A STRATEGIC GUIDE FOR SWITCHING: FROM INTEL® XEON® TO AMD EPYC™ SERVER CPUS

WHITEPAPER | 2026

A pragmatic guide to surpassing the status quo with AMD EPYC™ Server CPUs: Address migration myths, slash licensing costs, and execute a strategic refresh that pays for itself.

Executive Summary: The Cost of Stagnation

The modern enterprise data center stands at a precarious inflection point. For the better part of two decades, deploying Intel® Xeon® server processor-based systems was the "safe bet"—a default choice that required no justification to the Board or the risk officer. This "status quo" bias emerged from an era of architectural homogeneity, where choices were very limited, performance gains were iterative and incremental, and power constraints were secondary.

Today, that paradigm has inverted. The landscape of data centers and energy economics has shifted so dramatically that the historical "safe" choice has effectively become an active liability. Sticking with legacy architectures now risks incurring quantifiable penalties in Total Cost of Ownership (TCO), power efficiency, physical density, and operational agility. There has never been a stronger need for change.

The Advantages of AMD EPYC Server CPU-Based Systems

Moving away from legacy architecture often triggers execution paralysis in IT leadership. It is not a question of superior performance or energy efficiency—the benchmarks for AMD EPYC™ Server CPU-based systems are undeniable. It is a question of migration complexity.

Most marketing glosses over this complexity. This paper will not.

We acknowledge the reality: migrating architectures requires effort. It requires some downtime and planning. The operational savings and performance gains are often so significant that they pay for the upfront lift many times over—not to mention that even transitions within legacy platforms involve some of these same hurdles.

This guide provides a strategic framework for a partial-to-full refresh and offers an honest look at how to execute the switch with predictability and confidence.

Section 1 - Overcoming the Inertia: Dismantling the Myths

The inertia preventing organizations from migrating is frequently rooted in outdated assumptions rather than current technical realities. These barriers often manifest as myths surrounding software compatibility, downtime, and enterprise readiness.

Myth #1: "It's a different architecture, so my software won't run."

HPC and AI	ALTAIR, Ansys, Dassault Systemes, deci, Neural Magic, ONNX, PyTorch, SIEMENS, TensorFlow
Database Analytics	CLOUDERA, Couchbase, databricks, DATASTAX, elastic, MarkLogic, mongoDB, splunk>, SingleStore, VERTICA
Database	Exasol, influxdata, Microsoft SQL Server, ORACLE, PostgreSQL, redislabs, TigerGraph, SAP, cassandra, MySQL, MariaDB
OS	CANONICAL, citrix, FreeBSD, Microsoft, NUTANIX, ORACLE, Red Hat, SUSE, vmware
HCI / Orchestration	docker, kubernetes, Microsoft, NUTANIX, Red Hat, simplivity, vmware
SDS	CLOUDIAN, Excelero, Pivot3, Quobyte, ceph, StorMagic, WEKA

Reality: A prevalent concern is that a changing silicon architecture means refactoring the software stack. That can be true when the change is to a different Instruction Set Architecture (ISA)—which is one reason many enterprises have moved cautiously with RISC platforms like Arm. Shifting ISAs can drive real work: rebuilding and retesting applications, validating third-party dependencies, and re-certifying operational tooling across the environment. But moving between x86 vendors is fundamentally different. The ISA stays x86, the Application Binary Interface (ABI) expectations remain consistent, and for most enterprise deployments, x86 looks like x86—regardless of the brand. Just as importantly, most hardware/software interaction happens at the OS or hypervisor abstraction layer, which largely insulates workloads from vendor-specific implementation details.

In practice, that x86 consistency shows up in the x86 Ecosystem Advisory Group (EAG)—a joint initiative including AMD, Intel, and key industry partners—has formalized interoperability going forward. Technical milestones like FRED (Flexible Return and Event Delivery) and standardized AVX10 ensure that the ecosystem remains unified at the instruction level.

Myth #2: "Migration causes unacceptable downtime."

Reality: You are likely already scheduling "downtime"; you just call it maintenance. There is a perception that upgrading old Intel systems-to-new Intel systems allows for seamless, zero-downtime upgrades forever. **This is false.**

- The Hard Truth: If you are upgrading from a 5-year-old Intel Xeon "Ice Lake" cluster to the latest Intel Xeon 6 environment, you cannot simply vMotion active memory without performance penalties. To get the benefits of the new silicon, you often need a restart (Cold Migration) anyway.
- The Fix: Cold migration is functionally equivalent to a standard maintenance reboot for OS patching. It is a predictable, managed event.

Myth #3: "I need to replicate the old design."

