PART 0 · JULY 2026 EDITION

# Foundations & How to Use This Guide

*The most expensive failures of this build-out are not engineering failures — they are vocabulary failures. Part 0 fixes the objective function, the reading protocol, and the words, before a single drawing is made.*

THE STAKES

# ~$6.7T

## Cumulative global data-center capex by 2030 — most of it AI-capable capacity

*At this scale, a coordination mistake is a stranded asset with nine zeros. The expensive failures are rarely bad engineering — they are subsystems that each passed their own test.*

MCKINSEY, 'THE COST OF COMPUTE' · AS-OF 2025

THE REFRAME

# An AI data center is not a building that happens to contain computers. It is one co-designed machine — and it has exactly one honest objective function.

*TCO per unit of useful work. Not PUE, not rack fill, not switch utilization. Optimizing any subsystem to its own local metric routinely degrades the machine.*

ACT I

# The machine and the inflection

*Why 2024–2026 broke the old playbook: rack density jumped, named frontier profiles moved beyond qualified air envelopes, and the binding constraint moved from chips to megawatts and time-to-power.*

THE DENSITY JUMP

# ~10–15 kW → 120–600 kW

## Rack power across the inflection: legacy racks, to GB200 NVL72, to the Kyber-class roadmap

*The top of that range is a 2027 roadmap figure, not shipping hardware — yet the hall you pour today must survive it. A jump this size obsoletes a building designed to the current spec.*

NVIDIA · AS-OF 2026-08

THE DISCONTINUITY

# 30–40 kW typical RDHx; >50 kW with active fans; not a universal limit

## One conventional-air reference case — context, not a universal cooling selector

*Rack kW is one input. Heat split/flux, airflow/inlet and water conditions, room rejection, climate, service/redundancy, and refresh determine whether air, RDHx/AALC, hybrid, or DLC closes.*

SEMIANALYSIS, DATACENTER ANATOMY PART 2 – COOLING SYSTEMS, 13 FEBRUARY 2025 · AS-OF 2025-02-13

THE CONSTRAINT MOVED

# Power-bound, not chip-bound — read the design backward from the substation

**~128–208 wk**

HV power transformer lead time — often the single long pole in the schedule

POWER MAGAZINE · AS-OF 2026-08

**~2/3**

Inference share of AI compute in 2026 — demand has shifted from training to serving

DELOITTE TMT PREDICTIONS 2026 · AS-OF 2026

WHERE PROJECTS FAIL

# The coupling seams between the facility track and the IT track

| SEAM | FACILITY SIDE OWNS | IT SIDE OWNS | IF THE SEAM IS IGNORED |
|---|---|---|---|
| **Equipment + facility envelope → cooling modality** | Cooling plant, water loop, heat rejection | Rack power, accelerator coolant spec | Power interconnected but heat unremovable; stranded MW |
| **Fabric topology → floor plan** | Slab, manifold runs, rack adjacency | Scale-up domain, scale-out blocking | Cable reach blown; copper turns to optics |
| **Capacity ramp → substrate** | Floor loading, water and electrical headroom | Generation-by-generation GPU/MW curve | A hall that cannot absorb the next-gen rack |
| **Required maintenance/fault outcomes → candidate topology** | Transfer, post-event loading, path and control independence | Recovery target, workload consequence, contract | A topology that does not prove the required whole-service outcome |

THE FOUNDING DECISION

# Five procurement silos, or one machine?

## Procurement silos

— Each subsystem tenders to its own metric and acceptance test

— Teams meet for the first time at integration — the costliest place to disagree

— Every subsystem passes; the machine underperforms its TCO model anyway

— The PUE looks great — because the cooling ran lean and the GPUs throttled

## Co-designed machine

— One objective: TCO per unit of goodput over the asset's economic life

— Subsystem-local losses traded for system-level gains, on purpose

— Grid-to-chip and chip-to-atmosphere held as continuous chains, not handoffs

— The whole design read backward from the substation, not forward from the GPU

SO WHAT

The locally-optimized silo is this era's signature failure mode — a facility where every subsystem passed its own acceptance test, and the machine still misses its TCO model.

