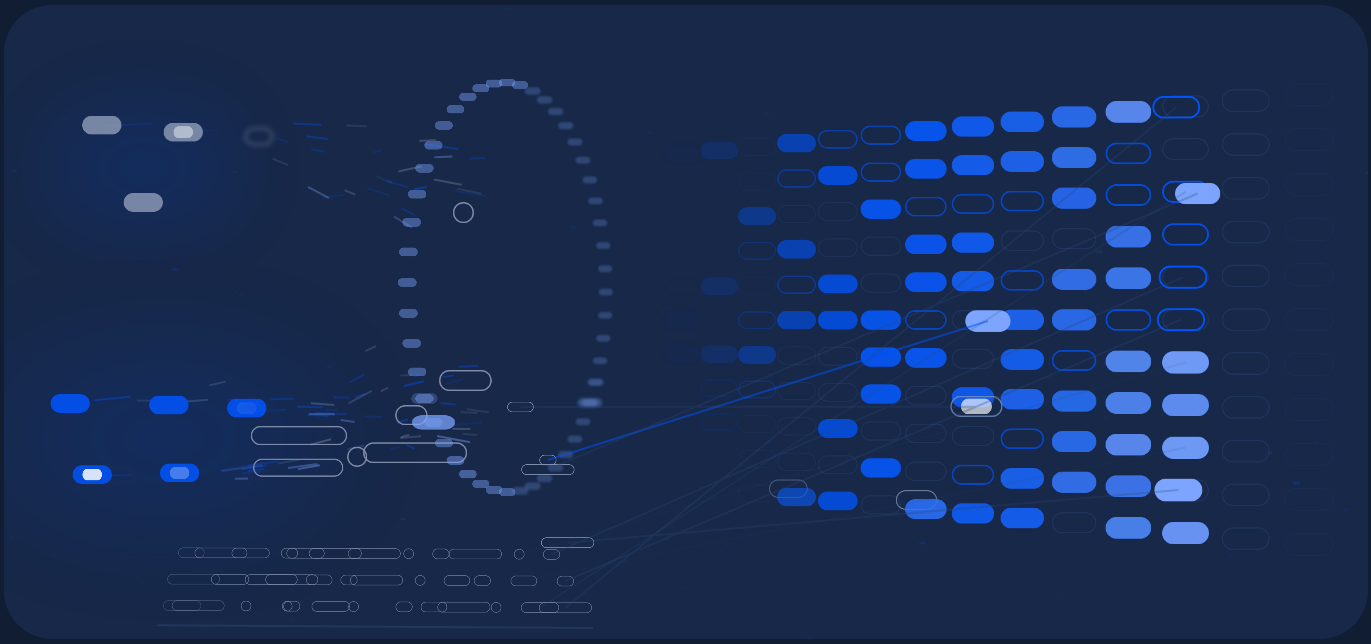
Rillor

EXPLAINER

GPU cloud costs beyond the hourly rate

The hidden costs of GPU cloud from dated price pages: data transfer, storage, IP addresses and support, and the switching charges the EU Data Act ends from 12 January 2027.



Rillor

Contents

01 Summary

02 Background

03 Method

04 Findings

05 Implications

06 Limits of this analysis

07 Sources

Summary

A GPU cloud bill carries line items that the GPU-hour price leaves out, and providers post each one on their own price pages.

- On 8 November 2026, AWS posted internet egress from US East (N. Virginia) at $0.09 per GB for the first paid tier, after a 100 GB monthly allowance shared across its services and regions. Transfer between its Availability Zones cost $0.01 per GB in each direction [1, 2].
- CoreWeave, Lambda, Crusoe and Nebius post no charge for network egress [19, 20, 22, 23]. Nebius still charges $0.015 per GiB for egress traffic from its Standard object storage class, so free network egress does not reach every service [22].
- From 12 January 2027, the EU Data Act bars charges for switching to another provider, data egress included. Egress outside a switch is not a switching charge. When a customer uses several providers at the same time, a provider may still pass on its egress costs without exceeding them [25, 27].
- A 70-billion-parameter checkpoint in MLPerf Storage's rules is 912 GiB, about 979 GB. Copying one from AWS US East to US West costs about $19.59, and sending one to the internet at the first paid tier about $88 (computed for this report from [2, 32]).
- Storage bills run while data sits idle. Ten 70-billion-parameter checkpoints, about 9.8 TB, cost about $134 to $783 a month across the object and block storage rates posted by Nebius, AWS and CoreWeave (computed for this report from [8, 9, 19, 22, 32]).
- AWS, Google Cloud and Azure credit egress for customers who move their data off, through a support request or an exit notice and within set periods, such as 60 days at Azure and 180 days for eligible AWS customers outside the EU [5, 14, 18].
- The UK Competition and Markets Authority (CMA) found in July 2025 that egress fees are a key commercial barrier to switching and multi-cloud, and that they weigh heavily on smaller customers and on those moving large shares of their stored data [29, 30].

