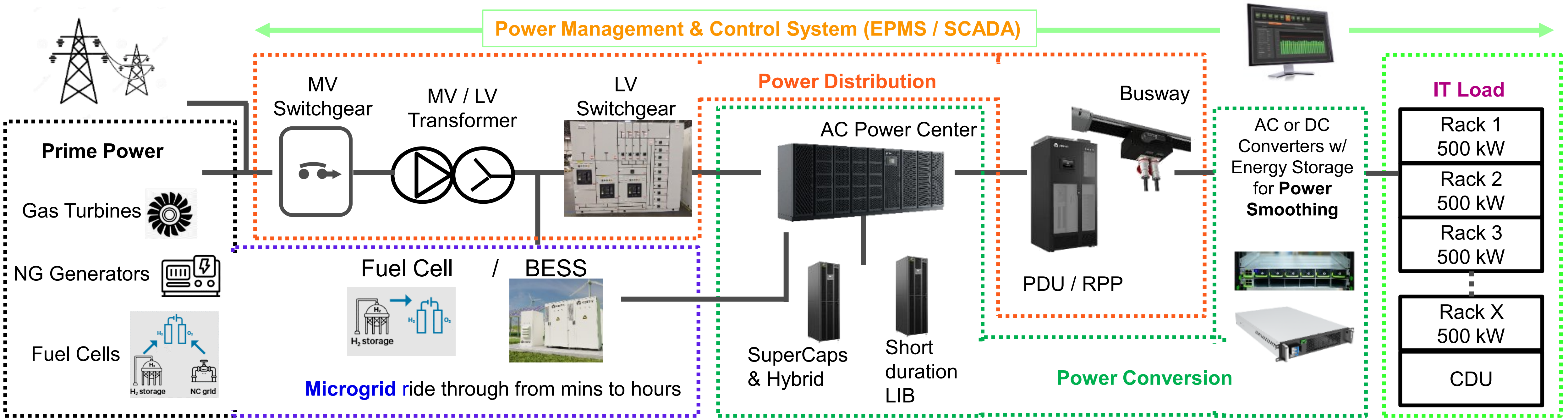
Breakout sessions
Power Management

Evolving Challenges/ Opportunities for Critical Power Train



Prime Power

- Integration of BYOP sources (FC, Gas turbines…)
- Support Renewables through grid services for frequency stabilization
- Mitigate permit challenges for extra power from grid

Backup Power

- Reduce/Eliminate diesel gen starts
- Integration of BYOP (FC, BESS)
- Mitigate permit challenges for fuel-based generators

Power Protection

- Larger Building Blocks (1.5→2→3MW)
- Support for new energy sources and transient loads (beyond VRLA / Li batteries)
- Management of AI loads

Power Distribution

- Distribution changing to accommodate higher voltages within LV limits →575V
- Enabling MV solutions (how far down the power train?)
- Deliver redundancy needed to support training and generative / inference

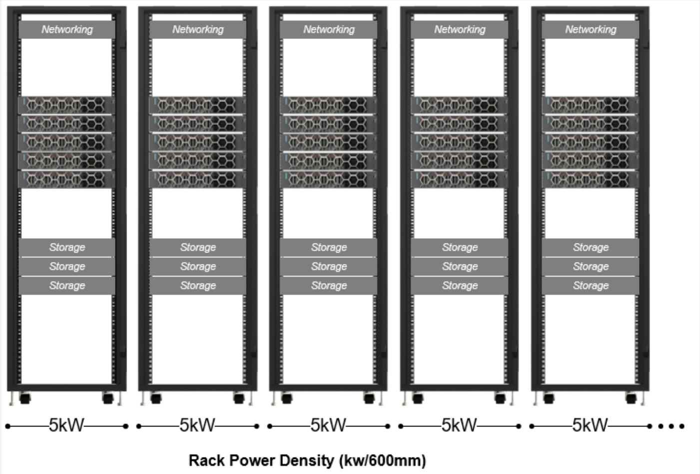
IT and Liquid Cooling Loads

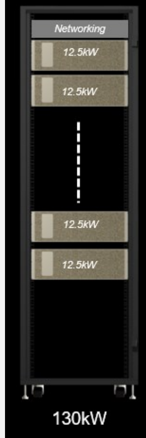
- Higher voltages to support higher rack power density (50→100→500kW)
- Optimize / integrate Power protection for Liquid Cooling
- Rack-based power shelf BBU

The rapid evolution of High Performance is driving data center power densification through all the Power Train components

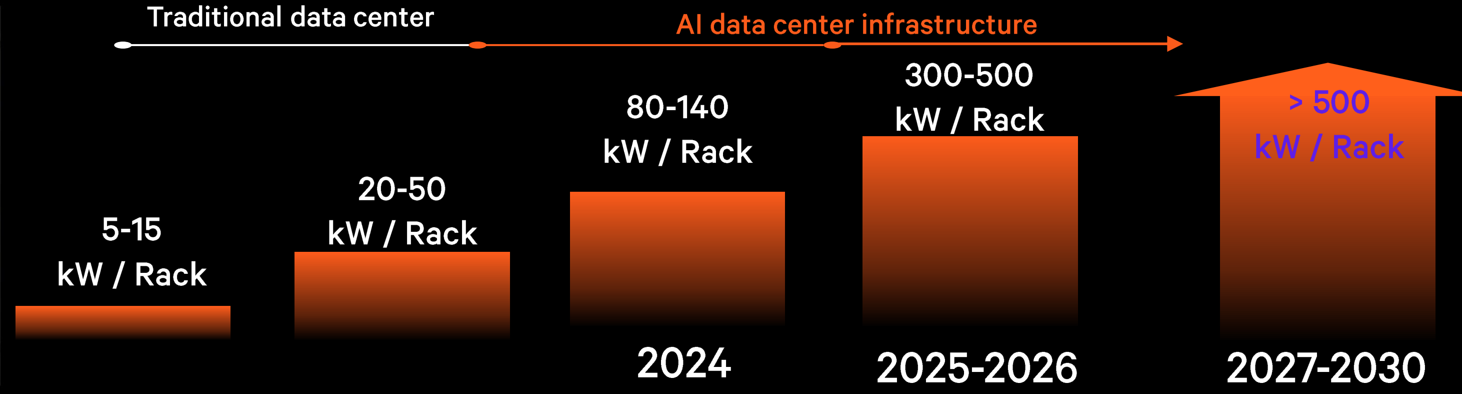
Data Center Evolution & Maturity

	Existing Data Centers	Next-Gen AI Data Centers	Design Principles for AI
Scale	• 3 MW 'capacity blocks' • 10 kW / rack	• 5 ⟶ 10 MW 'capacity blocks' • ⟶130 ⟶ 300⟶500⟶1000+ kW / rack	• Optimize distribution and building block sizes • Fundamentally re-size system components • Advance sustainability (including heat reuse)
Complexity	• Air-cooling only; independent power & cooling designs • Steady state loading • Established 'Tier'-based redundancy architectures	• Liquid + air cooling; highl interdependent with power • AI power demand spikes • Redundancy bifurcation	• Comprehensive power + cooling reference designs: rack ⟶ cluster ⟶ data center • Design range covering maximum capacity (N) & maximum availability (N+)

