Appendix L: Egress fees - background

L.1 This appendix sets out some background information for our analysis of egress fees in chapter 6. As set out in that chapter, we have focussed our analysis on egress fees incurred during the switching and/or multi-cloud process. This appendix provides background information on data transfer fees (eg egress and internal data transfer fees) in general.

Types of interconnections

L.2 There are different ways through which customers can transfer data into, out of and within a cloud provider's infrastructure. The fee that a customer incurs depends on the type of route used.[1]

L.3 Table L.1 sets out the different routes through which customers can transfer data in, out and within a cloud provider's infrastructure.

Table L.1: Data transfer routes

Transfer route	*Internal/External*	*Fees*
Intra-virtual network transfer	Internal	X
Inter-virtual network transfer	Internal	✓
Internet transfer	External	✓
Direct connection	External	✓
Physical transfer	External	✓
Content delivery networks	External	✓

Source: CMA analysis of publicly available information (Microsoft website, Pricing – Bandwidth \| Microsoft Azure; AWS website, EC2 On-Demand Instance Pricing – Amazon Web Services; Google website, All networking pricing \| Virtual Private Cloud \| Google Cloud, all accessed on 14 November 2023).

L.4 Table L.1 shows that there are two types of internal cloud transfers:

(a) intra-virtual network transfers are transfers of data within a virtual network, eg when customers transfer data from a database to a virtual machine hosted within the same virtual network to run analytics workloads. Such transfers are free of charge within an availability zone, otherwise a fee applies.

[1] For completeness, we note that there are also services that can be added onto transfer routes and generally do not change the underlying egress fees. For example, Network Address Translation ('NAT') gateways add security features to network connections over the internet. Customers buying NAT gateways incur a running cost that is listed separately, but the level of egress fee paid for the transfer remains the same with or without the service (see Google website, Pricing; Microsoft website, Azure NAT Gateway - Pricing - Microsoft; AWS website, NAT gateways - Amazon Virtual Private Cloud, accessed on 14 November 2023.

(b) inter-virtual network transfers are transfers of data within a cloud that is not within a virtual network. For instance, transfer of a customer's data between storage in US and UK regions, or transfer of data between two companies' virtual networks within the same cloud. Internal transfer fees apply to such transfers.

L.5 Table L.1 also sets out the four main routes we have identified that customers can use to transfer data out of a cloud provider's infrastructure:

(a) Internet transfer is the transfer of data from a public cloud over the internet. This may be to another resource using a public IP address in the same cloud, to the customer's on-premises network, or to end users. Egress fees are charged on such transfers. Internet transfers can be standard or premium internet transfers.

(i) standard internet transfers typically occur when the data exits the cloud provider's network close to the source and is then carried over the public internet by an internet service provider ('ISP') to the destination.

(ii) premium internet transfers occur when the data is carried over the cloud provider's network backbone until close to the destination. The ISP will still carry the data for the final stretch over the public internet.

(b) Direct connection[2] is a way to transfer data between multiple public clouds or on-premises networks without using the internet. Data is routed over the network backbone of the cloud provider to co-location centres, typically internet exchange points ('IXPs') run by third parties. We note that there is a distinction between direct connections to on-premises and to other clouds:

(i) to on-premises, direct connections can be dedicated (on a private physical cable) or via a connectivity partner (on shared private cables);

(ii) to other clouds, direct connections can be via on-premises, via IXP or via cloud provider (ie Google's recent Cross-Cloud Interconnect).

(c) Direct connections are generally used in cases of high volume or latency-sensitive traffic. Egress fees are charged on such transfers, but at a lower rate than for internet transfer. Hourly running fees are also charged.

(d) Physical device transfer occurs when data is uploaded and downloaded from a physical device which is shipped back and forth between the customer's on-premises and public cloud data centre. Such transfers are generally used

[2] We note that the terminology differs across providers. For instance, AWS refer to direct connections as 'Direct Connect', Microsoft as 'ExpressRoute' and Google as 'Dedicated Interconnect'.

for the migration of data between on-premises and the cloud where the network connection is poor, or the volume of data is very large.

(e) Content delivery networks (CDNs) are a specialised type of transfer where data is transferred first to a geographically distributed network of data centres and then to end users over the internet. The goal is to provide high availability and performance by distributing the service close to the end user. Cloudflare is an example of a content delivery network. These networks are typically used to distribute a relatively small number of files to a large number of users where latency is important (eg streaming media).[3] In niche cases they might be used to switch or facilitate hybrid clouds.

L.6 Based on the characteristics and common use cases for each of these types of transfers, as described above, we consider that (i) internet transfer and (ii) direct connections are the most relevant external transfer routes for the purpose of our assessment of egress fees.