Reality: The significant advances in processor and system architecture open new options for demanding business workloads. Where dual-socket, low-core count systems were the de-facto standard 5-10 years ago, today's many-core, single-socket systems offer more than enough performance and scalability to be a very strong alternative option—often allowing for significant energy savings versus a 2-socket system.

The cloud runs on single socket. In their relentless quest for growing capacity while controlling energy use, many of the world's largest cloud providers have shifted to single socket AMD EPYC architectures for their general-purpose fleets. One example of this transition is Cloudflare's Gen X server platform. By switching from dual-socket legacy architectures to a single-socket AMD EPYC™ 7642, Cloudflare reported a 36% increase in request capacity and a 50% reduction in NGINX latency, all while lowering power consumption per core by 25%. Hyperscale leaders like AWS also have leveraged AMD EPYC to redefine their general-purpose fleets. The Amazon EC2 M7a instances, powered by 4th Gen EPYC Server CPUs, provide up to 50% more compute performance and 2.25x the memory bandwidth of prior generations. By utilizing high-density 1:1 core mapping, AWS enables customers to consolidate demanding workloads onto fewer instances, maximizing rack density and energy efficiency.

- No-Compromise Single Socket: Unlike legacy single-socket chips, which crippled I/O and offered more limited memory expansion, AMD EPYC supports all 128 PCIe lanes, full memory bandwidth and up to 3TB memory capacity on a single chip-based system. This allows you to cut your physical CPU count in half—slashing per-socket licensing costs—without sacrificing connectivity.

Section 2 - The "Pre-Flight" Checklist

Trust is built on transparency. Here is the friction you may encounter, and the specific tools to mitigate it.

The Elephant in the Room: Live Migration Compatibility

Let's be clear: Live migration (vMotion) isn't guaranteed across dissimilar CPUs. Cross-vendor moves (Intel to AMD, or AMD to Intel) can't live-migrate active RAM because the CPU identity and feature exposure to the guest OS differ—so the VM must be powered down, registered on the new host, and restarted (a cold migration). And this isn't unique to switching between processor manufacturers: you can hit the same requirement when moving from a legacy Intel platform to a newer Intel host (or legacy AMD to much newer AMD) if the compatibility baseline can't be maintained.

- The Tool: VMware Architecture Migration Tool (VAMT). This automated tool handles the cold migration workflow—scripting the shutdown, register, and startup process—turning a manual headache into a batch process.
- Other Routes:
  - Nutanix: Use Nutanix Move.
  - Azure: Use Azure Migrate.
  - HPE: Use HPE Morpheus VM Essentials.
  - Red Hat: Use the Migration Toolkit for Virtualization (MTV).
  - Partner-Led Migrations: Major OEMs, integrators, and VARs offer proven services to plan, execute, and validate these moves—often bundling assessment, migration factory workflows, and post-move optimization to reduce downtime risk and accelerate time-to-value.

Validation & Day 2 Operations

Before moving a single production VM, ensure your operating model is ready for a mixed fleet. In most real-world deployments, servers with CPUs from Intel and AMD will coexist for a period of time (sometimes alongside other platforms), and that's normal. With the right baselines, tooling, and validation, day-to-day operations can remain consistent even as the underlying hosts diversify.

- BIOS Strategy: While the OS is largely hardware agnostic, BIOS tuning matters.
  Understand the difference between NPS1 (General Compute) and NPS4 (HPC/NUMA optimized) settings.
- Configuration Mapping: Consult the AMD Tuning Guides to map these settings
  to your specific OEM (Dell/HPE/Lenovo) profiles.
- Monitoring: Verify that your agents (e.g. Datadog, Prometheus, SolarWinds)
  are updated to recognize AMD telemetry, including thermal sensors.

Drivers: Modern Linux kernels (RHEL, Ubuntu) and Windows Server builds include in-box drivers for current AMD EPYC Server CPUs. The old “driver hunt” ritual formerly involved in upgrades is largely a thing of the past—still, confirm your Standard Operating Environment (SOE) image is current. Once the tooling, monitoring, and operational baselines are in place, the remaining question is sequencing: what to move first, what to leave for later, and how to keep risk contained while the fleet becomes mixed. That’s where a phased execution plan is essential.

Operational Foundations for a Mixed Fleet

Phase 1 should establish the repeatable operating pattern you’ll reuse for every later phase. In most environments, you’ll run Intel and AMD side-by-side for a while. The goal is to use the early phase to make that coexistence predictable: maintaining consistent monitoring, standardizing failover behavior, and codifying a playbook your team can execute flawlessly.

