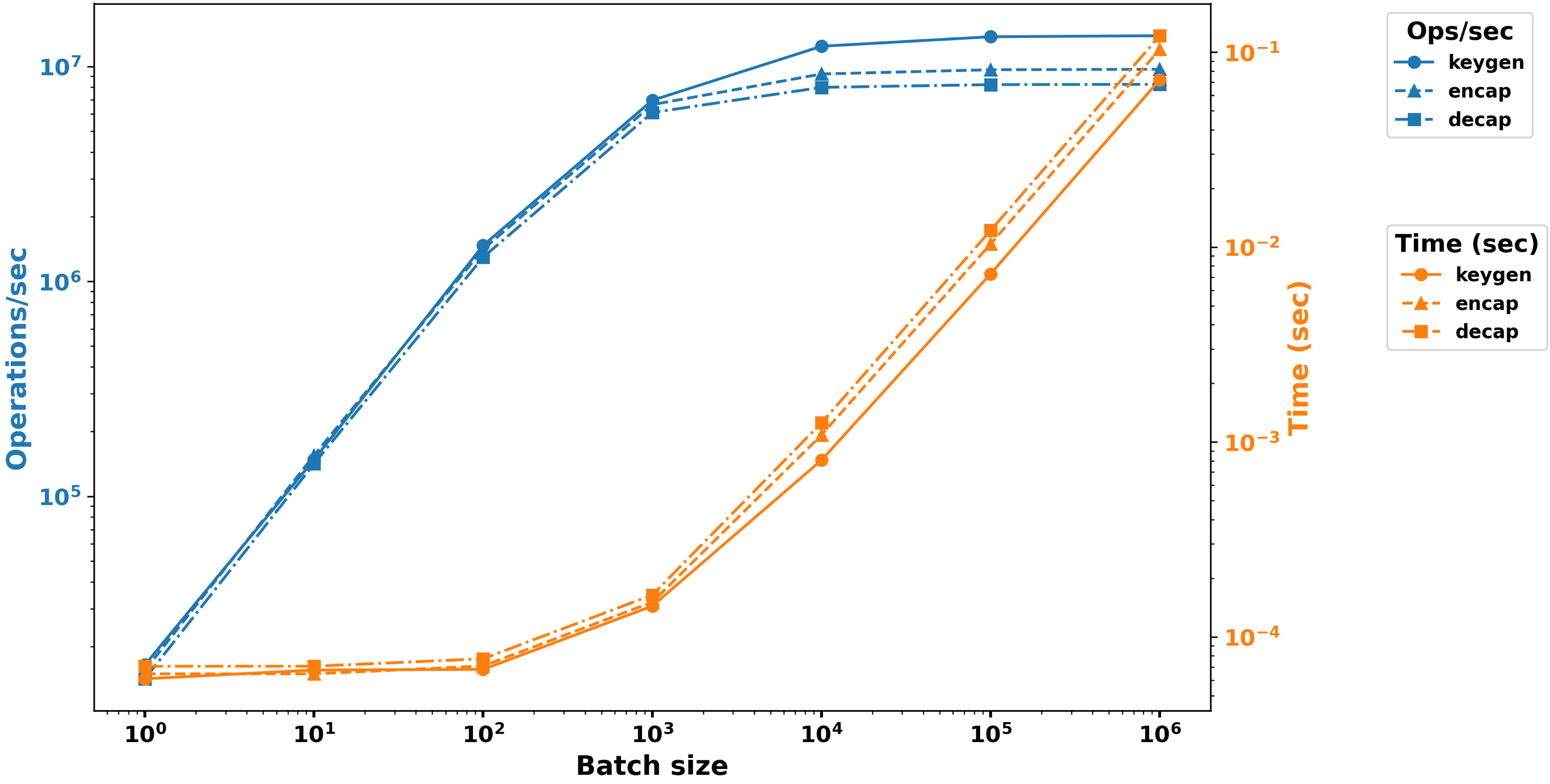
Countermeasures:
randomizing the algorithm's behavior
masking the sensitive data





ML-KEM Benchmarks

ML-KEM-768



ML-KEM
(Batched in 1,000,000)

Keygen

Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-KEM-512	0.049	20.36
ML-KEM-768	0.072	13.88
ML-KEM-1024	0.088	11.35

