Market Publishers
+357 96 030922
info@marketpublishers.com

The Global Market for High Performance Computing (HPC) and AI Accelerators 2025-2035

https://marketpublishers.com/r/GF412D239F1CEN.html
Date: April 2025
Pages: 920
Price: US$ 2,000.00 (Single User License)
ID: GF412D239F1CEN

Abstracts

The high-performance computing (HPC) and AI accelerator market is experiencing unprecedented growth, driven primarily by the surge in generative AI applications across industries. This sector has transformed from a specialized niche into a cornerstone of modern computing infrastructure, with data center processors forming the backbone of this revolution. The global data center processor market neared $150 billion in 2024 and is projected to expand dramatically to >$370 billion by 2030, with continued growth expected to push the market well beyond $500 billion by 2035. This growth trajectory is primarily fuelled by specialized hardware designed to handle the massive computational demands of AI workloads. Graphics Processing Units (GPUs) and AI Application-Specific Integrated Circuits (ASICs) have emerged as the dominant forces in this landscape, experiencing double-digit growth as they power the most demanding generative AI systems.

While traditional Central Processing Units (CPUs) and networking processors like Data Processing Units (DPUs) continue to play essential roles in data center infrastructure with steady growth, they no longer represent the cutting edge of AI computation. Field-Programmable Gate Arrays (FPGAs), once considered promising for AI applications, have seen a significant decline and are expected to remain flat through 2035 as purpose-built AI accelerators have proven more efficient for specific workloads.

The competitive landscape has shifted dramatically since OpenAI's breakthrough innovations in 2022. Nvidia has established clear market dominance with its advanced GPU offerings, particularly the Hopper (H100/H200) and newer Blackwell (B200/B300) architectures. These chips incorporate cutting-edge features like increased on-chip memory capacity—over 250GB of high-bandwidth memory (HBM)—enabling larger AI models with more parameters.

However, major cloud service providers like Google and AWS are pursuing strategic independence through partnerships with Broadcom, Marvell, and Alchip to co-design custom AI ASICs. These systolic array-based custom chips offer advantages over GPUs, including lower total cost of ownership, reduced vendor lock-in risk, and specialized optimization for specific workloads like transformers and recommender systems.

The market has also attracted numerous innovative startups such as Cerebras, Groq, Graphcore, SambaNova, and Untether AI, who are pioneering novel architectures including dataflow-controlled processors, wafer-scale packaging, spatial AI accelerators, and processing-in-memory technologies. This innovation wave has triggered significant merger and acquisition activity as established players seek to incorporate cutting-edge technologies. Technology trends driving this market include the shift toward multi-chiplet architectures, which optimize manufacturing yield while enabling larger dies, and the rapid adoption of advanced process nodes. Current leading-edge CPUs utilize 3nm technology, while GPUs and AI ASICs typically employ 4nm processes, with 3nm expected to arrive as early as 2025 in products like AWS Trainium 3, and sub-1nm nodes projected to emerge by 2035.

Compute performance has grown eightfold since 2020, with ambitious roadmaps like Nvidia's Rubin Ultra targeting 100 PetaFLOPs in FP4 for inference by 2027. By 2035, industry projections suggest that leading AI processors could deliver exascale performance in compact form factors, representing a thousand-fold increase over today's capabilities. Memory technologies have become increasingly critical as AI models expand, with High-Bandwidth Memory (HBM) currently serving as the standard for high-performance AI systems, though several startups are exploring SRAM-based alternatives to further improve performance.

The industry is also witnessing architectural shifts, with Arm-based CPUs gaining momentum against the traditional x86 architecture dominated by Intel and AMD. Meanwhile, cryptocurrency mining operations, with their expertise in cooling solutions and high-power infrastructure, are diversifying into AI by hosting powerful GPU clusters. Looking ahead to 2035, the market is expected to maintain a compound annual growth rate of 10-12% between 2030-2035, with revenues likely exceeding $800 billion as frontier AI development continues to drive demand for exceptional volumes of specialized chips within AI data centers worldwide, and as these technologies become increasingly embedded in critical infrastructure across all sectors of the global economy.

+357 96 030922
info@marketpublishers.com

The Global Market for High Performance Computing (HPC) and AI Accelerators 2025-2035 provides an in-depth analysis of the rapidly evolving high-performance computing landscape, with particular focus on the transformative impact of artificial intelligence technologies. This comprehensive report examines market dynamics, technological advancements, competitive strategies, and future trends that will shape this critical sector over the next decade. With detailed revenue forecasts from 2025 to 2035, the report offers essential intelligence for investors, technology providers, data center operators, and enterprise decision-makers navigating the complex intersection of traditional high-performance computing and cutting-edge AI acceleration technologies. Contents include:

Market Size and Growth Projections (2025-2035): Detailed forecasts for AI chips, GPUs, CPUs, AI ASICs, DPUs, network ASICs, and crypto ASICs, covering both shipments and revenues

Key Technology Inflection Points: Analysis of next-generation node transitions, advanced packaging technologies, and memory system innovations

Strategic Market Drivers and Challenges: Deep dive into generative AI computing requirements, energy efficiency imperatives, and supply chain vulnerabilities with mitigation strategies

Investment Outlook and Opportunities: Examination of high-growth market segments and emerging technology areas

Introduction to HPC and AI: Evolution from historical HPC systems to the exascale computing era, supercomputer vs. hyperscale data center comparisons, and AI computing fundamentals

Processor Technologies and Architectures: Comprehensive analysis of CPUs (x86, ARM, RISC-V), GPUs, AI ASICs, FPGAs, DPUs, and cryptocurrency computing hardware

Enabling Technologies: Detailed examination of advanced semiconductor manufacturing, packaging technologies, memory innovations, cooling solutions, networking advancements, and storage technologies

Market Analysis and Forecasts: Segment-specific projections for GPUs, AI ASICs, CPUs, FPGAs, and DPUs, with end-user segment analysis across cloud

+357 96 030922
info@marketpublishers.com

providers, HPC centers, enterprises, and telecommunications

Supply Chain Analysis: Deep dive into semiconductor manufacturing, advanced packaging services, memory suppliers, cooling solution providers, power delivery components, system integrators, and supply chain vulnerabilities

Technology and Market Trends: Performance scaling trajectories, memory bandwidth challenges, software ecosystem developments, and energy efficiency initiatives

Application Segments and Use Cases: Specialized infrastructure requirements for AI training, inference deployment, traditional HPC applications, cloud provider infrastructure, and enterprise computing needs