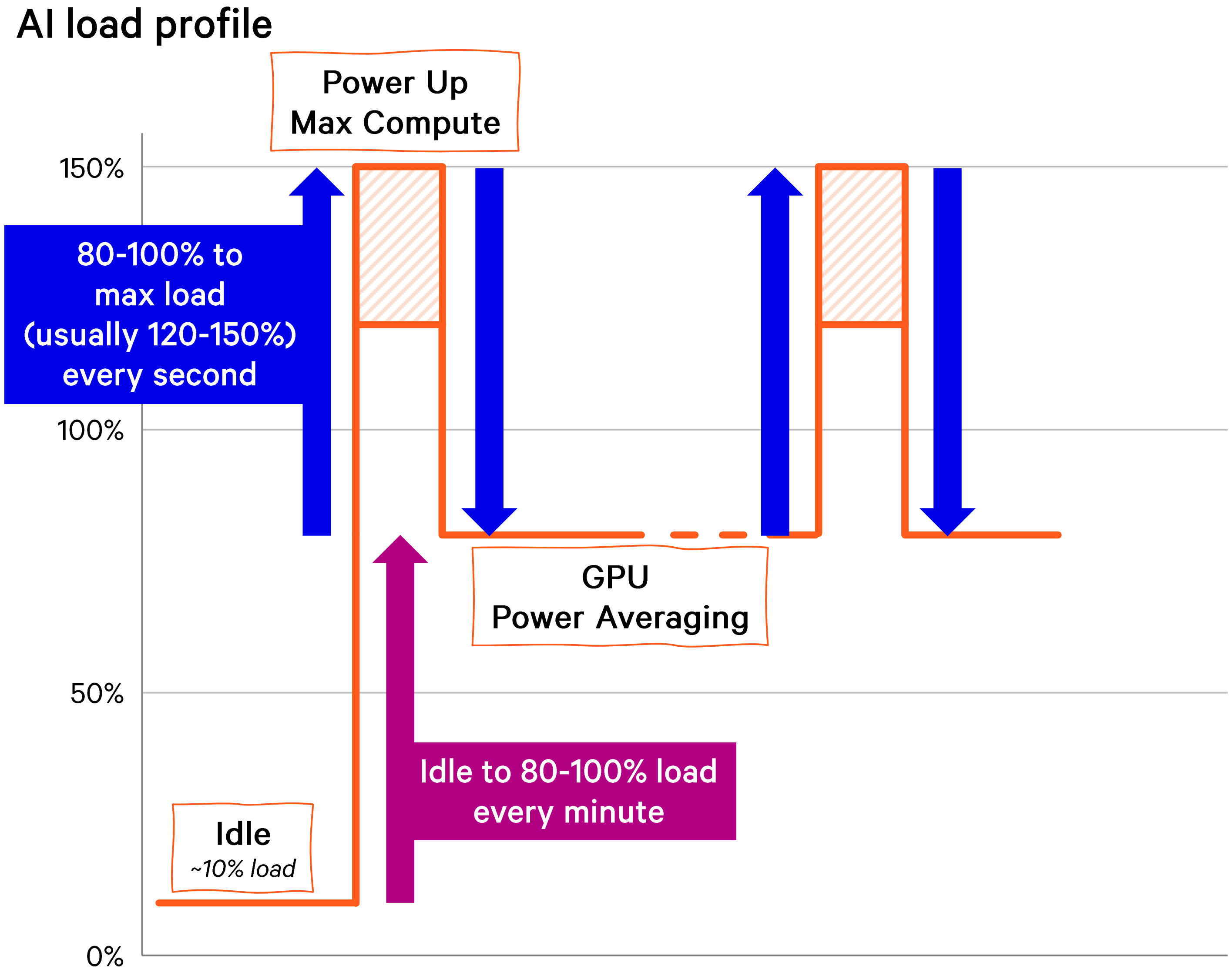




Rack power density growth is accelerating



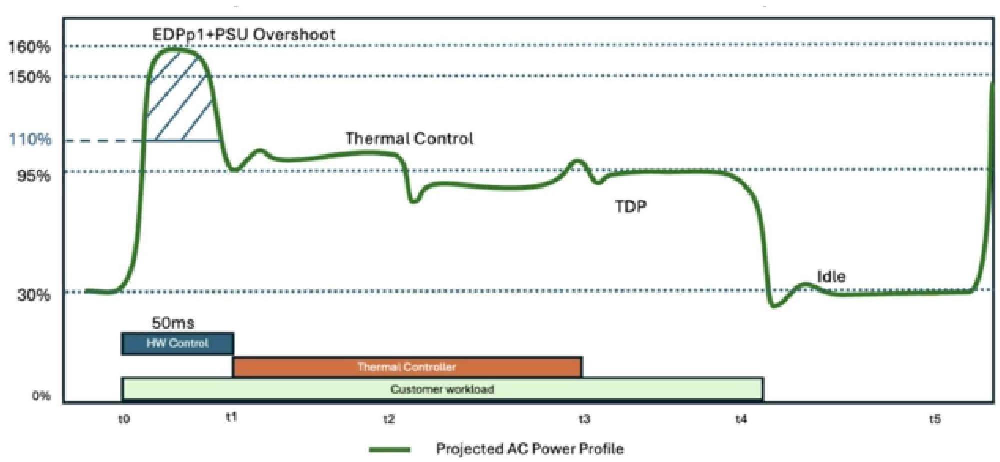
As GPUs become ubiquitous in data halls, they bring in **challenging load profiles** for supporting power infrastructure.



Load profile of GPUs running AI workloads is unique and very different from traditional IT equipment.

