Glossary of digital infrastructure sustainability

Key terms used by those defining, regulating and applying sustainability strategies

Uptime Institute | INTELLIGENCE

This high-level glossary of terms accompanies the Uptime Institute Intelligence report series, 'Digital infrastructure sustainability — A manager's guide' (see **Appendix**).

Entries are selected for their relevance to the sustainability topic. A list of legislation and standards related to sustainability is included separately with the report *Navigating regulations and standards.*

Terms in this glossary are listed alphabetically. Terms in the explanatory text that are highlighted in **bold** are also defined with their own separate entry.

Uptime Institute Intelligence is an independent unit of Uptime Institute dedicated to identifying, analyzing and explaining the trends, technologies, operational practices and changing business models of the mission-critical infrastructure industry. For more about Uptime Institute Intelligence, visit uptimeinstitute.com/ui-intelligence or contact research@uptimeinstitute.com.

Additionality. When a newly initiated or planned renewable energy generation project is enabled by a direct or **virtual power purchase agreement (VPPA)** — meaning the project would not have gone ahead without the signed **power purchase agreement (PPA)**. Additionality is a condition that can also be applied to **carbon offset** projects, denoting where a specific investment by a participating party was instrumental in project execution and completion.

Blowdown. The process of discharging a portion of water and associated total dissolved solids (TDS) from a cooling system to minimize equipment wear and tear. The blowdown is replaced by an equal volume of fresh water to maintain the concentration of TDS below a specified, safe level. The amount of blowdown can be minimized using chemical additives and filtration methods. The water discharged from the blowdown process typically requires treatment before it is discharged to a receiving water or septic system.

Bundled renewable energy certificates. **Renewable energy certificates (RECs)** that are separated from the renewable energy generation from which they are created and "bundled" with and matched to purchased grid mix energy to create a 100% renewable energy supply for the facility. This conforms with the RE100, green energy coefficient (GEC) and renewable energy factor (REF) standards. The generated RECs and the operating facility are in the same grid region, which means the avoided emissions for the bundled RECs match the greenhouse gas emissions for the grid mix generation, enabling the facility to claim **net-zero emissions** for its electricity supply.

Capacity. Used in two different contexts:

1. Power demand capacity represents the maximum amount of power that a data center can use (i.e., that it has provisioned), usually expressed in megawatts (MWs).
2. IT equipment capacity indicates the maximum amount of work (transactions per second), data (terabytes of storage) or network traffic (gigabytes per second) for the equipment to operate effectively. The capacity is measured by standard industry benchmarks (i.e., the **server efficiency rating tool, SERT**). Generally, equipment is more energy-efficient, sustainable and economical if it operates close to its maximum capacity.

Carbon neutral. See **Net-zero emissions.**

Carbon offset. Activities that result in avoided or removed emissions of **global warming potential** gases are documented and the carbon (or carbon equivalent) savings quantified. If investments are made in these activities, the carbon emissions avoided or captured can then be used to offset carbon emissions elsewhere. Emissions avoidance can be achieved through activities such as energy conservation / efficiency projects, carbon capture and storage, and emissions removal through forest protection, the planting of trees or other related activities that capture or sequester carbon.

Carbon usage effectiveness (CUE). A metric used to assess the "carbon efficiency" of a data center facility. It is calculated by dividing the carbon dioxide (CO2) emissions (kilograms of CO2e / year) generated by the data center by the energy used by its IT equipment, i.e., kilowatt hour (kWh) / year. It does not account for life cycle emissions such as construction or dismantling.

Chiller coefficient of performance (COP). A metric used to evaluate the efficiency of a data center's water-based cooling system. It is calculated by dividing the chiller's delivered cooling capacity by the electrical power input to the system.

Circular economy (CE). An economic system that aims to minimize or eliminate external inputs through reusing, repairing, repurposing, remanufacturing, or recycling existing products and components. Data center operators often incorporate CE principles alongside sustainability initiatives by designing buildings, products/equipment and operations for improved longevity and efficiency. Application of CE principles to data center operations is sometimes referred to as "circularity".

Comatose servers. Sometimes referred to as "zombie servers", comatose servers are energized and operational but no longer run a workload and have zero **utilization** (except for the operating system or updates) for several weeks or longer.

