

AWS Customer Carbon Footprint Methodology

October 2025

Table of Contents

Figures	2
Overview	3
1 Scope	1
1.1 Underlying Standards.....	1
1.2 Definitions	1
2 System Boundary	4
2.1 Attributable vs Non-Attributable Emissions.....	4
2.2 Scope 1 Emissions	5
2.2.1 Inclusions.....	5
2.2.2 Exclusions.....	5
2.3 Scope 2 Emissions	5
2.3.1 Inclusions.....	5
2.3.2 Exclusions.....	6
2.4 Scope 3 Emissions	6
2.4.1 Inclusions.....	6
2.4.2 Exclusions.....	6
3 Input Data	7
3.1 Scope 1 Emissions	7
3.2 Scope 2 Emissions	7
3.2.1 Overview of Data Sources and Transformations.....	7
3.2.2 Estimated Load	8
3.2.3 Grid Mix and Emission Factors	8
3.2.4 Estimate and Allocate Energy Attribute Certificates.....	8
3.3 Scope 3 Emissions	9
3.3.1 Fuel and Energy Related Activities (FERA).....	9
3.3.2 Embodied Carbon in IT Hardware	11
3.3.3 Embodied Carbon in Buildings & Equipment	20
3.3.4 Amortization of Embodied Carbon	27
4 Allocation	31
4.1 Cluster to Server Racks	32
4.2 Server Racks to Cloud Services.....	33
4.3 Cloud Services to Customer Accounts.....	35
4.3.1 Foundational Services.....	36
4.3.2 Non-Foundational Services	37
4.3.3 Aggregating All Together	37
4.3.4 Allocation of Residual Carbon Emissions.....	38
Bibliography	39

Figures

Figure 1 – System boundary in corporate carbon footprint reporting according to the GHG Protocol (Source: GHG Protocol)	4
Figure 2 – End-to-end process flow diagram for IT hardware data.....	11
Figure 3 – Generation of IT hardware component-level carbon data	12
Figure 4 – Aggregation of IT hardware carbon data from component-level to server-rack level.....	18
Figure 5 – End-to-end process flow diagram for building and equipment data.....	22
Figure 6 – System boundary for building and equipment carbon data generation, per EN 15978 and ISO 21930 (Source: National Research Council of Canada).....	23
Figure 7 – Process steps and data sets for generating embodied carbon emissions of data center buildings and equipment.....	25
Figure 8 – Amortization of capital emissions	27
Figure 9 – Allocation of amortized capital emissions	30
Figure 10 – Conceptual workflow for the allocation of carbon emissions in the customer carbon footprint model ...	32
Figure 11 – Allocation of cluster-level emissions to the underlying server racks.....	32
Figure 12 – Allocation of resources to cloud services.....	33
Figure 13 – Allocation of emissions to customer accounts	36

Overview

This document outlines the methodology that governs AWS' carbon footprint products (model version 3.0). AWS may reasonably update this document upon material changes to the model logic and calculations.

The document is structured as follows:

1. **Scope** – drivers and scope, along with a look into underlying standards governing the customer carbon footprint model.
2. **System Boundary** – insight into attributable vs non-attributable emissions, inclusions and exclusions.
3. **Input Data** – sources of data used in the customer carbon footprint model.
4. **Allocation** – how AWS carbon emissions are allocated to customer accounts.

1 Scope

Through its sustainability products, such as the Customer Carbon Footprint Tool (CCFT) and Data Exports, AWS publishes the estimated carbon emissions associated with customers' AWS service usage. Customers often use a wide array of AWS services across multiple Regions making it complex to track and apportion carbon emissions based on workloads. The customer carbon footprint model aims to address such complexity, using the best available data and methodology to date. This document outlines the logic and calculations for version 3.0 of the model.

The customer carbon footprint model quantifies customer-specific greenhouse gas (GHG) emissions associated with the use of AWS cloud services. It covers the full range of cloud products, from Amazon EC2 virtual and physical servers, to Amazon Storage Service range (S3), Amazon Elastic Block Store (EBS), Lambda, Sagemaker, and the rest of the AWS services. The model includes Scope 1, Scope 2, and Scope 3 data.

The results are provided as Metric Tons of Carbon Dioxide equivalents ($MTCO_2e$), accounted on a monthly basis for the previous 38 months, broken down by AWS Region, and by service. Historical customer carbon footprint data will be recast at least annually upon the completion of third-party assurance of Amazon' footprint or when significant methodology updates are introduced.

1.1 Underlying Standards

The customer carbon footprint model methodology is informed by the [Greenhouse Gas \(GHG\) Protocol's](#) emissions grouping into Scope 1 (direct emissions), Scope 2 (indirect emissions from purchased energy) and Scope 3 (other value chain indirect emissions). See section 2 'System Boundary' for more information on included emissions and exclusions. Along with the GHG Protocol, the methodology adopted in the model is based on the guiding principles outlined in the following standards/protocols:

- [GHG Protocol Corporate Accounting and Reporting standard](#) [1] and [ISO 14064-1](#) [2]
- [GHG Protocol Product Life Cycle Accounting and Reporting Standard](#) [3] and associated [Information and Communication Technology \(ICT\) sector guidance](#) [4]
- [ISO 14040](#) [5] and [ISO 14044](#) [6] for Life Cycle Assessment (LCA)

1.2 Definitions

Terminology	Description
Activity data	Activity data are quantitative measures of a level of activity that results in GHG Emissions. It is a measure that models what is taking place during a given period of time (e.g., amount of diesel consumed, gigabytes of storage used, dollars spent, etc.).
Allocation	The action of apportioning emissions of a given activity to a product/service/other activity based on an established weighting factor: e.g., usage-based ('physical allocation') or equivalent revenue-based ('economic allocation').
Amortization	The process of normalizing non-operational emissions that occur across the lifecycle of cloud services (outside of the use phase), e.g., the embodied carbon within server racks ($kgCO_2e$ /server rack) is amortized over the service life of the rack and processed in the allocation model as monthly emissions per server rack ($\frac{kgCO_2e}{rack\ asset\ x\ month}$) during each reporting month in order to enable allocation of the entirety of embodied carbon from rack assets to customers over their service life.
AWS carbon footprint products	Products that offer carbon footprint estimates at the customer level, such as the Customer Carbon Footprint Tool (CCFT) and Data Export.
CCFT	Customer Carbon Footprint Tool.
Cluster-level carbon footprint	The carbon footprint in Metric Tons of Carbon Dioxide equivalents ($MTCO_2e$) for a particular AWS cluster. An AWS cluster can be an AWS Region (e.g. Frankfurt, Ohio), or an AWS CloudFront Edge Cluster (e.g. CF in North America, CF in South America).

Customer accounts (CA)	<p>Customer accounts used to access cloud services. These accounts can be associated to different customers:</p> <ol style="list-style-type: none"> 1. AWS internal teams - accounts owned by AWS service teams such as EC2, Lambda, etc. 2. Amazon businesses - accounts owned by teams in other Amazon business units beyond AWS. 3. External - accounts used by external customers.
Customer-segment-level carbon footprint	The carbon footprint in MTCO ₂ e for a particular group of customers that are using a specific service.
EACs	<p>Energy Attribute Certificates are certificates that serve as an instrument to represent exclusive rights to the environmental, social, and other non-power attributes of renewable electricity generation. EACs demonstrate that procured electricity originated from a renewable resource and can be utilized to reduce Scope 2 emissions under the market-based method (MBM). EACs generally equate to one certificate for each 1 MWh of renewable energy produced by a renewable energy project. EACs are called Renewable Energy Certificates (RECs) in the US but have other names in other geographies (e.g., Guarantees of Origin (GOs) in the EU). Nuclear Emission-Free Energy Certificates (EFECs) are included in the list of EACs accounted.</p> <p>To claim the emissions reductions from renewable energy purchases, EACs need to be generated and allocated within the same GHG defined boundary.</p>
Edge sites	Types of data centers situated close to end customers, to enable ultra-low latency. Points of Presence (PoP), Cloudfront, and Transit sites are considered Edge.
Emission factors	An emission factor is a coefficient that quantifies the emissions (e.g., greenhouse gases) released from a given activity (e.g., production of 1 kWh of electricity).
Equivalent revenue	The equivalent revenue is the weighting factor leveraged when applying ‘economic allocation’. It is a standardized measure of usage-based revenue that excludes pricing variation factors such as discounts, and other adjustments. It reflects service usage for each customer account, allowing for a consistent comparison across customer segments, time periods and regions.
Foundational service (FS)	AWS services that have dedicated server racks in AWS data centers.
GHG	Greenhouse Gases. The array of gases in the atmosphere that absorb some of the heat radiated by our planet, resulting in a greenhouse effect, thus contributing to global warming. In alignment with GHG Protocol and Life Cycle Assessment (LCA) standards, the effect of greenhouse gases is expressed as carbon dioxide equivalent (‘CO ₂ e’), namely the equivalent amount of CO ₂ e emissions with the same global warming potential as the considered greenhouse gases emissions.
LBM	Location-based method - GHG Protocol method used in Scope 2 carbon emissions accounting that reflects the average emissions intensity of grids on which energy consumption occurs.
LCA	Life Cycle Assessment (LCA) is a standardized process to quantify the environmental impact of a product (or a service) throughout its life cycle.
MBM	Market-based method - GHG Protocol method used in Scope 2 carbon emissions accounting that reflects supplier-specific emissions intensity after accounting for EACs.
MEC	Month End Close: An operational monthly process that AWS takes to release updated carbon footprint data.
MTCO₂e	Metric Tons of Carbon Dioxide equivalent.
Non-foundational service (nFS)	AWS services with no dedicated server racks, which rely on foundational services to provide functionality to end customers.
PCR	Product Category Rules – set of standardized rules that provide the instructions for how the life-cycle assessment (LCA) should be conducted.
PPA	Power Purchase Agreement is a long-term contract between an electricity generator and an electricity customer to supply a certain amount of electricity at a pre-negotiated price (either fixed or indexed). PPA terms can vary in duration.
PPD	Planned Power Draw. The expected average power draw during peak usage conditions for a given server rack. The primary purpose of PPD values is to drive planning decisions to reserve sufficient power for each server rack deployed in AWS’s data centers.

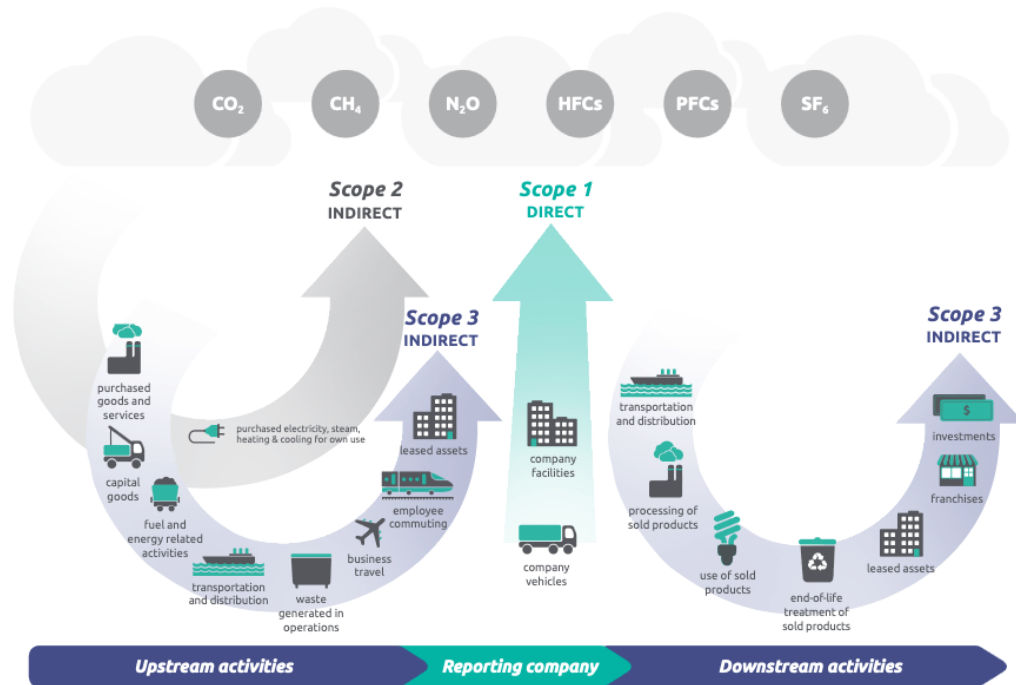
Recast	The recalculation of historical customer emissions estimates using updated methodology and/or inputs.
Residual grid mix	The emissions rate left after energy attribute certificates (e.g. EACs), direct contracts such as Power Purchase Agreements, supplier-specific emission rates are removed from the system.
Service-level carbon footprint	The carbon footprint in MTCO ₂ e for a particular AWS service. This can be a footprint for a foundational or a non-foundational AWS service in any given cluster.

2 System Boundary

The system boundary defines what activities and related emissions are accounted for in the customer carbon footprint model calculations. The model is based on the GHG Protocol's classification of emissions, which breaks down a company's emissions into three 'scopes'.

- **Scope 1** emissions are direct emissions from owned or controlled sources.
- **Scope 2** emissions are indirect emissions from the production of purchased energy.
- **Scope 3** emissions are all indirect emissions (not included in scope 2) that occur in the value chain of the reporting company, including both upstream and downstream emissions (e.g., manufacturing of hardware, end-of-life emissions, etc.).

Figure 1 – System boundary in corporate carbon footprint reporting according to the GHG Protocol
(Source: GHG Protocol)



2.1 Attributable vs Non-Attributable Emissions

'Attributable emissions' are those emissions arising from activities that directly support the provision of cloud services for AWS customers – e.g. production of electricity consumed in data centers, fuel combustion from generators, etc. Conversely, any services, materials, or energy flows not directly connected to AWS cloud services because they do not become, make, or directly carry the product/service through its lifecycle are 'non-attributable emissions'. This classification governs the list of emissions included in and excluded from the customer carbon footprint model.

AWS established guidelines for the classification and inclusion of attributable emissions:

1. Follow the guidelines outlined in the GHG Protocol and ISO standards (see section 1.1) and determine if the activity generating emissions directly supports the provision of cloud services for AWS.
2. Prioritize inclusion of direct attributable emissions from owned and controlled sources (Scope 1) and indirect emissions from the production of purchased energy (Scope 2).
3. Prioritize inclusion of attributable indirect emission in AWS's value chain (Scope 3) based on significance.

In version 3.0 of the customer carbon footprint model, AWS includes:

- Scope 1 – direct emissions from AWS operations
- Scope 2 – indirect emissions from purchased energy

- Scope 3 – indirect emission in AWS’s value chain
 - Fuel & Energy Related Activities (referred to as Scope 3 ‘FERA’ as per the GHG Protocol).

Scope 1, Scope 2 and Scope 3 FERA are collectively referred to as ‘operational emissions’ throughout the document. For more information about Scope 1 and 2 inventory best practices refer to [EPA Scope 1 and Scope 2 Inventory Guidance](#) [7].

- Embodied carbon of IT hardware (i.e., server racks)
- Embodied carbon of data center buildings
- Embodied carbon of non-IT equipment (e.g., generators, air handling units, etc.)
- Emissions from upstream transport and distribution (e.g., emission arising from transporting server racks to the data center sites).

These emissions are referred to as ‘capital emissions’ throughout the document.

The sections below explain the breakdown of emissions included in and excluded from the customer carbon footprint model, for each GHG Protocol’s scope.

2.2 Scope 1 Emissions

Scope 1 emissions are direct emissions from sources owned or controlled by an organization – e.g., direct carbon emissions from fuel combustion tied to operations in AWS data centers.

2.2.1 Inclusions

The customer carbon footprint model includes emissions from fuel combustion in emergency back-up generators, and emissions from use of refrigerants and natural gas from AWS owned or controlled facilities – including where AWS has operational control on the server racks deployed (such as in ‘colo’ data centers) – that support cloud services. The model also includes emissions from certain edge sites (e.g., CloudFront emissions are included).

2.2.2 Exclusions

The customer carbon footprint model excludes:

- Emissions associated with fuel combustion in fleet vehicles. These emissions are deemed non-attributable.
- Emissions associated with operations in facilities beyond data centers, such as warehouses, offices, or sites run in customer facilities (Amazon Cloud Extension, Embedded Points of Presence, Outpost sites). These emissions are deemed non-attributable.

