CERN IT Department

Computer Virtualization from a network perspective

Jose Carlos Luna Duran - IT/CS/CT

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Agenda

CERN IT Department

- Introduction to Data Center networking
- Impact of virtualization on networks
- VM machine network management

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Part I

CERN IT Department

Introduction to Data Center Networking

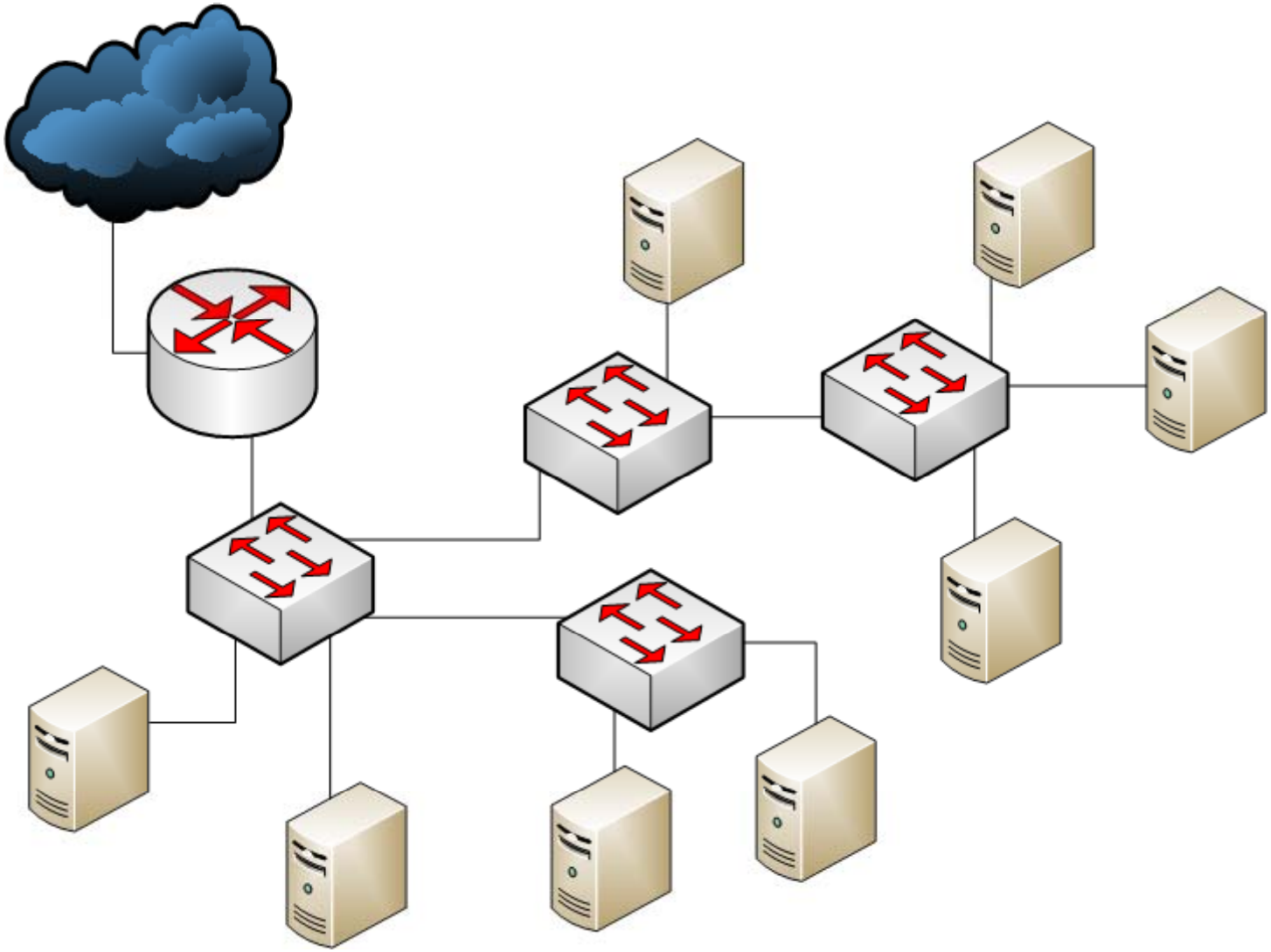
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Data Centers

