“Unit of Compute”

Sandia Power Electronics and Energy Conversion Workshop

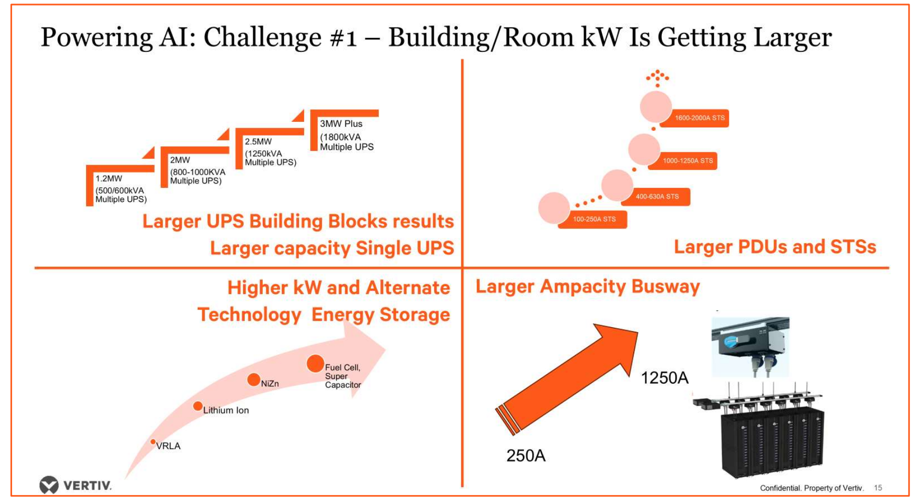
Peter A. Panfil

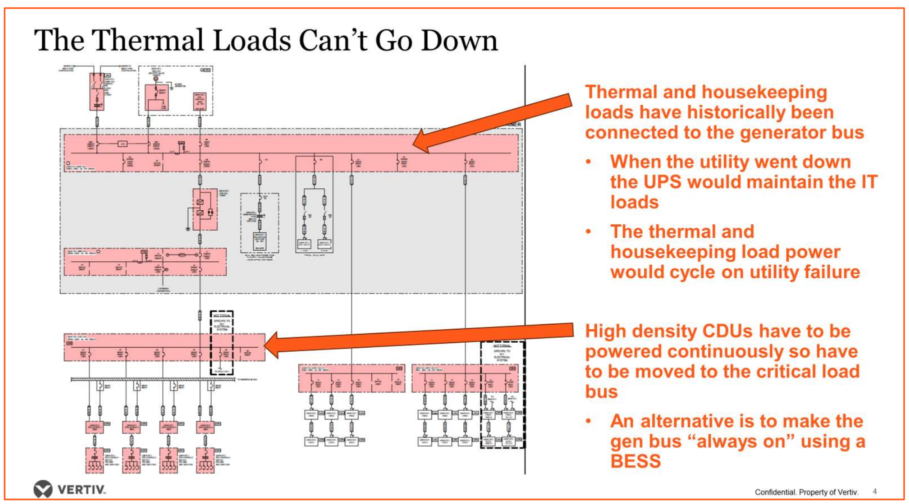
VP Global Power – Vertiv

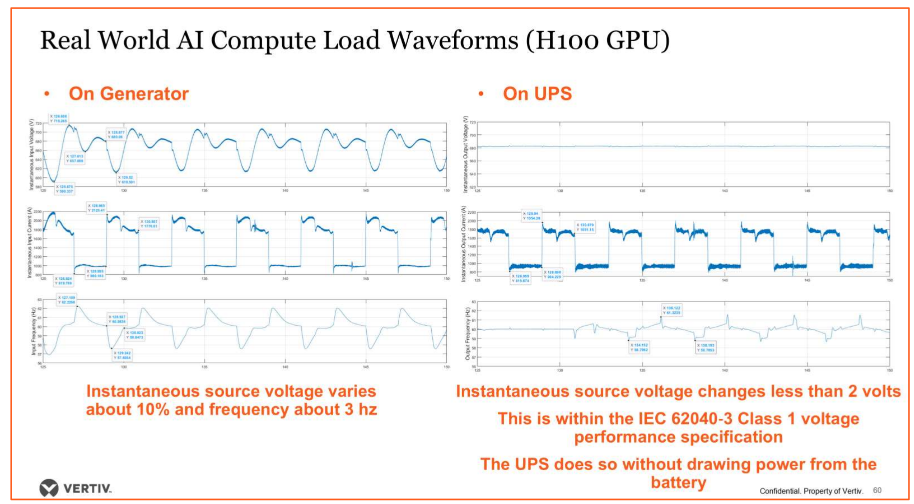
July 2025

AI Impact On Data Center Power

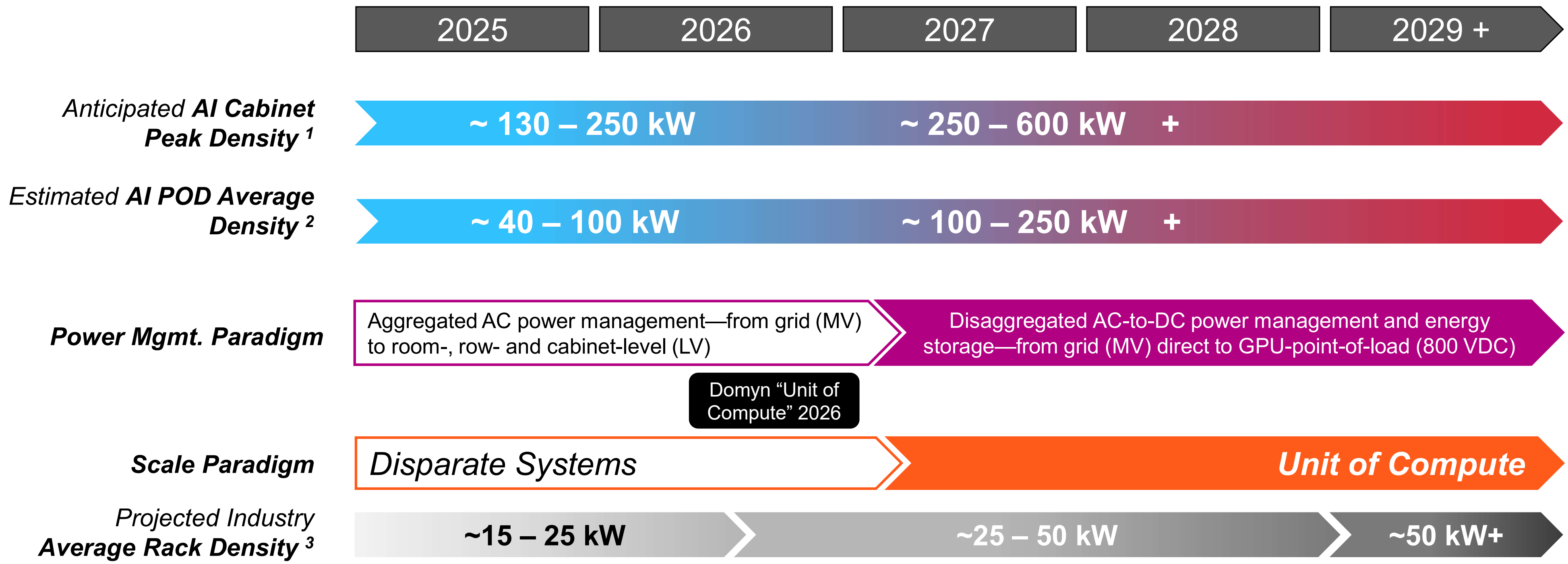
- **Size: everything gets bigger - rack loads,** Busway, PDUs, **UPSs,** Gear, **CDUs**
- **Scale: facilities are getting bigger – 3 MW blocks are projected to go to 20MW**
- **Always o**n power: liquid cooled GPUs must be continuously cooled
- **The AI compute load profile is demanding**







AI Factory Densification Roadmap



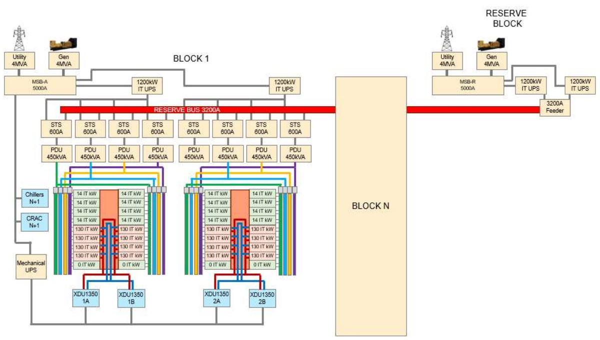
