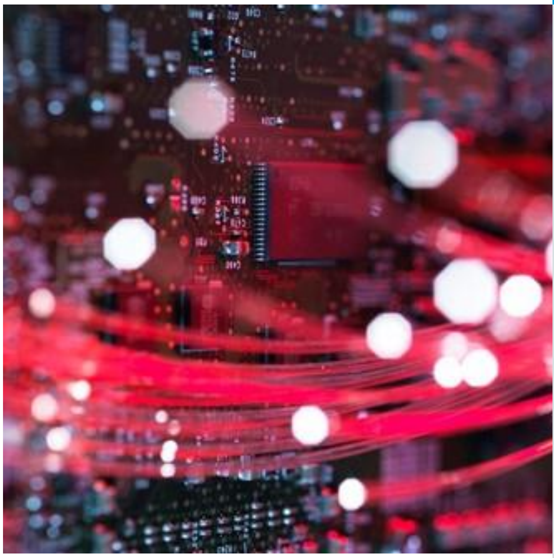


Study for the review of Commission Regulation 2019/424 (Ecodesign of servers and data storage products)

Task 3 Users – DRAFT v3

June 2024

Submitted to:

Davide POLVERINI

Policy Officer, DG GROW

Avenue d'Auderghem 45, 1040 Brussels, Belgium

Task 3 Users – Draft v3

A report submitted by ICF S.A.

Date: June 2024

Tom Lock
ICF S.A.
Avenue Marnix 17
Brussels
B-1000
Belgium
T +32 (0) 2 275 01 00
www.icf.com

Document Control

Document Title	Task 3 Users – Draft v3
Prepared by	Laurent Petithuguenin, John Clinger, Todd Leddy
Checked by	Eirini Passia, Tom Lock
Date	June 2024

ICF has used reasonable skill and care in checking the accuracy and completeness of information supplied by the client or third parties in the course of this project under which the report was produced. ICF is however unable to warrant either the accuracy or completeness of such information supplied by the client or third parties, nor that it is fit for any purpose. ICF does not accept responsibility for any legal, commercial or other consequences that may arise directly or indirectly as a result of the use by ICF of inaccurate or incomplete information supplied by the client or third parties in the course of this project or its inclusion in this project or its inclusion in this report.

ICF

Contents

Table of Figures

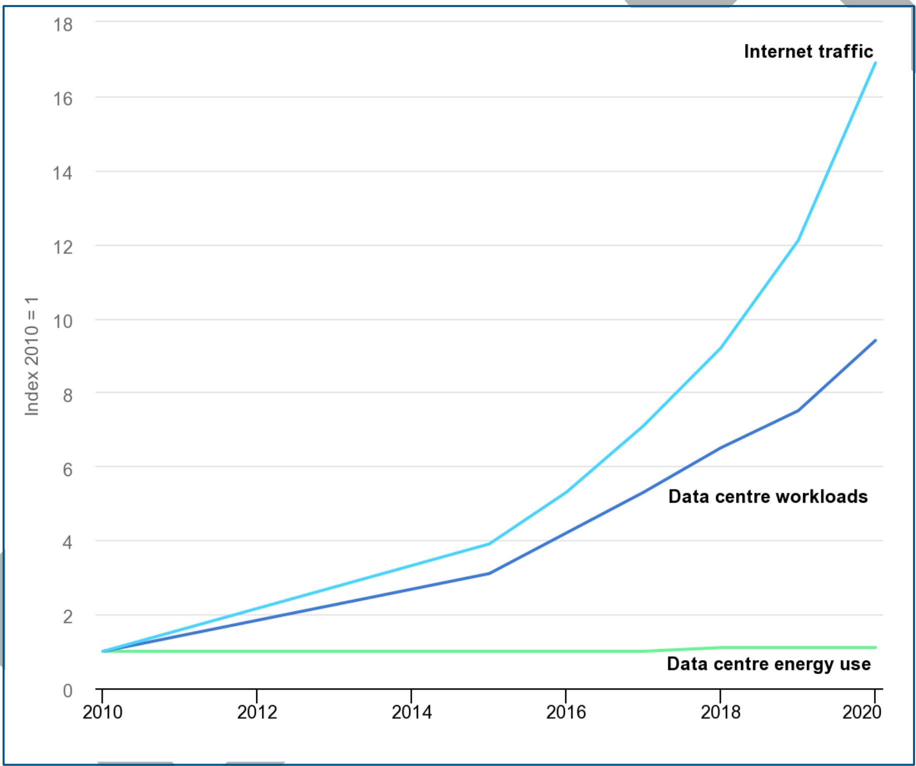
Table of Tables

3 Introduction to Task 3 Users

3.1 System aspects of the use phase for ErPs with direct energy consumption

Computer servers and data storage products underpin the digital service economy. Since 2010, the number of internet users worldwide has more than doubled, while global internet traffic has expanded 20-fold (as shown in Figure 3.1). Table 3.1 further emphasises this by demonstrating how from 2015 to 2021, with 1.9 billion more users on the internet (60% increase), there was 440% increase in global internet traffic. Not only are there more users connecting to the internet, but operations, storage, and services are relying more and more on the internet.

Figure 3.1 Global trends in internet traffic, data centres workloads and data centre energy use, 2010-2020[1]



Datacentre workloads have followed a similar curve of development as internet traffic from 2010 to 2020, although at a less intense rate. This amounts to a 260% increase in workloads from 2015 to 2021. Yet the energy consumption of data centres has only grown by 60% in the same period (excluding crypto). Energy efficiency improvements have reduced the growth in energy demand from data centres, thanks to improvement in IT hardware capabilities and cooling technology. There has also been a shift away from small, inefficient data centres towards more efficiency cloud and hyperscale providers. Global data centre electricity use in 2021 was 220-320 TWh, or around 0.9-1.3% of global final electricity demand. This excludes energy used for cryptocurrency mining, which was 100-140 TWh in 2021.

[1] Nov 2021, IEA report, Global trends in internet traffic, data centres workloads and data centre energy use, 2010-2020 – Charts – Data & Statistics - IEA

Table 3.1 Global trends in digital and energy indicators, 2015-2021[2]

	2015	2021	Change
Internet users	3 billion	4.9 billion	+60%
Internet traffic	0.6 ZB	3.4 ZB	+440%
Data centre workloads	180 million	650 million	+260%
Data centre energy use (excluding crypto)	200 TWh	220-320 TWh	+10-60%
Crypto mining energy use	4 TWh	100-140 TWh	+2 300-3 300%
Data transmission network energy use	220 TWh	260-340 TWh	+20-60%

Although the global trend for data centre energy consumption has been moderate, as data services can be provided on a global scale, there can be strong impacts for their needs at the local level. For example, Irish data centre electricity use more than tripled since 2015, representing 14% of total electricity consumption in 2021. Similarly, in Denmark, data centre energy use is projected to triple by 2025 to account for around 7% of the country's electricity use.

This report will review how servers and data storage products, which are the primary service provider equipment in data centres (and for internet and data services at large), are being used, delivering their work outputs, and the associated energy consumption this work requires. Networking equipment will not be covered by this Task 3 report.

3.1.1 Strict product / component scope

3.1.1.1 Product introduction and work delivered

Servers

Enterprise servers deliver tasks similar to personal computers, however the distinction is that they are described as "servers" as they do these tasks in response to the request from another computer. The only way to request this is through their Input/Output socket, which is to mean the internet. There is no screen, keyboard or mouse directly associated to the computer server.

A computer server may be dedicated to a particular task (and hence have specific hardware to deliver it) or perform a variety of tasks as required. The software can also determine the functions of the server. However, this depends on the user specification and the nature of the workloads. An example of some typical tasks include:

- Mail servers: They move, store, and send email. Typical software platforms are Microsoft Exchange, Gmail, Yahoo.
- Web servers: They host the content of a website to a user's Web browser utilising Hypertext Transfer Protocol (HTTP). When a device connects to a website, these servers will respond to show the related content on the page.
- File servers: They store and manage data files for access between computers. Transfer over the internet is done using FTP (File Transfer Protocol) or Secure FTP.

[2] Sep 2022, Data Centres and Data Transmission Networks – Analysis - IEA

- Database servers: They provide database services to client computers (users). Typical software platforms include SQL, SAP, and Oracle.
- Application servers: They are dedicated to the execution of programs, routines, scripts and work. These are sometimes referred to as “middleware” servers as they provide a connection between database servers and the user. The application server is programmed via a software platform such as Java, PHP and Microsoft.
- Terminal servers: They support today dedicated remote (virtual) desktop services including graphical user interfaces (GUI). Typical software platforms are from Microsoft and Citrix.
- Proxy / communication / VPN servers: They are dedicated to filter communication requests (gateway/firewall), share connections, and improve and monitor performance.
- DNS: Standing for Domain Name Server, these have an enlarged database of IP addresses with their linked hostnames. They enable communication between devices across the network.
- Dedicated servers: they provide a service to a specific client, and nobody else. This ensures prioritisation of the client’s task.