Pricing

L.7 Data transfer fees are generally volume-based and region-specific.[4] The total charge incurred when a customer egresses data also depends on a number of factors, including:

(a) where the data is stored. Cloud providers set data transfer fees differently across their regions.

(b) where the data is going. For example, in some cases data transfer 'within' a region is free, while data transfer to an endpoint not associated with the original cloud provider, over the public internet, is not free.[5]

(c) whether the cloud provider offers free tiers and/or rates that reduce with increased data transfer volumes (eg Microsoft, AWS, Google, Oracle and IBM all offer a free tier for certain services).[6]

[3] 'What is a CDN? – Cloudflare', accessed on 15 October 2024.
[4] Typically, the cloud infrastructure services available from Microsoft, AWS, and Google allow the customer to select a geographic 'region' in which the customers' data is stored and/or processed (for example, see 'Regions and zones – Google Cloud', accessed on 15 October 2024). We note however that with some cloud infrastructure services the customer does not have such a choice and the cloud provider may distribute processing or storage across multiple regions simultaneously (see 'Global services - AWS', accessed on 15 October 2024).
[5] 'Amazon EC2 On-Demand Pricing– AWS', accessed on 15 October 2024.
[6] Microsoft website, Pricing – Bandwidth | Microsoft Azure; AWS website, Amazon S3 Simple Storage Service Pricing - Amazon Web Services and EC2 On-Demand Instance Pricing – Amazon Web Services; Google website, All networking pricing - Virtual Private Cloud - Google Cloud; Oracle website, Cloud Networking Pricing - Oracle, all accessed on 12 March 2024; [✂] response to the CMA's information request [✂].

(d) whether a CDN is used. For transfers to end users via the public internet, the level of the data transfer fees also depends on whether the provider's own or a third party CDN is used.

(e) whether the cloud provider has different routing offerings – Google and Microsoft offer two egress tiers – one where data is transferred through their own network and one where data is routed via transit (internet service provider, 'ISP') networks. Microsoft describes ISP network routing as a 'cost-optimized' option – it avoids some of the internal transfer fees that would be incurred by routing through their backbone network.[7] AWS also offers a similar 'premium' internet transfer using its accelerator services. This is structured differently to the Google and Microsoft offerings but serves a similar purpose.[8]

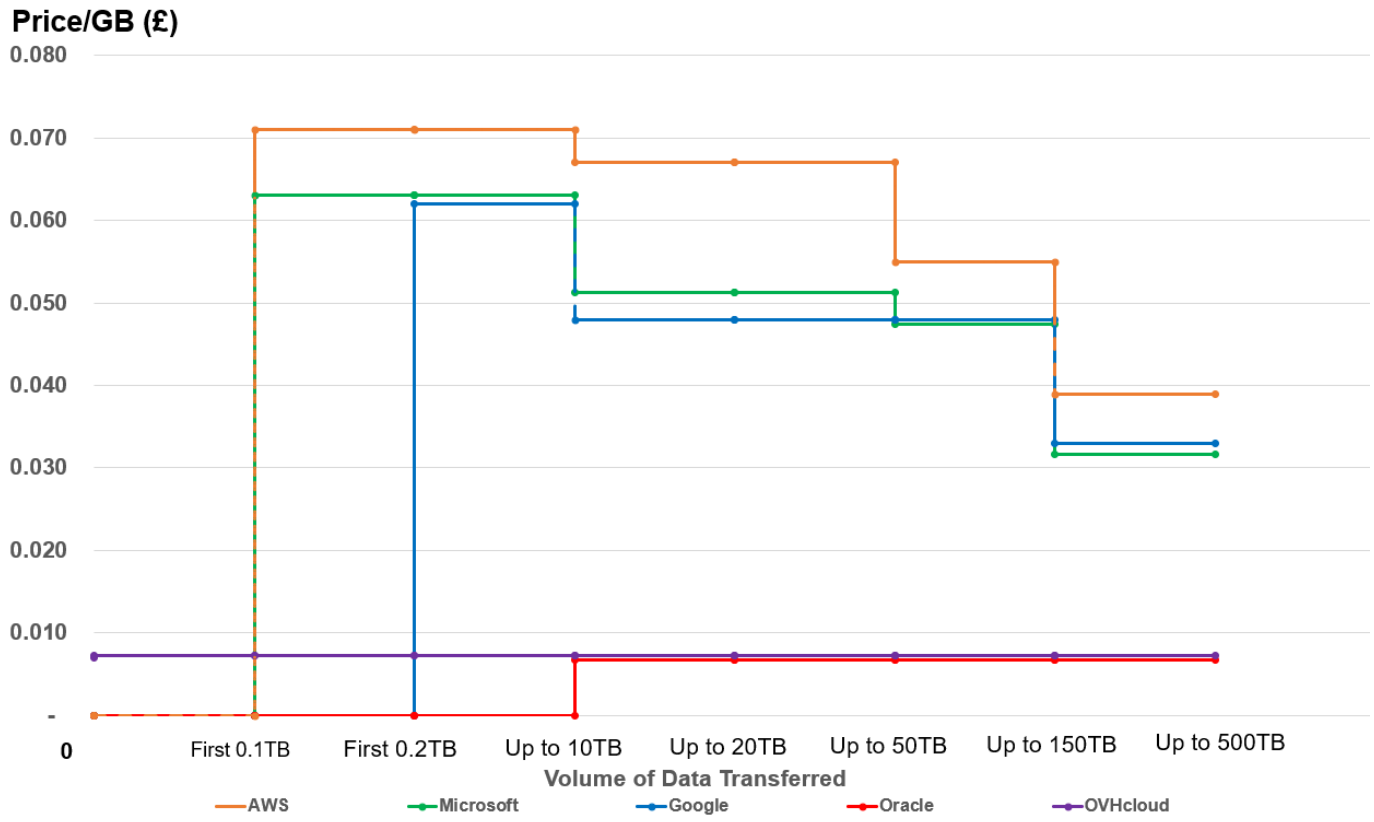
(f) which cloud infrastructure service(s) the customer is using to send the data. For example, the cost of data transfer can be different using AWS EC2 (a virtual machine IaaS service) compared to AWS Lambda (a serverless PaaS service).[9]