Note: Projections. Rack density varies by application. [1] Management Estimates, assuming a multi-node cabinet is a system domain of multiple interconnected GPUs and CPUs in a cabinet-scale footprint ; [2] Management Estimates, assuming an AI POD consisting of 18 cabinets (9 per row) including 8 GPU multi-nodes at higher peak density and 10 networking racks; [3] Management Estimates; average densities of data center cabinet installations across Cloud, Colocation and Enterprise/Distributed IT.

Climbing compute roadmaps are reshaping data center infrastructure requirements across all workload types requiring disparate power, cooling and compute systems designed as ONE

Market Adoption of GPUs Across Enterprise, MTDC and Hyperscale

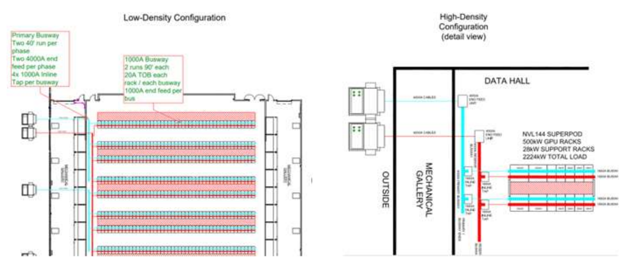
Conventional Compute Data Center Deploying GPUs