Company Profiles: Detailed analysis of over 200 companies spanning semiconductor manufacturers, AI chip startups, cloud service providers, and system integrators with tables of other companies in the supply chain also. Companies covered Accelsius, Achronix, Advanced Micro Devices (AMD), AheadComputing, AiM Future, Aistorm, AI21labs, Ambient Scientific, Amlogic, Ampere Computing, Anaflash, Analog Inference, Apple, AONdevices, Arm, Astrus, Atos, Amazon Web Services (AWS), Axelera AI, Axera Semiconductor, Azure Engine, Baidu, Baya Systems, Biren Technology, Bitmain, Blumind, Brainchip Holdings, ByteDance, Cambricon Technologies, Canaan, Celestial AI, Cerebras, Ceremorphic, CIX Technology, Clouder, Cognifiber, Cohere, Corerain Technologies, Corigine, CoreWeave, Cornami, DeepL DeepSeek, Deepx, Deezer, DeGirum, Denglin Technology, Digital Reality, d-Matrix, Eeasy Technology, EdgeCortix, Efinix, EnCharge AI, Enflame, Equinix, Epic Semiconductors, Esperanto Technologies, Etched, Eviden, Evomotion, Expedera, Flex Logix, Fulhan, Fujitsu, Fungible, Furiosa, GlobalFoundries, GigaByte, Google, Gowin, GrAI Matter Labs, Graphcore, GreenWaves Technologies, Groq, GUC, Guoxin Micro, Gwanak Analog, Gyrfalcon Technology, Habana, Hailo, HiSilicon, Hitachi, Hewlett Packard Enterprise, Horizon Robotics, Houmo.ai, HjMicro, Huawei, Hygon, IBM, Iluvatar CoreX, Icubecorp, Inflection AI, Innatera Nanosystems, Innosilicon, Intel, Inventec, Intellifusion, Intelligent Hardware Korea (IHWK), Inuitive, InspireSemi, iPronics, Jingjia Micro, Kalray, Kinara, Kneron, Knuedge, Krutrim, Kunlunxin Technology, Lattice Semiconductor, Lightelligence, Lightmatter, LiSuan Tec, Loongson Technology, Luminous Computing, Lynxi, Marvell Technology, MediaTek, Mellanox, MemryX, Meta, Metax-tech, Microsoft, Mistral AI, Mobilint, Modular,

Moffett AI, Moonshot AI, Moore Threads, Mythic, Nano-Core Chip, NebulaMatrix, Neuchips, Neuroblade, Neureality, Netronome, Nextchip, NTT Communications, Nuovoton, Nuvia, Nvidia, NXP, OpenAI, Oracle, Optalysys, Panmnesia, Penguin Computing, Pensando, Perceive, Pezy Computing, Phytium, Positron, PyTorch, Qilingxin, Quadric, Quanta Cloud Technology, Quanta Computer, Qualcomm, Quillion, Rackspace, Rain, Rapidus, Rebellions, Recogni, Renesas, Resnics, Retym, Rivai, Rockchip, Roviero, Salience Labs, SambaNova, Samsung and more......

The high-performance computing and AI accelerator market is experiencing unprecedented transformation driven by the explosive growth of generative AI, increasingly complex computational workloads, and technological innovations across the computing stack. This report provides essential insights into market dynamics, competitive positioning, and strategic opportunities that will define success in this rapidly evolving landscape. From cutting-edge semiconductor technologies to novel architectures and deployment strategies, this comprehensive analysis equips stakeholders with the knowledge needed to navigate the complex interplay of technical capabilities, market demands, and competitive pressures that characterize this vital sector.

+357 96 030922
info@marketpublishers.com

Contents

1 EXECUTIVE SUMMARY

1.1 Market Overview and Key Findings
1.1.1 Critical Market Shifts in Data Center Computing
1.1.2 Convergence of AI and Traditional HPC Workloads
1.1.3 Impact of Generative AI on Infrastructure Requirements
1.2 Market Size and Growth Projections (2025-2035)
1.2.1 AI Chips
1.2.1.1 Shipments
1.2.1.2 Revenues
1.2.2 Graphics processing units (GPUs)
1.2.2.1 Shipments
1.2.2.2 Revenues
1.2.3 Central processing units (CPUs)
1.2.3.1 Shipments
1.2.3.2 Revenues
1.2.4 AI ASICs
1.2.4.1 Shipments
1.2.4.2 Revenues
1.2.5 DPU
1.2.5.1 Shipments
1.2.5.2 Revenues
1.2.6 Network ASIC
1.2.6.1 Shipments
1.2.6.2 Revenues
1.2.7 Crypto ASIC
1.2.7.1 Shipments
1.2.7.2 Revenues
1.3 Key Technology Inflection Points
1.3.1 Next-Generation Node Transitions
1.3.2 Advanced Packaging Technologies
1.3.3 Memory System Innovations
1.4 Strategic Market Drivers and Challenges
1.4.1 Generative AI Computing Requirements
1.4.2 Energy Efficiency Imperatives
1.4.3 Supply Chain Vulnerabilities and Mitigations
1.5 Investment Outlook and Opportunities

+357 96 030922
info@marketpublishers.com

1.5.1 High-Growth Market Segments
1.5.2 Emerging Technology Areas

2 INTRODUCTION TO HIGH-PERFORMANCE COMPUTING AND AI

2.1 Defining High-Performance Computing (HPC)
2.1.1 Historical Evolution of HPC Systems
2.1.2 The Exascale Computing Era
2.1.3 TOP500 Analysis and Performance Metrics
2.1.4 Supercomputers vs. Hyperscale Data Centers
2.2 HPC Architectures and Infrastructures
2.2.1 Distributed Computing Models
2.2.2 On-Premises Deployments and Dedicated Infrastructure
2.2.3 Cloud-Based HPC Services (HPC-as-a-Service)
2.2.4 Hybrid and Multi-Cloud Approaches
2.2.5 Edge-HPC Integration Frameworks
2.3 Artificial Intelligence Computing Fundamentals
2.3.1 AI Algorithms and Computing Requirements
2.3.1.1 Deep Learning Architectures
2.3.1.2 Transformer Models and Attention Mechanisms
2.3.1.3 Reinforcement Learning Approaches
2.3.2 Training vs. Inference Workload Profiles
2.3.2.1 Training Infrastructure Requirements
2.3.2.2 Inference Optimization Strategies
2.3.2.3 Batch vs. Real-Time Processing
2.3.3 Precision Requirements for AI Computing
2.3.3.1 Numerical Formats (FP32, FP16, BF16, INT8)
2.3.3.2 Mixed Precision and Quantization Approaches
2.3.3.3 New Formats (FP8, FP4) and Implications
2.4 Large AI Models and Computing Requirements
2.4.1 Evolution of Model Scale and Complexity
2.4.1.1 Parameter Count Progression
2.4.1.2 Compute Requirements Scaling Trends
2.4.1.3 Memory Footprint Challenges
2.4.2 Language Models (GPT, LLaMA, Claude, Gemini)
2.4.3 Multimodal Models (Text, Image, Audio)
2.4.4 Domain-Specific AI Models
2.5 Market Convergence: HPC and AI Computing
2.5.1 Overlapping Hardware Requirements

+357 96 030922
info@marketpublishers.com

2.5.2 Shared Software Ecosystems
2.5.3 Dual-Purpose Infrastructure Deployments
2.5.4 Unified Management and Orchestration
2.6 Benchmarking Methodologies for HPC and AI
2.6.1 MLPerf Benchmarking for AI Workloads
2.6.2 HPC-Specific Benchmarks (HPL, HPCG)
2.6.3 Green500 and Power Efficiency Metrics
2.6.4 Real-World Application Performance Analysis

