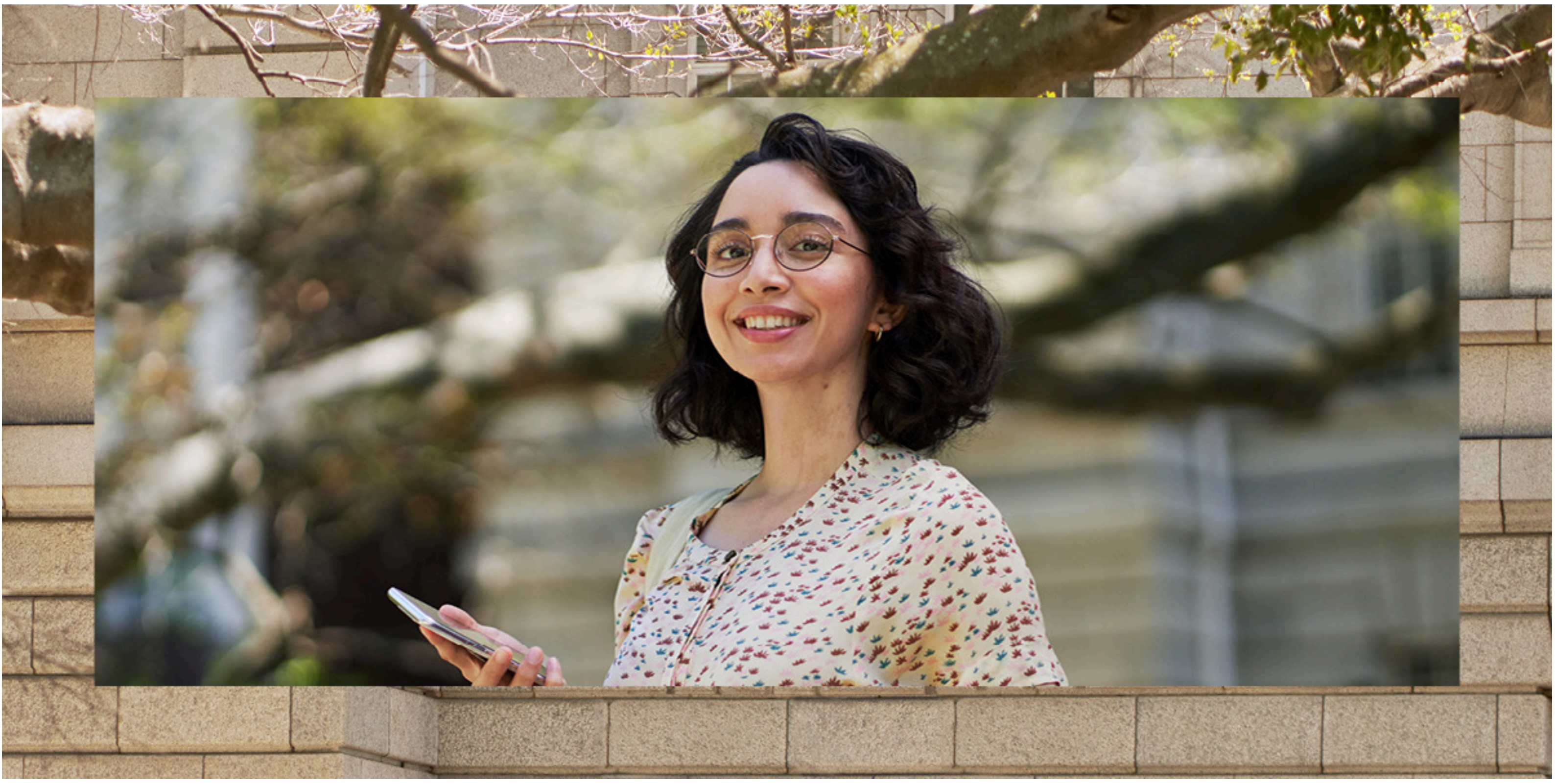
Hewlett Packard Enterprise

Technical white paper

# Efficiency benefits of IT lifecycle strategies



# Contents

## Abstract

This paper aims to determine the optimal cadence for information technology (IT) refresh cycles of mainstream data center equipment, considering both the environmental impact of equipment creation, its impact during use, and related economic cost implications. A case study is performed to evaluate different strategies for compute assets while relevant considerations are discussed for storage and network equipment. Recommendations for lifecycle strategies are suggested together with a conclusion. While many of the same principles might apply, specialized networking equipment and high-performance computing (HPC) is outside the scope of this paper, which discusses mainstream data center equipment: compute, storage, and networking.

### Key takeaways

- Server replacement projects must be combined with workload consolidation onto fewer devices to achieve efficiency benefits.
- Server efficiency ratings are rising, driven by improvements in CPU technology and increase in memory and PCI bus bandwidth.
- Aligning server replacement cycles with technology improvements of CPUs and other components is beneficial for maximizing environmental and economic efficiency over time.
- Efficiency gains for storage are mostly driven by architectural- and software innovations rather than hardware advancements. Data retention strategies, and tiering, play significant roles in driving efficiency.
- Like servers, the network footprint is most effectively reduced by hosting workloads with fewer devices. Refresh cycles are mostly driven by financial depreciation.

