Design and Engineering of Computer Systems

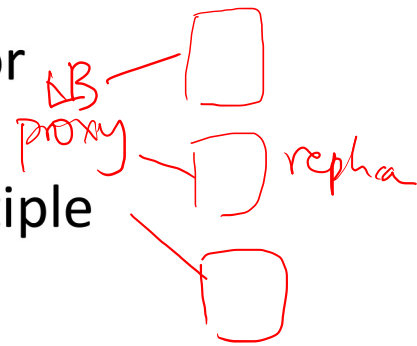
Lecture 30: Deployment of computer systems

Mythili Vutukuru

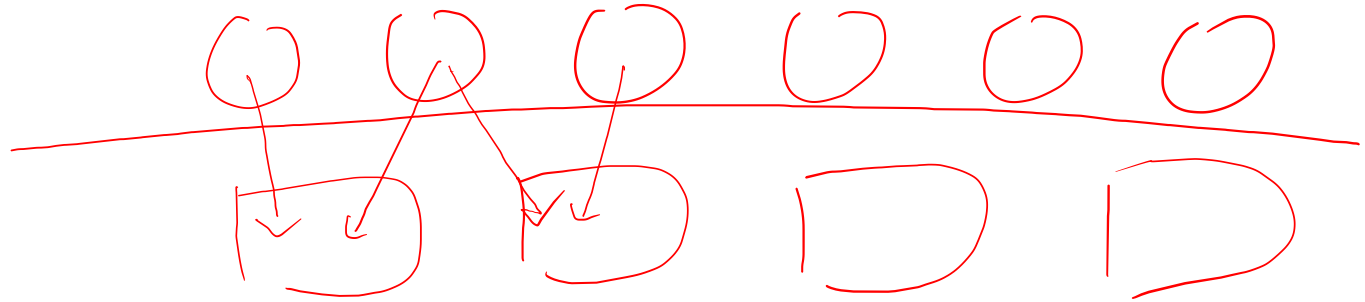
IIT Bombay

Deployment and maintenance

- Real-life computer systems are complex
  - Many logical components (frontend, backend, web servers, application servers, databases, ..)
  - Each component can have multiple micro-services/processes/threads for different purposes
  - Some additional components like proxy server to distribute load to multiple replicas for load balancing (more later)
- How are such large systems deployed and managed in real life?
  - How to run and manage so many components on multiple machines in a cluster in a production system?