3 PROCESSOR TECHNOLOGIES AND ARCHITECTURES

3.1 Central Processing Units (CPUs)
3.1.1 x86 Architecture Evolution
3.1.1.1 Intel Xeon Processor Family
3.1.1.2 AMD EPYC Processor Family
3.1.2 ARM-Based Data Center CPUs
3.1.2.1 AWS Graviton Processors
3.1.2.2 NVIDIA Grace CPU
3.1.2.3 Ampere Altra Family
3.1.2.4 Fujitsu A64FX for HPC
3.1.3 RISC-V and Other Instruction Set Architectures
3.1.3.1 Open-Source Ecosystem Development
3.1.3.2 Commercial RISC-V Server Initiatives
3.1.3.3 Market Positioning and Future Prospects
3.1.4 CPU AI Acceleration Technologies
3.1.4.1 Vector Processing Extensions
3.1.4.2 Neural Processing Units in Server CPUs
3.1.4.3 Matrix Multiplication Acceleration
3.1.5 CPU-GPU Hybrid Architectures
3.1.5.1 AMD APU Approach
3.1.5.2 Memory Coherency Benefits
3.1.5.3 Integrated vs. Discrete Solutions
3.2 Graphics Processing Units (GPUs)
3.2.1 GPU Architecture Evolution for AI and HPC
3.2.1.1 Streaming Multiprocessors (SMs)
3.2.1.2 Tensor Cores and AI-Specific Units
3.2.1.3 Ray Tracing Cores and Specialized Functions
3.2.2 NVIDIA Data Center GPUs
3.2.2.1 Ampere Architecture (A100)

+357 96 030922
info@marketpublishers.com

3.2.3 Hopper Architecture (H100, H200)
3.2.3.1 Blackwell Architecture (GB200)
3.2.3.2 Future GPU Roadmap and Performance Scaling
3.2.4 AMD Data Center GPUs
3.2.4.1 CDNA Architecture Evolution
3.2.4.2 Instinct MI Series (MI200, MI300)
3.2.4.3 Competitive Positioning and Performance
3.2.5 Chinese GPU Manufacturers
3.2.5.1 Biren Technology (BR100)
3.2.5.2 Moore Threads (MTT S4000)
3.2.5.3 MetaX (MXC500)
3.2.5.4 Iluvatar CoreX (Tianyuan/Zhikai)
3.2.6 Multi-GPU Systems and Scaling
3.2.6.1 Interconnect Technologies (NVLink, Infinity Fabric)
3.2.6.2 GPU-to-GPU Communication Optimization
3.2.6.3 Rack-Scale GPU Architecture
3.2.7 GPU Software Ecosystems
3.2.7.1 CUDA and CUDA-X Libraries
3.2.7.2 ROCm Platform
3.2.7.3 OneAPI and Industry Standards
3.3 AI Application-Specific Integrated Circuits (ASICs)
3.3.1 Cloud Service Provider Custom ASICs
3.3.1.1 Google Tensor Processing Units (TPUs)
3.3.1.2 AWS AI Accelerators
3.3.1.3 Microsoft Maia AI Accelerator
3.3.1.4 Meta MTIA Architecture
3.3.2 Matrix-Based AI Accelerators
3.3.2.1 Intel Habana Gaudi Architecture
3.3.2.2 Huawei Ascend AI Processors
3.3.2.3 Qualcomm Cloud AI 100
3.3.2.4 Chinese AI Accelerators
3.3.3 Spatial AI Accelerators
3.3.3.1 Cerebras Wafer-Scale Processors
3.3.3.2 SambaNova Reconfigurable Dataflow Architecture
3.3.3.3 Graphcore Intelligence Processing Unit (IPU)
3.3.3.4 Groq Tensor Streaming Processor (TSP)
3.3.4 Coarse-Grained Reconfigurable Arrays (CGRAs)
3.3.4.1 Academic Research and Commercial Applications
3.3.4.2 Dataflow vs. Control Flow Architecture Comparison

+357 96 030922
info@marketpublishers.com

3.3.4.3 Reconfigurability and Programming Models
3.3.4.4 Energy Efficiency Advantages
3.4 FPGAs and Other Programmable Solutions
3.4.1 FPGA Architecture for Data Center Applications
3.4.1.1 DSP Slices for AI Computation
3.4.1.2 Hard IP Blocks for Specialized Functions
3.4.2 Major FPGA Vendors and Products
3.4.2.1 Intel Agilex FPGA Family
3.4.2.2 AMD/Xilinx Versal Platform
3.4.2.3 Other Vendors
3.4.3 Adaptive Computing Solutions
3.4.3.1 ACAP (Adaptive Compute Acceleration Platform)
3.4.3.2 Hybrid FPGA-ASIC Approaches
3.4.3.3 Partial Reconfiguration Capabilities
3.4.4 FPGA Programming Models
3.4.4.1 High-Level Synthesis
3.4.4.2 OpenCL and Other Standards
3.4.4.3 AI Framework Integration
3.4.5 Market Position and Future Relevance
3.4.5.1 FPGA vs. GPU vs. ASIC Tradeoffs
3.4.5.2 Prototyping and Time-to-Market Advantages
3.4.5.3 Specialized Workload Optimization
3.5 Data Processing Units (DPUs) and SmartNICs
3.5.1 Network Interface Architecture Evolution
3.5.1.1 Network Acceleration Functions
3.5.1.2 Programmable Packet Processing
3.5.1.3 ARM Cores and Acceleration Engines
3.5.2 Major DPU Providers and Products
3.5.2.1 NVIDIA BlueField DPU
3.5.2.2 AMD/Pensando DPU
3.5.2.3 Intel Infrastructure Processing Unit (IPU)
3.5.2.4 Marvell OCTEON
3.5.3 Function Offloading Capabilities
3.5.3.1 Storage Processing
3.5.3.2 Security Functions
3.5.3.3 Virtualization Support
3.5.4 Integration with Computing Infrastructure
3.5.4.1 Software-Defined Networking (SDN)
3.5.4.2 Composable Infrastructure Models

+357 96 030922
info@marketpublishers.com

3.5.4.3 Management and Orchestration
3.6 Cryptocurrency and Blockchain Computing
3.6.1 ASIC Mining Hardware Architecture
3.6.2 GPU Mining Applications
3.6.3 Energy Efficiency in Crypto Mining
3.6.4 Overlap Between Crypto and AI Infrastructure

4 ENABLING TECHNOLOGIES

4.1 Advanced Semiconductor Manufacturing
4.1.1 Process Node Evolution
4.1.1.1 7nm and 5nm Technologies
4.1.1.2 3nm and 2nm Development
4.1.1.3 Sub-2nm Research and Innovations
4.1.2 Transistor Architecture Advancements
4.1.2.1 FinFET Technology
4.1.2.2 Gate-All-Around (GAA) Transistors
4.1.2.3 Nanosheet and Nanowire Approaches
4.1.2.4 Future Transistor Design Concepts
4.1.3 Leading-Edge Foundries and Capabilities
4.1.3.1 TSMC Technology Roadmap
4.1.3.2 Samsung Foundry Services
4.1.3.3 Intel Foundry Services (IFS)
4.1.3.4 Chinese Foundry Landscape
4.1.4 Semiconductor Design Scaling Challenges
4.1.4.1 Power Density and Thermal Constraints
4.1.4.2 Lithography Innovations (EUV, High-NA EUV)
4.1.4.3 Yield Management at Advanced Nodes
4.1.4.4 Cost Escalation and Economic Considerations
4.2 Advanced Packaging Technologies
4.2.1 2.5D Integration Approaches
4.2.1.1 Silicon Interposers
4.2.1.2 Organic Substrates
4.2.1.3 Fanout Wafer Level Packaging (FOWLP)
4.2.2 3D Integration Technologies
4.2.2.1 Through-Silicon Vias (TSVs)
4.2.2.2 Die-to-Die and Die-to-Wafer Bonding
4.2.2.3 Hybrid Bonding Technologies
4.2.3 Chiplet Architectures and Standards

