

# HPN 7.0

## A Data Center Network for Large Language Model Training

Jiaqi Gao

Staff Engineer @ Alibaba Cloud



OCT 15-17, 2024  
SAN JOSE, CA





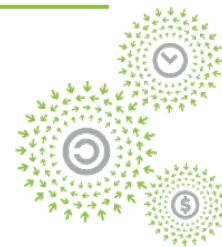
NETWORKING

# HPN 7.0

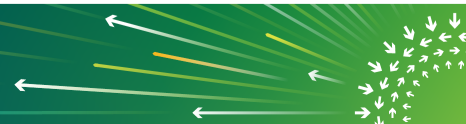
## A Data Center Network for Large Language Model Training

Jiaqi Gao

Staff Engineer @ Alibaba Cloud

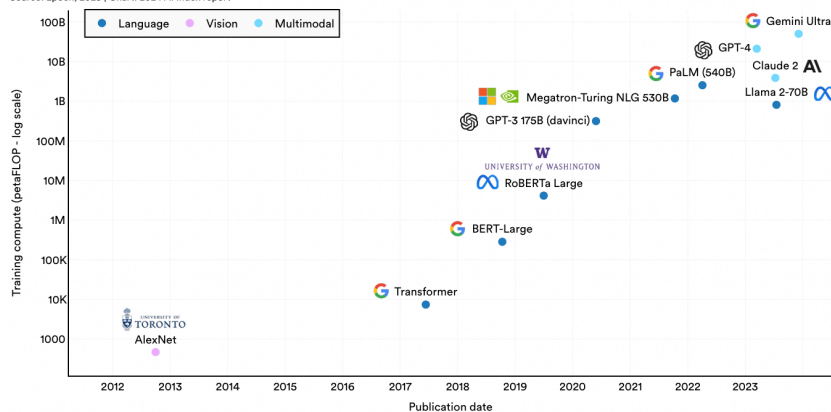


**OPEN**  
PLATINUM™



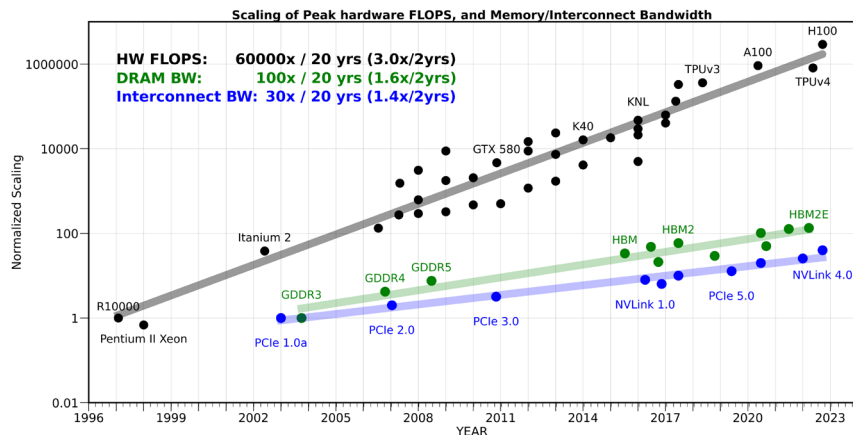
## Training compute of notable machine learning models by domain, 2012–23

Source: Epoch, 2023 | Chart: 2024 AI Index report



Transformer: 750x / 2yrs  
Moore's Law: 2x / 2yrs

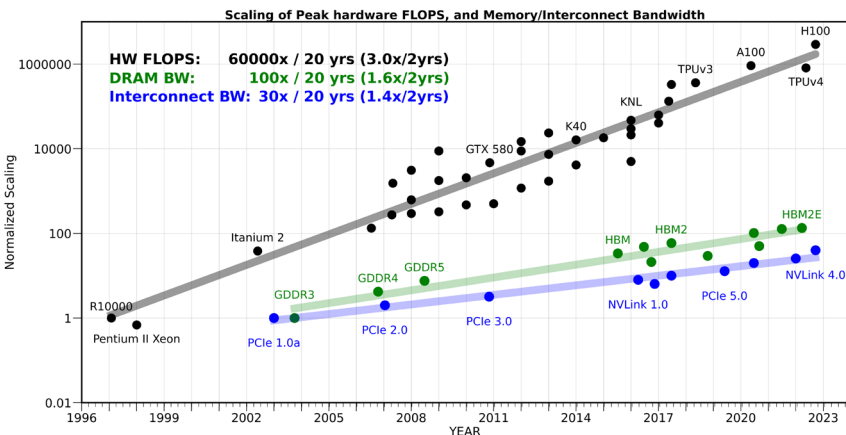
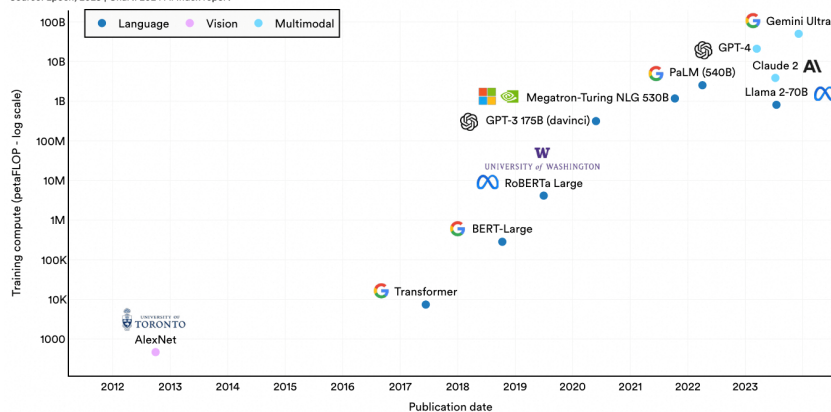
In the age of LLMs,  
Exascale Compute Power  
is the core competitiveness