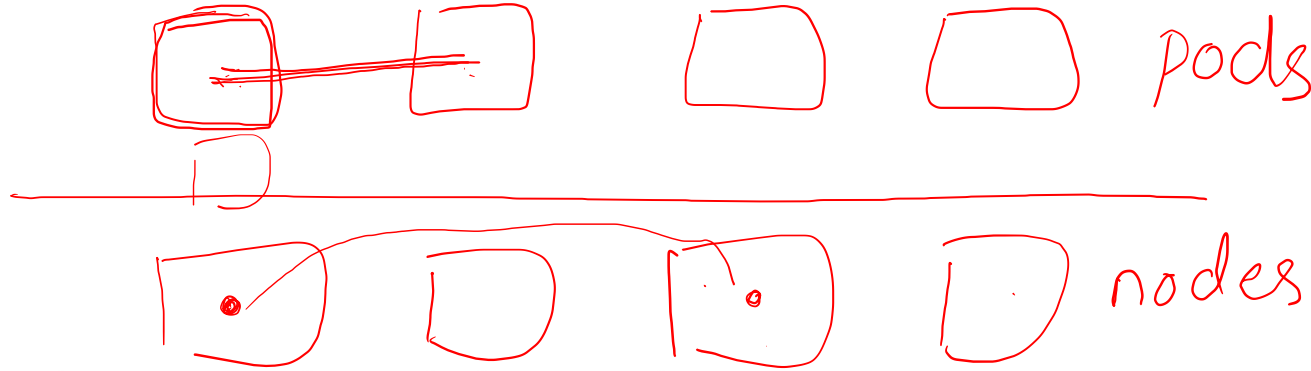


Cloud deployment



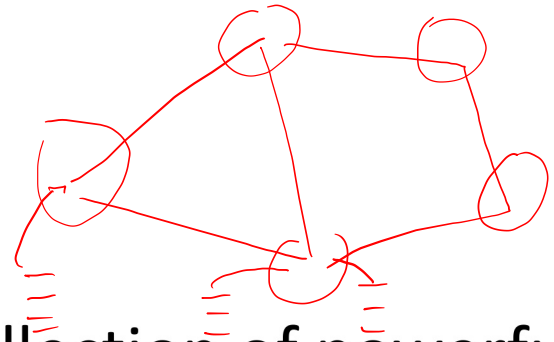
- Components can be run directly (bare metal) as processes on OS and hardware, or inside a VM or container
- VMs and containers can be hosted on a private cloud (setup by an organization) or on a public cloud managed by cloud service providers
- Why deploy using virtualization on cloud?
  - Efficient sharing of hardware resources across VMs / containers
  - Easy to package a component and all its dependencies in a VM / container
  - Cloud management/orchestration software makes it easy to manage multiple VMs / containers
- Further benefits of public clouds
  - Hassle free maintenance of hardware and system software
  - Public clouds provide various components like load balancers and data stores directly for use
  - Platform as a service model makes it easy to develop applications, developers can just focus on application logic

Cloud orchestration

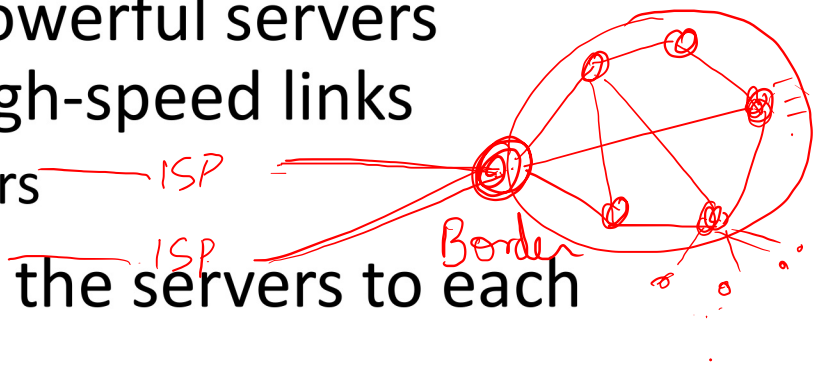


- Cloud orchestration software (e.g., Kubernetes) widely used to manage deployments of computer systems. What functionality does orchestration software provide?
- Example: Kubernetes applications consist of multiple “pods” which contain different application components / micro-services
  - A pod is one or more containers with application components that should be located together (along with all the dependencies)
  - Pods are instantiated on multiple physical “nodes” based on resource availability
  - The network connectivity between various pods is suitably configured by setting up routing/forwarding rules in the underlying network
  - Orchestration software manages the entire lifecycle of the pods (creation, configuration, execution, restart on crash, …)
  - Monitoring software monitors the usage of various resources (e.g., CPU utilization) and takes suitable actions (e.g., auto-scaling) as required

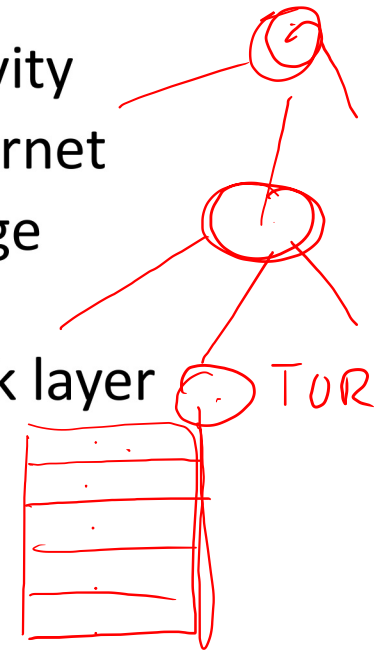
Network architecture



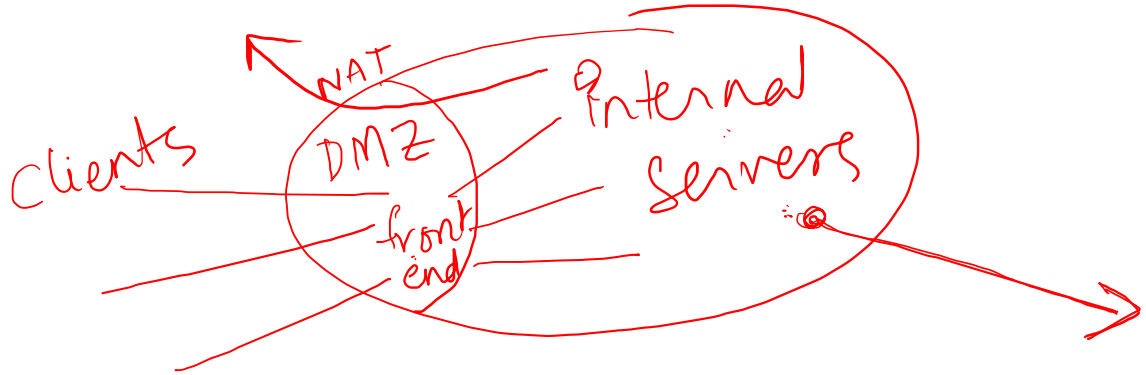
- Computer systems run on a cluster/collection of powerful servers connected by networking switches/routers over high-speed links
  - Cluster of servers hosted in private or public data centers