## Introduction

An IT lifecycle management strategy can yield measurable benefits in power savings, space reduction, cooling efficiency, IT staff productivity, and innovation opportunities. A 2023 Uptime Institute study reveals that a 2-to-1 server upgrade from servers running Intel® 2019 CPUs to Intel 2021 CPUs can provide a significant 40% improvement in data center energy efficiency within existing facilities without the need to migrate software to a cloud architecture.[1] Fewer physical devices result in a lower management burden and fewer service contracts and software licenses, given that it's licensed per CPU or the total CPU core count decreases. Yet, multiple disruptions often interfere with the initial lifecycle management strategy: acquisitions, divestitures, reorganizations, digital transformations, cloud migrations, and the emergence of new technologies.

Professionals implementing lifecycle strategies should understand the benefits while acknowledging the challenges of operationalizing technology refresh plans. IT lifecycle strategies should also embrace the inherent environmental benefits included by considering the entire product lifecycle. Manufacturing, transportation, usage, and disposal of equipment all contribute to the environmental impact of IT equipment. Optimizing the technology refresh lifecycle needs to be done in harmony with consolidating workloads from existing devices onto fewer new devices where possible, increasing equipment utilization, and rightsizing redundancy requirements. This leads to a reduction of the total number of devices needed, mitigating the environmental impact. Additionally, this paper aims to shed light on the right balance between manufacturing and usage emissions, which is expected to depend on the carbon intensity of the electricity available.

[1] "Efficient servers hold the key to energy efficient data centers," Rakhonen, Uptime Institute, 2023

## Benefits of lifecycle management

Various IT efficiency sources, including manufacturers, efficiency organizations, and consulting organizations, have published resources for considering the total lifecycle costs of IT products to define optimal lifecycles and asset management practices. These resources often propose utilizing the total cost of ownership (TCO) as a primary metric. However, the definition of TCO is often inconsistent, as sources may or may not consider implications for the entire product lifecycle, such as performance, staff implications, and environmental impacts, or do not discuss barriers to adoption. While financial cost is a relevant metric, focusing primarily on cost optimization fails to consider the benefits and opportunities associated with a strategic lifecycle management approach.

There are four primary benefits of optimizing technology refresh lifecycles: power and space reduction; operational efficiencies; residual value return; and reduced risk. These benefits will be discussed in the following sections. Each of these benefits also provides an associated cost benefit.

### Performance, power, and space benefits

**Compute**

The most apparent benefits of server technology refresh:

- Improved application performance, requiring fewer servers, due to associated increased core counts, more work done per core and clock cycle, higher memory/cache densities, more bandwidth on internal buses, and, to some extent, faster clock speeds
- Expected reduced heat generation and associated cooling costs from fewer servers
- Expected reduced data center space and electrical requirements from smaller cooling, electrical, and UPS capacity demanded by fewer servers
- Lowered environmental impacts associated with fewer servers, reduced power demand, and less support equipment

Server performance per watt, known as energy effectiveness, increases with each new generation of processors from AMD, Ampere®, and Intel. One method to assess the improvement in server energy effectiveness is reviewing results from the Standard Performance Evaluation Corporation (SPEC). The SPEC Power committee developed the Server Energy Efficiency Tool, named SERT™ Suite, a next-generation toolset for measuring and evaluating the energy efficiency of servers.

The SERT 2 Suite uses a set of synthetic worklets to test discrete system components such as processors, memory, and storage, providing detailed power consumption data at different load levels to simulate actual use cases.[2] It has been adopted by the U.S. Environmental Protection Agency (EPA) for assessing eligibility for the ENERGY STAR® server program.

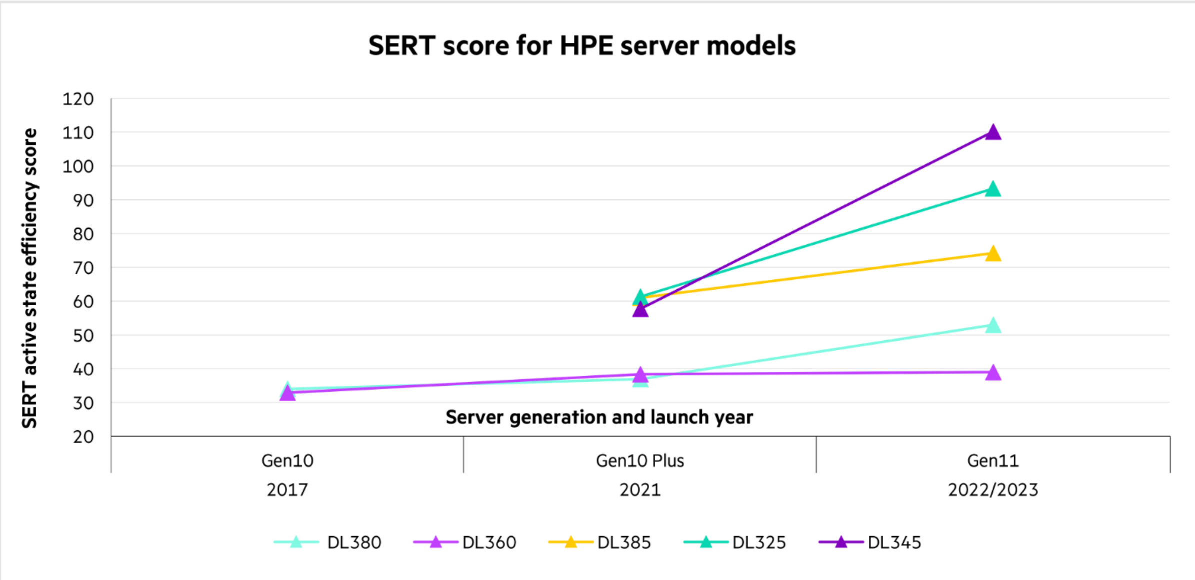



**Figure 1.** SERT scores for recent HPE server models[3] (No SERT 2 Suite tests were done for Gen10 models of HPE DL325, HPE DL345, and HPE DL385.)