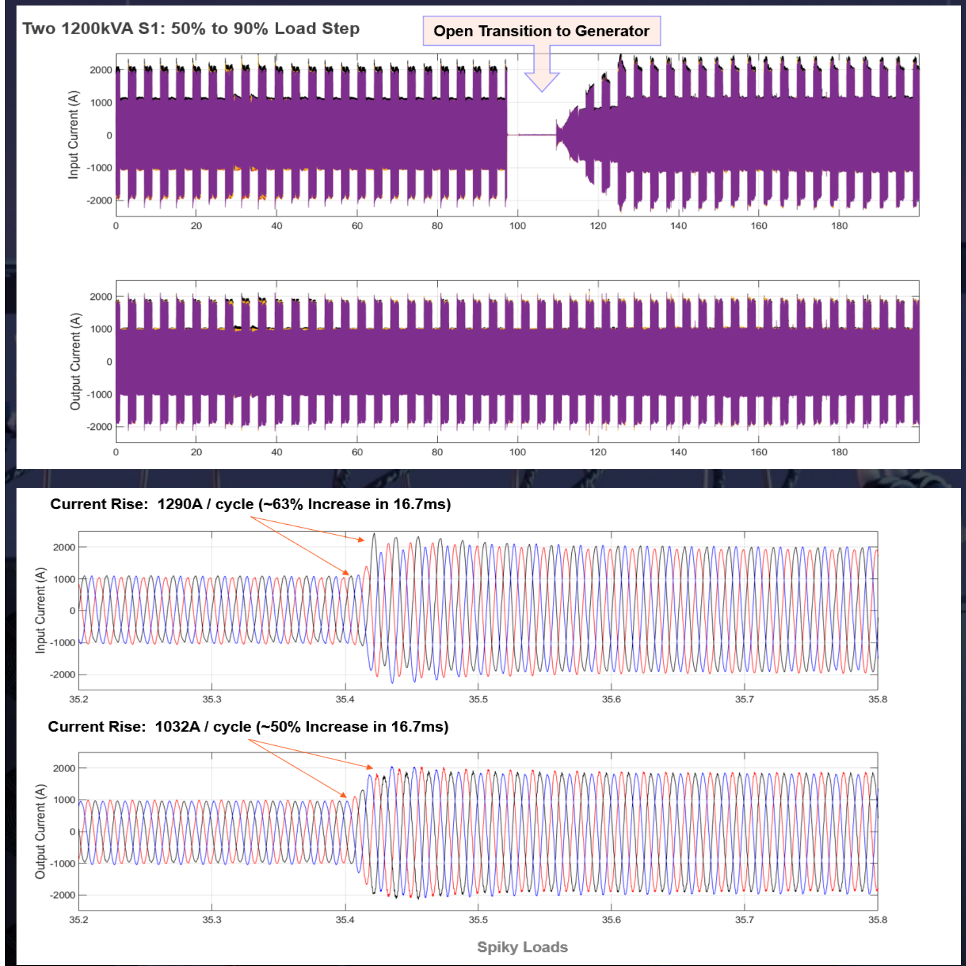
Vertiv Trinergy, an AI ready UPS

Artificial Intelligence (AI) has a unique load profile that presents challenges for the data center powertrain.



To properly protect critical IT loads in AI applications, Vertiv Trinergy:

1. Is able to handle the load profile without performance degradation and component degradation
2. Acts as a shield to protect the source (utility and generator) from seeing step loads
3. Is designed to support large increase in rack power density by working at higher voltages



The testing of Liebert EXL S1 UPS demonstrated the system's capability to provide **reliable power supply to AI power loads** across multiple scenarios, including operation under generator, utility, and transition conditions.

Vertiv Trinergy workloads management

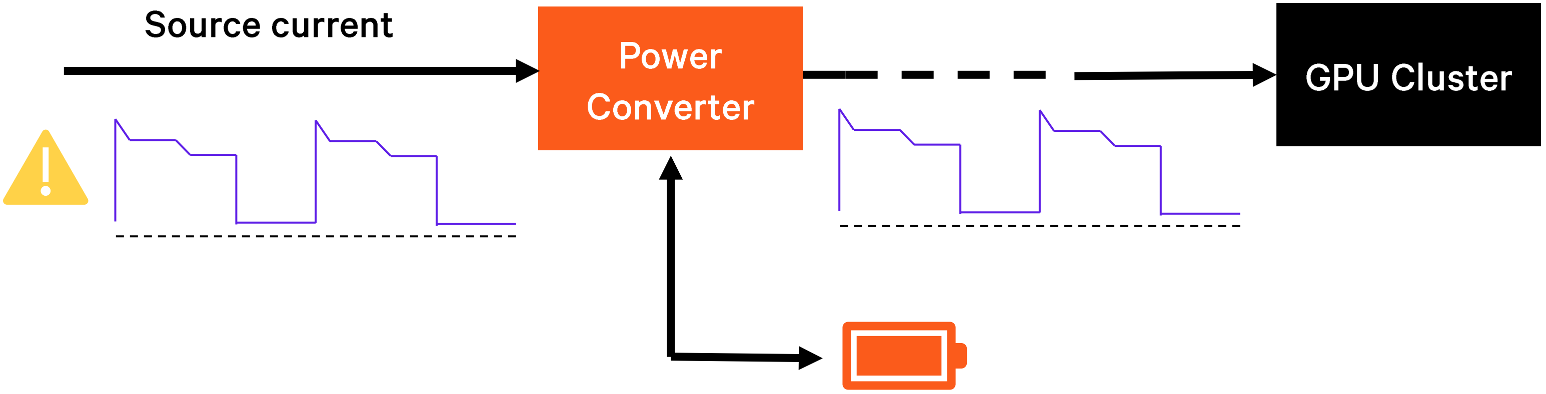
Minimize impact batteries

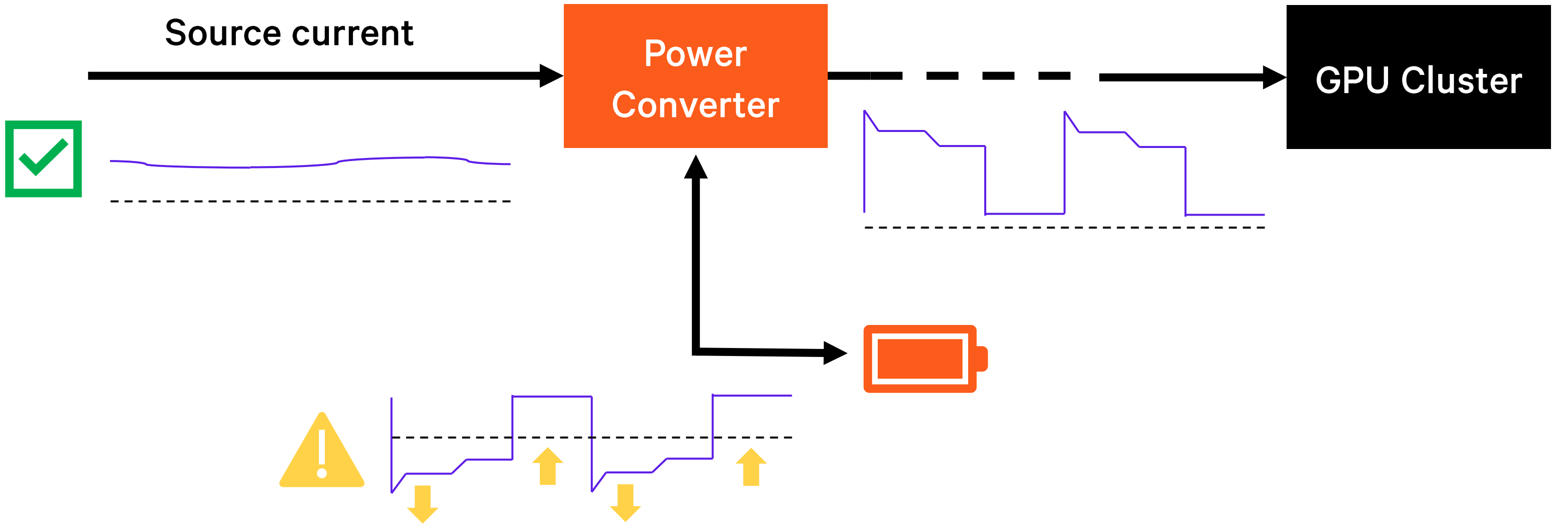
UPS FW is modified to handle load steps >100%