- Distribution from 4 separate distribution paths
  - 4 x Block Reserve power buses are woven across 4 Static Transfer Switches per UPS in a 4:3 BR weave
- No IT bus SPOFs



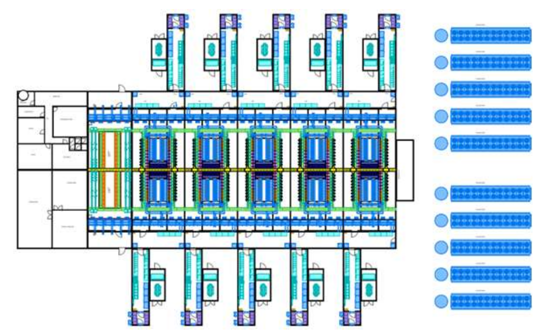
Next Gen Dater Center Deploying Either CPUs or GPUs

- Capacity CPU racks direct distribution from 4kA busway
  - 4x1000A IT distribution buses
- No central power protection is standard
  - If power protection is needed UPSs are mounted at the end of the rack row

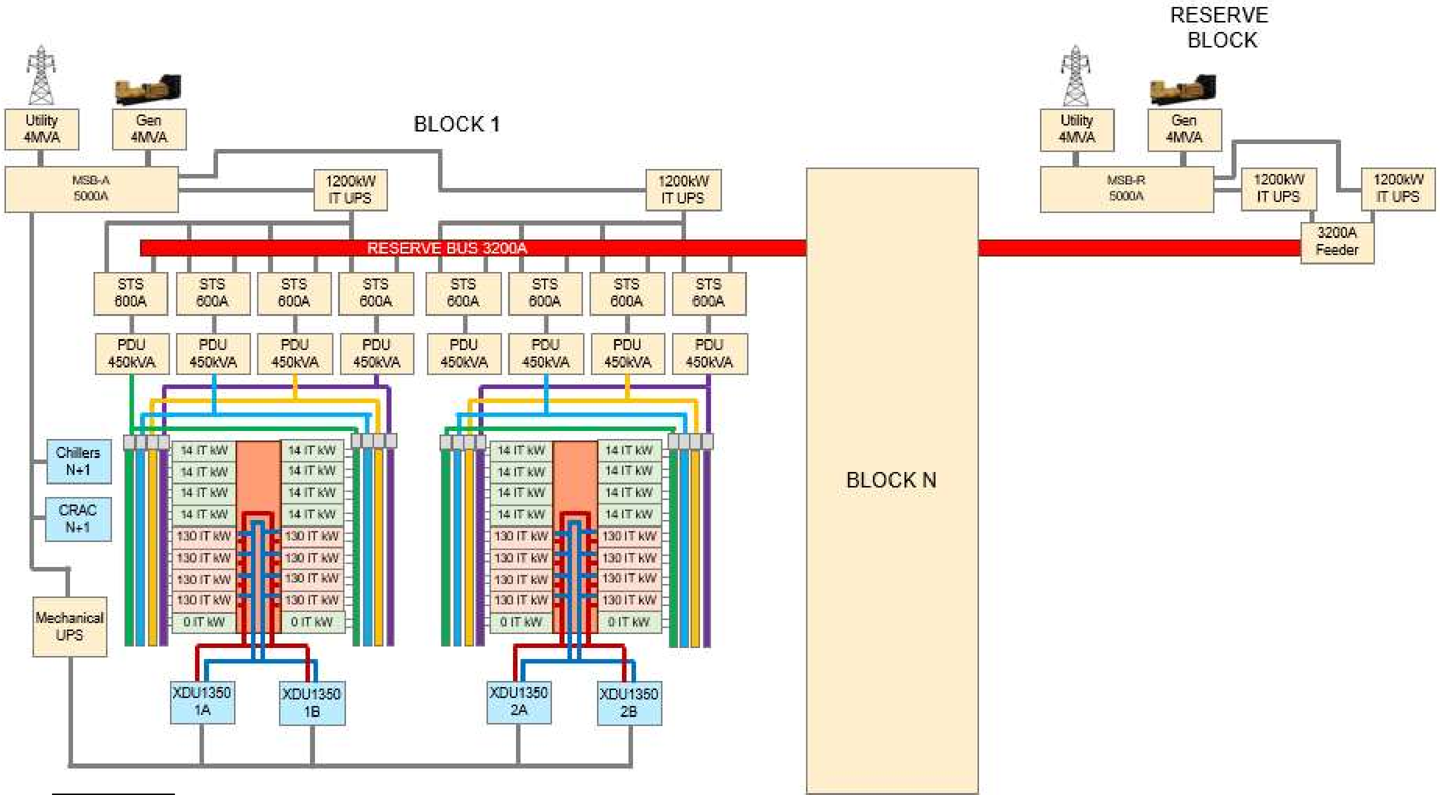


AI Factory GPU Dedicated Data Center

- Dedicated AI factory with modular PODs, electrical rooms and MV TX
  - 5xDR power buses with 4 applied per POD: A-B-C-D, B-C-D-E, etc.
- No IT bus SPOFs