Computation fluid dynamics (CFD) software. Software that models air flow and temperature distribution in the data center IT space. It is calibrated by taking temperature and airflow measurements at selected points in the data center. The software is then recalibrated when air flow is adjusted, when IT equipment is removed, installed or replaced, and / or when computer room air conditioner / handling units and / or slotted floor tiles are added, removed, moved or shut down. CFD software is used to identify hotspots that can damage equipment and to improve energy efficiency.

Congestion. Caused by limits to the transmission capacity, typically between electricity grid regions or zones within a grid region. It occurs when supply in one grid region exceeds demand, while demand in another grid region exceeds supply, but there is not sufficient transmission capacity on the interconnects to transport the excess power from one region to the other. Congestion can prevent the interconnected grid from running efficiently or sustainably.

Container (server). Used in two different contexts:

1. A software container is a package of software that includes an application programming interface, libraries, configurations and other functional software that can easily and reliably be moved across computing environments. The use of software containers, a form of **virtualization**, has many advantages and can be very energy-efficient.
2. A data center container is a small, prefabricated data center, or an electrical or cooling facility, that is designed and supplied in a shipping container (or similar design). Containerized data centers are usually designed to be very energy-efficient.

Cooling distribution unit (CDU). A preconfigured liquid distribution cooling system consisting of a heat exchanger, pumps, valves and associated controls. The cooling medium can be water or a nonconductive dielectric cooling fluid. Heat is exchanged from the data center cooling liquid loop to the facilities cooling water loop. The CDU separates the two cooling medium systems, minimizing the risk of significant leaks and breaks inside the data center and improving maintainability.

CO2e. An abbreviation for CO2 equivalent that represents the global warming impact, expressed in kilograms (kg) or metric tons of CO2, for emissions of a gas (such as hydrofluorocarbon refrigerants or perfluorinated compounds) with a **global warming potential (GWP)**. CO2e is calculated by multiplying the kg of emitted gas times the gas' GWP in kg of CO2 per kg of gas. For example, 1 kg of methane (with a GWP of between 28 and 36 kg of CO2 over 100 years) has a global warming impact equivalent to between 26 and 32 kg of CO2.

Central processing unit (CPU) power management functions. Consists of a set of hardware- and software-initiated functions that aim to reduce CPU voltage and / or frequency when workload demands are low or absent. Power management can reduce server energy use in times of low utilization or when idle at the cost of higher latency and slower response times.

C-states. A function of **CPU power management** that reduces the processor voltage during periods when the processor has no work to perform. During no-work periods, C-states sequentially reduce the power demand in time-defined steps. A lower C-state value indicates a shorter recovery time and less energy savings than higher C-state values.

Curtailment. One of the ways that electricity grid managers, referred to as independent system operators (ISOs) in the US, balance supply and demand in a grid region. During periods of high renewable energy output, the grid manager can direct specific generation facilities to disconnect from the grid for a specific period to prevent oversupply, power / voltage surges and system failures.

Data center infrastructure management (DCIM). A set of tools (mostly software and some cloud-based) used to manage and monitor both a data center's IT systems and the facility itself. DCIM can be used to manage assets, measure and analyze power demand and energy consumption, measure temperatures and utilization rates, and track many other data points. It can also be used to create, report and forecast metrics and key performance indicators, and some DCIM systems can be used as control systems. DCIM is useful — if not essential — for managing sustainable data centers.

Direct air cooling. A system that introduces cool air from the outside directly into the data center to cool IT equipment. Temperatures can be managed by premixing the outside air with warmer inside air. Ancillary cooling may also be used when outside temperatures are high. Filtration may be required to remove airborne solids or corrosive contaminants. Direct air-cooling systems, which are not suitable for use in many data centers, are very energy-efficient.

Direct liquid cooling (DLC). A system that uses the thermal properties of liquids to eject heat directly from the IT equipment. Direct cooling technologies include water or dielectric-cooled cold plates and full and partial dielectric immersion systems. DLC systems are best suited for power dense, high performance compute systems, which can require a lot of heat rejection. DLC systems are energy- and water-efficient but can be expensive to buy and maintain.