Minimize impact on grid/generators

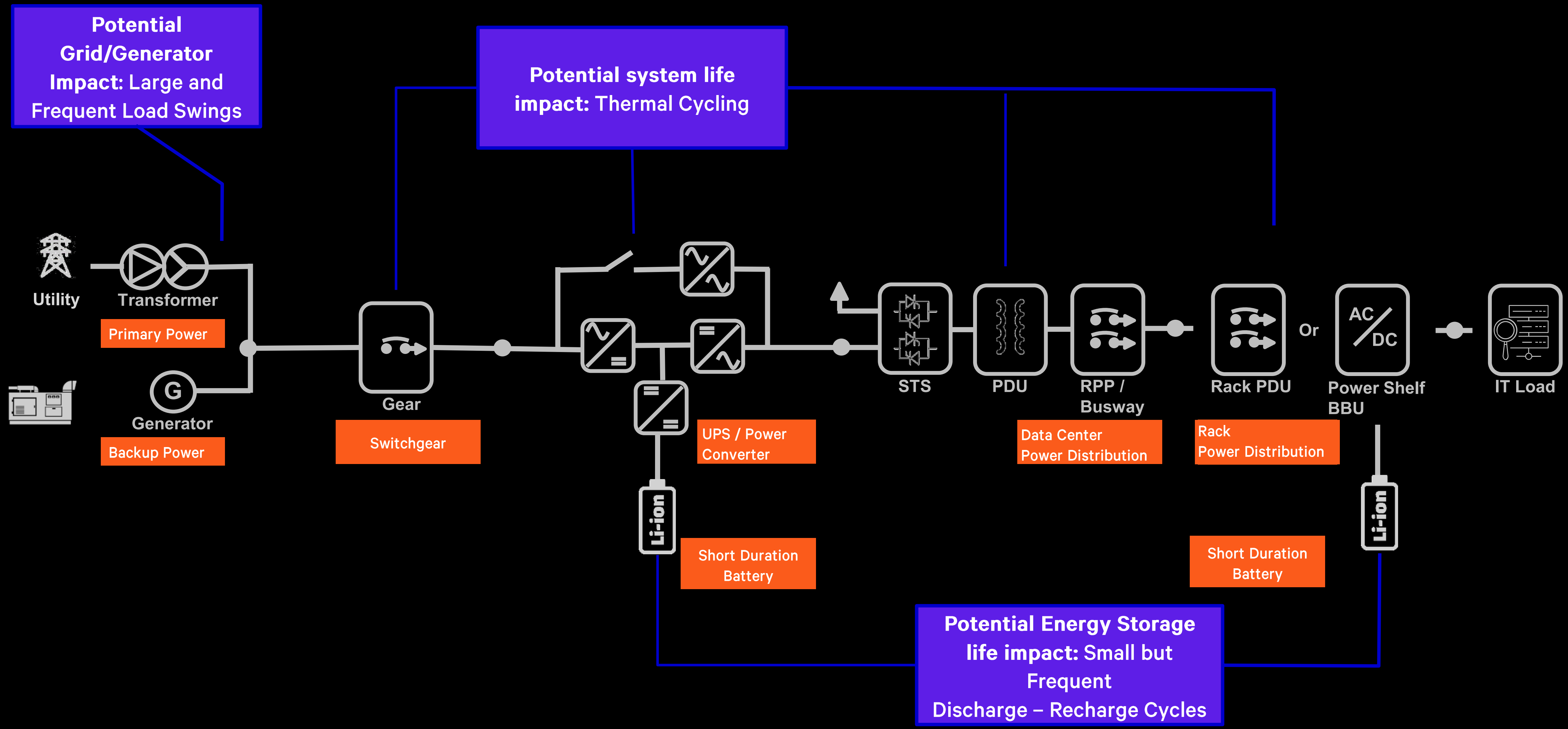
UPS FW is modified to use Batteries for Power Smoothing (Input Load Averaging)

At UPS level, two approaches can be taken to eliminate/minimize the AI load impact on the power train or batteries





AI Loads Effects on Data Center Powertrain Components



AI Loads Impact on the Data Center Powertrain

Artificial intelligence (AI) has a unique load profile which creates many challenges due to the rapidly fluctuating power demand, power density, and cooling requirements.

- **Overloads** as high as 150% causing voltage sags sending equipment to back up power.
- **UPS/BBU battery discharge & recharge cycles** impacting battery life.
- **Thermal cycling** of powertrain components impacting equipment life.
- **Generator voltage & frequency oscillations** causing power quality and stability issues within the data center.
- **Power fluctuations** being reflected to the grid causing power quality and grid stability risks.
- Higher rack power density for AI loads requiring **higher voltage.**
- Liquid cooling for GPUs have very **low thermal inertia** (1-10 secs), such that the CDU may need to run continuously.
- **Monitoring** and measurement equipment are "**too slow**" to **characterize** the **Fast dynamic AI load**.

Powertrain Architecture & Redundancy Strategies

Availability Design (COLO & Enterprise Interest)

- ✓ Redundancy included
- ✓ Concurrent maintainability

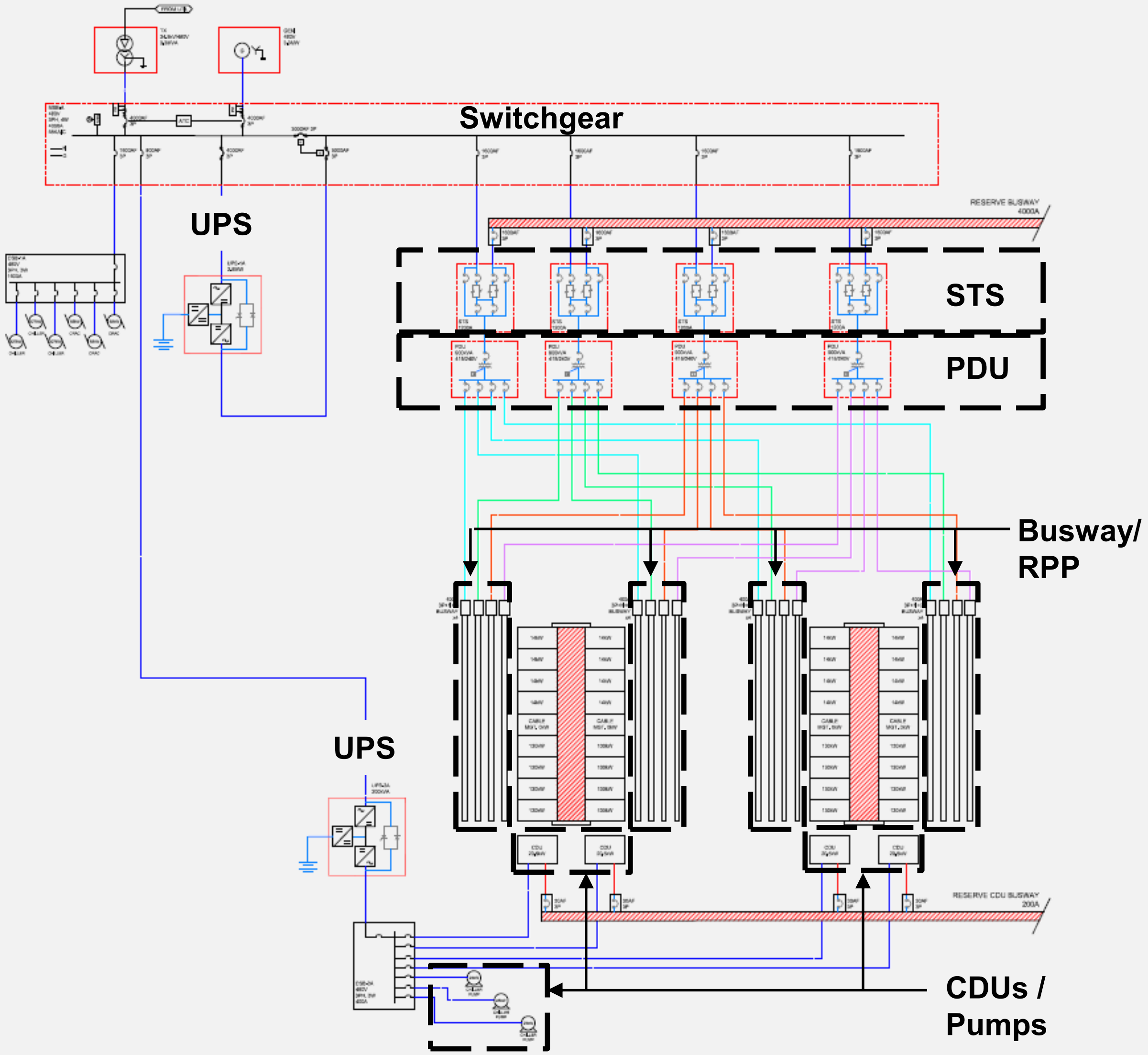
Capacity Design (Hyperscale Interest)

