Overcome Real World Challenges between Data and AI

Steven Yuan

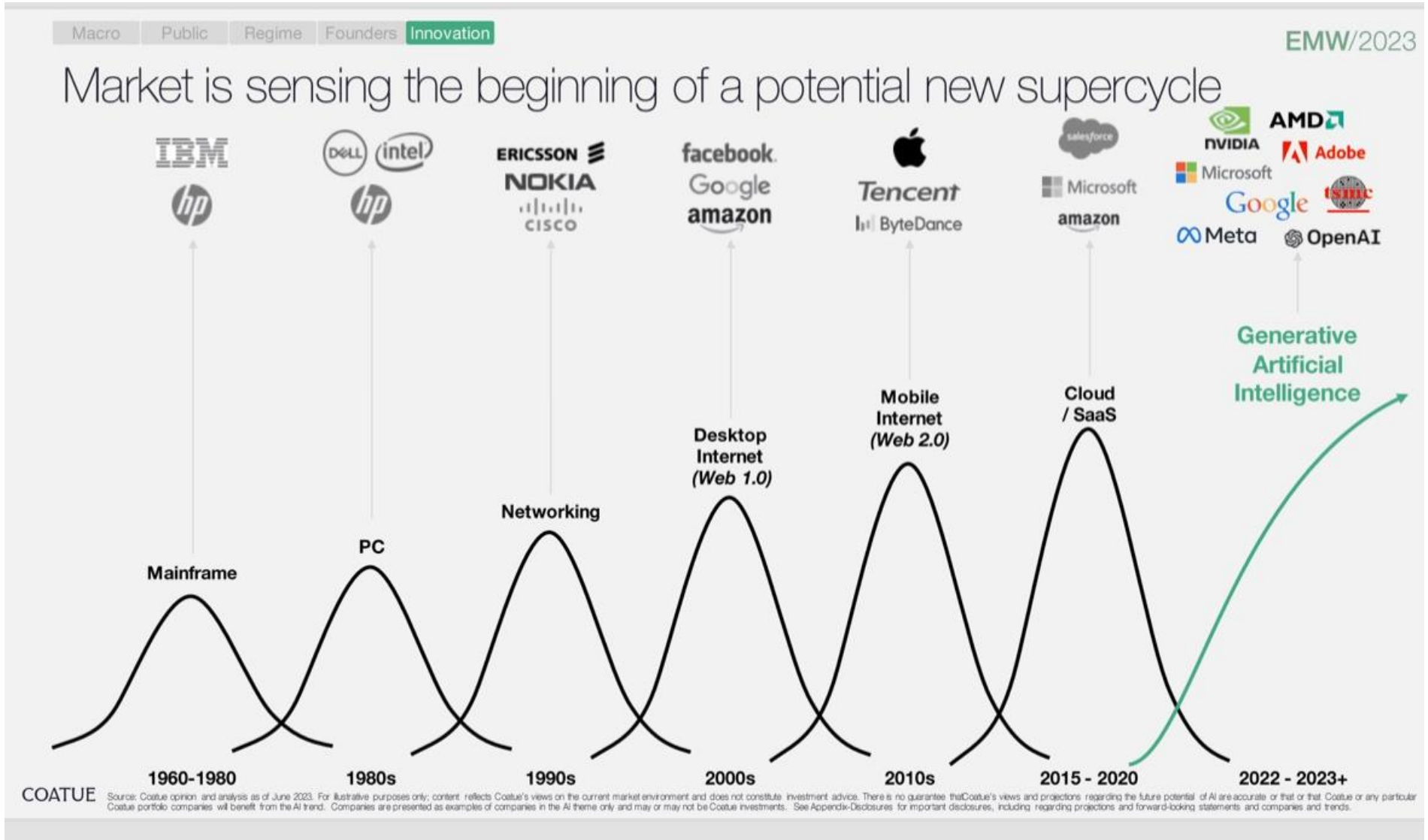
CEO of StorageX.ai

SNIA COMPUTE, MEMORY, AND STORAGE SUMMIT

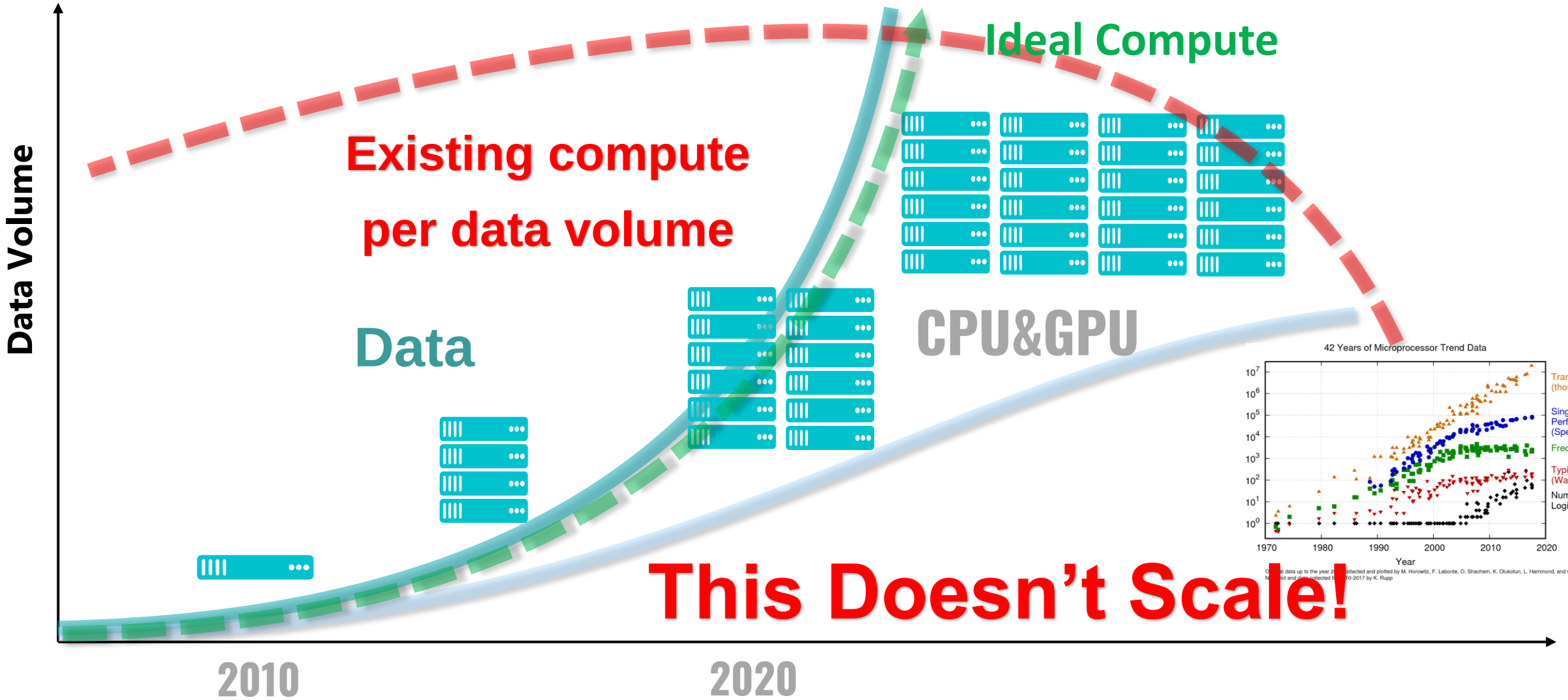