Other server types of mention are online gaming servers, chat servers, groupware servers.

Server operations are needed for all sorts of business operations, such as financial services, telecommunications, internet services and media providers, but also providing computing facilities for sectors such as private businesses, government, health, education and industry.

The work delivered by computer servers is defined by the calculations and storage operations delivered. This work is varied and can be difficult to define. Unlike some other devices, the operation of a server is not as simple as describing when it is on or off. Servers are technically always “on”, however that is not the same as when they are delivering useful work. When the server is not delivering work, this is defined as “idle”. These are the different factors that define the energy consumption of a sever:

- The time spent in idle phase and at different operating load levels. This is averaged into the “utilisation” level of the server. This is developed further in section 3.1.1.2.
- The operations the server is delivering in that time.
- The efficiency of the server at delivering these tasks.

To determine the efficiency of a server at delivering various tasks, SPEC (Standard Performance Evaluation Corporation) have developed SERT, the Server Efficiency Rating Tool. The tool defines the different workloads under worklets. The workloads are defined as the different operations relating to CPU, Memory and Storage. The worklets are the different operations taken within these workloads. They are tested under different load levels. These worklets and the load levels of operation are set out in Table 3.2.

Table 3.2 SPEC SERT worklets

Workload	Load levels	Worklet
CPU	25% / 50% / 75% / 100%	Compress
CPU	25% / 50% / 75% / 100%	CryptoAES
CPU	25% / 50% / 75% / 100%	LU

Workload	Load levels	Worklet
CPU	25% / 50% / 75% / 100%	SHA256
CPU	25% / 50% / 75% / 100%	SOR
CPU	25% / 50% / 75% / 100%	SORT
CPU	25% / 50% / 75% / 100%	XMLValidate
Memory	Full/Half	Flood
Memory	4 / 8 / 16 / 128 / 256 / 512 / 1024 GB	Capacity
Storage	50.0% / 100.0%	Random
Storage	50.0% / 100.0%	Sequential
Hybrid	12.5% / 25.0% / 37.5% / 50.0% / 62.5% / 75.0% / 87.5% / 100.0%	SSJ
Idle	No load	Idle

SERT uses these results to provide the average performance of a server. The different workloads are weighted under a geomean to provide a representative score of the average time a server spends in each workload.

As the SERT test can be time consuming to perform, and servers are subject to a high variation of configurations within the same server family, SERT has determined five configurations to provide a representation of server family performance: minimum power, maximum power, low-end performance, typical performance, and high-end performance configuration. It is important to note that minimum and maximum power configurations are not analogous to the minimum and maximum energy efficiency scores. Indeed, these configurations are defined under their total energy consumption, which doesn't include the factor of the amount of work delivered. Therefore, the minimum power configuration may consume less power, but does not deliver the same amount of work, and hence may be less efficient on a work delivered per Watt basis.

SERT metric is currently used by Energy Star, Blue Angel and Ecodesign as a performance requirement of active efficiency as discussed in Task 1. For server families, Blue Angel and Energy Star require for the criteria to apply to minimum performance, maximum performance and typical configuration servers. The Ecodesign regulation 2019/424 only applies this requirement on the low-end and high-end performance configurations.

ICF collaborated with Information Technology Industry Council (ITI) and The Green Grid, to create a database of SERT output results for 575 unique server families from 2017 to 2022. From this database, the Figure 3.2 was completed on the 2-socket rack servers server category types. These were chosen as they are the most common server configuration. Figure 3.2 shows how the average SERT score has continued to go up, indicating that the server market continues to increase in efficiency. This is driven by an increase in performance capability.

Figure 3.2 Average SERT server score and the average idle consumption from 2016-2021 for 2-socket rack servers[3]

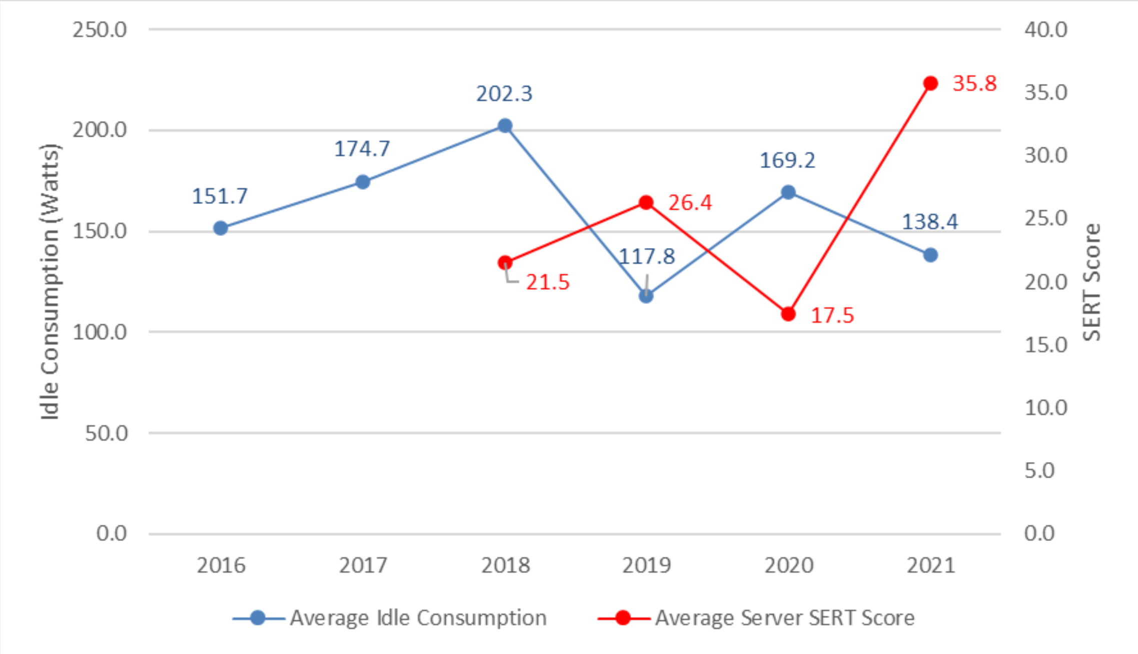


Figure 3.2 shows how the idle consumption is correlated with the server SERT score. Where the server active efficiency score is increased, the idle consumption decreases, as the improved idle score is a component of the final active efficiency score. It should be noted that although the SPEC SERT database is the most complete server energy efficiency database available, in some years there are a limited number of data points, and which types and numbers of servers are in the database significantly vary by year. For example, one year may include a much larger proportion of servers with high performance CPUs, and other years may mostly include servers with lower performing CPUs. Two factors which increased the year-to-year variation in Figure 3.2 are systems with large numbers of storage devices, systems with 1 of 2 CPUs installed and resilient servers[4]. For example, the SERT average in 2020 is artificially low because the database lacks AMD based servers (except one 8-core) and high-end Intel CPU based servers (only two), both of which had high active energy efficiency. All of the other servers included only had medium and lower performing CPUs with lower active energy efficiency[5].

The SSJ (server-side java) operation is a hybrid worklet designed to exercise the CPU and memory activities. Although this is not as comprehensive as the SERT score, it is therefore often used to have an indication of server capabilities.

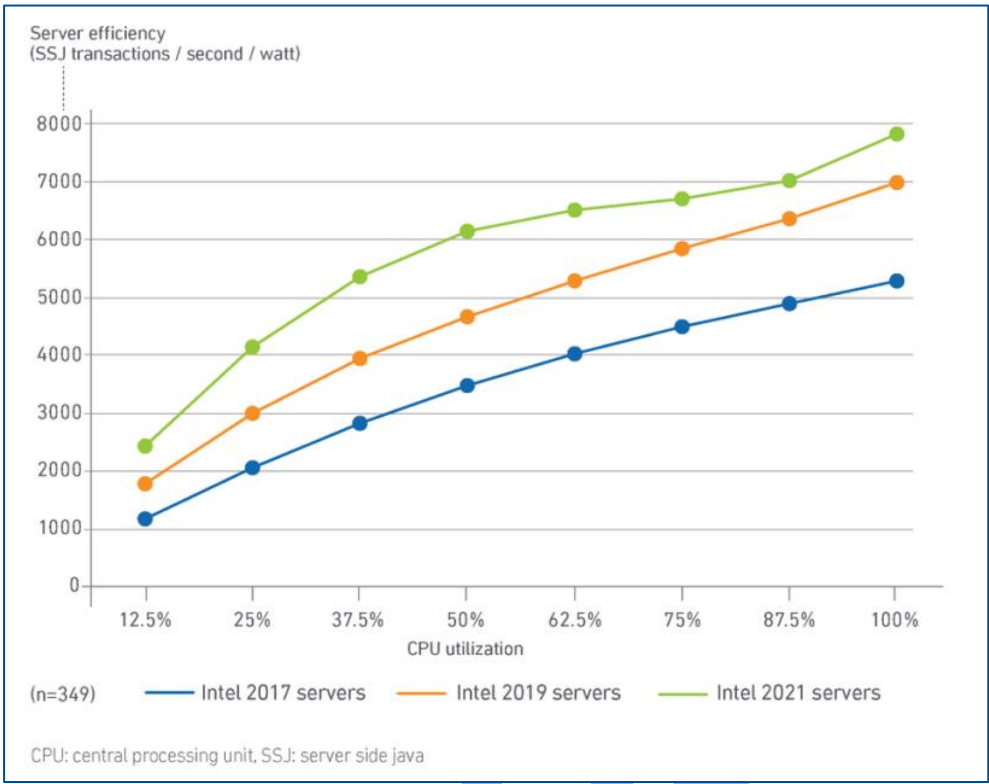
Figure 3.3 shows the trend for energy performance for Intel servers. There is a clear improvement in efficiency from the 2017, to 2019, to 2021 servers. Moreover, as the graph is set on an axis of CPU utilisation, it shows how servers are more efficient when operating at higher utilisation rates. Figure 3.3 also demonstrates that as the CPU utilisation rate grows the linear increase in server efficiency decreases.