Background

What a GPU-hour price covers

A posted GPU-hour price covers one instance or node for one hour: the GPUs, and with them a fixed amount of CPU, host memory and, at some providers, local disk [3, 6, 19, 20, 22]. AWS posts its p5.48xlarge, with eight H100 GPUs, at $55.04 an hour in US East [3, 6], which is $6.88 per GPU-hour (computed for this report). Moving data out, keeping data after the instance stops, a public address and support each carry their own rates, on separate pages [1, 4, 8, 10].

Units differ. AWS and Azure state transfer in GB and TB, and Azure notes that 1 TB is 1,000 GB [1, 16]. Google states network prices per GiB [12]. CoreWeave counts storage in binary gigabytes of 2^30 bytes [19], and Nebius, Crusoe and Lambda state GiB [21, 22, 23]. One GiB is 1.073741824 GB, so the same rate per GiB and per GB differ by about 7% (computed for this report).

Why transfer and storage reach the bill

A 2022 systems paper on bulk transfers found that egress is priced by the volume moved, whatever the speed, with no corresponding charge for data coming in [34]. In its example, a VM sending at 1 Gbps for an hour on AWS at $0.09 per GB ran up $40.50 in egress against $1.50 for the VM [34]. Training data is large: the FineWeb web corpus takes 50,446.9 GB on disk [33], and large models' checkpoints run to hundreds of gigabytes or more [32].

The UK CMA found that egress fees reduce customers' ability and incentive to switch or use several providers [29]. The EU Data Act entered into force on 11 January 2024, applies since 12 September 2025 and phases out switching charges, egress included, by 12 January 2027 [25, 26].

This report is for teams budgeting training or inference on rented GPUs who want the line items beyond the GPU-hour price and the rules on egress fees.

Method

What was examined. The hardware bundled with an instance, internet egress, transfer between zones and regions, storage, public IP addresses, NAT gateways, support, billing increments and tax. Then the exit waivers of AWS, Google Cloud and Azure, the Data Act's rules on switching and egress, and the CMA's findings.

Sources. Price pages and price files of AWS, Google Cloud, Azure, CoreWeave, Lambda, Nebius, Crusoe and Oracle [1, 2, 3, 4, 8, 9, 10, 11, 12, 16, 19, 20, 22, 23, 24], provider documentation and posts [5, 6, 7, 13, 14, 15, 17, 18, 21], the Data Act and Commission documents [25, 26, 27, 28], the CMA [29, 30], and papers and public files for checkpoint sizes, dataset sizes and transfer costs [31, 32, 33, 34, 35].

How. Every page and file was opened on 8 November 2026. Rates are list prices in U.S. dollars as posted that day, each with its region, tier and unit. AWS rates for US East (N. Virginia) come from the price files that fill its pricing pages [2, 3, 8, 9]. Providers are named only as publishers of their own pages. Computed figures are labelled, with the arithmetic shown. A month is 730 hours. AWS does not say whether a TB in its transfer tiers is 1,000 or 1,024 GB, so examples that cross a tier show both.

What is left out. Negotiated discounts, credits, tax amounts and any provider's own cost of egress, which none of the pages opened states. The Data Act and exit terms are described as their texts state them. This report is not legal advice.

Findings

1. What the instance price bundles

Provider and node	Price as posted	vCPUs	Host memory	Local storage, as named	Source
AWS p5.48xlarge, US East, on demand	$55.04 an hour, $6.88 per GPU	192	2 TiB	8 x 3.84 TB NVMe SSD, instance store	[3, 6]
CoreWeave HGX H100, North America, on demand	$49.24 a node-hour, $6.155 per GPU	128	2,048 GB	61.44 TB "Local Storage"	[19]
Lambda 8x NVIDIA HGX H100	$4.59 per GPU-hour, posted to take effect on 16 November 2026	208	1,800 GiB	22 TiB SSD	[20]
Nebius HGX H100, on demand	$4.50 per GPU-hour	16, listed with the per-GPU price	200 GB, listed with the per-GPU price	No local disk listed with the GPU, local SSD at $0.065 per GiB-month	[22]

Table 1. Posted on 8 November 2026. AWS's and CoreWeave's per-GPU prices are node prices divided by eight, computed for this report. Lambda directs users to log in for current pricing until 16 November 2026 [20].

At AWS, local disks are scratch space. An instance store volume "exists only during the lifetime of the instance to which it is attached", and its data survives a reboot. When the instance is stopped, hibernated or terminated, every block is cryptographically erased, and AWS's guide says not to rely on instance store volumes for valuable, long-term data [7]. A checkpoint that must outlive the instance goes to paid storage (Finding 4).