Solutions, Architectures, and Community

VIRTUAL EVENT, MAY 21-22, 2024

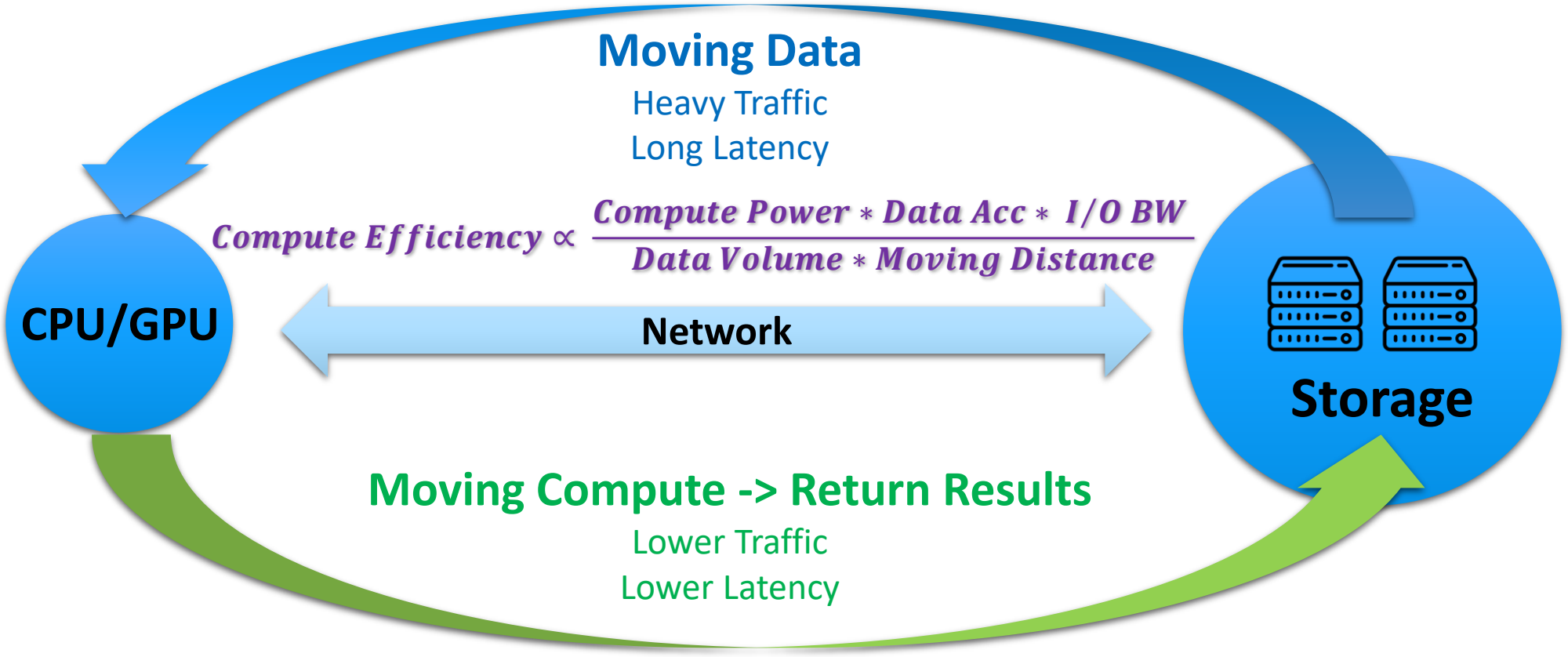
AI2.0: AGI is driving next big cycle



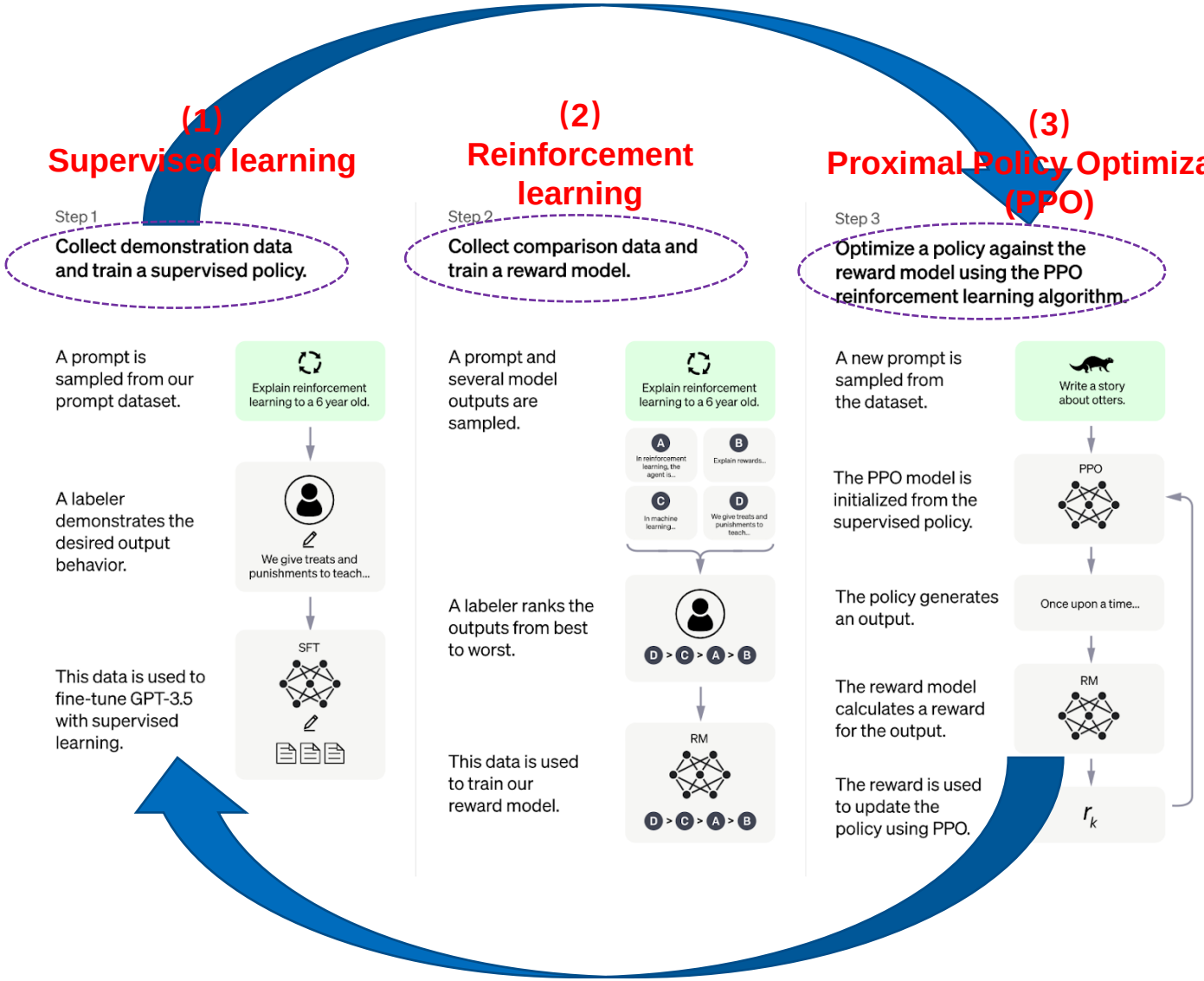
Scaling Issues : Compute + Network + Storage Bottleneck



The ‘Dumbbell effect’ causing high compute cost
Data Centric Computing is a key solution



Learning Process of ChatGPT



Data

Data

Data

Data Movements + Compute

Smart Factory: Intelligence needed for efficiency

Automation

Automation tools to reduce labor cost

Key Factor

- Cost reduction
- Production Cycle Time
- Factory Capacity

Digitization

use of data to achieve the intended production goals, or operational efficiency.