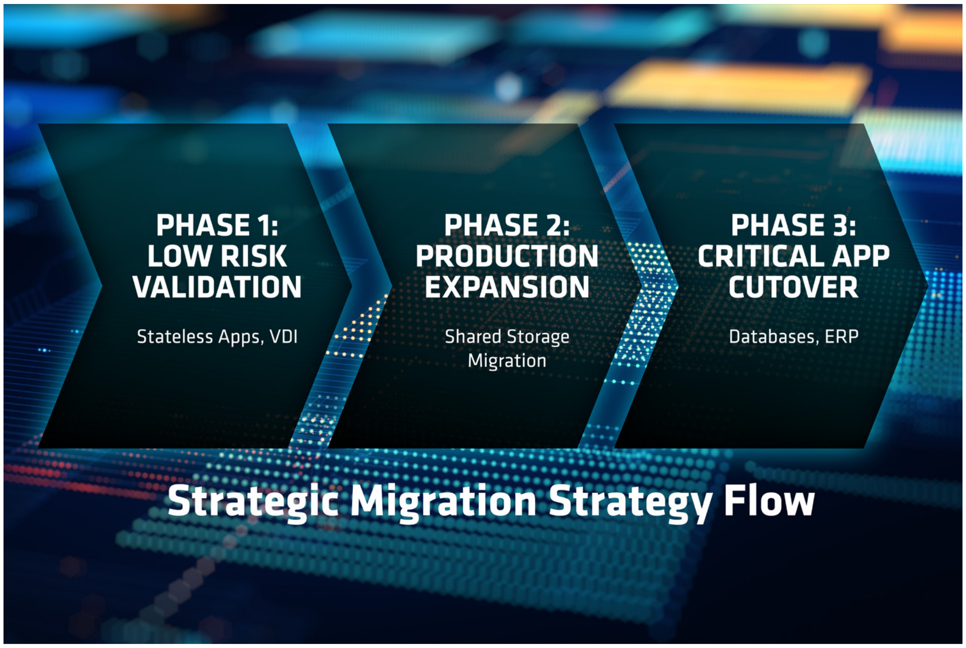
Use Phase 1 to establish the operational blueprint you will scale across the rest of the organization.

- Failover and recovery patterns: Validate what “good” looks like for your environment—cluster failover behavior, restart sequencing, maintenance windows, rollback steps, and how long critical workloads takes to return to service.
- Monitoring baselines: Treat Phase 1 as the moment you standardize dashboards, alerts, and thresholds across vendors. Confirm that host telemetry, VM performance counters, and capacity signals behave as expected, and update operational procedures so on-call responders have validated, vendor-specific procedures.
- Internal knowledge base: Capture technical specifications during this initial pilot. While operational procedures provide instructions for action, this knowledge base serves as a configuration reference, documenting BIOS profiles, firmware/driver versions, and “known-good” settings to guide later phases.
- Automation and standardization: Once validated, bake the new reality into your deployment system—golden images, configuration management, patching cadence, and compliance checks—so later phases feel like scaling a process, not repeating an experiment.
- Facilities and capacity planning: Consolidation changes the shape of your data center. Fewer servers can mean different power provisioning requests, different rack densities, different cooling assumptions, and newly available headroom. Use Phase 1 results to update power/cooling models, rack layouts, and expansion plans so later phases don’t get blocked by facilities constraints.
- Organizational impact: If consolidation reduces hardware sprawl, operational capacity often increases. Plan for what you’ll do with it—accelerate refresh timelines, reallocate staff to higher-value work, or repurpose reclaimed power and space for new initiatives (including AI pilots).

With these foundations in place, the execution becomes a phased rollout rather than a one-time migration event.

Section 3 - Phases of the Switch: A Strategic Execution Plan

Modernization is not all-or-nothing. Whether your strategy calls for a complete rip and replace to maximize immediate consolidation, or an expansion of your fleet alongside legacy infrastructure, the migration path is fully mapped. You can move at the speed of your budget, resource availability and risk tolerance.



Phase 1: Low Risk (The "Canary in the Coal Mine")

Start with workloads that, if they fail, do not stop business operations. This builds operational muscle memory for your team without risking revenue, while exposing potential issues before they become problems with more critical applications.

- Candidates: Low-risk containerized apps (Docker/Kubernetes), stateless web apps, internal test/dev environments, VDI (Virtual Desktop Infrastructure).
- Cloud Strategy ("Try Before You Buy"): Use Azure Migrate to lift-and-shift a low-priority on-prem workload to an AMD based Azure instance (e.g., AWS M8a or Azure Dasv5). This allows you to validate application behavior with zero hardware investment.
- Resources: Cloud Migration Guide

Phase 2: Medium Risk (Production Expansion)

This is where the bulk of your workloads live, and you can make the move as your resources and business allow.

