RESEARCH REPORT / VOLUME 02

Q4 2026
OCTOBER SNAPSHOT

STATE OF GPU SERVICES

SUPPLY / PRICING / PROVIDER EVIDENCE



Q4 2026 / MARKET UPDATE

Evidence captured 25-27 October / Early-Q4 snapshot

872
CANDIDATE DOMAINS REVIEWED

339
CONFIRMED GPU-SERVICE PROVIDER DOMAINS

4,511
SCOPED PRICE RECORDS

00 / MARKET RESEARCH

Inside this edition

A concise early-Q4 research briefing, using the October 2026 evidence refresh.

This Q4 2026 early-quarter briefing uses October supplier and GPU-model evidence and a strictly matched Q2-to-October pricing study. It is not a full-quarter market census.

Contents

Page	Focus
03	The October audit: supplier and pricing denominators
04	Which GPU models appear in public price evidence
05	Rates and terms: when prices are comparable
06	Q2-to-October: verified changes in matched offers
07	Advertised GPU quantities and supplier definitions
08	Findings, evidence limits and source notes

The principal findings

339 explicitly confirmed GPU-service-provider domain rows in the October audit. Separately, 221 domains contributed 4,511 heterogeneous public-price records. Among 60 historically matched price offers, 41 stayed unchanged, 15 rose and four fell.

Two dates, different purposes

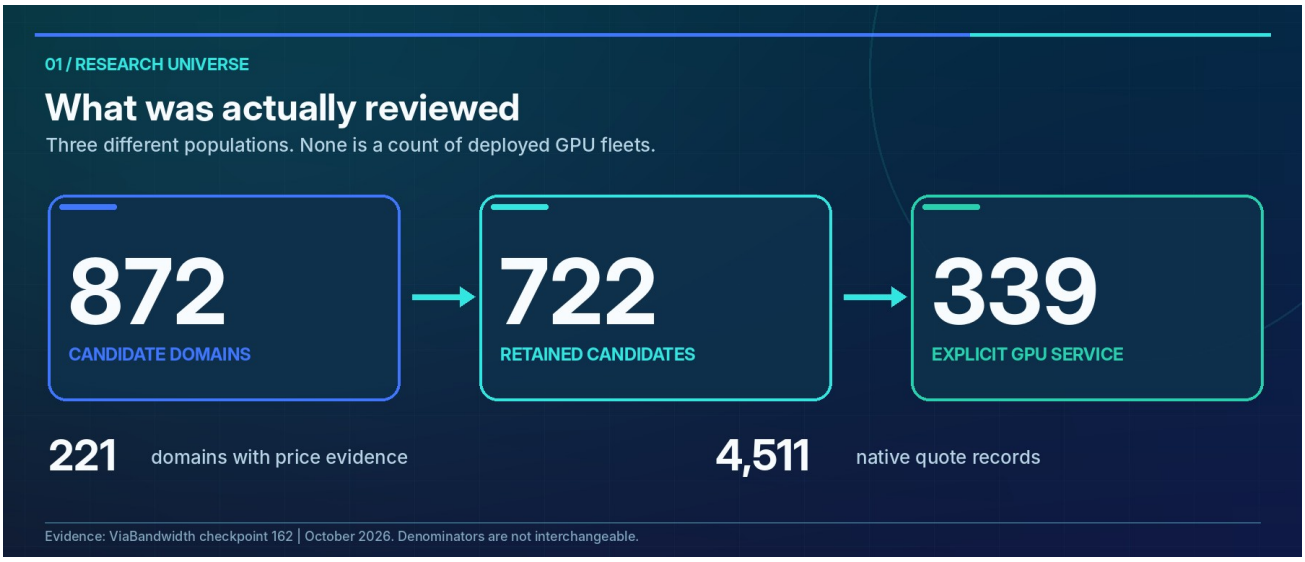
This Q4 2026 edition is an early-quarter snapshot, with evidence captured 25-27 October rather than a full October-December review. Its 60-offer price comparison uses the Q2 baseline and October observations. The July Q3-branded report used a different directory population.

Read every chart with its denominator: candidate domain, confirmed GPU-service-provider domain, price-bearing domain, posted configuration and same-offer matched price are different measurement units.

01 / MARKET RESEARCH

The October audit in one view

The research population is larger than the confirmed GPU-service-provider set.



The audit reviewed 872 candidate domains and excluded 150, leaving 722 active candidates with documented dispositions. Within those, 339 rows were explicitly labeled GPU service providers. A separate count of 216 retained domains had numeric or indicative publicly reported pricing.

