INTELLIGENCE UPDATE

# How scale and occupancy shape data center and colo economics

**Owen Rogers**
6 Jan 2026

Choosing whether to place workloads in a colocation facility or an enterprise-owned data center involves many considerations: capability, risk posture, operating model and strategic priorities. Cost is never the sole factor — but it is a significant one.

This report presents a simplified, normalized cost model that enables an enterprise to compare the cost of building a new center against the cost of a colocation facility with the same characteristics.

The intent is not to predict an exact price for a particular site. Instead, the model provides a directional benchmark that highlights two of the most important drivers of data center cost efficiency:

1. **Occupancy** — the average proportion of a data center's maximum designed critical power capacity that is actively used by IT load.
2. **Scale** — the size of the data center, measured in megawatt of IT load.

By expressing costs as ratios rather than absolute values, many complexities cancel out. The comparison focuses solely on occupancy and scale to illustrate relative unit costs. If other factors — such as resiliency, financing, tax conditions or geographic location — differ between the colocation and enterprise facility being compared, the results will vary.

The result is a practical model that highlights the main cost drivers without unnecessary complexity.

## Background

Data center cost structures differ sharply between colocation providers and enterprise-owned facilities. Colocation providers such as Equinix and Digital Realty operate at large scale, achieve high utilization, and deploy optimized designs. As a result, they typically report:

- Net profit margins of around 8-15% over the past several years. Equinix has reported a net margin of 7-13% between 2021 and 2024, while Digital Realty has typically reported margins in the 6-15% range.
- Near-zero vacancy in many markets. The CBRE Group reports global vacancy of around 5-6%, with metropolitan areas such as Northern Virginia, Frankfurt, Singapore and Tokyo frequently below 3%.
- Scale efficiencies from 20-100 MW campuses.
- Standardized designs and procurement advantages.

Enterprises, by contrast, often operate multiple smaller sites characterized by:

- Lower occupancy. The Uptime Institute Global Data Center Survey 2024 reveals that one in four data centers is typically under 40% utilized.
- Higher overhead per MW compared with colocation providers.
- Less predictable demand profiles.
- Legacy infrastructure and site-specific constraints.

This model focuses specifically on how occupancy and scale drive these differences.

## The model

Rather than attempting to rebuild a full data center cost simulation — involving land, construction, mechanical and electrical systems, operational staffing, power, depreciation, tax and financing — we instead use normalized financial data from publicly traded colocation providers.

In summary, the model compares the cost of building and running an enterprise-owned data center with the cost of renting capacity from a colocation provider. It starts with an assumed "ideal" data center — a highly occupied, large-scale, high-efficiency colocation facility — and then adjusts that baseline cost for an enterprise data center using two key factors — how full the facility is (occupancy) and how large it is (scale). These adjustments reflect the reality that smaller sites are usually more expensive per unit of capacity and that unused space drives up costs per unit. By expressing all results as ratios rather than absolute numbers, many other differences — such as labor costs, local energy prices, or specific equipment choices — mathematically cancel out.

A colocation provider's net margin indicates what share of revenue is profit and therefore what share represents cost. This model uses publicly reported colocation margins to establish an objective, auditable baseline of a data center cost structure. It is not intended to represent the "best possible operator," nor a normative target for enterprises. Rather, it provides a standardized and transparent cost basis against which variations in scale and occupancy can be compared consistently.

**If a provider has:**

- Revenue: $R$
- Net Margin: $M$

**Then the total cost can be expressed as:**

- Total cost: $R \times (1 - M)$

This provides a reliable, externally validated anchor point for estimating what it costs a highly efficient operator to plan, build and operate a data center — including the capital, finance, design resources (internal or external), power and operations staff.

To build such a data center (either internally or through a third party), an enterprise would need to perform similar tasks as a colocation provider, albeit at a much smaller scale.

## Variables and definitions

A detailed mathematical reasoning is included here for completeness.