Encapsulation

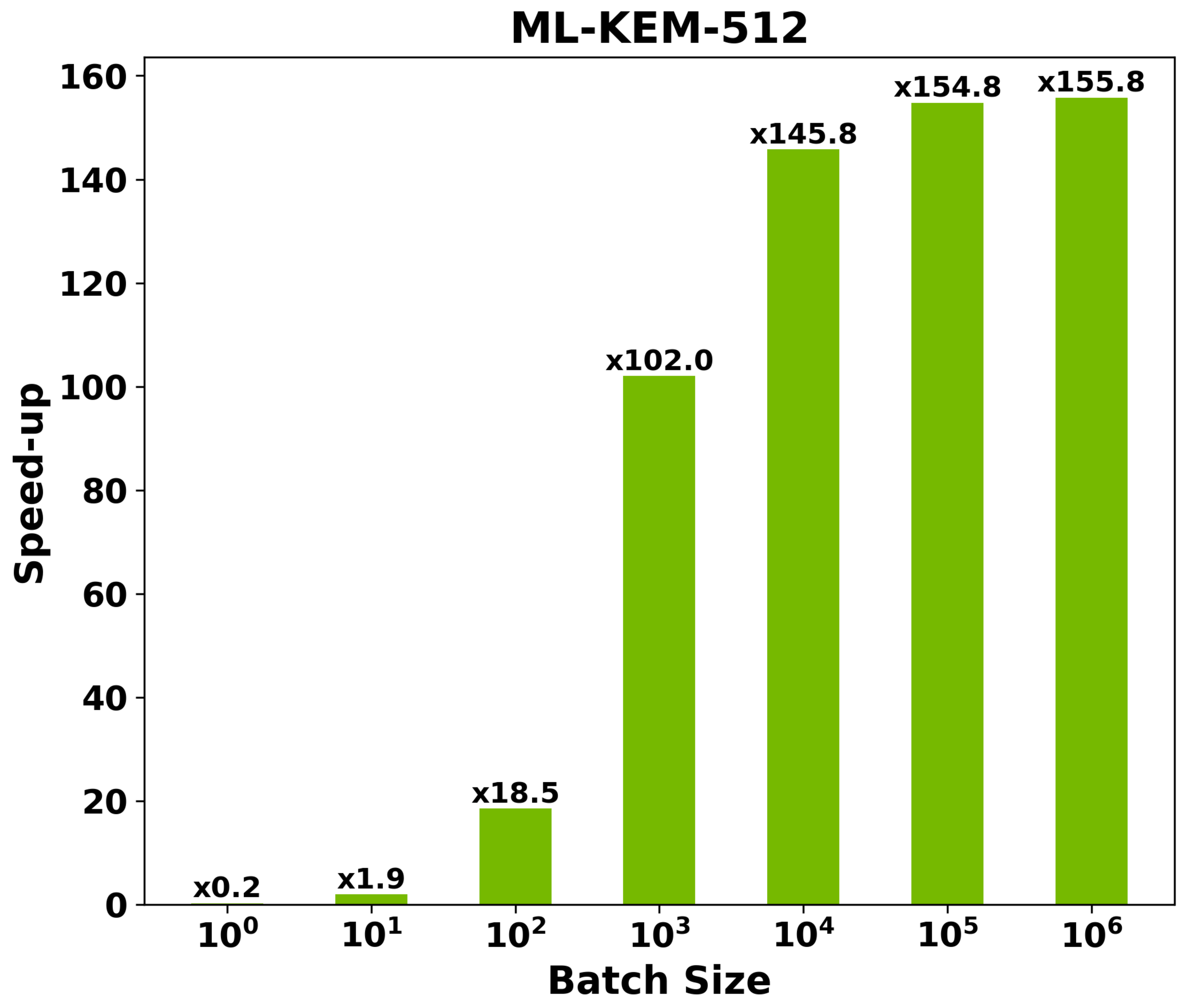
Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-KEM-512	0.051	19.35
ML-KEM-768	0.103	9.69
ML-KEM-1024	0.125	7.95