+357 96 030922
info@marketpublishers.com

4.2.3.1 Disaggregation Benefits and Challenges
4.2.3.2 Inter-Chiplet Interconnect Standards (UCIe)
4.2.3.3 Integration with Different Process Nodes
4.2.4 System-in-Package Solutions
4.2.4.1 Heterogeneous Integration Approaches
4.2.4.2 Co-Packaged Optics
4.2.4.3 Embedded Power Delivery
4.3 Memory Technologies
4.3.1 High Bandwidth Memory (HBM) Evolution
4.3.1.1 HBM2E and HBM3 Specifications
4.3.1.2 HBM3E Performance Enhancements
4.3.1.3 HBM4 Development and Roadmap
4.3.1.4 HBM Suppliers and Manufacturing Capacity
4.3.2 DDR Memory Advancements
4.3.2.1 DDR5 for Server Applications
4.3.2.2 LPDDR5/5X for Power-Constrained Designs
4.3.2.3 GDDR6/7 for Graphics and AI
4.3.3 Memory Hierarchy and Tiered Approaches
4.3.3.1 CXL Memory Expansion
4.3.3.2 Memory Pooling Technologies
4.3.3.3 Tiered Storage-Memory Systems
4.3.4 Emerging Memory Technologies
4.3.4.1 Phase Change Memory (PCM)
4.3.4.2 Resistive RAM (ReRAM)
4.3.4.3 Magnetic RAM (MRAM)
4.3.4.4 Near-Memory Computing Approaches
4.4 Cooling and Thermal Management
4.4.1 Air Cooling Technologies and Limitations
4.4.2 Liquid Cooling Solutions
4.4.2.1 Direct-to-Chip Cooling Systems
4.4.2.2 Cold Plate Technologies
4.4.2.3 Coolant Distribution Units (CDUs)
4.4.3 Immersion Cooling Technologies
4.4.3.1 Single-Phase Immersion Systems
4.4.3.2 Two-Phase Immersion Systems
4.4.3.3 Coolant Chemistry and Environmental Considerations
4.4.4 Thermal Interface Materials
4.4.4.1 TIM Performance Characteristics
4.4.4.2 Application-Specific TIM Solutions

+357 96 030922
info@marketpublishers.com

4.4.4.3 Next-Generation Thermal Materials
4.4.5 Energy Recovery and Efficiency Approaches
4.4.5.1 Waste Heat Utilization
4.4.5.2 Heat Pump Integration
4.4.5.3 Combined Cooling and Power Solutions
4.5 Networking and Interconnects
4.5.1 Data Center Network Architectures
4.5.1.1 Spine-Leaf Topologies
4.5.1.2 Fat Tree Networks
4.5.1.3 Clos Networks and Variations
4.5.1.4 High-Speed Interconnect Standards
4.5.1.5 Ethernet Evolution (100G to 800G)
4.5.1.6 InfiniBand HDR and NDR
4.5.1.7 OmniPath and Proprietary Interconnects
4.5.2 Optical Interconnects
4.5.2.1 Pluggable Optical Transceivers
4.5.2.2 Co-Packaged Optics (CPO)
4.5.2.3 Silicon Photonics Integration
4.5.3 Network-on-Chip Designs
4.5.3.1 On-Chip Interconnect Architectures
4.5.3.2 Chiplet-to-Chiplet Communication
4.5.3.3 Memory-to-Compute Interfaces
4.6 Storage Technologies for HPC and AI
4.6.1 Flash Storage Solutions
4.6.1.1 SSD Technology Evolution
4.6.1.2 Form Factors and Interfaces
4.6.1.3 Performance Characteristics
4.6.2 Storage Server Architectures
4.6.2.1 All-Flash Arrays
4.6.2.2 Hybrid Storage Systems
4.6.2.3 Scale-Out Storage Architecture
4.6.3 High-Performance File Systems
4.6.3.1 Parallel File Systems
4.6.3.2 Object Storage Solutions
4.6.3.3 AI-Optimized Storage Software
4.6.4 Storage Tiering for AI and HPC Workloads
4.6.4.1 Data Locality Optimization
4.6.4.2 Cache Hierarchy Design
4.6.4.3 Storage Class Memory Integration

+357 96 030922
info@marketpublishers.com

5 MARKET ANALYSIS AND FORECASTS

5.1 Overall Data Center Processor Market
  5.1.1 Global Market Value (2025-2035)
  5.1.2 Annual Revenue Projections
  5.1.3 Unit Shipment Analysis
5.2 Average Selling Price (ASP) Trends
5.3 GPU Market Segment
  5.3.1 Unit Shipment Analysis
  5.3.2 Average Selling Price Trends
5.4 AI ASIC Market Segment
  5.4.1 Revenue Forecast (2025-2035)
  5.4.2 Unit Shipment Analysis
  5.4.3 Vendor-Specific vs. Third-Party ASICs
5.5 CPU Market Segment
  5.5.1 Revenue Forecast (2025-2035)
  5.5.2 Unit Shipment Analysis (2025-2035)
  5.5.3 Architecture Market Share (x86, ARM, Others)
5.6 FPGA and Alternative Processor Segment
  5.6.1 Revenue Forecast (2025-2035)
  5.6.2 Unit Shipment Analysis (2025-2035)
5.7 DPU and Networking Processor Segment
  5.7.1 Revenue Forecast (2025-2035)
  5.7.2 Unit Shipment Analysis (2025-2035)
  5.7.3 Integration Trend Analysis
5.8 Market Analysis by End-User Segment
  5.8.1 Cloud Service Providers
    5.8.1.1 Spending Forecast by Processor Type
    5.8.1.2 Infrastructure Expansion Analysis
    5.8.1.3 In-House vs. Third-Party Hardware Strategy
  5.8.2 HPC and Supercomputing Centers
    5.8.2.1 Spending Forecast by Processor Type
    5.8.2.2 Government vs. Commercial Investment
    5.8.2.3 System Architecture Trends
  5.8.3 Enterprise Data Centers
    5.8.3.1 Spending Forecast by Processor Type
    5.8.3.2 Industry Vertical Analysis
    5.8.3.3 On-Premises vs. Cloud Migration Impact

