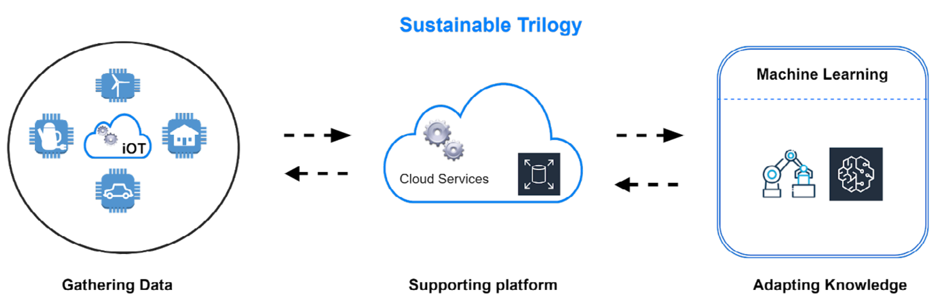
GlobalLogic®
A Hitachi Group Company

PRACTITIONER PERSPECTIVES

by Nicolas Cieri, Solution Architect, and Álvaro Soria, Solution Architect; reviewed by Harsimrat Singh, Technology Director

Achieving Energy Efficiency by Combining IoT, AI, and Cloud Computing

Together, Artificial Intelligence (AI), the Internet of Things (IoT), and cloud computing are the triumvirate of IT's future.



The IoT can gather data from various devices and platforms in various ways, and AI can use this knowledge to adapt to new inputs and carry out jobs in ways that mimic human performance. Cloud computing provides the supporting platform and IT infrastructure that makes these tasks possible at scale.

This trident is also the most promising approach to tackling the persistent issue of the CO_2 footprint created by large-scale IT infrastructure, and can enable innovative, effective solutions for reducing energy consumption.

Using sensors and IoT devices to collect data on energy usage and environmental factors, AI algorithms can analyze the data and provide insights for making more informed decisions about energy usage.

Cloud platforms provide the computing power and storage needed to process and store large amounts of data, while also enabling real-time monitoring and control of energy usage. Together, these technologies enable energy-efficient practices such as predictive maintenance, demand response, and automated energy management.

The goal of this whitepaper is to provide an overview of the importance of the trilogy for achieving energy efficiency. This includes cloud-based services and solutions created by major cloud providers to address energy consumption, along with the steps being taken to ensure the continued development of sustainable IT solutions.

Energy-Saving Techniques

A smart economy and environment are vital in enabling the widespread development of smart cities.[1] Consequently, governments, companies, and scientists have identified human necessities, water waste, and greenhouse gas emissions as critical issues.

Responsible use of energy sources is imperative, and the following approaches are part of the solution.

Sustainable Cloud Computing

The expansion of scientific computing, AI data processing, and IoT operations has driven increasing demand for cloud computing services in recent years.

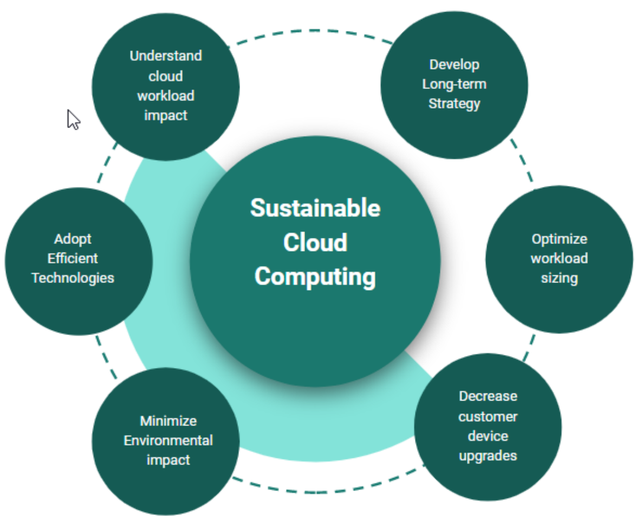
It takes a significant amount of energy from the cloud to provide efficient and reliable services; in fact, cloud computing is responsible for 2% of CO_2 emissions into the atmosphere.[2]

It's a problem that must be addressed to ensure the sustainable growth of cloud computing. Implementing the following best practices and design principles in sustainability application design helps to reduce energy consumption.

Understand Your Organizational Impact

The first step is to develop a comprehensive understanding of your cloud workload's impact by measuring its current and future impact. This includes evaluating all sources of impact, including that from customers using your products, and the eventual decommissioning and retirement of these products.