Let:

$R_c$ = Colo revenue
$M$ = Colo profit margin
$C_d$ = Data center capacity
$O$ = Average occupancy (0-100%)
$C_p$ = Ideal cost (cost of an ideal data center)
$U_e$ = Unit cost of enterprise resource
$U_c$ = Unit price of colo resource
$RI$ = Relative Index (cost-efficiency ratio)

The cost of the ideal data center is the colo's revenue minus its margin:

$$C_p = R_c \times (1 - \mathrm{M})$$

This represents the ideal cost structure for a perfectly sized, fully utilized data center.

An enterprise operating at a similar scale pays a unit cost based on capacity and occupancy. The cost per unit — whatever that unit may be — is the total cost divided by capacity used (note used, not available).

$$U_e = \frac{C_p}{C_d \times \mathrm{O}}$$

The colo unit price is revenue over capacity (assuming full capacity due to high occupancy):

$$U_c = \frac{R_c}{C_d}$$

To compare unit prices, divide:

$$RI = \frac{U_e}{U_c}$$

Substitute enterprise and colo unit equations:

$$RI = \frac{C_p / (C_d \times O)}{R_c / C_d}$$

Simplify (cancel $C_d$):

$$RI = \frac{C_p}{O \times R_c}$$

Substitute ideal cost:

$$RI = \frac{R_c \times (1 - M)}{O \times R_c}$$

Simplify (cancel $R_c$):

$$RI = \frac{1 - M}{O}$$

This formula shows that a comparison between colocation and on-premises costs can be calculated based on just two data points: colo margin and occupancy. Next, the formula can be adjusted to factor in the economies of scale that a colocation provider has over an enterprise.

## Scale adjustment

Construction cost per megawatt varies significantly by facility size. Costs are lower at a larger scale because fixed expenses — such as power infrastructure, cooling systems and staffing — are spread across more capacity, reducing the cost per megawatt. Furthermore, larger infrastructure purchases often result in a discount.

Colocation providers tend to have lower average build costs per megawatt because they construct facilities at a high scale. To normalize this cost base so that it is more realistic for an enterprise data center, it is scaled by a factor based on average build costs.

Using Cushman & Wakefield's Data Center Development Cost Guide 2025, average US construction costs range from:

- $13 million per MW (1-5 MW sites)
- $11.7 million per MW (5-20 MW sites)
- $10.3 million per MW (20 MW+ campuses)

Normalizing the 20 MW+ cost to 1.0 gives the following:

- Small (1-5 MW): S = 1.26
- Medium (5-20 MW): S = 1.14
- Large (20+ MW): S = 1.0

These values align closely with empirical industry benchmarks: smaller data centers cost materially more per megawatt due to less efficient cooling, power, labor and redundancy scaling.

The Relative Index can be scaled up or down using these scaling factors to reflect the higher build and operating costs faced by enterprises compared with colocation providers:

$$\mathrm{RI}_{\mathrm{s,o}} = \mathrm{S} \times \frac{1 - \mathrm{M}}{\mathrm{O}}$$

## Occupancy

Colocation providers typically operate at very high occupancy because capacity is sold only when customers commit. Public data from CBRE, Equinix and Digital Realty shows:

- Vacancy is often below 5% in key markets.
- Effective utilization near 100% once stabilized.

Enterprises, by contrast, tend to operate data centers with:

- Reserved space.
- Hardware refresh cycles.
- Fragmented demand across sites.
- Deliberate overhead to support resiliency.

As a result, realistic enterprise occupancies are far lower, on average, than those of colocation facilities.

Occupancy is the largest single cost driver in the model.

## Results

The model is unit-agnostic: the "unit" may be a rack, a square meter, a kilowatt, or a megawatt of capacity — it does not matter, as long as the same unit is used consistently for both the colocation provider and the enterprise. This ensures a true like-for-like comparison in which differences arise only from scale and occupancy, not from the choice of measurement. However, power is the primary engineering and commercial constraint in modern data centers, and it implicitly captures the associated scaling of space and cooling for a given design. Therefore, it is best to calculate comparison units and occupancy using power.

Unit cost is calculated based on the capacity that is actually used to deliver value, rather than capacity that sits idle. As a result, lower occupancy directly increases the effective unit cost by spreading the facility's fixed cost over a smaller amount of productive space.