+357 96 030922
info@marketpublishers.com

5.8.4 Telecommunications and Edge Computing
5.8.4.1 Spending Forecast by Processor Type
5.8.4.2 5G/6G Infrastructure Requirements
5.8.4.3 Edge AI Deployment Trends
5.9 Specialized Market Segments
5.9.1 Cryptocurrency Mining Infrastructure
5.9.1.1 ASIC Mining Hardware Market
5.9.1.2 GPU Mining Dynamics
5.9.1.3 Energy Efficiency and Regulatory Impact
5.9.2 AI-as-a-Service Providers
5.9.2.1 Hardware Investment Patterns
5.9.2.2 Infrastructure Scale Requirements
5.9.3 Specialized AI Hardware Providers
5.9.3.1 Custom AI Appliance Market
5.9.3.2 Edge AI Hardware
5.9.3.3 Integrated Solutions Growth
5.10 Competitive Strategy Analysis
5.10.1 Product Development Strategies
5.10.1.1 Architectural Innovation Approaches
5.10.1.2 Performance vs. Energy Efficiency Focus
5.10.1.3 Specialized vs. General-Purpose Design
5.10.2 Market and Channel Strategies
5.10.2.1 Direct vs. Indirect Sales Models
5.10.2.2 Cloud Service Integration Partnerships
5.10.2.3 OEM and System Integrator Relationships
5.10.3 Software and Ecosystem Strategies
5.10.3.1 Developer Tool Investments
5.10.3.2 Library and Framework Support
5.10.3.3 Open Source vs. Proprietary Approaches
5.10.4 Manufacturing and Supply Chain Strategies
5.10.4.1 Foundry Partnership Models
5.10.4.2 Advanced Packaging Collaborations
5.10.4.3 Component Sourcing Security
5.11 Investment Landscape
5.11.1 Venture Capital Funding Trends
5.11.1.1 Early-Stage Investment Analysis
5.11.1.2 Late-Stage Funding Rounds
5.11.1.3 Regional Investment Distribution
5.11.2 Strategic Investments and Corporate Ventures

+357 96 030922
info@marketpublishers.com

5.11.2.1 Semiconductor Industry Investments
5.11.2.2 Cloud Provider Strategic Investments
5.11.2.3 OEM and System Vendor Investments

6 SUPPLY CHAIN ANALYSIS

6.1 Semiconductor Manufacturing Supply Chain
6.1.1 Foundry Landscape and Capabilities
6.1.1.1 Leading-Edge Node Production Capacity
6.1.1.2 Technology Leadership Assessment
6.1.1.3 Regional Manufacturing Distribution
6.1.1.4 Capacity Expansion Investments
6.2 Advanced Packaging Services
6.2.1 OSAT (Outsourced Semiconductor Assembly and Test) Providers
6.2.2 Integrated Device Manufacturers (IDM) Capabilities
6.2.3 Advanced Packaging Technology Providers
6.3 Memory Supplier Ecosystem
6.3.1 DRAM Manufacturers and Market Share
6.3.2 HBM Production Capabilities
6.3.3 Supply Constraints and Expansion Plans
6.3.4 Price Trend Analysis and Forecast
6.4 Cooling Solution Providers
6.4.1 Air Cooling Component Manufacturers
6.4.2 Liquid Cooling System Suppliers
6.4.3 Immersion Cooling Technology Providers
6.4.4 Integration with Data Center Design
6.5 Power Delivery Components
6.5.1 Power Supply Manufacturers
6.5.2 Voltage Regulator Module (VRM) Suppliers
6.5.3 Power Distribution Solutions
6.5.4 Energy Efficiency Technologies
6.6 System Integrators and OEMs
6.6.1 Server Manufacturer Landscape
6.6.2 HPC System Specialists
6.6.3 AI Infrastructure Providers
6.6.4 Custom System Design Services
6.7 Supply Chain Risks and Resilience
6.7.1 Raw Material Constraints
6.7.1.1 Critical Minerals and Materials

+357 96 030922
info@marketpublishers.com

6.7.1.2 Substrate and Packaging Materials
6.7.1.3 Supply Diversification Strategies
6.7.2 Manufacturing Capacity Limitations
6.7.2.1 Leading-Edge Node Constraints
6.7.2.2 Advanced Packaging Bottlenecks
6.7.2.3 HBM Supply Challenges
6.7.3 Logistics and Distribution Challenges
6.7.3.1 International Shipping Dependencies
6.7.3.2 Inventory Management Strategies
6.7.3.3 Just-in-Time vs. Resilience Trade-offs
6.7.4 Supply Chain Risk Mitigation Strategies
6.7.4.1 Multi-Sourcing Approaches
6.7.4.2 Strategic Inventory Positioning
6.7.4.3 Long-Term Supply Agreements

7 TECHNOLOGY AND MARKET TRENDS

7.1 Performance Scaling Trends
7.1.1 Beyond Moore's Law Scaling
7.1.1.1 Transistor Density Evolution
7.1.1.2 Clock Frequency Plateaus
7.1.1.3 Architecture-Driven Performance Gains
7.1.2 AI Compute Growth Trajectories
7.1.2.1 Training Compute Requirements Growth
7.1.2.2 Inference Compute Scaling Patterns
7.1.2.3 Model Size vs. Compute Relationship
7.1.3 Performance per Watt Evolution
7.1.3.1 Process Technology Contributions
7.1.3.2 Architecture Optimization Impact
7.1.3.3 Energy Proportional Computing Progress
7.1.4 Performance Density Trends
7.1.4.1 Rack-Level Compute Density
7.1.4.2 Data Center Floor Space Efficiency
7.1.4.3 Performance per Square Foot Metrics
7.2 Memory Bandwidth and Capacity Challenges
7.2.1 Memory Wall Considerations
7.2.1.1 . Compute-to-Memory Performance Gap
7.2.1.2 Bandwidth vs. Latency Requirements
7.2.1.3 Model Size vs. Memory Capacity Needs

+357 96 030922
info@marketpublishers.com

7.2.2 HBM Adoption and Evolution
7.2.2.1 HBM Bandwidth Growth Trajectory
7.2.2.2 Integration Challenges and Solutions
7.2.2.3 Cost Structure and Scaling Economics
7.2.3 Memory Hierarchies and Tiering
7.2.3.1 Multi-Level Memory Architectures
7.2.3.2 CXL Memory Expansion Adoption
7.2.3.3 Software-Defined Memory Management
7.2.3.4 Processing-in-Memory Technologies
7.2.3.5 Computational Storage Approaches
7.2.3.6 Accelerator-Memory Integration
7.3 Software Ecosystem Development
7.3.1 AI Frameworks and Libraries
7.3.1.1 PyTorch Ecosystem Evolution
7.3.1.2 TensorFlow Development Patterns
7.3.1.3 JAX and Emerging Frameworks
7.3.1.4 Hardware-Specific Optimizations
7.3.2 Compiler and Optimization Technologies
7.3.2.1 MLIR and Multi-Level IR Approaches
7.3.2.2 Hardware-Specific Code Generation
7.3.2.3 Automatic Optimization Capabilities
7.3.2.4 Quantization and Model Efficiency Tools
7.3.3 Hardware-Software Co-Design
7.3.3.1 Algorithm-Hardware Optimization
7.3.3.2 Domain-Specific Languages
7.3.3.3 Software-Defined Hardware Approaches
7.3.3.4 Integrated Development Environments
7.4 Energy Efficiency and Sustainability
7.4.1 Power Usage Effectiveness Metrics
7.4.1.1 Data Center PUE Benchmarks
7.4.1.2 Infrastructure Efficiency Improvements
7.4.1.3 Total Energy Attribution Models
7.4.2 Renewable Energy Integration
7.4.2.1 On-Site Generation Approaches
7.4.2.2 Power Purchase Agreements (PPAs)
7.4.2.3 24/7 Carbon-Free Energy Models
7.4.3 Circular Economy Approaches
7.4.3.1 Hardware Lifecycle Extension
7.4.3.2 Component Recycling and Recovery