[3] SPEC SERT server data set

[4] Stakeholder Feedback

[5] Stakeholder Feedback

Figure 3.3 Efficiency improves with each Intel server generation[6]



Data storage products

Data storage systems provide data storage services for devices either directly connected (host) and/or to remote computing devices (client) via a network connection. This service is meant to supplement the internal storage of servers.

The following insight on the classification of data storage products under SNIA taxonomy has remained largely constant since the 2015 Lot 9 Preparatory study[7]. These devices are mainly specified according to their capacity and access criteria, which would include latency and reliability. The Storage Networking Industry Association (SNIA) defines the storage taxonomy as follows:

- Access pattern (random or sequential);
- The maximum time to first data (max. TTFD in ms), required to start receiving data from a storage system;
- The requirement for user access;
- Connectivity over network or direct connection to a single or multiple hosts;
- Integrated storage controller (optional or integrated);
- The status (optional or required) of storage protection, non-disruptive serviceability, no single point of failure, and storage organisation;
- Maximum supported disk configuration.

These functionalities and features have resulted in SNIA creating six product group categories:

[6] Transactions per megawatt-hours: Keys to increasing data centre efficiency, Uptime intelligence, 27 June 2023

[7] Source: Preparatory study for implementing measures of the Ecodesign Directive 2009/125/EC, DG ENTR Lot 9 - Enterprise servers and data equipment, Task 3: User, July 2015: Final report, bio by Deloitte & Fraunhofer IZM

- Online: Storage system for very fast random or sequential I/O request. The main distinction criteria is a maximum TTFD of >80 ms.
- Near Online: Storage system for moderate response time with maximum TTFD of >80 ms.
- Removable Media Library: System for sequential I/O request with long response time. This is an automated or manual media loader such as tape or optical library.
- Virtual Media Library: System for very fast sequential I/O request with maximum TTFD of <80ms. The media are not removable and intended for long-term data storage.
- Adjunction Product: Special purpose storage service, dedicated data path from host to storage device, no end-user access, maximum TTFD of <80 ms.
- Interconnect element: Managed interconnect elements within a storage area network such as switch or extenders.

Figure 3.4 provides a visual overview on how these product categories are distributed. For the full taxonomy detail please review Annex 1.

Figure 3.4 SNIA Taxonomy overview table[8]

Category Level	**Online** (see 5.3)	**Near Online** (see 5.4)	**Removable Media Library** (see 5.5)	**Virtual Media Library** (see 5.6)
Consumer/ Component[1]	Online 1	Near Online 1	Removable 1	Virtual 1
Low-end	Online 2	Near Online 2	Removable 2	Virtual 2
Mid-range	Online 3	Near Online 3	Removable 3	Virtual 3
	Online 4			
High-end	Online 5	Near Online 5	Removable 5	Virtual 5
Mainframe	Online 6	Near Online 6	Removable 6	Virtual 6

The enterprise sector storage products are mainly found in the low-end to mid-range Online 2, Online 3, and Online 4 (and Near Online 2 and 3 to a lesser extent). They are designed for random and partially sequential I/O requests. Storage media are typically more economical HDDs or for certain purposes SSDs. The low to mid-range online systems are utilised in storage pools with defined redundancy (RAID) and respective control.

Other categories of note are:

• Online 5 and 6 are higher performing (specialised) storage systems. These have high performance requirements on capacity, computation and controls, which are used in specialised applications.

• Removable media libraries and virtual media libraries including tape libraries are data back-up systems. These have a small power usage and market share.

[8] Taxonomy | SNIA

3.1.1.2 Utilisation rates for servers

A key metric for the energy consumption efficiency of a server is its utilisation rate. Indeed, as servers are always "online", they are technically always "on" and consuming power. The utilisation level is a metric calculated as the average operational level of a server, including the time spent idle and the time at varying operating levels.

This utilisation metric is important as it provides clarity on the level of operational capacity servers are delivering. Servers may always be switched "on", but the questions to be asked here are: how much power is being consumed in this time? and how much work is being delivered in this time?

A higher utilisation rate of servers results in a more efficient system, as fewer servers would be required to deliver the same amount of work (saving the cost of the resource to build those devices), but also as shown in Figure 3.3, servers at a higher utilisation rate tend to be more energy efficient on a watt per work delivered basis.

Current estimates for the average utilisation rates of servers are quite low, with the Uptime Institute Intelligence survey results showing that *"at least 40% of servers operate at <30% utilisation"*[9]*.* This low rate is generally justified by operators due to an abundance of caution to ensure that there is capacity to respond to peak demand times. Statements from IBM correlate this figure stating that the average rate of server utilisation is of 12-18% capacity.[10] Using a normal bell curve time distribution for a utilisation level of 12-18% would imply that the average server is in idle mode between 10-25% of the time. However, the Uptime Institute Intelligence indicates that for Enterprise and office applications, utilisation rates can be increased to 50% without any loss in performance. 50% utilisation rates should be used as a target best practice as increasing to higher utilisation levels may result performance failures.[11]

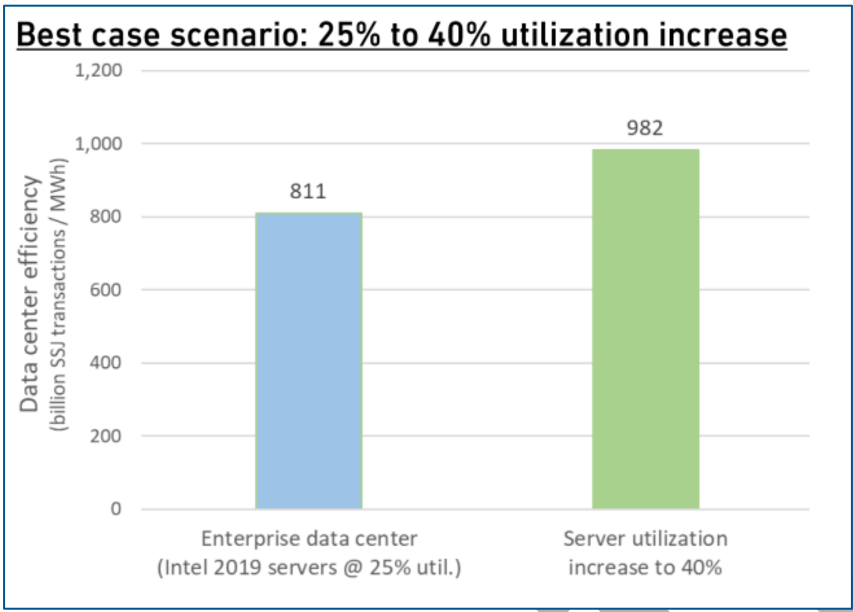
Furthermore, the Uptime Institute calculated in Figure 3.5 the efficiency benefits of shifting utilisation rates on the same hardware from 25% to 40%. This shows a 21% increase in efficiency of the SSJ transactions.

[9] Transactions per megawatt-hours: Keys to increasing data centre efficiency, Uptime intelligence, 27 June 2023

[10] Are Your Data Centers Keeping You From Sustainability? - IBM Blog

[11] Transactions per megawatt-hours: Keys to increasing data centre efficiency, Uptime intelligence, 27 June 2023

Figure 3.5 Efficiency savings from increased utilisation



The main concern from users with increasing the utilisation rates of an average server, is that in times of peak demand, the system would not cope with the additional demand and would crash. Therefore, additional buffer capacity is placed onto the system is provided to ensure the system is resilient. However, techniques are now emerging to ensure resilience at higher utilisation rates. For example, Intel has developed the concept of “effective resource utilisation”, which reviews not only if devices are being used, but also if that use is effective. For example, if computational programs on a system have been terminating before completion, that utilization is useless for the operator. Targeting those programmes, to ensure they are appropriately delivered (rather repeated) will reduce the total demand for processing.[12]

3.1.1.3 Data storage products utilisation

Storage products have distinctive sub-systems to bear in mind when discussing utilisation, namely the distinction between the storage controller and the storage devices.