<https://github.com/swyxio/ai-notes/blob/main/INFRA.md>

## Training compute of notable machine learning models by domain, 2012–23

Source: Epoch, 2023 | Chart: 2024 AI Index report



Transformer: 750x / 2yrs  
 Moore's Law: 2x / 2yrs

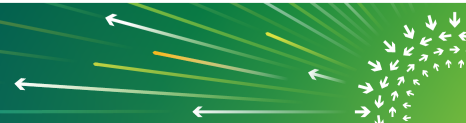
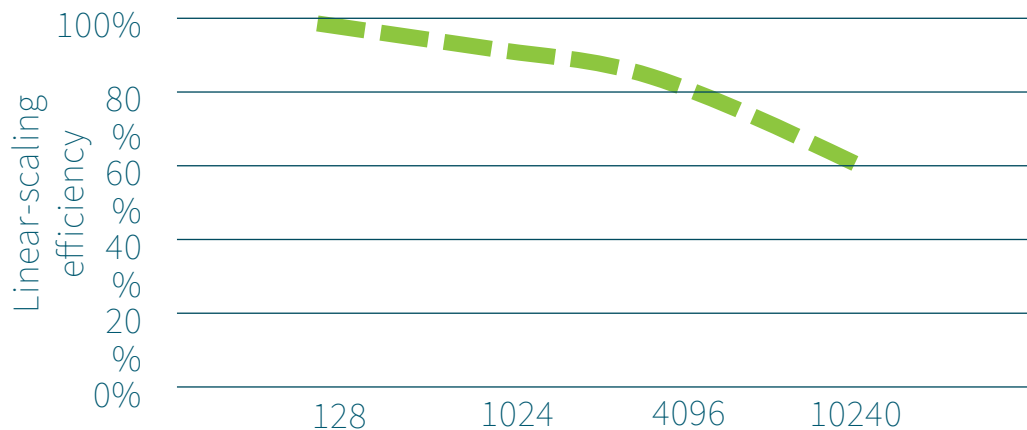
In the age of LLMs,  
 Exascale Compute Power  
 is the core competitiveness