This sustainability assessment should be used to establish key performance indicators, identify opportunities to improve productivity while minimizing impact, and forecast the outcomes of potential changes over time.



1. Bonetto R, Rossi MMachine learning approaches to energy consumption forecasting in households CoRR. arXiv:abs/1706.09648, 2017.

2. M. Hussain, L. F. Wei, A. Lakhan, S. Wali, S. Ali, and A. Hussain, "Energy and performance efficient task scheduling in heterogeneous virtualized cloud computing," Sustainable Computing: Informatics and Systems, vol. 30, 2021.

Set Sustainability Targets

Based on these opportunities, the next step is to develop a long-term strategy by creating sustainability targets for each cloud workload. This could mean decreasing the computing and storage resources necessary per transaction, for example.

This implies an analysis of the return on investment for sustainability enhancements to current workloads and provides the resources required for owners to invest in sustainability goals.

Optimize Workload with Efficient Design

In terms of hardware resources, optimizing workload sizing and adopting an efficient design are key. Two hosts running at 30% utilization consume more power than a single host running at 60% utilization, due to the baseline power usage per host.

Additionally, plan to minimize or shut down idle resources, processing, and storage to lower the energy required to power your workload.

This can help reduce the carbon footprint of a workload by using resources that would otherwise go unused.

A good example of this is EC2 Spot instances on AWS, which allow users to take advantage of spare EC2 capacity at a lower cost.

Modernize and Digitize for Energy Efficiency

Innovative and more effective hardware and software offerings must be embraced to anticipate and adapt to them.

- Encourage your partners and suppliers to make upstream enhancements that assist you in minimizing the impact of your cloud workloads.
- Continuously scrutinize and appraise new, more efficient hardware and software offerings.
- Create designs that allow for flexibility, allowing you to rapidly adopt new, efficient technologies.

Current cloud providers offer shared services that are available to a wide range of customers to minimize the infrastructure required to support cloud workloads.

For example, by migrating workloads to the AWS Cloud and leveraging managed services such as AWS Fargate for serverless containers, customers can share the effects of common data center components like power and networking, which increases resource utilization.

Utilize managed services that can reduce your environmental impact, such as Amazon S3 Lifecycle Configurations. It can automatically move infrequently accessed data to cold storage, or Amazon EC2 Auto Scaling, which can adjust capacity to meet demand.

Look for Energy Efficiency Gains Downstream

As a final recommendation, you can minimize the environmental impact of your cloud workloads downstream by lowering the energy or resources necessary to utilize your services. Decrease or eliminate the need for customers to upgrade their devices to access your services, for example.

You could also utilize device farms to evaluate the anticipated impact and work with customers to determine the actual impact of using your services. Cloud providers support multiple geographic regions and provide tools that measure each region's footprints, for example.

Google Cloud Platform (GCP) provides a Carbon-Free Energy Percentage (CFE%) dashboard that presents users with the percentage of their cloud usage powered by carbon-free energy sources, such as wind or solar. It is used by the tech industry, including companies like Salesforce and Box, to track their carbon footprint and set goals for reducing it.

Azure offers a Carbon Efficiency Calculator that can help users understand the carbon impact of their cloud deployments and identify ways to improve efficiency.

These guidelines and existing tools from major cloud providers enable you to advance your overall sustainability objectives, recognize areas that need improvement, and prioritize areas with the greatest growth potential.

These solutions span a broad range of industries. The automotive industry has used examples of this to run simulations for crash tests, which can help reduce the need for physical testing and the associated carbon emissions.

The financial industry, including companies like HSBC and BNP Paribas, is measuring the carbon impact of their cloud deployments and identifying ways to reduce it.

Support Sustainable Solutions with AI and Cloud Computing

AI technology has evolved beyond its initial hype phase and is now widely used to learn, adapt, recognize patterns, and simulate human intelligence at scale.

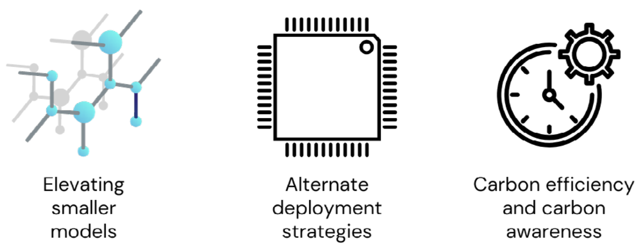
The application of AI in environmental applications has the potential to contribute up to $5.2 trillion USD to the global economy by 2030, resulting in a 4.4% increase compared to the business-as-usual scenario.