Billing increments differ. AWS bills On-Demand capacity "by the hour or second (minimum of 60 seconds)" from launch until the instance is terminated or stopped [1]. Lambda bills on-demand instances in one-minute increments, from launch and health checks until termination, "regardless if they're actively being used", and bills 1-Click Clusters in weekly increments [21]. An instance idle between jobs costs the same as one at work at both [1, 21]. CoreWeave offers up to 60% off on-demand prices for committed usage [19], and this report keeps to on-demand list prices.

2. Data transfer out to the internet

Provider and route	Free allowance	First paid tier	Next tiers	Unit	Source
AWS, from US East (N. Virginia)	100 GB a month, pooled across services and regions	$0.09 for the next 10 TB	$0.085 next 40 TB, $0.07 next 100 TB, $0.05 above 150 TB	GB	[1, 2]
Google Premium Tier, to North America	0 to 1 GiB	$0.12 to 1,024 GiB	$0.11 to 10,240 GiB, $0.08 above	GiB	[12]
Google Standard Tier	200 GiB a month per account	$0.085 to 10,240 GiB	$0.065 to 153,600 GiB, $0.045 above	GiB	[12]
Azure, Microsoft Premium Global Network, from North America or Europe	100 GB a month	$0.087 for the next 10 TB	$0.083 next 40 TB, $0.07 next 100 TB, $0.05 next 350 TB	GB	[16]

Table 2. Rates in U.S. dollars per unit, posted on 8 November 2026.

AWS's free tier "is global and does not apply separately or individually to AWS Regions" [1]. Premium Tier is Google's default, Standard Tier must be chosen explicitly, and the Premium table above is the page's default table, with rates that depend on the source region [12]. Azure's transit ISP routing option charges $0.08 per GB after the same free 100 GB [16]. Oracle lists the first 10 TB a month of outbound transfer as free [24].

This report prices one transfer of FineWeb, a corpus of more than 18.5 trillion tokens of English web text at 50,446.9 GB on disk [33], sent once to the internet in a month with no other egress.

Route	Arithmetic	Cost
AWS US East, 1 TB = 1,000 GB	100 GB free, then 10,000 GB x $0.09, 40,000 GB x $0.085 and 346.9 GB x $0.07	About $4,324
AWS US East, 1 TB = 1,024 GB	100 GB free, then 10,240 GB x $0.09 and 40,106.9 GB x $0.085	About $4,331
Azure, from North America	100 GB free, then 10,000 GB x $0.087, 40,000 GB x $0.083 and 346.9 GB x $0.07	About $4,214

Route	Arithmetic	Cost
Google Premium Tier, to North America	46,982 GiB: 1 free, 1,023 x $0.12, 9,216 x $0.11, 36,742 x $0.08	About $4,076
Google Standard Tier	46,982 GiB: 200 free, 10,040 x $0.085, 36,742 x $0.065	About $3,242

Table 3. Computed for this report from [1, 2, 12, 16, 33]. The AWS figure is about 10.8% of one p5.48xlarge month of $40,179.20 [3].

Four GPU-focused providers post no network egress charge. CoreWeave lists "Data transfer between CoreWeave and internet" and "Data transfer within CoreWeave" as free, and its storage carries "no ingress, egress, or transfer fees" [19]. Lambda advertises "Transparent pricing with no egress fees" [20]. Crusoe's FAQ says: "At this time, Crusoe Cloud does not charge for network ingress or egress, either within a VPC or to/from the public internet" [23]. Nebius lists "Egress/Ingress traffic" as free [22].

Free network egress does not reach every service. Nebius charges $0.015 per GiB for egress traffic from its Standard object storage class, while its Intelligent and Enhanced classes list egress as free [22]. Reading FineWeb's 46,982 GiB once out of Standard object storage would cost about $705 (computed for this report from [22, 33]).

3. Transfer between zones and regions

AWS charges $0.01 per GB in each direction for data across Availability Zones in one region, a rule its EC2 pricing page sets out for EC2 and the other services it names, and S3 is not among them [1]. Data to or from a public or Elastic IPv4 address in the region is charged the same [1]. Google charges nothing inside one zone on internal IPv4 or any IPv6 address in the same VPC network and $0.01 per GiB to another zone in the region. Its page adds that all traffic to and from external IPv4 addresses leaves the zone [12]. Azure lists transfer within one Availability Zone as free [16].