- ✓ Infrastructure optimized for IT compute load
- ✗ No redundancy
- ✗ No concurrent maintainability

AI Power Architecture – Availability Design

- No single point of failure in source, distribution or rack power path
- Can Utilize Block Reserve or Distributed Reserve Redundancy
- Infrastructure includes redundancy to achieve IT availability
- Central UPS protects IT & mitigates AI compute load behavior
- CDU/Chilled Water pumps protected by dedicated UPS; redundant power from block reserve
- Size power components to minimize stranded capacity

Availability Architecture - Block Reserve Example

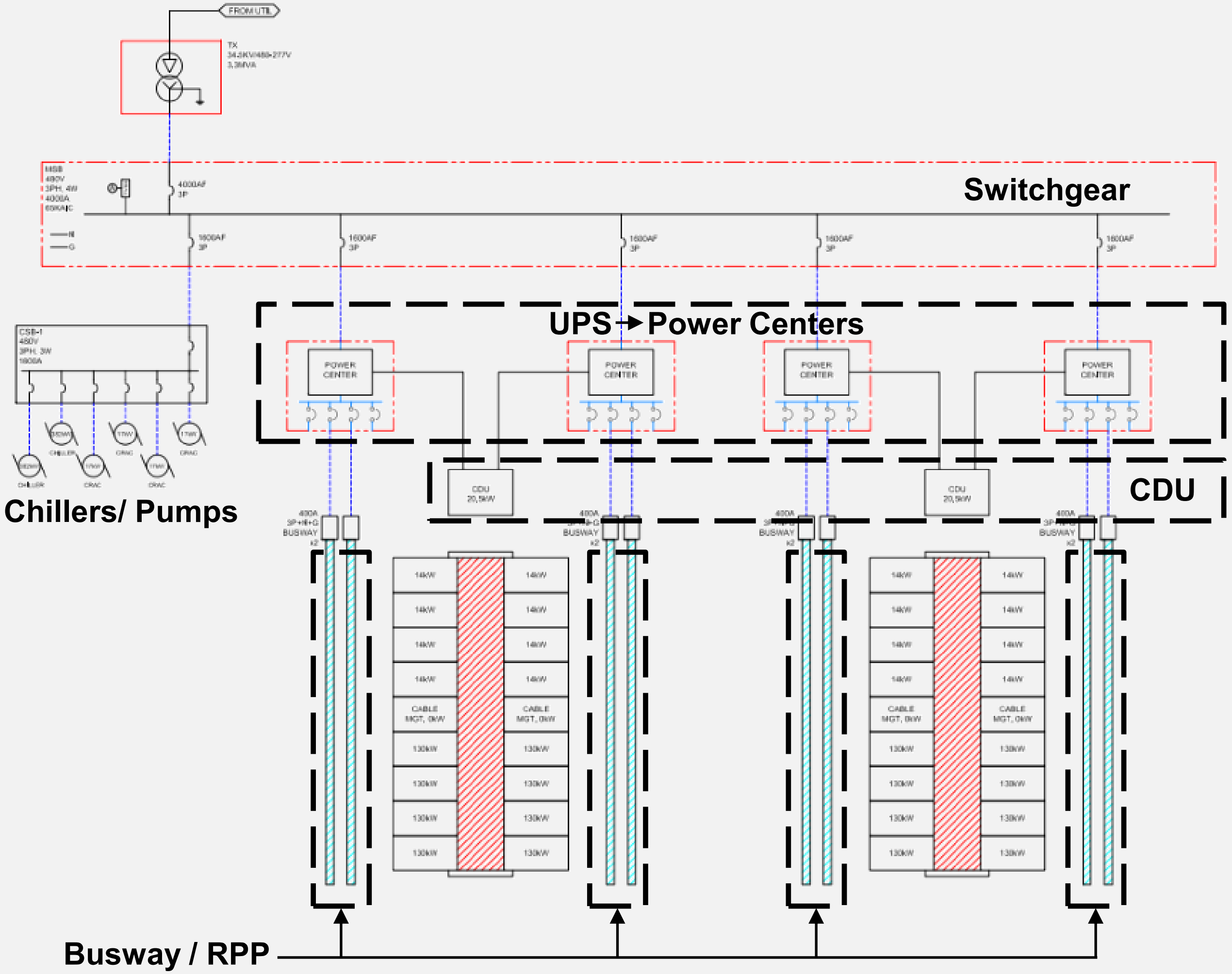


Nvidia DGX SuperPOD with GB200 - Blackwell NVL72 (130kW)

AI Power Architecture – Capacity Design

- All infrastructure supports AI IT; no reserve bus
- Failure at source, distribution or power shelf results in IT load loss
  - AI workloads are dynamically rerouted when issues occur
- Power bus and power shelfs sized only for needed capacity; no redundancy included
- No central UPS; CDU/CW pumps require separate protection
- AI Compute load behavior mitigation at the rack or row level using UPS->Power Center
- Power components sized to eliminate stranded capacity

Capacity Architecture Example



Nvidia DGX SuperPOD with GB200 - Blackwell NVL72 (130kW)

Close coupled system increase reliability, saves space, reduces installation and commissioning time

Site Architecture - PowerNexus is is a close coupled solution that combines the robust power of Vertiv™ Trinergy™ and Switchboard, reducing equipment footprint, cabling materials, and installation labor costs