Storage controllers are either installed internally or externally as an extra controller enclosure (CE) with attached disk enclosures (DE) which can provide extra scalability. These systems are capable of organising hundreds to thousands of attached HDDs and SDDs. Therefore, the utilisation of controllers is likely to be higher than that of storage devices. Although the controller has the higher utilisation rates, the storage devices have the higher energy consumption, which can make up to 80% of the total device consumption. It is important to note that unlike with servers, tuning for the utilisation of storage systems for idle is not a priority. Higher controller utilisation may result in performance below the required 20ms latency, hence a balance needs to be struck between controller usage and the number of drives serviced.

Storage products are organised into two types: transaction and streaming products.

[12] IT@Intel: Data Center Strategy Leading Intel’s Business Transformation, White paper, July 2020

Transactional data storage is optimized for running production systems (everything from websites to banks to retail). These are designed to read and write data quickly whilst maintaining integrity. Streaming data storage is optimised to better read and emit data (rather than quickly write). This data will be emitted at high volume in a continuous, incremental manner. The data storage therefore needs to be high volume, but not operation will not be as frequent. Both have varying use rates, though transactional systems typically have higher utilisations.

3.1.2 Extended product and systems approach

3.1.2.1 Presentation of server and data storage operation environments

Servers and data storage products can be operated in multiple different location types. These types of premises will influence the operations of the devices, their lifecycle and maintenance, along with the supporting infrastructure.

The two main categories to be aware of are data centre premises and distributed IT (also known as 'on-premise' datacentres or 'embedded data centres'). Distributed IT, also referred to as "closet IT" or "embedded IT", will host servers and data storage products in locales not dedicated to their use, such as a side room in an office, or a closet. These systems rely on the infrastructure facilities of the rest of the building, meaning that their usage and energy patterns may not be as closely monitored.

Data centres are premises dedicated to the operation of servers and data storage products. These will have supporting infrastructure designed for these devices: HVAC, power supplies, network equipment.

Data centres can be broken down into three different general types of premises: enterprise, managed services, colocation and hypercloud.

Enterprise data centres are built, owned and operated by companies, and optimised for their end users.

Managed services data centres are managed by a third company (or a managed services provider) on behalf of a company. The company leases the equipment and infrastructure instead of buying it.

A colocation data centre provides a space in a datacentre for a user to rent, where they can install their data equipment. The colocation data centre hosts the infrastructure: building, cooling, bandwidth, security, etc., while the company provides and manages the components, including servers, storage, and firewalls.

Hypercloud centres are off-premises form of data centre, where data and applications are hosted by a cloud services provider such as Amazon Web Services (AWS), Microsoft (Azure), or IBM Cloud or other public cloud provider.

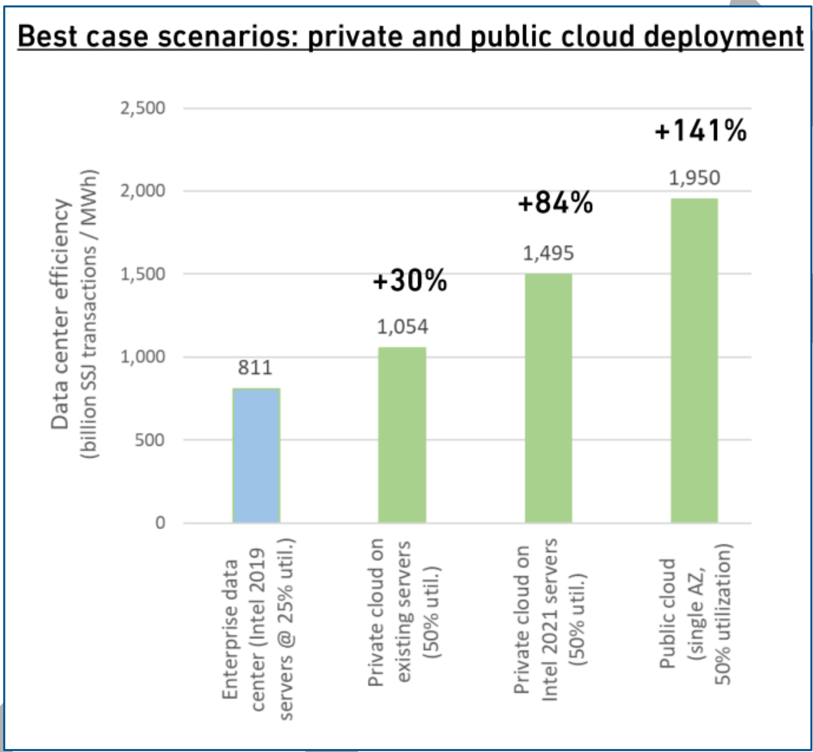
The distinction between these premises is required as the more specialised systems are most likely to be more efficient. Indeed, data centres with their dedicated power, connectivity and cooling facilities are likely to have the more efficient systems in that manner (more information on this developed in Section 0).

Locations such as enterprise datacentres or hypercloud centres, where the equipment is run by one entity, are also more likely to increase efficiency, as the server operation will be optimised, through IT equipment efficiency (improved server performance) and IT operational efficiency, such as using the latest by prioritising workloads to more efficient units (e.g. virtualisation). This is evidenced by Microsoft

Cloud facilities being between 22 and 93 percent more energy efficient than traditional enterprise data centres, depending on the specific comparison being made.[13]

The Uptime intelligence Institute has also exemplified these improvements in their analysis in Figure 3.6. This shows how the shift to a cloud system, where utilisation is increased from 25% to 50%, can result in an increased efficiency of 30%. Furthermore, these gains are compounding when shifting to newer, more efficient equipment, from Intel 2019 to Intel 2021, resulting in an increase efficiency of 84%. Moreover, this increase in utilization and server energy efficiency would also result in decreased space and IT equipment footprint by up to 50%.[14]

Figure 3.6 Model of efficiency gains through cloud deployment[15]



Stakeholder feedback indicate that the current trend is that users are currently migrating their primary operations away from cloud providers back to private enterprise structures. This drive is primarily driven by safety concerns. However, increases in the need for processing or storage capability, beyond core delivery operations, would be purchased on the cloud.

3.1.2.2 Discussion on consolidation and purchasing decisions

As quoted in the previous section, the workload of a server is something which can be redistributed to maximise asset utilisation and minimise the energy and IT equipment costs. This knowledge can be missed by a datacentre when equipment updates occur. Users will typically update their fleet of server equipment on a regular basis in relation to their age, rather than if they have technically failed. This

---

[13] The carbon benefits of cloud computing- A study on the Microsoft Cloud in partnership with WSP, 2020, [Download The Carbon Benefits of Cloud Computing: a Study of the Microsoft Cloud from Official Microsoft Download Center]

[14] Transactions per megawatt-hours: Keys to increasing data centre efficiency, Uptime intelligence, 27 June 2023

[15] Transactions per megawatt-hours: Keys to increasing data centre efficiency, Uptime intelligence, 27 June 2023

regular update ensures that the latest more powerful, and efficient, servers are used. However, users often will purchase replacement servers on a 1-to-1 basis. This buying behaviour is driven by a concern over avoiding failure risks. Yet this means that although the new equipment is more powerful, the same number of servers are installed. If the workload hasn't changed, this results in a lower utilisation rate on a per server basis.

Upon server fleet update, there is an opportunity to consolidate workloads. With new, more powerful servers installed, then the same amount of work can be consolidated onto fewer servers, resulting in higher utilisation rates for the servers (hence a higher energy efficiency for the work delivered), and lower IT equipment costs.

This approach can also be applied to storage equipment, where upon refresh, the storage footprint can be reduced by 25-50%. This can be done with higher capacity drives, and the use of capacity optimization methods (such as data deduplication and data compression).

In addition, to purchasing new equipment there is the option for a data centre operator to purchase refurbished systems. The exact size of this market in the EU is not clear and the study team is not aware of any research that has been done to look into this purchasing decision.

3.1.2.3 Explanation of the business models for data services

Servers and data storage products provide computing and data services to users. The base market for a user to have these services is to own the equipment themselves to deliver the work. However, the market for these services has now evolved. Rather than purchasing equipment, one can now purchase computational capacity directly. These are covered in three main groups: IaaS, PaaS and SaaS.

IaaS or Infrastructure as a Service refers to a system where one can hire the control over a virtual machine, with set storage and workload capabilities. This virtual machine can be accessed anywhere, and the contract would specify the storage and workload capability provided. These parameters can be flexed if the user finds they have an increase in their demand needs. This market structure is the closest thing to a remote data centre for business users.

PaaS or Platform as a Service relies on cloud computing to ensure that a particular application, or software, can be developed and hosted on the web. This can sometimes also be referred to aPaaS (application platform as a service). This market can be used to host a website online, without the need to have dedicated equipment to the task.