[2] spec.org/benchmarks.html#power
[3] energystar.gov/productfinder/product/certified-enterprise-servers/results

The SERT results in Figure 1 do not suggest the 2-to-1 server consolidation opportunity described by the Uptime Institute between Intel 2019 and 2021 CPU models. If the CPU utilization level is maintained, work done per core and clock cycle is assumed to be constant, such a consolidation rate would require a doubling in core count, that is, 100% more cores. The 2021 Intel CPUs have 43% more cores and utilization has increased only 3.8%, which means the 2-to-1 consolidation rate is achieved also due to advancements in system and CPU technology, enabling each core to perform more work per clock cycle.[4] This extra server performance capacity is not fully reflected in the SERT 2 Suite, calling for IT decision-makers to consider more than one source of knowledge before making decisions. The development of CPU core counts can be seen in Table 1.

**Table 1.** Core counts in CPU generations[5, 6]

| | AMD | | Intel | |
|---|---|---|---|---|
| **Year** | **Generation** | **Core count** | **Generation** | **Core count** |
| 2019 | 'Rome' and 'Milan' | 64 | Cascade Lake | 28 |
| 2021/22 | 'Genoa' | 96 | Ice Lake | 40 |
| 2023 | 'Bergamo' | 128 | Sapphire Rapids | 64 |
| 2024 | 'Turin' | 192 | Sierra Forest | 144 |

The advancements in system and CPU technology, rise in SERT scores, and rise in core counts seen in Table 1, suggests an opportunity to achieve significant improvements in energy effectiveness by replacing ageing servers. However, a 1-to-1 server replacement without consolidation efforts will increase the absolute electricity consumption, due to the increase in power draw in newer CPUs and servers. Therefore, to achieve improvements in energy effectiveness, the technology refresh must be combined with workload consolidation, resulting in improved CPU utilization and deploying fewer servers.

Another vital point to consider is the emissions stemming from the creation of servers. Like monetary costs, there is an environmental cost of manufacturing all necessary components, often referred to as embodied carbon footprint. This means that the optimal cadence for server replacements from an environmental perspective depends on the balance between the embodied carbon footprint and use phase emissions, as discussed in the introduction. Use phase emissions are determined by lifetime electricity use and the grid emission factor of generating that electricity. Grid emission factors vary greatly in different geographic locations, with some electricity grids relying heavily on natural gas, oil, and coal while others rely mostly on low-carbon sources such as hydro, nuclear, solar, and wind power.

In their 2023 Q&A, "IT sustainability: Achieving more transactions per megawatt hour", the Uptime Institute investigates the balance between embodied emissions and use-phase emissions.[7] Table 2 shows the results of an HPE case study carried out for this paper, building on Uptime's methodology, addressing further considerations that can be found in Appendix 1.

The following case study considers three cases all spanning over 16 years. The difference between the cases is how often server technology is being replaced: every four, six, or eight years. These frequencies are also the names given to identify each scenario in the columns in Table 2. A graphical representation of the moments in time replacements are made in the three different scenarios is shown in Figure 2 below. Two key assumptions for this study are that a new server generation is assumed to be released every four years, with a 75% increase in workload capacity per server with a 40% increase in energy consumption. The sources of these numbers along with further economic and environmental assumptions are found in Appendix 1.

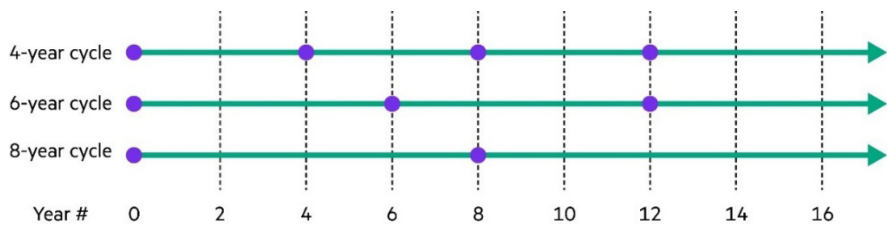



**Figure 2.** Timing of technology replacement points for strategies considered in the HPE case study

[4] Calculation on the basis of the Intel CPUs in question having 28 and 40 cores, which is a 43% increase.
[5] Intel Unveils Future-Generation Xeon with Robust Performance and Efficiency Architectures," Intel, 2023
[6] Confirmed by AMD, June 2024
[7] "IT sustainability: Achieving more transactions per megawatt hour," Jay Dietrich, Uptime Institute, 2023

**Table 2.** Comparing GHG emissions and the cost of different refresh cycle lengths
Refer to Total servers purchased, Total cost and Total emissions mixed grid rows.