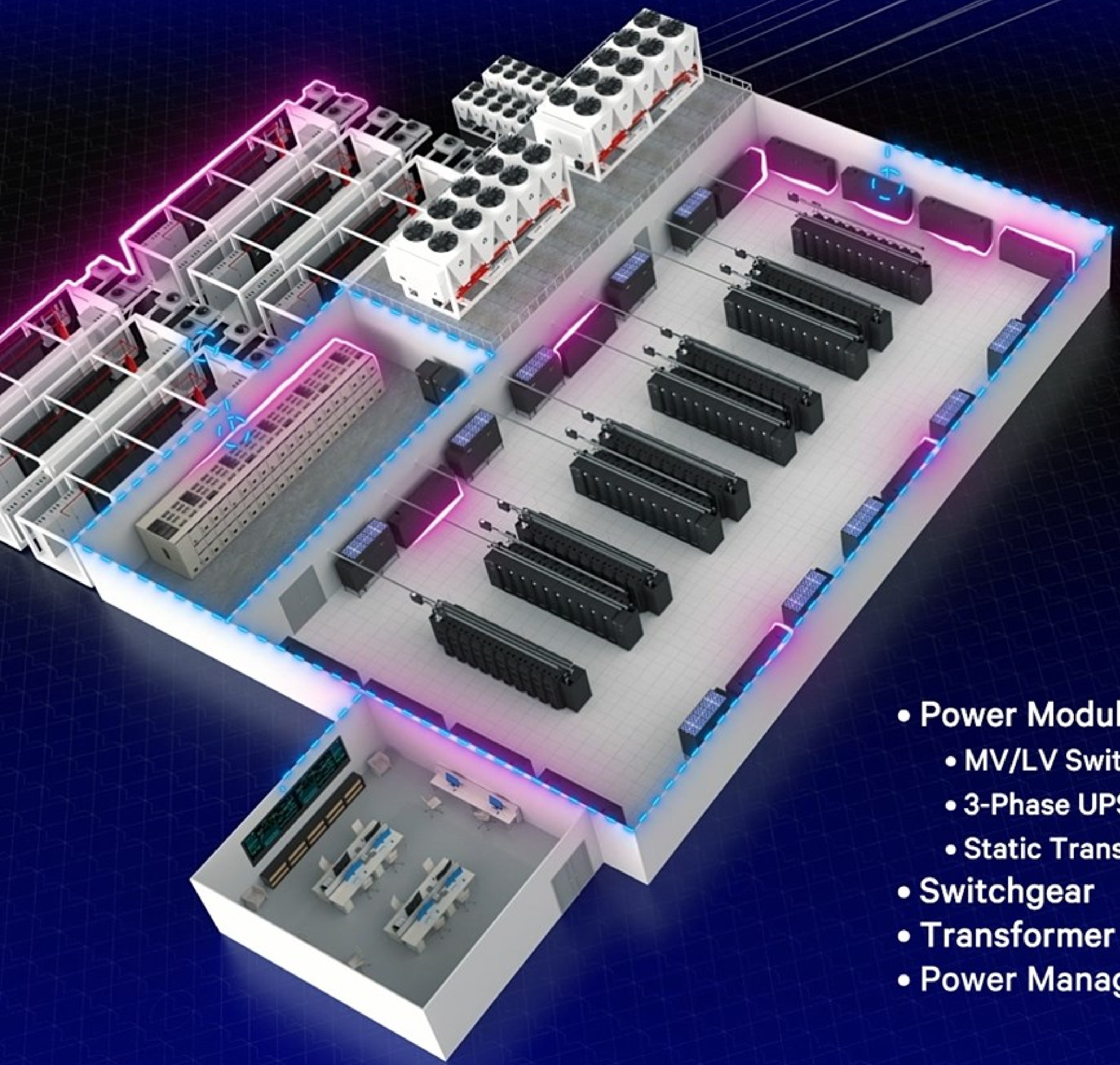
Speed up installation – Installation time in factory and on site is significantly reduced comparing with traditional build

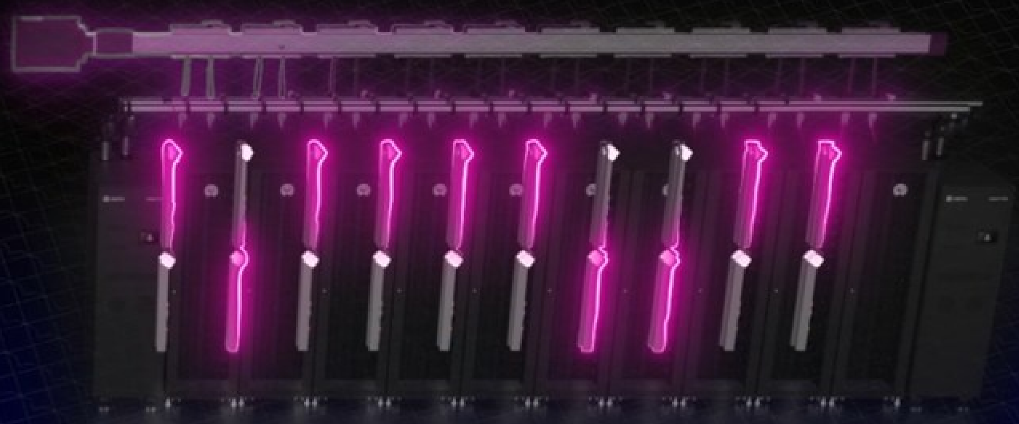
Increased reliability – Minimum quantity of interconnections reduces risk of faults and simplifies installation and maintenance

Install on site, install in factory – PowerNexus design enables different build strategies and elimininates compromises

High Density Compute

IT Power Capacity





- Remote Power Panels
- High-Capacity Busbar
- High-Capacity Rack PDU

- Power Modules
  - MV/LV Switchgear
  - 3-Phase UPS / Energy Storage
  - Static Transfer Switches
- Switchgear
- Transformer / Distribution Units
- Power Management and Controls