Route	Rate	Unit	Source
AWS, US East (N. Virginia) to US East (Ohio)	$0.01	GB	[2]
AWS, US East (N. Virginia) to US West (Oregon), Frankfurt, Ireland or Tokyo	$0.02	GB	[2]
Google, within North America or within Europe	$0.02	GiB	[12]
Google, North America to Europe, to Asia, to South America	$0.05, $0.08, $0.14	GiB	[12]

Route	Rate	Unit	Source
Azure, between regions within North America or within Europe	$0.02	GB	[16]
Azure, from North America or Europe to other continents	$0.05	GB	[16]

Table 4. Outbound transfer between regions, posted on 8 November 2026.

Training moves checkpoints across these lines. The MLPerf Storage rules list checkpoint sizes of 105 GB for an 8-billion-parameter model, 912 GB for 70 billion, 5.29 TB for 405 billion and 18 TB for 1 trillion, in binary units despite the GB and TB labels [32]. The 70B checkpoint is 912 GiB, or 979.3 GB (computed for this report).

Move of one 70B checkpoint	Arithmetic	Cost
AWS, US East to US West (Oregon)	979.3 GB x $0.02	$19.59
AWS, US East to the internet, first paid tier	979.3 GB x $0.09	About $88
Google, to another zone in the region	912 GiB x $0.01	$9.12
Google, North America to Europe	912 GiB x $0.05	$45.60

Table 5. Computed for this report from [2, 12, 32]. The AWS internet figure assumes the 100 GB allowance is used up. It equals about 1.6 hours of the p5.48xlarge. A 405B checkpoint, about 5,816 GB, costs about $523 on the same route.

A January 2026 preprint on multi-region spot jobs notes that a training job must move its checkpoint whenever it migrates to a new region, and that egress costs differ by up to 7 times across regions [35]. Moving about 500 GB of checkpoints for a 14-billion-parameter fine-tuning job costs roughly $10 from the US to Europe and about $40 from Asia to the US [35]. In its runs on AWS L4, A10G and A100 instances, egress for cross-region migration made up to 16% of total cost, and for the A100 job the 500 GB checkpoint led the system to switch zones inside one region [35].

4. Storage for datasets and checkpoints

A checkpoint is many times the weights. The ZeRO paper counts 16 bytes per parameter for mixed-precision training with Adam: 2 each for fp16 parameters and gradients, and 4 each for fp32 parameters, momentum and variance [31]. ZeRO counts memory, not files. The MLPerf Storage 70B checkpoint works out to about 14 bytes per parameter (computed for this report from [32]), against 2 bytes for fp16 weights alone [31].

Provider and class	Rate	Unit	Source
Amazon S3 Standard, US East	$0.023 for the first 50 TB, then $0.022 and $0.021	GB-month	[8]
Amazon EBS gp3, US East	$0.08, snapshots $0.05	GB-month	[9]
CoreWeave object storage	Hot $0.06, warm $0.03, cold $0.015	GB-month, binary	[19]
CoreWeave Distributed File Storage	$0.070	GB-month, binary	[19]
Nebius Standard object storage	$0.0147, plus $0.015 per GiB of egress	GiB-month	[22]
Nebius Shared Filesystem	$0.0800	GiB-month	[22]
Crusoe persistent disks, shared disks, object storage	$0.08, $0.07, $0.06	GiB-month	[23]

Table 6. Storage rates posted on 8 November 2026.

Ten 70B checkpoints kept one month	Arithmetic	Cost
Nebius Standard object storage	9,120 GiB x $0.0147	$134.06
Amazon S3 Standard	9,792.5 GB x $0.023	$225.23
CoreWeave object storage, hot	9,120 GiB x $0.06	$547.20
Amazon EBS gp3	9,792.5 GB x $0.08	$783.40

Table 7. Computed for this report from [8, 9, 19, 22, 32]. Keeping FineWeb's 50,446.9 GB in S3 Standard costs about $1,160 a month (computed for this report from [8, 33]).

Requests and provisioning add to the rate. S3 charges $0.005 per 1,000 PUT, COPY, POST or LIST requests and $0.0004 per 1,000 GET requests (computed for this report from [8]). EBS is charged "by the amount of GB you provision per month until you release the storage" [9]. Lambda bills filesystems per GiB used per month, and billing "continues as long as a filesystem exists, even if it's not mounted to an instance" [21].

5. Public IP addresses, NAT gateways and other network items

Provider	Public IPv4 address	One address for a month	Source
AWS	$0.005 an hour, in use or idle	$3.65	[4]
Google Cloud	$0.005 an hour in use on a standard VM, $0.01 an hour for a static address left unused	$3.65 in use, $7.30 unused	[12]
CoreWeave	$4.00 per address a month	$4.00	[19]
Nebius	Free	$0	[22]

Table 8. Monthly amounts computed for this report at 730 hours.