- Strategy – Shared Storage Migration provides a relatively simple option: You do not need to move the data, only the compute. Connect the new AMD rack to your existing SAN/NAS. Point the VMs to the new compute nodes. Reboot.
- Case Study (Dell Technologies): In a demonstration of a large-scale migration to Dell PowerEdge servers with AMD EPYC Server CPUs, Dell showed the ability to migrate 380 Virtual Machines in less than one hour.
  - Note: The downtime per VM was only minutes (the time to reboot). The hour was the total time to execute the batch.
- Third-Party Testing (Principled Technologies): In a third-party evaluation of the VMware Architecture Migration Tool (VAMT), testing demonstrated simple, predictable, and reversible migration from legacy servers to the newest AMD EPYC 9175F powered Dell PowerEdge servers in less than 15 minutes.

- **Load Balancing:** For web servers, keep the old Intel cluster running while spinning up the new AMD cluster behind a load balancer. Gradually shift traffic (10%, 20%, 50%, etc) to the new nodes and retire the old systems as utilization drops. No downtime is perceived by the user.

Phase 3: Critical Apps (The "Crown Jewels")

These are your high-value databases (Oracle, SQL) and ERP systems.

- Strategy: High planning, surgical execution. This level of rigor is required for any critical application migration—regardless of CPU vendor—because the real risk lives in data integrity, dependencies, and cutover sequencing.
- Validation: Build a sandbox that mirrors production as closely as practical (data volume characteristics, interfaces, auth, backup/restore paths). Run at least one full dress rehearsal: timing, rollback, and a documented go/no-go checklist.
- Tools: Leverage best knowledge and best tools to facilitate best results. Use vendor-specific tools like Oracle GoldenGate or Microsoft's Database Migration Guides (e.g., Oracle to SQL) to handle data consistency and minimize cutover risk.
- The Payoff: By this phase, you are moving your most resource-hungry apps to your most capable hardware —where consolidation and per-core/per-socket licensing efficiency benefits compound.

Section 4 - Deployment Models

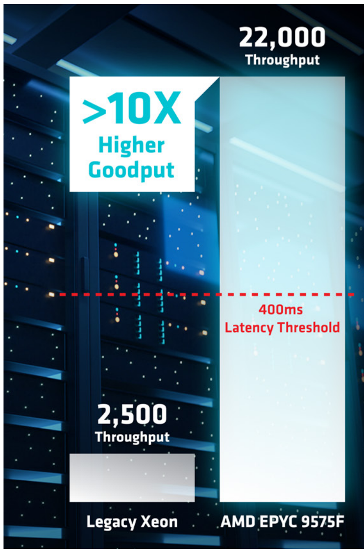
The AMD EPYC portfolio provides a consistent hardware foundation that simplifies how enterprises deploy workloads across diverse environments, from data centers to the public cloud.

1. On-Premises Refresh

- Strategy: Replace aging Xeon fleets (3-5 years old) to reclaim rack space.
- The "Seed Unit" Program: It is not just Cloud that has "Try before you buy." Customers can work with their trusted OEM partners (Dell, HPE, Lenovo) to secure seed units. These are pilot servers shipped to your facility, allowing you to run your actual production clones on AMD EPYC Server CPU-based systems before signing a purchase order.

2. Hybrid Cloud

- Strategy: Use Spot Instances of current-generation AMD processor-based servers in AWS/Azure to run low-cost validations, further de-risking the migration.



Challenge the GPU Monopoly

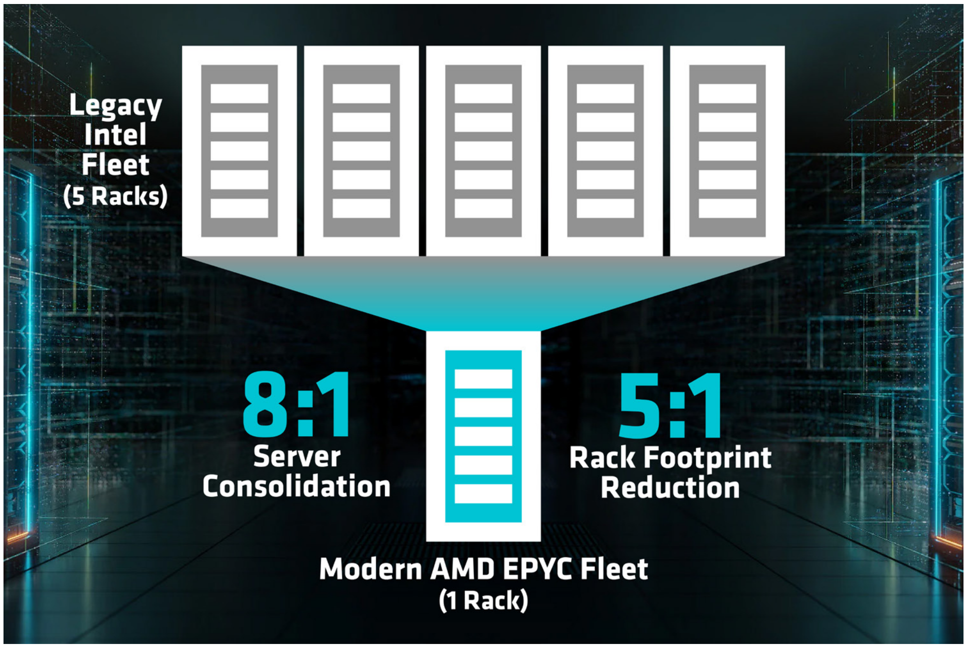
A pervasive myth is that all AI workloads require expensive, power-hungry GPUs. This leads to "AI FOMO" and rapid budget exhaustion. The strategic pivot to AMD EPYC Server CPU systems reveals that the CPU is a highly capable inference engine.