Egress fees pricing

L.8 Figure L.1 below shows different providers' egress fee list prices[10] for data transfer out via the public internet[11] as of June 2024.

[7] Google website, Network Service Tiers overview - Google Cloud; Microsoft website, Routing preference in Azure - Azure Virtual Network, accessed on 14 March 2024.
[8] See AWS website, S3 Transfer Acceleration; AWS website, AWS Global Accelerator features - Amazon Web Services, accessed on 19 March 2024.
[9] See 'AWS Data Transfer Charges for Server and Serverless Architectures, accessed on 15 October 2024.
[10] Note: All prices have been converted to GBP from USD where applicable, based on the GBP/USD exchange rate as of 19 March 2024. Google's prices have been converted from per-GiB to per-GB.
[11] 'Standard tier' egress via the internet, or equivalent.

Figure L.1: Egress fee list prices for data transfer out via the public internet (June 2024, GBP per GB)



Sources: Microsoft website, Pricing – Bandwidth | Microsoft Azure; AWS website, Amazon S3 Simple Storage Service Pricing - Amazon Web Services and EC2 On-Demand Instance Pricing – Amazon Web Services; Google website, All networking pricing | Virtual Private Cloud | Google Cloud; Oracle website, Cloud Networking Pricing | Oracle United Kingdom,; OVHcloud's website, Price list: a comparison of our Public Cloud offers | OVHcloud UK, all accessed on 19 June 2024.

L.9 Figure L.1 shows that list prices for AWS, Microsoft and Google may include:

(a) a higher marginal price for low volumes of data transfer outside of the free tier; and

(b) declining marginal prices for higher volumes of data transfer.

L.10 AWS, Microsoft, Google, Oracle and IBM, for certain services, provide a 'free tier' whereby customers are able to egress a certain volume of data to the internet without paying any fees.[12]

L.11 Finally, Figure L.1 shows that Oracle and OVHcloud also have a flatter fee structure – outside of the Oracle free tier, they charge one price irrespective of the amount of data transferred. Oracle and OVHcloud also charge materially lower fees than other providers included in the comparison.

[12] Microsoft website, Pricing – Bandwidth - Microsoft Azure; AWS website, Amazon S3 Simple Storage Service Pricing - Amazon Web Services and EC2 On-Demand Instance Pricing – Amazon Web Services; Google website, All networking pricing; Oracle website, Cloud Networking Pricing - Oracle, all accessed on 12 March 2024; [✂] response to the CMA's information request [✂].

L.12 When comparing list prices across providers, we must exercise a degree of caution. Cloud providers' prices may differ because of their underlying costs or quality differences in the network offerings, or because providers charge for network usage differently (eg by incorporating charges into the prices of other services). Cloud providers' costs are analysed in detail in appendix Q.

L.13 Figure L.1 is based on list prices, though some customers may not be subject to list prices for egress fees via the public internet as depending on the cloud provider they may be able to negotiate:[13]

(a) private discounts on egress fees specifically; or

(b) cross-product discounts (ie committed spend agreements) that could be applied to egress fees.

L.14 We have reviewed internal documents from the largest cloud providers relating to the negotiation of recent contracts with UK customers for the provision of public cloud infrastructure services where the providers agreed to offer discounts on egress fees. In the sample of documents reviewed, the customers had all negotiated discounts across products rather than discounts specifically related to egress fees (we consider discounts across products in chapter 8).[14]

L.15 Microsoft submitted that the effective price paid for egress fees over the internet has declined. Microsoft also submitted that, on average, real prices for both 'metered and unmetered' egress plans have fallen since 2018.[15] We note, however, that even if egress fees were shown to be declining over time, the current level of egress fees may still constitute a barrier to switching and multi-cloud, which is the focus of our theory of harm.

