Home   >   Cloud and data center networking

≡ | **Networking**   🔍

Clos network

By **Robert Sheldon** | **Jessica Scarpati** | **Tom Nolle,** Andover Intel

What is a Clos network?

A Clos network is a type of nonblocking, multistage switching network used today in large-scale data center switching fabrics. The concept was introduced in the 1950s to increase the efficiency of telephone switching networks and help lower their costs. The Clos architecture proved highly effective in routing calls on the available mechanical telephone systems at the time, although it eventually gave way to technologies that supported more modern systems.

However, the principles behind the Clos network reemerged in data center networks in the 1990s, offering an effective topology for optimizing the switch operations in Ethernet data communications. To this day, the Clos network continues to serve as a foundation for many interconnected fabrics, including those in large-scale data centers that drive the cloud.

What is Networking (Computer Networking)?

History of the Clos network

Clos networks are named after Bell Labs researcher Charles Clos, who first proposed his network design in 1952. He proved that his model could address the reliability and cost challenges that came with the electromechanical switches used in telephone networks at the time. The concept of a Clos network was introduced by Edson Erwin in 1938, although it was Charles Clos who demonstrated its practical application in [telephone networks](#).

Clos came up with a network design for delivering nonblocking, any-to-any connectivity that minimized the number of crosspoints required to support call switching. Crosspoints are the electromechanical relay mechanisms in a crossbar switch, a type of matrix switch once used extensively to route phone calls. Before the Clos network was introduced, the number of crosspoints had to equal the number of inputs multiplied by the number of outputs. This approach, known as *n-squared* or n^2, could quickly add up to a massive number of crosspoints, resulting in significant costs.

Clos used mathematical theory to prove that it was possible to achieve nonblocking connectivity in a switching array, now known as a *fabric*. To achieve this connectivity, the switches were organized into a three-stage architecture that included a middle stage sandwiched between the ingress (input) and egress (output) stages. In this design, each input switch connects to each middle switch, which, in turn, connects to each output switch, as shown in Figure 1. The result was a nonblocking topology that required fewer crosspoints than the more conventional switch networks at the time.
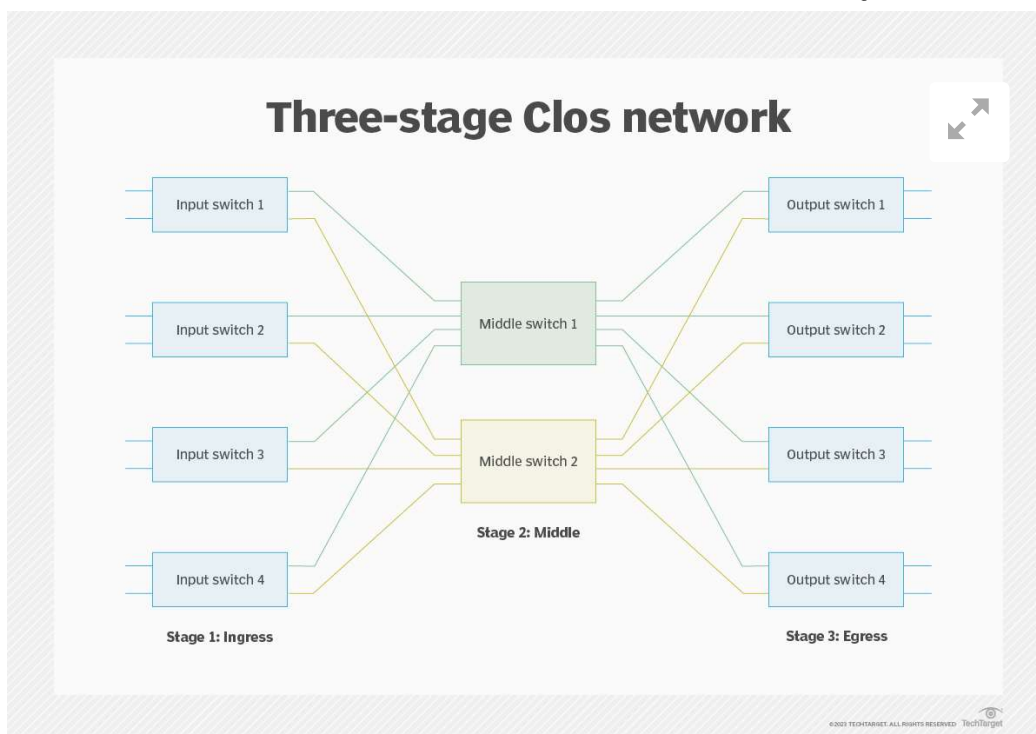
Figure 1. Diagram illustrating three-stage Clos network architecture

The Clos network made it possible for telephone calls to travel different paths and avoid being blocked by other calls, which often occurred in older networks that relied on dedicated connections. In addition, the Clos topology could be scaled out to accommodate greater call volumes, while still delivering the same level of service.