AWS charges NAT gateways per gateway-hour, with partial hours billed as full hours, plus a charge per GB processed "regardless of the traffic's source or destination" and standard data transfer [4]. At the $0.045 rates of AWS's US East (Ohio) example, one gateway costs $32.85 a month before data and $77.85 with 1 TB processed (computed for this report from [4]). CoreWeave lists its NAT gateway as free [19]. AWS Interconnect - multicloud bills each interconnect to another cloud by the hour with "no per-gigabyte data transfer charges" [11].

6. Support plans and tax

AWS Business Support+ costs the greater of $29 a month per account or 9% of monthly charges up to $10,000, 7% from $10,000 to $80,000, 5% to $250,000 and 3% above. Enterprise Support starts at the greater of $5,000 a month or 10% of the first $150,000 of monthly charges [10]. Fees use "each month's gross AWS charges (before any discounts or credits are applied)" [10]. On one p5.48xlarge month of $40,179.20, Business Support+ is 9% x $10,000 + 7% x $30,179.20 = $3,012.54, about 7.5% of the instance (computed for this report from [3, 10]).

AWS prices are "exclusive of applicable taxes and duties, including VAT and applicable sales tax" [1, 9]. Lambda's prices carry "Plus applicable sales tax/VAT/GST" [20], and its charges include sales tax based on the billing location [21].

7. Exit waivers

Provider	How to ask	What is credited	Time allowed	Source
AWS	Contact AWS Support	Credits for the data being migrated, once approved	180 days for eligible customers outside the EU. EU customers switching under the Data Act follow AWS's EU Data Act Addendum	[5]
Google Cloud	Exit Notice, then a Billing Support agent within 14 days	A credit on the final invoice for the transfer of workloads and production data	A 30-day Initiation Period, a Migration Period of 30 days or more, a Completion Notice up to 180 days after it	[14]
Azure	Support request with the subscription ID, start date and data estimate	Invoice-level credit for up to 60 calendar days of egress	60 days from the start date, 180 days for UK billing addresses leaving UK datacenters	[18]

Table 9. Exit terms as posted on 8 November 2026.

AWS's waiver covers moving all data off AWS from all Eligible Accounts, payer accounts in good standing, or all data off one service [5]. It needs no account closure, is open worldwide and follows "the direction set by the European Data Act" [5]. An update of 30 September 2025 dropped the need for more than 100 GB of monthly egress before contacting AWS Support, and one of 8 October 2026 dropped account-level review and gives eligible customers outside the EU 180 days [5]. AWS's EC2 page adds that EU customers may request reduced data transfer rates for eligible use cases under the Data Act [1].

Google made transfer out free for leaving customers globally on 11 January 2024 [13], for those who migrate all workloads and data from the relevant service [14]. Azure asks leaving customers to cancel all subscriptions after egressing their data [18].

Each waiver covers leaving. Ongoing egress to another provider is billed at the rates in Finding 2, apart from the in-parallel offers in Finding 8.

8. The EU Data Act: switching charges and in-parallel use

Article 2(35) defines data egress charges as "data transfer fees charged to customers for extracting their data through the network" from a provider to a different provider or to on-premises infrastructure [25]. Article 2(36) defines switching charges as charges, "other than standard service fees or early termination penalties", for the actions the regulation mandates for switching, "including data egress charges" [25].

Period	What the regulation says	Article
11 January 2024 to 12 January 2027	Providers may impose reduced switching charges, capped at the costs directly linked to the switching process	29(2), 29(3)
From 12 January 2027	Providers shall not impose any switching charges on the customer for the switching process	29(1)
Before and after 12 January 2027	For services used in parallel, providers may impose data egress charges only to pass on egress costs incurred, without exceeding them	34(2)

Table 10. Restated from the Official Journal text [25] and the Commission's FAQ [27].

Recital 89 says standard service fees "are not switching charges" and stay applicable until the contract ends [25]. Recital 99 says in-parallel egress "can be an ongoing activity, in contrast with the one-off egress required as part of the switching process", so providers may keep charging it at cost after three years [25]. Article 29(4) requires providers to tell a prospective customer, before contract, the standard fees, early termination penalties and reduced switching charges that might apply [25].

Article 25 sets a maximum transitional period of 30 calendar days for a switch, a notice period of two months at the outside and a data retrieval period of 30 calendar days or more. Where 30 days is technically unfeasible, an alternative period can run up to seven months [25]. Article 31 takes services mostly custom-built for one customer, and not offered at broad commercial scale, out of Article 29, and takes time-limited non-production versions for testing out of the switching chapter [25].