Do not read 722 as distinct, confirmed GPU suppliers. The audit includes reseller, marketplace, managed-service, infrastructure-adjacent and access-limited cases. It closed 79 manual-access exceptions without treating them as verified offers.

The evidence base

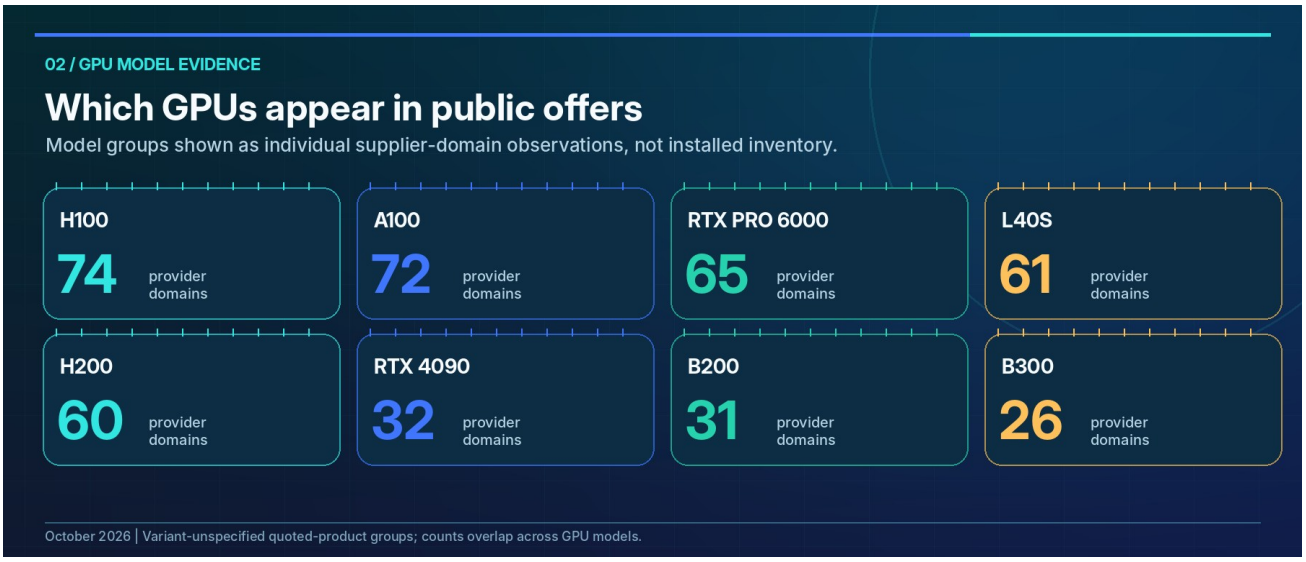
Across 221 price-bearing domains, the ledger contains 4,511 native price representations. Those records vary in object, currency, GPU count, billing unit, commitment and region. They are not 4,511 interchangeable per-GPU-hour quotes.

Source: ViaBandwidth checkpoint-162 aggregate, with October capture evidence. The underlying domain and offer ledgers remain private.

02 / MARKET RESEARCH

The GPUs found in public rate evidence

Model-by-model October observations, with clear eligibility and unit boundaries.



The largest groups shown here are H100 (74 price-bearing domains), A100 (72), RTX PRO 6000 (65), L40S (61) and H200 (60). B200 and B300 appear in 31 and 26 price-bearing domains respectively; MI300X appears in 11. These are conservative groups where the precise GPU variant was not specified.

This chart draws from October native-price evidence in a 221-domain ledger, NOT all 339 confirmed GPU-service-provider domains. Provider counts overlap across models. Neither published quotations nor their domain counts prove stock or installed GPUs.

What about the original July report?

The earlier public snapshot tracked 588 indexed operators and more than 45 accelerator models. It listed 229 operators for H200, 191 for B200, 91 for MI300X and 31 for MI325X. Those figures use different inclusion rules, so they are historical context, not bars in a Q2-to-October growth chart.

Sources: published_pricing_by_model.csv (2 November export); July State of GPU Cloud page. Both have distinct populations and methods.

