Realizing a more productive environment for quantitative analysis in the financial services industry

Improving the economics of analytics, risk management, and AI infrastructure with HPE Cray XD2000 systems powered by AMD EPYC processors

Introduction

Financial institutions (FIs), including investment banks, hedge funds, and private equity firms, face unprecedented challenges. With increasing customer expectations, the emergence of alternative banks, and born-in-the-cloud fintech, business models are constantly disrupted. FIs also need to deal with unprecedented levels of oversight and regulation. Along with insurers, proprietary trading firms, and market makers, banks increasingly compete on the scale and effectiveness of their analytics. Organizations that can develop more effective trading strategies, accurately price portfolios, and compute risk metrics for hedging quickly experience a substantial advantage versus their competition.

In this paper, we discuss the critical role of **quantitative analysis** and explain how the latest technologies from Hewlett Packard Enterprise and AMD can help FIs improve the scale, capacity, and efficiency of their analytics. HPE Cray XD2000 systems powered by 4th Gen AMD EPYC™ processors offer several advantages:

- **Unrivaled application performance**—The HPE Cray XD2000 systems powered by the latest 4th Gen AMD EPYC processors deliver higher throughput than servers powered by competing processors. With fast interconnects and exceptional I/O, these servers are ideally suited to quantitative analytics environments. Running a realistic mixed pricing quantitative analysis workload, the latest AMD EPYC processors delivered a dramatic 1.8x performance gain vs. 3rd Gen AMD EPYC-based servers.[1]
- **Built-in security features**—The HPE Cray XD2000 systems build on the integrated chip-level security features of AMD EPYC processors with integrated hardware root of trust functionality for secure operations, ensuring that proprietary and sensitive IP is protected in colocation and shared hosting environments.
- **Superior investment protection**—AMD EPYC processors are fully instruction-set compatible with other x86 processors helping most customers achieve impressive performance gains without recompilation or special tuning.
- **Exceptional TCO reduction**—The HPE Cray XD2000 systems shared infrastructure, supporting up to 80 servers per data center rack. And the latest AMD EPYC processors have better throughput per watt compared to 3rd gen AMD EPYC CPUs, resulting in dramatic savings enabling financial firms to support more compute capacity in power-constrained data center environments.[2]
- **Flexible deployment and procurement options**—With the HPE GreenLake edge-to-cloud platform, financial firms can realize a cloud experience in their own data centers with self-service, pay-per-use,* and the ability to scale capacity up and down. Together with HPE Pointnext Services, financial organizations can use HPE GreenLake to free up capital, boost operational and financial flexibility, and use talent to accelerate other important initiatives.

* May be subject to minimums or reserve capacity may apply

[1] Based on AMD internal QuantLib benchmark, November 2022.

[2] SP5-011C: SPECpower_ssj®2008 comparison based on published 2U, 2P Windows results as of 11/10/2022. Configurations: 2P AMD EPYC 9654 (30,602 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2022q4/power_ssj2008-20221204-01204.html) vs. 2P Intel® Xeon® Platinum 8480+ (16,653 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2023q1/power_ssj2008-20221207-01216.html). 2P AMD EPYC 7763 (23,505 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2021q2/power_ssj2008-20210324-01091.html) shown at 1.4x for reference. SPEC® and SPECpower_ssj® are registered trademarks of the Standard Performance Evaluation Corporation. See spec.org for more information.

Quantitative analysis

Quantitative analysis refers to the use of mathematical and statistical modeling, measurement, and research to understand and quantify market behaviors. By studying these behaviors, analysts can develop models that predict the prices of financial instruments under different market scenarios.

The main computational workload in investment banking is overnight risk and hedging calculations. Banks need to compute sensitivities and prices for all instruments on their books and do so before markets open the next day. Accurately calculating risk across the entire trading book is a massive computing problem involving millions of tasks. Calculations typically run across distributed computing environments comprising hundreds or even thousands of servers. Pricing and risk calculations need to run intra-day as well.

Having a more capable analytics infrastructure is advantageous for several reasons. Organizations with higher-performance analytics can:

- Price securities more accurately and get to market faster
- Improve profitability with a more thorough analysis of positions pre-trade
- Optimize portfolio composition to maximize performance and minimize risk
- Quantify and manage risk to minimize required reserves and keep capital productively employed
- Ensure compliance and meet new financial disclosure and reporting requirements
- Realize a healthier bottom line with more predictable financial performance

