virtana

From Storage to GPU and Back:
A Deep Dive into AI Data Center Innovations

Introduction

The landscape of artificial intelligence (AI) is evolving at an unprecedented pace, driven by advances in technology and the ever-increasing demand for intelligent solutions across various sectors. Traditional AI methods, which have been utilized in healthcare, equipment monitoring, logistics, marketing, and financial systems for over a decade, rely heavily on feature learning, labeling, and other techniques that demand significant human effort and often result in process delays. However, the advent of generative AI is revolutionizing the field by bringing data, policies, and systems closer to humans, thereby reducing search and correlation efforts and creating superior human interfaces. This breakthrough is opening new horizons for AI applications.

The Shift to Unsupervised Learning

AI is rapidly transitioning towards unsupervised learning as algorithms become more adept at feature learning and quality assurance. Generative AI techniques, in particular, enable foundation models to learn autonomously with minimal human intervention. While human red teams are occasionally deployed to set guardrails, the bulk of the learning process is automated. This shift enhances efficiency and paves the way for innovative AI applications.

Select AI Index technical performance benchmarks vs. human performance
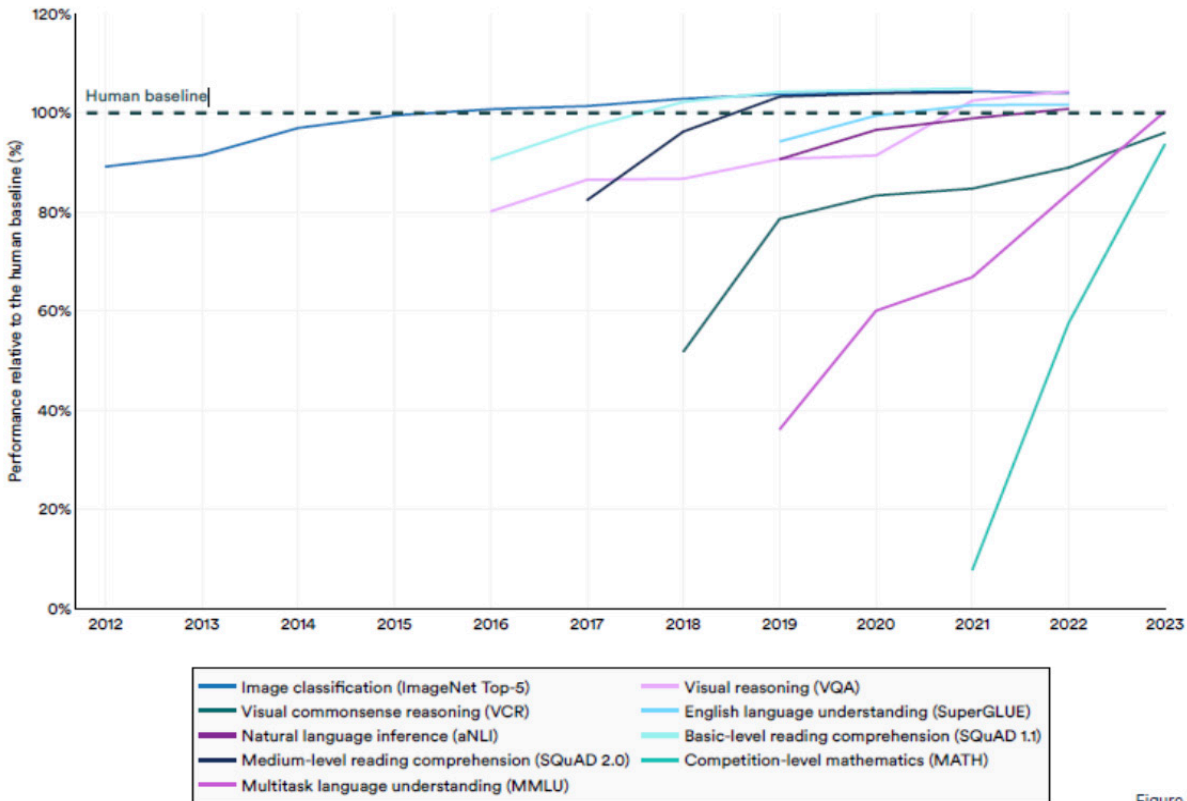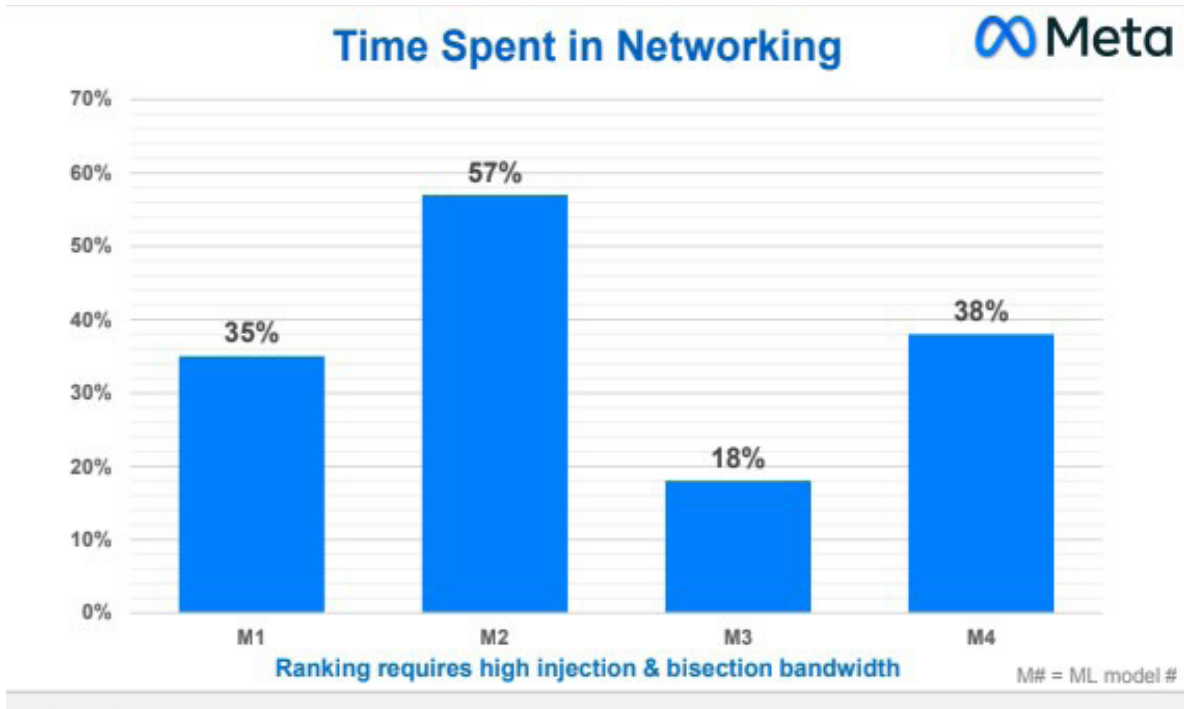Source: AI Index, 2024 | Chart: 2024 AI Index report

Figure 2

The Rise of AI in Computing Resources

AI and machine learning (ML) are increasingly becoming integral components of the computing resources budget. Integrating AI into applications is now commonplace, with virtual assistants and co-pilots being ubiquitous. This trend is driving applications and data centers closer to the data, necessitating the adoption of edge technology. Unlike the traditional approach of centralized data centers, the future lies in a hybrid model that combines cloud, data center, and edge pods. Projections indicate a 10% growth in data center systems during 2023-2024, with an accelerated pace anticipated in the coming years.

Navigating the AI Data Supply Chain

Building an AI Data Center (AIDC) presents numerous challenges, including asset utilization, power consumption, heat output, intellectual property (IP) leaks, IP liability, and substantial capital equipment costs. These concerns have prompted enterprises to house critical AI functions internally, especially when handling enterprise data. Optimizing GPU (Graphics Processing Unit) resources, which are pivotal for AI operations. Software optimizations can enhance efficiency by 5X to 10X, with improved schedulers. Identifying the causes and addressing the bottlenecks in this process is crucial for optimizing performance.