SaaS or Software as a Service is often called web services. This makes applications available to the end user via the internet. This could be for example through a subscription service, and would be the one which users are the most familiar with, and applies to services such as: Microsoft 365, Gmail, OneDrive, Dropbox, Netflix, etc.

The relevance of these different data service market models is that these may each result in more or less effective use of hardware. For example, IaaS are often under-utilised, with customers over-purchasing device capacity, meaning that utilisation is likely to stay around 20%. The lower utilisation rates also results in lower efficiencies for the work delivered.

For PaaS, the advantage is that one can develop the software or application, and not have any knowledge of the IT equipment required. The PaaS provider would then need to determine the machines required to provide the adequate computational power and data storage capacity. The PaaS provider can optimise this service over multiple servers to ensure a higher efficiency.

3.1.2.4 Wider systems considerations

Enterprise servers and data storage products have multiple impacts during their use phase. As we described in 3.1.1 and 3.1.2, their activity requires them to consume electricity as they operate. However, this consumption is only the first impact during their use phase, the other is the impact of the auxiliary infrastructure ensuring the appropriate working conditions for the IT equipment. This includes keeping the equipment cool at the right temperature, ensuring the power supply is of good quality (providing protection from any power surges), and providing adequate telecoms connection.

The efficiency of the IT equipment is doubly important as whatever electricity is consumed by the server is eventually converted into heat. This means that data centres create a lot of heat, requiring cooling facilities to ensure the equipment is kept at the right temperature.

The industry has developed the Power Usage Effectiveness metric to monitor how efficient the supporting cooling systems are. This is monitored under ISO 30134-2 and is calculated as the ratio of the total data centre energy consumption over the total IT equipment energy consumption. This provides a figure which has a theoretical minimum of 1.0, which is when all the datacentre consumption is equal to the IT equipment energy consumption only.

Table 3.3 shows how the energy consumption of the ICT sector in the EU27 has developed and is expected to the develop by 2025. Although server and storage consumption has gone up in real terms, the total data centre consumption has gone down in 2020 driven by the gains from improved facility infrastructure capabilities. These gains were pushed by a voluntary agreement of the industry to reduce the PUE figure, which has gone from 2.1 in 2010, to 1.46 in 2020 on average. This trend is expected to continue to reduce to 1.3 in 2025. The BAT for PUE is currently at 1.1, with some specialised data centres capable of going lower.[16]

Table 3.3 ICT Electricity Use EU27[17]

	2010	2015	2020	2025
Servers	18.66	18.66	22.05	27.24
Storage	1.80	1.80	4.35	4.45
Networks	0.53	0.53	0.74	1.06
Facility infrastructure (Cooling, Power protection, etc.)	23.74	23.74	12.4	10.07
Total Data Centre consumption	44.73	44.73	39.54	42.82

[16] The Idle Coefficients, EDNA 2021

[17] VHK and Viegand Maagoe. (2020). ICT impact Study.

	2010	2015	2020	2025
PUE	2.1	2.1	1.46	1.3

The gains in the PUE figure have been reached with investments into more efficient cooling units, fans, optimised ventilation layouts, heat recovery systems, improved power supplies and facilities lighting.

However, it should be noted that the PUE metric is only effective as a comparison to the IT equipment consumption. The PUE metric does not encourage the reduction in the IT equipment. Therefore, other metrics are required to ensure that the entire data centre system becomes more efficient.

Another considered metric is the WUE, Water Usage Effectiveness. This has been developed for data centres with regards to their consumption of water. This is because some cooling systems are reliant on water evaporative technologies. This consumption of water is of concern in water scarce regions, notably in western USA.

Beyond the scope of the PUE, users are considering the recuperation the heat generated by IT equipment under waste heat recuperation techniques. In section 2.12 of the Phase 1 report, we detail how this heat is considered “low grade” as the liquid recuperated from the cooling systems will reach a maximum of approximately 30°C. This is because the data centres aim to keep their premises at the low end of the ASHRAE A1 recommended temperature conditions 18 – 27°C (the justification to keep this temperature has been made in 3.3.2). This limits the direct application of this heat to residential, commercial areas or agricultural facilities: providing direct floor heating to residential or commercial buildings using 30°C water temperature (offices, houses, common spaces, etc.) or within agricultural facilities (greenhouses, fish farming). This requires for these applications to be close as the transport of such low-temperature water would incur losses as the distance grows. For other applications, the heat can be upgraded to higher temperatures through heat pump systems.

Waste heat recovery on datacentre could be a sustainable solution to heating, however it needs to be appropriately monitored in order not to encourage the development of inefficient datacentres. Although increasing the datacentre operating temperature would allow for more waste heat recovery applications, this temperature is recommended to be kept under 27°C as detailed in 3.3.2. as it would increase inefficiency in the total system. Another solution is to develop direct liquid-to-chip cooling systems which could recover heat better and operate at higher temperatures. These are discussed further in Task 4, under 4.1.3.4.

3.2 System aspects of the use phase for ErPs with indirect energy consumption effect

This part of the MEErP is not relevant for servers and data storage products. All indirect effects are investigated through the systems approach section.

3.3 Maintenance, Repairability and End-of-Life Behaviour

3.3.1 Product use and Stock life

The product use and stock life are defined within the MEErP as the time between purchase and disposal. The technical lifetime is defined as the time the device will last without need for repair. The economic lifetime is defined as the time the device will be used before it is replaced, which may be shortened from the technical lifetime by early replacement (due to newer more performant devices) or extended through repair.

As servers occupy valuable space in data centre facilities, it is assumed that there is no additional time lag between the end of the operational life of the asset and the disposal by the customer.

Table 3.4 indicates the estimated lifetimes for the servers and data storage products. This has been recovered from Table 2.6 of the Task 2 report.

Table 3.4 Average lifetime, by type of equipment[18]

Equipment type	Average economic lifetime (in years)	Average technical lifetime (in years)
Rack-blade, rack-mounted, tower/ standalone and multi-node servers	3 for lease 3 to 5 for primary users 5 to 7 for secondary user	5 - 12
Mainframe servers	7 - 15	20
Data storage products (HDD, SSD and hybrid drives)	5 - 7	HDD: 10-15 SDD: 7-10 Storage shelf: 15-20

Within section 2.7 of the Phase 1 report, the study team also indicate how providing an information sheet with the technical lifetime of the asset could be beneficial. Indeed, this could serve to increase the economic lifetime of the asset.

3.3.2 Good & bad practice in product use

Good practice for servers and data storage products can be broken down into 3 sections: maintaining appropriate ASHRAE operating conditions, providing power surge protection and maintaining firmware updates.

ASHRAE sets out the operating conditions of temperature and humidity for servers and data storage equipment. These operating conditions are set out in Table 3.5.

[18] According to DIGITLEUROPE in Lot 9 Ecodesign Preparatory study. 2015

Table 3.5 ASHRAE Operating condition classes[19]

	Dry bulb temp °C		Humidity range, non-condensing			
Operating condition class	Allowable range	Recommended range	Allowable range	Recommended range	Max dew point (°C)	Maximum rate of change (°C/hr)
A1	15- 32	18-27	– 12 °C Dew Point (DP) and 8 % relative humidity (RH) to 17 °C DP and 80 % RH	– 9 °C DP to 15 °C DP and 60 % RH	17	5/20
A2	10-35	18-27	– 12 °C DP and 8 % RH to 21 °C DP and 80 % RH	Same as A1	21	5/20
A3	5-40	18-27	– 12 °C DP and 8 % RH to 24 °C DP and 85 % RH	Same as A1	24	5/20
A4	5-45	18-27	– 12 °C DP and 8 % RH to 24 °C DP and 90 % RH	Same as A1	24	5/20

The recommended range indicates the conditions in which a server or data storage product should be operated in. The allowable ranges can be used but are intended to be temporary operating conditions.

Within the Phase 1 report the study team have discussed how the energy consumption of a data centre can be managed by maintaining the set temperatures within the server rooms. In Section 2.10 of the Phase 1 report, we recommend not to increase servers operating temperatures above 27°C, as this can result in an increase in total data centre consumption caused by an increase in individual server fan energy consumption. This follows the advice from ASHRAE which shows how for every degree increase in the air inlet temperature from 17.7°C, 4% can be saved on cooling costs. However, beyond 27°C, although cooling costs continue to decrease, and PUE value decreases, the energy consumption of the IT equipment increases, resulting in a total data centre energy increase.

The datacentre company Equinix typically run their colocation datacentres at 23 degrees Celsius and has been moving operation towards 25 degrees Celsius without issue. In 2023, the company is now moving to trial a colocation datacentre at 27 degrees Celsius, which should be safe and within the ASHRAE recommended range. The colocation industry cannot push beyond the ASHRAE range, as their

[19] Commission Regulation (EU) 2019/424 of 15 March 2019 laying down ecodesign requirements for servers and data storage products

servers are owned by others, and hence they cannot take too many risks.[20] However, there are efforts to push the operating temperature range beyond this, notably for Meta who has reportedly operated a datacentre at 32degrees Celsius. This is possible for Meta as they are operating their own servers.[21] Within section 2.7 of the Phase 1 report, the study team also indicate how operating conditions could be better labelled for datacentre operators, along with ASHRAE guidance. This can ensure that IT equipment operating temperature is not set too high (possibly damaging the asset) or too low (resulting energy loses) by the datacentre operator.