Key Factor

Use digital tools for business decision making.
Example:

- ERP, CIM, MES etc
- IoT devices
- Edge Cloud

Artificial Intelligence

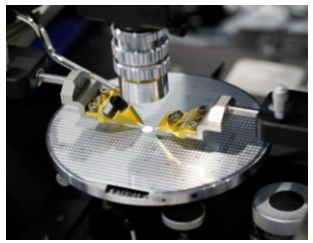
Use intelligence for self monitoring, automated problem identification, unmanned management etc.

Key Factor

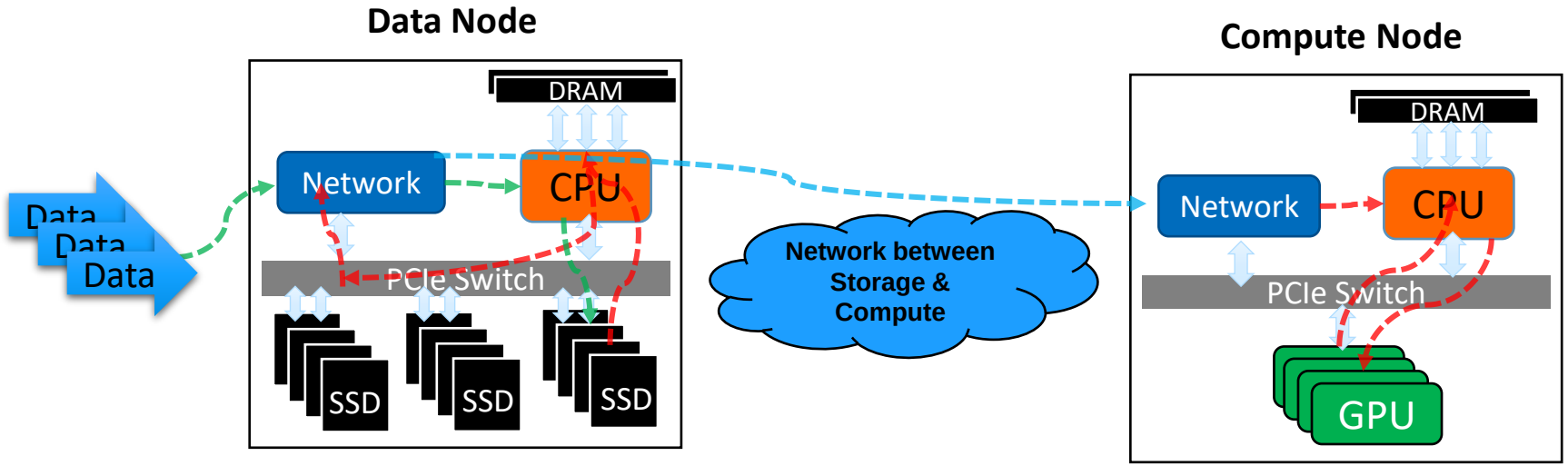
Use AI to achieve lighthouse factory:

- Higher Capacity
- Higher yield
- Faster problem solving
- Shorter cycle time
- More flexibility

Real World Challenges for Data and AI



AOI or other automation tools



Real world Requirements

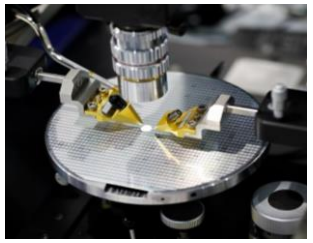
- **Big Data Set:** Large amount of data generated from AOI machinery, Million files per day, could accumulate to PB level data per week.
- **Large files:** can be more than GB per file
- **Faster cycle time:** less than seconds for decision making