stems from

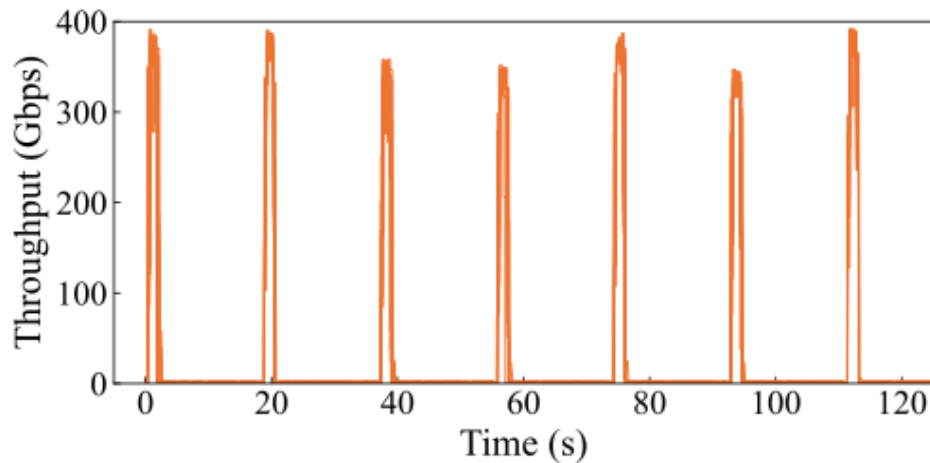
Accelerator Scale-up & out

<https://github.com/swyxio/ai-notes/blob/main/INFRA.md>

# Large Scale Accelerator Cluster != Exascale Computing Power



## New challenges for LLM training workload

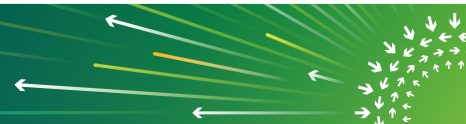


Low entropy

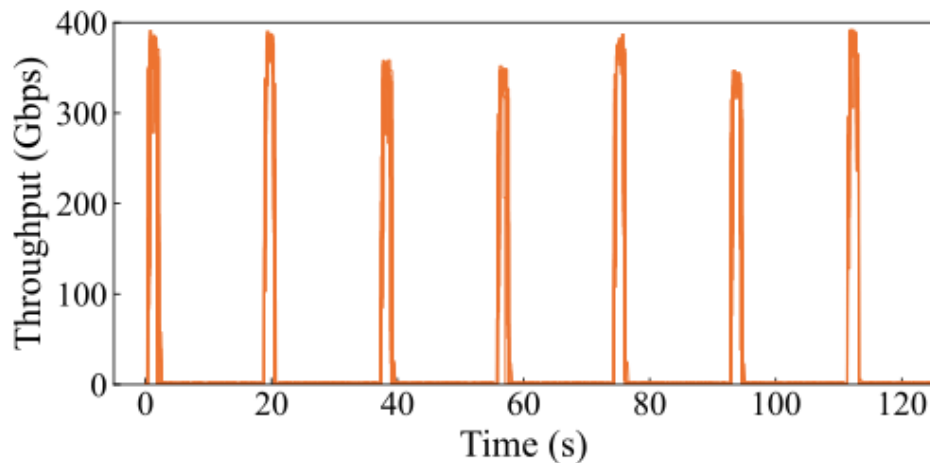
High volume

Globally synchronized

Large scale & Long-lasting



## New challenges for LLM training workload



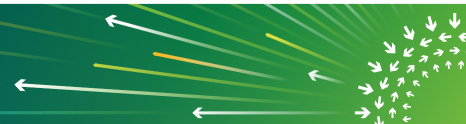
General-purpose Compute Cluster

Hash imbalance

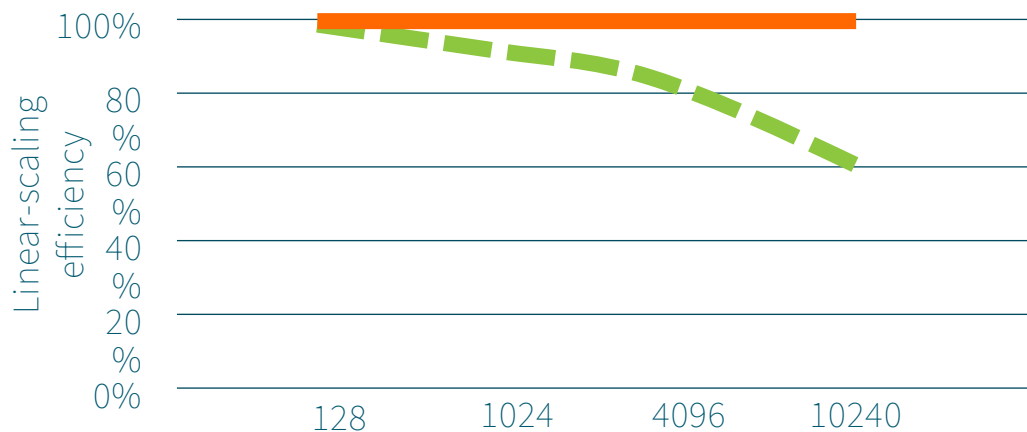
- Low link utilization
- Longer flow completion time

Failure-prone

- Job frequently restarted
- Low cluster availability



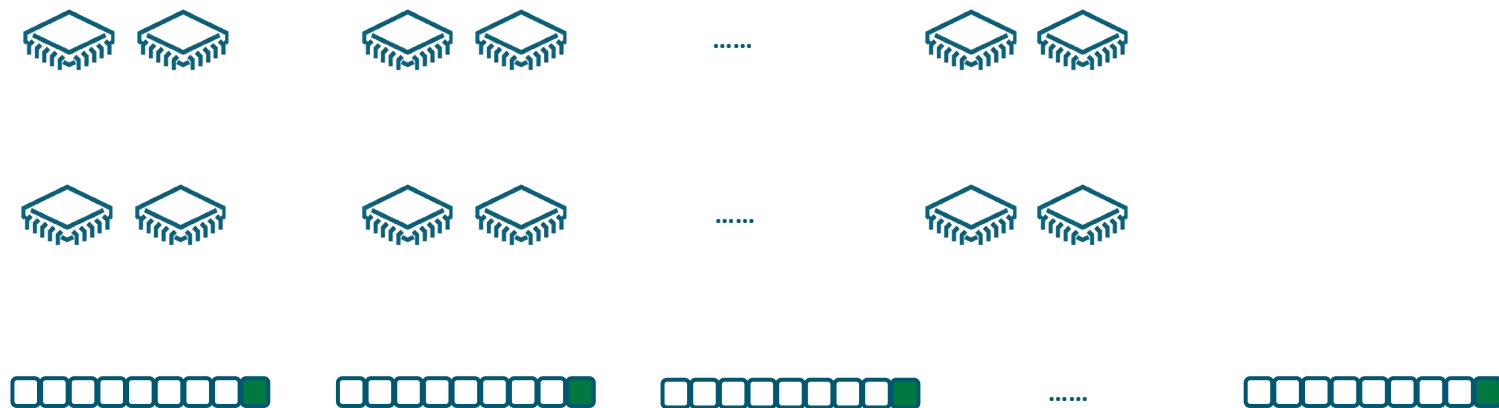
## Large Scale **Accelerator** Cluster == Exascale Computing Power