Surge protection for servers can be provided to ensure that electrical power supply is stable and does not damage equipment. Servers are equipped with internal power supplies in order to convert supplied alternative current to direct current, which provides some protection to the product. Additional grid surge protection can be provided by installing an uninterruptible power supply (UPS) at a data centre, in an on-line protection configuration. These UPS will provide filtering to smooth surges, spikes and dips in the power supply, and will also provide short-term power to critical loads to ensure alternative supply can be turned on in time.

Lastly, firmware updates are required to ensure that servers and data storage products continue to be operated in best conditions. The current Ecodesign regulation 2019/424 requires for these to be made available for products for a set period of time after being placed on the market. Updating these will ensure that product operation is appropriate.

3.3.3 Maintenance practices

Maintenance practices for servers and data storage products are typically covered by disk cleanup and scans (to detect potential for hard drive crashes), cleaning tape drives, monitoring fans and temperature systems performance (which is critical for CPU operation) or upgrading drivers and firmware.

The frequency is dependent on the components and sub-systems considered, the utilisation rates and the operating conditions (temperature and humidity). Hard disk drives, power supplies and memory components have a higher rate of failure and would therefore be maintained more often. Similarly, devices under high utilisation or operated at extreme temperatures are more likely to fail.

Maintenance rates can be included as part of vendor contracts between vendors and consumers. Most manufacturers would provide a three-year warranty in purchase prices, but some may go down to only one year. Third party maintenance is also a common practice in the EU, this entails that the user sources the up-keep of their products internally or via an engineer who does not work for the manufacturer. It has been communicated with the study team that third party maintenance has a substantial contribution to the overall market as users seek more cost-effective measures to extend the life of their products.

3.3.4 Repairability

As mentioned under maintenance practices, vendors may provide repair services as part of the original sales contract. The components with higher rate of failure would be replaced. This is mainly the case for memory and storage components. Other

[20] Is Raising Temperatures in Data Centers Good for Hardware? (datacenterknowledge.com)

[21] Meta thinks it has a great new way to save water in its data centers | TechRadar

elements in a server which can fail are the fan, CPUs and power supply, there tends to be multiple of each of these components therefore, if one fails the systems tends to remain functioning and a repair will likely be undertaken. However, these components can all be replaced. it is noted that servers have a much shorter economic life expectancy than technical one, therefore faulty products are more likely to be disposed of and upgraded rather than repaired by users. Repair or part recuperation is most likely to occur for servers during second-hand use by refurbishers.

Data from Intel shows that their HPC server fleet has an overall annualized failure rate of <1.37%, with only 1.22% in the first 4 years, and up to 1.56% in the year 4 to 8 of life. The main component failures seem to be from the PSU at 0.32% across the 0 to 4th year and 4th to 8th year brackets; and the motherboard which has a high rate of failure in the first 4 years of 0.55%, then subsequently dropping to 0.16%. Drives were third placed with failures of 0.18% across their lifetime. All other components had annualized failure rates below 0.1%.[22]

For data storage products, components are more likely to be repaired and replaced. Hard disk drives (HDD) are easily replaced with measures in place to ensure failures are managed as business-as-usual metrics, with repairs proceeding swiftly without any change in the quality of service delivered.

Solid State Drives (SSD) have a lower rate of failure than HDD, estimated failure rate after 2 million hours, versus 1.5 million hours for HDDs.[23]

Common warranty for drives is five years (43.800 hours).

Stakeholders indicate that spare parts are usually available for five years after the product is manufactured. The study team has recommended that this should be included in the Ecodesign regulation to ensure higher repair rates. This recommendation has been discussed in more detail within Section 2.8.4.4 of the Phase 1 report.

3.3.5 Second-hand use

Server products are generally disposed of before the end of their technical lifespan, which means they can be recuperated and repurposed for a second life. However, these practices are not standardised across the industry, and users have security concerns around their data, making it unclear to review what is the rate of recuperation and reuse of devices.

As mentioned above, servers and data storage products can be repaired and refurbished to provide a new life on their asset. Most hardware manufacturers have end-of-life facilities to adequately collect, repair, refurbish, recuperate parts and/or recycle servers and data storage products. These services are provided, but data on their implementation rate is unclear. There is also an existing IT Asset Disposal (or ITAD) third party industry who will recuperate these devices and guarantee that any data has been adequately erased, which ensures security for original device users.

Security of data can be ensured through two methods, either with software overwrite (known as “erasure”, which is distinct from “deletion”), or physical destruction of the drive. Software overwrite can be safe, notably using rigorous erasure services, as guaranteed by standards such as ADISA. However, there is a general mistrust in

[22] IT@Intel: Green Computing at Scale, August 2021

[23] Storage Review: SSD vs HDD. http://www.storagereview.com/ssd_vs_hdd

industry around these practices, meaning that software overwrites are used generally for internal reuse or non-critical data. Physical destruction is generally used for all other purposes, with customers insisting on physical destruction to ensure their data cannot be recovered. Recently there has been a clear movement to accept Secure Erasure which provides certificates of the removal of data, this ensures that destruction remains an alternative and only for those drives that have no reuse option.

Where some products may have been disposed of, as their users have upgraded to more performant products, recuperators will ensure data eradication, and either reuse, or refurbish devices by upgrading particular components (such as by replacing for a more powerful CPU). The product can then be re-sold, usually at a lower price. Stakeholder data shows how in Europe, the components recuperated and reused are: RAM memory (47% of units by weight) and processor (39.7% of units by weight).

3.3.6 Recycling, collection and disposal

Servers and data storage products are in the scope of the WEEE Directive. WEEE (or Waste electrical and Electronic Equipment), sets out communal collection and recycling targets under "extended producer responsibility". This means that producers must take responsibility for the equivalent amount of waste they place on the market, and finance dismantling, depolluting, recycling and disposal of WEEE. Individual manufacturers/vendors may go beyond these requirements. This WEEE metric is measured based on total mass collected and recycled, and therefore does not prioritise the recovery of critical materials, but rather bulk components. Eurostat data tracks the recuperation rates for WEEE streams as a whole but does not distinguish between different electronic devices.

There is a clear boundary of the jurisdictions of WEEE and Ecodesign products with regards to reuse and recovery. Until a final decision on the recycling of a product has been made then its life-cycle continues and it remains under the jurisdiction of the Ecodesign regulation. Therefore, if a product is re-used rather than recycled the Ecodesign requirements remain applicable to it. However, when a product reaches end of life, it would be defined as "waste" which brings jurisdiction over to the WEEE. After its listing as WEEE some components of the product can still be re-used thus labelling it as REEE (recycled electronic and electrical equipment) and under the procedures associated with this.

Most hardware manufacturers have end-of-life mechanisms to recuperate products when customers no longer require these (whether due to failure or upgrade). These mechanisms will aim to reuse (or resell) the product, repair (or refurbish), recuperate functioning components and recycle remaining materials. Due to the modularity of the products, many of the components may still be functional and be removed directly at the customers location. However, the rate at which this is done is unclear.

Critical raw materials are of particular concern for these products. In particular, the recuperation of strong permanent magnets used in HDDs. These strong magnets use rare earth elements, which are critical materials. The use of a shredder is not appropriate as the Neodymium magnets would crack and stick to the shredder itself. Therefore, manual recovery is required for adequate material recovery. They can be recovered manually by dismantling the HDDS but requiring special fine mechanical tools to dismantle the device, whilst careful not to have the strong magnetic force not affect the tools themselves. This recovery is difficult and not always commercially viable. As IT products can be categorised under WEEE norms, it is the total mass of product recovered which is legally required, hence simpler bulk materials (such as the metal frames) are prioritised for recovery.

Table 3.6 below shows the breakdown of material usage, recycling, energy recovery and landfill for enterprise servers provided by a European waste disposal stakeholder in 2023. This data shows how 81% of collected server mass is either re-used or materially recycled. Including waste heat recovery, up to 99% of server materials value is recovered. Only 0.16% of material by mass is incinerated or sent to landfill.

Table 3.6 Inputs in the end-of-life phase of collected enterprise servers[24]

	Plastics	Metals	Electronics	Misc.
Mass ratio within server	0,97%	67,88%	30,88%	0,28%
Re-use	0%	0%	1%	0%
Material recycling	98%	98%	43%	50%
Heat recovery	0%	2%	56%	0%
Non-recovery incineration	0%	0%	0%	50%
Landfill	2%	0%	0%	0%
Total	100%	100%	100%	100%

Table 3.7 shows the inputs in the end-of-life phase of the storage systems assumptions from the 2015 Preparatory study.