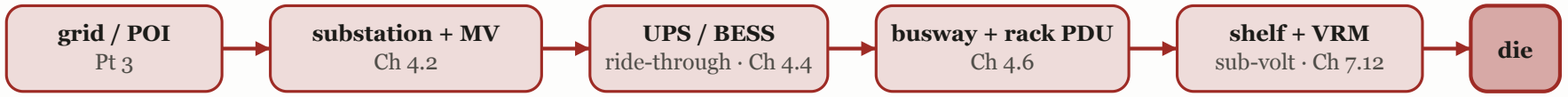
ONE MACHINE, TWO SPINES

# Electrons flow grid to chip; the same joules walk back, chip to atmosphere

## One machine, two spines — electrons in, heat out

The two physical chains every Part of this guide returns to. Every stage has an efficiency and a failure mode.

**GRID → CHIP · the power spine** Parts 3 · 4 · 7




cumulative loss = the gap between the MW you buy and the MW that reach silicon:
legacy AC chain **~82%** typical utility→VRM · 800 VDC / SST chain **~87%** (SemiAnalysis 2025) — the Ch 4.1 re-architecture case

**the objective function: megawatts → intelligence, at the highest sustained efficiency physics allows**
every stage above and below is one machine — co-designed, or mis-designed

**CHIP → ATMOSPHERE · the heat spine** Part 5


every joule that entered as electricity leaves as heat — the same number, walked the other way

**Why the spines can't be owned separately: the rack profile crosses both.**
HPE's cited GB200 profile sends ~115 kW to liquid and ~17 kW to air; that named heat split governs the facility design.
Other racks need their own airflow, liquid-capture, TCS/FWS, climate, rejection, redundancy, and service envelope.

Measured/sourced · Source: SemiAnalysis · as of August 2026

*Every watt bought is a watt to reject — power and heat are one chain read in both directions. Split the spines across separate owners and one rack decision rewrites the other team's building.*

WHAT CO-DESIGN BUYS

# ~87% vs ~82% AC

## End-to-end electrical-chain efficiency: a collapsed DC chain vs a typical AC path

*A system-level gain only co-design captures — a bolt-on retrofit cannot recover it. Every lost watt is paid for, every hour, for the life of the asset.*

SEMIANALYSIS, DATACENTER ANATOMY PT 1 · AS-OF 2025

ACT II

# How to use the instrument

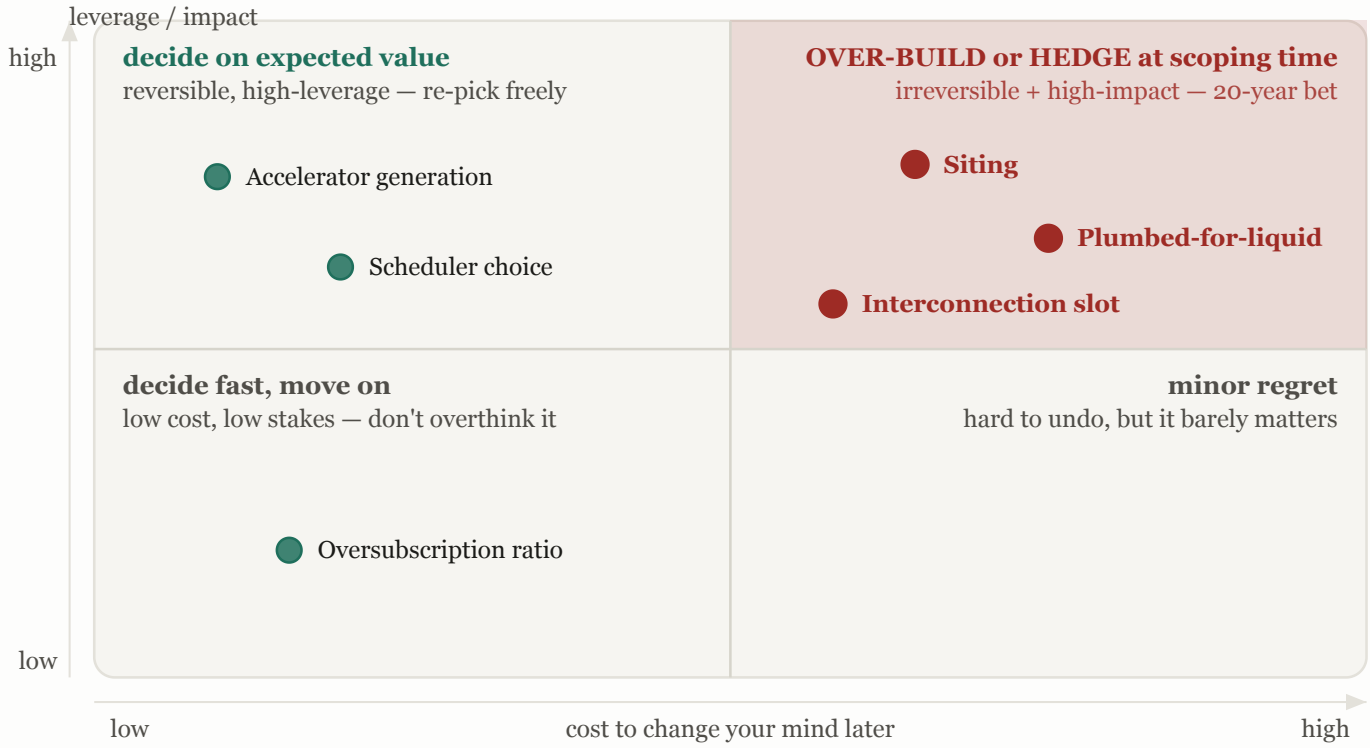
*This guide is a decision instrument, not a front-to-back read. Every chapter names one fork, the cost of each branch, and the date-stamped numbers you need to choose.*