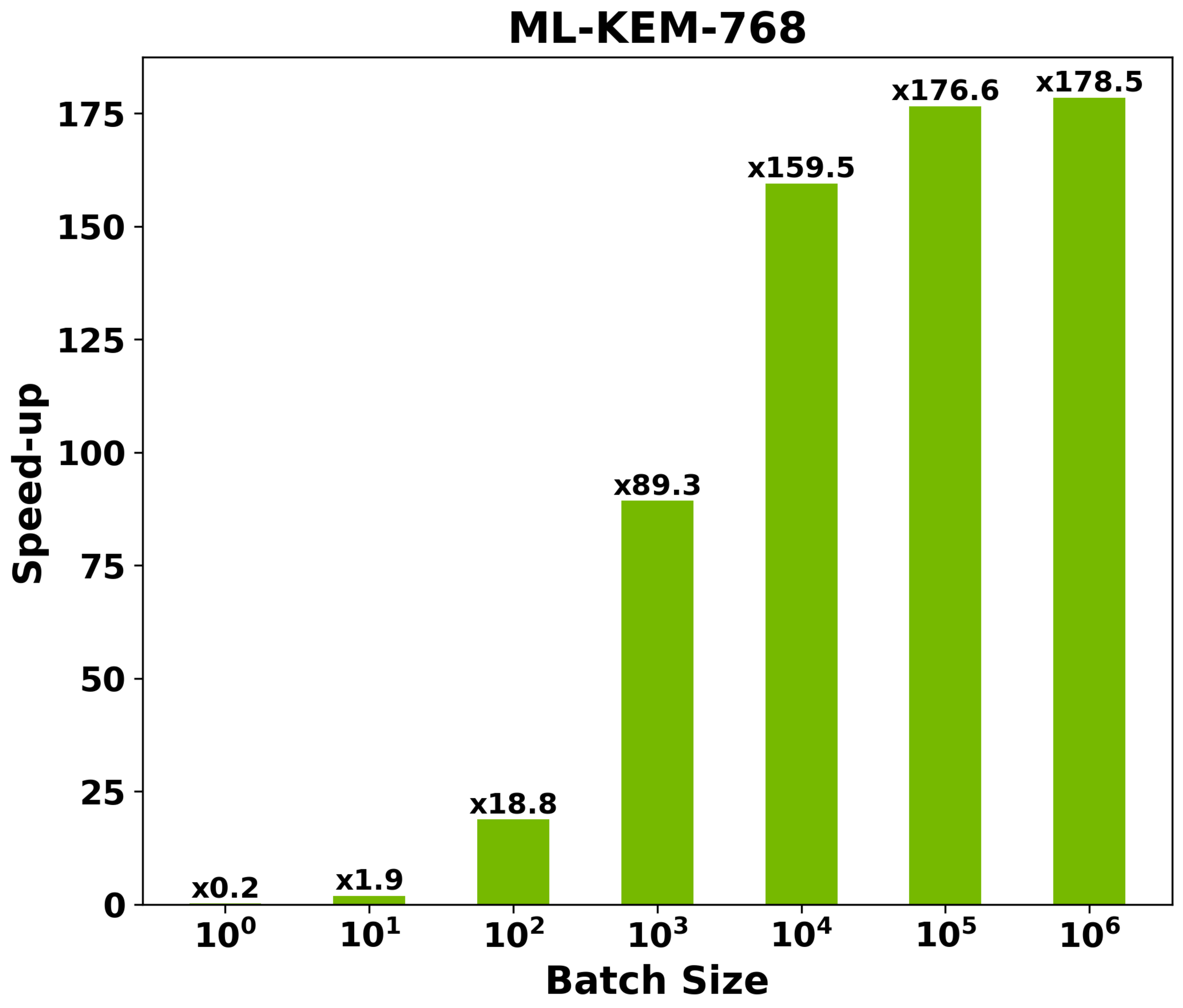
Decapsulation

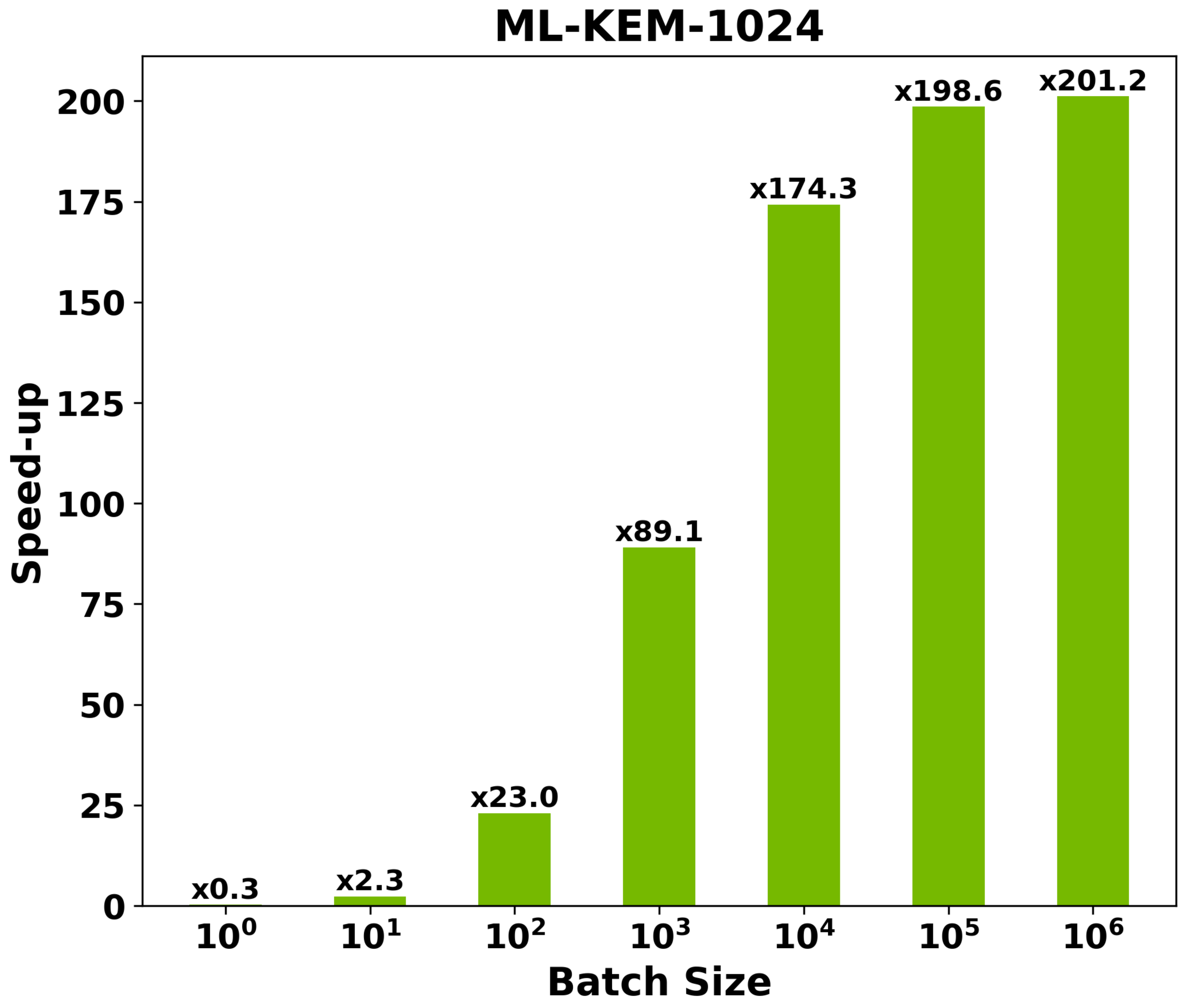
Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-KEM-512	0.055	18.21
ML-KEM-768	0.121	8.25
ML-KEM-1024	0.133	7.47