The Commission's FAQ of 22 January 2026 says providers may not charge for switching, egress included, from 12 January 2027, and may still bill in-parallel egress at cost after that date, because a multi-cloud deployment may imply constant egress [27]. Its fact page says the Act removes switching charges, including data egress charges, from 12 January 2027 [26]. That sentence is about switching, and both documents say they do not represent the Commission's official position [26, 27].

Providers have published in-parallel offers. Google's Data Transfer Essentials, announced on 10 September 2025, carries qualifying multicloud traffic for EU and UK customers at a zero charge, metered separately once a customer opts in, while other traffic stays at existing rates [15]. Azure offers at-cost transfer for in-parallel use by organizations with billing addresses in the EEA, EFTA or the UK, through a support request, refunded on the next bill. It covers internet egress routed via its transit ISP network, and for UK billing addresses moving data from UK datacenters also the Microsoft Premium Global Network. The data must be processed by the same organization at the destination, and Azure says it might change the policy [17].

On 19 November 2025 the Commission proposed Data Act amendments in its Digital Omnibus, whose explanatory memorandum says they maintain the ambition of removing vendor lock-in, particularly switching and egress charges [28].

9. What regulators found about egress fees

The CMA's final decision of 31 July 2025 found that competition in cloud services is not working well [29]. It called egress fees "a key commercial barrier" to switching and multi-cloud, particularly for smaller customers and those with large proportions of stored data, and found that they also reduce suppliers' incentive to compete for rivals' customers [29]. It found committed spend agreements widespread and, in their current form, not harmful to competition [29].

The final report states its estimates as ranges, each standing for a redacted figure [30]. To move all their data, 0-5% to 5-10% of customers would have had to pay egress fees above 5-10% of annual cloud spend per year if they spread the cost over a typical contract, and 10-20% to 20-30% would have had to pay that much in one year if they did not [30]. Smaller customers, and those egressing larger proportions of their data, are likely to pay higher egress fees as a proportion of total cloud spend [30].

The report estimates egress fees at 0-5% of AWS's, Microsoft's and Google's UK cloud revenues [30]. It reports AWS's submission that the effective price of egress, net of discounts, fell by almost 40% between 2018 and 2024, and records that the three providers cited larger free tiers as evidence of price reductions [30]. The CMA's work covers UK cloud customers across all workloads, and none of these figures is specific to GPU instances.

Implications

The line items scale with different things: egress and storage with data moved and kept, support with the whole bill, addresses and gateways with their count. Table 11 sets them against one GPU node for one month.

Line item, one month, AWS US East	Arithmetic	Amount	Share of the instance
p5.48xlarge, 730 hours	730 x $55.04	$40,179.20	100%
FineWeb sent once to the internet	Table 3	About $4,324	10.8%
Business Support+	Finding 6	$3,012.54	7.5%
FineWeb kept in S3 Standard	Finding 4	About $1,160	2.9%
FineWeb read once from EC2 instances in another zone	50,447 GB x ($0.01 out + $0.01 in)	About $1,009	2.5%
Ten 70B checkpoints in S3 Standard	Table 7	$225.23	0.6%
One NAT gateway with 1 TB processed	Finding 5	$77.85	0.2%
One 70B checkpoint copied to US West (Oregon)	Table 5	$19.59	0.05%
One public IPv4 address	730 x $0.005	$3.65	0.01%

Table 11. Computed for this report from [1, 2, 3, 4, 8, 10, 32, 33]. The cross-zone row applies the EC2 rule on [1] to data served from instances, never to reads from an S3 bucket.

Size changes the shares. The same FineWeb egress is 86% of a month on the single-GPU p5.4xlarge at $6.88 an hour, $5,022.40 for 730 hours (computed for this report from [1, 2, 3, 6, 33]). That matches the CMA's estimate that smaller customers pay higher egress fees as a share of total cloud spend [30].

- **Each route has its own meter.** Internet egress, transfer between zones and regions, and NAT processing are billed separately [1, 2, 4, 12, 16], so a budget lists data moved per month by route.

- **Free egress needs a check per service.** Nebius posts free network egress and charges egress on Standard object storage [22], and Crusoe dates its free egress with "At this time" [23].
- **Checkpoints set the size of many moves.** A 70B checkpoint is about 979 GB and a 405B one about 5,816 GB (computed for this report from [32]), and checkpoint egress reached 16% of total cost in one multi-region study [35].
- **Exit waivers cover leaving.** They require moving all data off a provider or service, a request or notice, and a set period [5, 14, 18].
- **The Data Act bars switching charges and caps in-parallel egress at cost.** Standard service fees stay outside switching charges, and in-parallel egress stays chargeable at cost [25, 27].