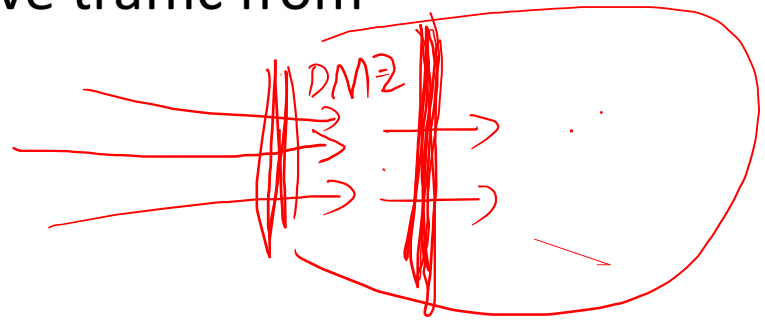
- Multiple IP routers and link-layer switches connect the servers to each other, and to clients over the Internet
  - Border (BGP) routers connected to one or more ISPs for Internet connectivity
  - IP prefixes of network are announced by BGP routers via ISP to rest of Internet
  - Internal IP routers connected to border routers and to each other, exchange routes to internal hosts
  - Traffic arrives at border routers, then forwarded via internal IP routers, link layer switches to servers
  - Servers in racks connect to top-of-rack (ToR) switch, and other switches



Firewall and DMZ

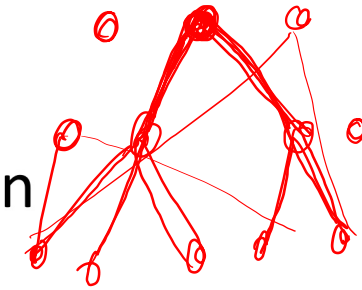


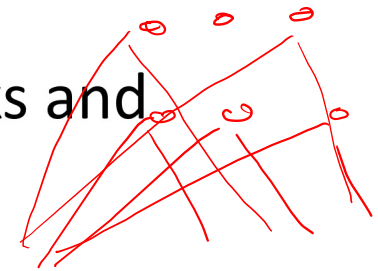
- DMZ (demilitarized zone) is the network edge that has public-facing servers which receive external traffic (e.g., web servers, mail servers, front end)
  - DMZ servers have public IP addresses, internal servers can use private IP addresses
  - Network Address Translators (NATs) assign temporary public IP addresses (and rewrite headers) on traffic coming from internal private IP addresses
- Firewalls at entry points filter unwanted incoming/outgoing traffic
  - More lenient firewall rules for servers in DMZ that receive traffic from external clients
  - More stringent filtering for non-DMZ internal servers
  - Firewalls can be software or hardware appliances

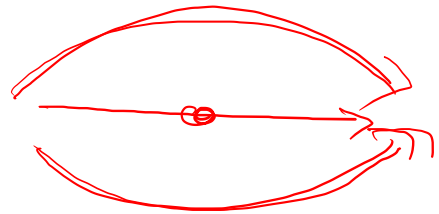


Network optimization

- Many techniques used to ensure good network connectivity from clients to servers, and between servers, in a computer system
- High-speed networking interfaces and switches (hundreds of Gbps) are used to connect hosts in a data center / cluster
- Network topology is carefully designed to maximize bandwidth between servers in a cluster
  - Simple tree topology may have bottleneck at top of tree, so “fat tree” topologies with multiple parallel paths between servers are used
- Careful traffic engineering to minimize network congestion bottlenecks and to distribute traffic across the multiple network paths
  - Techniques like SDN, label switched routing can be used for traffic engineering
- Special transport protocols optimized for high BDP networks
- Offload static content to content distribution networks (CDNs)







Summary

- In this lecture:
  - How computer systems are deployed and managed
- Install any container orchestration framework (e.g., Kubernetes) on your computer. Start a few containers with it and explore what functionality you get with such frameworks.