

The Stranded Power Problem

Why every data center is operating below capacity it has already paid for — and the three-state model that surfaces the gap.

Executive Summary

Most data center management systems track power capacity with a single number: how much is in use today. That number is necessary, but it is not sufficient. Modern facilities operate in three operational realities simultaneously — power **sold** to customers under contract, power **provisioned** to specific equipment in the design system, and power **consumed** as measured at the outlet. A facility that conflates these three quantities into one capacity number routinely makes the same three classes of mistakes:

1. **It oversells capacity** because the management system reports “available” based on consumed-only metrics, hiding contracted capacity that customers have not yet deployed.
2. **It strands capacity** because provisioned allocations stay locked to the nameplate rating of installed equipment, even though that equipment is typically drawing 40–60% of nameplate under real load.
3. **It misrepresents headroom** because consumption that exceeds the provisioned allocation goes unflagged, leaving operators to plan against falsely optimistic numbers.

This paper introduces the **three-state capacity model** that Rackvio implements as part of its Cloud Growth and Enterprise tiers, covered by US Provisional Patent Application No. 64/022,477 (filed March 2026; patent pending). The model maintains sold, provisioned, and consumed as first-class, separately stored quantities at every node of the power distribution hierarchy — from utility feed down to the device outlet — and computes three inter-state gap signals continuously: deployment lag, stranded capacity, and provisioning error.

The operational consequence: a typical operator who deploys the three-state model surfaces **15–35% recoverable capacity** in existing facilities, without new construction, new switchgear, or any change to physical infrastructure. For a 1 MW facility, that is between 150 and 350 kW of capacity that was already paid for but invisible. At the average colocation wholesale rate of \$150–\$200 per kW per month, that is **\$270,000 to \$840,000 per year** of revenue or deferred CapEx hiding in the difference between three numbers your management system is currently representing as one.

This paper explains why the binary model is computationally insufficient, defines the three states and their gaps precisely, shows how to quantify stranded power in a facility you already operate, and outlines the operational workflows the three-state model enables for capacity planning, sales/billing reconciliation, and sustainability reporting.

The Problem: Three Realities, One Number

Conventional DCIM and capacity management systems use a binary capacity model:

```
available_capacity = total_capacity - used_capacity
```

This model is operationally adequate only when “used” has a single, unambiguous meaning. In modern data centers — colocation facilities, managed service environments, and enterprise sites where capacity is committed to internal business units — “used” can mean any of three different things, depending on who is asking and what they need to decide.

Reality 1: Sold

A colocation provider sells a tenant 12 kW under a multi-year contract. From the moment the contract is signed and a specific circuit is anchored, that 12 kW is **committed**. It cannot be sold to anyone else. Whether or not the tenant has racked any equipment yet, the capacity is gone. A management system that does not represent this state will, on day one of the contract, report the 12 kW as “available” — and the sales team will sell it to a second customer.

Reality 2: Provisioned

The same tenant racks a Dell R750 with two 1100W power supplies. The DCIM records 2,200 W provisioned to that asset, derived from nameplate. From the perspective of the capacity planner reviewing the layout, that 2,200 W is allocated — it is part of the design and cannot be reassigned to another device without conflict. Provisioned capacity expresses the system’s *expectation* of consumption.

Reality 3: Consumed

The R750 actually draws 380 W under typical load. The PDU outlet metering reports 380 W. This is the *physical* reality. It is what the utility bill, the HVAC sizing, and the failover load calculation all care about.

These three quantities — sold, provisioned, consumed — are operationally distinct, and at any given instant they almost always disagree.

Node	Sold	Provisioned	Consumed
Tenant contract	12.0 kW		
Allocated assets		8.6 kW	
Outlet meter			3.4 kW

A binary model has one column. The three-state model has three. The information that the binary model cannot represent — *the gaps between the columns* — is precisely where stranded power, oversell risk, and provisioning errors live.

The Three Classes of Computational Error

A management system using a binary model produces three predictable errors:

Error 1 — Phantom availability. The system reports consumed-only headroom. The 8.6 kW that the tenant has provisioned but not yet powered on appears as available. The sales team commits it to another customer. Six months later, the first tenant racks more equipment and the system over-commits the circuit. This is a **deployment lag** error.

Error 2 — Invisible recovery opportunity. The same R750 was provisioned at 2,200 W but draws 380 W. That **1,820 W per device of stranded capacity** is multiplied across thousands of devices in a facility. The binary model, which only tracks “consumed,” cannot see the gap between the allocation and the actual draw. The capacity is recoverable but computationally invisible.

Error 3 — Silent provisioning drift. Over time, devices are added, removed, or moved without their provisioning records being updated. Eventually consumed exceeds provisioned for some circuits. A binary system records actual draw and reports headroom against total. There is no signal that the provisioning record is *wrong* and that planning calculations based on it will be wrong too.