End-of-life (EOL) product. A product that has reached the end of its planned life cycle or is no longer useful or economic. Organizations should have a plan for responsible refurbishment, reuse, recycle and disposal of EOL equipment (or its reusable components).

Energy management plan. Developed by a data center operator to monitor, control and conserve energy use in a data center and other digital infrastructure. The energy management plan should identify and drive continual improvements in energy use performance by facility and IT systems.

Energy Star. A certification program developed by the US Environmental Protection Agency that identifies and labels energy saving tools, technologies and systems in order to decrease overall energy use and increase efficiency. Australia, Canada, Japan, New Zealand, Switzerland and Taiwan are Energy Star partners that recognize or apply Energy Star programs.

Environmental, social and governance (ESG). A term used in the investment community to designate a set of nonfinancial criteria used to assess the degree to which an organization is responsibly operated. ESG criteria are not strictly defined but typically cover a range of factors, such as carbon emissions, water use, pollution, waste and recycling, treatment of workers, diversity, and a transparent and accountable corporate structure. Digital infrastructure can be designed and operated to meet certain ESG metrics, with a focus on environmental factors.

E-waste. Electronic products that have reached the disposal stage of their life cycle. The improper disposal of e-waste can lead to environmental health and safety problems and may result in fines or penalties.

Firming. The process of supporting or mitigating the intermittency of renewable (electricity) energy by supplying energy from other sources, which may or may not be renewable. The goal is to provide a reliable supply of electricity at the meter of an operating data center.

Free cooling. The use of the outside (ambient) air temperature to cool a data center. The most efficient form of free cooling is **direct air cooling**. Water and air-side economizers utilize a heat exchanger between the cold outside air or water and the data center cooling loop to protect the data center from outside contaminants. Many modern data centers use a large amount of free cooling to save energy and reduce carbon emissions.

Global warming potential (GWP). A measurement of the global warming effect of a gas, relative to CO2 (which has a GWP of 1). GWP represents how much heat in the atmosphere is absorbed or trapped by a given gas. For example, methane has a GWP of between 28 and 36 times that of CO2 (over 100 years) — although it disperses from the atmosphere much faster.

Green bonds. Fixed-interest financial instruments that may offer favorable interest rates to those companies that wish to finance energy efficiency and environmental improvement projects. Green bonds usually set out environmental or **ESG** criteria that must be met and may require the achievement of specific environmental outcomes to receive a lower interest rate.

Green tariff. A renewable energy contract offered by energy suppliers that is usually less risky to the buyer than a **power purchase agreement**. These contracts are usually short term (two to six years), and the energy costs and volumes are made clear at the outset. A green tariff contract can be used to procure reliable renewable power for a data center. The contracts place the management and the rate risks associated with procuring the energy **RECs** with the utility or retailer, rather than the purchaser.

Grid emission factor. The average emissions generated, expressed in metric tons (MT) of **CO2e** per megawatt-hours (MWh) supplied in a grid region. This value is typically calculated and published for a calendar year. Grid regions dependent on fossil fuels will have high emission factors: for example, Australia had an emission factor of 0.66 MT CO2 / MWh in 2020. Grid regions with significant renewable energy generation assets will have a lower emission factor. An extreme example is Quebec (Canada) that had an emission factor of 0.0013 MT CO2 / MWh in 2018. Grid emission factors are an important consideration when **unbundled renewable energy certificates** are purchased.

Guarantee of origin (GO). A tradeable commodity that tracks, accounts and assigns ownership of an amount of renewably generated electricity. GOs are mainly traded within the European Union through the European Energy Certificate System and are equivalent to r**enewable energy certificates** traded in the US and other markets.

Heat reuse. The transfer of heat generated by data center activity to a heat-consuming operation, rather than releasing it into the atmosphere as waste. Heat reuse is most common in cold climate locations. Heat reuse is becoming an increasingly important metric to track when data centers are assessed for sustainability.