Key applications and use cases

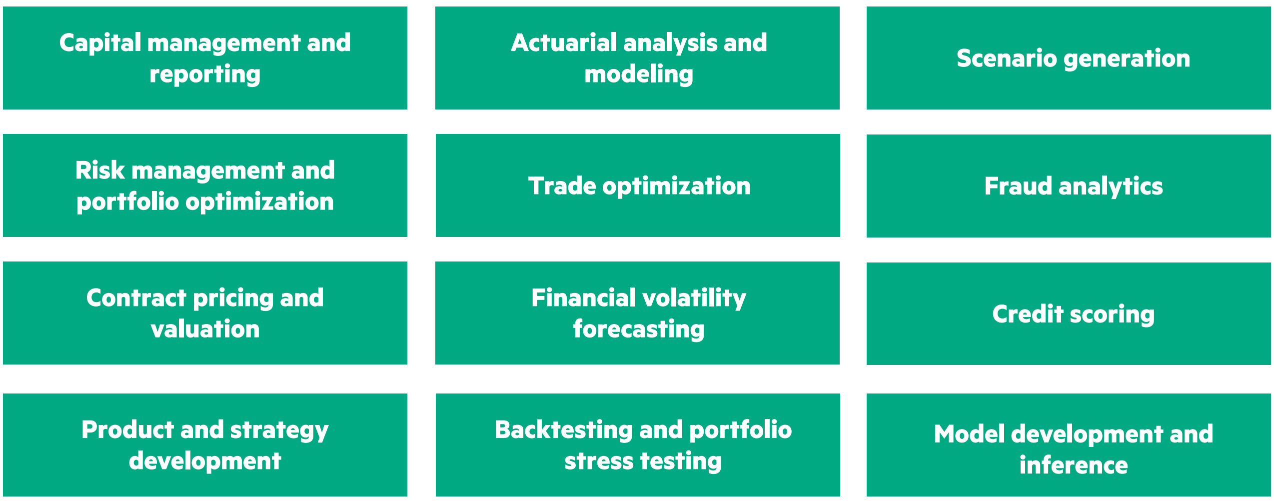


Figure 1. Key applications and use cases for quantitative analysis

There are dozens of use cases for quantitative analysis, including those illustrated in Figure 1. Computer simulation is used for applications that include contract pricing and valuation, risk management and portfolio optimization, scenario generation, and financial volatility forecasting.

A key workload is Monte Carlo simulation for financial workloads, which uses stochastic models to compute future values of portfolios over multiple scenarios and time steps. The Monte Carlo simulation is used to run scenarios, which is important because that is how financial institutions compete. They also compete based on their ability to accurately price complex financial instruments, including credit and interest rate derivatives and variable annuities.

Banks also employ backtesting to test the predictive quality of models by running them against historical market data. Insurers and re-insurers run similar compute-intensive actuarial analysis and modeling workloads to understand their risk exposures, project profit and losses, and develop new financial products.

Evolving business requirements

Following the financial crisis of 2008, banks faced a "tsunami" of new regulations around modeling liquidity risk and counterparty exposures. To follow these regulations, investment banks, fund managers, and proprietary trading firms had to significantly improve their risk management environments. Examples of these regulations include Dodd-Frank and Basel III, both enacted in 2010. Basel introduced Comprehensive Capital Analysis and Review (CCAR) guidelines, along with stress testing and liquidity requirements. It also introduced the phased implementation of its Fundamental Review of the Trading Book (FRTB), expected to come fully into effect in 2023. Additional regulations, including Solvency II, impose similar requirements on the insurance industry.

A key lesson from 2008 was the need for a stronger focus on credit risk. Regulators began to require that banks compute a variety of valuation adjustments collectively referred to as X-Value Adjustment (XVA) to reflect the risk-adjusted value of portfolios more accurately. These included metrics such as credit value adjustment (CVA), funding value adjustment (FVA), and capital value adjustments (KVA). Today, all positions need to be risk managed, and all trades need to be risk offset, which requires substantially more computing capacity.

Diverse applications

Another challenge associated with quantitative analysis workloads is their sheer diversity. FIs typically run a mix of commercial and in-house developed applications, increasingly leveraging open-source software. Figure 2 shows examples of these applications. Large firms often develop proprietary pricing models as a source of competitive advantage. In-house developed models are built using various tools, programming languages, and libraries. These models work alongside commercial software or open-source tools and frameworks for results aggregation, analysis, and reporting.

- Financial firms run diverse commercial and open-source applications for pricing, hedging, and risk management.
- Calculations are frequently distributed across large-scale grid computing environments.
- Computer systems must exhibit excellent performance across a range of workloads.

Figure 2. Financial firms run diverse quantitative analysis

FIs need servers and processors that perform well across all their workloads to support these applications. Banks typically have large sunk investments in pricing and analytic models and are reluctant to change them. In addition to the applications shown in Figure 2, quantitative analysts use hundreds of open-source analytic libraries for in-house pricing and risk models, particularly in proprietary trading. Examples include QuantPy, QuantLib, vollib, pynance, pyfin, qrmtools, quantmod, and ffn.