None of these errors are operator mistakes. They are mathematical consequences of using a data model that cannot represent the three quantities the business actually depends on.

The Three-State Capacity Model

Rackvio's three-state model replaces the single capacity number at each node in the power hierarchy with three explicitly typed quantities, each measured in kilowatts, each stored independently:

```
capacity_summary(node) = {  
  sold:      kW contractually committed to customers,  
  provisioned: kW allocated to assets recorded in the system,  
  consumed:  kW measured at the outlet (or estimated from model),  
  total:    kW physical rated capacity of this node  
}
```

The model is maintained at every level of the power distribution hierarchy — utility feed, generator/UPS, floor PDU, circuit, rack PDU, outlet, and device — and aggregated upward by simple summation:

```
parent.sold      = SUM(child.sold      for each child)  
parent.provisioned = SUM(child.provisioned for each child)  
parent.consumed  = SUM(child.consumed  for each child)
```

This aggregation is incremental: a value change at a leaf propagates upward through ancestor nodes in $O(\text{depth})$ time — typically four to seven nodes — without recomputing the entire tree. The system can therefore maintain accurate aggregate capacity values in real time across hierarchies of thousands of nodes.

Consumed Capacity Has a Provenance

Not all “consumed” values are equally reliable. The patent-pending model distinguishes consumed values by source:

Source	How produced	Confidence
Metered telemetry	Live readings from intelligent PDU outlets or branch-circuit meters	High
Equipment-model typical draw	Catalog “typical draw” value for the installed model	Medium
Vendor nameplate rating	Manufacturer worst-case rating	Low

Operators should weight gap calculations by source confidence. A 600 W gap grounded in metered telemetry is an actionable recovery opportunity. A 600 W gap grounded in nameplate-only values is a hypothesis that needs metering before it should drive an asset move. Rackvio carries the underlying provenance through power-node and outlet records; surfacing a typed confidence field on aggregate gap reports is on the product roadmap.

The Three Gaps and What Each Reveals

The value of the three-state model is the three inter-state gaps it makes computable. Each gap is a continuously maintained signal exposed at every node in the hierarchy, with its own operational meaning and its own threshold-based alerting.

Gap 1 — Deployment Lag (sold – provisioned)

$$\text{deployment_lag} = \text{sold} - \text{provisioned}$$

A positive value means a customer is paying for capacity they have not deployed. Two operational consequences:

- **Billing accuracy.** The customer is paying full rate for capacity that draws no marginal energy and produces no marginal cooling load. This is generally desired behavior in a colocation contract, but it should be **visible** for renewal conversations, customer success workflows, and capacity right-sizing offers.
- **Deployment-window monitoring.** Many contracts include language requiring the customer to deploy by a date. The system can flag contracts where the start date has passed and the deployment lag is still high.

A negative deployment lag means the customer has provisioned *more* than they have contracted — an **oversell condition** that warrants immediate operator review.

Gap 2 — Stranded Capacity (provisioned – consumed)

$$\text{stranded_capacity} = \text{provisioned} - \text{consumed}$$

This is the headline gap. A positive value means power has been allocated to specific assets that are not drawing it. The aggregate across a facility is the **recoverable headroom** — capacity already paid for by past CapEx that can be redeployed to new customers, new equipment, or denser configurations without new construction.

Stranded capacity recovery is not theoretical. It is grounded in three empirical realities:

1. **Nameplate ratings are worst-case.** A 1100W PSU is sized for power-up surge, full-load fault tolerance, and the vendor's certification regime. Steady-state draw is consistently 35–55% of nameplate in production workloads.
2. **Redundant supplies double the count, not the draw.** A dual-PSU server with two 1100W supplies records 2,200 W of provisioned capacity but draws roughly the same as a single-PSU equivalent (with the second PSU sharing the load in N+N or idle in N+1).
3. **Equipment lifecycle gaps.** A rack designed for a 24-server build often runs with 16–20 servers actively powered, leaving 4–8 provisioned slots stranded.

Independent academic studies have repeatedly found provisioned-to-actual ratios of 1.8x to 3.0x in operational data centers, with average IT utilization clustered in the 30–50% range. The Lawrence Berkeley National Laboratory's *United States Data Center Energy Usage Report* (Shehabi et al., LBNL, 2016) and subsequent updates establish the canonical academic baseline. Uptime Institute's *Global Data Center Survey* (annual) corroborates the pattern from operator-self-reported data — average PUE has remained near 1.55–1.60 across the industry for over a decade, even as IT power density has grown, indicating that the gap between provisioned and consumed has not narrowed organically. The recoverable fraction — what the

three-state model surfaces as actionable — is typically **20% to 50% of provisioned capacity, or 15% to 35% of total facility capacity** in a mixed-load facility.