Time Spent in Networking ∞ Meta

Ranking requires high injection & bisection bandwidth

M# = ML model #

2022 OCP keynote by Alexis Bjorlin, VP, Infrastructure at Meta

Understanding AI Data Patterns

The last few years have seen a dramatic increase in dataset sizes driven by AI. The nature of the data has also evolved, presenting distinct patterns compared to traditional application workloads. Key observations include:

- **Data-intensive Operations**: High injection and bisection bandwidth with elephant flows (gradient and weight exchange) and high su

- **Low Entropy Messages**: Data formats remain consistent, characterized by periodic peaks and valleys, and are highly ordered and predictable.

- **Lossless Network Requirements**: Training processes are sensitive to tail latency and job completion time and are drop- and order-sensitive.

- **Latency Sensitivity for Inference**: While training and fine-tuning are more tolerant, inference processes demand low latency.

Checkpointing large and complex models exemplifies the pressure exerted on network and storage infrastructure. These checkpoints are essential for recovery during crashes or when the error gradient stabilizes, but they also temporarily halt model progression until completion.

Understanding the Types of Data

The cornerstone of artificial intelligence is data—its abundance, variety, and quality directly impact the effectiveness and scope of AI applications. When delving into AI, it's essential to grasp the different kinds of data that fuel these intelligent systems:

Primary and Secondary Sources of Data

- **Primary Data**: Generated directly by the enterprise, providing unique and proprietary insights.
- **Secondary Data**: Public data or data generated by third parties, often used to supplement primary data sources.

Quantitative and Qualitative Data

- **Quantitative Data:** Numerical data that quantifies elements, facilitating statistical analysis and pattern recognition.
- **Qualitative Data**: Descriptive data that provides in-depth insights into specific problems, often used in initial exploratory phases of AI projects.

Data Types

- **Structured Data:** Organized in tabular formats, easy to analyze and use in machine learning models.
- **Unstructured Data**: Raw data in its most unprocessed form, such as text, images, and videos, requiring more sophisticated processing techniques.
- **Semi-Structured Data**: A hybrid of structured and unstructured data, with some organizational properties that facilitate easier analysis than purely unstructured data.

Historical and Real-Time Data

- **Historical Data**: Used for building strategic insights and long-term models, crucial for understanding trends and making predictions.
- **Real-Time Data**: Processed immediately as it is generated, essential for applications requiring instant responses and up-to-date insights.

Internal and External Data

- **Internal Data:** Collected within the organization, often highly relevant and reliable.
- **External Data**: Sourced from outside the organization, used to enhance and corroborate internal data, providing a broader context.


Availability and Performance

Effectively managing AI data patterns demands a blend of large and varied data types, high streaming bandwidth for training jobs, optimal write performance for handling high volumes, and robust random read performance for inference tasks. Crucially, these capabilities must be reliable 24×7, accessible across diverse silos and applications. However, this process often encounters challenges such as:

- **Floods:** Data surges that overwhelm the system.
- **Transient Over-subscriptions**: Temporary spikes in demand that exceed capacity.
- **Flow Collisions**: Conflicts in data transmission paths.
- **Incast Scenarios**: Congestion due to simultaneous data requests.