New challenges for IT

Quantitative analysis teams have an insatiable appetite for computing power. However, data center managers face several constraints, as shown in Figure 3. Data center managers must balance growing analytic requirements with limited data center capacity, power and density constraints, and finite budgets.

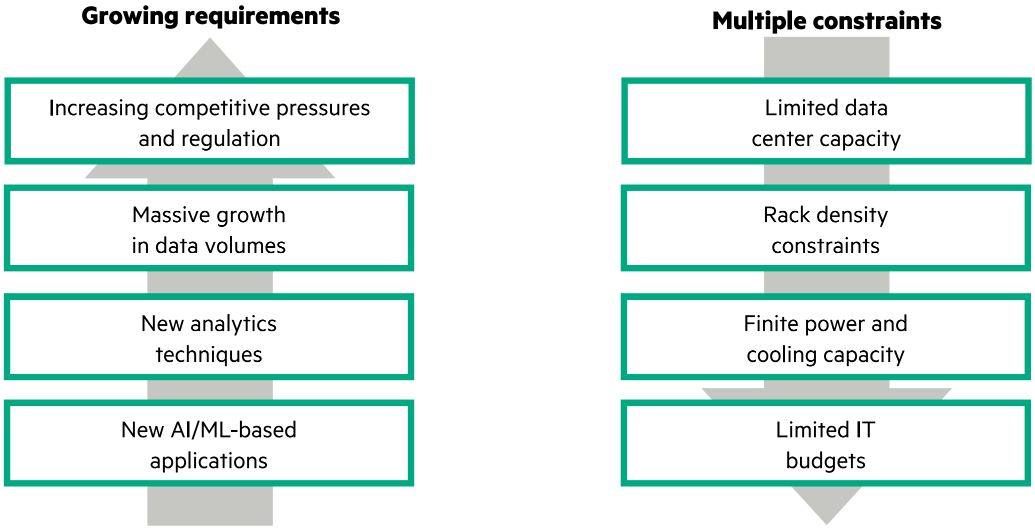


Figure 3. Complex requirements and challenges for data center

As a result, data center managers are looking for solutions that can deliver more throughput and capability per dollar invested in IT infrastructure.

Throughput and efficiency are key

For the simulation-intensive workloads described earlier, a key metric for FIs is throughput per dollar spent. The individual calculations that comprise large-scale workloads tend to be distributed across grid computing environments. These tasks are often single-threaded and favor multicore processor architectures that can run many tasks in parallel.

A related metric that is also critical is throughput per watt. Data centers often have limited power and cooling capacity that limit server density. Therefore, servers delivering higher throughput per watt are favored because they provide more capacity per data center rack and are typically cost-effective.

For quantitative analysis workloads, key metrics are throughput per dollar spent and throughput per watt. Financial firms compete based on their computing infrastructure's speed, capacity, and efficiency.

HPE solutions for quantitative analysis

HPE offers a variety of servers and managed hybrid cloud solutions that can help financial services customers accelerate digital transformation and deliver a more effective infrastructure for quantitative analysis workloads. Popular solutions for customers needing dense, high-performance solutions that fit standard data center environments include HPE GreenLake, and the latest HPE Cray supercomputers designed for corporate data centers.

The HPE Cray supercomputers were created from the ground up to handle demanding modeling, simulation, AI, and analytics workloads. The latest generation of supercomputers incorporate several of the pioneering design features of supercomputers built by HPE and powered by AMD but are designed to run in standard data center environments.[3]

[3] hpe.com/psnow/doc/a00127725enw?from=app&section=search&isFutureVersion=true

HPE Cray XD2000 system

The HPE Cray XD2000 system is a dense, multi-server platform that packs exceptional performance and workload flexibility into a small data center space while delivering the efficiencies of a shared infrastructure. Each HPE Cray XD2000 2U chassis supports up to four AMD EPYC-powered HPE Cray XD225v 1U servers. Alternatively, it can support two half-width 2U 2P HPE Cray XD295v servers. Each of the servers can be serviced without impacting the operation of other servers in the same chassis, increasing server availability. HPE Cray XD2000 delivers up to 4x the density of a traditional rackmount 2U server in standard racks and provides rear-aisle serviceability access.[4]

The HPE Cray XD2000 systems offer a complete, scalable solution for high-performance computing (HPC) customers, with flexible power and cooling options, including direct liquid cooling (DLC), delivering superior performance while reducing TCO. Up to 20 HPE Cray XD2000 chassis can be installed in either 42U or 48U HPE standard racks delivering up to 160 servers and 320 x 4th Gen AMD EPYC processors per data center rack, subject to power and cooling considerations.