Conventional Compute Data Center Deploying GPUs



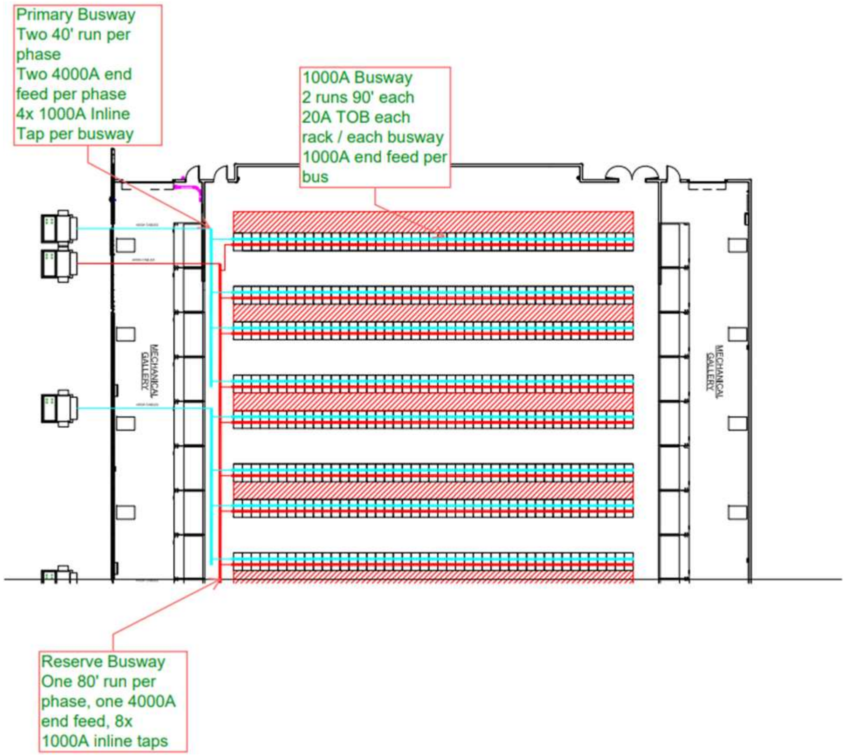
Partner

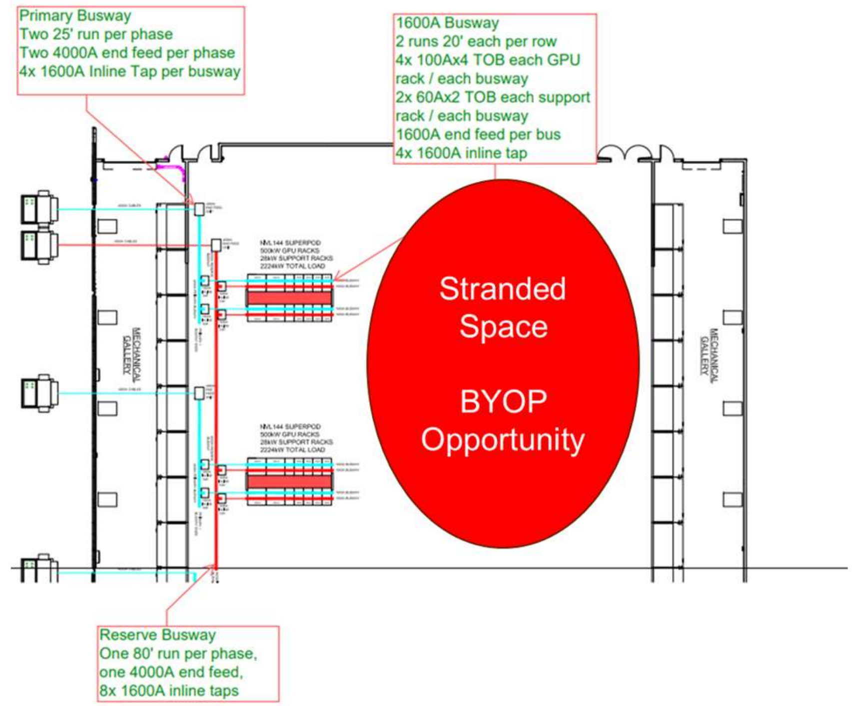
Modern enterprise and MTDC high availability architectures closely mirror the high availability refence one lines developed in conjunction with a major GPU mfr:

- **4 power buses are woven across 4 Static Transfer Switches per UPS in a 4:3 Block Redundant weave or 4 distribution buses in a Distributed Redundant weave**
- **No IT bus Single Points of Failure**
- **Vertiv developed a full data center behavioral model that can verify the site can support the AI dynamic load power**