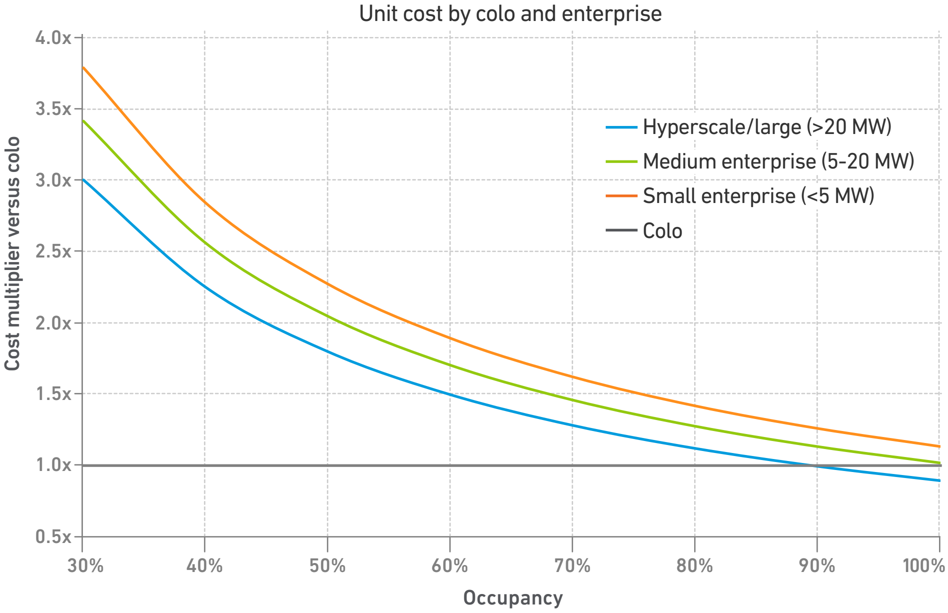
Using a typical net margin of 10% (based on historical results from Equinix and Digital Realty), the model yields:

**Relative Index =** $S \times 0.9 / O$

This financial margin is a reasonable average globally, although it will vary by area depending on the supply and demand for capacity.

**Figure 1** illustrates the relative unit cost of enterprise data centers compared with a colocation baseline (indexed at 1.0). The unit cost of a colocation facility is constant as the enterprise is not affected by the occupancy — the colocation provider themselves absorbs this in their costs.

*Figure 1 Colocation vs enterprise costs*




UPTIME INSTITUTE 2026

This comparison is intended to provide a like-for-like view of cost. It assumes that the colocation provider and the enterprise operate their data centers in broadly similar ways and that the "customer" in each environment — an internal IT team within an enterprise or a paying tenant in a colocation facility — is procuring capacity on comparable terms. This comparison only applies to net new builds, not to retrofitted facilities or written-off/depreciated assets.

Without these common assumptions, differences in operating model or procurement behavior would distort the results.

## Analysis

Use of the model is best illustrated using a hypothetical example. An enterprise is considering building a new fault-tolerant data center in Cardiff, Wales, to support a growing portfolio of digital services. An alternative option is to deploy the same workloads within an existing fault-tolerant colocation facility in the same metropolitan area, where capacity is currently available. Both facilities are assumed to have similar PUE. The enterprise is a large multinational organization with low-cost access to finance. The enterprise keeps regulated workloads and those requiring high performance or low latency in its own facilities. The Cardiff location will be used for non-regulated workloads.

The proposed enterprise facility is relatively small, at 4 MW of critical IT load capacity. Demand is expected to ramp gradually over a five-year period, reaching 1.6 MW in Year 1 (40% occupancy), 2.4 MW by Year 3 (60%) and 3.2 MW by Year 5 (80%).

Under the model, this places the enterprise in the "small (<5 MW)" scale band (orange line), with a corresponding scale factor applied to the cost base. At 40-60% occupancy in the early years, the enterprise facility would operate with materially higher unit costs than a colocation provider (dark gray line). By contrast, the colocation option enables the enterprise to contract only the required power capacity (e.g., 1.6-3.2 MW) at each stage, avoiding stranded capacity as demand ramps.