| Metric | Unit | 4-year cycle | 6-year cycle | 8-year cycle |
|---|---|---|---|---|
| Total servers purchased | n/a | 109 | 90 | 76 |
| Total HW cost | $ | 654,000 | 540,000 | 456,000 |
| Total embodied carbon emissions | $kgCO_2e$ | 260,401 | 215,010 | 181,564 |
| Electricity consumption | MWh | 4775 | 5177 | 5274 |
| Electricity cost | $ | 716,219 | 776,498 | 791,173 |
| Product support cost | $ | 1,149,818 | 1,412,450 | 1,195,654 |
| Software license cost | $ | 827,869 | 1,016,964 | 981,083 |
| Use-phase carbon emissions clean grid (50g/kWh) | $kgCO_2e$ | 238,740 | 258,833 | 263,724 |
| Use-phase carbon emissions mixed grid (200g/kWh) | $kgCO_2e$ | 954,959 | 1,035,330 | 1,054,897 |
| Use-phase carbon emissions fossil-heavy grid (500g/kWh) | $kgCO_2e$ | 2,387,398 | 2,588,326 | 2,637,242 |
| Total cost | $ | 3,347,906 | 3,745,912 | 3,423,910 |
| Total emissions clean grid | $kgCO_2e$ | 499,141 | 473,843 | 445,288 |
| Total emissions mixed grid | $kgCO_2e$ | 1,215,360 | 1,250,340 | 1,236,461 |
| Total emissions fossil-heavy grid | $kgCO_2e$ | 2,647,799 | 2,803,336 | 2,818,806 |

Replacing equipment often can reduce the electricity demand and use phase carbon emissions but increases the total embodied carbon emissions of the data center. This trade-off is seen clearly when looking at emission results in Table 2 for different grids—as the grid emission factor increases, longer refresh cycles result in higher emissions. Economically, the four-year scenario outperforms the other scenarios. Staff costs due to a higher server management burden and floor space costs are not included in this study. These costs would increase the economic benefit of shorter refresh cycles. Additionally, for those located in grids with emission factors above $100gCO_2e/kWh$, a shorter refresh cycle is the best choice, both environmentally and economically. In situations where the grid emission factor is low and economic considerations are critical, one must consider what weighs more between reducing emissions and reducing costs. For those prioritizing environmental considerations higher than economic cost who are also located in regions with low emission factors, a longer refresh cycle is the best choice, despite being economically unattractive.

An important acknowledgement is the assumption that each new server generation has a workload capacity improvement of 75% compared to the previous generation. The workload capacity improvement will depend on each case and configuration, influencing the results heavily. Therefore, it is recommended for the readers to leverage the methodology in this study and apply it to their situations. Readers are also encouraged to investigate how often technological advancements relevant to them are launched, as this can be more often than every four years. The results in this paper suggest that if electricity is generated by renewable and clean sources, a study should be made to evaluate the environmental and economic trade-offs. If the electricity is generated by fossil sources, a shorter refresh cycle will most likely result in lower emissions and lower costs.

Finally, IT professionals should acknowledge that the most beneficial way to reduce emissions globally relies on the combination of efforts to improve how efficient energy is used and heavy investments in low-carbon energy, not just one or the other. All parties are committed to a tripling in low-carbon energy investments and a doubling of energy efficiency under the Global Renewables and Energy Efficiency Pledge at COP28.[8] As this study shows, shorter refresh cycles reduce the use phase carbon emissions but increase the embodied carbon emissions. The opposite is true for longer refresh cycles, creating a trade-off determined by the grid emission factor of the electricity. In both these cases, efforts to consolidate workloads on fewer devices when replacing infrastructure combined with eliminating unnecessary applications and hardware will reduce both economic cost and total carbon emissions.

### Storage

The quantity of data generated by people and things is growing exponentially while applications are increasingly deployed on hybrid cloud architectures, raising new challenges for enterprise data strategies. To define a storage lifecycle management strategy that will yield measurable efficiency benefits, it's critical to first define a data lifecycle strategy. This first step will better inform the technology refresh cycle for the storage infrastructure.

Refreshing storage technology can bring about improvements in performance, power efficiency, and space utilization. Consideration of the following key points is essential for making well-informed decisions:

- **Data reduction techniques:** Advances in data reduction techniques contribute to more efficient storage media usage. Recognizing that data reduction activities have an associated carbon footprint due to processing requirements, opting for a storage system that allows leveraging these continuous advancements while replacing little or no hardware is likely to result in the lowest environmental footprint.
- **Technology continuum:** Realizing improved operational and power efficiency through a continuum of tape, hybrid, and flash technology aligned with availability tiers and the data's lifecycle stage.
- **Prioritizing data locality:** The closer that data can be kept to where it is being processed, the less need there will be to transfer it for processing. This approach is called prioritizing data locality, and it has the obvious benefit of speeding up data access. Organizations can also reduce data transfer energy consumption by optimizing data transfer protocols to reduce the amount of data that needs to be moved. Techniques include data compression and efficient routing.
- **SSDs vs. HDDs:** Choosing SSDs over HDDs reduces absolute power demand by approximately two-thirds, leading to enhanced application performance and improved power efficiency in both capacity and IOPS per watt. However, it is crucial to acknowledge that for the same capacity, SSDs carry embodied emissions approximately eight times higher than those of HDDs, resulting in a nearly doubled overall carbon footprint.[9] Hence, storage systems operating in locations with lower grid emission factors will achieve their lowest environmental impact when leveraging HDDs where use cases allow.
- **Space utilization:** Achieving reduced rack and floor space requirements through enhanced data reduction technology performance and higher-density media.

Data center storage capacity grew by 2500% over the last decade,[10] yet there is no industry standard benchmark providing power efficiency insights for storage arrays. The ENERGY STAR label, which relies on a standard test suite developed by the Storage Networking Industry Association (SNIA), partially addresses this gap, but does not disclose detailed benchmark results. This void requires IT departments to make decisions based on vendors' dimensioning tools or the advice of storage consultants. When assessing storage efficiency, evaluators should consider unique workload requirements such as the size and capacity of the data footprint, energy performance, and data redundancy objectives. Storage footprint and capacity for a given rack space vary widely depending on the RAID level of storage, compression, and deduplication technologies in use, if the data is reduceable, and the capacity and type of media used.