Challenges to Infrastructure

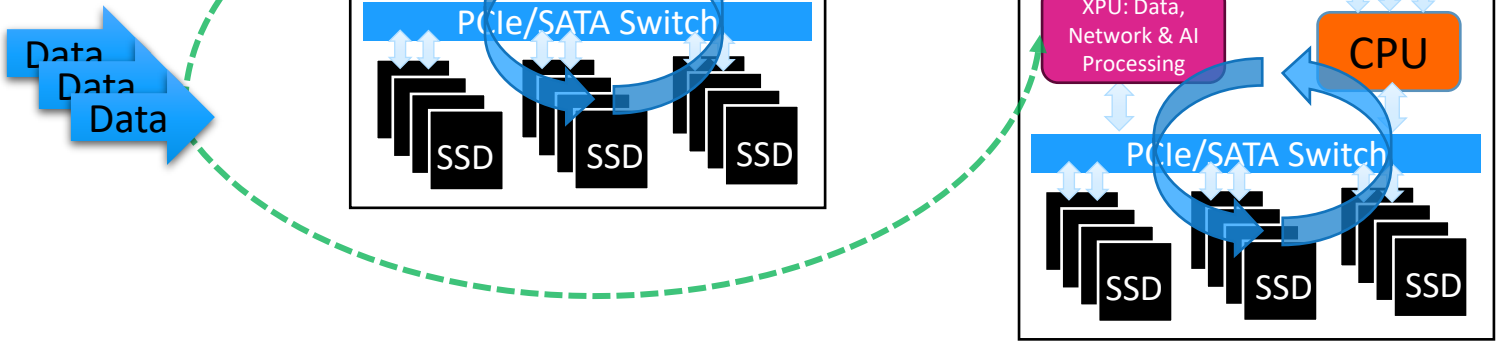
Impact to decision making

- **Storage**: Write time, IOPS, Capacity
- **Network**: Latency, throughput, moving distance
- **Compute**: Process time

Solutions to solve these challenges



AOI or other automation tools



Real world Requirements

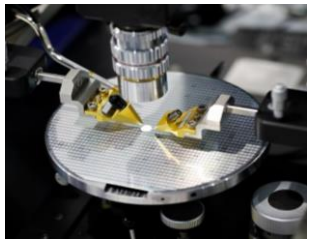
- **Big Data Set:** Large amount of data generated from AOI machinery, Million files per day, could accumulate to PB level data per week.
- **Large files:** can be more than GB per file
- **Faster cycle time:** less than seconds for decision making