THE MASTER LENS

# Sort every decision by the cost of changing your mind

## Sort decisions by how hard they are to undo

The same fork is trivial or fatal depending on where it sits — sort first, decide second.




**The fork is the same — the axis decides the stakes.**
A choice on the right you can't take back; sort there first, then spend your attention.

Guide scenario · Guide synthesis · as of August 2026

*Reversible forks: decide fast, on imperfect data. Irreversible, high-leverage forks — the slab, the voltage, the cooling modality — get over-built or hedged at scoping time.*

WHY THE CAVEATS MATTER

**2–3** yr bear case vs 4–6 yr estimate

## GPU economic life vs book life — a figure this guide flags as contested

*Contested means: run irreversible decisions across the whole range, never a point estimate. Every figure here carries an as-of date and a confidence caveat for exactly this reason.*

CNBC · AS-OF 2026

THE DATING RULE

# Never let a stale level drive an irreversible fork.

*Durable figures (physics, standards): design to the level. Semi-durable hardware: use the level, confirm it ships on your timeline. Volatile market numbers: direction only — re-pull before committing.*

ACT III

# A shared vocabulary

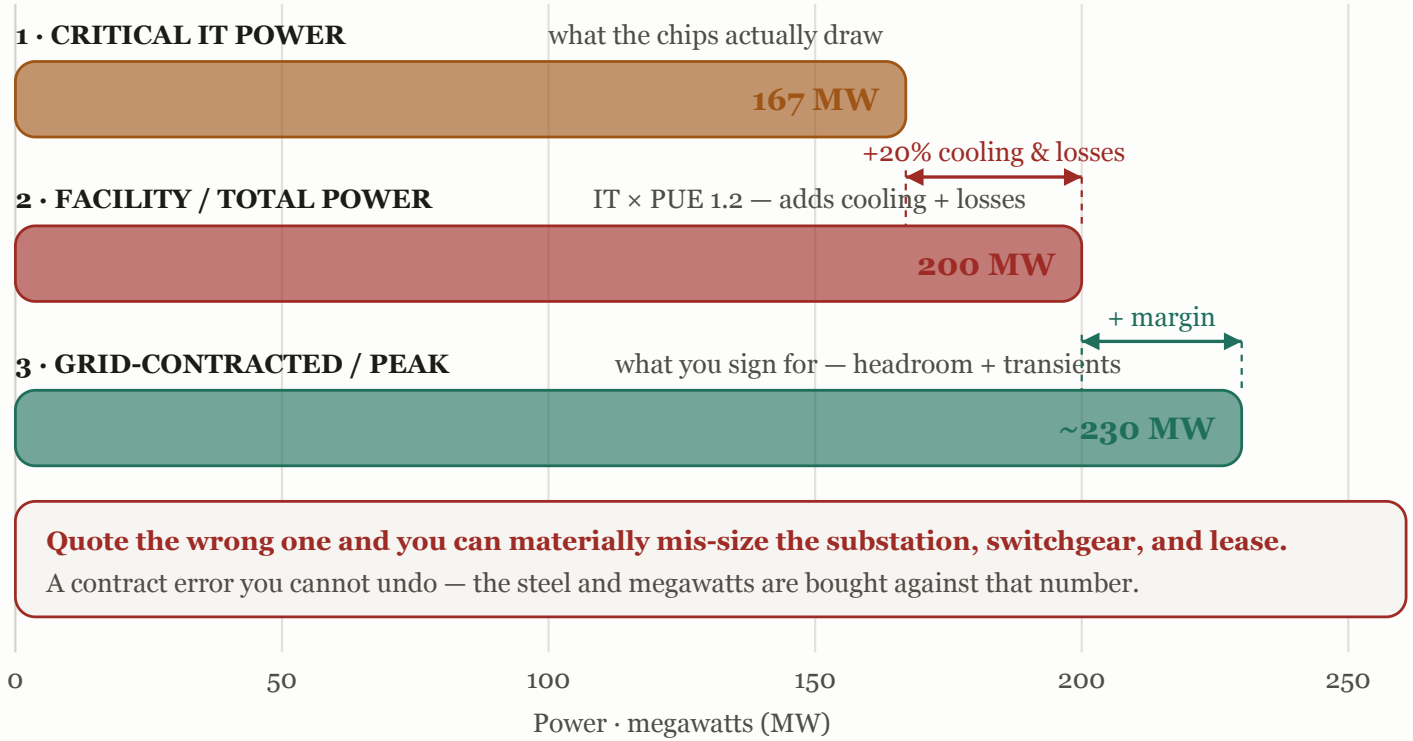
*Before any engineering: name the power number, the capacity unit, the metric you optimize, and what a redundancy promise actually promises. Misname any one and you misprice the build.*