2.3 Scope 2 Emissions

Scope 2 emissions are indirect GHG emissions from the generation of acquired and consumed electricity, steam, heat, or cooling. These emissions are considered an indirect emissions source, because the emissions are a consequence of activities of the reporting organization but actually occur at sources owned or controlled by another organization.

2.3.1 Inclusions

The customer carbon footprint model accounts for Scope 2 emissions from AWS owned or controlled facilities – including where AWS has operational control on the server racks deployed (such as in ‘colo’ data centers) – that support cloud services, as well as certain edge sites (e.g., CloudFront emissions are included), using both the location-based method (LBM) and the market-based method (MBM).

2.3.2 Exclusions

The customer carbon footprint model excludes any Scope 2 emissions associated with operations in facilities beyond data centers, such as AWS warehouses, offices, or sites run in customer facilities (same as in 2.2.2). These emissions are deemed non-attributable.

2.4 Scope 3 Emissions

Scope 3 emissions are indirect GHG emissions that stem from activities in the reporting organization's value chain, both upstream and downstream, but are not directly controlled by the organization. Such emissions are a consequence of the organization's activities but occur from sources owned or controlled by other entities.

2.4.1 Inclusions

The customer carbon footprint model accounts for:

- Emissions from Fuel & Energy Related Activities (referred to as Scope 3 emission category 'FERA' as per the GHG Protocol). These include the following emissions that are associated with fuel and energy used in the facilities pertaining to the customer carbon footprint model's system boundaries (see Scope 1 and 2 sections):
 - Upstream emissions of purchased fuels – i.e., extraction, production, and transportation of fuel used in emergency back-up generators, and for heating/cooling in AWS's owned or controlled data center facilities.
 - Upstream emissions of purchased electricity – i.e., extraction, production, and transportation of fuels consumed in the generation of electricity used in AWS's owned or controlled data center facilities.
 - Emissions from transmission and distribution (T&D) losses – i.e., generation of electricity that is lost in T&D to AWS's owned or controlled data center facilities.
- Embodied carbon of IT hardware – 'cradle-to-gate' emissions of AWS's server racks deployed in AWS's owned or controlled data center facilities. Cradle-to-gate emission include activities spanning from the extraction of raw materials and the manufacturing of server racks' components (e.g., CPUs, SSDs, motherboards, GPUs, RAM modules), the assembly process at AWS's integrator sites, where the server components are assembled into server racks and tested, and the transportation of server racks to AWS's data center facilities.
- Embodied carbon of data center buildings – 'cradle-to-gate' emissions of AWS's owned or controlled data center buildings. These include the outer 'shell' of the data center facility, as well as the data center rooms and the associated mechanical, electrical and plumbing (MEP) system. Cradle-to-gate emissions include activities spanning from the extraction of raw materials and manufacturing of building components, to the transportation to AWS's data center sites, and all the emission arising from the construction.
- Embodied carbon of non-IT equipment – 'cradle-to-gate' emissions of AWS's non-IT equipment deployed in AWS's owned or controlled data center facilities. Non-IT equipment refers to all the auxiliary equipment, beyond server racks, that is installed in AWS's owned or controlled data center facilities. The list includes equipment such as air handling units, generators, power distribution cabinets, HVAC, etc. Cradle-to-gate emission include activities spanning from the extraction of raw materials and the manufacturing of such equipment, to the transportation to AWS's data centers.

2.4.2 Exclusions

The customer carbon footprint model excludes any embodied carbon associated with facilities beyond data centers, such as AWS warehouses, offices, or sites run in customer facilities (same as in 2.1.2). These emissions are deemed non-attributable. The Scope 3 emissions associated with the end-of-life treatment of IT hardware and non-IT equipment, and the end-of-life of data center buildings are not currently included in the model version 3.0.

3 Input Data

The customer carbon footprint model calculates different categories of emissions and subsequently allocates them to customers based on their use of cloud services.

Scope 1, 2, and Scope 3 FERA emissions are referred to as '**operational emissions**' as these are directly linked to the consumption of electricity, fuels, and refrigerants in each data center. Such emissions are calculated at the **cluster and month level** and subsequently allocated to customers based on their use of cloud services.

Scope 3 embodied carbon of IT hardware, buildings, and equipment are referred to as '**capital emissions**' as these emissions are generated at a point in time during the lifetime of a given asset (e.g., server rack). Any AWS customer who uses the asset over its lifetime – through cloud services – shares a portion of the burden. Therefore, similar to financial amortization of CapEx, the Scope 3 embodied carbon emissions are amortized – i.e., determining the monthly equivalent emissions based on the service life of the asset. Such emissions are hence calculated at the **asset and month level** and then allocated to customers.

This section outlines the sources of data, related assumptions, and transformations that occur upstream of the customer carbon footprint model to define Scope 1, 2 and 3 carbon emissions for each AWS cluster. The allocation methodology is detailed separately in Section 4.

3.1 Scope 1 Emissions

Amazon generates and assures Scope 1 activity data for its annual footprint every year. See [Amazon's carbon footprint methodology](#) [8]. To bridge the gap between Amazon's annual reporting and customer carbon footprint model's monthly cadence, AWS uses unassured primary Scope 1 activity data to determine monthly emissions for the current month. Some of the activity data might not be available at the time of publishing the monthly report, therefore translating in an underestimation of Scope 1 emissions¹. We update our estimates when recasting, to align Scope 1 emissions reported in the customer carbon footprint model with the assured data.

Note that the sites included in the customer carbon footprint accounting differ from the boundary used for the corporate carbon footprint as some emissions sources are not directly attributable to providing a cloud service to customers (e.g., direct emissions associated with warehouses, see section 2.2.2).

3.2 Scope 2 Emissions

Similar to Scope 1, the quantification of Scope 2 emissions associated with each AWS cluster aligns with [Amazon's carbon footprint methodology](#) [8]. In line with Amazon's approach, we prioritize quality of data at the time of publishing customer carbon footprint data, falling back to other sources (e.g., estimated load) when the primary source of data (e.g., actual load from power invoices) is not reasonably available. We correct our estimates when recasting, to align Scope 2 emissions reported in the customer carbon footprint model with the assured data.

3.2.1 Overview of Data Sources and Transformations

3.2.1.1 Location Based Method

The steps for estimating LBM Scope 2 emissions are:

1. Estimate Load (MWh) at cluster-month level. See 3.2.2 for details.
2. Collect LBM emission factors (ef) ($\frac{MTCO_2e}{MWh}$) associated with the generation of electricity, for all clusters. See 3.2.3 for details.
3. Calculate LBM Scope 2 emissions ($MTCO_2e$) as:

¹ Scope 1 data is prone to month-over-month fluctuations due to factors such as increased backup generator usage for outages relating to extreme weather events.

$$E_{Scope\ 2,LBM} = Load * ef_{Generation,LBM}$$

Eq 1

3.2.1.2 Market Based Method

The steps for estimating MBM Scope 2 Emissions are:

1. Estimate Load (*MWh*) at cluster-month level. See 3.2.2 for details.
2. Allocate EACs at a cluster-month level based on:
 - a. Estimated Load.
 - b. MBM grid mix for all clusters (i.e. % renewable energy in the grid). See 3.2.3 for details.
 - c. Estimated Energy Attribute Certificates EACs (*MWh*) at a GHG Protocol's boundary and month level. See 3.2.4 for details.
3. Collect MBM emission factors (*ef*) ($\frac{MTCO_2e}{MWh}$) for all clusters. See 3.2.3 for details.
4. Calculate MBM Scope 2 emissions (*MTCO₂e*) as:

$$E_{Scope\ 2,MBM} = [(Load - Allocated\ EACs) * ef_{Generation,MBM}] + (Allocated\ EACs * ef_{EACs})$$

Eq 2

3.2.2 Estimated Load

The primary activity data for LBM and MBM Scope 2 Emissions is energy consumption (i.e. 'load'). When estimating load, the first step is to decide which facilities (i.e. sites) should be included (see section 2.1.1). Note that the system boundary differs from the corporate carbon footprint boundary as we exclude emissions sources not directly related to providing cloud services to customers (such as electricity consumption in offices).

The load data is estimated from utility power invoices, historical data – e.g., deployed power – and forecasted data – e.g., planned server racks and YoY growth rate estimates. The 'composition' of the load data changes over time, based on how timely the data sources (e.g., power invoices) are available for the sites in scope. We update the estimated load when recasting, to align it with assured data used in Amazon's carbon footprint.

3.2.3 Grid Mix and Emission Factors

As outlined earlier, there are two types of electricity emission factors:

- Location-based (LBM) emission factors measure the average emissions intensity of the grids where the electricity consumption occurs.
- Market-based (MBM) emission factors measure the emissions intensity based on the electricity that organizations have chosen to purchase. The customer carbon footprint model follows the Scope 2 Quality Criteria outlined in the [GHG Protocol Scope 2 Guidance](#) [9], which defines how emission factors should be derived based on the underlying market-based contractual instrument, and the hierarchy for the emission factor selection. Following such hierarchy, the MBM emission factors used in the model derive from energy attribute certificates (such as Renewable Energy Certificates, Guarantees of Origin, etc.), direct contracts (such as Power Purchase Agreements) where electricity attribute certificates do not exist or are not required for a usage claim, supplier specific emission rates, and other default emission factors representing the untracked or unclaimed energy and emissions (termed "residual mix") if a company does not have other contractual information that meets the GHGP's Scope 2 quality criteria. Finally, when it is not possible to define specific residual mix emission factors, grid-average emission factors are used. Refer to the [GHG Protocol Scope 2 Guidance](#) [9] for additional details.

AWS seeks to use the latest available emission factors. The list of emission factors is third-party reviewed every year upon generation of carbon data for Amazon's corporate carbon footprint. See [Amazon's carbon footprint methodology](#) [8].

3.2.4 Estimate and Allocate Energy Attribute Certificates

AWS allocates EACs to clusters to calculate MBM Scope 2 emissions. Due to the timing lag between EAC creation, delivery, and retirement, we estimate EACs using the below logic:

1. Calculate monthly EACs from expected annual EAC volume (provided by carbon-free energy project developer).

2. Calculate the expected carbon-free energy using estimated load, MBM grid mix, and procured EAC volume.
3. Determine remaining EAC volume needed (unprocured EAC) by comparing the carbon-free energy to AWS objectives.
4. Add up procured and unprocured EAC volume to obtain the total EACs.

The allocation model used to calculate MBM Scope 2 monthly estimates aligns with the model used to calculate MBM Scope 2 emissions for Amazon’s corporate carbon footprint that is assured every year.

3.3 Scope 3 Emissions

The customer carbon footprint model calculates Scope 3 emissions from the sources outlined in section 2.4. Section 3.3 outlines the sources of data and the transformations that occur upstream of the allocation model to calculate such emissions. The allocation methodology is detailed separately in Section 4.

3.3.1 Fuel and Energy Related Activities (FERA)

This section provides an overview of the estimation of Scope 3 FERA Emissions at a cluster-month level. These estimates are a key input in the allocation model, which attributes the cluster-month level emissions to customers. FERA emissions include the following sources that are associated with fuel and energy used in the facilities pertaining to the customer carbon footprint model’s system boundaries, namely: upstream emissions of purchased fuels; upstream emissions of purchased electricity; emissions from transmission and distribution (T&D) losses.

3.3.1.1 Upstream Emissions of Purchased Fuels

Fuel activity data is collected for each cluster in scope as explained in section 3.1, and emission factors associated with the extraction, production, and transportation of fuels are applied to calculate the upstream emissions of purchased fuels.

3.3.1.2 Upstream Emissions of Purchased Electricity

3.3.1.2.1 Location Based Method

A key input is the load (*MWh*) for each cluster in scope in the customer carbon footprint model, covered in detail in section 3.2.2. The calculation steps are as follows:

1. Collect:
 - i. Load (*MWh*) at cluster-month level.
 - ii. LBM upstream emission factors (*ef*) ($\frac{MTCO_2e}{MWh}$) for all clusters in scope.
2. Calculate LBM upstream emissions (*MTCO₂e*) as:

$$E_{Upstream,LBM} = Load * ef_{Upstream,LBM} \tag{Eq 3}$$

3.3.1.2.2 Market Based Method

The steps for estimating upstream MBM emissions of purchased electricity are:

1. Collect:
 - i. Load (*MWh*) at cluster-month level.
 - ii. Allocated EACs (*MWh*) at a cluster-month level from Scope 2 calculations.
 - iii. MBM upstream emissions factors (*ef*) ($\frac{MTCO_2e}{MWh}$).
2. Calculate MBM upstream emissions (*MTCO₂e*) as:

$$E_{Upstream,MBM} = [Load - Allocated EACs] * ef_{Upstream,MBM} \tag{Eq 4}$$

3.3.1.3 T&D Loss Emissions

3.3.1.3.1 Location Based Method

The calculation steps are as follows:

1. Collect:
 - i. Load (*MWh*) at cluster-month level.
 - ii. LBM T&D loss emission factors (*ef*) ($\frac{MTCO_2e}{MWh}$) for all clusters in scope.
2. Calculate LBM T&D loss emissions (*MTCO₂e*) as:

$$E_{T\&D\ Loss,LBM} = Load * ef_{T\&D\ Loss,LBM} \quad \text{Eq 5}$$

3.3.1.3.2 Market Based Method

The steps for estimating T&D loss MBM emissions of purchased electricity are:

1. Collect
 - i. Load (*MWh*) at cluster-month level.
 - ii. Allocated FERA EACs (*MWh*).
 - iii. LBM T&D loss, upstream, and generation emission factors (*ef*) ($\frac{MTCO_2e}{MWh}$).
2. Calculate T&D loss ratio as:

$$T\&D\ Loss\ Ratio = \frac{ef_{T\&D\ Loss,LBM}}{ef_{Generation,LBM} + ef_{Upstream,LBM} + ef_{T\&D\ Loss,LBM}} \quad \text{Eq 6}$$

3. Calculate T&D Loss Load (*MWh*) as:

$$T\&D\ Loss\ Load = \frac{Load}{(1 - T\&D\ Loss\ Ratio)} - Load \quad \text{Eq 7}$$

4. Calculate MBM T&D loss emissions (*MTCO₂e*) as:

$$E_{T\&D\ Loss,MBM} = (T\&D\ Loss\ Load - Allocated\ FERA\ EACs) * (ef_{Generation,LBM} + ef_{Upstream,LBM}) \quad \text{Eq 8}$$

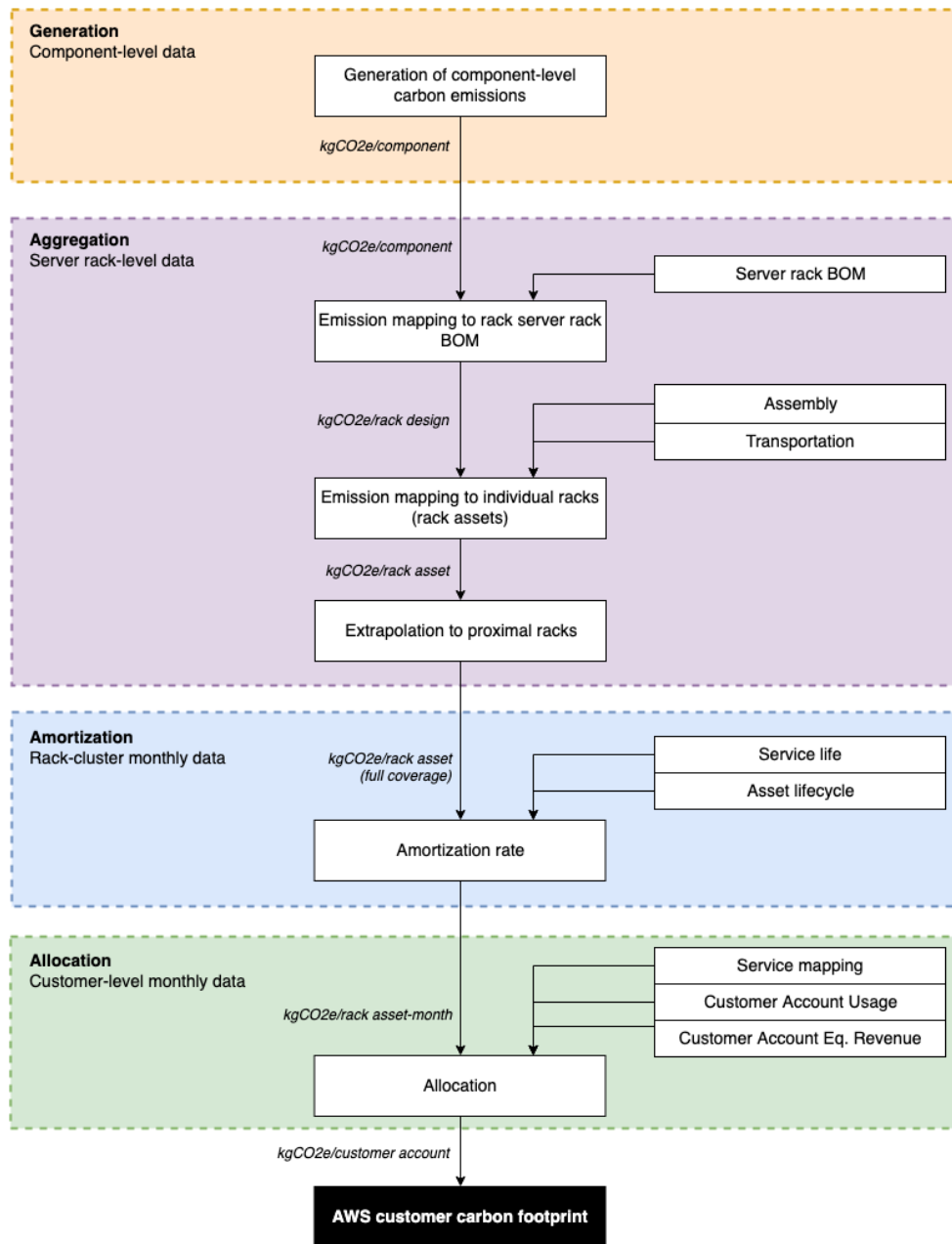
3.3.2 Embodied Carbon in IT Hardware

We refer to the embodied carbon of IT hardware as the ‘cradle-to-gate’ emissions of AWS’s server racks deployed in AWS’s owned or controlled data center facilities. Cradle-to-gate emission include activities spanning from the extraction of raw materials required for the manufacturing of server racks’ components (e.g., CPUs, SSDs, motherboards, GPUs, RAM modules) to the transportation of assembled server racks to AWS’s data center facilities. The scope includes all server racks – including networking – racks deployed in AWS owned or controlled facilities.