- The "Goodput" Metric: In rigorous tests using the Llama-3.3-70B model, the AMD EPYC 9575F host CPU delivered[i] over 10 times the "goodput" (throughput under a specific 400ms time-to-first-token constraint) compared to competitive Intel Xeon processors.
- Open Source Optimization: When utilizing vLLM optimization, AMD EPYC processors achieved a 1.56x uplift[ii] in token throughput compared to Intel Xeon, making them viable for real-time applications.
- Strategy: Use your AMD fleet for business apps during the day and batch AI inference processing at night or other off-peak periods.

Section 5 - The Business Case: Financial Ammunition

To get the CIO and the CFO to sign off on the perceived friction of a migration, you need to show them the "layer cake" of savings. It is not just cheaper or even "faster" hardware; it is compounding efficiency.

Layer 1: The "Density Dividend" – High-performance density allows organizations to replace sprawling legacy fleets with compact, efficient clusters.



- The Math: A single AMD EPYC 9005 CPU-based server can perform the work of more than eight 2019-era Intel Xeon Platinum servers.[iii]
- The Impact: Replacing eight legacy servers with one modern AMD EPYC server results in a 5:1 physical footprint reduction. This consolidation liberates expensive floor space, reduces colocation fees, and simplifies cabling and switching It also reshapes your power profile: overall facility draw may drop, but watts per rack often rise—so validate rack-level power delivery (PDUs/breakers/UPS) and cooling as you increase density. Finally, as utilities and regulators tighten requirements for large loads (new rate classes and, in some markets, emergency curtailment or even onsite generation/storage expectations), reducing total load and improving efficiency can lower both cost and operational friction.