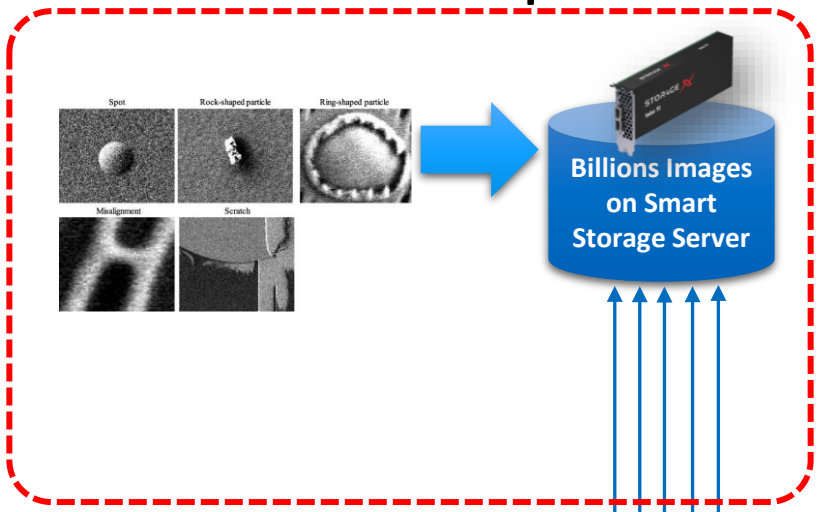
Challenges to Infrastructure

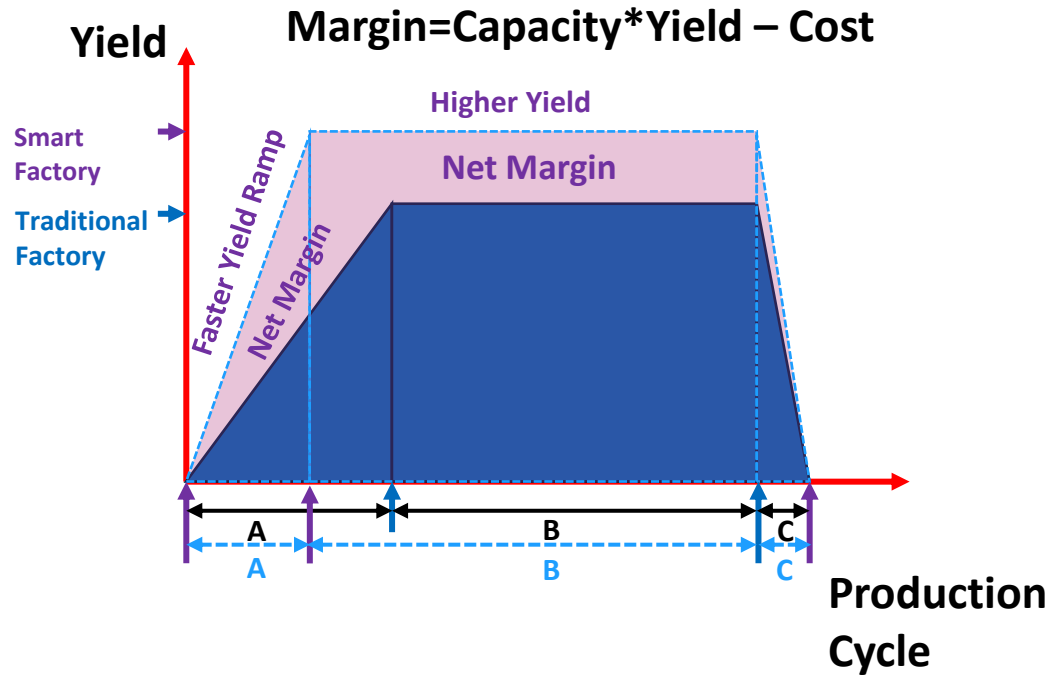
Faster decision making

- **Storage**: Faster write time & shorter data path
- **Network**: Lower Latency, Less data movements
- **Compute**: Faster processing time

Real Time Data Analytics: Impact to the Smart Factory







- **Key Factors**
  - Real time data analytics
  - Low latency data fetching
  - Pattern analysis
  - Faster Commonality Study

Faster
Data Driven Decisions

- **Results**
  - Quicker Failure Identification
  - Faster Problem Solving
  - Higher Yield
  - Faster ramp up time

Smart Factory generates large volume data & requires Low Latency & Real Time Analytics

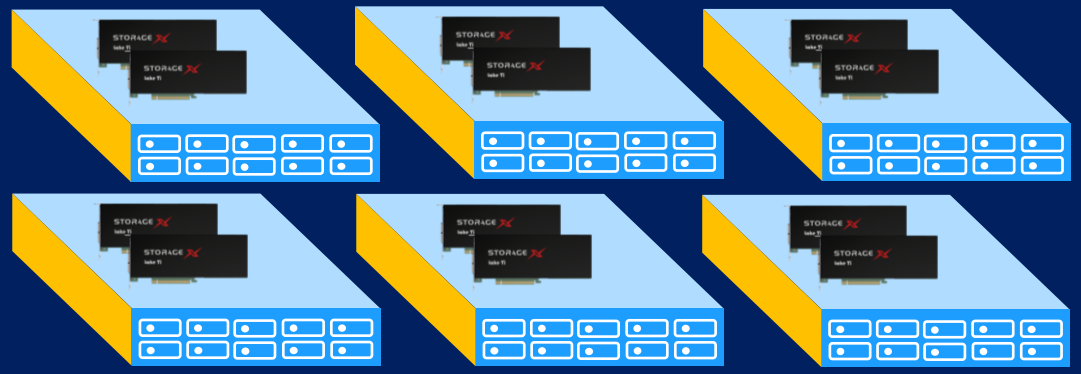
Industrial Data mining & Analytics Platform