Virtana Data Fabric Map

To navigate these complexities, Virtana's AIDC Fabric Map provides a comprehensive topology of data fabrics, illustrating their roles and interactions within an AI data center.
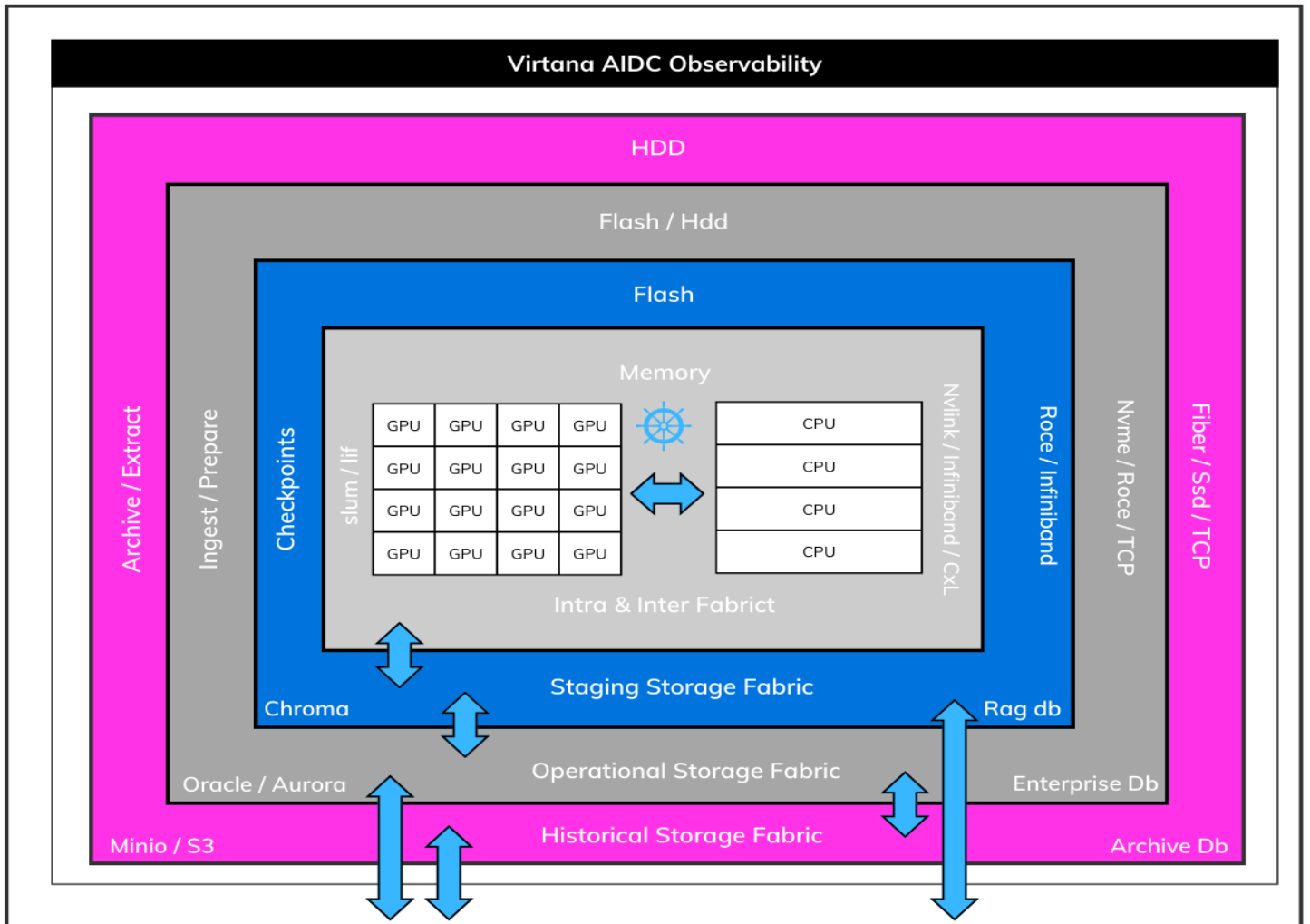
Data resides on various storage fabrics, categorized by their proximity to the GPU:

Memory: The closest layer to the GPU, utilizing fast interconnects like NVLink for efficient data access and processing. The optimal use of the GPU/CPU assets is to keep them fed with data efficiently. GPUs can have 100s of kernels that perform various functions and multiple cores that they run in. Maintaining the necessary data close to the GPU in the memory the GPU can access is referred to as 'memory-centric.' This innermost memory fabric uses fast interconnects like NVLink, allowing GPUs to access, process, and transfer data to multiple kernels and other GPUs.

Staging: Handles processed data inputs, checkpoints, and model outputs, using technologies like Flash storage and RoCE for quick data movement. Rag Databases are candidates for this layer. This layer must be fast and move data in and out of the Memory layer. Flash storage and interconnects like RoCE (RDMA Over Converged Ethernet) and InfiniBand are optimal.