IT workload shape curve. A graphic that shows time intervals of server processor and / or memory **utilization** in a data center. The graphic can be used to show equipment or data center level utilization either as raw utilization (that does not consider the varied workload capacity of different servers) or **capacity** utilization (that incorporates workload capacity in the calculation). The IT workload shape curve reflects the time-based workload demands on a server or groups of servers. Low utilization throughout the day indicates there is over-provisioning of IT and an opportunity to improve energy efficiency. Ideally, workloads can be distributed so that the shape curve for utilization is consistently high and relatively flat.

Life cycle analysis (LCA). A holistic assessment of a product's environmental impact throughout its life cycle, from raw material extraction to its disposal.

Negative emissions (carbon negative). The status achieved when an organization's digital infrastructure or data center operations achieve **net-zero emissions** and further actions are taken by the operator to reduce carbon emissions beyond the scope of its business activities. For example, the organization can buy extra offsets and / or perform or invest in activities that sequester or remove carbon from the atmosphere.

Net-zero emissions. The point at which an organization's operations have no net impact on the greenhouse gas concentration in the atmosphere. For data centers, a net-zero emissions goal may be achieved through a combination of renewable energy use, energy optimization and efficiency strategies, and the purchase of **unbundled and bundled renewable energy certificates, virtual power purchase agreements** and **carbon offsets.** Uptime Institute uses the terms net-zero emissions and **carbon neutral** interchangeably.

Power capping. A technique that can be applied to IT equipment processors and other components to limit the maximum power consumption of the equipment. The most common way to control power consumption is to adapt the **P-state** settings to ensure that a maximum voltage or frequency (power demand) is not exceeded.

Power purchase agreement (PPA). An agreement between the buyer and renewable energy supplier, typically lasting eight to 20 years or more, where a fixed price is set for the supply of electricity and its associated **renewable energy certificates.** PPAs provide a means to procure the output of a specified MW of generation capacity or specified MWh of energy. PPAs carry financial risks as spot market prices can fall below the contract price and create an economic loss for the buyer / data center operator.

P-state. Modulates the power consumption of an IT equipment processor during periods when it is active (as opposed to C-states). P-states use predetermined operating frequencies and associated operating voltages to match processor work capability and power demand to workload demand, thus reducing and optimizing power consumption.

Power supply unit (PSU) efficiency. The PSU is an internal hardware component of IT equipment that converts outlet power to the appropriate voltage for IT equipment function. PSU efficiency is calculated by dividing the power supplied to the IT equipment by the power drawn from the outlet.

Power usage effectiveness (PUE). A ratio of all power used by a data center over a given time, divided by the amount used by IT equipment alone. For example, if a data center uses 2,000 MWh for a given operating period, and the IT equipment uses 1,000 MWh, the PUE is 2. The PUE number shows how much energy is used or "lost" to support the IT equipment, relative to that used by the IT equipment itself. A PUE number is always more than 1, and a low PUE (e.g., 1.2) is better than a higher one (e.g., 1.9). PUE can be measured at any time, over any length of time, retrospectively or forward looking (prospective PUE). A maximum allowable PUE number can be required by regulators. PUE can also be expressed as a percentage of data center energy used by the IT equipment, known as data center infrastructure efficiency (DCIE).

Renewable energy certificate (REC). Tradeable commodities that represent 1 MWh of zero emissions of renewable energy generation. RECs can be obtained through direct and **virtual power purchase agreements,** as part of a retail contract, or they can be purchased as **unbundled renewable energy certificates** from brokers to match nonrenewable electricity purchases and offset their associated emissions. This allows a business or corporation to claim 100% renewable energy use. In Europe, the RECs may be referred to as **guarantees of origin**.

Renewable energy shape curve. A representation of intermittent renewable energy output data (in MWh) at a generation facility or on the grid as a function of time (hourly, daily, weekly, monthly or annually). The shape of the curve varies by facility or region, the type and amount of renewable generation assets, time of day and the month of the year and can be difficult to predict.

Renewable energy-shaped price. The average price for power when a wind turbine, solar panel or other renewable generation asset is generating. The shaped price is calculated using the MWh production data from the **renewable energy shape curve**, which is matched in time with the energy spot market. The shaped price will be averaged over a defined period, such as a month or a year, and will typically be less than the average spot market price. The shaped price is key for assessing the economic viability of a renewable energy project.