Consistent high performance interconnect



# HPN 7.0: Predictable Data Center Network for AI Infra

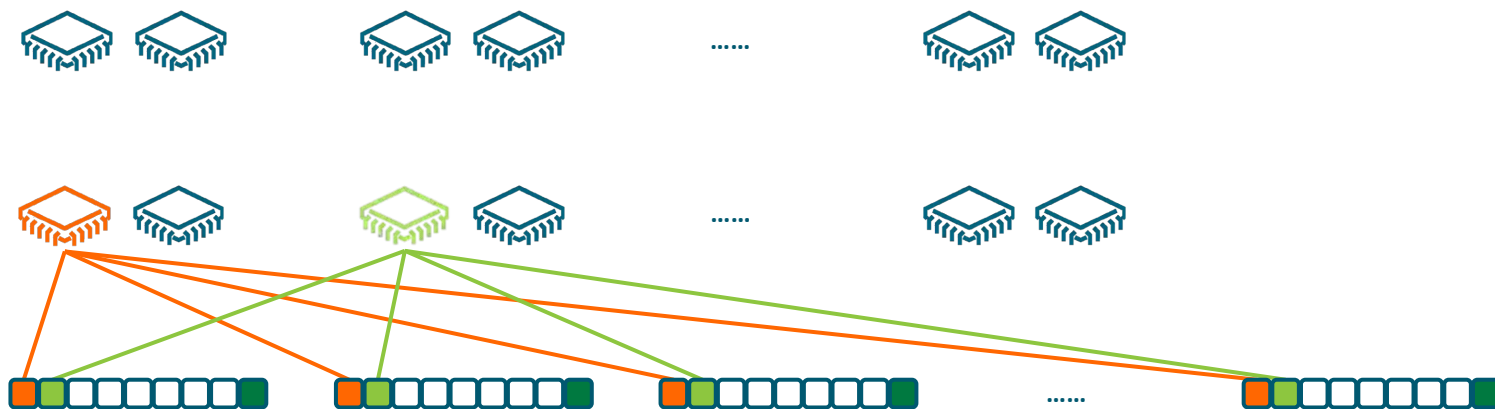


A Two-tier

topology

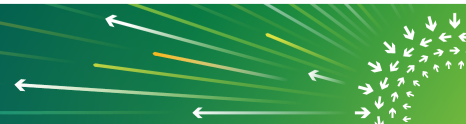


# HPN 7.0: Predictable Data Center Network for AI Infra

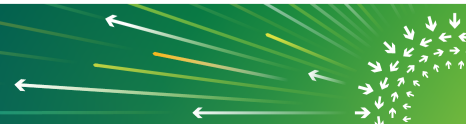
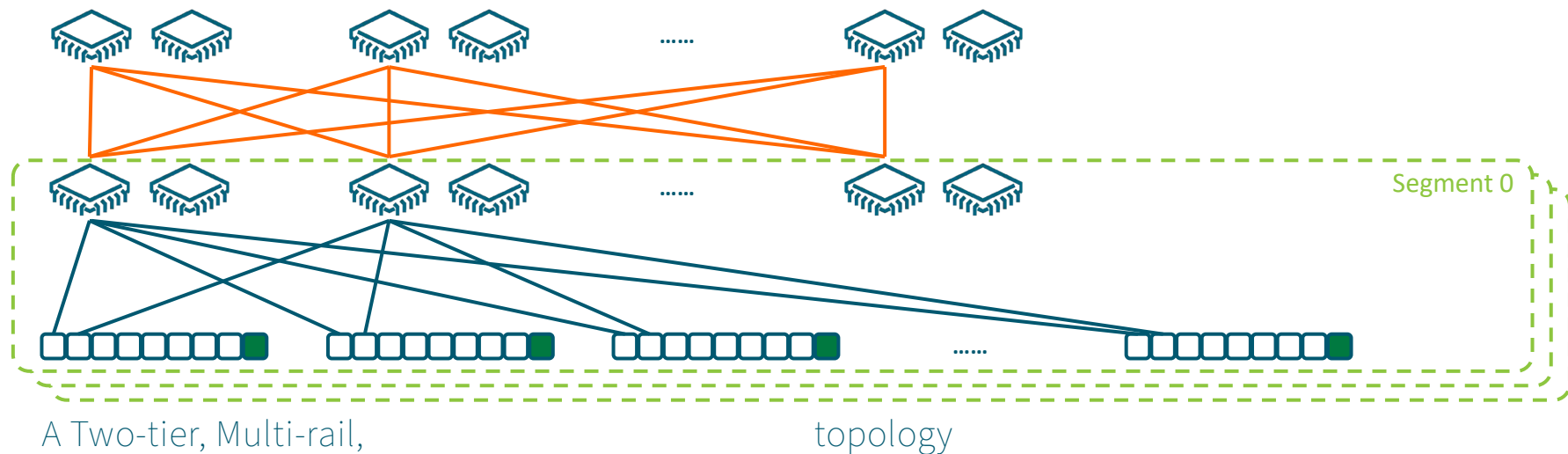