SaaS
(CIM/MES)

- Data Collection
- Process Monitoring
- Data Retrieval
- Yield Analysis
- Image Classifications
- Data Comparison
- Defect Analysis
- Process Improvement

PaaS

- Data Collection
- Data Mining
- Data Analytics
- AI Inference
- Smart Data Lake API

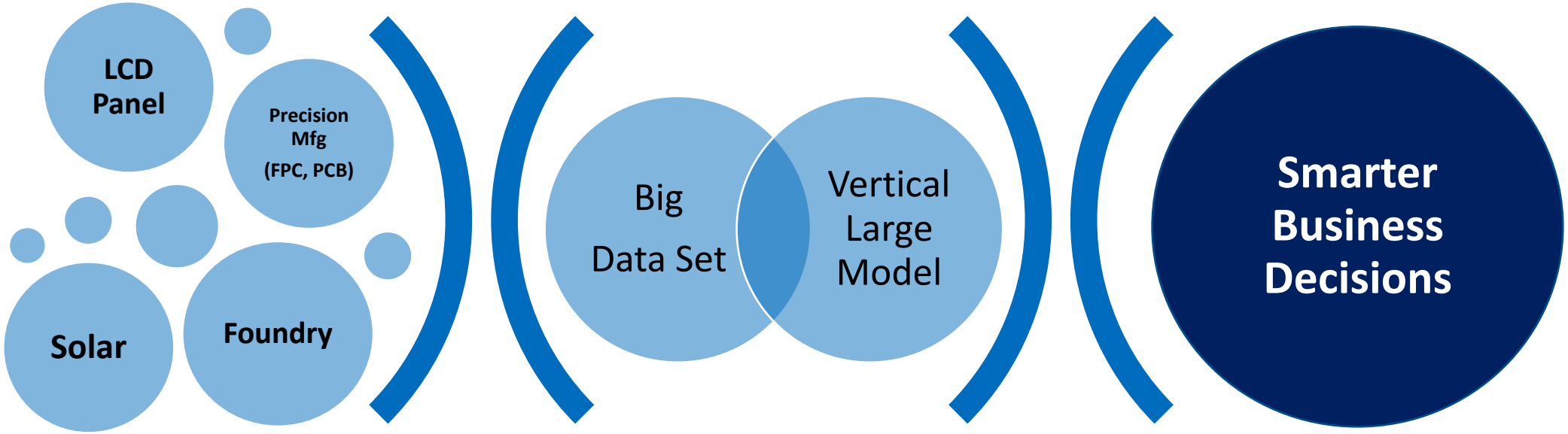


Infrastructure

Smart Factories for different Industry

- LCD/OLED Panel Manufacture
- Foundry
- Solar
- Battery
- 3C Products (PCB, Cellphone, Auto)

Vertical Model leads to better business decisions



Conclusion: Data centric computing is important for data intensive workload

- We live in a world with full of data & growing even faster.
  - Those data be used to create better productivity, for smart and intelligent factories.
- Data can be very “big and fast”, we need better architecture to handle this.
  - Post big challenges to infrastructure, **Compute, network and storage**
- Compute near the data can solve many issues.
  - Lower latency leads to faster cycle time for the factory
  - Then leads to faster business decisions, better efficiency.
- Next, Vertical Large Models can be very important for domain specific areas.
  - Data will be accumulated, analyzed, not sitting around.
  - AI powered commonality studies represents a significant advancement in identifying patterns and similarities within large datasets.
- Faster processing becomes more important than ever.

Please take a moment to rate this session.

Your feedback is important to us.