THE COSTLIEST CONFUSION

# One facility, three different 'power' numbers

## Three numbers, all called "the power"

Illustrative scenario: keep IT, facility, and contracted-peak boundaries explicit.




Guide scenario · Source: guide power-boundary model · as of August 2026

*Grid draw, total facility power, and critical IT power can sit tens of percent apart. Before any term sheet, force every party to name which one they mean.*

THE METRIC STACK

# Three metric families that pull in different directions

## 1.54

Industry-average PUE, flat for a sixth year — overhead you buy but never sell

UPTIME INSTITUTE GLOBAL DATA CENTER SURVEY 2025 · AS-OF 2025

## 90% vs 96% scenario

Illustrative 90%-versus-96% goodput sensitivity — replace with fleet measurement

SEMIANALYSIS CLUSTERMAX 2.0 PROVIDER TIERS AND COREWEAVE EXAMPLE · AS-OF 2025

## ~1.8 TB/s vs ~400 Gb/s

Scale-up vs scale-out bandwidth — the cliff a job falls off when it spills a rack

NVIDIA · AS-OF 2025

THE FRAME SHIFT

# Training economics often make goodput the governing workload metric.

*Facility continuity remains a design input. Checkpoint/restart changes interruption cost; it does not choose N, N+1, 2N, or a Tier. Select from required maintenance and fault outcomes.*

STANDARDS RUN ON TWO CLOCKS

# v0.7.0

## The version of OCP's Diablo 400 sidecar-power spec — a timestamp, not a settled fact

*Open-hardware specs now re-version in months; resilience and thermal standards move in multi-year cycles. The question is never just which standard — it is which version, and how soon it drifts.*

OCP (GOOGLE/META/MICROSOFT) · AS-OF MARCH 1, 2026

READING A PROMISE

# What a redundancy claim actually claims

| THE SPEC SAYS | WHAT TO VERIFY | IF YOU DON'T |
|---|---|---|
| **2N UPS** | Independent all the way to the rack, or a shared bus and transfer switch downstream? | A shared throat is a single point of failure behind a '2N' label |
| **N+1 cooling** | Redundant CDUs and pumps and heat rejection — or just one stage of the loop? | The unredundant stage caps continuity for the whole chain |
| **Tier III certified** | Design, constructed facility, or operations? Three separate certificates | A design certificate says nothing about the as-built or how it is run |
| **Concurrently maintainable** | Every distribution path serviceable live, or only the components? | Path-level work forces the load down |

THE FAILURE RECORD

# What actually takes facilities down

## 45%

Share of impactful outages root-caused to power — most often the UPS itself

UPTIME INSTITUTE ANNUAL OUTAGE ANALYSIS 2025 · AS-OF 2025

## 58%

Of human-error outages: staff not following procedure — process, not topology

UPTIME INSTITUTE ANNUAL OUTAGE ANALYSIS 2025 · AS-OF 2025

## ~7 days / one 512-H100 cluster

Observed interval on one mature 512-H100 cluster — not a per-GPU rate

SEMIANALYSIS, AI NEOCLOUD PLAYBOOK AND ANATOMY · AS-OF 2024-10

THE PREMIUM IN QUESTION

# ~25–40% facility

## The capital premium for fault tolerance over concurrent maintainability

*A dated McKinsey rule of thumb, not a quote. Value the premium from required maintenance/fault states, interruption, post-event load, recovery, workload consequence, and contract—not a workload label.*

SAVILLS · AS-OF 2024-05

THE THROUGH-LINE

# The most expensive redundancy is the kind the workload does not value — and the most expensive number is the one nobody defined before it went into the contract.