3.3.2.1 Overview

This section provides an overview of the framework adopted, from the generation of carbon data to the allocation to AWS customers. The end-to-end process flow diagram is reported in **Figure 2**.

Figure 2 – End-to-end process flow diagram for IT hardware data



Each step is summarized below and described in detail in the respective section of this document.

Generation: Component-Level Carbon Data – The process starts with the generation of component-level cradle-to-gate carbon data obtained via Life Cycle Assessment (LCA), with focus on high-priority server components deployed in the AWS fleet. AWS prioritizes LCAs of server components deployed in the AWS fleet based on volume and expected impact to maximize coverage of total carbon footprint,

and to also inform strategic carbon reduction initiatives across AWS. For other components, and whenever detailed Bill of Materials (BOMs) or key input data is not available, emissions are estimated using parametric models or Economic Input-Output LCA.

Aggregation: Server Rack-Level Carbon Data – The server rack-level embodied carbon emissions are estimated in this step. The model leverages information on the composition of the server racks from their bill of material (BOMs), maps it to the manufacturing carbon data associated with the underlying components from the previous steps, and integrates it with the emissions associated with assembly and transportation of server racks to AWS data centers. Carbon emissions associated with server racks with incomplete data are estimated by analyzing comparable racks with known carbon footprints. This step is necessary to enable full coverage, yielding a carbon data point for each given server rack asset in the AWS fleet, and hence complete coverage in the customer carbon footprint model.

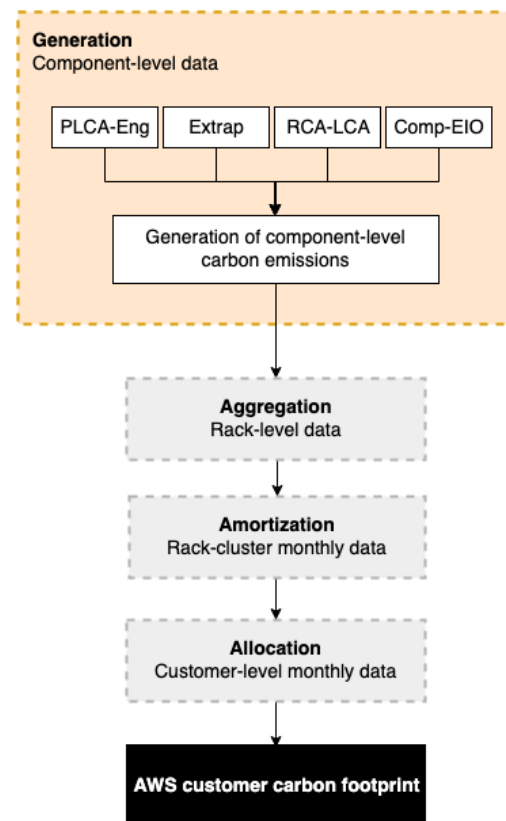
Amortization: Rack-Cluster Monthly Data – This step amortizes server rack hardware embodied carbon emissions across their useful life, quantifying the equivalent embodied carbon emissions to allocate to AWS customers in each reporting month. It determines the service life of each rack asset deployed in AWS's data centers and then amortizes embodied carbon emissions to yield rack-cluster level monthly emissions so that the entirety of the embodied carbon of each rack asset is allocated to their users over their service life, accounting for edge cases like early or late rack retirements.

Allocation: Customer-Level Carbon Data – In this step the amortized emissions are fed into the allocation model and allocated to AWS services and customers' accounts.

3.3.2.2 Generation of Carbon Data

In this section we describe the process of generating carbon data for server components of the AWS fleet. The process involves four different pathways to estimate the GHG emissions associated with components. The choice of pathway is based on the materiality of a component to the overall AWS footprint, as well as data availability. The methodology employs a hierarchical approach using four distinct calculation pathways in descending order of accuracy and detail: Process based Life Cycle Assessment with Engineering Attributes (PLCA-Eng), Extrapolation, Representative Category Average Life Cycle Assessment (RCA-LCA), and Economic Input-Output Life Cycle Assessment (Comp-EIO). The process is represented in Figure 3 below:

Figure 3 – Generation of IT hardware component-level carbon data



3.3.2.2.1 Scope and System Boundary

This LCA methodology estimates life cycle GHG emissions associated with IT hardware in data centers. It focuses on the cradle-to-gate emissions of IT hardware. The lifecycle phases considered span from the extraction of raw materials to the manufacturing of components, the server racks assembly process at AWS's integrator sites where the components – printed circuit boards assemblies (PCBAs), CPUs, memory, cables, network equipment, power supplies, etc. – are assembled into server racks, and finally, the transportation of server racks to AWS's data centers. The LCA methodology aligns with the international standards mentioned in section 1.1.

For each activity pertaining to the system boundary of the LCAs, we apply relevant global warming potential (GWP) factors corresponding to the characterization factors outlined in the Intergovernmental Panel on Climate Change (IPCC)'s Sixth Assessment Report (AR6).² For some aggregated data sources, AR6 values may not be available, and in these cases, we will default to the latest available assessment method and/or modify values when feasible.

3.3.2.2.2 Input Data

AWS leverages granular IT hardware bill-of-material (BOM) data from internal repositories and product lifecycle management systems. For each IT hardware component, AWS collects engineering, manufacturing, and technological (EMT) attributes data. The EMT attributes refer to physical dimensions, mass, manufacturing processes, manufacturing energy intensity, country of origin, and other relevant technical specifications. AWS also utilizes component-level cost data. The IT hardware BOM and the corresponding EMT and cost attributes are the main activity data for the IT hardware LCA. For emission factors, a range of data sources are used based on inventory data availability – primary data provided by suppliers/manufacturers, US EPA's USEEIO, ecoinvent [10], GaBi [11], imec.netzero [12] (Interuniversity Microelectronics Center), and academic/independent studies.

3.3.2.2.3 Methodology Overview

To estimate Scope 3 emissions for IT hardware, AWS adopts a flexible bottom-up LCA framework. This approach builds upon emissions inventories for individual hardware components and sub-component parts, utilizing primary data wherever possible.

The methodology combines process-based LCA and economic input-output LCA (EIO-LCA), which is referred to as a "component-level hybrid LCA approach." This approach helps navigate the ambiguity and data availability challenges often encountered in Scope 3 emissions estimation.

The component-level hybrid LCA is implemented through a "waterfall" model with four methodological pathways:

- **Process-based LCA using manufacturer's data (PLCA-Eng):** Whenever possible, we collect data and employ a process-based LCA approach. We prioritize the use of primary data, like manufacturing energy intensity, used directly as inputs in LCA. When primary data is not available, AWS collects engineering and technical attributes for specific high emission-intensity components to model their carbon footprint using process-based models.
- **Interpolation/extrapolation using process-based LCA data points (Extrap.):** In this pathway we use PLCA-Eng data points, LCA modelling techniques and machine learning tools to predict the LCA results (in terms of kgCO₂e per component) as a function of the key input attribute parameters. This approach is applicable to certain component types, whose technological features can be relatively easily characterized by a few predictors, and whose emissions are believed to make up a significant portion of the overall AWS hardware emissions. The component categories currently included are CPU, GPU, SSD, HDD, Motherboard, RAM, Dongle, and Chassis.
- **Representative Category Average LCA (RCA-LCA):** In this pathway a representative emission factor from LCA databases and the component measured/estimated mass are used to calculate the emissions. This pathway is preferred for long-tail of component categories – with low frequency in the fleet – where PLCA-Eng, Extrap or cost are not available.
- **Component-level EIO-LCA (Comp-EIO-LCA):** If none of the three methodological options listed above are feasible, AWS then utilizes a component-level EIO-LCA approach. This is based on sector-level emissions factors (in kgCO₂e per \$) and the component unit cost (\$).

² [IPCC Sixth Assessment Report.](#)

3.3.2.2.4 Process-based LCA using Engineering, Manufacturing and Technical (EMT) attributes (PLCA-ENG)

This pathway starts with a bill of materials (BOM) from AWS Product Lifecycle Management (PLM) system. For each given part, the main material or electric/electronic component type, manufacturing processes, and mass (or other physical attribute) are specified. For manufacturing and assembly processes, whenever primary data is not available, industry average estimates are used to determine yield loss, energy consumption, water use, and direct greenhouse gas (GHG) emission in production. For a component with n sub-components the cradle-to-gate emissions are calculated as follows:

$$E_{Component} = \sum_{n=1}^N E_{part_n} \quad \text{Eq 9}$$

The following sub-sections expand on the emission calculations for each type.

Materials: A material may be composed of one or more constituent materials (j). The associated impact is equal to the sum of the mass of the constituent materials (m_j), multiplied by an emissions per mass material emission factor (ef_j), divided by a loss factor that accounts for waste generated in making the constituent material L_j :

$$E_{material} = \sum_{j=1}^J \frac{m_j * ef_j}{L_j} \quad \text{Eq 10}$$

Manufacturing Process: The manufacturing impact accounts for the emission-generating activities associated with transforming input materials into a finished good. Within each manufacturing process, there may involve primary material inputs, energy, water, and direct GHG emissions in the facility. Therefore, the process impact for a component is equal to the emissions from electricity (E_{elect}), natural gas (E_{gas}), water (E_{water}), and direct GHG emissions ($E_{ghg,direct}$).

$$E_{process} = E_{elect} + E_{gas} + E_{water} + E_{ghg,direct} \quad \text{Eq 11}$$

Integrated Circuits (IC): The model integrates data from imec.netzero model, which covers the most used process technologies, with legacy LCA datasets for older nodes (N45+), ensuring comprehensive coverage across different semiconductor generations. Die size emerges as the most critical and traceable parameter because it directly correlates with fabrication energy consumption and fluorinated gas emissions, which are the primary contributors to an IC's carbon footprint. When direct measurements aren't available, the model employs a die-to-package ratio system derived from TechInsights [13] database analysis, providing reasonable estimates based on package type and dimensions. The model implements a hierarchical decision tree that first considers technology node information, matching to exact process nodes when possible and defaulting to more conservative (smaller node) estimates when necessary. With regards to the package, the model checks for a direct match and falls back to an approximation system that maps similar package types.

Printed Circuit Board (PCB): The PCBs are modeled based on their type: rigid (PCB) and flexible (FPC) circuit boards. The model requires five essential parameters: board type (rigid/flex), layer count, active area, and two utilization factors (shipping panel and material). Depending on the input the model chooses the appropriate emission factor and then scales based on the area of the board. When layer count information is missing, the model defaults to the most common option in the fleet.

Capacitor: Different types of capacitors are considered, including ceramic, aluminum, tantalum, and other. For ceramic capacitors, the model considers several emission factors that account for both material differences (e.g., x5r vs. X7r) and physical size. The methodology employs different scaling approaches based on capacitor type: e.g., area-based scaling for ceramic capacitors, per-piece or mass-based scaling for other types.

Resistor: The model leverages emission factors for different resistor sizes based on inch code. Depending on the size of the resistor, we select the most representative emission factor. Impact is scaled by area or mass accordingly.

Other electronic components: For other electronic components, the model scans for the most representative emission factor, and the impact is then scaled by mass or by area based on the component type and data availability.

If scaled by mass, the component impact is equal to the mass of the component (m), multiplied by an emission factor (ef) per reference mass (m_{ref}) divided by a loss factor (L):

$$E_{part,EE} = \frac{m * ef}{m_{ref}L} \quad \text{Eq 12}$$

If scaled by area, the component impact is equal to the area of the component (a), multiplied by an emission factor (ef) per reference mass (a_{ref}) divided by a loss factor (L):

$$E_{part,EE} = \frac{a * ef}{m_{ref}L} \quad \text{Eq 13}$$

3.3.2.2.5 Interpolation/Extrapolation using Process-based LCA data points (Extrap.)

We collect LCA BOM and EMT attributes to calculate component emissions using the PLCA-Eng pathway where possible. We use the results of the previous pathway, as well as other sources such as third-party certified Product Carbon Footprint (PCF) reports, relevant studies in the literature, to build parametrized models to estimate emissions when primary data is not available.

Component-level extrapolation is employed to estimate embodied carbon emissions through a parameterized relationship between process-based LCA results and components' key technological or engineering characteristics. The parametric relationship between inputs and outputs is then examined and established for extrapolation. This approach allows maximizing the utilization of process-based LCA results while addressing data availability challenges.

CPU: The parametric model is based on total die area and process node. The total silicon die area is estimated either based on CPU generation or number of cores. Process node is also estimated using specifications published by the manufacturers. Whenever no estimation is possible, the core count and process node from the later CPU generation from the specific manufacturer is assumed. imec.netzero model is used to predict the emissions based on total die area and process node.

RAM (DIMM): The first step involves extracting RAM specific parameters such as rank, bit per chip, capacity, and speed, from the part description using JEDEC (Joint Electron Device Engineering Council) Standards. JEDEC Standards are comprehensive technical specifications that define requirements for electronic components, including their manufacturing, testing, packaging, and form factors. Subsequently, the average die dimensions are estimated based on the capacity of the RAM using RAM BOMs collected for the PLCA-Eng pathway as well as RAM BOMs in the TechInsights tear down registry. Imec.netzero model is then used to calculate the emissions of the RAM IC die. Finally, the impact of the PCB and Registered Clock Driver (RCD) are calculated based on dimensions of an average DIMM using the PLCA-Eng pathway.

GPU and GPU Board Assembly: The GPU IC emissions are estimated based on GPU die specifications. The GPU memory emissions calculation follows two paths based on memory type, Graphics Double Data Rate (GDDR) and HBMs (High Bandwidth Memories). In both cases, the final emissions are calculated by scaling an average emission intensity from PLCA by memory size. The impact of GPU Board Assembly is calculated based on the number of GPU cores, the mass of heatsinks and area of the PCB baseboard.

$$E = n \times \underbrace{(E_{die} + E_{HS1} + E_{Mem} + E_S)}_{GPU} + \underbrace{(E_{Motherboard} + E_{HS2})}_{Board Assembly} \quad \text{Eq 14}$$

where E_{Die} is the impact of the GPU chip, E_{HS1} is the GPU heat sink impact, E_{Mem} is the memory impact, and E_S is the GPU substrate impact. This sum is multiplied by the number of GPUs (n) mounted on the board assembly. The board assembly emission includes the motherboard ($E_{Motherboard}$) and heat sink impact (E_{HS2}).