Power Train

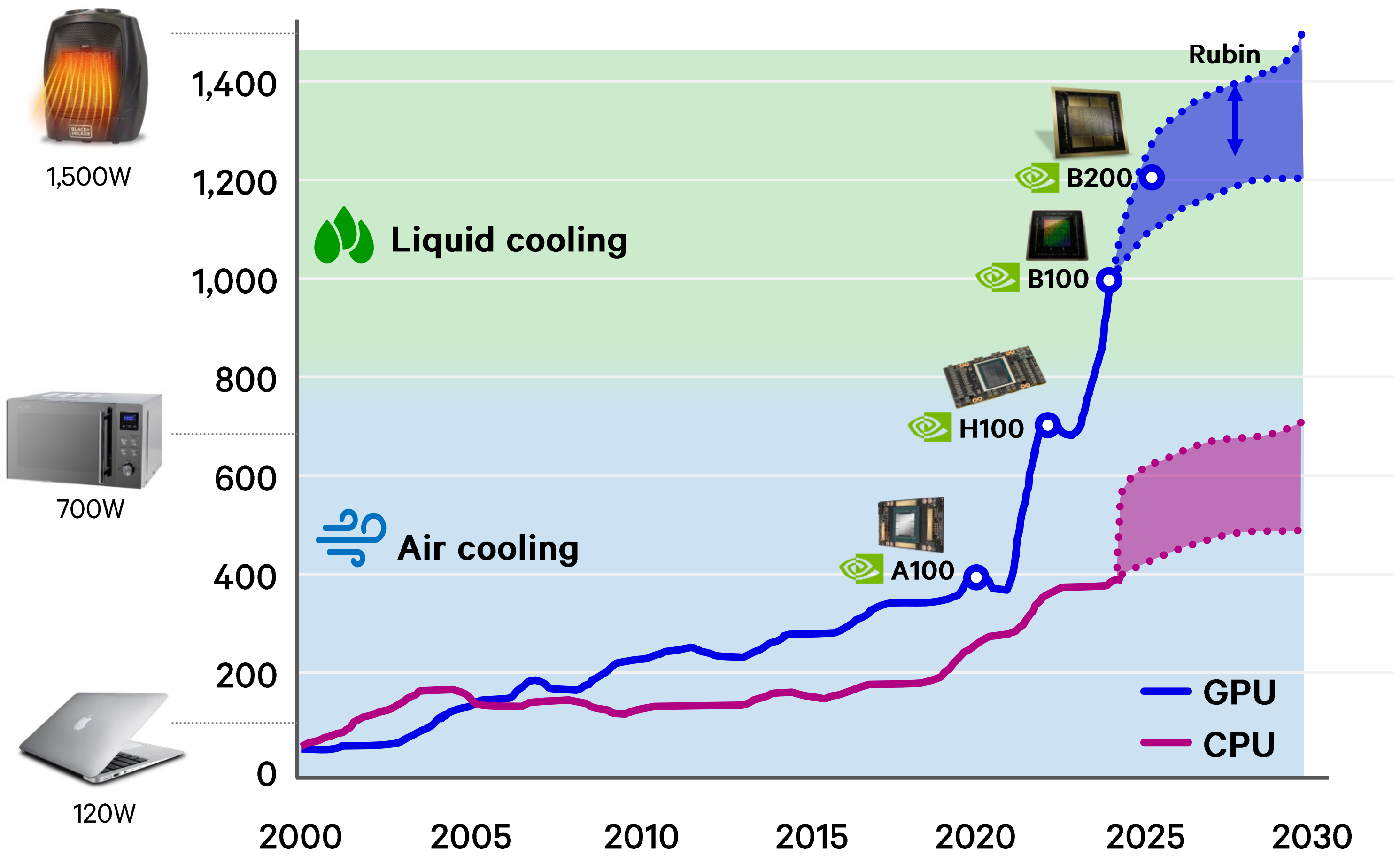
Power Management and Controls

Higher thermal design power (TDP) of AI chips is a key factor driving adoption of liquid cooling technology.

CPU and GPU power consumption forecast

Thermal Density Power - TDP (watts)

OMDIA



Above 700-800W TDP per chip, liquid cooling quickly becomes a necessity.

Source: Omdia, NVIDIA, Data Center Dynamics

With liquid cooling, thermal loads cannot go down. What options are there to keep these loads up?

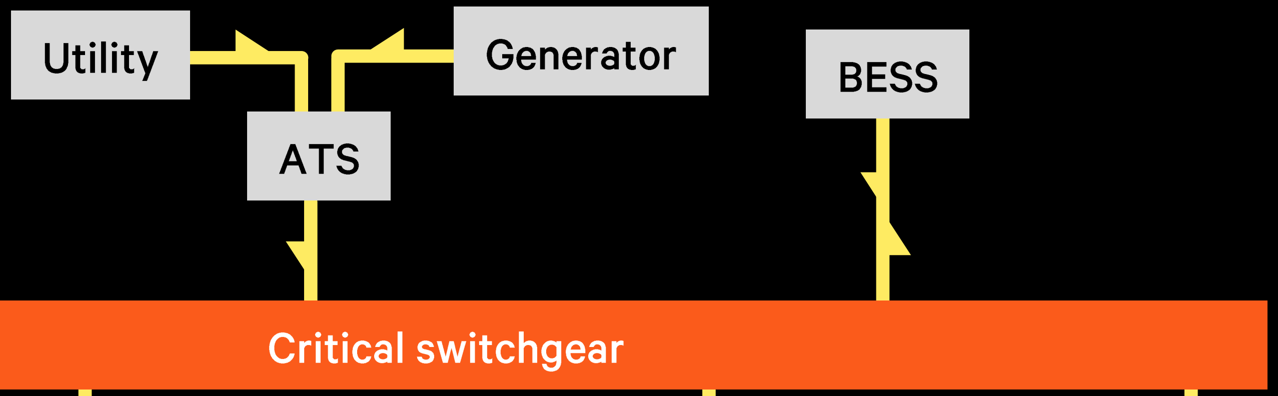
Air cooling

Thermal and housekeeping loads have historically been connected to the mechanical bus.

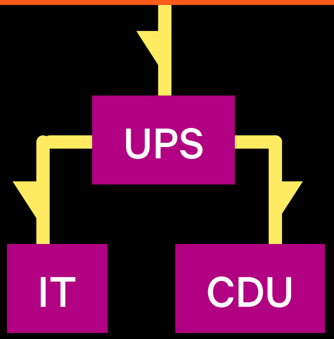
- When the utility went down, the UPS would maintain IT loads.
- Thermal and housekeeping load power would cycle (go off briefly before turning back on) on utility failure.

Liquid cooling

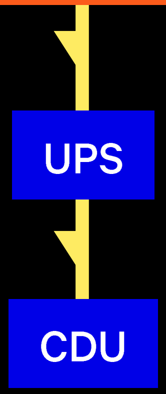
Coolant distribution units (CDUs) must be **powered continuously** just like IT loads.



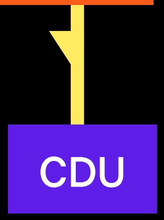
1. Plug CDU in share same UPS system used to power IT load?



2. Segregate the CDU into a dedicated mechanical load UPS?



3. Avoid the need of a UPS making the gen bus “always on” using a BESS?



Effects of climate change: as data and digital infrastructure scales, it must do so responsibly for the planet and its people.

Ever growing demand for compute and data: AI leading to an increase consumption of energy.

Increase need for dynamic power solutions:

- Fuel cells and local generation
- Battery storage solutions
- Grid support services

Our industry sits at the center of two major forces that must be at the core of our design considerations.

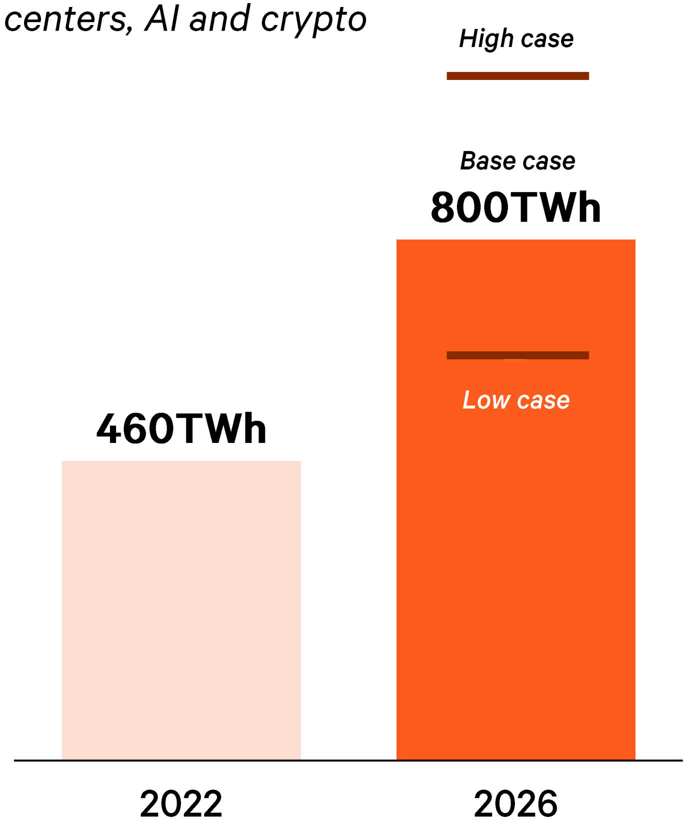
AI is expected to **accelerate energy consumption** by data centers adding pressure to already constrained power grids.