CERN IT Department

- Typical small data center: Layer 2 based



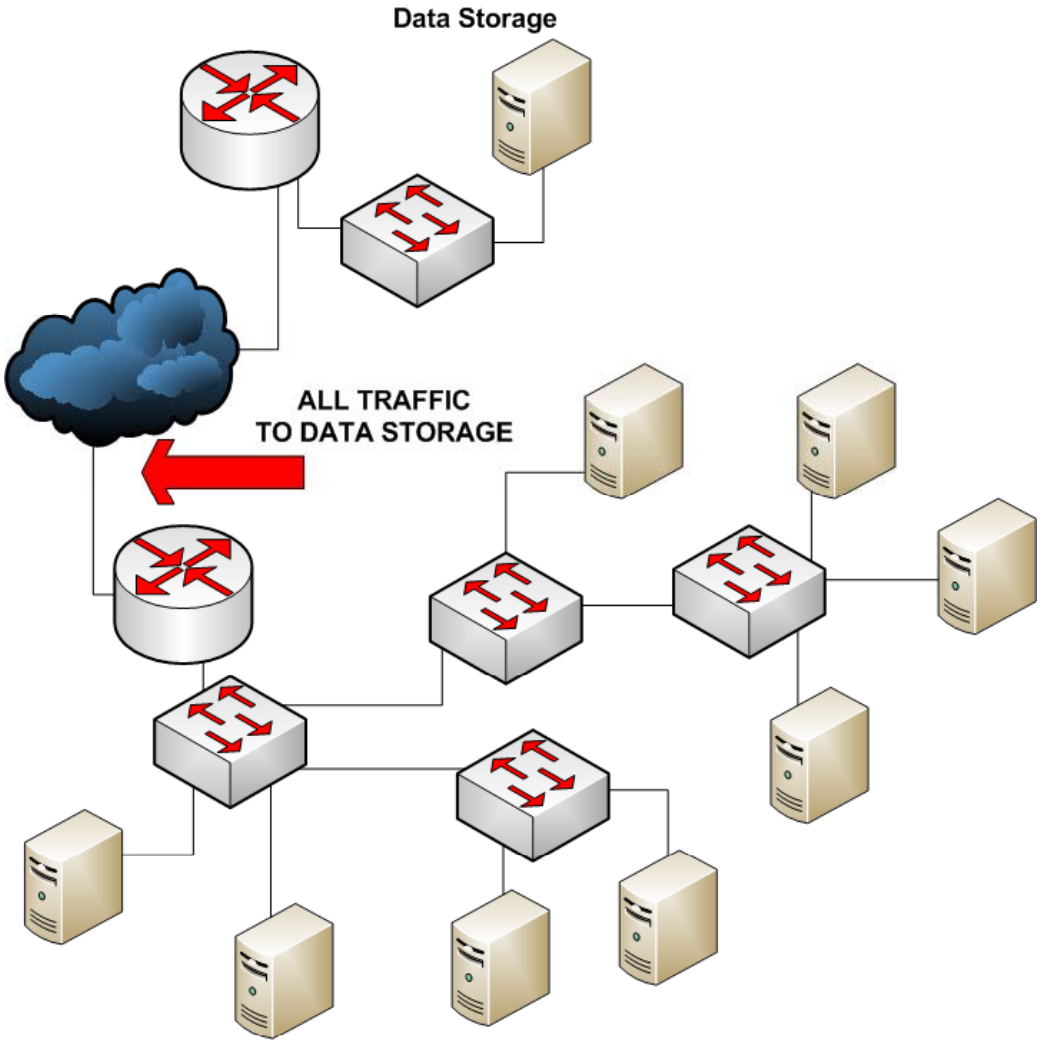
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Layer 2 Data Center

- Flat layer two Ethernet network: same broadcast domain.
- Appropriate when:
  - Network traffic is very localized.
  - Same responsible for the whole infrastructure
- But…
  - Uplink shared with a big number of host.
  - Noise from other nodes (broadcast): problems may affect the whole infrastructure.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Data Center L2: Limitations