Gap 3 — Provisioning Error (consumed – provisioned)

`provisioning_error = consumed – provisioned`

A positive value means actual consumption is **exceeding** the recorded allocation. Always a data-integrity signal. Either:

- The provisioning record is wrong (an asset was added, replaced, or upgraded without updating its allocation), or
- The equipment is failing or misconfigured (a server in a fault state can draw substantially above expected), or
- The model “typical draw” used to estimate consumed is below reality and needs adjustment.

Whichever is the case, an operator who sees this gap can act on it. An operator who only sees consumed and total cannot.

The model supports threshold-based alerting on the provisioning-error gap (default: 10% of provisioned). Alert generation on this gap is a roadmap item — the underlying values are computed continuously, so an operator can build a report against the gap field today, and threshold-driven push alerts ship with the broader alerting framework.

Three Available Capacity Values, Each for a Different Role

A binary model derives one “available” capacity. The three-state model derives three:

Available	Formula	Question it answers
Available to sell	<code>total – sold</code>	What can the sales team commit to a new customer right now?
Available to provision	<code>total – provisioned</code>	What can the operations team allocate to new equipment right now?
Available headroom	<code>total – consumed</code>	What is the actual physical safety margin under current load?

These three numbers are routinely different by tens of percentage points in a real facility, and they answer the questions of three different operational roles. A sales engineer planning a new tenant contract needs *available to sell*. A capacity planner adding rack space needs *available to provision*. A facilities engineer sizing a UPS failover scenario needs *available headroom*. A binary model forces all three roles to make decisions from the same single number — and that number is, by definition, wrong for at least two of them.

Quantifying Stranded Power in Your Facility

The three-state model becomes immediately useful when applied to your actual data. The first-pass calculation is straightforward, requires no software install, and can be done with two CSV exports from your current systems.

Inputs

1. **Provisioned (from your DCIM or asset register):** the sum of nameplate or estimated power values for every active device across the scope you want to analyze. Roll up to the room or facility level.
2. **Consumed (from your power monitoring or BMS):** the sum of metered branch-circuit or outlet draw across the same scope, averaged over a representative 30-day window (avoid month-end batch jobs and known maintenance windows).

Calculation

$$\text{stranded_kw} = \text{provisioned_kw} - \text{consumed_kw}$$
$$\text{stranded_ratio} = \text{stranded_kw} / \text{provisioned_kw}$$

A typical result for a managed enterprise data center or a colocation hall lands between 0.35 and 0.55. A typical result for an HPC or high-density AI training pod lands between 0.10 and 0.25 — these workloads run closer to nameplate.

Dollar value

Translate to dollars using your facility's cost-of-capacity. Two common methods:

- **Colocation revenue equivalent:** $\text{stranded_kw} \times \text{monthly_rack_rate} / \text{typical_kw_per_rack} \times 12$. At \$1,500/rack/month and 6 kW/rack, each stranded kW represents \$3,000 of annual revenue capacity.
- **Deferred CapEx:** $\text{stranded_kw} \times \text{build_cost_per_kw}$. At \$10,000–\$15,000 per kW for new colocation build-out, each stranded kW represents \$10,000–\$15,000 of avoided CapEx that the three-state model recovers without breaking ground.

For a 500 kW facility at the median 40% stranded ratio:

- Stranded capacity: **200 kW**
- Annual colo revenue equivalent: **~\$600,000**
- Deferred CapEx equivalent: **\$2.0M–\$3.0M**

These numbers are not Rackvio-specific. They are inherent in the gap between provisioned and consumed, which the three-state model surfaces and the binary model cannot.

How Rackvio Implements the Three-State Model

Rackvio's Cloud Growth and Enterprise tiers maintain the three-state capacity model continuously across the full power hierarchy of every site in your tenancy. Specifically:

- **Sold capacity** is bound to customer contract records. Creating a contract in Rackvio with a power entitlement automatically creates the sold capacity row on the anchored circuit. Contract start/end dates flow into deployment-lag alerting.
- **Provisioned capacity** is bound to asset records. Every device placed in a rack carries its provisioned power value (operator-entered, equipment-model typical, or nameplate), and aggregates upward through the rack PDU, circuit, panel, UPS, and feed.
- **Consumed capacity** is bound to telemetry sources. Where a PDU is integrated (intelligent rack PDU with outlet metering, BMS branch-circuit meters, or third-party power telemetry feeds), consumed reflects the current measured draw. Where telemetry is absent, the equipment-model typical-draw values are used. The first-class, typed confidence field for downstream filtering is on the product roadmap; today the data is carried at the source records.