Motherboards and Dongles: The calculation follows a hybrid approach that combines area-based PCB emissions with component-specific CPU/SoC impacts to avoid double-counting. The model first parses the PCB description to extract physical properties including the number of layers and dimensions, and then calculates the PCB area. Simultaneously, it analyzes the main description to identify CPU sockets and SoC counts, while also checking whether the motherboard has a CPU attached. The core calculation uses the following formula:

$$E = (area \times avg_{ci}) + E_{CPU} + E_{SoC} \quad \text{Eq 15}$$

where avg_{ci} is the average carbon intensity per cm^2 derived from historical PLCA-Eng data. The CPU impact is added if CPU is found on the PCB and no SoC are present, while SoC impact accounts for explicitly mentioned system-on-chip components by multiplying their count with average CPU emissions.

SSD: First the total IC die area, and the 3D NAND layer count of the SSD is estimated based on SSD BOM data from Techinsights. Then imec.netzero model is used to estimate the footprint of the SSD NAND ICs with layer count and die area as input. An average impact for the non-NAND electronics (e.g., PCB and interface) and the hardware is calculated and added.

HDD: We use a regression model to estimate the embodied carbon of an HDD based on its capacity. The model leverages available product carbon footprint (PCFs) published for different hard drives (e.g., 18TB, 7200 SATA 3.5" from a given manufacturer), it then parses the drive capacity from the part description, and finally applies a regression to estimate the embodied carbon.

Chassis and Enclosure: The model parses chassis descriptions to extract the physical rack unit size (e.g., '1U', '2U' etc.) and uses a regression to calculate the impact. This parameterized approach assumes chassis emissions scale linearly with physical size:

$$E = E_{base} + E_{slope} \times (U_{size} - 1) \quad \text{Eq 16}$$

where E_{base} represents baseline emissions for a 1U chassis and E_{slope} accounts for additional emissions per rack unit above 1U. The baseline for a chassis with 1U is derived from PLCA-Eng. Similarly, for enclosures the impact is modeled based on size. In this case the model uses a lookup table approach where the system parses the description to extract two key parameters: U_{size} (rack units) and aspect ratio.

Battery Back Up Units (BBU): We have a product carbon footprint (PCF) for each BBU design based on third party conducted and verified LCA studies. Depending on the BBU design we use the corresponding impact value from the PCFs.

AWS Cables: The extrapolation relies on component descriptions to determine the total mass and cable type. This mass is then multiplied by the relevant emission factor to derive the total impact for the cable. This category covers cable located into AWS server racks, such as ethernet, optical, SATA, or PCIe cables.

3.3.2.2.6 Representative Category Average LCA (RCA-LCA)

RCA-LCA (Representative Category LCA) is a pathway designed specifically for long-tail server rack components (low quantity component categories) where traditional Comp-EIO methodology does not produce accurate results due to missing or inaccurate cost estimates. Instead of relying on cost-based emissions estimation, RCA-LCA uses mass-based calculations combined with K-Nearest Neighbors (KNN) classification to map components to appropriate emission factors. KNN classification is a simple but effective machine learning algorithm that classifies new data points based on their proximity to existing labeled data points in the feature space. The algorithm works by first storing all the training data, and when a new data point needs to be classified, it looks at the 'k' closest neighbors to that point (where k is a user-defined number) and decides based on majority voting among these neighbors.

The RCA-LCA process follows three main steps:

- Generate a component inventory of rack asset inventory components in the component-level EIO pathway (see next section) characterized by inaccurate cost data.
- Apply KNN classification to match component descriptions to emission factors, then assign category-average mass estimates.
- Integrate the results into the "waterfall" model (see section 3.3.2.2.3).

3.3.2.2.7 Component-level EIO-LCA (Comp-EIO-LCA)

Environmental Input-Output (EIO) analysis estimates carbon emissions by linking economic activity to environmental impacts through sector-specific emission factors. In the context of IT infrastructure, when analyzing components like racks and their associated parts, we use cost data associated with AWS's rack aggregators as input. These costs are mapped to their corresponding North American Industry Classification System (NAICS) codes, which represent specific industry sectors. Each NAICS code is associated with an emission factor that represents the average carbon emissions per dollar of economic output for that sector. By multiplying the component costs by their respective emission

factors, we estimate the embodied carbon emissions associated with manufacturing and supplying these components.

Cost estimation logic – Component cost data is obtained from quotes provided by rack aggregators. When data is not readily available, the cost is estimated based on other available information, such as the component identifiers or its description. The cost estimation methodology employs a “waterfall-logic” that attempts to find the finest granularity of cost estimate that can be applied to a given component. Such estimates are obtained as an average of available data for different component identifier granularities (e.g., unique ID, component type, etc.) and temporal granularities (e.g., month, year, etc.). For example the trailing-twelve-month average for a unique ID purchased in a given month or the all-time average for a component type purchased in a given month. Furthermore, the logic leverages an automated machine learning (AutoML) framework to predict the cost of a component based on its description and other relevant features, so that a cost estimate is provided for each given component.

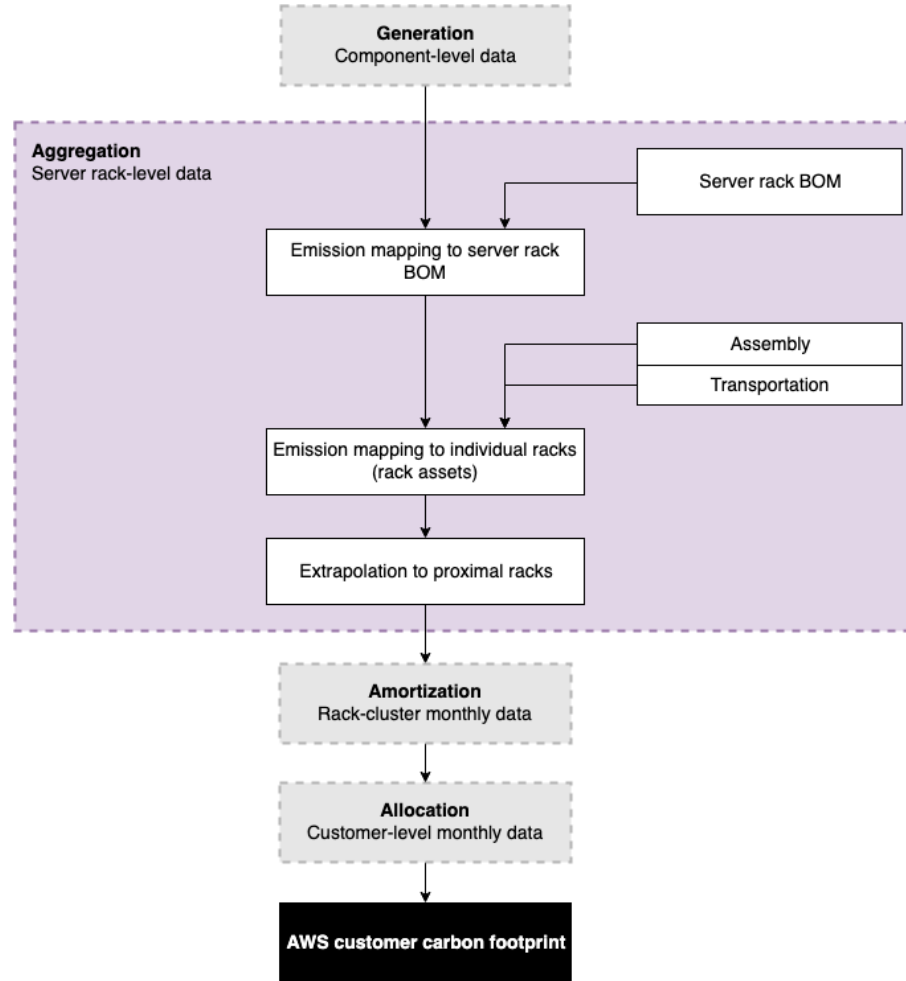
3.3.2.2.8 Data and Methodology Quality

We use data quality scoring – Composite Quality Score (CQS) – and validation checks to track and help improve the overall quality of the data pipeline. The quality scoring evaluates the footprint calculation pathways based on three dimensions: 1) activity measurement method; 2) granularity; and 3) impact factor representativeness. This approach allows us to track the hierarchy of a given calculation method to help inform auditing, and track data quality improvement. Furthermore, we perform data validation at different critical stages in the pipeline, spanning from raw data ingestion to output generation. This validation framework enables early detection of potential issues, drives consistency across our carbon accounting framework, and maintains the standards required for third-party assurance of our emissions data.

3.3.2.3 Aggregation

“Aggregation” refers to the process of estimating rack-level embodied carbon emissions. This estimation leverages the bill of materials (BOMs) for server rack composition, carbon data from component manufacturing, and emissions from rack assembly and transportation to AWS data centers. The output of this step is a single embodied carbon value for each given server rack asset in the AWS fleet. The diagram in **Figure 4** below outlines the end-to-end data flow. The process quantifies server rack emission ($\frac{kgCO_2e}{rack}$), iterating the process each month.

Figure 4 – Aggregation of IT hardware carbon data from component-level to server-rack level



The following sections lay out the logical steps and calculation to produce server rack-level carbon emissions.

3.3.2.3.1 Server Rack BOM

This step defines a BOM for each server rack leveraging AWS’s Product Lifecycle Management (PLM) system that stores BOM data for each server rack design. A server rack design comprises parts that are considered 'fungible'. For instance, an assembly can be composed by a different combination of orderable components that are considered fungible. This means that each unique server rack asset in the fleet that refer to a server rack design will have a specific composition of that given assembly. This step is hence necessary to produce and store the BOM in a workable format that enables mapping of each constituent part/assembly to best available carbon data for every server rack.

3.3.2.3.2 Emission Mapping

This step calculates carbon emissions associated with each server rack BOM defined earlier, by assigning the best available carbon data from the component-level carbon data generation step to each relevant element of the BOM and subsequently aggregating the calculated carbon emissions to rack-level.

Selection of Best Available Carbon Data – This step scans the carbon data made available from the component-level LCAs, selecting the carbon data with the highest Composite Quality Score (CQS) score for each given element i_n of the BOM.

$$E_{i_n} = \left\{ E_{i_n}^k \mid CQS_{E_{i_n}^k} = \max \left(CQS_{E_{i_n}^k} \right) \right\} \quad \text{Eq 17}$$

, where k is the carbon data generation pathway:

$$k \in \{PLCA - Mfg; PLCA - Eng; Extrap; RCA - LCA; Comp - EIO - LCA\}$$

Pathway stitching – Once the best available component-level carbon data is mapped to each element of the BOM, the data is further processed to determine what is the best carbon data for each given assembly. This is required to: determine whether the mapped carbon data of an assembly i_n is of higher quality than the aggregated carbon data calculated from its underlying sub-components j_{n+1}^i ; minimize carbon data gaps at any given level of the BOM.

The process reconciles carbon data availability for a given assembly i_n from a given generation pathway k (see section 3.3.2.2) with that of its underlying ‘children’ j_{n+1}^i . The guiding principles for the stitching are completeness of data, quality and granularity:

1. Completeness – If there are carbon data gaps associated with children, then the carbon data of the ‘parent’ assembly is used:

$$\exists j_{n+1}^i \cdot (E_{j_{n+1}^i} = 0) \Rightarrow E_{i_n} = \left\{ E_{i_n}^k \mid CQS_{E_{i_n}^k} = \max(CQS_{E_{i_n}^k}) \right\}$$

2. Quality – Compare parent’s CQS with the aggregated CQS of children (ACQS). Use parent’s carbon data if CQS is higher than the ACQS of its children, otherwise use the summation of carbon data of its children:

$$E_{i_n} \neq 0 \wedge \forall j_{n+1}^i \cdot (E_{j_{n+1}^i} \neq 0) \Rightarrow E_{i_n} = E_{i_n} \in \left\{ E_{i_n}^k \mid CQS_{E_{i_n}^k} = \max(CQS_{E_{i_n}^k}); \sum_j (E_{j_{n+1}^i} \times qty_{j_{n+1}^i}) \right\} \mid \max(CQS_{E_{i_n}^k}; ACQS_{E_{i_n}})$$

, where $ACQS_{E_{i_n}}$ is calculated based on the CQS of each child ($CQS_{E_{j_{n+1}^i}^k}$) weighted on carbon emissions $E_{j_{n+1}^i}$ and $qty_{j_{n+1}^i}$:

$$ACQS_{E_{i_n}} = \frac{\sum_j (E_{j_{n+1}^i} \times qty_{j_{n+1}^i} \times CQS_{E_{j_{n+1}^i}^k})}{\sum_j (E_{j_{n+1}^i} \times qty_{j_{n+1}^i})} \quad \text{Eq 18}$$

$$E_{j_{n+1}^i} = E_{j_{n+1}^i}^k \mid CQS_{E_{j_{n+1}^i}^k} = \max(CQS_{E_{j_{n+1}^i}^k}) \quad \text{Eq 19}$$

3. Granularity – If parent’ carbon data has the same CQS as the children’s carbon data ACQS, use children’s carbon data:

$$E_{i_n} \neq 0 \wedge \forall j_{n+1}^i \cdot (E_{j_{n+1}^i} \neq 0) \wedge \max(CQS_{E_{i_n}}) = ACQS_{E_{i_n}} \Rightarrow E_{i_n} = \sum_j (E_{j_{n+1}^i} \times qty_{j_{n+1}^i})$$

This condition prioritizes granularity of output data.

Mapping to Unique Rack Assets – In this step each unique server rack asset found in AWS’s asset lifecycle management data is assigned the relevant carbon emissions E_i based on its mapping to rack design. Any data gaps resulting from this step are filled via rack-level extrapolation – the process of inferring carbon data from proximal server racks based on shared characteristics. The process is described in the next sections.

3.3.2.3.3 Rack-Level Extrapolation

There are cases where it is not possible or where we elect not to calculate the carbon emissions for a rack asset based on a BOM definition. Examples of this include cases when a missing BOM data, or cases where aggregate CQS score of the BOM does not meet our quality thresholds (e.g., substantial portion of component emissions in the BOM calculated using the EIO-LCA pathway), and hence an estimate derived from a rack-level model is expected to be more accurate. This section provides insights into the rack extrapolation methodology.

Rack Identifiers – Racks are associated with identifiers that reflect their designation and function (e.g., unique ID, type, family, etc.). These identifiers provide indirect information about a rack’s composition and hence carbon emissions. Racks are also associated with physical characteristics such as the number of servers contained in each rack, weight, dimensions, and total cost of the rack.

Rack Emissions Modelling – To predict the emissions of a rack asset, we use AutoGluon Tabular (AG). Let x_j be the feature vector for $rack_j$, and E_j^{rack} be the emissions for that rack. Then the estimated emissions \hat{E}_j^{rack} is a function of the AG model:

$$\hat{E}_j^{rack} = AG(x_j) \quad \text{Eq 20}$$

and the error ϵ_j equal to the difference

$$\epsilon_j = E_j^{rack} - \hat{E}_j^{rack} \quad \text{Eq 21}$$

AG is configured to optimize for root mean squared error (RMSE). Performance is evaluated based on RMSE, mean absolute error (MAE), and explained variance (R^2).

Input data quality and coverage – We continuously monitor the quality of carbon emission data fed into the extrapolation model, to avoid contaminating the training dataset with low quality data. The current threshold translates into >80% of a racks' carbon data obtained from process-based LCA. With regards to coverage, less than 30% (by carbon emissions) of the server rack fleet is currently estimated via rack extrapolation. The portion of assets for which extrapolation is required will vary over time and be reduced as the availability and quality of data inputs in the LCA process increases.

3.3.2.3.4 Assembly & Transportation

At this stage, each server rack asset tracked in AWS rack asset lifecycle management system has been assigned the best available estimate of embodied carbon emissions associated with the manufacturing of its underlying components. For the sake of simplifying the notation, we have referred to such emissions with the term E_i , instead of E_i^{mfg} . In fact, the embodied carbon emissions associated with a given server rack asset i should also account for the emissions stemming from the assembly and testing of the individual components at AWS's server rack aggregators sites, and from the transportation of the assembled server racks to the data center where such racks are deployed. The following sections describe how such emissions are calculated.

Assembly – The emissions arising from the assembly and testing of a given server rack i :

$$E_i^{assembly} = \zeta \times \bar{e}_{server} \times e_{grid}^i \quad \text{Eq 22}$$

, where ζ is the server density (number of servers in the server rack i), \bar{e}_{server} is the average energy (kWh) required for a server assembly, and e_{grid}^i is the country-specific emission factor based on the location of the server integrator.