Figure 4. Density-optimized HPE Cray XD2000 chassis supporting up to 4x HPE Cray XD225v 1U servers or 2x HPE Cray XD295v 2U servers

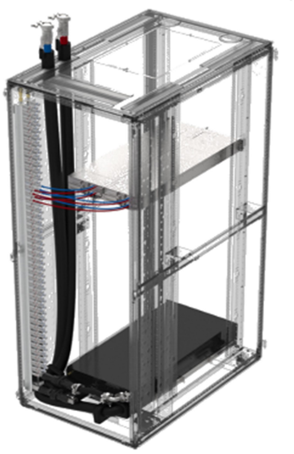


Figure 5. HPE data center rack with optional Direct Liquid Cooling (DLC)

Table 1. HPE Cray XD2000 system with 4th Gen AMD EPYC, HPE Cray XD225v (1U) and HPE Cray XD295v (2U) specifications

Processor	Up to two 4th Gen AMD EPYC processors—up to 96 cores / 192 threads per socket
OS support	Windows, Red Hat® Enterprise Linux®, SLES, Ubuntu, HPE Cray OS, VMware®
Memory	12 channels DDR5 per CPU @4800 MT/s (up to 24 total DIMMs for 2P)
Nodes	1U/2P half-wide; up to 4 per chassis 2U/2P half-wide (GPU support); up to 2 per chassis
PCIe lanes	65 PCIe 5.0 lanes with 4-link XGMI + 6 bonus PCIe 3.0 lanes per CPU
Storage	Supports up to 2x NVMe SSDs per node (4 for 2U) and/or M.2 2280/22110
Expansion slot	2x PCIe Gen5 x16 slot 1x PCIe Gen5 x16 MCIO slot 1x PCIe Gen5 x16 MCIO cable slot 1x PCIe Gen3 x4 M.2 Mezz
Power supply (hot plug)	CRPS PSU support 2400W 240V
Fans	4x 4056 fans per node
Cooling	Air-cooled or optional DLC

Customers who wish to boost compute density can deploy the HPE Cray XD2000 with up to 4 1U HPE Cray XD225v servers per chassis. For customers running GPU workloads, up to 2 HPE Cray XD295v servers can be installed in each chassis supporting the latest GPUs, including AMD Instinct™ MI210 PCIe Accelerators.

[4] See HPE Cray XD2000 QuickSpecs

Latest technologies for accelerated computing

The HPE Cray XD2000 supercomputers also feature a variety of technologies to support compute- and data-intensive workloads, such as high-performance analytics, AI, and machine learning (ML). While financial services customers may only need some of these advanced technologies at this time, they are available as requirements evolve. These technologies include:

- **Choice of fabrics:** Customers can choose from InfiniBand, standard Ethernet options, or HPE Slingshot—an Ethernet interconnect to address demands for high speed and better congestion control, helping complex applications run more smoothly. Fast interconnects can help maximize simulation performance in large-grid environments managed by TIBCO DataSynapse GridServer® Manager, IBM Spectrum Symphony®, Kubernetes, and other workload management solutions commonly used in financial services.
- **Cray ClusterStor E1000:** An optional high-performance storage solution that provides intelligent tiering support for data-intensive workloads, such as AI model training to make storing and accessing data in large-scale distributed environments easy and efficient.
- **System management tools:** Customers can choose HPE High Performance Cluster Manager or NVIDIA® Bright Cluster Manager to manage their high performance computing environments.
- **HPE Cray Programming Environment:** A fully integrated software suite with compilers and developer tools allowing developers to optimize applications for high-performance modeling, simulation, and AI applications.
- **MPI implementation:** For distributed parallel workloads implemented using the Message Passing Interface (MPI), customers can choose from HPE Cray MPI, Open MPI, and Mellanox HPC-X.

The AMD "EPYC" advantage

First introduced in June of 2017, AMD EPYC processors bring together high core counts, large memory capacity, extreme memory bandwidth, large cache sizes, and massive I/O with the right ratios to enable exceptional performance for analytic workloads. For financial services customers, this can translate into faster pricing calculations and more thorough simulation, delivering a significant competitive advantage.

4th Gen AMD EPYC processors offer several advantages over the previous generation. Among these are:

- 2P 96 cores EPYC 9654-powered servers deliver 1.45x the floating-point throughput of alternative processors[5]
- A faster Infinity Fabric™, delivering 2x the speed of the prior generation[6]
- 50% more memory channels than any other x86 architecture CPU[7]
- Industry-leading I/O capacity with up to 4x the I/O of competing processors[8]
- Advanced chip-level security enhancements (SME, SEV-ES, and SEV-SNP)

The latest EPYC processors deliver a 1.9x top-of-stack generational performance uplift in Java applications and deliver approximately 1.8x the throughput per-watt based on the SPECpower_ssj 2008 benchmark.[9, 10] This means that financial services customers can achieve dramatic performance and efficiency gains by thoughtful selections of their server infrastructure.