Renewable generation capacity factor (CF). The percentage of the potential generating **capacity** of an intermittent source that will be converted to MWh of energy over a year. The CF is computed by calculating the annual MWh generated (accounting for intermittency) and dividing it by the MWh that would be delivered if the generation unit ran at full capacity for the full year. This enables an operator to estimate the power that a generating asset will provide. For example, a 10 MW solar facility operating at a 20% CF will deliver 17,520 MWh / year (0.20 * 10 MW * 24 hours / day *365 days / year *).

Renewable energy procurement criteria. A set of requirements governing renewable energy purchases. Setting the criteria will require collaboration with facilities, procurement and legal departments, and potentially cloud and / or colocation providers. There is no one-size-fits-all approach. Procurement criteria are particular to each operator and help organizations achieve their business and sustainability goals.

Renewable portfolio standard (RPS). A requirement that a minimum percentage of electricity consumption in state, province or country be sourced from renewable generation assets or from specific renewable generation types, such as solar. Governing bodies may use a RPS to force utilities to invest more in renewable energy. The term is most used in the US.

Resiliency or resilience. A data center’s ability to continue delivering its services through natural disasters, power outages, security breaches or any incident that could result in downtime is known as its resiliency. Resilient data centers may incorporate an uninterruptible power supply, extensive security measures, back-up storage and / or redundant equipment or systems. In some situations, but not all, achieving a high level of resiliency may make it more difficult (or expensive) to reach a high level of sustainability.

Science-based target (SBT). Carbon emissions reduction targets for a business or corporation that measures operational emissions against global reduction goals set by the 2015 Paris Agreement (with the aim to be net-zero by 2050). An operator will set an interim greenhouse gas emissions reduction target — typically with a duration of five to eight years — to move toward the net-zero goal. A science-based target can be self-certified or validated through a certifying body.

Scope 1 emissions. The direct emissions of CO2 or chemicals with a **GWP** (CO2 equivalents). These emissions result from the burning of fossil fuels, fugitive emissions from fossil fuel extraction and management systems or the use and emission of gases with GWP in manufacturing, agriculture or other business operations.

Scope 2 emissions. Indirect greenhouse gas emissions associated with the generation of electricity, steam, heat or cooling that is purchased for business operations. Scope 2 emissions are equal to the Scope 1 emissions generated in the production of the consumed energy commodity. The carbon associated with a data center's purchased energy consumption is classed as Scope 2 emissions.

Scope 3 emissions. Consist of greenhouse gas emissions that occur in a company's value chain and that are not included in Scopes 1 or 2. These include emissions from suppliers and from customers using a company's products or services. Scope 3 emissions, although often the hardest to account for, often constitute most of a data center operator's emissions.

Secondary water use. Water use associated with operations that supply energy or other commodities, primarily electricity, to a data center facility. A facility's overall impact on regional water supplies should be measured using the quantity of both primary and secondary water used.

Server Efficiency Rating Tool (SERT). A performance / power measurement tool set created by the Standard Performance Evaluation Corporation (SPEC). The software tool set consists of seven CPU worklets, two storage worklets and two memory worklets. A worklet is a special test program that mimics a specific server operation, such as data compression or a data set sort. As the worklets are run, a server's work **capacity** and power consumption are measured at defined **utilization** points. The individual worklet measurements are combined using the geometric mean function to calculate a single server workload / power efficiency measure. A higher SERT score indicates a more efficient server.

Storage capacity optimization methods (COMs). Software and / or hardware tools used to reduce the storage **capacity** required to store a given quantity of data and improve the **utilization** of the available storage capacity. Storage COMs include data deduplication, thin provisioning, compression and delta snapshots. COMs can reduce energy use and save on capital equipment.

Sustainability strategy. A flexible set of goals, practices, processes and policies intended to achieve meaningful improvements in the environmental performance of a data center or critical infrastructure. Typical goals include the reduction of greenhouse gas emissions, increased energy efficiency relative to the workload, reductions in water use and increased use of renewable energy.