+357 96 030922
info@marketpublishers.com

7.4.3.3 Design for Disassembly and Reuse
7.4.4 Carbon Footprint Reduction Strategies
7.4.4.1 Embodied Carbon in Hardware
7.4.4.2 Operational Carbon Reduction
7.4.4.3 Carbon Accounting Methodologies

8 APPLICATION SEGMENTS AND USE CASES

8.1 AI Training Infrastructure
8.1.1 Large Model Training Requirements
8.1.1.1 Foundational Model Training Infrastructure
8.1.1.2 Distributed Training Architectures
8.1.1.3 Training Cluster Design Principles
8.1.2 Training Methodology Evolution
8.1.2.1 Pre-Training Approaches
8.1.2.2 Fine-Tuning Infrastructure Requirements
8.1.2.3 Reinforcement Learning from Human Feedback (RLHF)
8.1.3 Training Efficiency Optimization
8.1.3.1 Data Pipeline Optimization
8.1.3.2 Distributed Training Techniques
8.1.3.3 Resource Utilization Management
8.1.4 Key Players and Market Leaders
8.1.4.1 Cloud Provider Training Services
8.1.4.2 Specialized AI Hardware Solutions
8.1.4.3 AI Research Lab Infrastructure
8.2 AI Inference Deployment
8.2.1 Cloud Inference Solutions
8.2.1.1 Large Model Serving Infrastructure
8.2.1.2 Inference Server Architectures
8.2.1.3 Multi-Tenant Inference Systems
8.2.2 Inference Optimization Techniques
8.2.2.1 Model Quantization and Compression
8.2.2.2 Batching and Throughput Optimization
8.2.2.3 Latency Reduction Approaches
8.2.3 Edge Computing Integration
8.2.3.1 Edge AI Hardware Requirements
8.2.3.2 Model Deployment Strategies
8.2.3.3 Edge-Cloud Collaborative Inference
8.2.4 Real-Time vs. Batch Processing

+357 96 030922
info@marketpublishers.com

8.2.4.1 Low-Latency Inference Requirements
8.2.4.2 Batch Inference Efficiency
8.2.4.3 Hybrid Processing Approaches
8.3 Traditional HPC Applications
8.3.1 Scientific Research and Simulation
8.3.1.1 Climate and Weather Modeling
8.3.1.2 Molecular Dynamics and Drug Discovery
8.3.1.3 Quantum Computing Simulation
8.3.2 Engineering and Design Simulation
8.3.2.1 Computational Fluid Dynamics (CFD)
8.3.2.2 Finite Element Analysis (FEA)
8.3.2.3 Electromagnetic Simulation
8.3.3 Financial Services Computing
8.3.3.1 Risk Analysis and Modeling
8.3.3.2 Algorithmic Trading Systems
8.3.3.3 AI-Enhanced Financial Models
8.3.4 AI-HPC Convergence Use Cases
8.3.4.1 Physics-Informed Neural Networks
8.3.4.2 AI-Enhanced Simulations
8.3.4.3 Hybrid Modeling Approaches
8.4 Cloud Service Provider Infrastructure
8.4.1 Hyperscale Data Center Architecture
8.4.1.1 Compute Infrastructure Design
8.4.1.2 Storage and Memory Hierarchy
8.4.1.3 Networking Architecture
8.4.2 Hyperscaler Technology Selection Strategies
8.4.2.1 In-House vs. Third-Party Hardware
8.4.2.2 Infrastructure Standardization Approaches
8.4.2.3 Specialized Hardware Integration
8.4.3 Total Cost of Ownership Analysis
8.4.3.1 Capital Expenditure Considerations
8.4.3.2 Operational Cost Structure
8.4.3.3 Energy and Cooling Economics
8.4.4 Cloud AI Services Architecture
8.4.4.1 AI Platform Service Design
8.4.4.2 Hardware Resource Allocation
8.4.4.3 Multi-Tenant Optimization
8.5 Enterprise Data Center Computing
8.5.1 AI Integration in Enterprise Computing

+357 96 030922
info@marketpublishers.com

8.5.1.1 On-Premises AI Infrastructure
8.5.1.2 Enterprise AI Appliances
8.5.1.3 Departmental AI Computing Resources
8.5.2 Private Cloud Solutions
8.5.2.1 Private AI Cloud Architecture
8.5.2.2 Resource Pooling and Virtualization
8.5.2.3 Self-Service AI Infrastructure
8.5.3 Hybrid Computing Models
8.5.3.1 Hybrid Cloud AI Deployment
8.5.3.2 Multi-Cloud AI Strategies
8.5.3.3 Cloud Bursting for AI Workloads
8.5.4 Industry-Specific AI Computing
8.5.4.1 Healthcare and Life Sciences
8.5.4.2 Manufacturing and Industrial
8.5.4.3 Retail and Consumer Services
8.5.4.4 Media and Entertainment
8.6 Future Outlook
8.6.1 Short-Term Market Dynamics (1-2 Years)
8.6.1.1 Supply-Demand Balance Forecast
8.6.1.2 Technology Deployment Trends
8.6.1.3 Pricing and Margin Expectations
8.6.2 Medium-Term Technology Evolution (3-5 Years)
8.6.2.1 Architecture Innovation Roadmap
8.6.2.2 Process Technology Progression
8.6.2.3 Memory and Storage Evolution
8.6.3 Long-Term Strategic Positioning (5-10 Years)
8.6.3.1 Post-Silicon Computing Potential
8.6.3.2 Radical Architecture Innovations
8.6.3.3 Compute Paradigm Shifts