A Two-tier, Multi-rail,

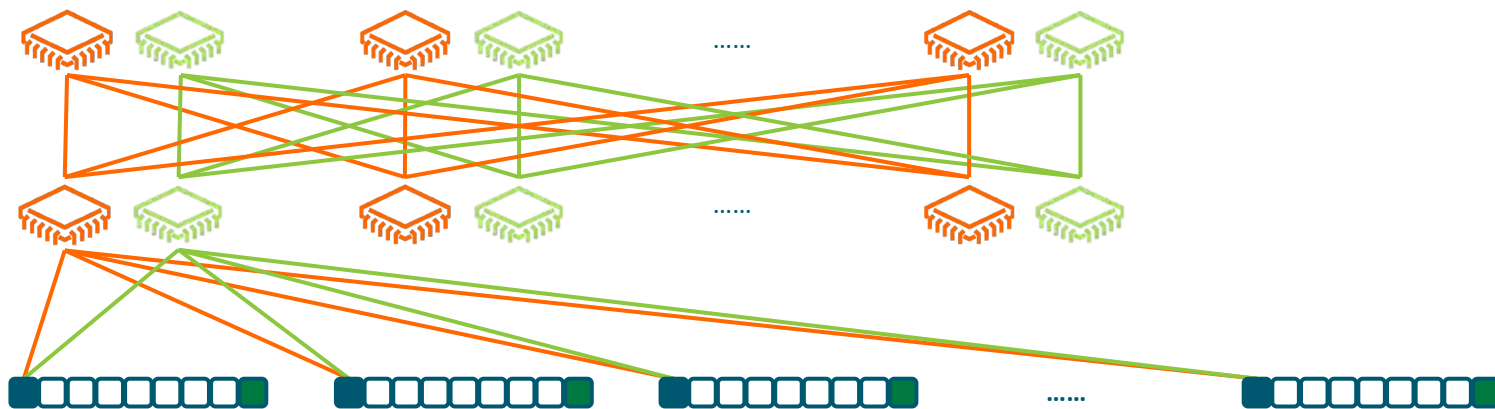
topology



# HPN 7.0: Predictable Data Center Network for AI Infra

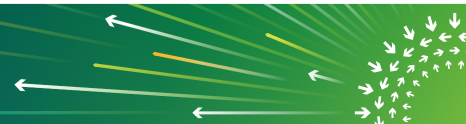


## HPN 7.0: Predictable Data Center Network for AI Infra

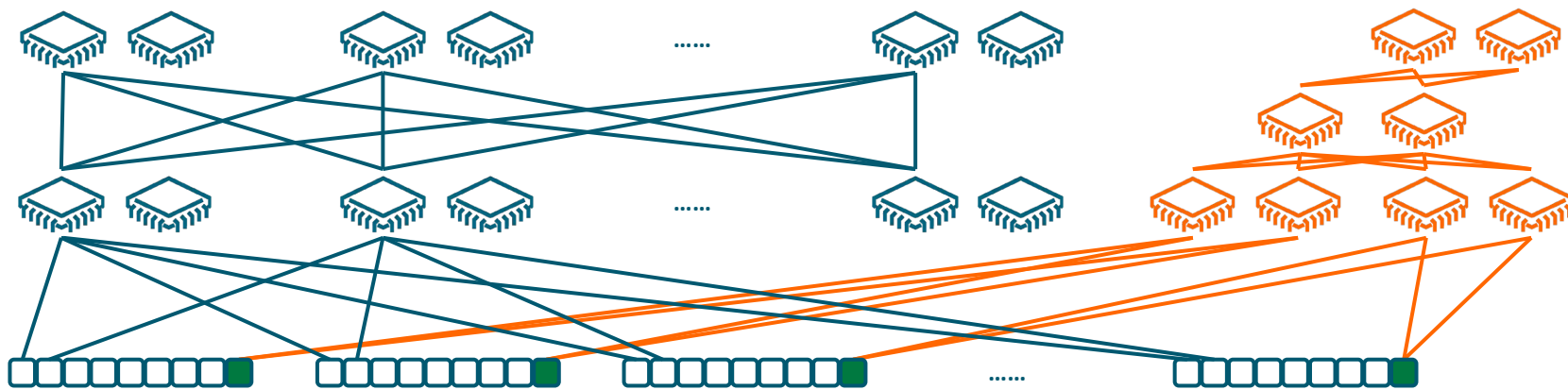


A Two-tier, Multi-rail, Dual-plane,

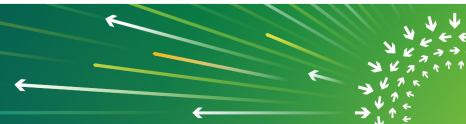
topology



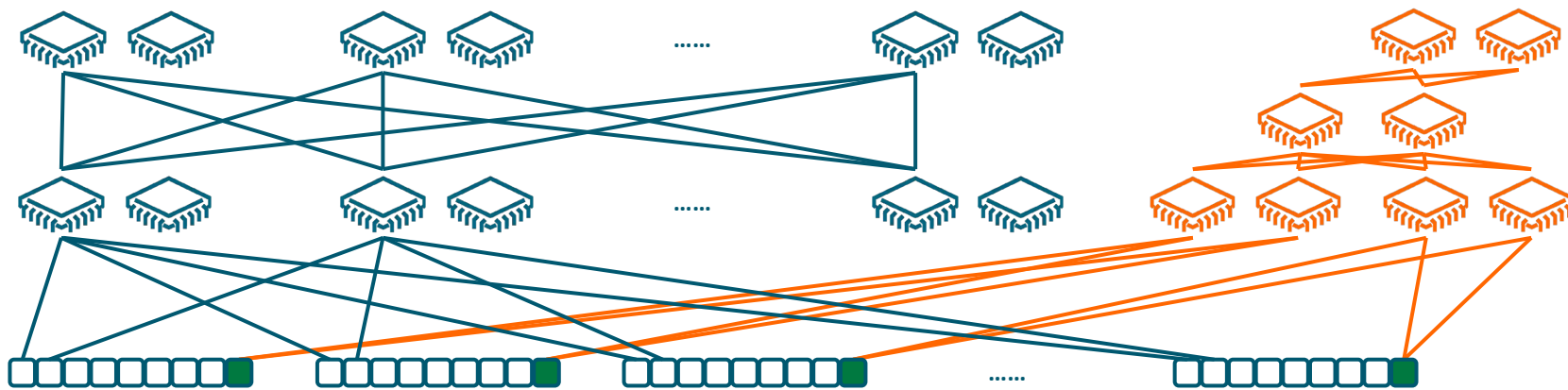
# HPN 7.0: Predictable Data Center Network for AI Infra



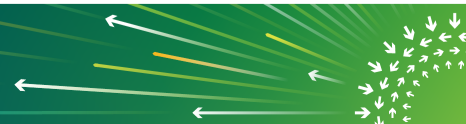
A Two-tier, Multi-rail, Dual-plane, Front/back-end topology



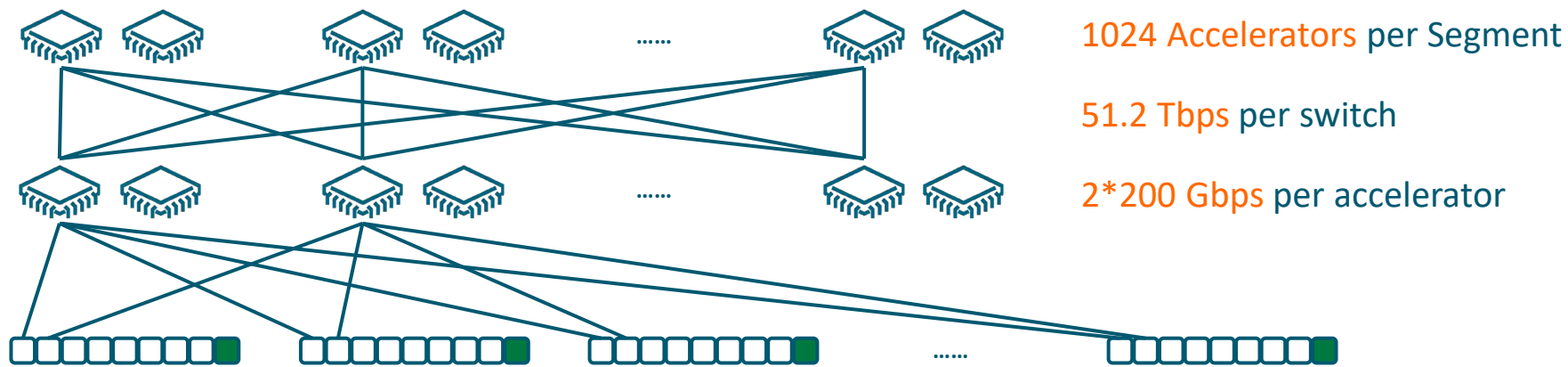
## HPN 7.0: Predictable Data Center Network for AI Infra



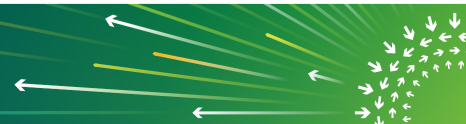
A Two-tier, Multi-rail, Dual-plane, Front/back-end topology that supports up to 100k accelerators