Table 3.7 Inputs in the end-of-life phase of storage systems[25]

	Plastics	Metals	Electronics	Misc.
Re-use	25%			
Material recycling	5%	70%	50%	68%
Heat recovery	69%	0%	24%	1%
Non-recovery incineration	0.5%	0%	0.5%	5%
Landfill	0.5%	5%	0.5%	1%
Total	100%	100%	100%	100%

[24] Feedback from European electronics waste disposal stakeholder 2023

[25] Source: Preparatory study for implementing measures of the Ecodesign Directive 2009/125/EC, DG ENTR Lot 9 - Enterprise servers and data equipment, Task 3: User, July 2015: Final report, bio by Deloitte & Fraunhofer IZM

3.4 Local Infrastructure

3.4.1 Energy

The overall energy consumption and related energy costs are an extremely important factor that businesses in the EU consider. Especially given the recent spike in energy prices in the EU caused by the Ukraine War.

Geographic location will impact the reliability, availability and source of electricity that is provided to data centres situated in the EU. For example, data centres within Northern Europe have a particular advantage because the lower annual temperatures mean that less energy is required for maintaining. Lower ambient temperatures allow for free cooling, which reduces the overall cooling capacity needed. This will improve the data centres thermal efficiency meanwhile, due to the temperate, moderate to high rainfall climates in Northern Europe, there are more renewable energy sources from wind or hydropower, which are often less intermittent than solar power. There have been instances where free-cooling utilisation rates of greater than 50% are achieved in temperate regions. This is achieved when the raised floor temperatures are moved to the ASHRAE A2 Standard, and the cooling system is run with optimisation software.

However, it should be noted that despite the availability of free-cooling in northern regions of Europe, there is still a need for data centres to install compression refrigeration system. This will ensure that when the ambient temperatures are warmer and above the data centres set point temperatures, it is still able to maintain suitable environmental conditions inside to ensure full functionality of the servers and data storage products. Therefore, preserving this set point temperature and environment in the data centre requires significant amounts of energy, up to 40% of a data centres total energy use in warmer climates[26]. No matter the location, all data centres will need to cover this redundancy to ensure the facility is able to function all year round. Climate change is also having a significant impact on data centres across the EU, with more and more centres requiring active cooling as a result. In order to meet businesses' net zero goals, some data centre operators may choose to buy for example, solar panels in order to reduce their facilities carbon emissions.

The transparent metering of energy consumption is a necessity for the businesses in the EU. Therefore, energy metering for data centre facilities is encouraged by using the EN 50001 standard. However, this can be dependent on the equipment ownership and business models within the data centre, as often these locations are shared with other businesses.

Electricity supply for servers and data storage products is usually ensured through UPS units. UPS units will provide filtering to smooth surges, spikes and dips in the power supply, and will also provide short-term power to critical loads to ensure alternative supply can be turned on in time. This ensure that no damage or data is lost if there is a power surge or loss within the data centre. In addition, grid surge protection can also be provided by installing an UPS unit at data centres, in an on-line protection configuration. Usually, UPS products supply electricity via batteries. However, rotary or dynamic UPS products can utilise stored kinetic energy through a flywheel to provide electricity in case of a power supply failure.

Surge protection for servers can also be provided to ensure that electrical power supply is stable and doesn't damage equipment. Servers are equipped with internal

[26] https://www.eolitservices.co.uk/2022/02/23/the-environmental-impact-of-our-data-storage/

power supplies to convert supplied alternative current to direct current, which provides some protection to the product.

3.4.2 Water

Water or another liquid can be used for cooling at both the product and system level. Liquid cooling of servers typically provides higher energy efficiency when compared to air cooled systems and could enable the driving of data centre industry forward. Manufacturers state that liquid cooling allows optimum energy use within the IT suite so that more power drives the applications on the servers rather than the cooling systems[27]. At the same time there are free air-cooled data centres in the US and adoption of free air cooling (at least part time) is preferred at Government data centres where possible.

Proposing requirements around proper deployment or quality of water-cooled solutions could be the way forward, however as these systems are still at their infancy, comprising of less than 5% of the server market, to impose requirements on water cooled solutions in a regulatory setting requires further review and analysis. In addition, liquid cooling introduces several new considerations including safety, proper fluid handling and maintenance (to avoid mould and/or corrosion etc.) and proper system design to minimise long term problems that specifically arise with the complications of liquid cooled solutions. Liquid cooling recommendations for this review study have been put forward in Section 2.11 of the Phase 1 report.

Product level liquid cooling

Indirect cooling is a system of liquid cooling where no liquid flows through the servers, but rather to the rack. These are usually set up as rear-door hybrid cooling, where a cooling liquid is set to flow at the back of the server rack. This allows for air to be cooled as is it removed from the server by the liquid flow. This liquid flow would then need to be cooled elsewhere, using technologies such as a chiller, or free cooling. These systems have a low to medium efficiency, require outlet water temperature to be under 50 ºC, are simple to install and have a medium total cost of ownership.

Another technology is immersion or submersion cooling. These systems have the entire server submerged in liquid which grants a high efficiency due to the liquid heat capacity. However, the installation is complex, and servicing the servers is difficult.

A direct-to-chip liquid cooling system will have liquid cooling flow provided directly in thermal contact with the hottest components of a server. These are highly efficient and allow for higher outlet temperatures, 60 – 75 ºC, which allows for more efficient waste heat recovery applications and the use of additional free cooling. These have low maintenance and low total cost of ownership. There are some installation concerns, but the main drawbacks are in the design of the specialised server to accommodate for the liquid heat exchange.

Since direct-to-chip liquid cooling can enable waste heat re-use above 60 – 70 °C, this waste heat could be used directly without the need for a heat pump cycle. Examples of such technologies and companies are Denmark based Asetek, and Canada based CoolIT.

[27] Solving Data Center hunger, EiBi, issue June 2022

System level liquid cooling

At a system level the utilisation of water in cooling systems is very common at data centres. However, the availability and efficiency of these systems can vary dramatically depending on geography.

Liquid cooling options discussed at the product level result in a more efficient cooling system, compared to air-based cooling solutions. Adopting more of the solutions described above would mean that fewer data centres have to use liquid cooled evaporative towers. Therefore, harnessing the high latent heat capacity of water to cool data centres, can mean that there is a lower water usage effectiveness score. This is major issue in areas where water scarcity is high and is vital to ensure the sustainable use of data centres in these areas. It is especially important with impacts of climate change beginning to mean more areas in the EU are experiencing longer drought periods. Particularly in regions where they usually do not occur, in the more temperate regions. Since data centres here are more likely to be less equipped for these types of environmental conditions.

In addition, to direct-to-chip liquid cooling providing cooling for the server, as described in the product level section. This type of liquid cooling would also provide a more efficient waste heat recovery solution. This could lower the data centre cooling energy consumption from a PUE of 1.2 to 1.12.