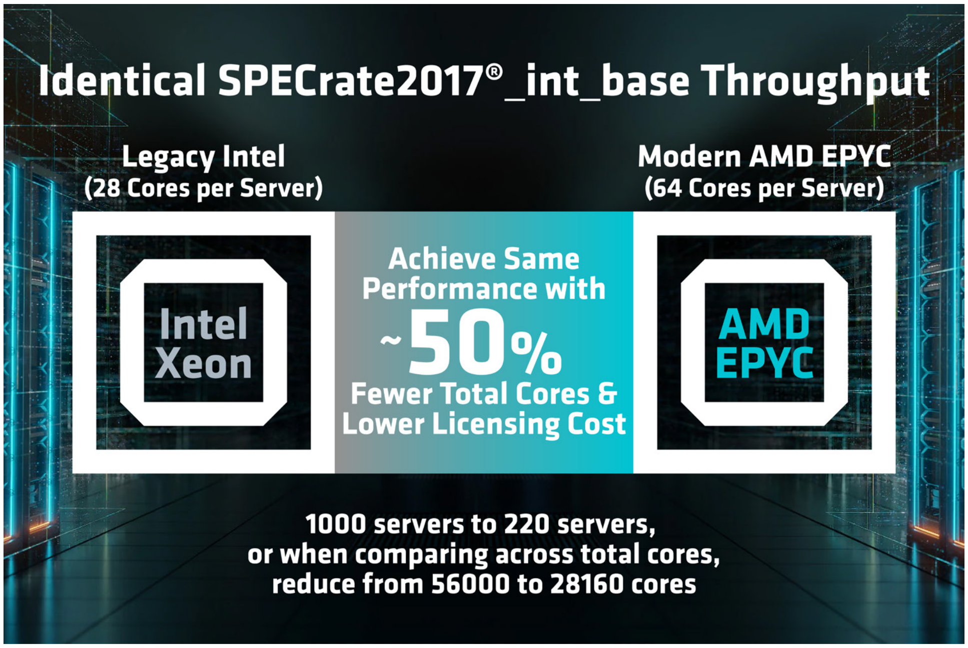
TABLE 1: SERVER FLEET COMPARISON

METRIC	LEGACY INTEL FLEET (EST.)	MODERN AMD EPYC FLEET (EST.)	STRATEGIC IMPACT
SERVER COUNT	1,000 Nodes (2P Xeon 8280)	122 Nodes (2P EPYC 9965)	**88% Reduction** in hardware management overhead.
POWER DRAW	High Baseline	70% Lower	**$2.5M Savings** (3-Year Energy Bill).
PERFORMANCE	391k Integer Units	391k Integer Units	**Parity** with fractional footprint.
SPACE	Multiple Rows (72x 8.6kw Racks)	25x 8.6kw Racks	Massive Real Estate Savings / Colocation Fee Reduction.

Layer 2: Cloud Economics and Infrastructure Benchmarking

- Market-Based Value: Even for organizations primarily operating on-premises, public cloud pricing serves as a transparent benchmark for hardware value and efficiency. On average, like-for-like x86 instances equipped with AMD EPYC Server CPUs—within the same cloud, region, and purchase model—are priced approximately 10% lower than comparable Intel instances across several major families.
- Operational Efficiency: For hybrid or cloud-native workloads, this lower hourly rate combined with high per-vCPU performance directly reduces the cost per unit of work, such as cost per transaction or per job. Because actual results vary by specific deployment, organizations should benchmark their applications to ensure optimal right-sizing.

Layer 3: The Licensing "Killer App"



Most enterprise software (VMware, SQL Server) charges per core or per socket.

- The Loophole: If you can achieve the same general purpose performance using half as many AMD cores instead of Intel cores, you have just cut your software licensing bill by 50%.[iv]

In the same scenario, compared to the newest 64C Intel Xeon 6767P powered servers, the AMD EPYC server has up to 23% lower software virtualization licensing costs. These savings often pay for the hardware migration labor in the first year.

Section 6 - Conclusion: Future-Proofing Your Fleet

The decision to switch to AMD is no longer about taking a risk; it is about maximizing the usefulness and return on your investment. Staying with legacy infrastructure translates to higher OpEx, higher licensing fees, and a larger carbon footprint – and it increases the likelihood that you won't have the ability to say "yes" when the business needs a new project.

Security by Design

Long-term viability also means security. AMD Infinity Guard[v] embeds security at the silicon level. AMD SEV is part of the AMD Infinity Guard security suite, which provides a portfolio of technologies to address different deployment models and threat vectors, including:

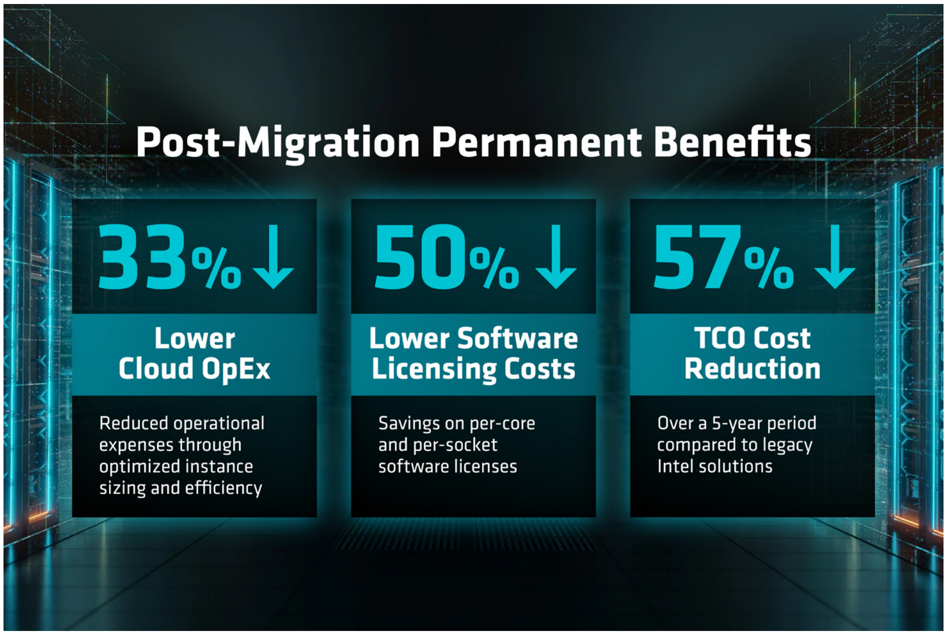
- Transparent Secure Memory Encryption: Encrypts system memory with a single key.
- Encrypted State: Encrypts CPU register contents to prevent leakage on VM context switches.
- Secure Nested Paging: Adds memory integrity protection to block attacks such as data replay or remapping.
- Trusted I/O: Extends the TEE to PCIe devices (NICs, accelerators, storage) using the PCI-SIG defined TDISP standard.