This example demonstrates how small-scale and gradual utilization growth structurally disadvantage a self-built facility in the early years, while a colocation model aligns cost more closely with deployed load. Only if the enterprise sustains high long-term utilization (above 85-90%) does the cost gap materially narrow.

More generally, across the full range of occupancy levels, colocation remains the lowest-cost option, with a constant cost index of 1.0. By contrast, enterprise facilities show materially higher costs, particularly at lower utilization levels:

- Small enterprises (<5 MW) experience the highest unit costs, reaching almost 4× the cost of colocation at 30% occupancy and remaining above 1.5× even at 80% utilization.
- Medium enterprises (5-20 MW) perform slightly better but still incur 2-3× the unit cost of a colocation provider at typical occupancies (40-70%).
- Large or hyperscale-sized facilities (>20 MW) narrow the gap, approaching colocation economics only when occupancy rises above 90%.

Enterprises only approach cost parity when operating at very large scale and sustaining exceptionally high occupancy — conditions that are rarely achieved in practice.

At typical enterprise occupancy levels (40-70%), unit costs remain 1.5× to 3× higher than colocation. This aligns with observed industry behavior, where colocation providers operate at near-full utilization and benefit from economies of scale that individual enterprises seldom realize.

## Limitations

This model is intentionally simplified and focuses on the structural cost drivers of scale and utilization. It does not attempt to represent the full total cost of ownership. Several factors sit outside scope:

- **Baseline simplification:** The colocation margin benchmark provides a transparent, normalized cost anchor, but it does not represent every possible enterprise design or an optimized greenfield build.
- **Operational differences:** PUE, staffing efficiency and process maturity vary widely. These are indirectly reflected in net margin and are excluded separately to avoid double-counting.
- **Scale effects:** The scale factor uses construction cost per megawatt as a conservative proxy. Real-world scale advantages in procurement and operations would likely widen the gap.
- **Colocation-specific charges:** Cross-connects, remote hands and migration costs vary greatly by deployment and are excluded for comparability, although they may narrow the difference in specific cases.
- **Capital structure:** Cost of capital and required profit differ between enterprises and real estate investment trusts but typically have less impact than scale and utilization.
- **Legacy assets:** The model focuses on forward-looking economics; fully depreciated enterprise facilities can be cheaper, but this reflects sunk-cost advantages rather than structural cost position.
- **Strategic factors:** Latency, security and control requirements are the real drivers of data center decisions but sit outside this cost-normalized comparison.

Despite this, the model shows how unit costs and cost efficiency of an enterprise data center vary with occupancy and scale. It also shows why colocation providers are an attractive proposition for enterprises. However, enterprises should conduct their own detailed analysis before pursuing or avoiding any course of action.

## Value over cost

In this report, we focus purely on the cost of a unit under comparable circumstances. There are valid economic reasons for using an enterprise data center, even if the costs may be higher. A self-owned data center may still be preferable when:

- Workloads are mission-critical and require full control.
- Security or data-sovereignty requirements are non-negotiable.
- Ultra-low latency to on-premises systems is mandatory.
- The enterprise already owns a depreciated site (sunk-cost advantage).
- Unique facilities engineering is required (e.g., for advanced AI infrastructure).
- Capacity is not available or not assured over a longer period.

### The Uptime Intelligence View

> Much of a colocation provider's value lies in acquiring land, obtaining finance, building facilities and operating them efficiently. By aggregating demand, organizations can rapidly and flexibly scale without substantial capital commitments. Given this, it is unsurprising that colos are generally cheaper than net-new private facilities on a per-unit basis. Organizations will continue to use a range of venues, colo and on-premises facilities alongside cloud, choosing the most appropriate venue for each workload.

*Thanks to Dr Kerrison Steven, James Cook University, Singapore*

## ABOUT THE AUTHOR

### Dr. Owen Rogers

Dr. Owen Rogers is Uptime Institute's Senior Research Director of Cloud Computing. Dr. Rogers has been analyzing the economics of cloud for over a decade as a chartered engineer, product manager and industry analyst. Rogers covers all areas of cloud, including AI, FinOps, sustainability, hybrid infrastructure and quantum computing.

orogers@uptimeinstitute.com