## HPN 7.0: Predictable Data Center Network for AI Infra

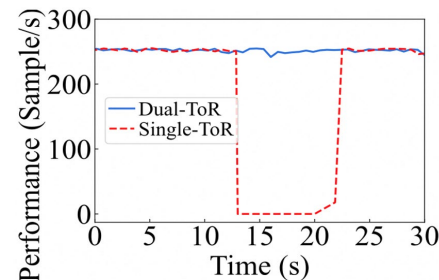
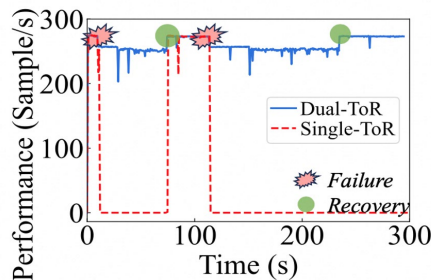
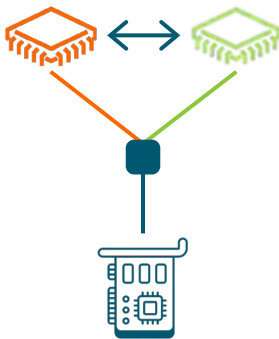


A Two-tier, Multi-rail, Dual-plane, Front/back-end topology that supports up to 100k accelerators



# HPN 7.0: Predictable Data Center Network for AI Infra

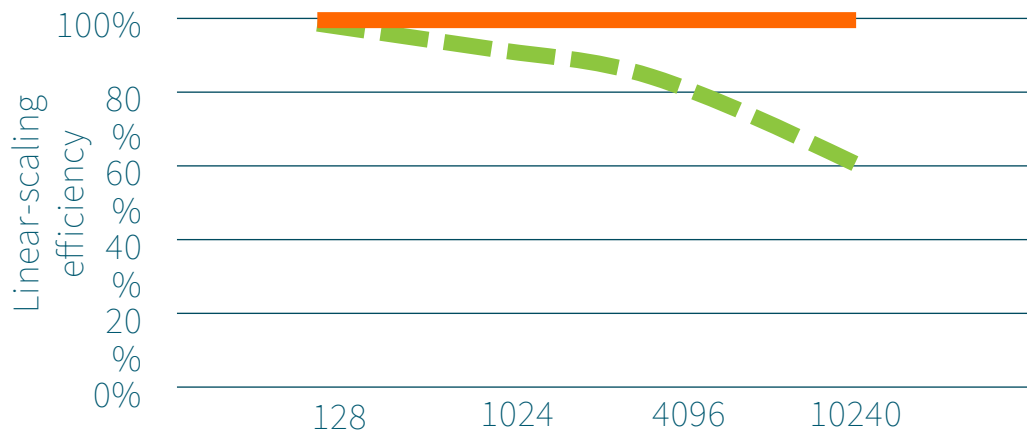
Non-stacked dual-ToR solution, one bond interface per accelerator



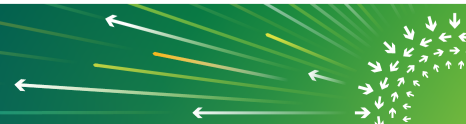
No single link failure disconnects accelerator from the network



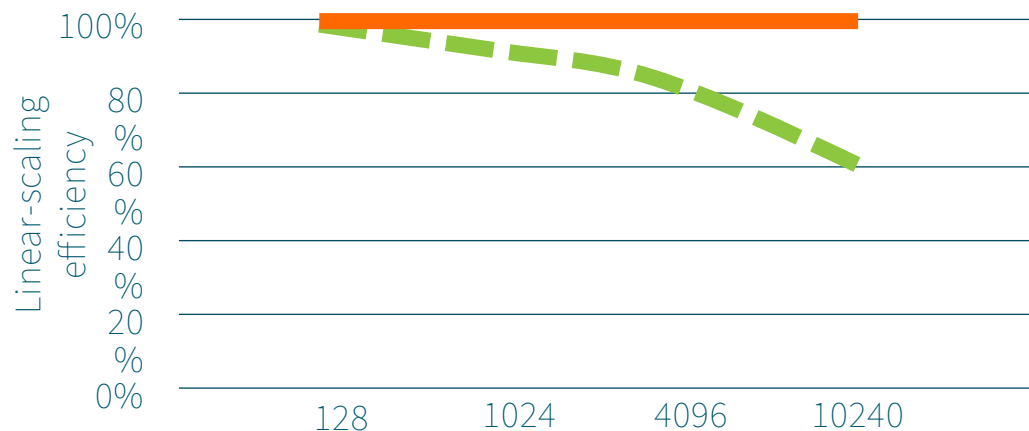
# Large Scale Accelerator Cluster == Exascale Computing Power



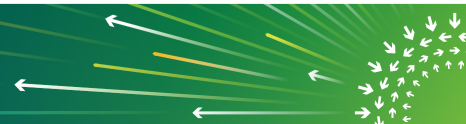
Consistent high performance interconnect



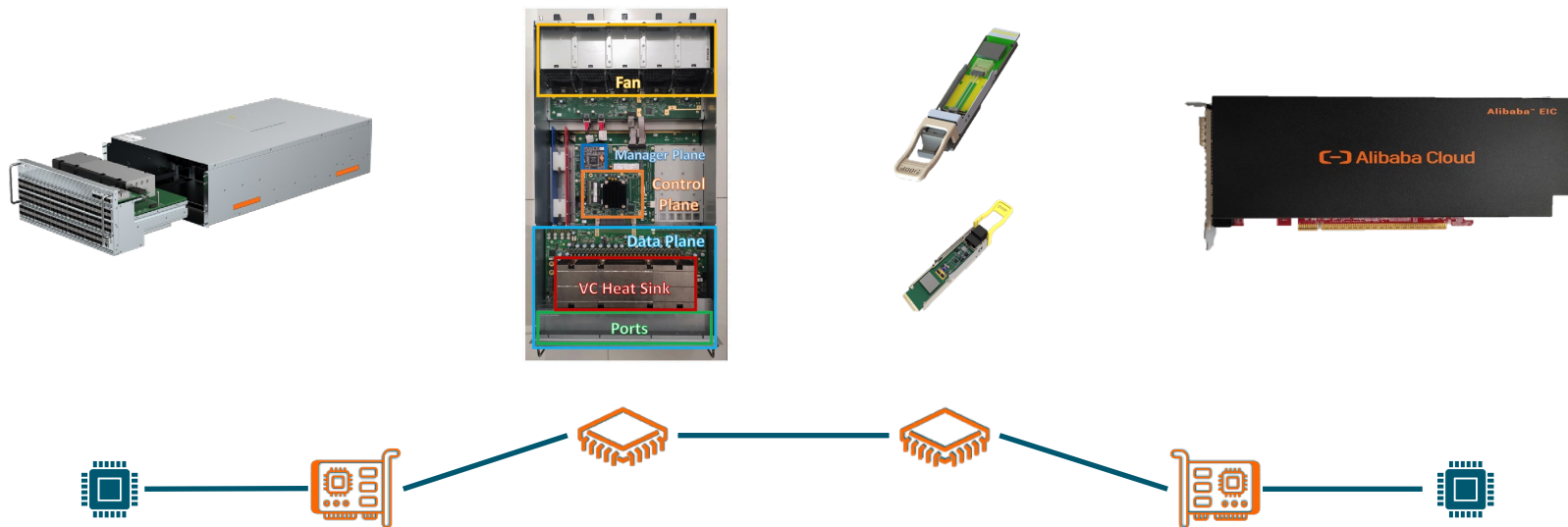
# Large Scale Accelerator Cluster == Exascale Computing Power