[5] SP5-009D: SPECrate®2017_fp_base based on published scores from spec.org as of 1/11/2023. Configurations: 2P AMD EPYC 9654 (1480 SPECrate 2017_fp_base, 192 total cores, spec.org/cpu2017/results/res2022q4/cpu2017-20221024-32605.html) is 1.45x the performance of published 2P Intel Xeon Platinum 8490H (1020 SPECrate 2017_fp_base, 120 total cores, spec.org/cpu2017/results/res2023q1/cpu2017-20221206-33040.html). SPEC, SPEC CPU®, and SPECrate® are registered trademarks of the Standard Performance Evaluation Corporation. See spec.org for more information.

[6] See the 4th Gen AMD EPYC Processor Architecture white paper, page 8.

[7] EPYC-033A: AMD EPYC 9004 CPUs support 12 memory channels. 3rd and 4th Gen Intel® Xeon® CPUs support 8 memory channels. 12 ÷ 8 = 1.5x the memory channels or 50% more memory channels per ark.intel.com/.

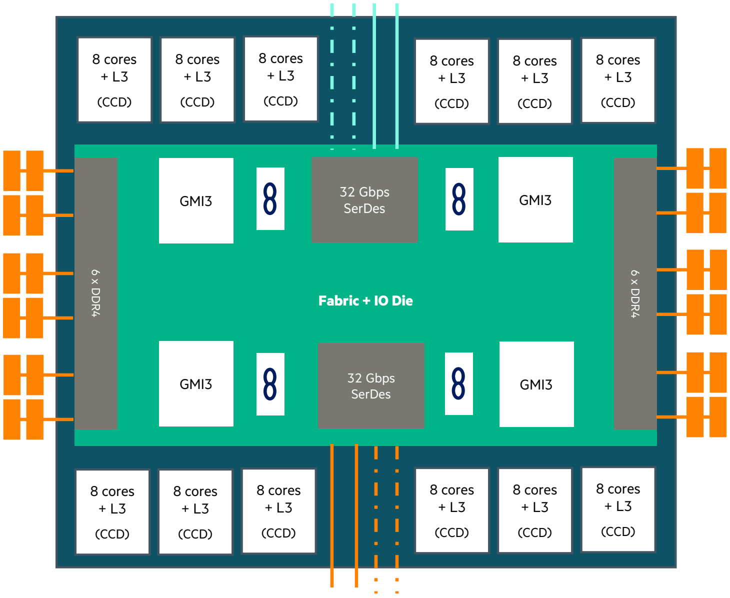
[8] EPYC-035A: One AMD EPYC 9004 CPU supports 128 PCIe® Gen5 lanes plus up to 8 PCIe3 lanes. One 4th Gen Intel Xeon Scalable CPU supports up to 80 lanes of PCIe5 per ark.intel.com/. 128 ÷ 80 = 1.6x

[9] spec.org/jbb2015/results/res2021q3/jbb2015-20210701-00692.html; SP5-005C: SPECjbb®2015-MultiJVM Max comparison based on published results as of 11/10/2022. Configurations: 2P AMD EPYC 9654 (815459 SPECjbb2015-MultiJVM max-jOPS, 356204 SPECjbb2015-MultiJVM critical-jOPS, 192 total cores; spec.org/jbb2015/results/res2022q4/jbb2015-20221019-00861.html) vs. 2P AMD EPYC 7763 (420774 SPECjbb2015-MultiJVM max-jOPS, 165211 SPECjbb2015-MultiJVM critical-jOPS, 128 total cores, spec.org/jbb2015/results/res2021q3/jbb2015-20210701-00692.html). SPEC and SPECrate are registered trademarks of the Standard Performance Evaluation Corporation. See spec.org for more information.

[10] SP5-011C: SPECpower_ssj 2008 comparison based on published 2U, 2P Windows results as of 11/10/2022. Configurations: 2P AMD EPYC 9654 (30,602 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2022q4/power_ssj2008-20221204-01204.html) vs. 2P Intel Xeon Platinum 8480+ (16,653 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2023q1/power_ssj2008-20221207-01216.html). 2P AMD EPYC 7763 (23,505 overall ssj_ops/W, 2U, spec.org/power_ssj2008/results/res2021q2/power_ssj2008-20210324-01091.html) shown at 1.4x for reference. SPEC and SPECpower_ssj are registered trademarks of the Standard Performance Evaluation Corporation. See spec.org for more information. 30,602 ssj_ops/W / 16,653 ssj_ops/W = 1.84 energy efficiency improvement.