AI could help reduce global greenhouse gas (GHG) emissions by 4% in 2030, equivalent to 2.4 Gt CO2e (the greenhouse gas total expressed in terms of billions of tons of global annual CO2 equivalent emissions), comparable to the annual emissions of Australia, Canada, and Japan combined. This transition to AI is also expected to create 38.2 million new jobs, with more opportunities for skilled occupations.

However, achieving these projections requires adopting a broader technology-focused strategy to scale, and this is where combining cloud computing and AI has become a powerful tool for enhancing IT operations.

It can create a vast network capable of learning, growing, and storing enormous amounts of data, enabling users to save, analyze, and draw insights from their data.

What is Sustainable AI?



In this context, software architects must abide by sustainability principles and guidelines to maximize utilization and minimize the resources needed to support their workloads.

Architects can follow several guidelines to reduce the carbon footprint of AI-Cloud-based solutions:

- Eliminate idle resources with serverless technologies and environment automation.
- Minimize unnecessary data processing and storage by utilizing publicly available datasets and models.
- Maximizing resource utilization by right-sizing environments and asynchronous processing.
- Enhance CPU efficiency with simplified versions of algorithms and model compilation.

These recommendations adhere to best practices and objectives aimed at reducing the environmental impact of AI-based cloud solutions. In alignment with these goals, major cloud providers offer the following services to optimize and reduce energy consumption.

Amazon SageMaker

AWS offers Amazon SageMaker [SageMaker], a fully managed service for building, training, and deploying machine learning models. The service includes features such as automatic model tuning and cost optimization, which can help reduce the energy consumption of training and deploying models.

The energy industry uses it to predict and optimize wind turbine performance, which can help increase energy production and reduce the need for fossil fuels.

Google Cloud AI Platform

Google Cloud AI Platform [CloudAIPlatform] provides various services for building and deploying machine learning models, including AutoML, which automates building custom models.

This can help reduce the compute resources required to train models and has been used by the transportation industry to optimize routing and reduce fuel consumption for fleets of vehicles, including taxis and delivery trucks.

AzureML

Azure Machine Learning [AzureML] is a cloud-based service for building, training, and deploying machine learning models.

The service includes features such as automated machine learning and model management, which can help reduce the energy consumption of building and deploying models.

The retail industry has used this to optimize inventory management and reduce waste, which can help reduce the environmental impact of overproduction and unsold goods.

A noteworthy example of this combination is the collaboration between Google and DeepMind to create an AI-based recommendation system for optimizing the energy efficiency of their data centers.

By using thousands of sensors to gather information on the cooling systems of the data centers and feeding this data into DeepMind's neural networks, the AI system can predict the impact of different actions on future energy consumption. This information is then used to determine the most energy-efficient action while meeting necessary safety requirements.

The future of IT systems is expected to focus on data generation, processing, and delivery at high speeds. Two significant trends in this regard are IoT and AI, which complement each other perfectly. IoT allows various devices to exchange data collected through different platforms, while AI leverages this data to make real-time and time-sensitive decisions. A hybrid multi-cloud platform is necessary to bring these technologies together effectively and fuel business advantages, innovation, and the future of the cloud.

The Role of Internet of Things (IoT) in Sustainability

Devices are constantly collecting, distributing, and processing data at the edge of both the enterprise and consumer spaces. This data must be quickly analyzed and transferred over a space beyond immediate control.

Meeting these requirements calls for distributed collection and storage capabilities closest to the source, with the IoT edge and computing events at the forefront of automation and emerging trends.

These advancements drive innovations in future computing architecture, in line with the ongoing growth of increasingly intelligent and interactive devices.

The IoT edge must allow instantaneous transactions without central control, instead utilizing distributed connections to validate, create, and remove connections.

The limitations of data transfer distances due to latency causing operational issues set the boundaries for the feasible edge. AI is the backbone of these elements, governing data life cycles, flow, classifications, reporting, and numerous other aspects of the IoT.

The utilization of IoT can significantly improve energy efficiency not only in energy consumption but also in energy production. By using smart grid sensors, energy networks can be monitored more accurately, allowing the allocation of resources to be better aligned with actual demand needs, which can lead to a reduction in energy waste.

