AI INFRA SUMMIT

AI DATA CENTER INSIGHTS

INCLUDING

BROADCOM Microsoft NVIDIA OpenAI Schneider Electric

KAREN PETERSBURG, VP, DATA CENTER DEVELOPMENT & CONSTRUCTION, POWERHOUSE DATA CENTERS

KEYNOTE PANEL
GETTING GROWTH RIGHT - INNOVATION & SUSTAINABILITY IN THE AI DATA CENTER

01 **Sustained Demand**
Unlike previous tech booms, current data center growth is characterized by zero vacancy and pre-payment for capacity, indicating immediate and sustained demand rather than speculative building.

02 **Key Headwinds**
Major challenges to continued growth include power availability (long utility lead times, complex relationships), supply chain constraints for components, and the ecosystem's ability to keep pace with rapid technological advancements.

03 **Timescale Misalignment**
A significant challenge is the mismatch between the slow, multi-year planning cycles of utilities and the rapid, annual (or even weekly) evolution of AI models and hardware.

04 **AI Workload Convergence**
The distinction between AI training and inference workloads is blurring, with continuous feedback loops and reinforcement learning driving a constant need for new models and evolving data center footprints.

05 **Sustainability Initiatives**
Microsoft aims for carbon negative, water positive, and zero waste by 2030, requiring innovation and ecosystem collaboration across carbon, water, waste, and ecosystem protection.

06 **Supply Chain Sustainability**
Microsoft mandates suppliers to reduce Scope 3 emissions, setting annual targets and requiring 100% clean energy by 2030, extending sustainability efforts throughout the entire supply chain.

JC Pan
VP, Cloud Supply Chain Sourcing
Microsoft

Dion Harris
Sr. Director, AI and HPC Infrastructure Solutions
NVIDIA

Laura Ortman
CEO
Cologix

Steven Carlini
Chief AI Advocate
Schneider Electric

KEYNOTE
SCALING UP TO AGI

RAM VELAGA, SVP & GM, CORE SWITCHING GROUP, **BROADCOM**

01 **Shift to Specialized Accelerators**
The industry is moving away from general-purpose GPUs towards specifically optimized accelerators (XPUs) to achieve the lowest cost and power per token.

02 **Need for Large-Scale Clustering**
Individual XPUs are insufficient for large language models, requiring thousands to hundreds of thousands of XPUs to be clustered and operate as a single supercomputer.

03 **Network Bottleneck**
XPUs are often idle 20-50% of the time waiting for data to traverse the network, making the network an extremely critical component in large-scale AI systems.

04 **Ethernet as the Solution**
Ethernet is presented as the reliable, fast, and open network choice for building the world's largest distributed systems, similar to Google's early data centers.

05 **Decoupling XPU and Network Layers**
Ethernet uniquely decouples XPU transaction management from the network layer, allowing hyperscalers to innovate their XPU architectures without being constrained by proprietary network specifications.

06 **Broadcom's Ethernet Innovations**
Broadcom has developed 100+ terabit switches and the Tomahawk Ultra (50 Tbps) chip, achieving sub-250 nanosecond latency and line-rate 64-byte packet handling, challenging previous limitations of Ethernet.

PRESENTATION
AGENTS, ARCHITECTURE AND ALIGNMENT - WHAT THE NEXT WAVE DEMANDS OF OUR INFRASTRUCTURE

RICHARD HO, HEAD OF HARDWARE, **OPEN AI**

01 **Global Scale Computing**
The industry is moving towards global-scale computing, necessitating many massive data center campuses spread worldwide to accommodate demand and user distribution.

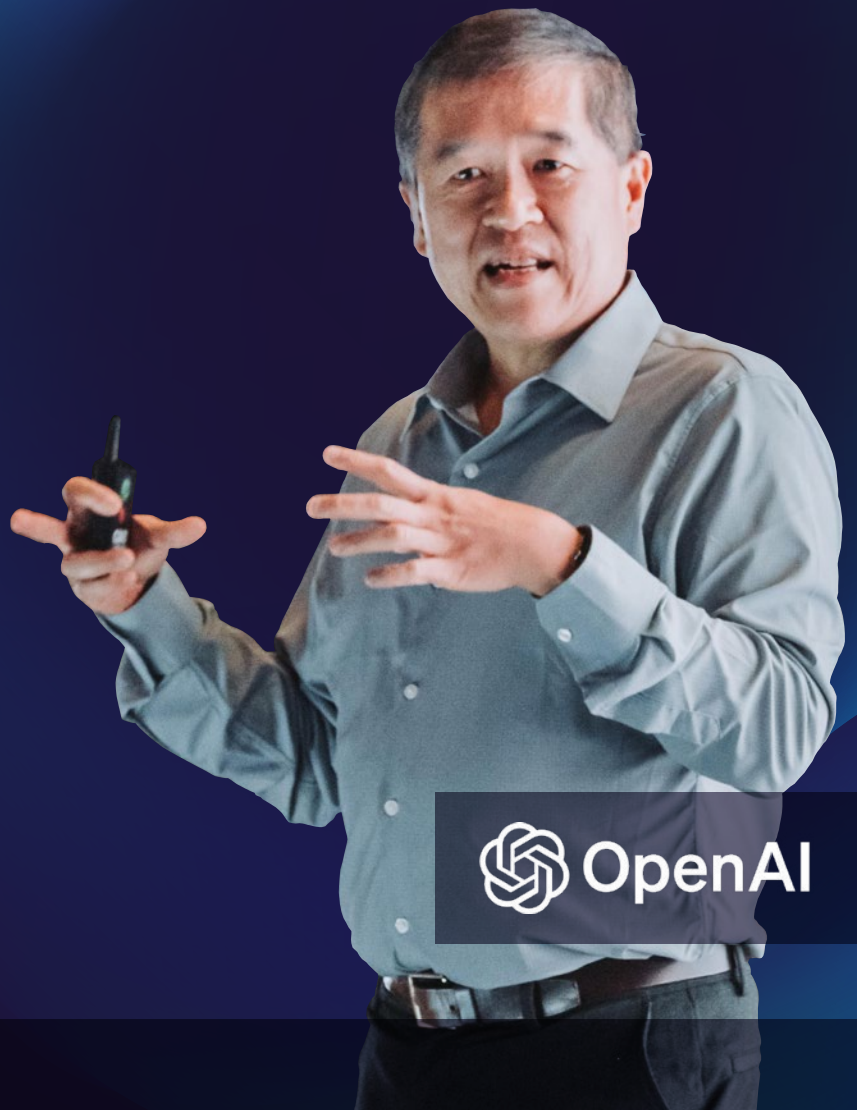
02 **Infrastructure Demands**
Agentic workloads require significant infrastructure evolution, including advanced memory architectures (HBM, disaggregated memory), low-latency, high-throughput, low-power networking, and efficient thermal management.