Transportation – We refer to the transportation of server racks from rack integrators sites to AWS's data center as "last-mile" transportation. Conversely, the component-level transportation is referred to as "first-mile", indicating the transportation of server rack components from the manufacturing sites to the rack integrators. The emissions associated with the first-mile transportation is handled as part of the component-level modelling.

Let $w_{rack i}$ be the total weight (kg) of rack asset i , $d_{rack i}^k$ the distance (km) per transport mode k travelled between rack integrators and the data centre where the rack is deployed, and $e_{rack i}^k$ the emission factor per transport mode $\left(\frac{kgCO_2e}{kg-km}\right)$; the emissions arising from the last-mile transportation of server rack i is:

$$E_i^{transportation} = w_{rack i} \sum_k (d_{rack i}^k \times e_{rack i}^k) \quad \text{Eq 23}$$

Putting All Together – The resulting embodied carbon emissions for a given server rack asset i are hence:

$$E_i = E_i^{mfg} + E_i^{assembly} + E_i^{transportation} \quad \text{Eq 24}$$

3.3.3 Embodied Carbon in Buildings & Equipment

The methodology outlined in this section outlines the approach for quantifying life cycle greenhouse emissions related to AWS' data center building shell, site development, room builds and long-lead equipment (LLE). The resulting embodied carbon footprint is then allocated to AWS customers (see section 3). **Table 1** provides a list of key terms with definitions that are used throughout this section.

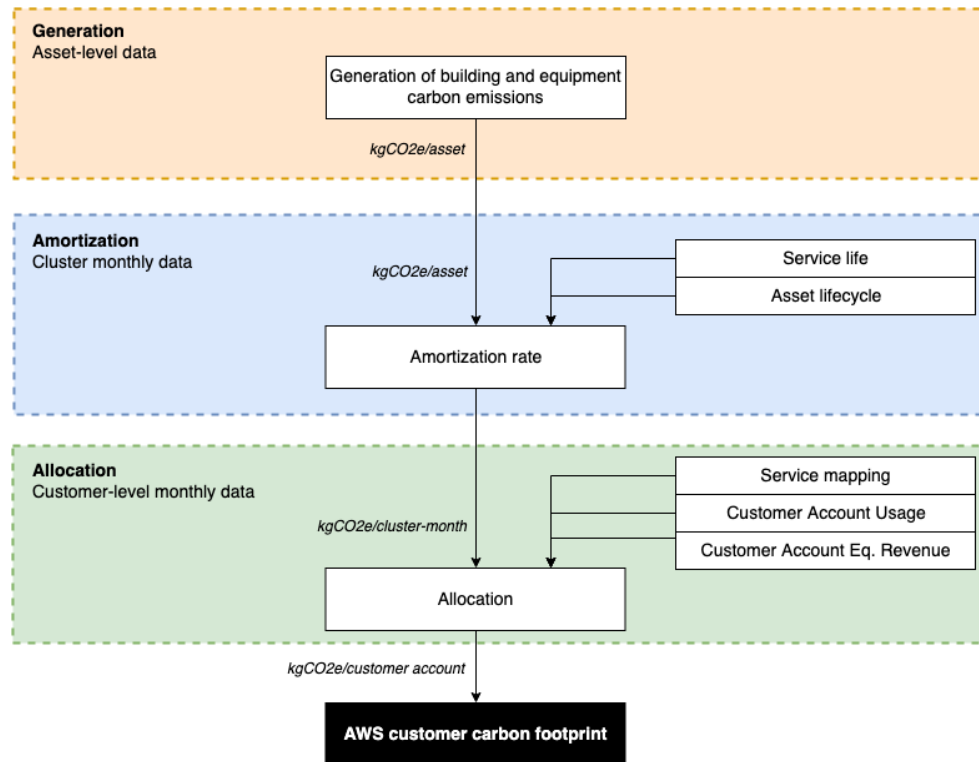
Table 1- Key Definitions and Acronyms for building and equipment carbon data generation

Terminology	Description
Baseline	A benchmark derived from a single building; may be derived from a theoretical design or a constructed building.
BOD	The BOD (Basis of Design) is a unified global collection of reference documents and drawings that outline global AWS requirements for the design and construction of AWS owned data centers .
Bill of flows	The list of all life cycle flow quantities included in the building model scope that are required to model the system boundary.
Bill of materials	The list of product flow quantities included in the building model scope that make up the physical building.
BIM	Building Information Modeling.
Building Model	The building design information used to derive flows for the purpose of wbLCA; e.g., a building information model (BIM), energy model, utility bill, etc.
EPD	Environmental Product Declaration.
GFA	Gross Floor Area measured in square feet (ft^2) or square meters (m^2).
LLE	Long-Lead Equipment are large custom-designed equipment.
LOD	<p>LOD stands for "Level of Detail" or "Level of Development" in BIM (Building Information Modeling) projects. It is a standard that defines the required amount of detail and information included in a Revit model at different stages of the design and construction process.</p> <p>The LOD standards are defined as follows:</p> <ul style="list-style-type: none"> • >LOD 100: Conceptual, high-level information, such as building massing and location • >LOD 200: Approximate, with general system components and approximate sizes or capacities • >LOD 300: Detailed, with specific system components and accurate sizes or capacities. • >LOD 400: Fabrication, with detailed information for construction and procurement. • >LOD 500: As-built, with comprehensive information reflecting the final, constructed project.
PCR	Product Category Rules are documents used for developing LCA models and EPD reports. They provide product category specific instructions on how to model and report an EPD.
Rooms	In the context of DCs, Rooms are considered an embodied carbon program and its scope constitutes all materials and equipment needed for building a room.
Shells	In the context of DCs, Shells are considered an embodied carbon program and its scope constitutes all materials required to building the shell, core and adjacent site of a DC.
wbLCA	Whole Building Life Cycle Assessment is an LCA model of the life cycle of a building.

3.3.3.1 Overview

This section provides an overview of the methodological framework, the standards and protocols considered, the emissions in scope and a high-level description of the constituent components of the end-to-end process. The end-to-end process is represented in **Figure 5** below, and aligns with the steps followed for IT hardware (section 3.3.2):

Figure 5 – End-to-end process flow diagram for building and equipment data



3.3.3.1.1 Underlying Standards

We follow the National guidelines for whole building LCA (wbLCA) published by the National Research Council of Canada (Bowick et al., 2022) as it is a comprehensive guideline which references industry leading normative standards. As a process-based wbLCA guideline, it heavily references and follows EN15978:2011 which is considered a leading international standard for conducting wbLCA. EN15978:2011 in turn references a standard related to LCA for products: EN15804, Sustainability of construction works – Environmental product declarations – Core rules for the product category of construction products. This is a standard for environmental product declarations (EPDs) of construction products. For a North American context, a different EPD standard for construction products is more relevant: ISO 21930:2017. This is because EPD practice in North America is typically in accordance with ISO 21930. Therefore, in this document, the provisions of ISO 21930:2017 are referenced in place of EN 15804, where appropriate. Beyond the specific wbLCA standards, it is also worth noting that the methodology is informed by and aligned with the standards and protocol outlined in section 1.1.

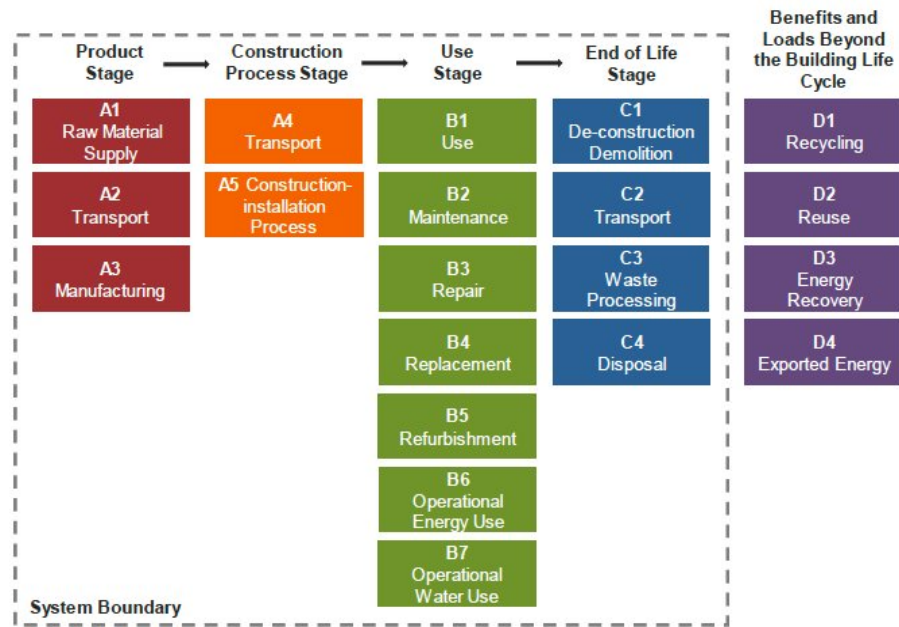
3.3.3.1.2 Attributable Emissions

The methodology excludes any emissions associated with facilities beyond data centers, such as AWS warehouses and manufacturing facilities, offices, or sites run in customer facilities. These emissions are deemed non-attributable.

3.3.3.1.3 System Boundaries

The system boundary defines which life cycle activities are to be included in the analysis. As illustrated in **Figure 6**, the system boundary according to EN15978 is characterized by the life cycle stages. The activities that occur at each stage are classified and grouped in information modules (or simply "modules"), labeled with alpha-numeric designations: the product manufacturing (modules A1 to A3) and transport to site (module A4), construction process (module A5), use (module B), and end-of-life (module C). Module D considers additional benefits beyond the scope of the building life cycle like recycling, reuse and energy recovery or other social or environmental benefits beyond the categories being considered—in this case we currently only consider life cycle carbon emissions across modules A to C and measured with characterization factors representing the GWP100 year time horizon.

Figure 6 – System boundary for building and equipment carbon data generation, per EN 15978 and ISO 21930 (Source: National Research Council of Canada)



The following modules are within quantification scope: A1, A2, A3, A4, A5, B2, B3, B4, B5, C1, C2, C3, and C4. It should be noted that operational energy use (B6) is accounted for as Scope 2 emissions (see section 3.2).

3.3.3.2 Generation of Carbon Data

This section provides an overview of the modeling framework from the generation of carbon data – derived from activity data and representative emission factors – to the detailed footprint results ($\frac{kgCO_2e}{asset}$). These are input into the amortization and allocation steps to generate monthly embodied carbon for AWS customers.

3.3.3.2.1 Activity Data and Emission Factor Mapping

At its foundation, the generation of carbon data relies on process-based LCA models for building shell/core/site and materials/equipment for Room and LLE.

The building shells and rooms LCA models are developed from detailed Bases of Design (BOD) templates made within an AutoDesk Revit design environment where material take-offs covering the building structure are accounted for (i.e., foundation, exterior enclosure, superstructure, interior, mechanical, electrical & plumbing (MEP) equipment and adjacent site works). Such LCAs rely mostly on third-party verified EPDs and on ecoinvent database to cover life cycle activities not quantified in existing EPDs. For LLE, the model leverage supplier-specific LCAs to cover various the critical non-IT equipment such as air handling units, diesel generators, power transformers, etc.

A spend-based approach is used as fall-back carbon data generation pathway when it is not possible to collect the required activity data for a process-based LCA. More information in 3.3.3.2.2.

Modules A1-A3: Product stage

Generally, the emission factors used to represent modules A1-A3 (i.e., cradle-to-gate manufacturing) either stem from an EPD, a contracted equipment-specific LCA study, or an industry average database like ecoinvent. For building shells, we rely entirely on EPDs. In general, we follow a hierarchical approach to determine what source of data to use for any given product or assembly, based on the how reflective available EPDs and LCAs are compared to a product’s technical/physical characteristics, performance. E.g., we prioritize the use of supplier-specific EPDs over industry-average based on relevant region.

Modules A4: Transportation

The module A4 transportation covers the transportation of construction materials and equipment assemblies from place of manufacture to construction site and are estimated using the following approach in a hierarchical order:

1. If the material/equipment is supplier-specific, the facility of manufacture is known (e.g., from a supplier-specific EPD), and transport mode(s) of delivery are known or can be inferred, then manufacturer-to-site logistics are directly modeled.
2. If an EPD is used, then these A4 impact per unit is used for the given material or product input. Generally, when an EPD does include A4 requirements, these requirements are either based on PCR-stated assumptions or otherwise, average distances and transport mode are assumed, and are stemming from third-party verified models and reports.
3. If the LCA model only includes cradle-to-gate (A1-A3) and A4 modeling was out of scope of the study, then we apply default assumptions based onecoinvent's global default data for freight transport per product group.

Modules A5: Construction site and waste treatment

Module A5 activity data covers construction including direct emissions due to fuel combustion or fugitive emissions, emissions associated with construction waste generation such as waste transport, treatment, and the need to procure equivalent quantities of wasted materials to site. A5 requirements are estimated in the following order of preference:

1. Model based on site and material specific estimates.
2. Model based on average data from surveys on actual construction waste of specific materials and fuel consumed across one or more sites.
3. Representative carbon intensity factor based on recommendations from ASHRAE 240P [14] guidance, which includes waste rates per material type.

Modules B2-B5: Maintenance, repair, replacement refurbishment

The use stage modules B2 and B3 relate to any materials and services required for maintenance and repair; B4 considers impacts due to the need to replace building components and equipment; B5 considers any refurbishment requirements. We use spend-based methods based on the spend associated with these modules in the month-year it occurs.

Modules C1-C4: End-of-life treatment

Module C1 includes emissions associated with deconstruction or demolition of the building at the end of its service life. We use the following methods in order of priority:

1. Infer emissions from the associated wbLCA.
2. Representative carbon intensity factor based on recommendations from ASHRAE 240P associated with on-site emissions due to demolition/deconstruction of the building.

Module C2 includes emissions due to transporting building waste materials to an end-of-life treatment center (i.e., landfill, recycling/reuse center, or incineration). We aim to use national or more local jurisdictional average assumptions for distance, transport mode and technology where available. The following assumptions are used in order of priority:

1. Infer emissions from the associated wbLCA.
2. Use a default assumption based on United States Environmental Protection Agency (EPA)'s report³.

Module C3 includes emissions due to waste sorting and processing for reuse, recycling or recovery. The following default assumptions are used in order of priority:

1. Infer emissions from the associated wbLCA.
2. Infer emissions from available EPDs or LCA studies, if applicable and within scope.

Module C4 includes emissions due to waste disposal via landfill and incineration without energy recovery. The following default assumptions are used in order of priority:

1. Infer emissions from the associated wbLCA.

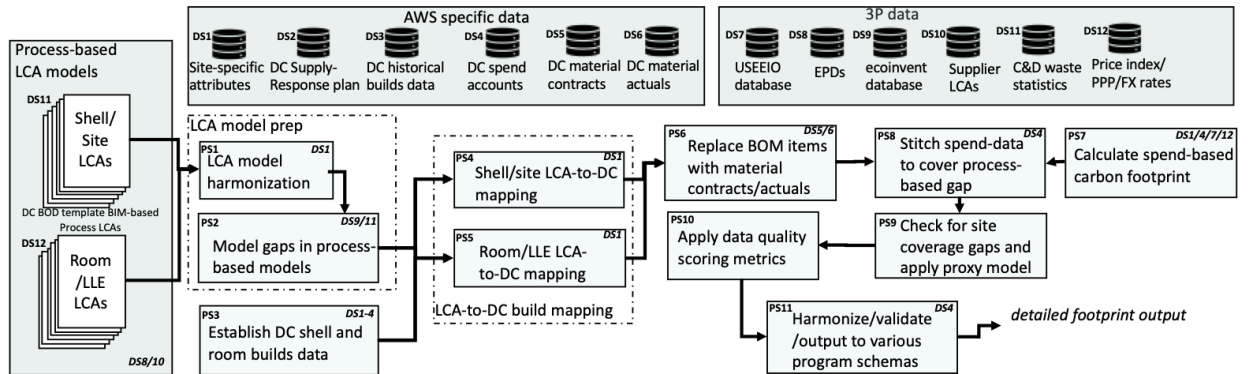
³ [EPA's Life Cycle Inventories of End-of-Life Pathways of Construction and Demolition Materials](#)

- Infer emissions from available EPDs or LCA studies, if applicable, within scope, and regionally relevant to where the waste materials are generated.
- Representative carbon intensity factors for landfilling- and incineration-specific material waste categories from United States (EPA)'s Waste Reduction model (WARM)⁴.