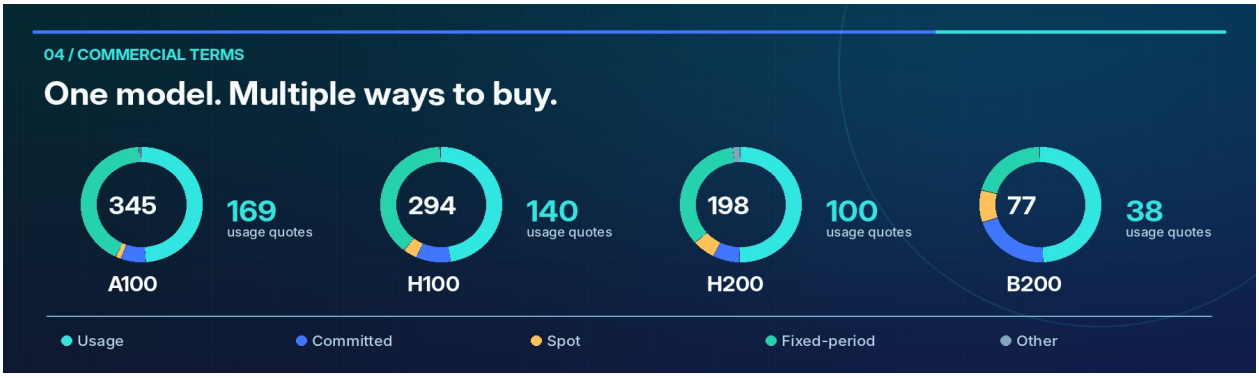
03 / MARKET RESEARCH

Public pricing is not one price

Rates, commitments and sold configurations create materially different comparisons.



In the narrow October subset explicitly denominated in native USD per GPU-hour, the observed median was $2.41 for H100 (8 offers, 5 suppliers), $3.84 for H200 (6, 3), $6.50 for B200 (4, 2) and $1.49 for A100 (5, 3). These small samples are descriptive, not a market-wide price index.



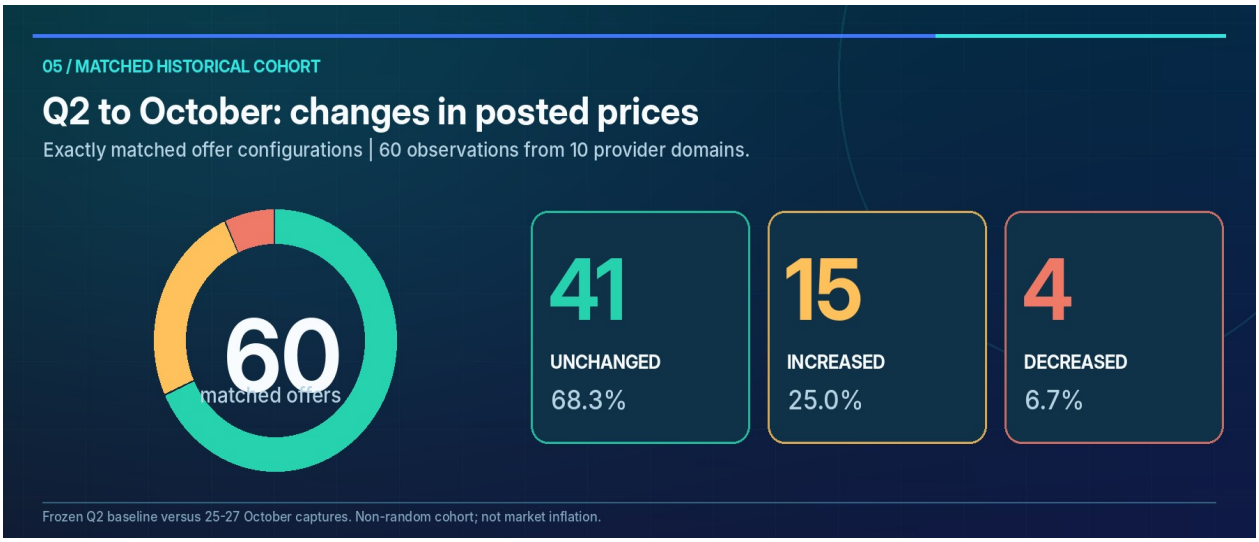
The same GPU model also appears in usage, commitment, spot and fixed-period quotations. The lower chart counts saved native quote records within four variant-unspecified model groups, not suppliers or comparable GPU-hour prices.

No blanket currency conversion or “price per GPU” normalization is appropriate across unlike products. Compare full configurations under the same contract, region and term.