Limits of this analysis

- Rates are list prices posted on 8 November 2026 for the regions named. Other regions, dates, discounts and commitments differ.
- The source region of Google's default Premium Tier table is not certain from the static page, so the report uses it for traffic to North America [12].
- AWS does not define a TB in its transfer tiers as 1,000 or 1,024 GB, so tiered examples show both.
- Lambda's instance prices take effect on 16 November 2026 [20], and its billing guide's filesystem rate is an example that may not reflect current pricing, so none is used [21].
- The report gives no Azure rate between Availability Zones and no Oracle rate above 10 TB a month, as neither appeared on the pages opened [16, 24].
- No page opened states a provider's reduced or at-cost Data Act egress rate. The AWS EU Data Act Addendum was not reviewed.
- Egress terms can change. Crusoe states its free egress "At this time", and Azure says it might change its egress credit policy [17, 23].
- The worked examples are illustrations, and no source reviewed gives a typical share of a GPU bill for these items.
- The CMA's figures are redacted ranges for UK customers across all workloads [30], the 16% share comes from one preprint's experiments [35], and the 1 Gbps example dates from 2022 [34].
- The Digital Omnibus is cited only as proposed on 19 November 2025 [28].

Sources

1. Amazon EC2 On-Demand Pricing, data transfer section. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/ec2/pricing/on-demand/

2. AWS data transfer price file (datatransfer.json), published 3 November 2026. Amazon Web Services, viewed 8 November 2026. https://b0.p.awsstatic.com/pricing/2.0/meteredUnitMaps/datatransfer/USD/current/datatransfer.json

3. Amazon EC2 on-demand price file, US East (N. Virginia), Linux, published 7 November 2026. Amazon Web Services, viewed 8 November 2026. https://b0.p.awsstatic.com/pricing/2.0/meteredUnitMaps/ec2/USD/current/ec2-ondemand-without-sec-sel/US%20East%20%28N.%20Virginia%29/Linux/index.json

4. Amazon VPC pricing. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/vpc/pricing/

5. Free data transfer out to internet when moving out of AWS. Sébastien Stormacq, AWS News Blog, Amazon Web Services, 5 March 2024, updated 30 September 2025 and 8 October 2026, viewed 8 November 2026. https://aws.amazon.com/blogs/aws/free-data-transfer-out-to-internet-when-moving-out-of-aws/

6. Amazon EC2 P5 Instances. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/ec2/instance-types/p5/

7. Data persistence for Amazon EC2 instance store volumes. Amazon EC2 User Guide, Amazon Web Services, viewed 8 November 2026. https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/instance-store-lifetime.html

8. Amazon S3 pricing, with the S3 Standard price file, published 4 November 2026. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/s3/pricing/ https://b0.p.awsstatic.com/pricing/2.0/meteredUnitMaps/s3/USD/current/s3-standard.json

9. Amazon EBS pricing, with the EBS price file, published 7 November 2026. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/ebs/pricing/ https://b0.p.awsstatic.com/pricing/2.0/meteredUnitMaps/ec2/USD/current/ebs.json

10. AWS Support Plan Pricing. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/premiumsupport/pricing/

11. AWS Interconnect - multicloud Pricing. Amazon Web Services, viewed 8 November 2026. https://aws.amazon.com/interconnect/multicloud/pricing/

12. Network pricing. Google Cloud, viewed 8 November 2026. https://cloud.google.com/vpc/network-pricing

13. Cloud switching just got easier: Removing data transfer fees when moving off Google Cloud. Amit Zavery, Google Cloud Blog, 11 January 2024, viewed 8 November 2026. https://cloud.google.com/blog/products/networking/eliminating-data-transfer-fees-when-migrating-off-google-cloud

14. Google Cloud Exit free data transfer request. Google Cloud, viewed 8 November 2026. https://cloud.google.com/exit-cloud

15. Introducing no-cost, multicloud Data Transfer Essentials for EU and U.K. customers. Jeanette Manfra, Google Cloud Blog, 10 September 2025, viewed 8 November 2026. https://cloud.google.com/blog/products/networking/new-for-the-uk-and-eu-no-cost-multicloud-data-transfer-essentials

16. Bandwidth pricing. Microsoft Azure, viewed 8 November 2026. https://azure.microsoft.com/en-us/pricing/details/bandwidth/

17. What data transfer fees are applied when data is transferred between Azure and an external endpoint? Microsoft Learn, Microsoft Cost Management documentation, last updated 5 May 2026, viewed 8 November 2026. https://learn.microsoft.com/en-us/azure/cost-management-billing/manage/data-transfer-fees

18. Cancel and delete your Azure subscription, section "What data transfer fees are applied when moving all data off Azure?" Microsoft Learn, last updated 7 June 2026, viewed 8 November 2026. https://learn.microsoft.com/en-us/azure/cost-management-billing/manage/cancel-azure-subscription