At the disk level, power efficiency data is available from manufacturers' data sheets. Some manufacturers offer disk-level power-throttling options to optimize power consumption in large deployments.

Energy effectiveness should be calculated considering the power requirements of all components in the storage system including drives, storage array controllers, network components, required cooling, and any other associated equipment. After total power demand is determined, power effectiveness can be determined based on metrics such as capacity per watt, IOPS per watt, or activity per watt. Capacity per watt will be highest when the array is fully or close to fully populated and being operated at the designed capacity level while avoiding over-dimensioning.

For more information on sustainable data storage practices, see the HPE whitepaper Data Efficiency: Empowering Sustainable IT.

[8] "EU negotiators secure agreement at COP28 to accelerate the global transition away from fossil fuels and triple renewables and double energy efficiency this decade," European Commission, 2023
[9] "The Dirty Secret of SSDs: Embodied Carbon," Swamit Tannu and Prashant J. Nair, Cornell University, 2022
[10] "Sources of data center energy estimates: A comprehensive review," David Mytton and Masao Ashtine, Joule, 6(9), 2032-2056, Sep 2022

### Network

Like storage, networks are not the main power consumer in the data center and there are no standardized third-party published test results to equip IT managers in helping assess the power effectiveness of various generations of networking equipment. More than power effectiveness, the driving factors behind technology refreshes of network elements are the depreciation of assets, the need to increase throughput, and designing inefficiencies out of the network.

New Ethernet and Wi-Fi standards bring performance improvements in speed, security, and wireless density. Favoring wireless networks in remote, branch offices, or campuses is an established trend that delivers a step change in efficiency by removing the need to lay cables and offering flexible role-based access control options. The more recent wireless networking standards, such as Wi-Fi 6, also address power management efficiency gain opportunities in both the network and connected devices.

Network upgrades can be disruptive and require downtime, leading to infrastructure management occurring at night or during weekends. Network operations support has not traditionally benefited from the widespread adoption of software automation. These combined factors make it operationally challenging to replace network elements.

Choosing technology that leverages a higher grade of automation during installation, provisioning, and updating will ease both the burden and risk of network migration projects, allowing businesses to be more nimble, secure, and quickly eliminate sources of inefficiency.

### Edge

Edge servers and gateways benefit from the same performance, power, and space improvements as data center servers and storage. These similarities are the result of components being leveraged for edge uses in more ruggedized form factors. However, additional considerations must be included when determining the optimal lifecycle of edge infrastructure:

- **Data gravity:** As compute power moves from the data center to where the data is being generated, planning the edge infrastructure refresh cycle requires an understanding of the data strategy and how the data will be used at the edge.
- **Operational technology equipment lifecycle:** Edge infrastructure has been deployed as part of dedicated technology systems historically. As these systems get modernized, opportunities will emerge to address new use cases, consolidate, and increase the utilization of equipment at the edge.
- **Internet of Things:** Components in the edge infrastructure are sometimes specific to the use cases they are serving. Choosing edge equipment that supports multiple protocols such as Ethernet, Zigbee, Bluetooth, and USB will be beneficial, extending use cases and coping with sensor technology evolution and updated networking standards.
- **Asset management:** Edge infrastructure often also presents unique asset management complexity associated with the logistics involved in physically reaching them.

As a result, the refresh strategy for edge technology supporting preventive maintenance of manufacturing equipment, retail video analytics, or autonomous driving would each have to be approached differently. As edge use cases rapidly evolve, and are required to support a broader data strategy, they inherently carry the risk of deploying infrastructure that may no longer be optimal or adequate much sooner than expected. Prioritizing versatile equipment with broad compatibility, while keeping options open to refresh early, will mitigate the technology risk—for both the edge devices and the network elements that connect and enable them.

## Operational efficiency benefits

In addition to the benefits noted earlier, a formal technology refresh strategy can also have significant operational and cost benefits.

Hardware-based improvements can be implemented into the infrastructure quicker when shorter refresh cycles are adopted. Remote monitoring, measurement, and administration hardware dramatically simplify management and increase device security. Hardware-based security, such as silicon root of trust from Hewlett Packard Enterprise, provides a series of trusted handshakes from lowest level firmware to BIOS and software to ensure a known good state.

Shorter refresh cycles also allow organizations to roll out and reap the benefits of software innovations faster. For example, further effectiveness can be gained by the adoption of:

- **Anticipatory and self-correction technology tools:** Self-correcting tools can minimize the number of redundant systems required. The combination of self-diagnosis and correction, along with effective maintenance servicing, drives a very efficient use of deployed resources.
- **Performance monitoring tools:** These inform pertinent technology updates or upgrades. Tools monitoring operational status and managing device configurations make administrators aware of system status, utilization levels, and available firmware updates. For example, AI Insights from HPE Aruba Networking Central automatically identifies issues, recommends potential solutions, anonymously benchmarks against peers, and can be configured by the admin to perform auto-remediation. HPE Sustainability Insight Center provides a current view of the total energy consumption across customer's HPE infrastructure and includes information about greenhouse gas (GHG) emissions and costs associated with energy consumption.

### Residual value return

Another benefit of shortening refresh cycles, particularly for servers, is that newer generation models may have residual value at their end of use. Savvy IT organizations have identified this as an additional opportunity associated with proactive server lifecycle management. Not only do organizations benefit from the power, space, and cooling efficiencies noted in the previous section, but IT refreshes also maximize residual value at the end of use. However, operationalizing this aspect of the technology refresh strategy requires coordination between IT, finance, facilities management, and procurement organizations to most effectively realize this value.