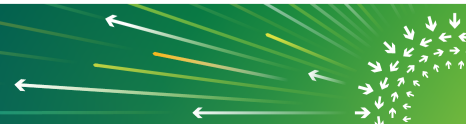
Consistent high performance interconnect & Computation communication Co-optimize



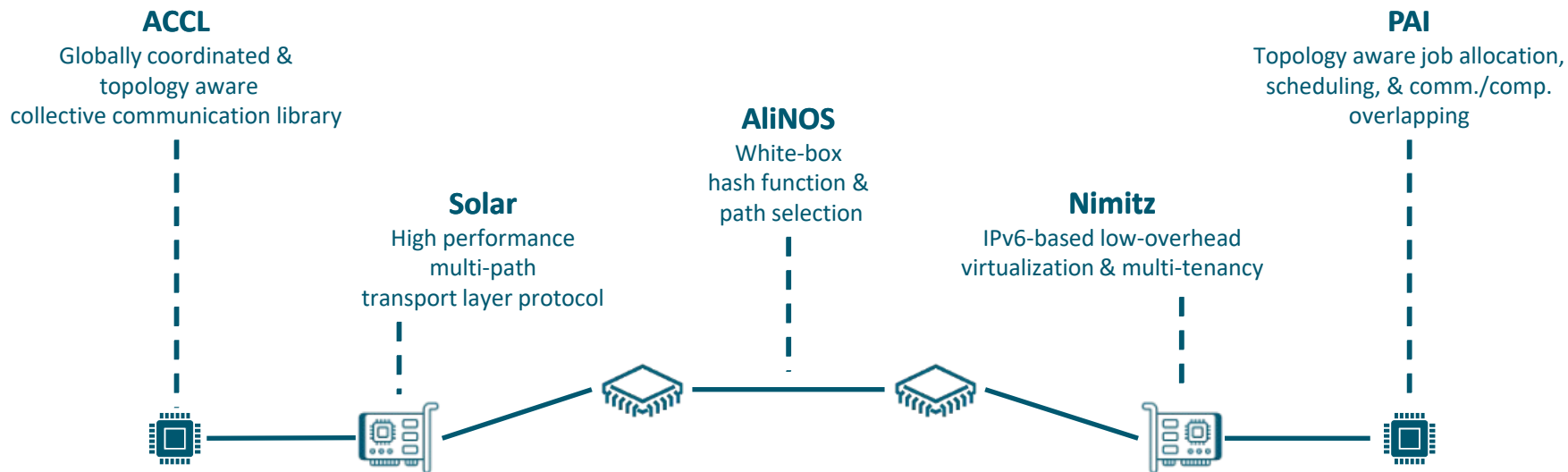
# HPN 7.0: Computation communication co-optimization



In house designed & developed hardware for customization, faster iteration, and stability



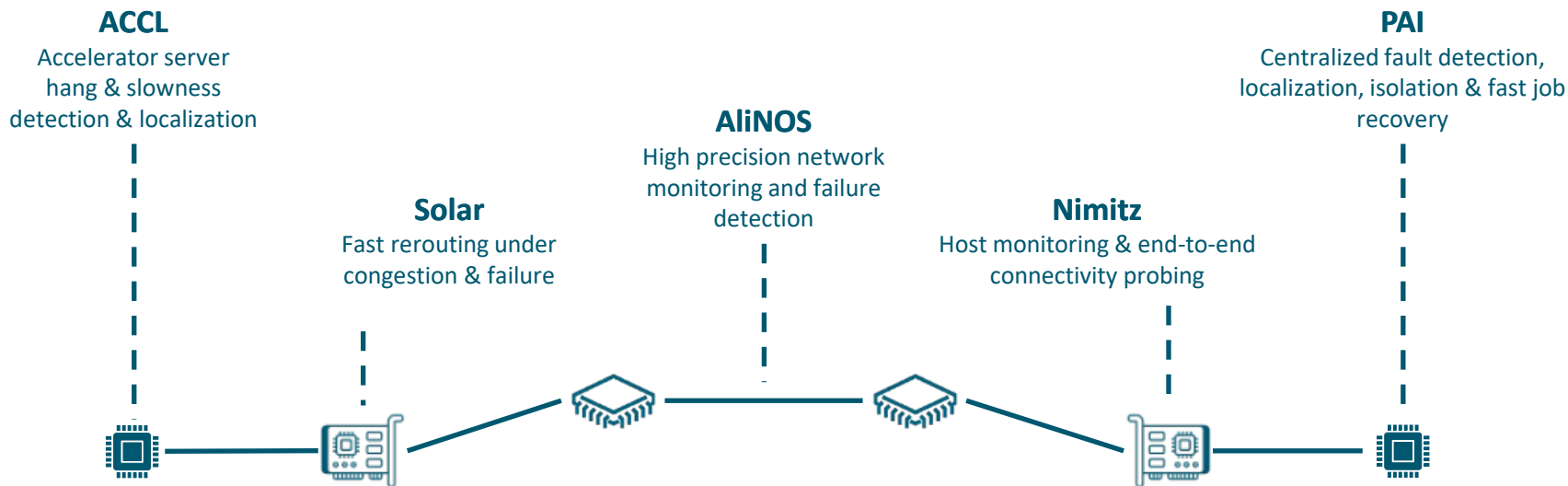
# HPN 7.0: Computation communication co-optimization