Unbundled renewable energy certificates (RECs). Traded separately, usually in a different grid region, from the energy that created them. Unbundled **RECs** may be purchased to enable a buyer to match / offset their electricity consumption with an equivalent amount of renewable energy. The avoided emissions factor of the unbundled RECs must typically be matched with the **grid emission factor** of the grid region in which the RECs are applied. Unbundled RECs do not guarantee that the renewable energy is consumed or that new or additional renewable energy capacity was built to match demand. See also **Bundled RECs**.

Utilization. The amount of work that is being performed relative to the total capacity of the IT equipment or data center capacity, expressed as a percentage. Utilization can also be applied to power and cooling system capacity. Tracking of average IT equipment utilization is a good sustainability metric, as the work delivered per unit of energy increases with an increase in IT equipment utilization.

Virtualization. A software-enabled technique (managed by a hypervisor) that permits many software programs to share access to CPU cores, memory, storage and network resources. The use of virtualization enables greater software portability, higher server **utilization** and increased energy efficiency.

Virtual power purchase agreement (VPPA). A type of **power purchase agreement** in which renewably generated energy is not physically delivered to a buyer but is instead sold into the local wholesale market. The buyer takes possession of the r**enewable energy certificates** generated under the VPPA. With a VPPA, the buyer agrees to a fixed "strike price" (purchase price) for the output from a fixed generation **capacity** for the full term of the VPPA. VPPAs are typically settled each month. If the aggregate revenue from the sale of the energy is higher than the aggregate cost of the energy to the buyer, the generator pays the difference to the buyer. If the aggregate revenue is lower, the buyer pays the difference to the generator. VPPAs can act as a partial hedge against market fluctuations, but they can also represent a significant financial risk as renewable generation capacity increases in a grid region.

Water usage effectiveness (WUE). A ratio of annual data center water use (in liters), divided by the annual IT equipment's energy use in kWh. Similarly to the power usage effectiveness metric for energy, WUE can show how efficiently the facility is managing its water consumption. The value of this metric depends on the type of cooling system used at the data center, with very different values for water- and air-based systems.

Appendix

Reports in the series

'Digital infrastructure sustainability — A manager's guide' is a series of Uptime Institute Intelligence reports that guide managers and responsible operators through the complex set of issues involved in creating an environmental sustainability strategy for data centers and associated digital infrastructure. The reports outline the strategies, the terminologies and the approaches involved; explain the key areas that must be covered by an environmental sustainability strategy and how progress should be measured and reported; and provide guidance on how to navigate some of the intricacies and challenges involved. These reports do not constitute a blueprint environmental strategy or a new standard but provide pragmatic, actionable advice for managers who must navigate this critical area.

The series includes the following reports:

Creating a sustainability strategy

All those who operate digital infrastructure must have a sustainability strategy that spans all facilities and IT operations and addresses the needs of all stakeholders.

Tackling greenhouse gases

Operators of digital infrastructure must have a greenhouse gas emissions reduction goal that takes into account Scope 1, 2 and 3 emissions — and they must report these reductions in accordance with agreed policy.

Reducing the energy footprint

The first objective of a sustainability plan is to minimize energy use through efficiency measures. Further benefits will be realized by replacing electricity from nonrenewable sources with renewably generated energy.

IT efficiency: the critical core of digital sustainability

A digital sustainability strategy should incorporate both the facilities and IT operations, even for colocation operators. This report covers strategies, software tools and metrics that can help drive up IT efficiency.

Three key elements: water, circularity and siting

This report discusses three important elements of the sustainability strategy: water use; siting, including design and certification; and reuse, disposal and recycling. Addressing these elements can significantly reduce the environmental impact of digital infrastructure.

Navigating regulations and standards

Critical digital infrastructure is subject to an expanding set of regulations, directives and standards, with varying levels of maturity and acceptance. Most are voluntary, but more are becoming mandatory.

Glossary of digital infrastructure sustainability

This document explains the key terms used by those defining, regulating and applying digital infrastructure sustainability strategies.

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers – the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions. With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.

Uptime Institute is headquartered in New York, NY, with offices in Seattle, London, Sao Paulo, Dubai, Singapore, and Taipei.

For more information, please visit **www.uptimeinstitute.com**

All general queries:

Uptime Institute
405 Lexington Avenue,
9th Floor, New York,
NY 10174, USA
+1 212 505 3030
info@uptimeinstitute.com