An ideal processor for quantitative analysis

The unique architecture shown in Figure 6 is the key to the 4th Gen AMD EPYC processor's throughput advantage. The 13-die system-on-a-chip (SoC) features 1 I/O die and 12 core complex dies (CCDs). The 32 MB of cache is shared across all cores on each CCD. The advanced 5 nm process enables clock frequencies to scale up to 4.40 GHz maximum boost, delivering exceptional single-threaded performance for time-critical risk and pricing calculations.[11]



- Next-generation 5 nm technology
- Up to 96 cores/192 threads
- Up to 4.40 GHz Max boost clock speed*
- Up to 384 MB L3 cache
- Up to 32 MB shared L3 cache per CCD
- Infinity Fabric—4 x 32 Gbps socket-to-socket connections
- 12 DDR5 memory channels per socket
- Up to 4800 MT/s DDR
- 128 PCIe Gen5 lanes per socket
- CXL™ 1.1+ memory expansion
- Embedded AMD Secure Processor

* EPYC-18: Refers to the AMD EPYC 9174F part. Max. boost for AMD EPYC processors is the maximum frequency achievable by any single core on the processor under normal operating conditions for server systems.

Figure 6. 4th Gen AMD EPYC high-level processor design

Processors tailored to your workload

AMD EPYC processors can be grouped into distinct categories depending on what attributes are important for a workload, as illustrated in Table 2. Customers requiring the highest per-core performance for time-critical pricing workloads can choose high-frequency parts. Customers can choose high-throughput parts with higher core and thread counts for quantitative analysis and other workloads where throughput and density are more important than single-core performance.

Table 2. Selected 4th Gen AMD EPYC processors

EPYC model	Cores/ threads	Base clock	Max. boost clock[12]	All core boost speed[13]	L3 cache	Power (Watts)	L3 cache per core	Workload type
AMD EPYC 9004 Series								
9654	96/192	2.40 GHz	Up to 3.7 GHz	3.55 GHz	384 MB	360	4 MB	Throughput
9554	64/128	3.10 GHz	Up to 3.75 GHz	3.75 GHz	256 MB	360	4 MB	Throughput
9474F	48/96	3.60 GHz	Up to 4.1 GHz	3.95 GHz	256 MB	360	5.3 MB	Core performance
9374F	32/64	3.85 GHz	Up to 4.3 GHz	4.1 GHz	256 MB	320	8 MB	Core performance
9274F	24/48	4.05 GHz	Up to 4.30 GHz	4.1 GHz	256 MB	320	10.7 MB	Core performance
9174F	16/32	4.10 GHz	Up to 4.40 GHz	4.15 GHz	256 MB	320	16 MB	Core performance

[11] Refers to the AMD EPYC 9174F part. Max. Boost for AMD EPYC processors is the maximum frequency achievable by any single core on the processor under normal operating conditions for server systems (EPYC-18).

[12] Max. boost speed for EPYC processors is the maximum speed achievable by any core on the processor under normal operating conditions for server systems (EPYC-18).

[13] All-core boost for EPYC processors is the average frequency of all processor cores running in performance mode while utilizing a low-activity workload. Actual achievable all-core boost will vary based on hardware, software, workloads, and other conditions (EPYC-021).

Putting 4th Gen EPYC processors to the test

In November 2022, AMD ran a synthetic benchmark based on QuantLib—an open-source library written in C++—to characterize the performance of 4th Gen EPYC processors running a real quantitative analysis workload. QuantLib is a widely used research tool with high-quality algorithms in several areas. Sometimes, smaller FIs use it in production, and large firms may use QuantLib for challenger models during model validation. In terms of performance, QuantLib serves as a reasonable proxy for a production quant risk library.

To simulate a realistic workload, AMD selected 21 tests from the QuantLib test suite covering various use cases.[14] To simulate a large workload, typical of what might be seen on a compute node in a grid computing environment, each of the 21 tasks was run 400x for a total of 8,400 tasks.

GNU parallel was used to queue the individual tasks so that the number of tasks executing in parallel per node was equal to the number of cores on each server to simulate the behavior of a scheduler in a large-scale grid environment. The 8,400 tasks were run on three different 2P servers having the same amount of physical memory and storage.[15]

For quantitative analysis workloads, 4th Gen AMD EPYC processors can deliver up to 2.33x more throughput while improving energy efficiency by ~1.31x compared to previous generations for dramatic improvements in data center efficiency.[16]

The results of this mixed pricing workload are shown in Figure 7.