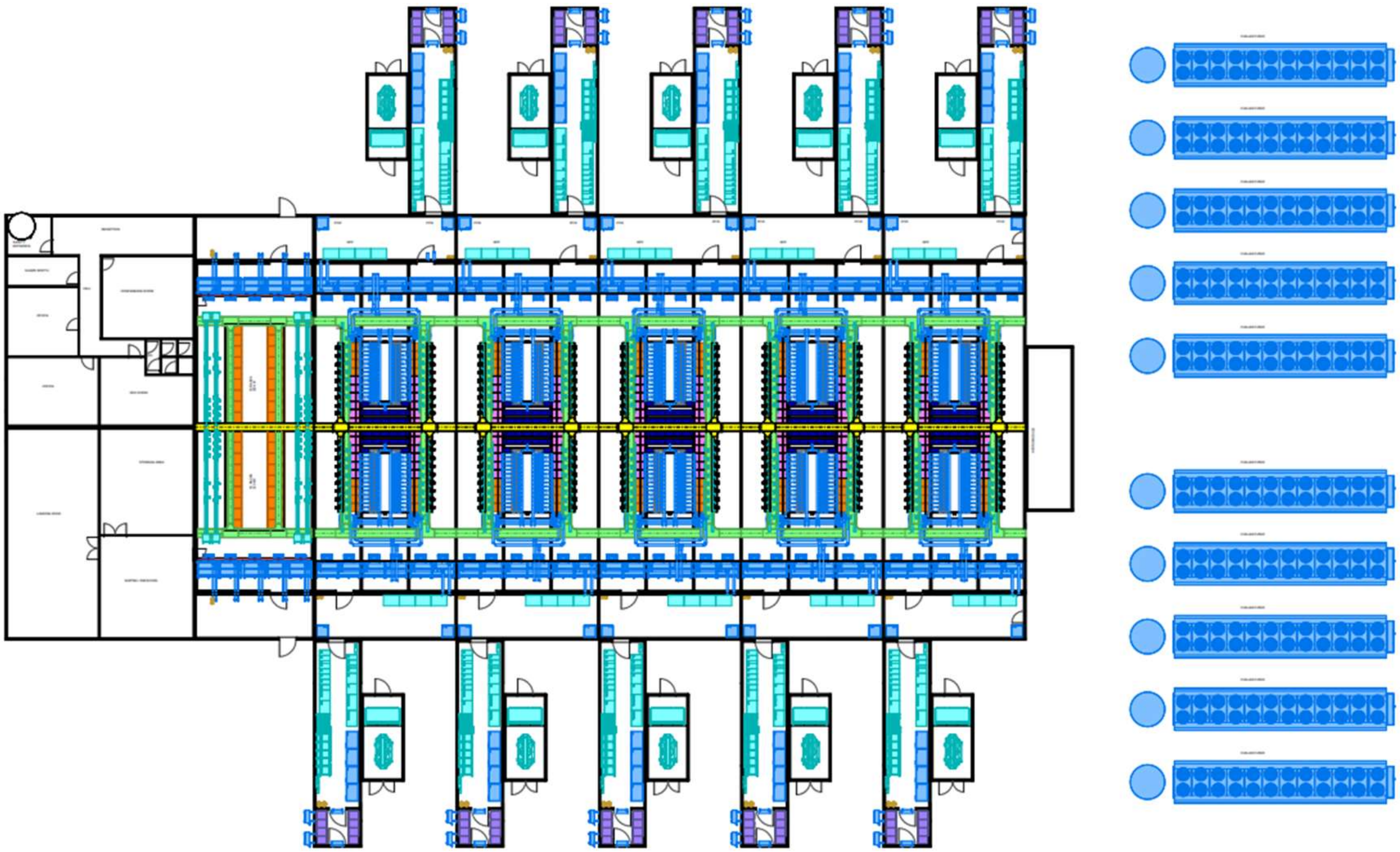
This “Step and Repeat” process can be applied at virtually all modern Enterprise, Colo and Hyperscale Data Centers

Next Gen Dater Center Deploying Either CPUs or GPUs





AI Factory GPU Dedicated Data Center



Safe Operation

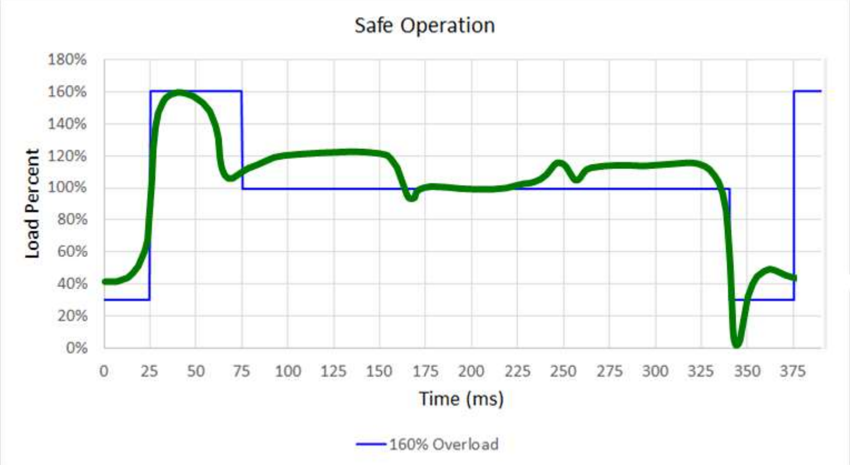
Achievable EDP Load Profile

Peak Overload (kW%)	Peak Overload Duration (ms)	Average Power Duration (ms)	Workload Idle Duration (ms)	Average Power (kW%)	Frequency (Workloads / second)	Workload Duty Cycle (%)
160	50	265	35	99	2	90
150	50	220	30	99	3	90
135	50	152	23	99	4	90
125	50	107	18	99	5	90
115	50	62	13	99	8	90

Peak Overload + Duration determines the maximum number of AI workloads per second.