ML-KEM Keygen Benchmarks

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5



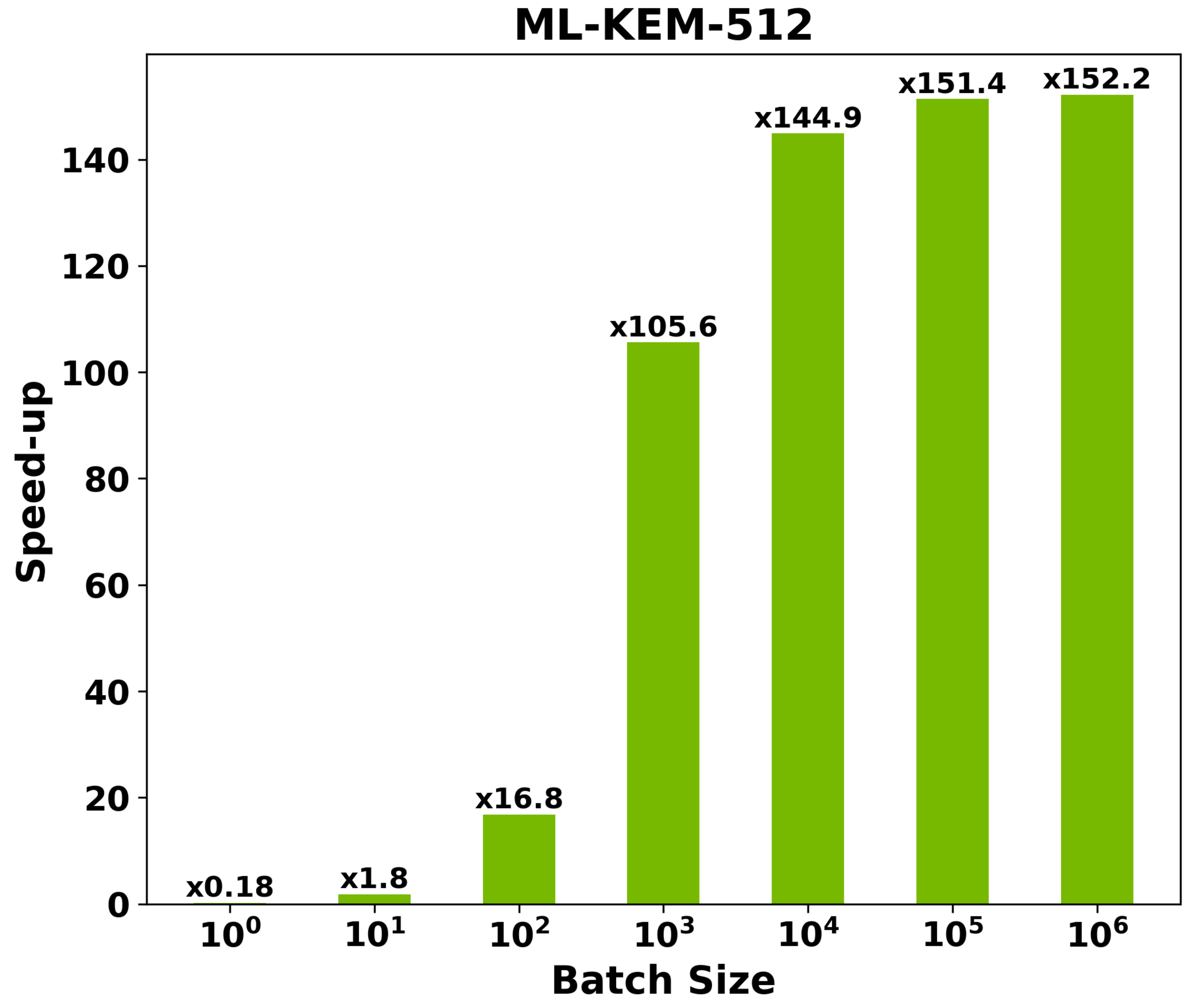


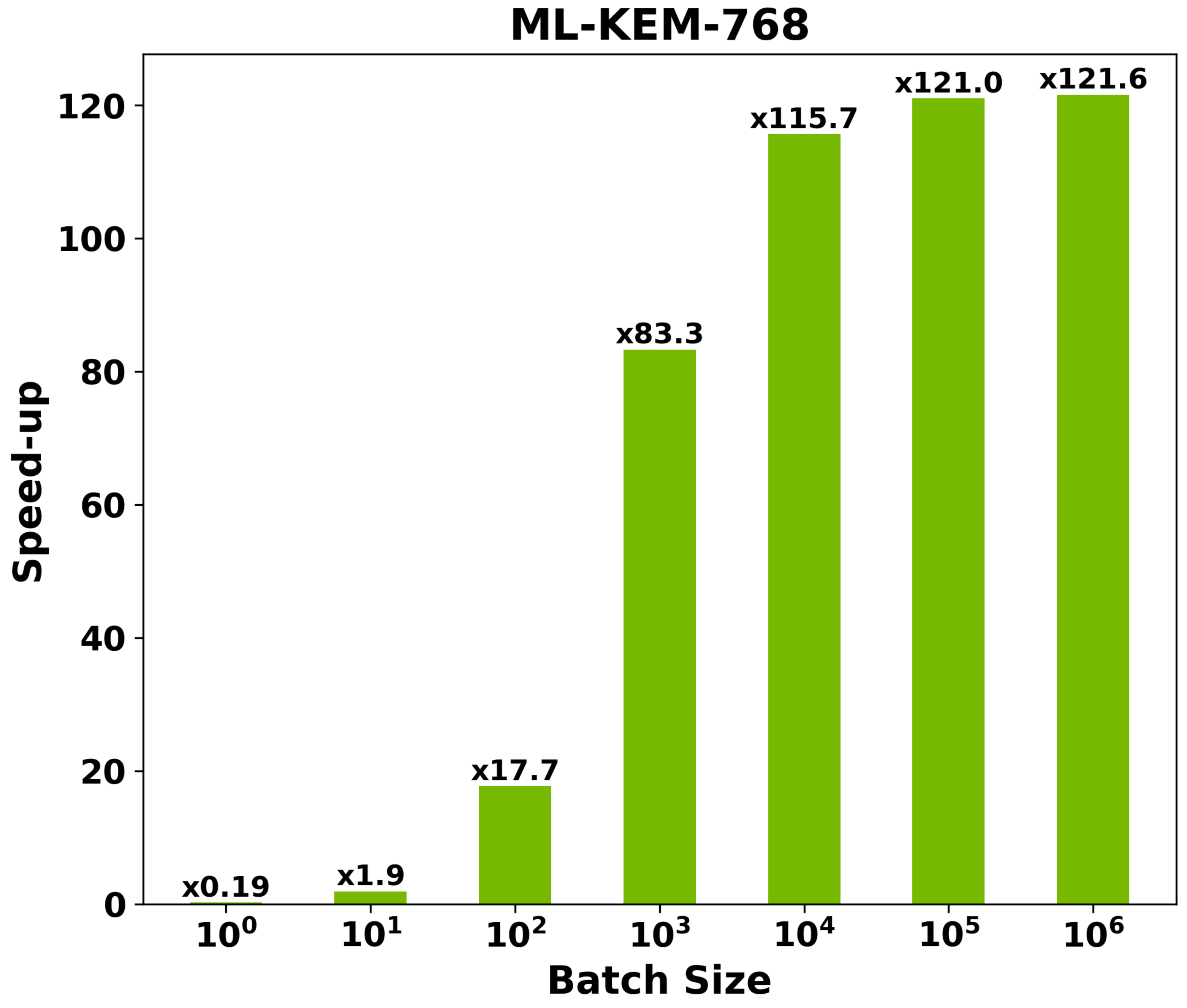


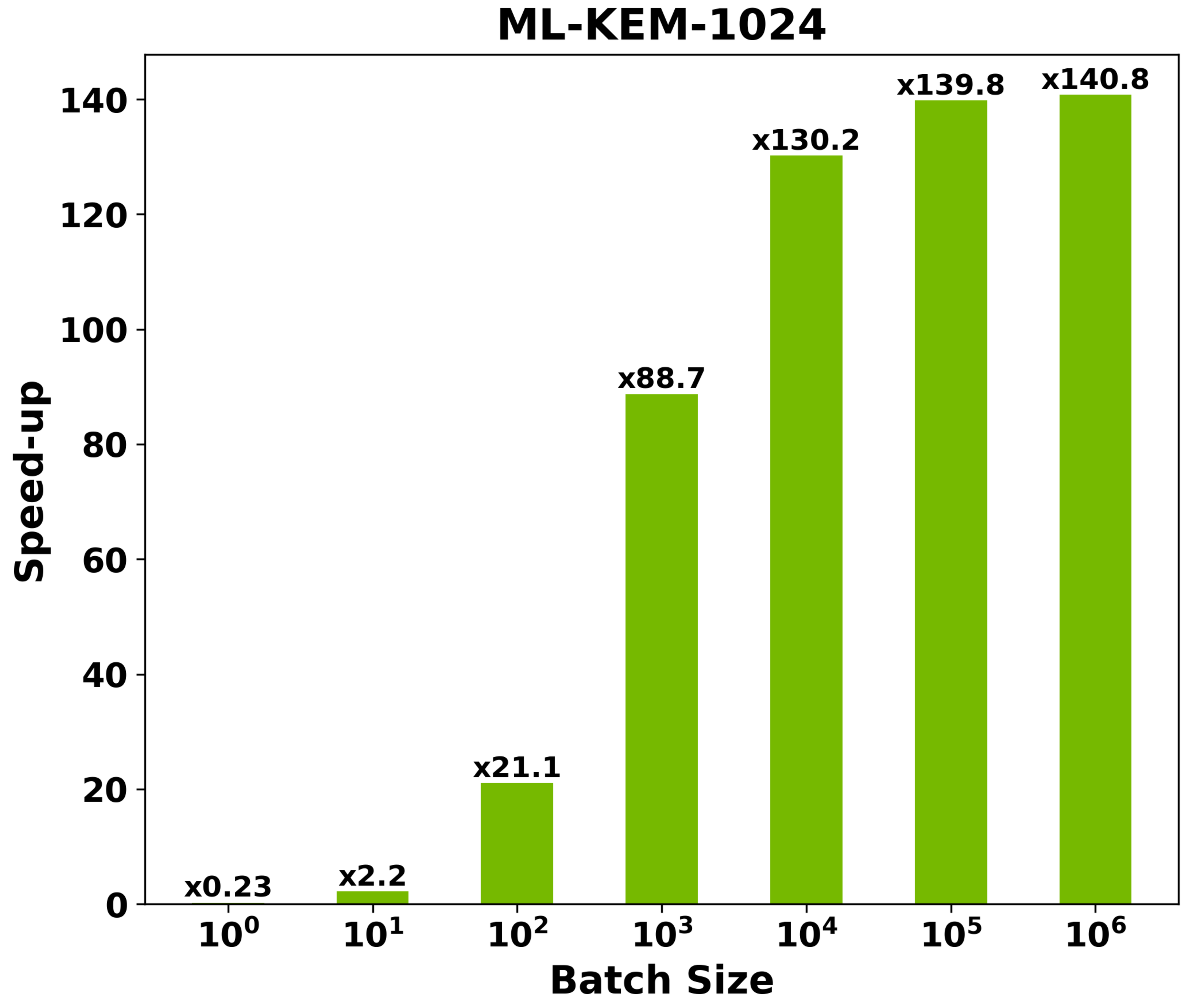
NVIDIA