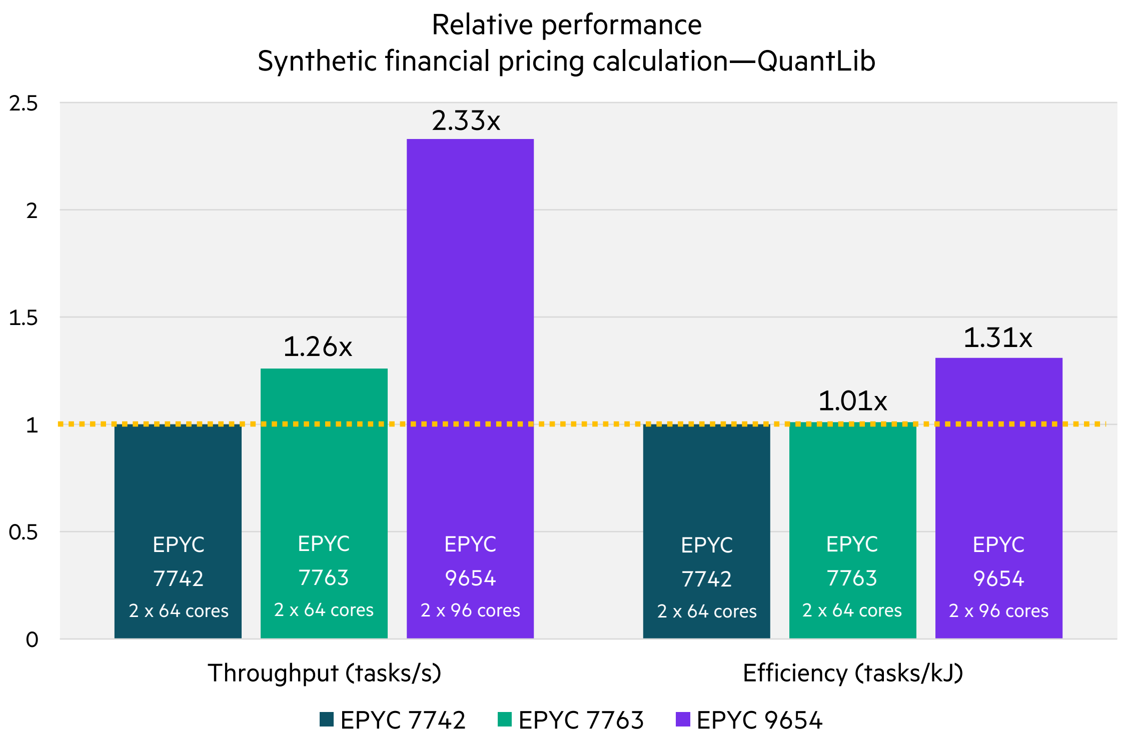


Figure 7. Comparing successive generations of EPYC processors in terms of throughput and efficiency

[14] The complete QuantLib test suite is available at quantlib.org/reference/test.html. Tests included pricing barrier options using Monte Carlo and PDEs, pricing Bermudan swaptions, calibrating and pricing Heston and Heston SLV models via PDEs and Monte Carlo, pricing Asian options and convertible bonds, Longstaff-Schwartz pricing of American options, and calibrating and pricing Heston-Hull-White models via Monte Carlo and PDEs.

[15] Server configurations were as follows: A dual-socket system with 2nd Gen AMD EPYC 7742 processors for a total of 128 physical cores and a Thermal Design Power (TDP) of 225W per processor. A dual-socket system with 3rd Gen AMD EPYC 7763 processors for a total of 128 physical cores and a TDP of 280W per processor. A dual-socket system with 4th Gen AMD EPYC 9654 processors with 192 physical cores and a TDP of 360W per processor.

[16] SP5-094: QuantLib v1.92 workload running 8400 tasks (21 jobs) comparison based on AMD measurements as of 12/13/2022. A TDP-kJ is computed as (2 * TDP-of-processor) * (benchmark runtime) / 1000. This is not an actual energy-consumed measurement but is a proxy for energy-consumed. Configurations: 2 x EPYC 9654@ 400W (Silicon TDP 360W) scoring 77.534 tasks/sec vs 2 x EPYC 7763@ 280W scoring 42.053 tasks/sec vs. 2 x EPYC 7742@ 225W scoring 33.319 tasks/sec. Results may vary based on factors such as software and driver versions, BIOS configurations, and benchmark size.

Superior investment protection

Because FIs have large sunk investments in model development and validation, compatibility with previous-generation processors is essential. The latest AMD EPYC 4th Gen processors are fully instruction set compatible with existing x86 processors helping ensure reliable execution without recompiling code.

For numerically intensive financial applications, AMD offers an updated version of its AMD Optimizing CPU Libraries (AOCL) optimized for the latest fourth-generation Zen cores. AOCL 4.0 provides tuned implementations of industry-standard math libraries commonly found in quantitative analysis applications.[17] Among these standard libraries is AOCL-LibM, providing scalar and vector implementations of standard math functions. Developers can optionally link to these new libraries to replace compiler-provided math functions. These libraries also offer updated implementations of key libraries used in FSI-specific workloads such as Black-Scholes and Monte Carlo simulations.[18] Also, to further improve performance, customers willing to recompile code can optionally take advantage of AMD Optimizing C/C++ compilers (AOCC) enhanced for the latest 4th Gen EPYC processors.