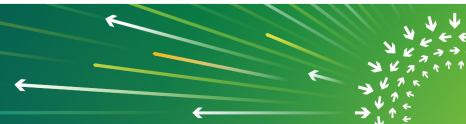
Optimized software stack & transport protocol for optimal performance



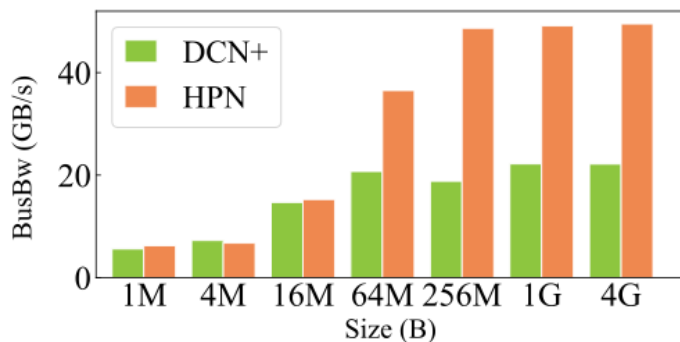
# HPN 7.0: Computation communication co-optimization



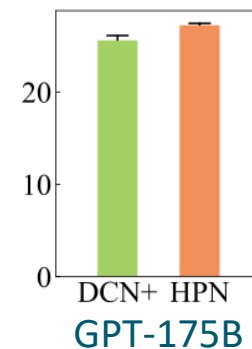
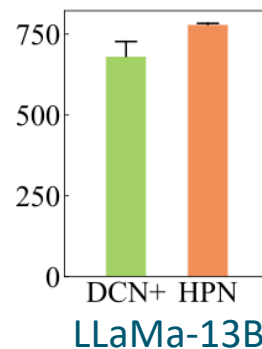
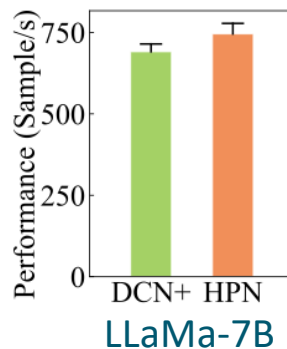
Host-network co-designed management system for fast failure detection, localization, and mitigation



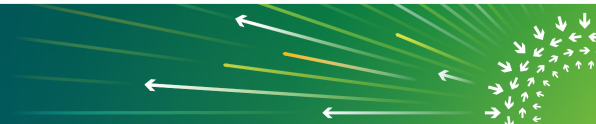
## HPN 7.0: Generation-leap in performance and stability



Increase CCL performance by **100%**



Increase training throughput by **14.4%**



# HPN 7.0: Generation-leap in performance and stability

96%

Linear-scaling efficiency

98.22%

Cluster availability

3 Weeks+

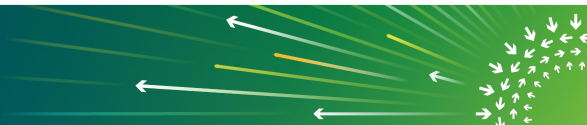
Continuous operation

Minutes

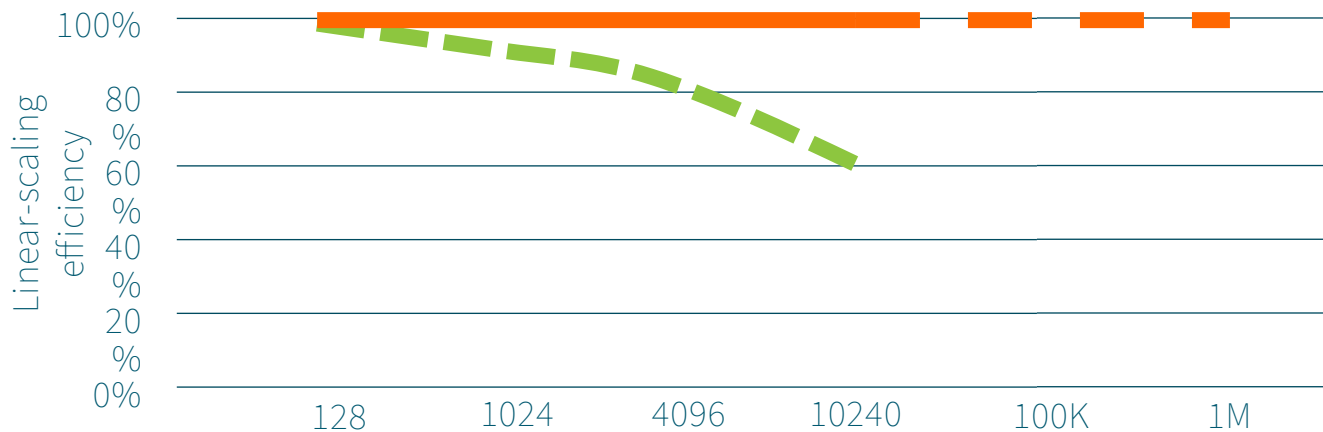
Failure recovery time

\* 5K Accelerators Ultra-Large AI Training Tasks

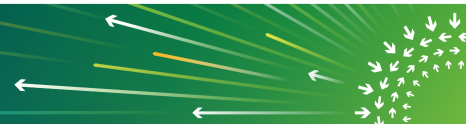
\* Source: Alibaba Cloud online cluster



## Call to Action



Call for talented people and close partnership!





# Thank you!



OCT 15-17, 2024  
SAN JOSE, CA