Source: IEA, Financial Times, KPMG, Data Center Frontier

Data center **power demand** is expected to grow at a **rapid pace** with AI.

iea International Energy Agency

Global electricity demand from data centers, AI and crypto



Authorities imposing **limitation on data center expansion**.

Connection **restrictions** in place since 2021.

Top colocators had permits rejected by Irish authorities 2023.

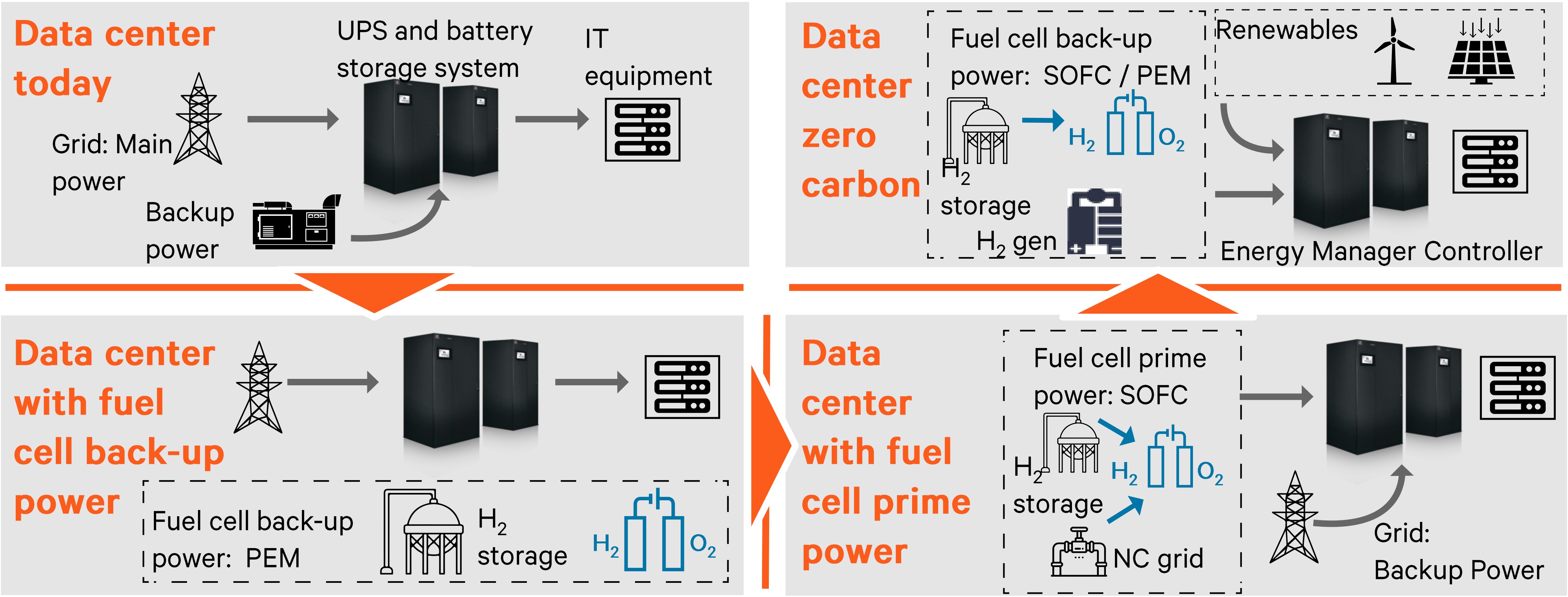
New data centers **banned** for three years.

Aimed to minimize impact on Singaporean grid and carbon footprint commitments, measure was lifted with introduction of stricter rules.

Restrictions on new data centers in Virginia.

New rules impacting construction along route 7 in Loudoun county, where power infrastructure is limited.

Micro-grids with local sustainable energy production are an option for decarbonization of data centers.

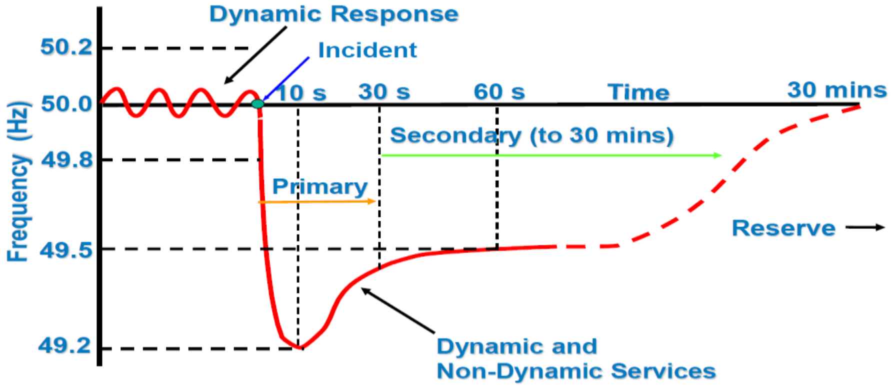


CONAPTO partners with Vertiv to provide services to the grid operator and achieve sustainability goals.

COLOCATION MADE IN SWEDEN

Frequency management

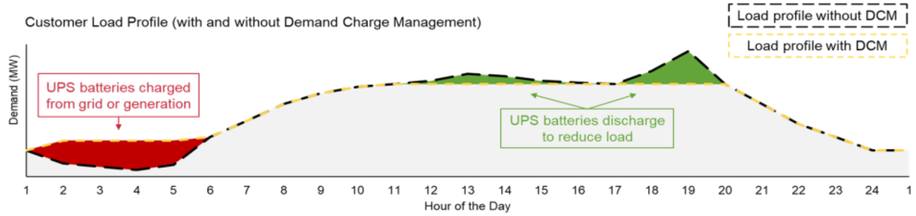
A fast-acting balancing system is needed to provide a quick response to sudden frequency variations and increase or reduce the electricity demand within a few seconds (fast frequency response and primary reserves) or minutes (secondary reserve).

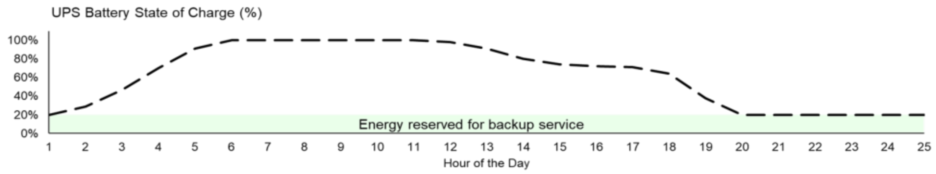


The faster the response, the higher the revenue opportunity.

Demand management (peak shaving)

In times of low demand or high supply, energy is stored, from which it is released at times of high demand or low supply.





Alternatively, consumers can price and carbon arbitrage adjusting consumption according to the changes in market price or renewable participation.

Battery storage and dynamic power are increasingly being monetized by data centers.

VERTIV™

NVIDIA | Partner