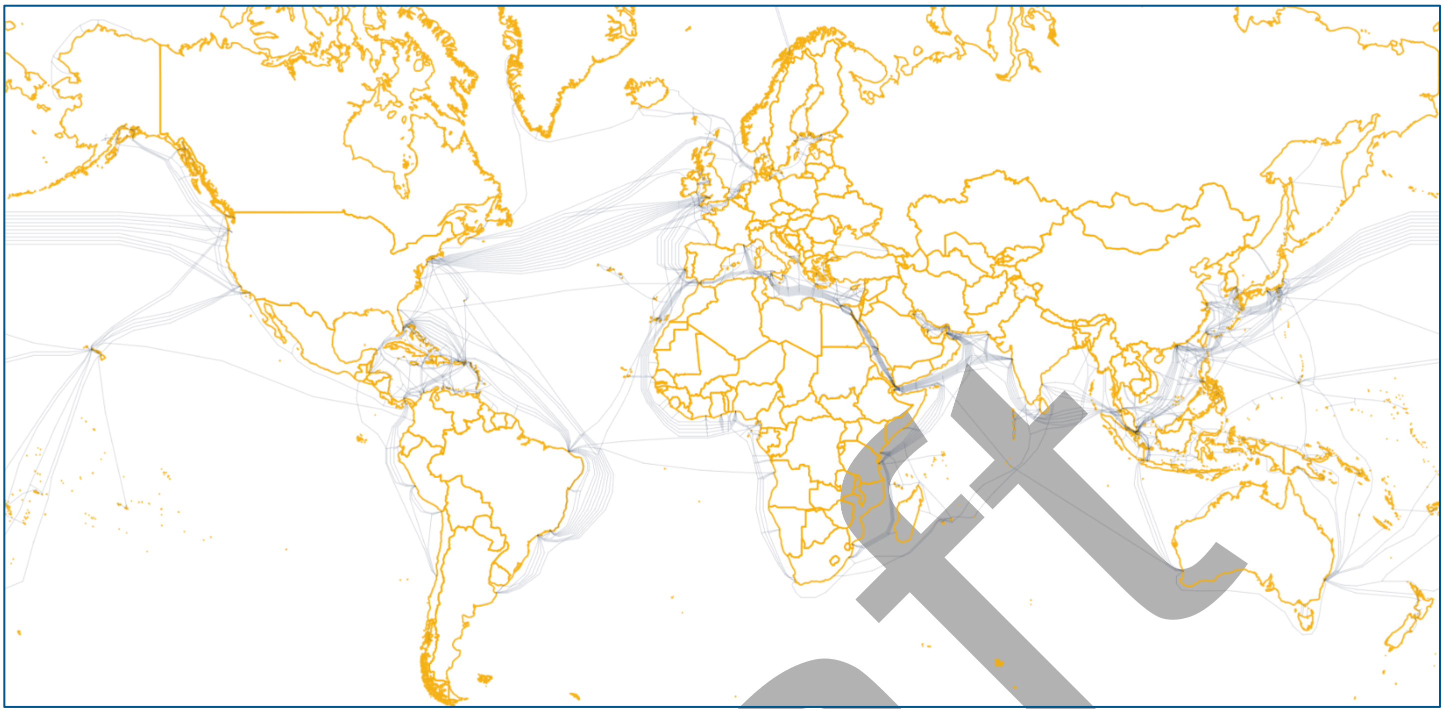
3.4.3 Telecommunications

A fixed-line broadband network access is crucial to the functionality of a data centre. This is has meant that over the years data centres have been specifically built closer to connection nodes or activity points to ensure connectivity is maintained Connection nodes are where intercontinental cables connect to the continent. For example, Figure 3.7 illustrates that there are many cables that go between the US and the UK, hence some may connect their data centres closer to that node. In this manner they can benefit from some of the higher bandwidth connections available closer to the connection nodes.

The previous preparatory study mentioned that wireless networks had not been implemented due to security and interference concerns. For servers and data storage products which fall within the scope of this Regulation this aspect has not changed. There has yet to be any serious development in connecting data centres wirelessly, due mainly to large amounts of data being processed. Therefore, larger data centres with will be connected via telecommunication cables. This continues trend is highlighted in Figure 3.7 which illustrates how widespread these cables are, connecting the EU with the North and South America, Africa, Asia, Middle East and Australia. The reliance on the use of the submarine telecommunication cables is emphasised by the fact that over 99% of international internet and telephone traffic passed through submarine cables in 2019[28].

[28] https://subtelforum.com/submarine-telecoms-industry-report-10th-anniversary-issue-now-available

Figure 3.7 Submarine telecommunications cables across the world[29]



Servers are heavily tied to the financial sector, which relies on servers for processing financial information such as updated stock prices. Therefore, this time sensitive data centre activity, such as this will need to be closer to connection nodes to ensure that data is received as quickly as possible. For this reason, data centres will be close to where the financial operations are taking place, to benefit from fastest response rates.

3.4.4 Installation

Installation of servers and data storage equipment is usually part of the service provided by the equipment vendor, typically this is charged as an additional fee. Task 2 describes this in more detail, with the contract often including warranties that can be between three to seven years depending on the manufacturer.

3.4.5 Physical environment

Servers and data storage products are installed and operated within server rooms. Server rooms are the individual rooms that servers operate within, at a data centre. The data centre represents the entire infrastructure in which a server room can be found in. Therefore, for data centres there are many more considerations to take into account, such as heating and cooling requirements, and power supplies and many more.

Servers will tend to be hosted in two types of locations: distributed IT (where servers are used in offices or other buildings which are not dedicated for servers), and data centres (which are specialised locations for server usage, with dedicated power and HVAC services).

[29] ITU: Committed to connecting the world (https://bbmaps.itu.int/bbmaps/)

The bulk of servers are operated within data centres. The trend has been to consolidate data operations in data centres as their operation is optimised, allowing for better tracking of energy usage across devices and support systems. These buildings can also be reinforced in terms of data connectivity and power availability.

The physical environment of a data centre is usually defined by the maximum energy density that can be safely handled. As discussed in Section 3.4.1 the location of the data centre has a significant impact on not only the energy consumption but also the environmental impact of a data centre. Geographic location will determine the data centres access to colder ambient temperatures, water, and renewable energy sources.

Not only will geography impact the amount of electricity consumed by a data centre, but since each EU Member State has its own unique electricity mix this means that the local physical geography of each member state plays a pivotal role in a data centres overall environmental impact[30]. For example, in 2022, 38.7% of the EU's electricity was generated from fossil fuels, such as oil, natural gas, and coal[31]. The geographic variation is highlighted by the fact that Malta generates 87% of its electricity from fossil fuels, whereas Denmark produced only 21% of its electricity from fossil fuels in 2022[32]. This is because Denmark harnesses a lot of its energy from offshore wind farms in the North Sea, with 79% of its electricity generated from renewable sources. Innovative solutions to reduce data centres energy consumption are seen across the EU. For example, in Nordic countries, district cooling facilities are installed to take advantage of the cold temperatures experienced here.

[30] https://iopscience.iop.org/article/10.1088/1748-9326/abfba1

[31] How is EU electricity produced and sold? - Consilium (europa.eu)

[32] How is EU electricity produced and sold? - Consilium (europa.eu)

Annex 1 SNIA Taxonomy detail

Table A1.1 SNIA Taxonomy detail for Online storage[33]

Attribute	Classification					
	Online 1	Online 2	Online 3	Online 4	Online 5	Online 6
Access Pattern	Random/ Sequential	Random/ Sequential	Random/ Sequential	Random/ Sequential	Random/ Sequential	Random/ Sequential
Connectivity	Not specified	Connected to single or multiple hosts	Network-connected	Network-connected	Network-connected	Network-connected
Consumer/ Component	Yes	No	No	No	No	No
FBA/CKD Support	Optional	Optional	Optional	Optional	Optional	Required
Integrated Storage Controller	Optional	Optional	Required	Required	Required	Required
Maximum Configuration[1]	≥1	≥ 4	≥ 12	> 100	>400	>400
MaxTTFD (t)	t < 80 ms	t < 80 ms	t < 80 ms	t < 80 ms	t < 80 ms	t < 80 ms
No SPOF	Optional	Optional	Optional	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Required	Required
Storage Protection	Optional	Optional	Required	Required	Required	Required
User-Accessible Data	Required	Required	Required	Required	Required	Required

[33] Taxonomy | SNIA

Table A1.2 SNIA Taxonomy detail for Near Online storage[34]

Attribute	Classification					
	Near Online 1	Near Online 2	Near Online 3	Near Online 4	Near Online 5	Near Online 6
Access Pattern	Random/ Sequential	Random/ Sequential	Random/ Sequential		Random/ Sequential	Random/ Sequential
Connectivity	Not specified	Network connected	Network connected		Network connected	Network connected
Consumer/ Component	Yes	No	No		No	No
FBA/CKD Support	Optional	Optional	Optional		Optional	Optional
Integrated Storage Controller	Optional	Optional	Required		Required	Required
Maximum Configuration[1]	≥ 1	≥ 4	≥ 12		> 100	> 1000
MaxTTFD (t)	t > 80 ms	t > 80 ms	t > 80 ms		t > 80 ms	t > 80 ms
No SPOF	Optional	Optional	Optional		Optional	Required
Non-Disruptive Serviceability	Optional	Optional	Optional		Optional	Required
Storage Protection	Optional	Optional	Required		Required	Required
User-accessible Data	Required	Required	Required		Required	Required

Table A1.3 SNIA Taxonomy detail for Removable Media Libraries[35]

Attribute	Classification					
	Removable 1	Removable 2	Removable 3	Removable 4	Removable 5	Removable 6
Access Pattern	Sequential	Sequential	Sequential		Sequential	Sequential
Maximum Drive Count	Not specified	4	≥ 5		≥ 25	≥ 25
MaxTTFD (t)	80ms < t < 5m	80ms < t < 5m	80ms < t < 5m		80ms < t < 5m	80ms < t < 5m
No SPOF	Optional	Optional	Optional		Optional	Required
Non-disruptive Serviceability	Optional	Optional	Optional		Optional	Required
Robotics	Prohibited	Required	Required		Required	Required
User-Accessible Data	Required	Required	Required		Required	Required

34 Taxonomy | SNIA

35 Taxonomy | SNIA

Table A1.4 SNIA Taxonomy detail for Virtual Media Libraries[36]

Attribute	Classification					
	Virtual 1	Virtual 2	Virtual 3	Virtual 4	Virtual 5	Virtual 6
Access Pattern	Sequential	Sequential	Sequential		Sequential	Sequential
User-accessible Data	Required	Required	Required		Required	Required
MaxTTFD (t)	t < 80 ms	t < 80 ms	t < 80 ms		t < 80 ms	t < 80 ms
Storage Protection	Optional	Optional	Required		Required	Required
No SPOF	Optional	Optional	Optional		Optional	Required
Non-Disruptive Serviceability	Optional	Optional	Optional		Optional	Required
Maximum Configuration[1]	12	>12	> 48		> 96	> 96

36 Taxonomy | SNIA