3.3.3.2.2 Modelling Framework

The framework for modeling embodied carbon emissions for AWS data center buildings (i.e., shells + site), rooms (ROM) and long-lead equipment (LLE) (i.e., major equipment that sits within or nearby the building structure) is depicted in **Figure 7**, where distinct process steps (PS#) and datasets (DS#) are stated.

Figure 7 – Process steps and data sets for generating embodied carbon emissions of data center buildings and equipment



The following sub-sections summarize each process step (PS#) in the framework:

PS1: LCA model harmonization – Shell/site and room/LLE LCA models are imported and harmonized. The BODs are design version, region and number of floor/room-specific architectural models. These BODs are considered as 60-80% (i.e., LOD350) detailed structural designs of DCs developed in a Revit model, which is then exported to generate a material take-off or bill of material (BOM) which is utilized in the LCA to map industry- and regional-average EPDs to each material input.

PS2: Missing process-based modules – The life cycle of a building is divided into specific life cycle stages and modules from raw material extraction and product manufacturing to transportation to site, construction, use and end-of-life (see **Figure 6** for details). We utilize default modeling assumptions to cover modules that may not be included in the available LCAs, like material transport to site (module A4), site construction (A5), replacement/refurbishment of equipment and building components during use phase (modules B4/B5), and end-of-life treatment requirements (modules C1-C4).

PS3: Establish DC shell and room builds dataset – The site-specific DC build dataset is established from several sources that document shell and room build. The main purpose of this step is for maximum coverage of builds that fall within the system boundary, as these builds are then mapped to representative process-based BOD LCA models.

PS4 and PS5: Mapping shell/site/room/LLE LCAs to DC – Each asset is mapped to the most representative LCA model from PS1 and PS2.

PS6: Replace BOM items with material contracts/actuals – The BOD-based LCAs is the foundation to the embodied carbon accounting of DCs. However, we prioritize the use supplier-specific contracts and actual quantities procured (e.g., >LOD 400) to replace both the activity data (i.e., BOM-level material inputs) and the EPDs when available.

PS7: Calculate spend-based carbon footprint – The model employs a proprietary advanced large language model (LLM)-based mapping system to accurately map expenditure descriptions to 6-digit NAICS (i.e., North American Industry Classification System) sectors. This system recommends appropriate 6-digit NAICS codes and corresponding BEA sectors, which are then used to generate sector-specific environmental impact factors through the EPA's USEEIO model. The spend-based footprint calculation incorporates these factors, adjusting for currency exchange rates, purchasing

⁴ EPA's Waste Reduction model (WARM)

power parity, and price indices to convert expenditures into carbon footprint estimates. This methodology is applied across all historical capital expenditures in AWS's financial accounts, providing a comprehensive spend-based footprint estimate for all directly paid capital infrastructure.

PS8: Stitch spend-data to cover process-based gaps – While our process-based carbon data generation is generally complete for the defined scope of any building element and equipment, there are building components/systems and service-oriented activities that inevitably do not get incorporated due to a lack of design specification in the Revit models, or a lack of representative and transparent peer reviewed LCA study to map the activity requirement to. We map spend accounts to our process-based LCA model coverage, and in turn, determine the required stitching – process of integrating spend-based data to process-based data – of spend categories or items that were not included in the process-based LCA.

PS9: Check for site coverage gaps and apply a proxy model – The model has a proxy processing module to address critical data gaps in calculating embodied carbon emissions for AWS's data center infrastructure portfolio. The module operates by analyzing well-documented sites to generate proxy footprint intensities, which are then applied to sites with data gaps, providing life cycle coverage across AWS's entire fleet. This approach enables consistent carbon footprint accounting despite varying historical record-keeping practices and data availability challenges. The proxy model follows a four-step process to address data any resulting data gaps in data center embodied carbon calculations from the previous steps. The four steps are:

1. **Filtering and aggregation:** selects sites with complete carbon coverage. The resulting sub-set of sites are used as the foundation for the subsequent step, to create a statistical baseline for proxy calculations.
2. **Time series development:** creates monthly time series, tracking infrastructure elements' carbon footprint and service life assumptions. The time series maintains separate tracking for the different building elements and life cycle modules. This granular approach allows for detailed analysis of how different infrastructure elements contribute to the overall carbon footprint over time.
3. **Infrastructure normalization and statistical analysis:** involves the utilization of a capacity metric to perform normalization. The process considers that data centers vary significantly in size, configuration, and operational capacity, necessitating a standardized approach to enable comparisons and proxy value generation. The primary normalization metric used is active rack count. Critical to this normalization process is the temporal relationship between infrastructure deployment and rack installation. Data centers are constructed in phases where the shell and first room are initially built and brought online while subsequent rooms may be built out in future months. The output of this process are region-specific carbon intensity estimates.
4. **Gap filling and harmonization:** this step categorizes each site into distinct cases based on two primary criteria: the existence of process-based LCA results and spending thresholds across the three key components – shell and site, rooms, and LLE. The system first establishes a continuous time series of cumulative rack counts for each site. This historical rack count data serves as the scaling factor for applying proxy values. The methodology then leverages the regional statistical results from the previous step to generate carbon footprint data for each site. Finally, the model integrates the obtained carbon footprint data with the process and spend-based footprint data obtained via PS1-PS8.

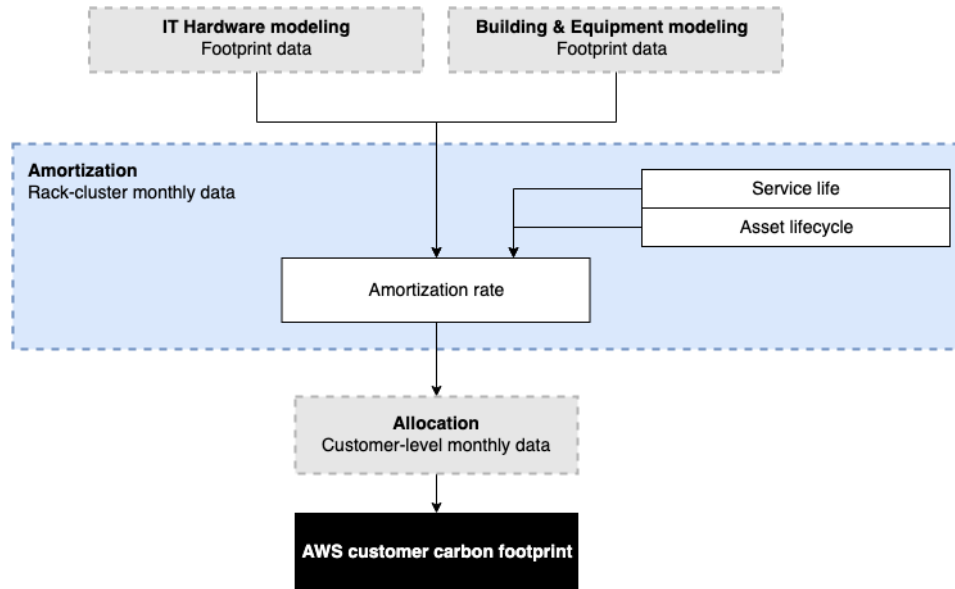
PS10: Data quality scoring metrics – In a similar fashion as what outlined for IT hardware, we developed a data quality scoring method called Composite Quality Score (CQS) which is based on three data quality dimensions, and it is applied to every footprint calculation line item in the model (i.e., BOM-item or LLE level). The data quality scoring approach allows us to track the hierarchy of a given footprint calculation method to facilitate auditing and enable us to track data quality improvements.

PS11: Configure/Validate/Output – This step configures the output based on the requirements of the allocation model – i.e., data schema, aggregation level, etc.

3.3.4 Amortization of Embodied Carbon

As outlined in the previous sections, portion of Scope 3 emissions associated with the provision of cloud services comes from the materials and manufacturing phase of AWS' asset such as – IT hardware, buildings and non-IT equipment. As opposed to 'operational emissions' – i.e., emissions that are directly linked to data center operations such as those arising from generation of electricity to power AWS services, or from fuel combustion in generators – these emissions are generated at a point in time during the lifetime of the given hardware (e.g., server rack). Any AWS customer who uses the asset directly or indirectly (e.g., supporting infrastructure) over its lifetime – through cloud services – shares a portion of the burden. Therefore, similar to financial amortization of CapEx, the Scope 3 emissions should be amortized and then allocated based on customer usage.

Figure 8 – Amortization of capital emissions



3.3.4.1 Overview

The purpose of this section is hence to establish a standardized approach for amortizing the embodied carbon of server racks, buildings, and non-IT equipment across their useful life, to quantify the equivalent embodied carbon emissions to allocate to AWS customers in each reporting month.

Through amortization the embodied carbon data of each asset $\left(\frac{kgCO_2e}{asset}\right)$ is amortized over its service life and passed on to the allocation step as monthly emissions per asset $\left(\frac{kgCO_2e}{asset \times month}\right)$ each reporting month so that the entirety of the embodied carbon of each asset is allocated to the users of cloud services over their service life. Essentially, the amortization methodology, and its underlying model, serve as a bridge between upstream carbon data generation – sections 3.3.2 and 3.3.3 – and downstream customer allocation model. The latter step 'allocation' is described in detail in section 4.

3.3.4.2 Carbon Accounting First Principles

We leverage the standards listed in section 1.1 to define the governing guidelines for the amortization of embodied carbon and allocation to AWS customers. Specifically, the ICT sector guidance states that: "All data center emissions should be allocated to the services that the data center delivers". We interpret "All data center emissions" as all attributable emissions, which include Scope 3 capital emissions, i.e., embodied carbon in IT hardware, data center buildings, and non-IT equipment.

The **overarching principle** derived from the standards is hence the following, and it is applied throughout the methodology, when amortizing and allocating the embodied carbon of such assets:

The entirety of the embodied carbon emissions of any active server rack, data center building, and non-IT equipment shall be allocated to the users of cloud services. In other words, the summation of the monthly embodied carbon of any given asset included in the system boundary over the reporting months in which the asset is active must be equal to the total embodied carbon of that asset.

3.3.4.3 Amortization Logic

The model calculates the amortized carbon for each rack asset found in AWS clusters in any given reporting month as explained below.

Amortization Rate – For each asset deployed in AWS’s clusters in a given reporting month, the methodology employs a linear depreciation to attribute emissions evenly across an asset’s service life. For each asset i , the monthly amortization factor ($amortization_factor_i$) is calculated as:

$$amortization_factor_i = \frac{1}{service_life_i} \quad \text{Eq 25}$$

where $amortization_factor_i$ is the fraction of embodied carbon of asset i to be allocated to each month j while the asset is assumed to be active; $service_life_i$ is the assumed default service life of asset i expressed in months of total expected service life.

Amortized IT hardware embodied carbon – For each given reporting month j , let $month_start_j$ and $month_end_j$ be the month start and end dates. Let also $deployment_date_i$ be the deployment date of the rack i , and $retirement_date_i$ the actual retirement date of the rack i . The logic performs a series of checks for each cluster. Note that the notation used below does not contain any reference to clusters. This is because each server rack asset i is univocally mapped to a given cluster – i.e., each server rack asset is unique – and hence referring to server rack i inherently contain information about the cluster where this is deployed.

Is rack i newly deployed?

$$month_start_j \leq deployment_date_i < month_end_j \Rightarrow \text{rack } i \text{ IS "newly deployed"}$$

Is rack i early retired this month?

$$retirement_date_i \neq "" \wedge month_start_j \leq retirement_date_i < planned_retirement_date_i \\ \Rightarrow \text{rack } i \text{ IS "early retired"}$$

, where:

$$planned_retirement_date_i = deployment_date_i + service_life_i$$

Is rack i beyond retirement this month?

$$retirement_date_i = "" \wedge planned_retirement_date_i < month_start_j \\ \Rightarrow \text{rack } i \text{ IS "beyond retirement"}$$

Subsequently, the logic calculates the amortized carbon emissions.

Amortized carbon emissions for rack i in the reporting month j :

$$amortized_E_i = E_i \times amortization_factor_i \times active_time_i^j \quad \text{Eq 26}$$

, where:

$$active_time_i^j = \min\{retirement_date_i; month_end_j\} - \max\{month_start_j; deployment_date_i\} \quad \text{Eq 27}$$

This means that $active_time_i^j =$

= $month_end_j - month_start_j$ for racks active at month start and month end.

= $retirement_date_i - month_start_j$ for active racks that are retired during the reporting month.

= $month_end_j - deployment_date_i$ for newly deployed racks that remain active at month end.

= $retirement_date_i - deployment_date_i$ for newly deployed racks that are retired in the same month.

Service life of server racks – The service life for the server racks aligns with the financial-based service life of 6 years. The list includes all server racks – including networking racks – in the data center facilities included in the customer carbon footprint model’s scope.

Handling of early retired server racks – There are cases where server racks are retired before their assumed service life, potentially translating into residual embodied carbon in such assets. Any residual carbon needs to be appropriately amortized and allocated to AWS service users in a fair way, in alignment with the carbon accounting first principles outlined in section 3.3.4.2.

Let $E_{i,r}$ be the residual embodied carbon from the early retirement of rack i :

$$E_{i,r} = E_i - \sum_{j=\text{deployment month}}^{\text{retirement month}} \text{amortized_}E_i^j \quad \text{Eq 28}$$

The residual carbon $E_{i,r}$ is calculated as the embodied carbon of rack i minus the summation of amortized carbon from deployment to actual retirement (early retirement in this case). We assume that the retirement date is equal to the planned retirement date calculated based on the predicted service life, disregarding the actual physical retirement date of the server rack.

$$\text{retirement_date}_i = \text{planned_retirement_date}_i = \text{deployment_date}_i + \text{service_life}_i \quad \text{Eq 29}$$

No residual carbon emissions are observed given that:

$$\sum_{j=\text{deployment month}}^{\text{retirement month}} \text{amortized_}E_i^j = E_i \quad \text{Eq 30}$$

Handling of late server racks retirements – As noted above, there are no residual emissions from early retirements as we assume that the retirement date is equal to the planned retirement date. In a similar fashion, this means that for rack assets whose actual service life extends beyond the planned retirement date, their Scope 3 emissions for the extended time is considered zero. This avoids overallocation/double counting.

$$\text{rack}_i \text{ IS beyond retirement} \Rightarrow \text{active}_{\text{time}_i^j} = 0 \Rightarrow \text{amortized}_{E_i} = 0$$

Amortized building and equipment carbon – Each building and equipment asset has an assumed deployment date (i.e., when the asset is actively utilized), and default service life duration (i.e., number of months in operation). Therefore, the amortized carbon due to asset i , is evenly distributed across the timeline between the year/month of deployment date and the year/month when the asset is assumed to be retired. The equations below represent the calculation of the amortized carbon due to asset i in a given reporting year/month j . Note, for simplicity the notation used does not contain any reference to site or cluster and this is because each asset i is mapped to a given site/cluster and hence referring to asset i inherently contain information about the site/cluster where the asset is deployed.

The amortized carbon emissions for asset i in the reporting month j are then calculated as follows:

$$\text{amortized_}E_i^j = \begin{cases} E_i \cdot \text{amortization}_{\text{factor}_i}, & \text{if } \text{yy_mm_start}_i \leq j \leq \text{yy_mm_end}_i \\ 0, & \text{if } j < \text{yy_mm_start}_i \text{ or } j > \text{yy_mm_end}_i \end{cases} \quad \text{Eq 31}$$

, with:

$$\text{yy_mm_end}_i = \text{yy_mm_start}_i + \text{service_life}_i$$

Where:

$\text{amortized_}E_i^j$ is the amortized embodied carbon of asset i allocated to reporting year-month j .

E_i is the total embodied carbon calculated for asset i .

yy_mm_start_i is the year and month of asset i deployment date.

yy_mm_end_i is the year and month that asset i is assumed to be retired.