### Technology risk mitigation

Because information technology evolves quickly, any technology solution carries an inherent risk of becoming obsolete sooner than expected. Technology refresh cycles are also a good opportunity to reassess requirements and make any adjustments to the type and quantity of equipment being procured. Composable systems, system-on-chip technology, hyperconverged infrastructure, liquid-cooled systems, edge compute devices, and AI performance recommendation tools have matured into general use. Assessing the merits of these various options is beyond the scope of this paper but would nonetheless be an important element to consider as part of the wider benefits of technology refresh. Each of these technologies provides an opportunity to drive innovation as well as power, cooling, space, and IT staff efficiencies. Proactive lifecycle management is particularly relevant as a risk mitigation strategy for edge and edge-network components.

## Business requirements and technologies driving shorter refresh cycles

Many organizations have already adopted lifecycle management strategies with shorter refresh cycles, particularly where IT systems are directly linked to business outcomes, such as research teams in energy generation or high-tech manufacturing, computational sciences, and high-performance data analytics. These business functions recognize the value that a technology refresh approach can bring when the most efficient and highest-performing IT assets are used.

Several other technology trends have increased the movement to shorter refresh cycles including:

- The widespread adoption of hybrid cloud architectures with containerized workloads detached from the infrastructure—thereby removing the risk of application migration between hardware versions
- The appeal of a step change in productivity granted by software-driven composable architectures
- The ability to upgrade equipment subassemblies, particularly on composable frames and storage systems, making it easier for IT staff to consider refresh initiatives rather than riskier migration projects
- Adoption of IT technologies with AI-driven self-diagnosis and healing capabilities, which dramatically accelerate incident troubleshooting and resolution time and therefore reduce operating expenses
- Data availability commitments and live upgrade capabilities. These features make it easier for IT staff to consider storage arrays and network elements refresh initiatives, which historically have required the undertaking of complex migration projects

## Barriers to the adoption of shorter refresh cycles

Despite the benefits, the reality is that server and storage infrastructure is not being refreshed at a rate that drives the highest efficiency benefits for organizations. Some organizations have successfully implemented efficient lifecycle strategies, allowing them to proactively set requirements and plan for the implementation of new technology features. These organizations typically know how long equipment will be in use at its inception. In other cases, there is a backlog of equipment to be installed, causing newly acquired hardware to wait in line for an undesirable amount of time before being operational.

There are notable barriers to technology refreshes:

- Technical debt in the application layer makes the asset migration complex with many legacy applications having legacy architecture requirements for specific OS and hardware versions. A hardware refresh would require application modernization, requiring investments at a different scale.
- Depreciation policies for IT assets and software are not necessarily tied to technology cycles and are often rooted in larger financial considerations governed by the industry the enterprise operates.
- For CapEx budget–driven organizations, technology refresh tends to be deprioritized in the face of more urgent projects and is only addressed once it has already become a major issue.
- IT organizations tend to be measured on application availability and do not have any incentive to reduce power consumption.
- Power-efficiency metrics are usually the responsibility of the facilities management team, making the benefits of refresh less clear to individual organizations. With an expected increase in focus on the environmental aspects of IT and related regulations, IT organizations should be prepared to understand and contribute to optimizing the useful work done per kWh of electricity.

- IT organizations often underestimate the benefits that power, performance, and operational improvements can achieve.

It is also important to acknowledge that newer IT infrastructure may put pressure on the data center infrastructure by driving higher power density requirements. The drastic advance in cores per CPU discussed previously and the increased adoption and performance of GPUs increase the demand for electricity and the amount of heat produced in data centers. With rising densities and resulting heat generation, air cooling will reach a point where it cannot remove enough heat to maintain hardware at operating temperature, requiring operators to rethink the design of their data centers. While a data center redesign or modular design update may have initial cost implications, it is also certainly an opportunity to drive higher standards of energy efficiency, which will reduce costs and environmental impact.[11]

## Asset end-of-use strategies

A critical part of any lifecycle management strategy is the appropriate disposition of end-of-use assets being replaced. Circular economy principles provide a helpful framework to inform end-of-use decisions to minimize the environmental impact. Technology refresh strategies need to consider three primary areas of risk:

- **Security and compliance:** Are the assets being disposed of following cybersecurity best practices and applicable government privacy regulations?
- **Environmental:** Is the equipment processed following labor, chemical substances, e-waste, and environmental regulations?
- **Economic:** Is the economic model implemented for asset disposition transparent, scalable, and replicable?

At end of use, assets can either be sent to landfill, recycled, or repurposed, in whole or in parts. Repurposing is often called upcycling. Without examining the specifics of recycling versus upcycling benefits, as a rule, upcycling is less energy-intensive than recycling. Upcycling opportunities for the system, or its parts, are correlated to the asset age, configuration, and associated secondary market demand.

This raises the paradox of reintroducing ageing, less efficient infrastructure into the market when the purpose of retiring it may have been eliminating sources of inefficiency. There are several ideal scenarios, for example, ageing CPUs used to host business-critical applications, running continuously, repurposed for application development and testing, where they are used sporadically. This mitigates the environmental downside of inefficient equipment while avoiding the environmental impact of manufacturing it from the ground up. Another strategy to mitigate manufacturing emissions is to use repurposed components for memory and storage capacity expansions. Furthermore, in some situations, second-hand equipment may be more suitable. This could be due to the need to run legacy applications that are only compatible with older equipment, procurement regulations mandating a portion of second-hand equipment, or markets where acquiring the newest equipment is financially challenging.