ML-KEM Encapsulation Benchmarks

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5

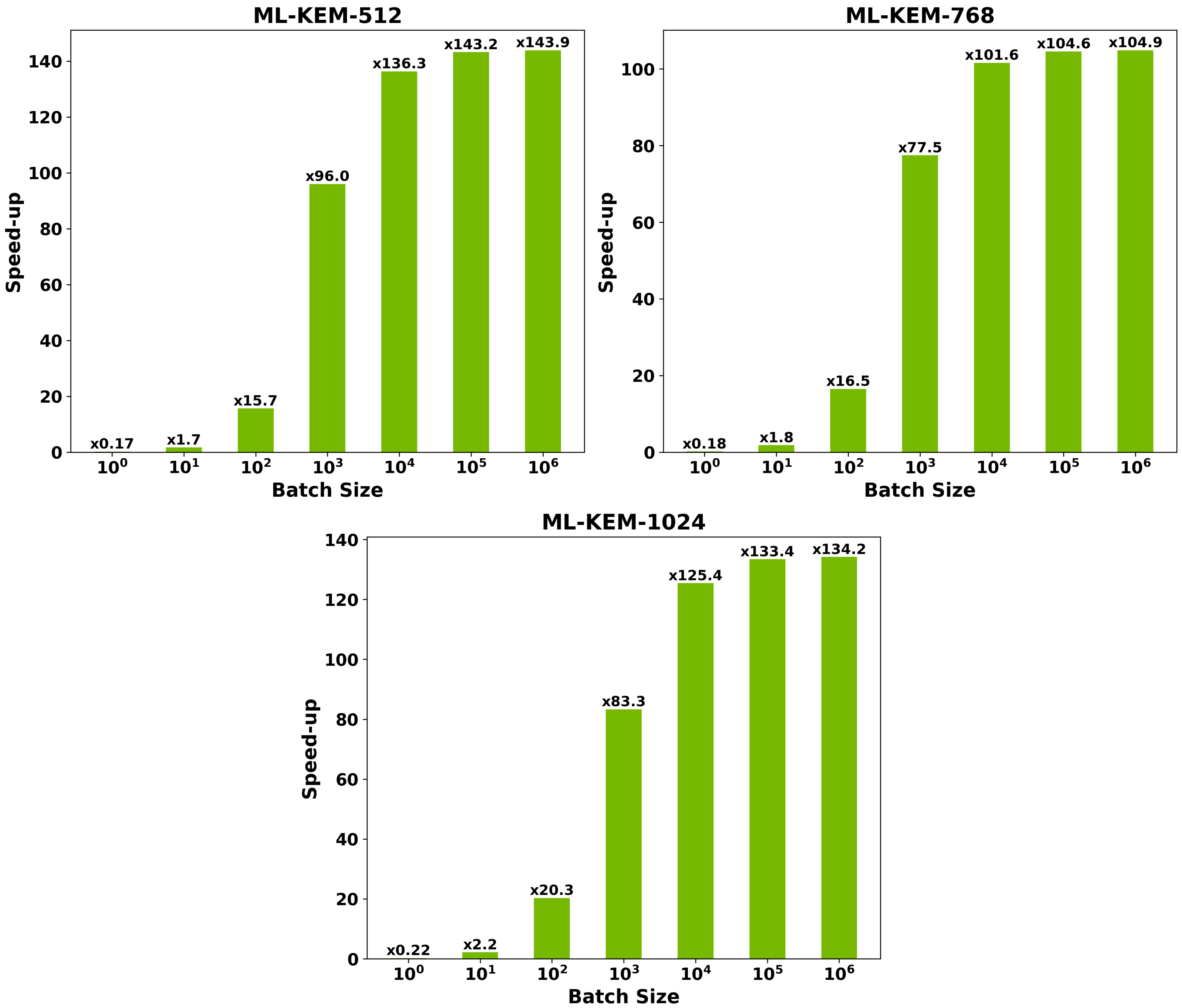






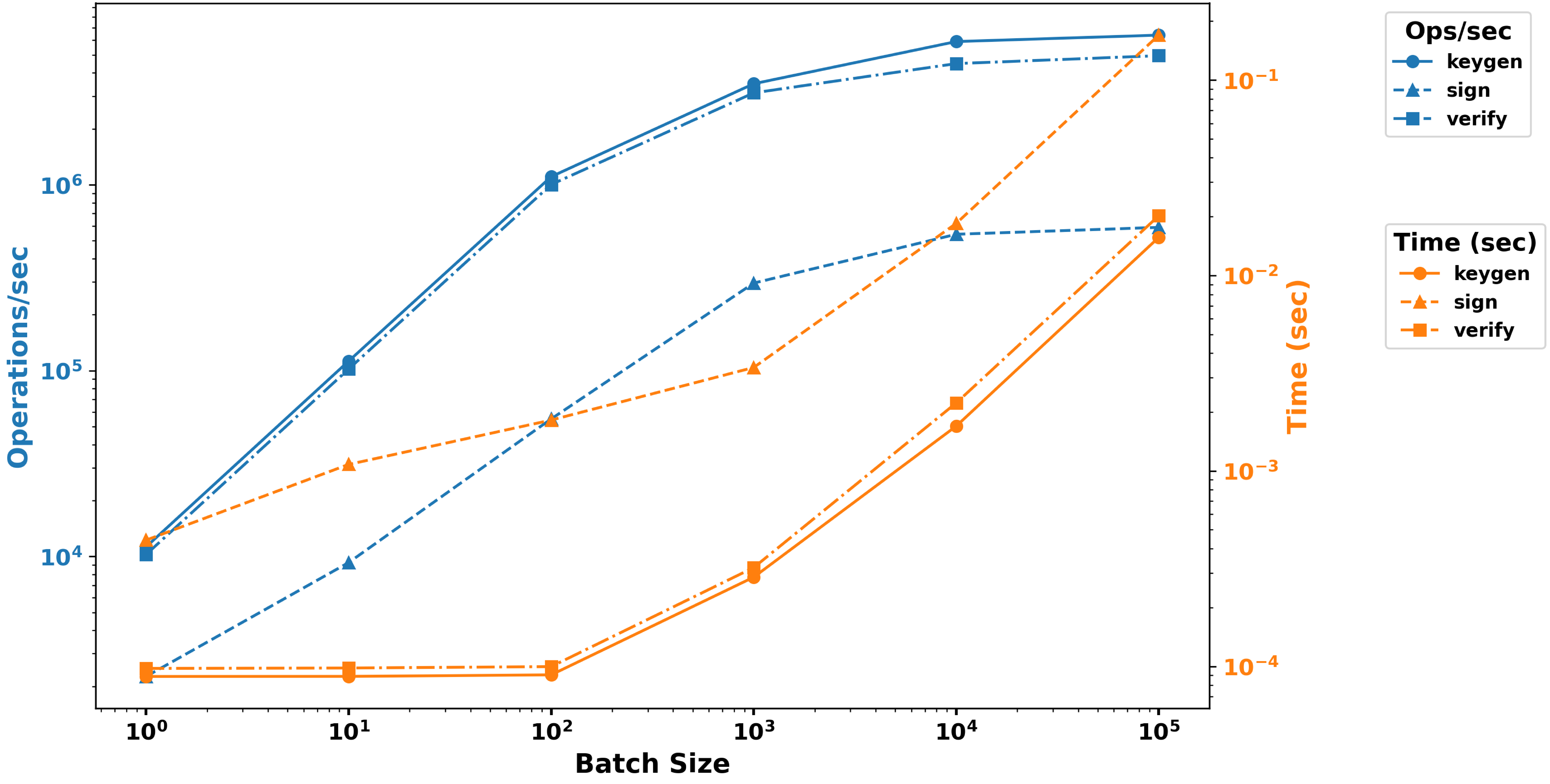
ML-KEM Decapsulation Benchmarks

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5



ML-DSA Benchmarks

ML-DSA-65



ML-DSA
(Batched in 100,000)

Keygen

Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-DSA-44	0.12	8.21
ML-DSA-65	0.16	6.38
ML-DSA-87	0.23	4.32