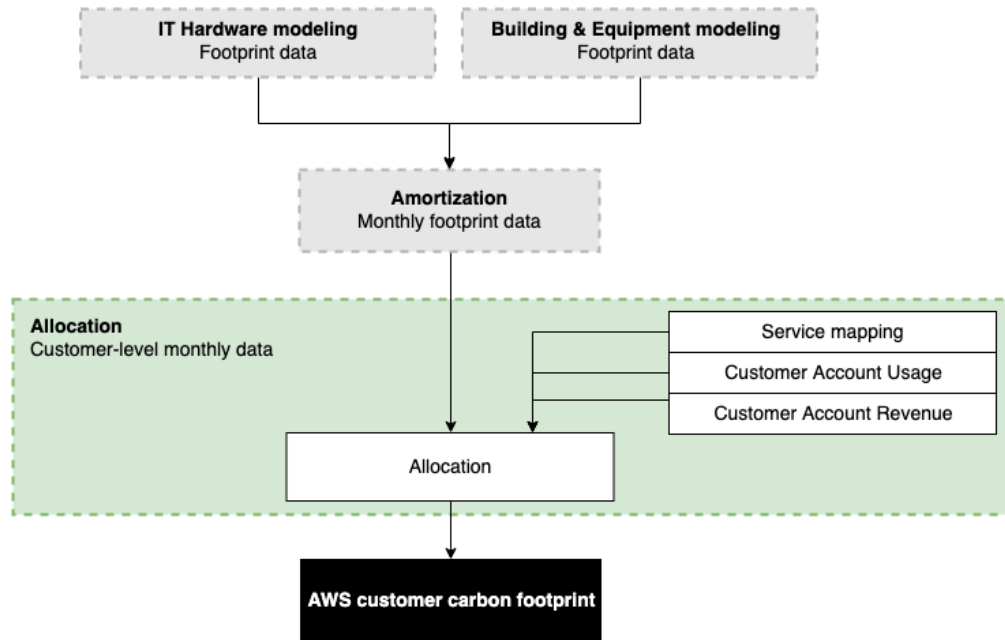
service_life_i is the assumed number of months that asset i is in operational service at an AWS data center.

Service life of data center buildings and equipment – The assumed service life for data center buildings and equipment is based on a combination of their engineering-based and financial-based service life. For instance, the assumed service life of data center buildings is 50 years which is a typical default service life for buildings that is recommend in key building LCA standards (EN 15978:2011).

3.3.4.4 Preparing Amortized Emissions for the Allocation Model

Through the previous step, amortization, the embodied carbon of each asset deployed in AWS' cluster is amortized to obtain the equivalent monthly emissions to allocate to the users of cloud services in any given reporting month. Allocation is the process of allocating such amortized emissions to AWS services, and customers' accounts.

Figure 9 – Allocation of amortized capital emissions



Server racks – The amortized emissions ($amortized_E_i^j$) of a given server rack i in the reporting month j are grouped by rack type k -cluster l pair, which is the granularity used in the allocation model to process server rack amortized embodied carbon. The amortized emissions associated with a given rack type (e.g., EC2.C5) and cluster (e.g., IAD) in the reporting month are hence:

$$amortized_E_{server\ rack\ type\ k}^{j,l} = \sum_i amortized_E_i^{j,l} \mid server\ rack\ asset\ i \in k\ in\ cluster\ l, month\ j \quad Eq\ 32$$

Data center building and equipment – The amortized emissions of a given building and equipment asset i in the reporting month j are grouped by cluster l . The amortized emissions associated with a given building or equipment asset and cluster in the reporting month are hence:

$$amortized_E_{building\ and\ equipment}^{j,l} = \sum_i amortized_E_i^{j,l} \mid building\ and\ equipment\ asset\ i \in cluster\ l, month\ j \quad Eq\ 33$$

Allocation model – The amortized emissions are then made available for allocation: amortized emissions from data center building and equipment are mapped to the associated cluster and added to the rest of cluster-level emissions – i.e., operational carbon; amortized emissions from IT hardware are mapped to each server rack type. Refer to section 4 for more information about the allocation model.

4 Allocation

The carbon emissions described in section 2 are processed in an allocation model that calculates customers' carbon footprint associated with the use of AWS's cloud services on a monthly basis. AWS prioritizes 'physical allocation' – i.e., usage-based allocation – and considers 'economic allocation' – i.e., based on equivalent revenue – as a fall-back option. Meaning, if physical data is not reasonably available, we rely on economic factors.

The model takes operational and capital emissions associated with each AWS cluster and performs a series of transformations to break down such emissions into several logical segments. Conceptually, the model works using the following logical transformation workflow:

1. Allocate cluster-level emissions – i.e., operational carbon emissions as well as building and equipment amortized embodied carbon – to the server racks in the cluster using the server racks' power draw; add the server racks amortized embodied carbon associated with each rack in that given cluster.

For instance, suppose a cluster is allocated 500 MTCO_{2e} and has a total power draw of 1000 KVa. If the total power draw of server rack 1 is 600 KVa, it will get allocated 60% (600/1000) of the carbon, which is 300 MTCO_{2e}. The amortized embodied carbon associated with server rack 1 (e.g., 100 MTCO_{2e}) is added to get the total monthly emissions for such server rack.

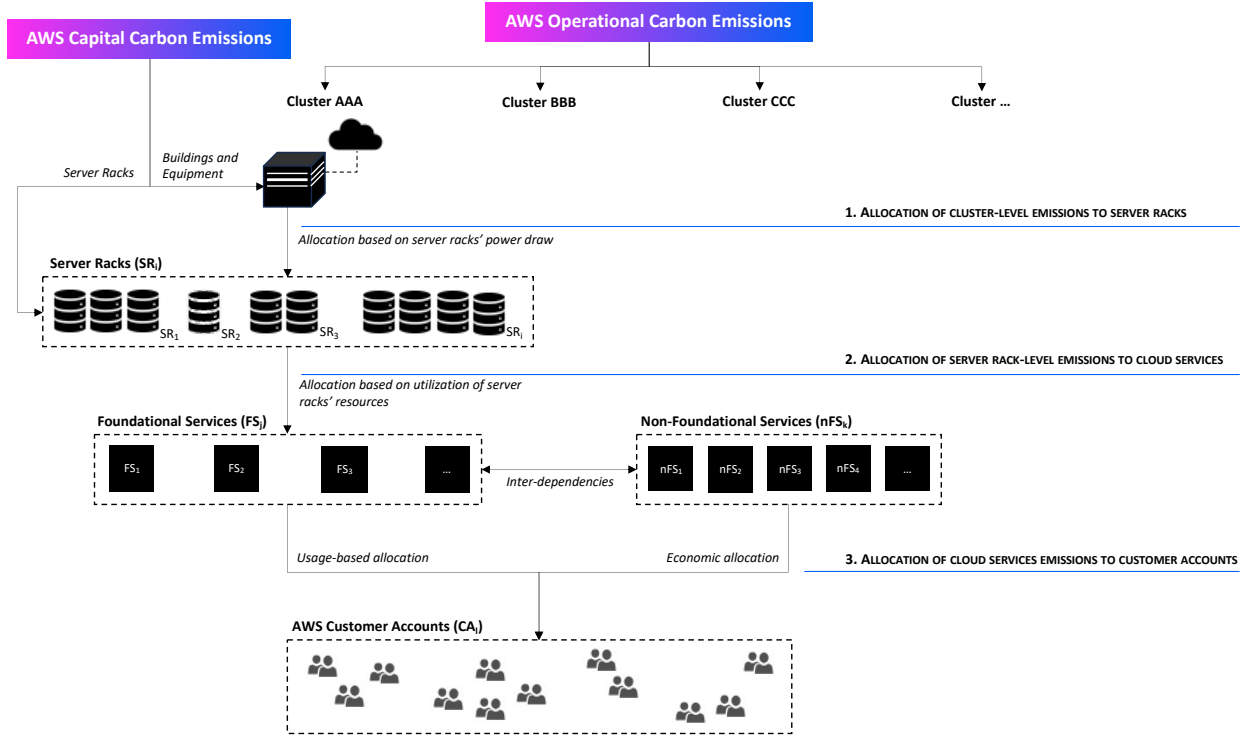
2. Allocate carbon emissions associated with server racks to AWS cloud services based on utilization of server racks, accounting for inter-dependencies between:
 - a) 'Foundational services', AWS services that have dedicated racks in a data center.
 - b) 'Non-foundational services' AWS services with no dedicated racks, which rely on other cloud services to provide functionality to end customers.

*Following our previous example, the model identifies that server rack 1 has a total usage of 1000 GB-month in that cluster. Scanning through the usage records the model observes that 250 GB-month is being used up by cloud service 1. The model assigns 25% (250/1000) of server rack 1's emissions to cloud service 1, which is 100 MTCO_{2e} (25% * 400 MTCO_{2e}).*

3. Allocate carbon emissions associated with each cloud service to individual customer accounts. For foundational services, this is done with physical allocation, while we use economic allocation for non-foundational services.

*Using the same example, the model identifies that cloud service 1 (which had a usage of 250 GB-month) is used across several customer accounts. Customer account 1 used 125 GB-month, and is therefore assigned 50% (125/250) of emissions, which equals to 50 MTCO_{2e} (50% * 100 MTCO_{2e}).*

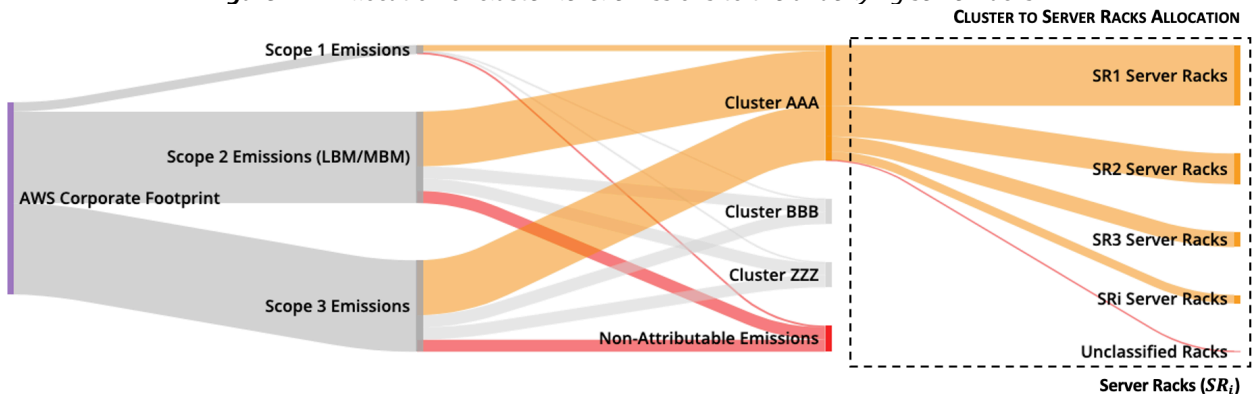
Figure 10 – Conceptual workflow for the allocation of carbon emissions in the customer carbon footprint model



4.1 Cluster to Server Racks

Intent	Allocate cluster-level emissions to the underlying server racks. Applicable for each AWS cluster.
Input	Cluster-level emissions: operational emissions and capital emissions from building and equipment. Rack-level emissions: embodied carbon.
Output	Emissions per server rack.
Type of allocation	Cluster-level emissions: physical allocation, based on Planned Power Draw (PPD). Rack-level emissions: no allocation, embodied carbon is mapped directly to the server racks.

Figure 11 – Allocation of cluster-level emissions to the underlying server racks⁵



In this first allocation step, the model allocates cluster-level emissions to the underlying server racks in the cluster. The weighting factor for the allocation is the share of the total server racks' planned power draw (PPD) in the cluster. Such share is obtained as follows: for each server rack type (e.g., EC2.C5), the model takes the number of server racks ('population'), multiplies it by the average PPD of that specific

⁵ Note: the Sankey diagram is for illustrative purposes only and does not reflect actual split of emissions across clusters, server racks, services and customers.

server rack type, and divides it by the total server racks' PPD in the cluster. This produces the weighting factor, i.e., a percentage based on the server rack type distribution of that cluster. Let % SR_i be the share of power draw attributable to server racks type i , n_{SR_i} their population in the cluster, and PPD_{SR_i} the respective PPD of the server rack type i :

$$\% SR_i = \frac{n_{SR_i} * PPD_{SR_i}}{\sum_i (n_{SR_i} * PPD_{SR_i})} \tag{Eq 34}$$

where:

$$\begin{aligned} i &= \text{type of server rack [EC2 (EC2.C5, EC2.C6N, ...); S3 (standard, glacier, ...); ...]} \\ \sum_i (n_{SR_i} * PPD_{SR_i}) &= (n_{SR_{EC2.C5}} * PPD_{SR_{EC2.C5}}) + (n_{SR_{EC2.C6N}} * PPD_{SR_{EC2.C6N}}) + (n_{SR_{S3\ standard}} * PPD_{SR_{S3\ standard}}) + \dots \\ &= \text{Total server racks' PPD in the cluster} \end{aligned}$$

The carbon emissions associated with a given server rack type i in the cluster are:

$$E_{SR_i} = (E_{cluster} * \% SR_i) + E_{SR_i,embodied} \tag{Eq 35}$$

where, $E_{SR_i,embodied}$ are equal to the amortized embodied carbon associated with the server racks type i that are active in the cluster – as calculated in section 3.3.2.

For instance, suppose a cluster has 500 MTCO_{2e} of associated carbon emissions and has a total planned power draw of 1000 KVa. If there is a total of 100 racks in such cluster with 80 SR_1 consuming a total of 150 KVa, 15 SR_2 consuming 50 KVa, and 5 SR_3 consuming 800 KVa, the model determines that:

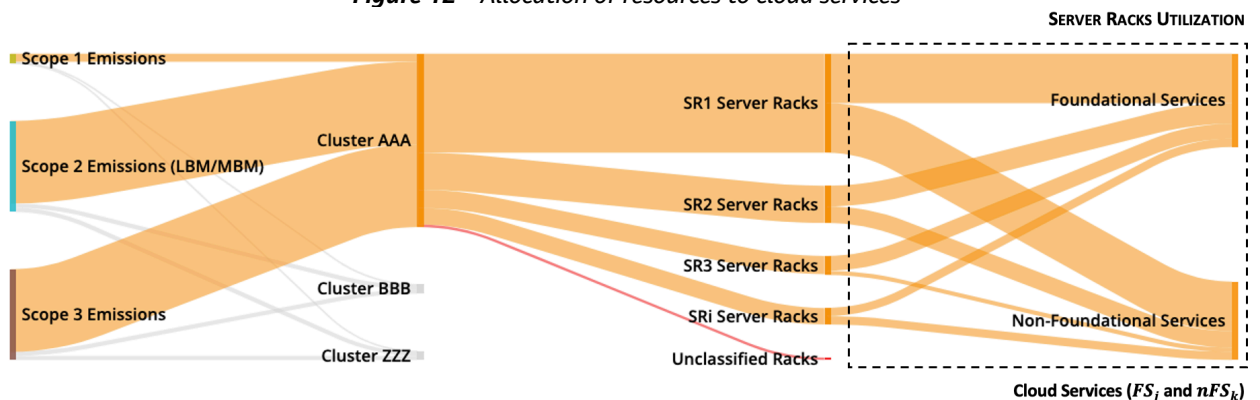
- 15% of the cluster resource allocation should be assigned to SR_1 , which is 75 MTCO_{2e}
- 5% of the cluster resource allocation should be assigned to SR_2 , which is 25 MTCO_{2e}
- 80% of the cluster resource allocation should be assigned to SR_3 , which is 400 MTCO_{2e}

The model subsequently adds the amortized embodied carbon associated with each rack to determine the total equivalent emissions per rack.

4.2 Server Racks to Cloud Services

Intent	Allocate emissions associated with server racks to the cloud services using them. Applicable for each AWS cluster.
Input	Emissions per server rack.
Output	Emissions per cloud service.
Type of allocation	Usage-based allocation.

Figure 12 – Allocation of resources to cloud services⁶



⁶ Note: the Sankey diagram is for illustrative purposes only and do not reflect actual split of emissions across clusters, server racks, services and customers.

In this allocation step, the model defines a unit of usage for each server rack that is used to calculate the % of usage from a given cloud service – foundational services j (% FS_j Use of SR_i) and non-foundational services k (% nFS_k Use of SR_i) – out of the total usage of a server rack i .

The units of usage for each server rack are determined and maintained by each AWS team managing the respective foundational service: e.g., normalized Instance-Hour (NIH) for EC2 rack types, gigabytes (GB) of storage for S3 rack types, etc.