When IT assets cannot be upcycled due to age or disrepair, they should be recycled in accordance with applicable regulations and best practices. The appropriate processing of e-waste is a growing challenge that has been acknowledged by the United Nations and many governments. The World Health Organization (WHO) states that “primitive recycling techniques, such as burning cables for copper recovery, expose both adult and child workers, as well as their families, to a range of hazardous substances.” Recycling or upcycling assets through an organization without adequate certifications, asset tracking, and assurance programs throughout the commodity lifecycle poses material, reputational risks for organizations.

## Lifecycle strategy recommendation

As part of a broader IT strategy, lifecycle management allows IT departments to be more intentional in the way they drive and implement their IT lifecycle decisions, taking advantage of the opportunities presented by performance improvements and newer technologies featured in current generation equipment.

### Environmental impact

As discussed, enterprise IT equipment has a significant environmental impact stemming from the materials, manufacturing, and logistics required to move the equipment to a customer, known as embodied impact. However, the greatest environmental impact results from the use phase, or when enterprises are using the asset. According to internal HPE server product carbon footprint (PCF) calculations for average electricity grids, 75% or more of the environmental impact of compute equipment is typically caused during the use phase. This fact is confirmed also over the long term by the server operations data in Table 2. This paper serves as a guideline to balance the embodied impact while minimizing the use-phase impact. Asset end-of-use options for assets must also be investigated, with readers encouraged to prioritize partners who can upcycle the equipment. In all cases, the total environmental impact depends on internal strategies and external factors such as geographic location, availability of upcycling partners, electricity grids, and logistics paths.

[11] “Could liquid cooling save the planet?,” TechHQ, 2023

Enterprises can evaluate currently available referenceable data to make informed lifecycle decisions, such as:

- Data center electricity consumption
- Mix of electricity sources
- Modularity and repairability
- Availability of spare parts
- Product material composition
- Recyclability of the various components
- Carbon footprint

Tracking these data points are simple steps that organizations can take today to start measuring and reporting on the environmental impact of their IT equipment, supporting the organization's larger business and environmental objectives.

## Determining the optimal lifecycle

The case study in Table 2 aims to uncover the economic and environmental trade-offs between refreshing often and refreshing rarely. Refreshing often reduces emissions from electricity consumption at the expense of emissions of manufacturing more equipment, versus the opposite, reducing emissions from manufacturing the equipment by keeping it longer and procuring less often, at the expense of emitting more $CO_2$ due to more electricity consumption in operation. The best strategy for minimizing emissions is found to depend on the grid emission factor of the electricity consumed, according to findings in Table 2. The break-even point at which longer and shorter refresh cycles yield similar total emissions is found to be around 100g$CO_2$e/kWh. Above the break-even point, a four-year cycle performs better. Below, an eight-year cycle performs better. Thus, IT refresh strategies must consider the location of operation and the emission factor of the grid in which they operate. Additionally, the manufacturing emissions can be mitigated to some extent by sending used equipment to be reused, as preowned products or spare parts, reducing the future demand for manufacturing equipment. When replacing older hardware, and running nonperformance-sensitive applications, one can also consider procuring preowned products, further mitigating the manufacturing emissions. A shorter refresh cycle is found to be the best economic strategy, with product support being the main cost driver, followed by software licenses and electricity consumption. IT operations in locations with an electricity price higher than that of this study will experience electricity cost to be the main driver, with the short-cycle strategy performing even better economically than in this study.

Edge servers should broadly follow the compute lifecycle and include additional considerations associated with the specific use case as noted earlier. To effectively mitigate the technology risk associated with some rapidly evolving use cases, information technology and operations technology departments should prioritize versatile equipment with broad compatibility while keeping options open to refresh earlier than data center compute assets and reconsider the assumption at that time. This is based on the evolution of sensors, standards, edge applications, bandwidth, and processing intensity requirements of the data handled by the edge infrastructure.

It is harder to draw a generic conclusion for storage. Efficiency is driven mostly by data strategy and software advancements. As noted earlier and similar to compute, the grid emission factor should be considered when establishing the strategy related to hardware. Operators benefiting from low-grid emission factors should assess the implications of deploying more storage tiers on HDD technology, considering the lower embodied carbon footprint. However, they should be aware of the trade-off between CapEx and OpEx, as well as the higher use phase emissions compared to SSDs. Nevertheless, considering the constant pressure on storage infrastructure associated with data growth, IT departments should monitor storage technology innovation developments and reassess every 3 to 4 years. This assessment should include an evaluation of the data and associated storage tiering they have implemented and if this level is still optimal. Storage lifecycle must align with the data lifecycle strategy it supports. Often, newer technology offers efficiency gains for specific data segments. Therefore, storage lifecycle management is less about a traditional technology refresh approach and more about augmenting existing capabilities with newer and more efficient capabilities for select sets of data.

The optimal lifecycle of a network device is a function of which part of the network the device belongs to, either core, data center, campus, or edge. Data growth, number of devices, and density of connected end points each impact network element's technology refresh cycles, particularly at the edge. For example, the adoption of Power over Ethernet (PoE) technology to power edge devices presents opportunities and challenges. PoE is an opportunity to reduce cabling and allow technology to be deployed at locations that were designed without power outlets. However planning and dimensioning PoE requirements is a challenge, as edge use cases, and the type and quantity of connected devices are still evolving rapidly. Power consumption and supported feature requirements may outpace original plans. The recommendation for edge and campus network element lifecycle management is therefore to consider the use cases that the network elements are supporting and consider the network as a service, as financial depreciation is often a main driver. As for other edge elements, assuming a shorter lifecycle at the onset while opting for versatile equipment and keeping options open to refresh early will mitigate the technology risk.