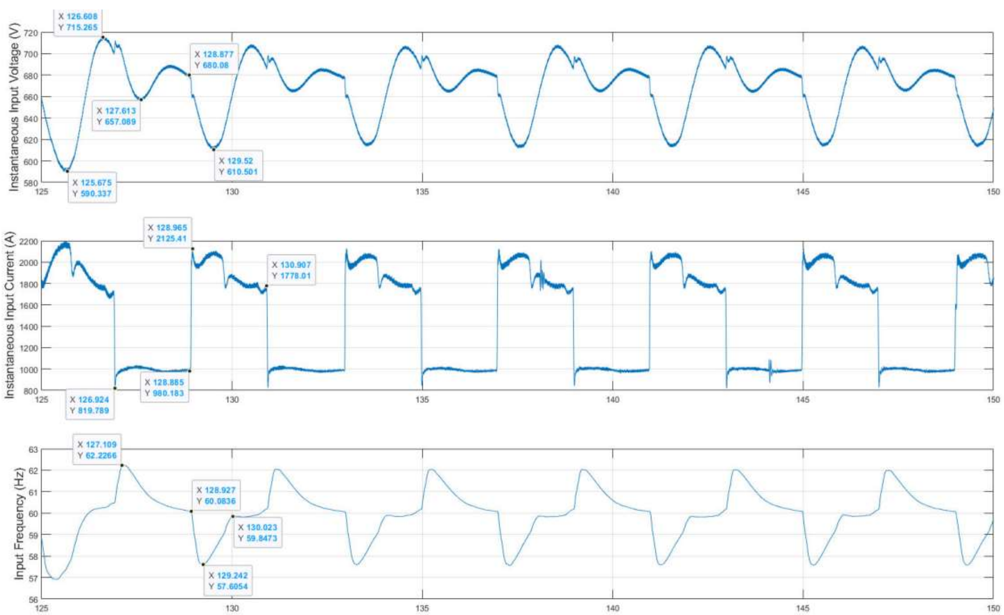
Average Power (P_{avg}) needs to be ***100% or less***.

Duty Cycle does not matter. The idle time could be zero (i.e., 100% duty cycle).



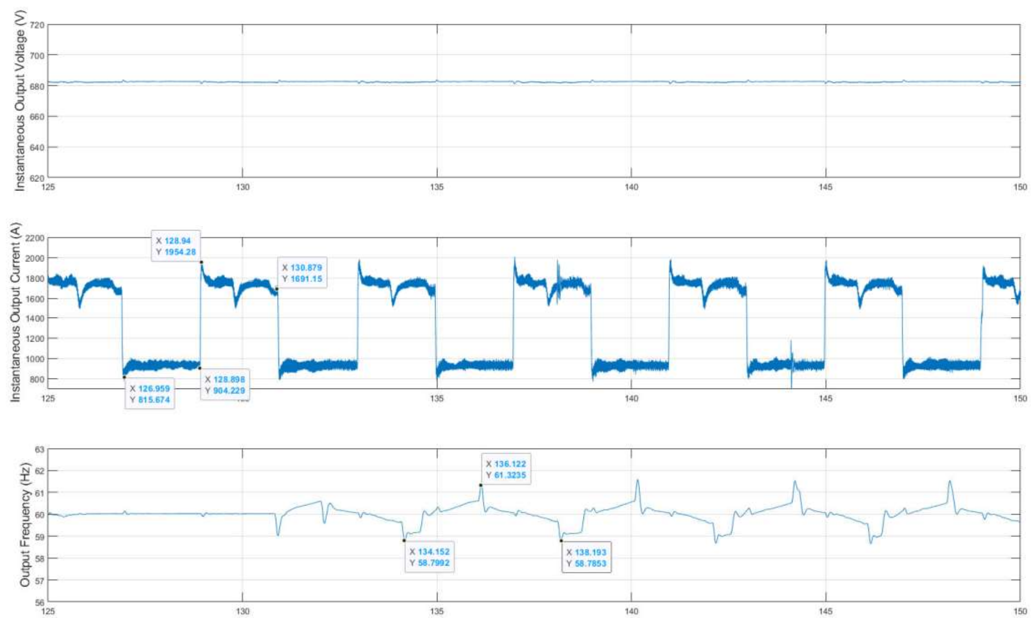
Real World AI Compute Load Waveforms (NVIDIA H100 GPU)