9 COMPANY PROFILES 701 (222 COMPANY PROFILES)

10 REFERENCES

+357 96 030922
info@marketpublishers.com

List Of Tables

LIST OF TABLES

Table 1. Total Market Value and Growth Rates (Billions USD), 2025-2035.
Table 2. AI chips shipments (2025-2035).
Table 3. AI chips revenues (2025-2035).
Table 4. Graphics processing units (GPUs) shipments (2025-2035).
Table 5. Graphics processing units (GPUs) revenues (2025-2035).
Table 6. Central processing units (CPUs) shipments (2025-2035).
Table 7. Central processing units (CPUs) revenues (2025-2035).
Table 8. AI ASICs shipments (2025-2035).
Table 9. AI ASICs revenues (2025-2035).
Table 10. DPU shipments (2025-2035).
Table 11. DPU revenues (2025-2035).
Table 12. Network ASIC shipments (2025-2035).
Table 13. Network ASIC revenues (2025-2035).
Table 14. Crypto ASIC shipments (2025-2035).
Table 15. Crypto ASIC revenues (2025-2035).
Table 16. Next-Generation Node Transitions
Table 17. Advanced Packaging Technologies.
Table 18. Generative AI Computing Requirements.
Table 19. Main HPC and Generative AI investments 2023-2025.
Table 20. High-Growth Market Segments.
Table 21. Emerging Technology Areas.
Table 22. TOP500 Analysis and Performance Metrics.
Table 23. Supercomputers vs. Hyperscale Data Centers.
Table 24. Distributed Computing Models.
Table 25. Hybrid and Multi-Cloud Approaches.
Table 26. Edge-HPC Integration Frameworks.
Table 27. Deep Learning Architectures.
Table 28. Deep Learning Architectures.
Table 29. Training vs. Inference Workload Profiles.
Table 30. Inference Optimization Strategies.
Table 31. Batch vs. Real-Time Processing.
Table 32. Mixed Precision and Quantization Approaches.
Table 33. Compute Requirements Scaling Trends.
Table 34. Memory Footprint Challenges.
Table 35. HPC and AI Computing Overlapping Hardware Requirements.

+357 96 030922
info@marketpublishers.com

Table 36. HPC and AI Computing Dual-Purpose Infrastructure Deployments.
Table 37. HPC-Specific Benchmarks .
Table 38. Green500 and Power Efficiency Metrics.
Table 39. Real-World Application Performance Analysis.
Table 40. Commercial RISC-V Server Initiatives
Table 41. Market Positioning and Future Prospects
Table 42. Vector Processing Extensions
Table 43. Neural Processing Units in Server CPUs
Table 44. Integrated vs. Discrete Solutions
Table 45. Streaming Multiprocessors (SMs)
Table 46. Tensor Cores and AI-Specific Units
Table 47. Ray Tracing Cores and Specialized Functions
Table 48. Chinese GPU Manufacturers.
Table 49. OneAPI and Industry Standards
Table 50. Chinese AI Accelerators .
Table 51. WSE Architecture and Manufacturing.
Table 52. Academic Research and Commercial Applications.
Table 53. Dataflow vs. Control Flow Architecture Comparison
Table 54. Reconfigurability and Programming Models
Table 55. Energy Efficiency Advantages
Table 56. Hybrid FPGA-ASIC Approaches
Table 57. Partial Reconfiguration Capabilities.
Table 58. OpenCL and Other Standards.
Table 59. FPGA vs. GPU vs. ASIC Tradeoffs.
Table 60. Prototyping and Time-to-Market Advantages.
Table 61. Composable Infrastructure Models.
Table 62. ASIC Mining Hardware Architecture.
Table 63. GPU Mining Applications.
Table 64. 7nm and 5nm Technologies
Table 65. 3nm and 2nm Development
Table 66. Sub-2nm Research and Innovations
Table 67. Nanosheet and Nanowire Approaches
Table 68. Future Transistor Design Concepts
Table 69. Samsung Foundry Services
Table 70. Intel Foundry Services (IFS)
Table 71. Chinese Foundry Landscape
Table 72. Power Density and Thermal Constraints
Table 73. Lithography Innovations (EUV, High-NA EUV)
Table 74. Yield Management at Advanced Nodes

+357 96 030922
info@marketpublishers.com

Table 75. Cost Escalation and Economic Considerations
Table 76. Hybrid Bonding Technologies
Table 77. Disaggregation Benefits and Challenges
Table 78. Inter-Chiplet Interconnect Standards (UCIe)
Table 79. Integration with Different Process Nodes
Table 80. Heterogeneous Integration Approaches
Table 81. HBM2E and HBM3 Specifications
Table 82. HBM3E Performance Enhancements
Table 83. HBM Suppliers and Manufacturing Capacity
Table 84. DDR Memory Advancements
Table 85. Memory Pooling Technologies
Table 86. Tiered Storage-Memory Systems
Table 87. Air Cooling Technologies and Limitations
Table 88. Direct-to-Chip Cooling Systems
Table 89. Cold Plate Technologies
Table 90. Coolant Chemistry and Environmental Considerations
Table 91. TIM Performance Characteristics
Table 92. Application-Specific TIM Solutions
Table 93. Next-Generation Thermal Materials
Table 94. Energy Recovery and Efficiency Approaches
Table 95. Combined Cooling and Power Solutions
Table 96. Ethernet Evolution (100G to 800G).
Table 97. SSD Technology Evolution.
Table 98. Flash Storage Solutions Performance Characteristics.
Table 99. Hybrid Storage Systems
Table 100. Object Storage Solutions
Table 101. Average Selling Price (ASP) Trends
Table 102. GPU Market Segment Revenue Forecast (2025-2035)
Table 103. GPU Market Segment Unit Shipment Analysis (2025-2035)
Table 104. GPU Market Segment Average Selling Price Trends
Table 105. AI ASIC Market Segment Revenue Forecast (2025-2035)
Table 106. AI ASIC Market Segment Unit Shipment Analysis (2025-2035)
Table 107. Vendor-Specific vs. Third-Party ASICs
Table 108. CPU Market Segment Revenue Forecast (2025-2035)
Table 109. CPU Market Segment Unit Shipment Analysis (2025-2035)
Table 110. FPGA and Alternative Processor Revenue Forecast (2025-2035)
Table 111. FPGA and Alternative Processor Unit Shipment Analysis (2025-2035)
Table 112. DPU and Networking Processor Revenue Forecast (2025-2035)
Table 113. DPU and Networking Processor Unit Shipment Analysis (2025-2035)

+357 96 030922
info@marketpublishers.com

Table 114. Government vs. Commercial Investment
Table 115. System Architecture Trends
Table 116. On-Premises vs. Cloud Migration Impact
Table 117. 5G/6G Infrastructure Requirements
Table 118. Edge AI Deployment Trends
Table 119. GPU Mining Dynamics
Table 120. Energy Efficiency and Regulatory Impact
Table 121. Hardware Investment Patterns
Table 122. Infrastructure Scale Requirements
Table 123. Architectural Innovation Approaches
Table 124. Performance vs. Energy Efficiency Focus
Table 125. Direct vs. Indirect Sales Models
Table 126. Regional Investment Distribution
Table 127. Semiconductor Industry Investments
Table 128. OEM and System Vendor Investments
Table 129. Regional Manufacturing Distribution
Table 130. Capacity Expansion Investments
Table 131. OSAT (Outsourced Semiconductor Assembly and Test) Providers
Table 132. Integrated Device Manufacturers (IDM) Capabilities
Table 133. Advanced Packaging Technology Providers
Table 134. Memory Price Trend Analysis and Forecast
Table 135. Air Cooling Component Manufacturers
Table 136. Liquid Cooling System Suppliers
Table 137. Immersion Cooling Technology Providers
Table 138. Power Supply Manufacturers
Table 139. Voltage Regulator Module (VRM) Suppliers
Table 140. Energy Efficiency Technologies
Table 141. HPC System Specialists
Table 142. AI Infrastructure Providers
Table 143. Critical Minerals and Materials
Table 144. Substrate and Packaging Materials
Table 145. Leading-Edge Node Constraints
Table 146. Advanced Packaging Bottlenecks
Table 147. HBM Supply Challenges
Table 148. Inventory Management Strategies
Table 149. Inference Compute Scaling Patterns
Table 150. Model Size vs. Compute Relationship
Table 151. Performance per Watt Evolution
Table 152. Performance Density Trends