- **Gap analytics** are computed against the live state of the power hierarchy. The three gaps are visible at every node in the UI. The stranded capacity recovery report identifies candidate nodes ranked by recoverable headroom.
- **Three-phase A/B/C circuit modeling** extends the three-state values to per-phase load. The per-rack phase-summary endpoint reports phase-A / phase-B / phase-C watts and percent-of-breaker for each leg, and flags imbalance when any leg exceeds 80% of the upstream breaker rating — surfacing circuits that look safe in aggregate but are over-provisioned on a single phase.
- **kWh-to-dollar translation** uses a per-organisation energy rate (configurable, with optional time-of-use blocks) to convert consumed kW into operating cost. The operator-side report that pairs stranded kW with the per-org rate to produce stranded-dollar value is on the near-term product roadmap; the underlying inputs are already stored.
- **Failover-overload detection** simulates a single-UPS failure across your topology and identifies the cabinets that would exceed surviving capacity. The algorithm walks the power chain to identify racks fed by ≥ 2 distinct UPS ancestors, simulates the worst-case single-UPS loss, and flags any rack whose load would exceed the surviving UPS's capacity. The output is the "X cabinets at risk if UPS Y fails" view operations teams act on directly.

The patent-pending status of the underlying capacity-state and gap-analytics architecture (US Provisional Patent Application No. 64/022,477) means the model is a Rackvio differentiator. Conventional DCIM systems offer two-column reporting (allocated vs. used) at best, and binary single-number availability at worst.

Operational Use Cases

The three-state model enables operational workflows that a binary system cannot support. Five examples that customers most commonly act on first:

Sales and Billing

- **Sales floor capacity dashboard.** Available-to-sell, not available-headroom, drives the floor map shown to the sales team. A 70% utilized facility on a binary model can be fully sold-out (and over-committed) on the three-state model.
- **Customer utilization reports.** Each tenant sees their own sold / provisioned / consumed breakdown. Right-sizing renewals, expansion conversations, and capacity true-ups all become quantitative instead of negotiated.
- **Deployment lag alerts.** Tenants who have not deployed contracted capacity within a configurable window receive customer-success outreach — protecting churn, surfacing expansion intent, identifying customers who are about to need help.

Capacity Planning and Operations

- **Stranded capacity recovery reports.** Ranked, per-device, scoped by room or building. The top of the list is typically tens of kW recoverable from a handful of over-provisioned assets, payable as deferred CapEx.
- **Provisioning audit.** The provisioning-error gap surfaces specific circuits where consumed exceeds provisioned — high-confidence data integrity work for the ops team.

Sustainability and Compliance

- **Per-site IT load tracking and per-tenant consumption breakdown** feed the IT-side inputs that both EU Energy Efficiency Directive Article 12 reporting and California SB 253 Scope 2 GHG calculations depend on. Hourly time-series storage for long-horizon market-based reporting is a Phase 5.5 roadmap item;

today operators pair Rackvio's per-site IT state with utility-meter or BMS-exported kWh totals to produce annual indicators. See the companion *EED Compliance Playbook* and *SB 253 Quick-Start* for the regulatory mapping.

Conclusion

The stranded power problem is not a deficiency of operators, vendors, or data centers. It is a deficiency of the **data model** that data centers have historically used to represent capacity. A binary single-quantity model cannot encode three operationally distinct realities, and the gaps between those realities — gaps that average 30–50% in real facilities — are where capacity, revenue, and reporting fidelity hide.

The three-state capacity model is a tractable, well-defined replacement. It changes the underlying data structure, computes the three gaps continuously, and surfaces them in dashboards and APIs that operators, sales teams, and sustainability functions all act on. The first-pass calculation can be done on a spreadsheet. The continuous, hierarchy-aware, alert-driven implementation is what Rackvio is built around.

If your facility is operating on a binary capacity model today, you are leaving capacity, revenue, and reporting accuracy on the table. The size of the gap is calculable. The path to recovering it is not theoretical.

Next Steps

- **Run the first-pass calculation** above with two CSV exports from your current systems. If your stranded ratio is below 0.20, the three-state model still helps you for billing and compliance. If it is above 0.30, the recovery economics are likely to justify a switch on their own.
- **Read the companion documents.** *EED Compliance Playbook* maps the three-state model to EU Directive 2023/1791 Article 12 reporting requirements. *SB 253 Quick-Start* maps it to California Climate Corporate Data Accountability Act Scope 2 disclosures.
- **Request a working session.** Bring an anonymized inventory CSV; we run the calculation with you and walk the resulting recovery candidates. Book at rackvio.com/demo.

© 2026 Rackvio. The three-state power capacity model and gap analytics framework described in this paper are the subject of US Provisional Patent Application No. 64/022,477 (filed March 2026). Numerical ranges for stranded capacity recovery are illustrative; actual results depend on workload profile, redundancy posture, and metering granularity in the facility analyzed.