- **On Generator**



Instantaneous source voltage varies about 10% and frequency about 3 hz

- **On UPS**

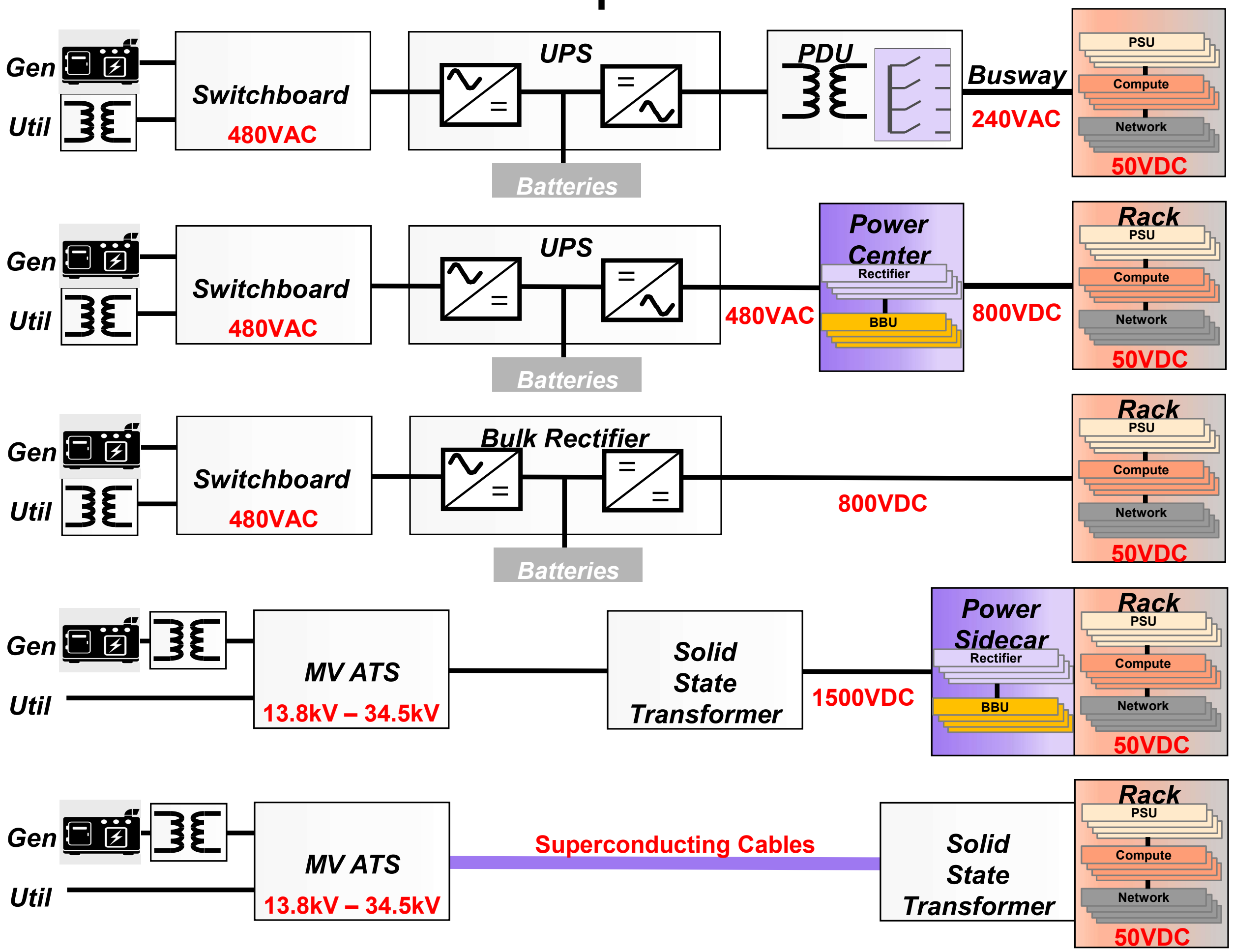


Instantaneous source voltage changes less than 2 volts

This is within the IEC 62040-3 Class 1 voltage performance specification

The UPS does so without drawing power from the battery

Source to Load Options



Architecture 1: Conventional Data Center

- Reserve or 2N Architecture
- Centralized UPS at 415VAC or 480VAC
- Direct or PDU to Busway Distribution Path at 415/240VAC
- PSU or Power Shelves in the rack at 50VDC

Architecture 2: HVDC at the Rack

- Reserve or 2N Architecture
- Centralized UPS at 415VAC or 480VAC
- Power Center converts 480VAC or 415VAC to 800VDC at the row
- PSU or Power Shelves in the rack at 50VDC

Architecture 3: HVDC Bulk Rectifier

- Reserve or 2N Architecture
- Bulk Rectifier converts 480VAC or 415VAC to 800VDC
- PSU or Power Shelves convert 800VDC to 50VDC

