COMPUTE

AI INFRA SUMMIT

Born from our legacy of launching AI hardware startups and bleeding-edge compute products, the compute track is our bread and butter. Join the world's leading AI architects and hardware pioneers making AI training and inference faster, more efficient, and more affordable for enterprises. This is the engine room of AI Infra Summit.

Key themes

- Efficient Inference Compute
- AI for Chip Design
- Scheduling & Orchestration Strategies
- Inference Cost Optimization Strategies
- Observability & Monitoring

Who's Attending?

aws | AMD | arm | BROADCOM | cadence | cerebras | cisco
CoreWeave | credo ai | d-Matrix | DELL | FURIOSA | Google | groq
Hewlett Packard Enterprise | HITACHI | IBM | intel | Lambda | Meta | Microsoft
NVIDIA | ORACLE | Qualcomm | rebellions_ | SambaNova | SIEMENS | SK hynix

Chief Technology Officer // Chief AI Officer // Chief Data Officer // Chief Information Security Officer // Chief Innovation Officer // SVP/EVP/VP/Head of AI // SVP/EVP/VP/Head of Infrastructure // SVP/EVP/VP/Head of Engineering // SVP/EVP/VP/Head of Hardware // SVP/EVP/VP/Head of Networking // Chief Architect // Director AI Systems // AI/ML Engineer/Architect // Systems Engineer/Architect // Hardware Engineer/Architect // Distinguished Engineer/Architect

Previous Sponsors

aws | AMD | BROADCOM | cadence | cerebras
Google | intel | Microsoft | NVIDIA | Qualcomm