The Clos design in the modern data center

Electromechanical switching eventually gave way to newer technologies, eliminating the need for the Clos design. However, because of the efficiency gains it offered, the Clos approach resurfaced in the 1990s with the emergence of high-performance switches in data center fabrics. A Clos network could provide nonblocking connectivity across an interconnected Ethernet switch fabric without the need for n-squared ports.

In many of today's data centers, the Clos architecture is implemented in a leaf-spine layout in which the spine layer represents the switches in the middle stage and the leaf layer represents the switches in both the ingress and egress stages, as shown in Figure 2. This design is often referred to as a *folded Clos network* because it is based on the idea of folding the middle layer along its axis, resulting in a conceptual layout of two tiers. This design is also sometimes referred to as a *fat-tree network*.
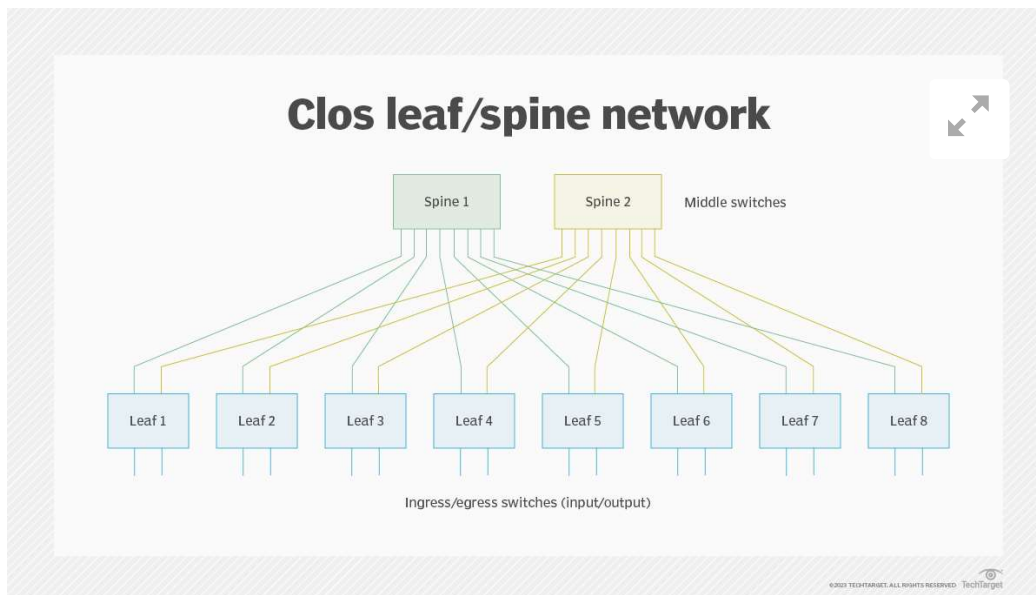
Figure 2. Diagram illustrating folded Clos leaf/spine network architecture

The efficiency of a Clos network is proportional to the number of ports in the network. Larger networks that contain thousands or tens of thousands of ports benefit the most from the Clos design. It is highly reliable and works exceptionally well for east-west communications. In east-west communications, data packets flow from server to server in the data center rather than switch to switch (north-south). It also supports massive scale-out growth, making it possible to add spine and leaf switches with minimal impact on the overall topology.

The Clos design is sometimes used to create five-stage networks, rather than three. In this scenario, each spine in the middle stage is broken into a smaller Clos network, with its own ingress, middle and egress stages. Although a five-stage design adds planning and maintenance overhead, it enables data centers to interconnect multiple pods into a single fabric.

The Clos architecture can also be extended beyond five stages. It's possible to create Clos networks with seven stages or even more, as long as there is an odd number of stages. However, these types of Clos networks are not as common as three-stage and five-stage networks.