03 **Hardware Alignment**
Safety and alignment must be embedded directly into hardware, not just software, to ensure reliable and secure AI operation.

04 **Security & Telemetry**
Critical hardware features include robust telemetry for anomalous events, secure enclaves to prevent malicious injection, and secure global networking against physical tampering.

05 **Resilience & Error Correction**
Infrastructure needs high resilience through network redundancy, self-healing protocols, on-chip error correction, and silent data correction for both training and inference.

06 **Industry Investment**
There's an urgent need for increased investment in infrastructure (fabs, wires, switches, racks) to meet the exploding demand driven by exponential AI growth.



PRESENTATION
REDEFINING THE AI DATA CENTER: FROM TOKENS TO WATTS, AND BEYOND THE RACK

CHRIS SHARP, CTO, **DIGITAL REALTY**

01 **Complexity of Tokenomics**

Not all tokens are equal; understanding total token production, token lifespan, and deployment (e.g., Mixture of Experts models) is crucial, advocating for private AI exchanges for critical connectivity.

02 **Power (Watts) as Product**

Power is a binary enabler; critical considerations include total capacity, power densification (e.g., NVIDIA DGX certification), and sustainable power sources like solar, wind, and small modular reactors (SMRs).

03 **Capital (Dollars) Challenges**

The capital required for AI infrastructure is monumental, leading to concerns about technological obsolescence (GPUs every 12 months), ROI, and the need for efficient operations (low PUE).

04 **Data Center Site Selection**

Key factors for data center selection include proximity to consumers and emerging AI ecosystems, high optical throughput for connectivity, modular facilities for retrofitting (air to liquid cooling), and supporting heavy rack floor loading.

05 **Hardware Evolution**

Silicon changes rapidly, unlike concrete, requiring careful consideration of where hardware is placed and how to evolve with new capabilities and efficiencies in the market.

06 **Software & Models**

Emphasizes understanding and leveraging various models (RAG, SLMs, hierarchical reasoning) to find cost-effective solutions, rather than solely relying on large parameter models.

PANEL
TECHNOLOGY ROADMAPPING - DUE DILIGENCE FOR NEAR TERM VS. HORIZON TECHNOLOGIES

01 **Rapid Adaptation**
The industry faces challenges integrating rapidly changing chip technology, especially AI hardware, into current operations and future planning due to the quick pace of innovation.

02 **Flexibility in Design**
A key strategy to combat obsolescence is building flexibility into electrical, mechanical, structural, and architectural designs, planning for "known unknowns" and a 15-20 year occupancy period.

03 **Proactive Planning**
Companies design 5-7 years ahead, use modular designs, and engage in supplier pre-purchase programs to adapt to customer needs and manage supply chain lead times.

04 **Power Constraints**
Power availability is a major concern, leading to exploration of alternative solutions like natural gas generation, small modular reactors (SMRs), and hydrogen, alongside traditional utility power.

05 **Super Capacitors**
Super capacitors are being explored for efficient mechanical ride-through and as a filter to manage how AI loads affect UPS systems.

06 **Upstream Collaboration**
There's a recognized need for better upstream partnerships and transparent conversations with chip designers and hardware manufacturers to proactively address challenges like cooling and power availability.

Marc Hourican
CTO
CleanArc Data Centers

Brittany Miller
Global SVP, infrastructure Development
NTT Data Centers

Ryan Mallory
President & COO
Flexential

Karen Petersburg
VP, Data Center Development & Construction
PowerHouse Data Centers

EXCLUSIVE INTERVIEW WITH...

PHILL LAWSON-SHANKS, CHIEF INNOVATION OFFICER, **ALIGNED DATA CENTERS**

REGISTER YOUR INTEREST IN ATTENDING IN 2026