

Rete nei data center

I data center sono un agglomerato di host posizionati vicini per poter supportare applicazioni e servizi in rete, le principali sfide sono:

- Gestire applicazioni con molti utenti che le usano
- Affidabilità del servizio
- Gestione del carico e bilanciamento del traffico

La rete di un data center

Un data center è composto da una rete con diversi elementi, principalmente riconosciamo:

- **Border router:** si trovano a livello più alto e sono responsabili con la comunicazione esterna del data center, possono gestire grande traffico in entrata e in uscita.
- **Tier-1 switch:** sono connessi direttamente con i border routers, servono per fornire ridondanza e distribuzione dei dati nell'intero data center, ogni tier-1 è connesso a circa 16 tier-2, chiamati anche **spine switch**.
- **Tier-2 switch:** servono per connettere i tier-1 switch con i TOR switches i quali sono fondamentali per l'interconnessione della rete, chiamati anche **fabric switch**.
- **TOR switch:** sono switch collegati alla cima del rack (Top Of Rack), servono come punto d'accesso al rack di server.
- **Rack di server:** ogni rack di server è composto da 20-40 server blade ovvero l'unità di calcolo di un rack, ogni blade è un dispositivo fisico che esegue applicazioni o salvano dati.

Multipath: ricca interconnessione tra switch e rack per consentire la presenza di strade multiple per lo stesso rack, così da non rendere SPOF il collegamento tra switch e rack.

Load Balancer: è una macchina che gestisce il routing a livello rete o applicativo, ovvero che verifica le richieste dall'esterno, le dirige nel datacenter e restituisce il risultato all'esterno.