Architecture 4: Medium Voltage to Row

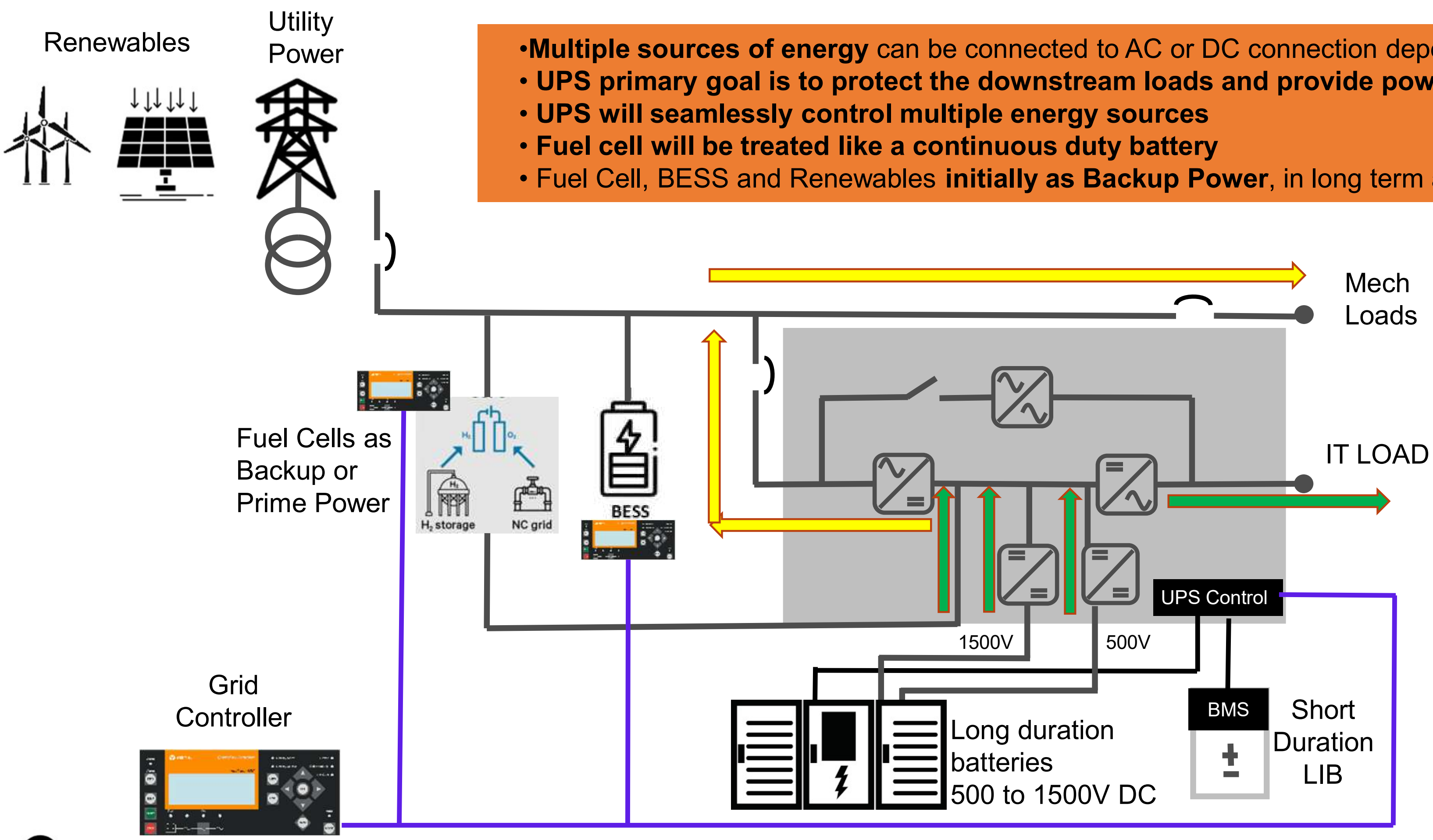
- Reserve or 2N Architecture
- Medium Voltage to Solid State Transformer (SST)
- SST Converts MV to 1500VDC
- Power Sidecar converts 1500VDC to 50VDC

Architecture 5: Superconductor to SST

- Reserve or 2N Architecture
- MV Superconductor Solid State Transformer (SST)
- SST Converts MV to LV
- PSU or Power Shelves in the rack at 50VDC

Evolution of the UPS Using Dynamic Energy Resources

- **Multiple sources of energy** can be connected to AC or DC connection depending on site needs.
- **UPS primary goal is to protect the downstream loads and provide power quality**
- **UPS will seamlessly control multiple energy sources**
- **Fuel cell will be treated like a continuous duty battery**
- Fuel Cell, BESS and Renewables **initially as Backup Power**, in long term also as **Prime power**



Fuel Cell and BESS based architectures are being developed and quoted today.

Grid Support Protocol via Modbus(tm) and CANbus, DNP, IEC61850

Questions?

Peter A. Panfil

VP Global Power

Peter.Panfil@vertiv.com

+1-614-314-4154 (M)

VERTIV™