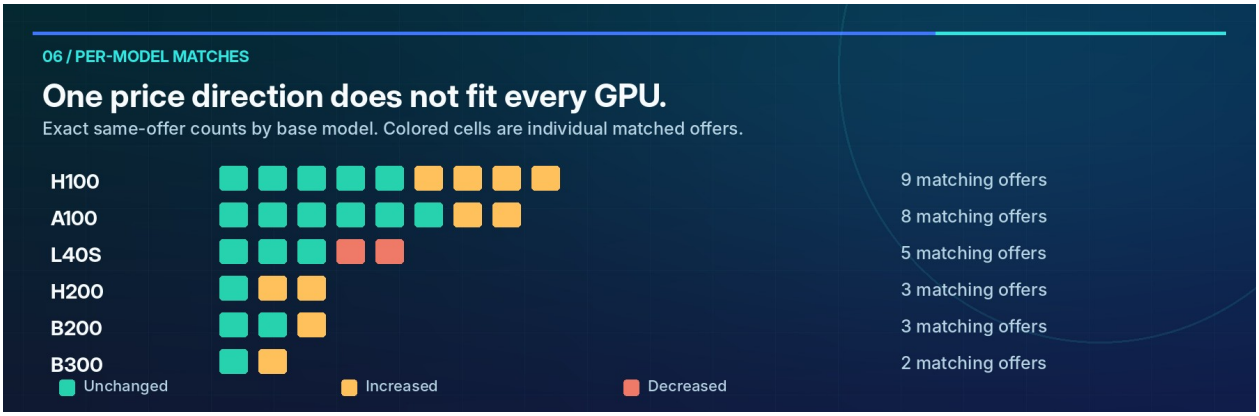
04 / MARKET RESEARCH

Q2 to October: what actually changed

A revalidated, same-offer comparison. Unlike the wider market snapshots, this is matched.



Codex recalculated the 60-offer private historical ledger: 41 unchanged (68.3%), 15 increased (25.0%) and four decreased (6.7%), across 10 provider domains. Every matched row passed validation; no arithmetic discrepancies were found.



Within the selected model groups above, H100 recorded four higher and five unchanged quotes (9 matches); A100 recorded two higher and six unchanged (8 matches). Two of five L40S matches declined. Small model samples are descriptive only.

Matched on provider, same sold object, GPU variant/count, region, currency, contract term and billing unit. No currency or unit conversion. This non-random cohort cannot establish GPU-market inflation or supplier-wide pricing trends.

05 / MARKET RESEARCH

GPU quantities are not GPU capacity

A genuinely new quantitative view, with an important distinction.



The audit can count GPU quantities explicitly stated inside distinct advertised configurations: for example, 270 A100 positions across 75 configurations from 19 domains, and 161 H200 positions across 38 configurations from 20 domains. These are sums of posted configuration sizes, not unique chips.

Physical deployed stock, immediately orderable capacity, firm commitments and announced future GPU capacity are NOT ESTABLISHED. A single underlying machine may support more than one published configuration.

The supplier comparison also changed definition

In the same 10-domain historical price cohort, the June classification had two “not GPU” entries that October evidence classified differently. That is a migration in verification methodology, not evidence of two new suppliers entering the market.

And geography?

None of the 60 strict matched-price rows contains a populated region field. Q2-to-October geographic changes are therefore not measurable from this cohort. A provider’s announced locations also must not be substituted for independently verified facility locations.

06 / MARKET RESEARCH

Five findings worth retaining

An evidence-led view of the GPU service market, without manufactured growth claims.

01 / A large candidate list is not a supplier census

The completed review retained 722 active candidates, but only 339 rows were explicitly labeled GPU service providers. Identity and eligibility need their own denominators.

02 / GPU-model evidence is substantial but uneven

Public-price records identify dozens of domains for H100, A100, H200 and other models. Variant-level transparency differs, and domain counts do not establish inventory.

03 / Prices require the entire contract to make sense

GPU count, sold object, currency, term and billing unit determine whether any two public quotations can be compared.

04 / The matched history is useful because it is narrow

The 60-offer Q2-to-October cohort showed 41 unchanged rates, but does not describe the price direction of the full market.

05 / Some important numbers are not yet measurable

Posted configurations are not fleet capacity. Geographic movement, installed GPUs and broad Q2-to-October market growth remain unestablished.

Methods and sources

[1] ViaBandwidth checkpoint-162 State of GPU Services audit. October evidence refreshed 25-27 October 2026.

[2] Codex read-only comparison export, 2 November 2026: published_pricing_by_model.csv; matched_prices_direction.csv; requested_gpu_capacity.csv; supplier_composition_comparison.csv; publication_readiness.md. Aggregates derived from private official-provider evidence, which is not reproduced here.

[3] ViaBandwidth, State of GPU Cloud, historical July 2026 report: https://viabandwidth.com/reports/state-of-gpu-cloud. July counts have a non-comparable directory denominator.

Editorial scope: aggregate, descriptive and date-bound. Not a live inventory feed, supplier-performance ranking, physical capacity estimate or general market inflation index.