Together, these capabilities make SEV the most mature and widely deployed confidential computing technology in the industry,[vi] protecting workloads across leading cloud providers and enterprise deployments.

The Final Verdict

The migration requires a cold restart, yes. But that is a temporary maintenance event and you are well versed in such processes. The benefits are more permanent:

- 33% lower Cloud OpEx[vii]
- 50% lower VMware licensing[iv]
- 57% TCO Reduction over a 5 year period compared to legacy Intel solutions or 26% TCO reduction compared to modern Intel solutions[iv]



Next Steps

1. Don't Guess, Validate: Ask your VAR/OEM for a seed unit to test one critical app on-prem.
2. Audit Your Licenses: Calculate how much you would save by reducing your core count by 30% while maintaining performance.
3. Start Phase 1: Spin up a single pilot instance on Azure/AWS/Google Cloud today.

DISCLAIMERS

[i] **9xx5-169:** Llama-3.3-70B latency constrained throughput (goodput) results based on AMD internal testing as of 05/14/2025. Configurations: Llama-3.3-70B, vLLM API server v1.0, data set: Sonnet3.5-SlimOrcaDedupCleaned, TP8, 512 max requests (dynamic batching), latency constrained time to first token (300ms, 400ms, 500ms, 600ms), OpenMP 128, results in tokens/s 2P AMD EPYC 9575F (128 Total Cores, 400W TDP, production system, 1.5TB 24x64GB DDR5-6400 running at 6000 MT/s, 2 x 25 GbE ConnectX-6 Lx MT2894, 4x 3.84TB Samsung MZWLO3T8HCLS-00A07 NVMe ; Micron_7450_MTFDKCC800TFS 800GB NVMe for OS, Ubuntu 22.04.3 LTS, kernel=5.15.0-117-generic , BIOS 3.2, SMT=OFF, Determinism=power, mitigations=off) with 8x NVIDIA H100 2P Intel Xeon 8592+ (128 Total Cores, 350W TDP, production system, 1TB 16x64GB DDR5-5600 , 2 x 25 GbE ConnectX-6 Lx (MT2894), 4x 3.84TB Samsung MZWLO3T8HCLS-00A07 NVMe, Micron_7450_MTFDKBA480TFR 480GB NVMe , Ubuntu 22.04.3 LTS, kernel-5.15.0-118-generic , SMT=OFF, Performance Bias, Mitigations=off) with 8 x NVIDIA H100. Results:CPU 300 400 500 600; 8592+ 0 126.43 1565.65 1987.19; 9575F 346.11 2326.21; 2531.38 2572.42; Relative NA 18.40 1.62 1.29. Results may vary due to factors including system configurations, software versions, and BIOS settings. TDP information from ark.intel.com

[ii] **9xx5-255:** Llama3.1-8B throughput results based on AMD internal testing as of 10/16/2025. Configurations: llama3.1-8B, vLLM, python 3.10, TPOT max 100ms, BF16, input/output lengths: [128/256, 256/512, 1024/1024] 1P AMD EPYC 9555 (64 Total Cores, reference system, 768GB 12x128GB DDR5-6400, BIOS RVOT1003C, Ubuntu® 22.04 | 6.8.0-84-generic, SMT=OFF, Determinism=power, Mitigations=off), vLLM 0.9.2, ZenDNN 5.1.0, NPS1 1P Intel Xeon 6767P (64 Total Cores, production system, 512GB 8x64GB MRDIMM at 8000 MT/s, BIOS IHE110U-1.20 (AMX on), Ubuntu 24.04 | 6.8.0-84-generic, SMT=OFF, Performance Bias, Mitigations=off), vLLM 0.9.1, IPEX 2.7, NPS1, SNC OFF Results: input output prompts6767P throughput6767P prompts9655 throughput9655 128 256 60 435.98 76 680.58 256 512 51 419.63 68 586.18 1024 1024 30 249.53 43 378.03 Results may vary due to factors including system configurations, software versions, and BIOS settings.