Sign

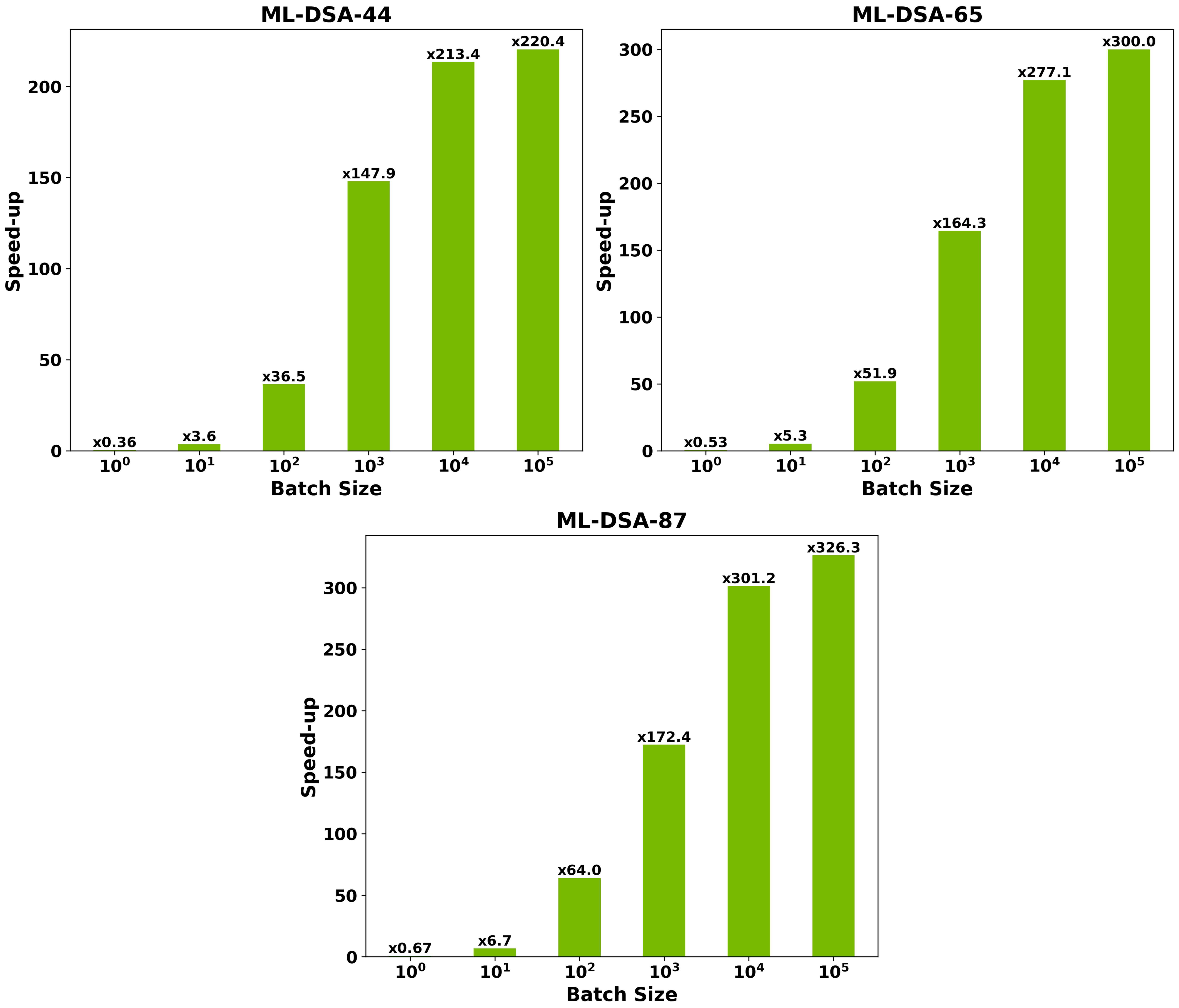
Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-DSA-44	1.30	0.76
ML-DSA-65	1.70	0.60
ML-DSA-87	1.76	0.57

Verify

Security Level	Amortized GPU Time (μsec)	Operations/sec (in Millions)
ML-DSA-44	0.136	7.35
ML-DSA-65	0.202	4.95
ML-DSA-87	0.285	3.50