Operational: Contains high-temperature and transactional data. Enterprise databases like Oracle, observability databases like Elastic, and cloud databases like AWS Aurora reside there. These databases have varying performance requirements and are migrating to NVMe and RoCE interconnects in preparation for performance requirements that are expected to grow.

Historical: Data that has started to cool down is typically stored in this layer. AWS S3 and its OnPrem equivalents, like Minio and Apache Iceberg, reside in this layer. Cost-effective storage like HDD and Fibre Channel connections are used.

Data flows bidirectionally between these layers at high speeds and volumes over a complex fabric and compute environment. **Ensuring optimal performance in an AIDC involves observing, proactively predicting issues, and swiftly troubleshooting data traffic between storage and compute assets.**

Addressing Performance Issues

Virtana has been at the forefront of tackling storage and network fabric challenges, extending its expertise to meet the expanded requirements of AI data centers. Here are a few illustrative examples:

Example 1: Congestion Problem

Fibre Channel (FC) employs a credit-based mechanism to handle congestion and minimize data loss. Virtana monitors these conversations, detecting anomalies and issuing alerts when credits cause deadlocks. This approach is mirrored in AI environments with faster interconnects, such as InfiniBand, which uses a credit mechanism to ensure efficient data transmission without interference.

Example 2: Load Balance Problem

Virtana monitors traffic loads from multiple servers, enabling quick detection of path congestion and failures. In an AIDC, congestion can occur across various layers. By monitoring numerous links and utilizing Priority Flow Control and Explicit Congestion Notifications, Virtana detects anomalies and raises alerts, ensuring smooth data flow across the infrastructure.

Example 3: Latency Issue in Real-Time Inference

In real-time inference applications, even minor latencies can significantly impact performance. Virtana's monitoring tools detect latency spikes by continuously analyzing data paths and identifying bottlenecks in real time. For instance, by implementing dynamic load balancing and prioritization mechanisms, latency-sensitive tasks are given precedence, ensuring seamless operation and quick response times.

Example 4: Data Integrity in High-Volume Transfers

High-volume data transfers are prone to integrity issues, especially in complex, multi-node environments. Virtana utilizes advanced error-checking algorithms and redundancy protocols to ensure data integrity. During extensive data migrations, Virtana's system checks for inconsistencies and automatically corrects them, maintaining data accuracy and reliability without manual intervention.

Example 5: Optimizing Data Placement for GPU Utilization

Effective GPU utilization hinges on optimal data placement. Virtana's AI-driven analytics identify patterns in data usage and automatically adjust data placement strategies. By placing frequently accessed data closer to GPUs and distributing less critical data to lower-tier storage, Virtana maximizes GPU efficiency and reduces access times. This approach also helps manage heat output and power consumption by optimizing the load across the infrastructure.

The Future of AI Data Centers: A Holistic Approach to Performance Optimization

The AI revolution demands a new breed of data center – one meticulously designed for the unique challenges of Generative AI and beyond. IT leaders must move beyond siloed monitoring and embrace a comprehensive approach that encompasses the entire AI data center ecosystem. By monitoring every critical component – from storage devices and networking fabrics to Kubernetes clusters and cloud deployments – IT leaders gain a unified perspective of their AI infrastructure.

This comprehensive monitoring empowers proactive management through:

- **Discovery and Inventory**: Gain a complete understanding of all resources at play.

- **Configuration Management**: Ensure consistent and optimized configurations across the infrastructure.

- **Detailed Topology Mapping**: Visualize data flow and identify potential bottlenecks.

- **Rich Telemetry**: Collect and analyze metrics, logs, profiles, and traces for deeper insights.

- **AI-powered Anomaly Detection**: Proactively identify performance issues before they disrupt operations.

- **Causal Analysis and Troubleshooting**: Pinpoint the root cause of problems for faster resolution

Adopting this holistic approach empowers IT leaders to build and maintain high-performing AI data centers. This translates to accelerated innovation, improved efficiency, and, ultimately, positive outcomes for stakeholders and customers. The future of AI is bright, and Virtana is here to guide you on the journey.