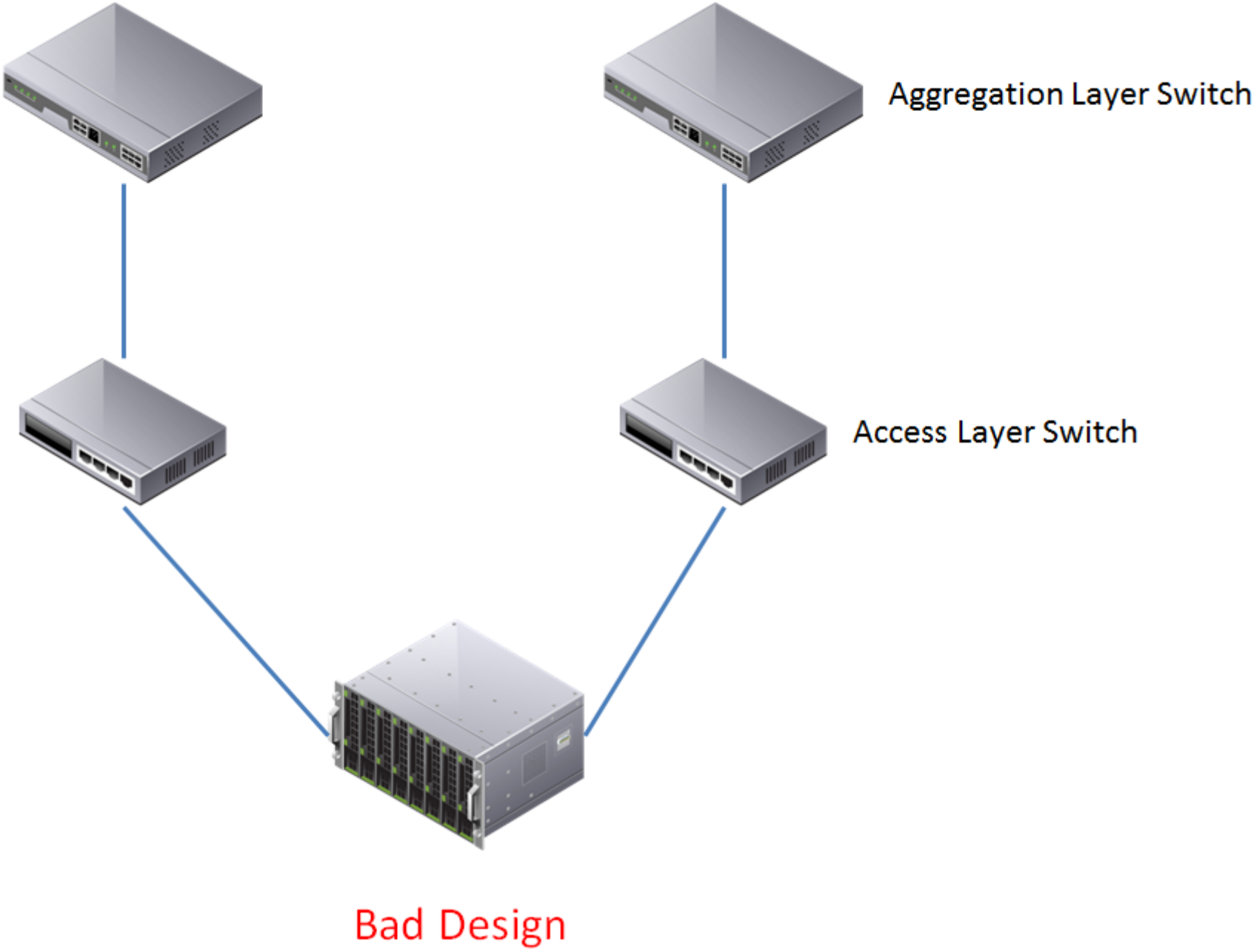
Design Considerations – Virtualized Datacenter Networking

When switching was introduced in Hypervisor Layer, inter VM switching started happening locally and there was no visibility from external network. Now let me introduce two very basic concepts and those are Access port and Trunk port in a Switch.

- Access Switch Port Configuration requires a physical NIC for each VLAN
- Trunk Port allows multiple VLANs on a single interface

If you do not have Links between Access Layer Switches then it will introduce variable performance due to multiple hops. However placing links in between Access Layer switch will complicate the STP environment. So you can see there is a Tradeoff.



If you have connectivity between Access and Aggregation Layer in a Active/Standby then STP disables one path to prevent loops. However, you still can manually load balance VLANs.

Consider Oversubscription in these layers:

- Access Layer to Aggregation Layer
- Aggregation Layer to Core
- Traffic Patterns

So lets get to bottom of these design considerations. When the classic hosts are replaced with hypervisors, another layer of switching is added. The hypervisor virtual switches extend the Access Layer and perform Layer 2 switching within the hypervisor itself. So VMs that are on the same hypervisor and VLAN will have traffic between them switched locally. VMs on different hypervisors, or on different VLANs will require data to be sent to the Access or Aggregation Layer before being sent back to the destination VM.

Because hypervisors have multiple VMs connecting to the network over the same physical infrastructure, you need to determine how to configure the switch ports on the Access Layer, as well as the networking on the hypervisor. If all of the VMs on a particular hypervisor reside on a single VLAN, then you can configure the switch ports as Access Ports, and completely ignore VLANs at this level. If, however, you have to support multiple VLANs on a hypervisor, you need to use either multiple Access Ports or a Trunk Port.

Since Access Ports can only support a single VLAN, a physical connection from each hypervisor to the switch is needed for each VLAN that you need to support. This configuration is viable if you are supporting only a small number of VLANs and that is up to the maximum number of NICs that your hypervisor can support. If HA is a concern, then you can only support half of that number. This approach provides a more customized configuration for different hypervisors.

Trunk ports allow you to have a large number of VLANs configured on a single interface. This allows for a more flexible configuration for each hypervisor. You can configure each hypervisor to support all applicable VLANs, and then there is no concern when a VM is moved from one host to another. This approach is much more desirable when using 10 Gigabit Ethernet networks, as deploying a NIC for each VLAN is not cost effective.

Many of the design principles for a classic data center still apply in a Virtualized Data Center. Those are:

- Do not establish links between Access Layer switches. Because these switches operate at Layer 2, placing links between the switches will complicate the STP environment, and ultimately result in most, if not all, of the links being disabled by STP.
- Connectivity between the Access and Aggregation layer is Active/Standby. Again, STP will disable one path to prevent loops, so in effect, the bandwidth between the layers is halved. If you want to utilize the full bandwidth, you will need to manually configure one link as primary for certain VLANs, and the other link as primary for other VLANs.
- Oversubscription also need to be taken into account. You will most likely not have a 1:1 ratio of host ports to uplink ports between the Access and Aggregation Layers, and this will be further consolidated between the Aggregation and Core Layers. You will need to understand the amount of traffic being sent and received by each host, as well as where the traffic is going to and coming from. Remember that traffic between devices on the same VLAN and connected to the same Access Layer switch is switched locally. If related systems (e.g., application and database server) are strategically placed, a significant amount of traffic may not even reach the Aggregation Layer.