[iii] **9xx5TCO-018:** This scenario contains many assumptions and estimates and, while based on AMD internal research and best approximations, should be considered an example for information purposes only, and not used as a basis for decision making over actual testing. The AMD Server & Greenhouse Gas Emissions TCO (total cost of ownership) Estimator Tool - version 1.53, compares the selected AMD EPYC™ and Intel® Xeon® CPU based server solutions required to deliver a TOTAL_PERFORMANCE of 391,000 units of SPECrate2017_int_base performance as of December 3, 2025. This analysis compares a 2P AMD 192 core EPYC_9965 powered server with a SPECrate2017_int_base score of 3230, https://spec.org/cpu2017/results/res2025q2/cpu2017-20250324-47086.pdf; compared to a 2P Intel Xeon 128 core Xeon_6980P based server with a SPECrate2017_int_base score of 2510, https://spec.org/cpu2017/results/res2025q2/cpu2017-20250324-47099.pdf; versus legacy 2P Intel Xeon 28 core Platinum_8280 based server with a SPECrate2017_int_base score of 391, https://spec.org/cpu2017/results/res2020q3/cpu2017-20200915-23984.pdf Environmental impact estimates made leveraging data from the 2025 International Country Specific Electricity Factors and can be found at https://www.carbondi.com/#electricity-factors/ and the US EPA Greenhouse Gas Equivalencies Calculator used in this analysis was sourced on 09/04/2024 and can be found at https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator.. For additional details, see https://www.amd.com/en/legal/claims/epyc.html#q=9xx5TCO-018.

[iv] **9xx5TCO-020:** This scenario contains many assumptions and estimates and, while based on AMD internal research and best approximations, should be considered an example for information purposes only, and not used as a basis for decision making over actual testing. The AMD Server & Greenhouse Gas Emissions TCO (total cost of ownership) Estimator Tool - version 1.55, compares the selected AMD EPYC™ and Intel® Xeon® CPU based server solutions required to deliver a TOTAL_PERFORMANCE of 391000 units of SPECrate®2017_int_base performance as of December 3, 2025. This analysis compares a 2P AMD 64 core EPYC_9555 powered server with a SPECrate2017_int_base score of 1780, https://spec.org/cpu2017/results/res2025q4/cpu2017-20251102-50212.pdf; compared to a 2P Intel Xeon 64 core Xeon_6767P based server with a SPECrate2017_int_base score of 1370, https://spec.org/cpu2017/results/res2025q3/cpu2017-20250825-49383.pdf; versus legacy 2P Intel Xeon 28 core Platinum_8280 based server with a SPECrate2017_int_base score of 391, https://spec.org/cpu2017/results/res2020q3/cpu2017-20200915-23984.pdf Environmental impact estimates made leveraging data from the 2025 International Country Specific Electricity Factors and can be found at https://www.carbondi.com/#electricity-factors/ and the US EPA Greenhouse Gas Equivalencies Calculator used in this analysis was sourced on 09/04/2024 and can be found at https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator.. For additional details, see https://www.amd.com/en/legal/claims/epyc.html#q=9xx5TCO-020.

[v] **EPYC-056:** Confidential Computing on EPYC is enabled by the SEV security feature, which was introduced with 1st Generation EPYC in 2017. 2nd Gen EPYC powered the first confidential computing cloud instance in Google Cloud in 2020. EPYC: powers the highest number of confidential VM options available on all major CSP; Supports both host and guest in the Linux® Kernel; Is available on all major Linux Distributions; Has support on VMware®; supports confidential containers.

[vi] **GD-183A:** AMD Infinity Guard features vary by EPYC™ Processor generations and/or series. Infinity Guard security features must be enabled by server OEMs and/or Cloud Service Providers to operate. Check with your OEM or provider to confirm support of these features. Learn more about Infinity Guard at https://www.amd.com/en/technologies/infinity-guard

[vii] **9xx5C-062:** Testing by AMD Performance Labs as of 10/25/2025. M8a.16xlarge score comparison to M8i-16xlarge running the following benchmarks: - SPECrate®2017_int_base - SPECrate®2017_fp_base - Server-side Java multi-instance max - Server-side Java composite max Performance differences (normalized to M8i): M8A.16XL 64 vCPU vs M8i.16XL 64 vCPU Perf Perf/$ OpEx Savings SPECrate®2017_int_base 1.78 est 1.55 est 36% SPECrate®2017_fp_base 2.36 est 2.06 est 51% Server-side Java multi-instance max 1.48 1.29 22% Server-side Java composite max with 1.52 1.32 24% Top 4 Workload Benchmarks 1.78 1.55 33% On-demand hourly pricing from https://aws.amazon.com/ec2/pricing/on-demand/ (us-east) as of 10/25/2025: M8A.16XL: $3.895, M8i.16XL $3.386 Cloud performance results presented are based on the test date in the configuration. Results may vary due to changes to the underlying configuration, and other conditions such as the placement of the VM and its resources, optimizations by the cloud service provider, accessed cloud regions, co-tenants, and the types of other workloads exercised at the same time on the system. Cloud OpEx savings calculated from relative runtime and instance cost ratio. Java is a registered trademark of Oracle and/or its affiliates. SPEC® and SPECrate® are registered trademarks of Standard Performance Evaluation Corporation. Learn more at SPEC.org.