Additionally, smart meters can offer valuable insights to minimize energy costs.

For example, automated billing based on network status or time of use, or enabling communication between appliances and the meter to help consumers make more informed decisions that ultimately reduce energy expenses.

In this context, devices provide capabilities to help them be environmentally friendly, and lean devices use minimal resources to perform their tasks. They are constructed to require fewer resources, reducing the environmental impact of their manufacture, use, and disposal.

Smartphones are an example of devices that use rare-earth metals in many components, which harm the environment during extraction and disposal. Reducing the quantity of these materials in the design of a device can promote sustainability.

Another property is the use of updated and secure software and improvements in code and data management.

Finally, durable devices remain in service for extended periods while delivering their intended purpose and value. They are adaptable to evolving business needs and can recover from operational failures. The longer a device is in operation, the lower its carbon footprint will be because the manufacturing, shipping, installation, and disposal processes will require less energy.

To provide support for the sustainability properties of devices, the cloud computing market offers a wide range of IoT solutions:

IoT Greengrass

AWS offers IoT Greengrass, which allows users to run IoT applications and machine learning models on connected devices. This can help reduce the amount of data sent to the cloud for energy-efficiency.

The manufacturing industry uses it to run machine learning models on connected devices in factories, which can help reduce the need to send data to the cloud and the associated energy consumption.

Google IoT Core

Google Cloud IoT Core [IoTCore] provides a fully managed service for connecting and managing IoT devices.

The service is designed to be highly scalable and efficient, helping to reduce the energy consumption of IoT deployments, and has been used by the agriculture industry to monitor crops and optimize irrigation, which can help reduce water usage and increase crop yields.

Azure IoT Hub

Azure's IoT Hub [IoTHub] provides a secure and scalable platform for connecting and managing IoT devices. The service includes device management, telemetry processing, and analytics. The healthcare industry uses it to monitor medical devices and equipment, which can help reduce downtime and the need for repairs or replacements.

By using cloud platforms, IoT can optimize energy consumption in multiple areas. The IoT can reduce electricity consumption during peak hours through the shutdown or reduction of equipment and production in industrial settings. Lighting, heating, and cooling can be improved based on operational conditions and the weather or time of day.

Advancements in device connectivity allow manufacturers to track energy consumption at the device level, giving managers visibility into energy consumption and actionable insight into waste and available efficiencies.

Another area where the IoT can be applied is in smart cities.

Using IoT technology, the different components of modern cities such as energy, mobility, buildings, water management, lighting, and waste management, can become part of an interconnected ecosystem. IoT technologies can gather information that was previously impossible or difficult to obtain, such as real-time location and utilization of public transport, environmental data, noise levels, waste bin monitoring, energy consumption in public buildings, and lighting.

Barcelona's Energy-Saving Smart Street Lights are an excellent example, where sensors installed in streetlights enable automatic brightness control by analyzing noise levels, air pollution, and population density, resulting in at least 30% energy savings per year.

Conclusion

As climate change becomes an increasingly serious concern, individuals and organizations such as executives, employees, investors, and customers are becoming more aware of the environmental impact of their actions and are investing in sustainable practices.

This has resulted in a growing interest in Sustainable IT and implementing environmentally friendly practices in information technology.

The examples of sustainability principles and guidelines in this paper can help achieve sustainability in ways that leading cloud providers, namely Amazon, Google, and Microsoft are already adopting.

These major cloud providers will continue to address their carbon footprint and offer more sustainable cloud services in the coming years.

The examples of sustainability principles and guidelines in this paper can help achieve sustainability in ways that leading cloud providers, namely Amazon, Google, and Microsoft are already adopting.

Amazon AWS has set a target to run its operations using 100% renewable energy by 2025. Likewise, Google's cloud platform has been carbon neutral since 2007 and utilizing solely renewable energy since 2017. Moreover, Google provides its users with Carbon Sense software to help monitor carbon emissions.

Microsoft's Azure has been carbon neutral since 2012 and aims to power its data centers with 100% renewable energy by 2025. Furthermore, they aim to move towards becoming carbon negative by 2030.

These major cloud providers will continue to address their carbon footprint and offer more sustainable cloud services in the coming years.

At the same time, the onus is also on the organizations and people within them responsible for the development and management of the software and infrastructure utilization to adopt sustainable IT guidelines and best practices.

If your organization is looking for innovative ways to practice sustainable IT, GlobalLogic can help.