

Data Center

A data center is a dedicated possibly distributed space hosting computational systems and associated components, such as telecommunications and storage systems. Since IT operations are crucial for continuity, it generally includes some redundant or backup components and infrastructure for power supply, data communication connections, environmental controls (air conditioning, fire suppression) and various security devices.

Data centers can vary widely in terms of size, power requirements, redundancy and overall structure. Four common categories used to segment types of data centers are

- onsite data centers
- colocation facilities
- hyperscale data centers
- edge data centers

Dynamic infrastructure provides the ability to intelligently, automatically and securely move workloads within a data center, anytime anywhere for migrations, provisioning, to enhance performance or building co-location facilities. It also facilitates performing routine maintenance on either physical or virtual systems all while minimizing interruption. A related concept is composable infrastructure, which allows for the dynamic reconfiguration of the available resources to suit needs, only when needed. Side benefits include the ability to enable business continuity and high availability of cloud and grid computing.

Micro data centers are access-level data centers which are much smaller in size than traditional data centers but basically provide the same features. They are typically located near the data source to reduce communication delays. Small size allows several micro data centers to be spread out over a wider area. Micro data centers are well suited to user-facing front end applications. They are commonly used in edge computing and other areas where low latency data processing is needed.

Edge computing is a distributed computing paradigm that brings computation and data storage closer to the sources of data, so that a user of a cloud application is likely to be physically closer to a server than if all servers were in one place. This is meant to make applications faster. More broadly, it refers to any design that pushes computation physically closer to a user, so as to reduce the latency compared to when an application runs on a single data center. In the extreme case, this may simply refer to client side computing.

See also SHDCS.