19. CoreWeave Cloud Pricing. CoreWeave, viewed 8 November 2026. https://www.coreweave.com/pricing

20. Rent NVIDIA GPUs on demand: H100, H200, and B200 (instance pricing table). Lambda, viewed 8 November 2026. https://lambda.com/instances

21. Billing overview. Lambda Cloud documentation, Lambda, viewed 8 November 2026. https://docs.lambda.ai/public-cloud/billing/

22. NVIDIA GPU Pricing, Nebius AI Cloud. Nebius, viewed 8 November 2026. https://nebius.com/prices

23. GPU pricing for AI compute and inference, Crusoe Cloud pricing. Crusoe, viewed 8 November 2026. https://www.crusoe.ai/cloud/pricing

24. OCI Price List, Networking section. Oracle, viewed 8 November 2026. https://www.oracle.com/cloud/networking/pricing/

25. Regulation (EU) 2023/2854 (Data Act), Official Journal text. European Parliament and Council of the European Union, OJ L, 2023/2854, 22 December 2023 (adopted 13 December 2023), viewed 8 November 2026. https://eur-lex.europa.eu/eli/reg/2023/2854/oj/eng

26. Data Act explained. European Commission, Shaping Europe's digital future, last updated 15 December 2025, viewed 8 November 2026. https://digital-strategy.ec.europa.eu/en/factpages/data-act-explained

27. Frequently Asked Questions: Data Act, Version 1.4. European Commission, 22 January 2026, viewed 8 November 2026. https://ec.europa.eu/newsroom/dae/redirection/document/108144

28. Proposal for a Regulation (Digital Omnibus), COM(2025) 837 final, procedure 2025/0360(COD). European Commission, 19 November 2025, viewed 8 November 2026. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52025PC0837

29. Cloud services market investigation: Summary of final decision. UK Competition and Markets Authority, 31 July 2025, viewed 8 November 2026. https://assets.publishing.service.gov.uk/media/688b20e6ff8c05468cb7b120/summary_of_final_decision.pdf

30. Cloud infrastructure services: Final decision report. UK Competition and Markets Authority, 31 July 2025, viewed 8 November 2026. https://assets.publishing.service.gov.uk/media/688b8891fdde2b8f73469544/final_decision_report.pdf

31. ZeRO: Memory Optimizations Toward Training Trillion Parameter Models. Samyam Rajbhandari, Jeff Rasley, Olatunji Ruwase and Yuxiong He, arXiv:1910.02054, version 3, 13 May 2020 (first version 4 October 2019), viewed 8 November 2026. https://arxiv.org/abs/1910.02054

32. MLPerf Storage Rules (Rules.md, rules edition 3.0). MLCommons, GitHub repository mlcommons/storage, viewed 8 November 2026. https://github.com/mlcommons/storage/blob/main/Rules.md

33. FineWeb dataset card (HuggingFaceFW/fineweb). Hugging Face, last modified 11 July 2025, viewed 8 November 2026. https://huggingface.co/datasets/HuggingFaceFW/fineweb

34. Skyplane: Optimizing Transfer Cost and Throughput Using Cloud-Aware Overlays. Paras Jain, Sam Kumar, Sarah Wooders, Shishir G. Patil, Joseph E. Gonzalez and Ion Stoica, arXiv:2210.07259, 13 October 2022, viewed 8 November 2026. https://arxiv.org/abs/2210.07259

35. SkyNomad: On Using Multi-Region Spot Instances to Minimize AI Batch Job Cost. Zhifei Li, Tian Xia, Ziming Mao and others, arXiv:2601.06520, 10 January 2026, viewed 8 November 2026. https://arxiv.org/abs/2601.06520

About this report

Published by Rillor on 9 November 2026. Superintelligence (SI) tools assist research and drafting, and every figure cites its source. Rillor builds and secures SI systems, supplies the data they learn from and delivers the infrastructure they run on. We apply SI to financial and compute markets.

Cite as: Rillor. GPU cloud costs beyond the hourly rate. 9 November 2026. https://rillor.com/insights/gpu-cloud-costs-beyond-the-hourly-rate

Notices

Rillor is not a registered investment adviser or commodity trading advisor and does not provide investment or trading advice. Research, forecasts, data and software described on this site are for research and engineering use. Nothing here is an offer or recommendation to buy or sell any security, commodity interest or digital asset. Past or simulated results do not indicate future results.

NVIDIA, AMD, Ceph, Kubernetes and related product names are trademarks of their owners and are used only to identify products. No affiliation or endorsement is implied.

Work with Rillor

Send the GPU classes, regions and pricing types you need, and how often you need them. Contact us at rillor.com/contact?type=index or contact@rillor.com. More on GPU and neocloud pricing: rillor.com/markets/gpu-pricing