TAKE IT HOME

# What to do Monday

01 Name the objective function in writing — TCO per unit of goodput — and score every subsystem tender against it, not local metrics.

02 Map the coupling seams between the facility track and the IT track, and give each seam a single named owner.

03 Classify open decisions on the reversibility quadrant; over-build or hedge only the irreversible, high-leverage forks.

04 Check the as-of date and caveat before any figure drives a fork; run contested figures as a range, never a point.

05 Force every contract to name its power number — grid draw, total facility, or critical IT — and its capacity unit.

06 Read every redundancy claim as three questions: what it survives, what it costs, and whether the workload values it.

THE DEFINITIVE GUIDE TO

# AI Data *Centers*

*This deck is the map. The guide is the territory — every number here is dated, sourced, and kept current in the full chapters.*



READ THE FULL PART, WITH EVERY SOURCE AND DERIVATION, AT AIDATACENTERGUIDE.COM

SOURCES

# Every figure in this deck traces to the guide's dated register (1/2)

McKinsey, 'The cost of compute' · as-of 2025 · https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers

NVIDIA / Tom's Hardware / The Register (CES 2026, GTC); GB300 rack TDP per Lenovo Press LP2357 and HPE QuickSpecs; VR200 ~190–230 kW is an analyst estimate (Kuo; Hashrate Index, 2026), not an NVIDIA specification · as-of 2026-08 · https://www.tomshardware.com/pc-components/gpus/nvidia-shows-off-rubin-ultra-with-600-000-watt-kyber-racks-and-infrastructure-coming-in-2027

SemiAnalysis, Datacenter Anatomy Part 2 – Cooling Systems, 13 February 2025; RDHx discussion citing nVent · as-of 2025-02-13 · https://newsletter.semianalysis.com/p/datacenter-anatomy-part-2-cooling-systems

POWER Magazine · as-of 2026-08 · https://www.powermag.com/transformers-in-2026-shortage-scramble-or-self-inflicted-crisis/

Deloitte TMT Predictions 2026; McKinsey · as-of 2026 · https://www.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2026/compute-power-ai.html

SemiAnalysis, Datacenter Anatomy Pt 1 / Inside the 800VDC Revolution · as-of 2025 · https://newsletter.semianalysis.com/p/inside-the-800vdc-revolution-part

SOURCES

# Every figure in this deck traces to the guide's dated register (2/2)

CNBC; theCUBE; DeepQuarry; Aravolta · as-of 2026 · https://www.cnbc.com/2025/11/14/ai-gpu-depreciation-coreweave-nvidia-michael-burry.html

Uptime Institute Global Data Center Survey 2025 · as-of 2025 · https://uptimeinstitute.com/resources/research-and-reports/uptime-institute-global-data-center-survey-results-2025

SemiAnalysis ClusterMAX 2.0 provider tiers and CoreWeave example · as-of 2025 · https://newsletter.semianalysis.com/p/clustermax-20-the-industry-standard

NVIDIA / SemiAnalysis · as-of 2025 · https://nvidianews.nvidia.com/news/nvidia-blackwell-platform-arrives-to-power-a-new-era-of-computing

OCP (Google/Meta/Microsoft) · as-of March 1, 2026 · https://www.opencompute.org/documents/ocp-specification-diablo-400-v0p5p2-2025-05-30-pdf

Uptime Institute Annual Outage Analysis 2025 · as-of 2025 · https://intelligence.uptimeinstitute.com/resource/uptime-institute-global-data-center-survey-2025

SemiAnalysis, AI Neocloud Playbook and Anatomy · as-of 2024-10 · https://newsletter.semianalysis.com/p/ai-neocloud-playbook-and-anatomy

Savills · as-of 2024-05 · https://pdf.euro.savills.co.uk/european/european-commercial-markets/spotlight-eu-data-centre---may-2024---final.pdf

SOURCES

# About these figures

*Figures are resolved by reference from the guide's claim register at build time (July 2026 edition) — a number appears here only as it appears, dated and sourced, at aidatacenterguide.com. Newer data may exist: check the live register.*