The latest 4th Gen AMD EPYC processors also provide built-in AVX-512 instruction support, so customers already running code optimized to take advantage of AVX-512 vector instructions can leverage these enhancements.

HPE GreenLake

Quantitative analysis and risk applications need high-performance computing, and IT departments are struggling to keep up with the necessary skills to make HPC broadly accessible and manage increased demand. With the HPE GreenLake platform, organizations can gain the agility to scale the environment easily and empower users with self-service.

The HPE GreenLake platform brings cloud experience, self-service, pay-per-use, and the ability to scale capacity up and down. It is fully managed by HPE and deployable in any location, including the network edge, colocation facilities, and financial data centers. HPE GreenLake can help banks and other financial institutions evolve toward as-a-service delivery models that support virtualized or containerized applications to improve utilization and efficiency.

Customers transitioning to HPE GreenLake from traditional CAPEX models can achieve a greater than 30% reduction in energy costs and total cost of ownership.[19] Multiple financial firms already realize the benefits of HPE GreenLake, including YF Life, au Kabucom Securities Co., Ltd., and MKB Bank.[20] According to a recent Forrester study, HPE GreenLake users can enjoy an ROI of up to 161%.[21]

[17] See developer.amd.com/amd-aocl/ for additional details

[18] See the Financial Services Industry (FSI) AMD EPYC 9004 tuning guide for details—amd.com/system/files/documents/58026-epyc-9004-tg-fsi.pdf

[19] "Hewlett Packard Enterprise Releases 2021 Living Progress Report; Accelerate Net-zero Climate Target by 10 Years," June 2022

[20] The future of financial services—HPE GreenLake for financial services

[21] "The Total Economic Impact™ Of HPE GreenLake," Forrester Consulting, May 2022

Delivering on climate commitments by advancing sustainable computing

A key concern for all businesses is the unsustainable carbon dioxide (CO_2) emissions and greenhouse gasses (GHG) driving climate change. Most large enterprises, including major banks, have made public commitments around climate initiatives to employees, customers, and shareholders. As such, meeting these commitments is a high priority. With its exceptional power efficiency and density, HPE Cray XD2000 supercomputers can play a significant role in helping FIs meet their climate commitments.

For additional details on how HPE servers powered by AMD EPYC processors can help financial services firms achieve their sustainability goals, read the white paper Addressing sustainability in the financial services industry.

Security features built-in

For security conscience financial services environments, security features are built into HPE Cray XD2000 systems from the hardware layer up, creating a chain of trust that spans from the lowest levels of silicon to the workloads and data residing on servers. The chain of trust from HPE is designed to secure HPE Cray XD2000 systems against malicious firmware attacks.

The HPE Cray XD2000 supercomputers build on the integrated chip-level security features of 4th Gen EPYC processors with integrated hardware root of trust functionality for secure operations of the computer system. The hardware root of trust with automated firmware recovery prevents systems from booting with compromised firmware and restores them to the most recent working state. Additional standard security features on HPE Cray XD2000 supercomputers include:

- UEFI Secure boot and Secure Start support
- HPE Cray XD hardware root of trust
- Tamper-free updates—components digitally signed and verified
- Secure recovery—recover critical firmware to a known good state on detection of compromised firmware

HPE Pointnext Services

HPE provides various services to help financial services organizations get the most out of their computing environments. HPE has decades of industry experience tailored to planning, deploying, and managing highly customized solutions and organizations can collaborate with technical and professional experts through HPE Pointnext Services.

HPE Pointnext Services are designed to help simplify the deployment of high-performance systems and boost operational flexibility—with offerings such as project management, on-site consulting, and solution architecture consulting. Other services include:

- Pre-installation activities and solution implementation
- System testing and performance validation
- Advanced hardware and software support
- Lifecycle services
- Customized customer training

HPE Pointnext Services can help determine an optimal infrastructure environment to support a quantitative analytics environment depending on the customer's mix of applications, help remove potential roadblocks, and even help manage the complete technology environment after deployment.

The HPE and AMD advantage

For financial services institutions running quantitative analysis and other workloads, the latest HPE Cray XD2000 supercomputers powered by AMD EPYC processors offer compelling advantages. These advantages include unrivaled application performance, built-in security features, investment protection including compatibility with existing x86 applications, exceptional TCO reduction, and flexible performance and procurement options.

Learn more at

hpe.com/us/en/solutions/amd.html

hpe.com/servers/crayxd2000

amd.com/epyc

Hewlett Packard Enterprise

a50008228ENW