+357 96 030922
info@marketpublishers.com

Table 153. Performance per Square Foot Metrics
Table 154. Bandwidth vs. Latency Requirements
Table 155. Model Size vs. Memory Capacity Needs
Table 156. Integration Challenges and Solutions
Table 157. Multi-Level Memory Architectures
Table 158. Near-Memory and In-Memory Computing
Table 159. MLIR and Multi-Level IR Approaches
Table 160. Software-Defined Hardware Approaches
Table 161. Data Center PUE Benchmarks
Table 162. Total Energy Attribution Models
Table 163. 24/7 Carbon-Free Energy Models
Table 164. Carbon Accounting Methodologies
Table 165. Cloud Provider Training Services
Table 166. Specialized AI Hardware Solutions
Table 167. Edge AI Hardware Requirements.
Table 168. Total Cost of Ownership Analysis
Table 169. AMD AI chip range.
Table 170. Evolution of Apple Neural Engine.

+357 96 030922
info@marketpublishers.com

List Of Figures

LIST OF FIGURES

Figure 1. Total Market Value and Growth Rates (Billions USD), 2025-2035
Figure 2. AI chips shipments (2025-2035).
Figure 3. AI chips revenues (2025-2035).
Figure 4. Graphics processing units (GPUs) shipments (2025-2035).
Figure 5. Graphics processing units (GPUs) revenues (2025-2035).
Figure 6. Central processing units (CPUs) shipments (2025-2035).
Figure 7. Central processing units (CPUs) revenues (2025-2035).
Figure 8. AI ASICs shipments (2025-2035).
Figure 9. AI ASICs revenues (2025-2035).
Figure 10. DPU shipments (2025-2035).
Figure 11. DPU revenues (2025-2035).
Figure 12. Network ASIC shipments (2025-2035).
Figure 13. Network ASIC revenues (2025-2035).
Figure 14. Crypto ASIC shipments (2025-2035).
Figure 15. Crypto ASIC revenues (2025-2035).
Figure 16. Historical Evolution of HPC Systems.
Figure 17. AMD EPYC Processor Family.
Figure 18. NVIDIA Grace CPU
Figure 19. Ampere Altra Family
Figure 20. A64FX for HPC
Figure 21. Ampere Architecture (A100).
Figure 22. Hopper Architecture (H100, H200)
Figure 23. Blackwell Architecture (GB200)
Figure 24. 3Future GPU Roadmap and Performance Scaling.
Figure 25. CDNA Architecture Evolution
Figure 26. Instinct MI Series (MI200, MI300)
Figure 27. Interconnect Technologies (NVLink, Infinity Fabric)
Figure 28. Rack-Scale GPU Architecture
Figure 29. Google Tensor Processing Units (TPUs).
Figure 30. Trainium Architecture.
Figure 31. AWS Neuron SDK.
Figure 32. Microsoft Maia AI Accelerator
Figure 33. Meta MTIA Architecture
Figure 34. Intel Habana Gaudi Architecture
Figure 35. Greco and Gaudi3 Roadmap.

+357 96 030922
info@marketpublishers.com

Figure 36. Huawei Ascend AI Processors
Figure 37. Da Vinci Architecture
Figure 38. Qualcomm Cloud AI 100.
Figure 39.Cerebras Wafer-Scale Processors
Figure 40. SambaNova Reconfigurable Dataflow Architecture.
Figure 41. Cardinal SN10 RDU
Figure 42. SN40L Next-Generation System
Figure 43. Dataflow Computing Model
Figure 44. Graphcore Intelligence Processing Unit (IPU).
Figure 45. Colossus MK2 Architecture
Figure 46. Groq Tensor Streaming Processor (TSP).
Figure 47. AMD/Xilinx Versal Platform
Figure 48. Network Interface Architecture Evolution.
Figure 49. NVIDIA BlueField DPU.
Figure 50. AMD/Pensando DPU
Figure 51. Intel Infrastructure Processing Unit (IPU)
Figure 52. Marvell OCTEON
Figure 53. FinFET Technology.
Figure 54. Gate-All-Around (GAA) Transistors
Figure 55. TSMC Technology Roadmap.
Figure 56. Silicon Interposers
Figure 57. Fanout Wafer Level Packaging (FOWLP)
Figure 58. HBM4 Development and Roadmap.
Figure 59. Coolant Distribution Units (CDUs).
Figure 60. Cloud Service Providers Spending Forecast by Processor Type
Figure 61. HPC and Supercomputing Centers Spending Forecast by Processor Type
Figure 62. Enterprise Data Centers Spending Forecast by Processor Type
Figure 63. Telecommunications and Edge Computing Spending Forecast by Processor Type.
Figure 64. Transistor Density Evolution
Figure 65.Supply-Demand Balance Forecast
Figure 66. Architecture Innovation Roadmap.
Figure 67. AMD Radeon Instinct.
Figure 68. AMD Ryzen 7040.
Figure 69. Alveo V70.
Figure 70. Versal Adaptive SOC.
Figure 71. AMD’s MI300 chip.
Figure 72. Cerebas WSE-2.
Figure 73. DeepX NPU DX-GEN1.

Market Publishers

+357 96 030922
info@marketpublishers.com

Figure 74. InferX X1.
Figure 75. “Warboy”(AI Inference Chip).
Figure 76. Google TPU.
Figure 77. Colossus™ MK2 GC200 IPU.
Figure 78. GreenWave’s GAP8 and GAP9 processors.
Figure 79. Journey 5.
Figure 80. IBM Telum processor.
Figure 81. 11th Gen Intel® Core™ S-Series.
Figure 82. Envise.
Figure 83. Pentonic 2000.
Figure 84. Meta Training and Inference Accelerator (MTIA).
Figure 85. Azure Maia 100 and Cobalt 100 chips.
Figure 86. Mythic MP10304 Quad-AMP PCIe Card.
Figure 87. Nvidia H200 AI chip.
Figure 88. Grace Hopper Superchip.
Figure 89. Panmnesia memory expander module (top) and chassis loaded with switch and expander modules (below).
Figure 90. Cloud AI 100.
Figure 91. Peta Op chip.
Figure 92. Cardinal SN10 RDU.
Figure 93. MLSoC™.
Figure 94. Grayskull.

+357 96 030922
info@marketpublishers.com

I would like to order

Product name: The Global Market for High Performance Computing (HPC) and AI Accelerators 2025-2035

Product link: https://marketpublishers.com/r/GF412D239F1CEN.html

Price: US$ 2,000.00 (Single User License / Electronic Delivery)

If you want to order Corporate License or Hard Copy, please, contact our Customer Service:

info@marketpublishers.com

Payment

To pay by Credit Card (Visa, MasterCard, American Express, PayPal), please, click button on product page https://marketpublishers.com/r/GF412D239F1CEN.html