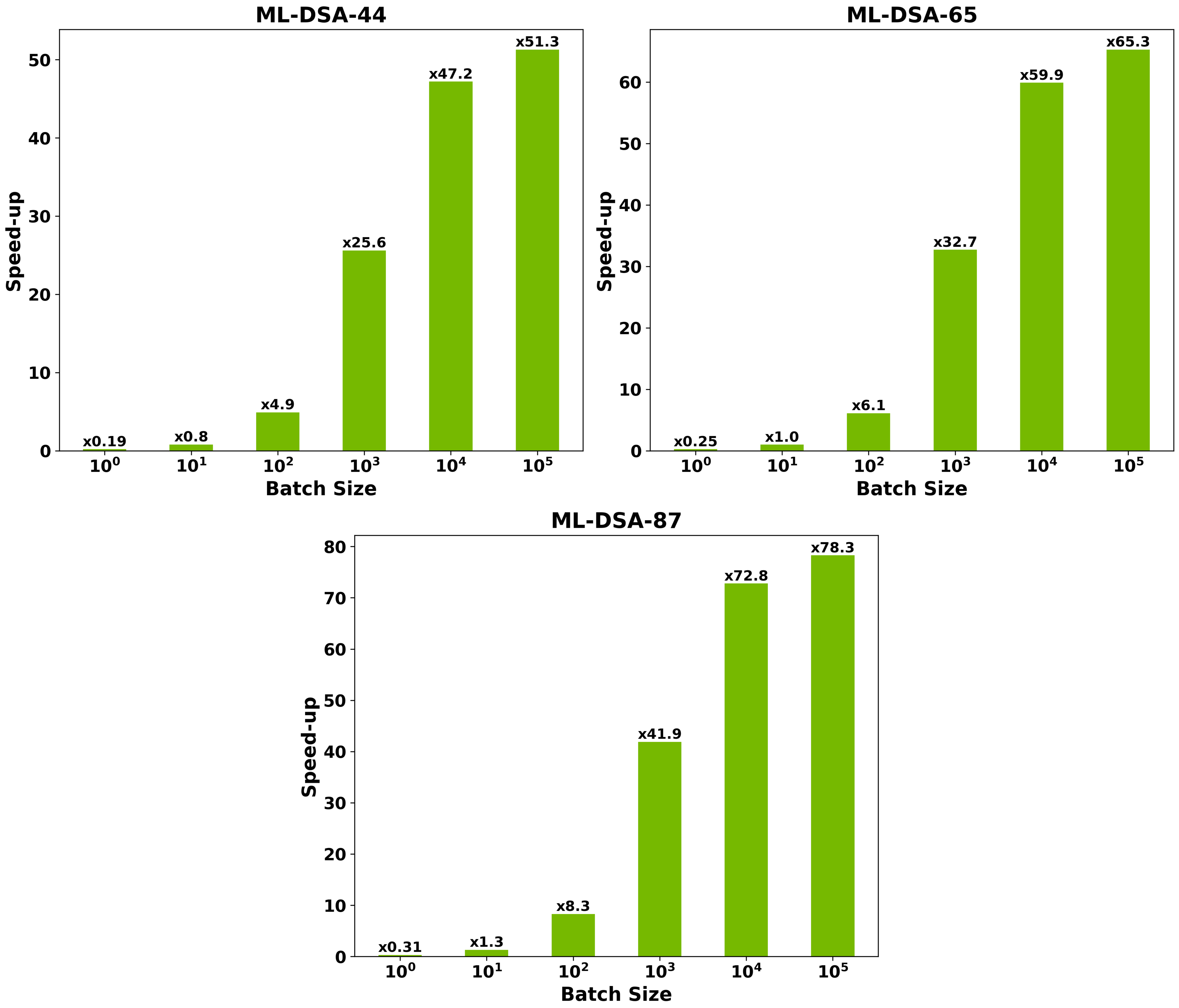
ML-DSA Keygen Benchmark

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5



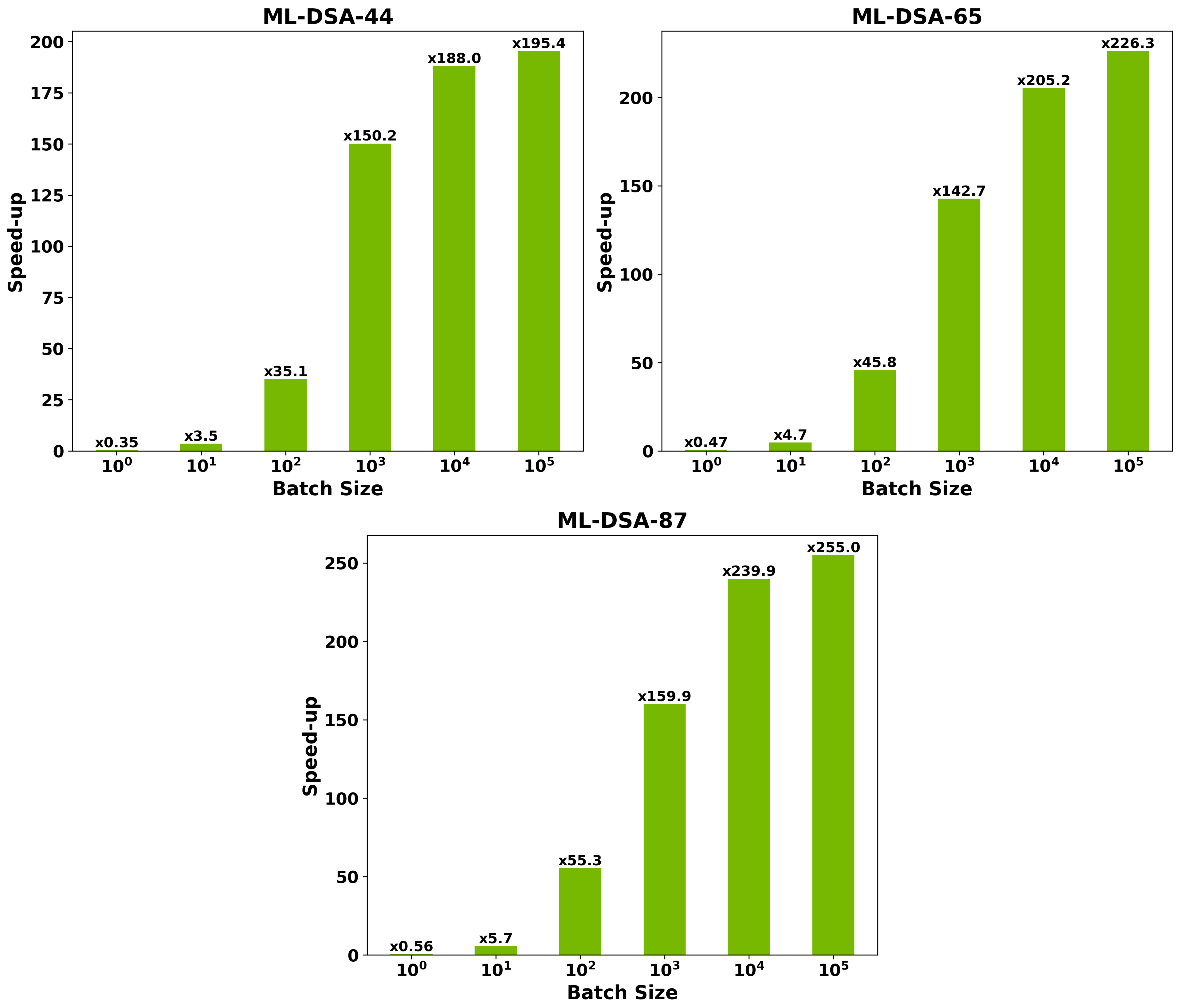
ML-DSA Sign Benchmark

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5



ML-DSA Sign Verify Benchmark

AMD EPYC 7313P 16-Core Processor vs NVIDIA H100 SXM5



Thank You!

Questions?