Assets used to provide data transfer services

L.16 Cloud providers have identified similar assets used in providing data transfer services – data centres and network equipment (including servers, routers and switches), and connection cables (short-haul and/or long-haul).

L.17 The key network assets identified are summarised below. We note that some of these may also be used in providing other cloud services such as storage.

Table L.2: Summary of key network assets used by cloud providers in transferring data

Networking assets

[13] Responses to the CMA's information requests [✂].
[14] [✂].
[15] Microsoft's submission to the CMA [✂].

Servers	Servers are used for receiving and transmitting data. Servers can be origin servers or edge servers – origin servers store the original version of the content, while edge servers store cached versions of the content.
Routers	Routers connect computers and other devices to the Internet. A router acts as a dispatcher, choosing the best route for information to travel. Routing is the ability to forward IP packets – a package of data with an internet protocol (IP) address – from one network to another. Boundary routers forward ingress traffic to internal routers and switches.
Switches	Internal switches send traffic to the cloud computing and storage resources for processing. The return egress traffic that contains customer data is traversed via internal switches to the boundary routers to go to the Internet or to the backbone network. Ethernet switches connect cabled devices, like computers, WiFi access points and IoT devices, and servers, in an Ethernet Local Area Network so they can communicate with each other and to the Internet. Once a device is connected to a port, the Ethernet switch manages the flow of data between the device and other devices, applications, data cloud services, and the Internet.
Optic fibre cables: Backbone	Backbone, also know as long-haul fibre, are optic fibre cables providing connectivity between regions and Points of Presence (POP) around the world. They may be used for inter-region connectivity, providing private direct connection connectivity to clouds, and connectivity to IP transit and peering across regions.
Optic fibre cables: Metro	Metro fibre, also referred to as short-haul fibre, are optic fibre cables which provide connectivity over shorter distances, eg region to region.

Source: Responses to the CMA's information requests [✂].

Cloud providers' connectivity arrangements

L.18 There are three main approaches for cloud providers to establish connectivity between workloads running across different infrastructures:

(a) over the public internet, using internet service providers (ISPs);

(b) through connectivity partner networks, using network service provider (NSP) services; and,

(c) direct connectivity at a colocation data centre facility.

L.19 Internet peering allows cloud providers to exchange traffic directly with other networks and can be done as private peering, using private network interconnections (PNIs), and public peering, through internet exchanges.

L.20 'IP Transit', via connection to the public internet, allows cloud providers to exchange traffic indirectly with other networks. AWS said this may be used because the other network is not otherwise available for peering or may be for redundancy purposes. IP Transit providers, by connecting with each other to send and receive data, offer access to every network on the internet (as opposed to peers, who offer access only to their own networks and those of their customers).[16]

[16] [✂] response to the CMA's information request [✂].

L.21 Google said private peering (which could be paid or settlement free) is the primary way it exchanges traffic with other cloud provider networks, major ISPs and CDN providers and is purpose built to exchange high volumes of traffic reliably.[17]

L.22 Some cloud providers said each participating member is responsible for the costs of connecting to third party networks, the public internet and/or public peering.[18] [✂].[19] [✂].[20] Microsoft said any connection supporting connectivity to another cloud provider's network is shared by both parties equally.[21] Google said costs incurred in peering between cloud providers are typically shared, but may be negotiated between providers when establishing the connection.[22]

L.23 Oracle said its public network is connected to [✂] public networks either directly or via internet exchanges (IXs). Oracle has a commercial agreement with Microsoft for connecting Oracle Cloud Infrastructure (OCI) and Azure. The interconnect allows customers to build a private interconnection between their OCI and Azure environments.[23] In both cases, each party pays their own hardware and operational costs (eg data centre space/ power/ cabling, labour).[24] Oracle has also recently announced an interconnection partnership with Google.[25]

L.24 [✂].[26]

[17] [✂] response to the CMA's information request [✂].
[18] Responses to the CMA's information requests [✂].
[19] [✂] response to the CMA's information request [✂].
[20] [✂] response to the CMA's information request [✂].
[21] Microsoft's response to the CMA's information request [✂].
[22] Google's response to the CMA's information request [✂].
[23] Oracle interconnect for Azure, accessed on 20 May 2024.
[24] [✂] response to the CMA's information request [✂].
[25] Oracle and Google Cloud Announce a Groundbreaking Multicloud Partnership, accessed on 30 September 2024.
[26] [✂] response to the CMA's information request [✂].