Note: the emissions associated with a type of server rack (E_{SR_i}) – e.g., EC2 types of server racks – is not equal to the emissions associated with the respective foundational service (E_{FS_j}) – e.g., EC2 – unless the server rack supports ‘bare metal’ types of instances, i.e., physical servers provided to customers rather than in a virtualized fashion. Outside of bare metal instances, server racks also support non-foundational services; thus, a portion of the carbon emissions associated with such racks is allocated to non-foundational services. Essentially:

$$\sum_i E_{SR_i} \neq \sum_j E_{FS_j}$$

The share % of usage of a server rack resources from a given cloud service is calculated as follows:

For foundational services (FS_j)

$$\% FS_j \text{ Use of } SR_i = \frac{FS_j \text{ Use of } SR_i}{\text{Tot Use of } SR_i} \quad \text{Eq 36}$$

For non-foundational services (nFS_k)

$$\% nFS_k \text{ Use of } SR_i = \frac{nFS_k \text{ Use of } SR_i}{\text{Tot Use of } SR_i} \quad \text{Eq 37}$$

where:

i = type of server racks [EC2; S3; ...]

j = type of foundational service

k = type of non foundational service

$$\text{Tot Use of } SR_i = \sum_j FS_j \text{ Use of } SR_i + \sum_k nFS_k \text{ Use of } SR_i$$

$$\sum_j FS_j \text{ Use of } SR_i = FS_1 \text{ Use of } SR_i + FS_2 \text{ Use of } SR_i + FS_3 \text{ Use of } SR_i + \dots$$

$$\sum_k nFS_k \text{ Use of } SR_i = nFS_1 \text{ Use of } SR_i + nFS_2 \text{ Use of } SR_i + nFS_3 \text{ Use of } SR_i + \dots$$

There are aggregations occurring in the model, which adjust the server racks usage (% FS_j Use of SR_i and % nFS_j Use of SR_i) to account for different usage pathways/scenarios of foundational and non-foundational services.

Finally, carbon emissions associated with foundational services j and non-foundational service k are obtained as:

For foundational services (FS_j)

$$\begin{aligned} E_{FS_j} &= \sum_i [(\% FS_j \text{ Use of } SR_i) * (E_{SR_i})] \\ &= (\% FS_j \text{ Use of } SR_{EC2}) * (E_{SR_{EC2}}) + (\% FS_j \text{ Use of } SR_{S3}) * (E_{SR_{S3}}) + \dots \end{aligned} \quad \text{Eq 38}$$

For non-foundational services (nFS_k)

$$\begin{aligned} E_{nFS_k} &= \sum_i [(\% nFS_k \text{ Use of } SR_i) * (E_{SR_i})] \\ &= (\% nFS_k \text{ Use of } SR_{EC2}) * (E_{SR_{EC2}}) + (\% nFS_k \text{ Use of } SR_{S3}) * (E_{SR_{S3}}) + \dots \end{aligned} \quad \text{Eq 39}$$

where:

i = type of server racks [EC2; S3; ..]

j = type of foundational service

k = type of non foundational service

$$E_{SR_i} = (E_{cluster} * \% SR_i) + E_{SR_i,embodied}; \text{ i.e., carbon emissions associated with server racks type } i \text{ in the cluster}$$

Following our previous example, suppose the model assigned 400 MTCO₂e to server racks 3 in a cluster. The model identifies that server racks 3 have a total usage of 1000 GB-month in that cluster. Scanning through the usage records the model observes that:

- 50 GB-month is being used up by cloud service 1
- 700 GB-month is being used up by cloud service 2
- 250 GB-month is being used up by cloud service 3

This means that the service utilization is:

- 70% to cloud service 2, equal to 280 MTCO₂e
- 25% to cloud service 3, equal to 100 MTCO₂e
- 5% to cloud service 1, equal to 20 MTCO₂e

Note: usage and emissions in this example are not based on real data.

Utilization cases – The allocation approach described above aligns with the guidelines outlined in the ICT sector guidance [4] which states that 'All data center emissions should be allocated to the services that the data center delivers' and that 'IT equipment that manages workloads for a service or that serves as reserve capacity for a service, should also be allocated to that service and its users'. Essentially, the guideline promotes completeness of allocation, and, in this step, it translates into allocating the entirety of the emissions associated with a given server rack to the underlying services it supports, including the emissions arising from server racks that are not utilized at full capacity. This is captured in the equations defined earlier for foundational and non-foundational services:

$$E_{FS_j} = \sum_i [(\% FS_j \text{ Use of } SR_i) * (E_{SR_i})]$$

$$E_{nFS_k} = \sum_i [(\% nFS_k \text{ Use of } SR_i) * (E_{SR_i})]$$

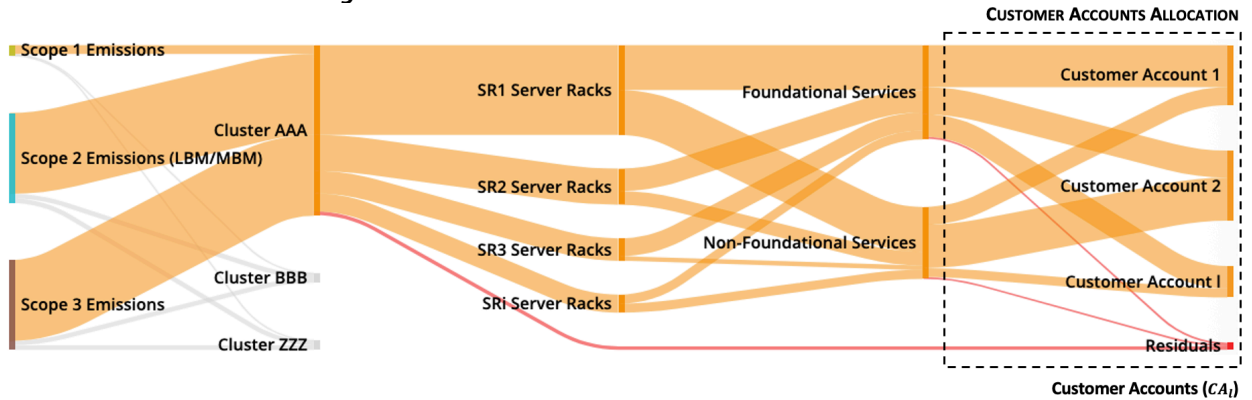
The term SR_i Emissions encompasses the entirety of a server rack's emissions in a given accounting month and cluster.

Building on this, the allocation logic accounts for varying levels of resource utilization across AWS infrastructure, depending on the maturity of the server racks, and the cluster. The logic enables more equitable allocation of emissions in those cases where the service infrastructure is in its early phases ('ramp-up' period). For infrastructure in the ramp-up period (regions and racks <12 months old), the model compares utilization rates against mature infrastructure (regions and racks ≥12 months old) of the same rack type to identify any utilization anomalies. When utilization is lower than the mature infrastructure average, the resulting additional emissions are distributed as a global overhead across all services using that specific resource rack type. This approach prevents overburdening of services and their users in a particular cluster by sharing the emissions associated with the underlying operational overhead in the ramp-up of new clusters/server racks pro rata across services and their users at global level.

4.3 Cloud Services to Customer Accounts

Intent	Allocate emissions associated with cloud services to individual customer accounts. Applicable for each AWS cluster.
Input	Emissions per cloud service.
Output	Emissions per customer account (CA_l where l = account ID).
Type of allocation	Usage-based allocation for foundational services; economic allocation for non-foundational services.

Figure 13 – Allocation of emissions to customer accounts⁷



The type of allocation method used to allocate cloud services' emissions to customer accounts depends on the data available to track the usage of such services from individual accounts. We prioritize 'physical allocation' when usage metrics are available (e.g., customer account l using X GB in the accounting month using $S3$ services in a given cluster). We consider 'economic allocation' as fall-back option when it is not possible to determine a normalized weighting factor of usage across services and customer accounts, or usage data is not complete.

4.3.1 Foundational Services

The share of emissions attributed to foundational services in a given cluster is split among the customer accounts hosting workloads on that cluster, via usage-based allocation. Essentially, each customer account l receives a share of foundational services emissions proportional to its use of a given foundational service j out of total usage of such foundational service for that cluster, using the units of usage identified in section 4.2.

Let $\% CA_l \text{ Use of } FS_j$ be the share of usage of foundational service j from customer account l :

$$\% CA_l \text{ Use of } FS_j = \frac{CA_l \text{ Use of } FS_j}{Tot \text{ Use of } FS_j} \quad \text{Eq 40}$$

Where:

j = type of foundational service

l = customer account ID [XXXXXXXXXXXXX;YYYYYYYYYYYY; ...]

$CA_l \text{ Use of } FS_j$ = Use of foundational service j by customer account l in a given cluster

$$Tot \text{ Use of } FS_j = \sum_l CA_l \text{ Use of } FS_j$$

Therefore, the customer account l emissions associated with the use of a foundational service j , ($E_{CA_l}^{FS_j}$) are:

$$E_{CA_l}^{FS_j} = (\% CA_l \text{ Use of } FS_j) * (E_{FS_j}) \quad \text{Eq 41}$$

Where E_{FS_j} are the carbon emissions associated with the foundational service j in the cluster, as calculated in section 4.2.

Let's assume that in a cluster, cloud service 1 has associated emissions equal to 300 MTCO₂e after step 4.2, and that the service is used across three customer accounts:

- Customer account 1, 400 GB-month
- Customer account 2, 200 GB-month
- Customer account 3, 200 GB-month

The total usage of cloud service 1 is 800 GB-month. Therefore, the model allocates emissions to these customer accounts based on their share of usage of cloud service 1:

⁷ Note: the Sankey diagram is for illustrative purposes only and do not reflect actual split of emissions across clusters, server racks, services and customers.

- Customer account 1 receives 50% of cloud service 1 emissions, i.e. 150 MTCO2e
 - Customer account 2 and 3, receive 25% of cloud service 1 emissions each, i.e. 75 MTCO2e
 Note: usage and emissions in this example are not based on real data.

4.3.2 Non-Foundational Services

The share of a cluster’s emissions attributed to non-foundational services is allocated to customer accounts via economic allocation. An account-level equivalent revenue (*eqRevenue*) is the established weighting factor for the allocation. The equivalent revenue provides a pure reflection of service usage for each customer account, allowing for a consistent comparison across customer segments, time periods and regions⁸. Each customer account *l* receives a share of emissions proportional to its associated equivalent revenue for a given non-foundational service *k*. Let % *CA_l Use of nFS_k* be the share of usage of non-foundational service *k* from customer account *l*:

$$\% CA_l \text{ Use of } nFS_k = \frac{CA_l \text{ eqRevenue for } nFS_k}{Tot \text{ eqRevenue for } nFS_k} \quad \text{Eq 42}$$

Where:

k = type of non foundational service
l = customer account ID [XXXXXXXXXXXX;YYYYYYYYYYYY; ...]
 Tot eqRevenue for nFS_k = $\sum_l CA_l \text{ eqRevenue for } nFS_k$

Therefore, the customer account *l* emissions associated with the use of a non-foundational service *k*, *E_{CA_l^{nFS_k}}* are:

$$E_{CA_l}^{nFS_k} = (\% CA_l \text{ Use of } nFS_k) * (E_{nFS_k}) \quad \text{Eq 43}$$

Where *E_{nFS_k}*

 are the carbon emissions associated with the non-foundational service *k* in the cluster, as calculated in section 4.2.

Let’s assume that in a cluster, the non-foundational service 1 (nFS₁) has associated emissions equal to 200 MTCO2e after step 4.2. The model scans the revenue stream associated with the use of nFS₁ and finds that, out of a total of \$100k equivalent revenue:

- Customer account 1, \$60k, 60% of total equivalent revenue
- Customer account 2, \$30k, 30% of total equivalent revenue
- Customer account 3, \$10k, 10% of total equivalent revenue

The model allocates the emissions associated with the cloud service nFS₁ (200 MTCO2) based on the customer accounts’ % of equivalent revenue:

- Customer account 1, 120 MTCO2e
- Customer account 2, 60 MTCO2e
- Customer account 3, 20 MTCO2e

Note: equivalent revenue and emissions in this example are not based on real data.

4.3.3 Aggregating All Together

The carbon emissions associated with a customer account *l* from the use of cloud services in a given cluster for the accounting month in the customer carbon footprint model are:

$$E_{CA_l} = \sum_j E_{CA_l}^{FS_j} + \sum_k E_{CA_l}^{nFS_k} \quad \text{Eq 44}$$

Where:

j = type of foundational service

⁸ The equivalent revenue is a standardized measure of usage-based revenue that excludes pricing variation factors such as discounts, and other adjustments.

$k = \text{type of non foundational service}$

$l = \text{customer account ID [XXXXXXXXXXXXX;YYYYYYYYYYYY; ...]}$

$$\sum_j E_{CA_l}^{FS_j} = CA_l \text{ total emissions from use of foundational services}$$

$$\sum_k E_{CA_l}^{nFS_k} = CA_l \text{ total emissions from use of non foundational services}$$

4.3.4 Allocation of Residual Carbon Emissions

The overarching principle in the allocation of carbon in the customer carbon footprint model is the 'conservation of carbon emissions', i.e., allocating the entirety of carbon emissions in scope across clusters to customer accounts. For every accounting month:

$$\sum_m E_{cluster_m} = \sum_{l,m} E_{CA_l}^{cluster_m} \tag{Eq 45}$$

Where:

$$E_{CA_l}^{cluster_m} = CA_l \text{ total emissions associated with cluster } m$$

As the final step, the model checks for any residual carbon emissions, and allocates them to customer accounts. Residual carbon emissions can stem from, for example, incomplete mapping of server rack types or services to accounts, or model processing errors. Following the conservation principle, the model reallocates the residual carbon emissions proportionally – based on customer accounts' shares of emissions – using two different level of aggregation:

- Cluster-level, if the root cause for the residual carbon can be traced back to a specific cluster. In this case, the residual carbon is reallocated pro-rata among customer accounts associated with that specific cluster within each scope.
- Global-level, if the root cause for the residual carbon is not directly attributable to a cluster or is not clearly identifiable. In this case, the residual carbon is reallocated pro-rata among all customer accounts at global level within each scope.

AWS conducts regular quality assessments to verify the residual allocation method is working as expected.

Bibliography

- 1 WRI GHG Protocol Corporate Accounting and Reporting Standard (2004)
<https://ghgprotocol.org/>
- 2 International Organization for Standardization ISO 14064, International Standard for GHG Emissions Inventories and Verification (2006)
- 3 WRI Product Life Cycle Accounting and Reporting Standard (2010)
<https://ghgprotocol.org/product-standard>
- 4 Carbon Trust, Global e-Sustainability Initiative (GeSI) ICT Sector Guidance built on the GHG Protocol Product Life Cycle Accounting and Reporting Standard (2017)
<https://ghgprotocol.org/sites/default/files/2023-03/GHGP-ICTSG%20-%20ALL%20Chapters.pdf>
- 5 International Organization for Standardization ISO 14040 (2006)
<https://www.iso.org/standard/37456.html>
- 6 International Organization for Standardization ISO 14044 (2006)
<https://www.iso.org/standard/38498.html>
- 7 United States Environmental Protection Agency (US EPA) Scope 1 and Scope 2 Inventory Guidance (2024) <https://www.epa.gov/climateleadership/scope-1-and-scope-2-inventory-guidance>
- 8 Amazon - Sustainability Carbon Methodology (2024)
<https://sustainability.aboutamazon.com/carbon-methodology.pdf>
- 9 World Resources Institute (WRI) GHG Protocol Scope 2 Guidance
https://ghgprotocol.org/sites/default/files/standards/Scope%202%20Guidance_Final_Sept26.pdf
- 10 ecoinvent: An international Life Cycle Assessment (“LCA”) database in many areas such as energy supply, agriculture, transport, biofuels and biomaterials, bulk and specialty chemicals, construction materials, wood, and waste treatment <https://ecoinvent.org/> published by ecoinvent
- 11 GaBi: A Life Cycle Assessment (“LCA”) software with a content database that details the costs, energy and environmental impact of sourcing and refining every raw material or processed component of a manufactured item <https://sphera.com/life-cycle-assessment-lca-database/> published by Sphera
- 12 Interuniversity Microelectronics Center: imec.netzero model <https://netzero.imec-int.com>
- 13 TechInsights: TechInsights Platform; The Trusted Source for Semiconductor Insights
<https://www.techinsights.com/> published by TechInsights
- 14 ASHRAE: Quantification of Life Cycle Greenhouse Gas Emissions of Buildings
<https://www.ashrae.org/>