### Operationalizing the decision

Once an optimal lifecycle has been defined, it should be implemented with other business and operational processes. The refresh process should be determined up front and integrated with either the equipment acquisition process or the equipment disposal process.

### Refresh options

**From the start:** Defining asset lifecycle at the time of equipment acquisition is an approach that allows IT departments to derive up-front economic benefit from the improved residual value of the equipment at the time of return. Aligning refresh decisions with the procurement process allows organizations to leverage a very well-defined and documented set of governance and compliance processes associated with lifecycle management of assets. It is also the most flexible option, as it allows organizations to delay the decision of the actual technology refresh options until a predefined time period at the end of use.

**After the fact:** Some organizations prefer an after-the-fact approach where assets due to be refreshed are systematically sold for upcycling or recycled if too old for upcycling. This approach may be harder to implement and monitor as initially there may be less governance attached to the end of use than to the procurement of assets. An after-the-fact approach may also negate the economic benefit if assets have little residual value and result in financial complexities, such as accounting for revenue from fully depreciated IT assets. Because the risks associated with IT asset disposal are material, an after-the-fact approach to asset upcycling requires the same robust governance processes as those attached to the asset procurement phase, including transparent reporting, governance, and compliance.

**As a service:** The emergence of as-a-service models, such as HPE GreenLake, embed technology refresh cycles, which include tools and metrics to evaluate equipment effectiveness and carbon emissions along with the necessary framework to evaluate optimal lifecycles and effectively manage end-of-use operations with the least cost and environmental impacts.

## Conclusion

A comprehensive IT lifecycle strategy that considers all phases of the lifecycle and incorporates business, technology, and environmental implications, will yield significant positive efficiency, cost, and environmental benefits. Organizations that seek input from technology partners, impacted internal organizations, and trusted advisors are best positioned to overcome common barriers to adoption. As the rate of technology improvement accelerates, effective IT infrastructures which include effective lifecycle management principles are a business imperative. Location, electricity grid emission factor, electricity price, needs within security, performance, licenses, supportability, redundancy, and manageability are all factors that must be considered. Emissions and costs have been investigated in this paper to enable the reader to consider them successfully when defining a lifecycle strategy. Thankfully, the information and resources needed to develop and implement effective lifecycle strategies are readily available, along with numerous examples of success from companies that have implemented these strategies.

## Appendix 1

Three sources of data are used in the case study: HPE product carbon footprint for HPE ProLiant DL380 Gen11 Server,[12] Uptime Institute's 2023 Q&A titled "IT sustainability: Achieving more transactions per megawatt hour,"[13] and Enter Enterprise Strategy Group's technical validation "Validating the Sustainability Impact of HPE ProLiant Compute Gen11 Servers".[14] Manufacturing emissions, electricity consumption, and PUE are gathered from the mainstream configuration in HPE PCF. Workload capacity improvement increase in each new server generation is gathered from Uptime Institute. This improvement assumption aligns precisely with that found by Enterprise Strategy Group from HPE Gen10 servers to HPE Gen11 servers. The increase in power draw per server for each new generation is gathered from Enterprise Strategy Group's paper.

- 100 servers are replaced in year 0.
- New server generations are released every four years.
- Each new server generation has a workload capacity improvement of 75% compared to the previous generation, which results in 43% fewer servers needed to host the same workload refreshes.
- Energy use per server will increase by 40% for each server generation, with each of the 100 original servers using 3224 kWh per year. This means that servers in the first generational refresh each use 4514 kWh, servers from the second generational refresh use 6319 kWh per year, and so forth.
- A PUE of 1.55 is assumed, which adds to the electricity consumption stated in the previous point.
- Manufacturing each new server results in 2389kg$CO_2$ for all servers throughout the entire study.
- Electricity price is set to $0.15/kWh to simulate a price point relevant to the United States.
- Each server costs $6000 in all the years of the study.
- Support cost is $2500 per server per year.
- Software license cost is $1800 per server per year.
- Support and software license costs increase by 2% every year.

**Further considerations of the case study**

- Electricity prices in various other countries can be substantially more expensive than in the U.S. Readers outside the U.S. should scale electricity costs accordingly for the relevancy of the results.
- Software licenses are commonly priced either by the number of CPUs or by the number of CPU cores. For those software licenses that follow a per-core pricing model, it is inaccurate to model software license price per server. This leads to some inaccuracy in the cost consideration, as future server generations will be available with many more cores per CPU, and more licenses will be needed for those specific applications. However, this is only a portion of the total software license cost, as some licenses still follow the per-CPU model. Additionally, as the cited Uptime Institute survey shows, CPU advancements also result in more work being done per core in newer generations, which yields economic cost savings despite software licenses being priced per core.

[12] "HPE product carbon footprint—HPE ProLiant DL380 Gen11 Server," HPE, 2024
[13] "IT sustainability: Achieving more transactions per megawatt hour", Jay Dietrich, Uptime Institute, 2023
[14] "Validating the Sustainability Impact of HPE ProLiant Compute Gen11 Servers," Tony Palmer, Enterprise Strategy Group, 2024

## Learn more at

HPE.com/us/en/living-progress.html

Explore HPE GreenLake

Chat now (sales)



a50002031ENW

Hewlett Packard Enterprise