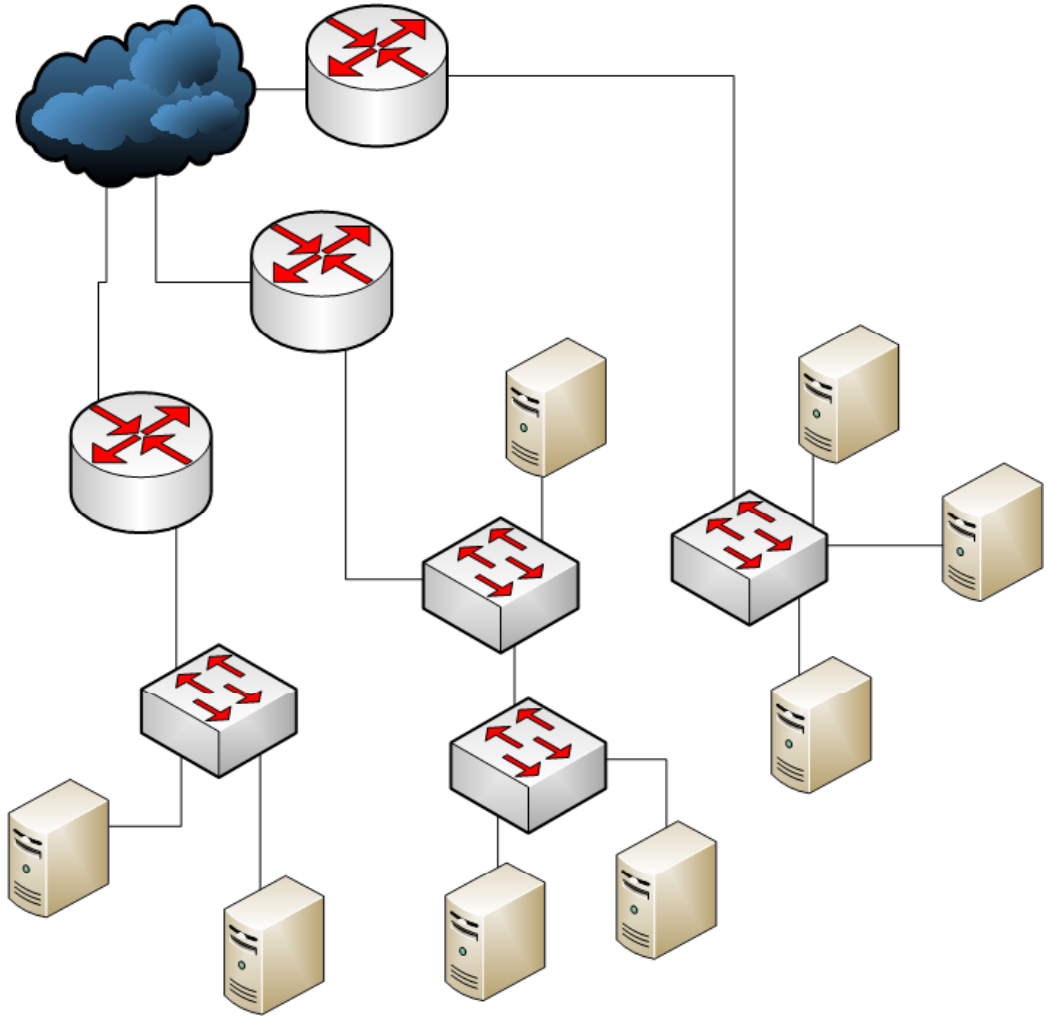


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Data Center L3

CERN IT Department

- Layer 3 Data Center



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Data Center L3

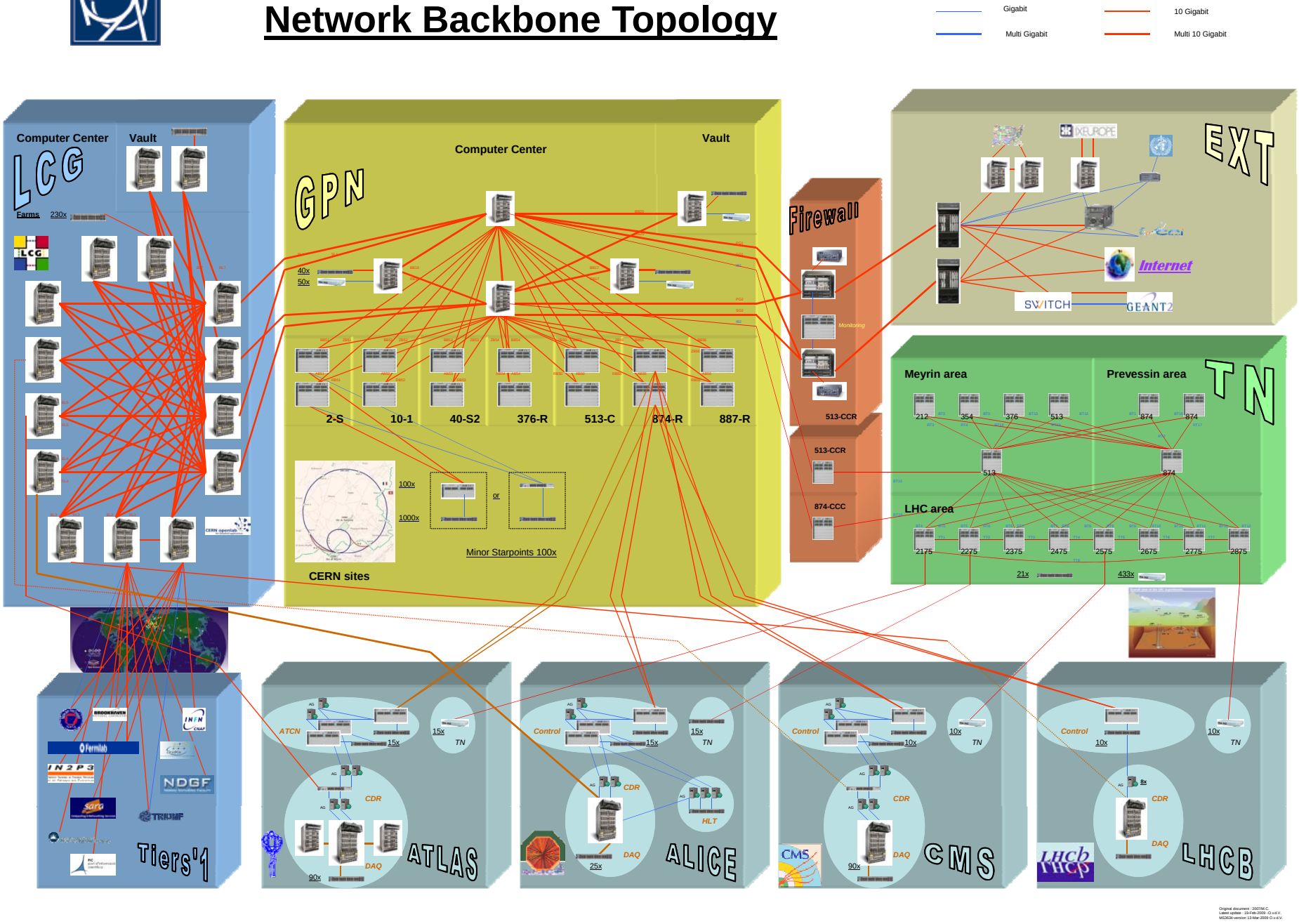
CERN IT Department

- Advantages:
  - Broadcasts are contained in small area (subnet)
  - Easier management and network debugging.
  - Promotes “fair” networking (all services point-to-point are equally important).
- But…
  - Fragmentation of the IP space.
  - Move from one area (subnet) to another requires IP change.
  - Needs a performing backbone.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Network Backbone Topology



CERN Network

CERN IT Department

- Highly Routed (L3 centred)
  - In the past several studies where done for localizing services -> Very heterogeneous behaviour: did not work out.
  - Promote small subnets (typical size: 64)
  - Switch to Router: 10 Gb uplink
- Numbers:
  - 150+ Routers
  - 1000+ 10Gb ports
  - 2500+ Switches
  - 70000+ 1Gb user ports
  - 40000+ End nodes (physical user devices)
  - 140 Gbps WAN connectivity (Tier 0 to Tier 1) + 20 Gbps General Internet
  - 4.8 Tbps at the LCG backbone CORE

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Part II

CERN IT Department

Impact of virtualization on networks

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Types of VM Connectivity

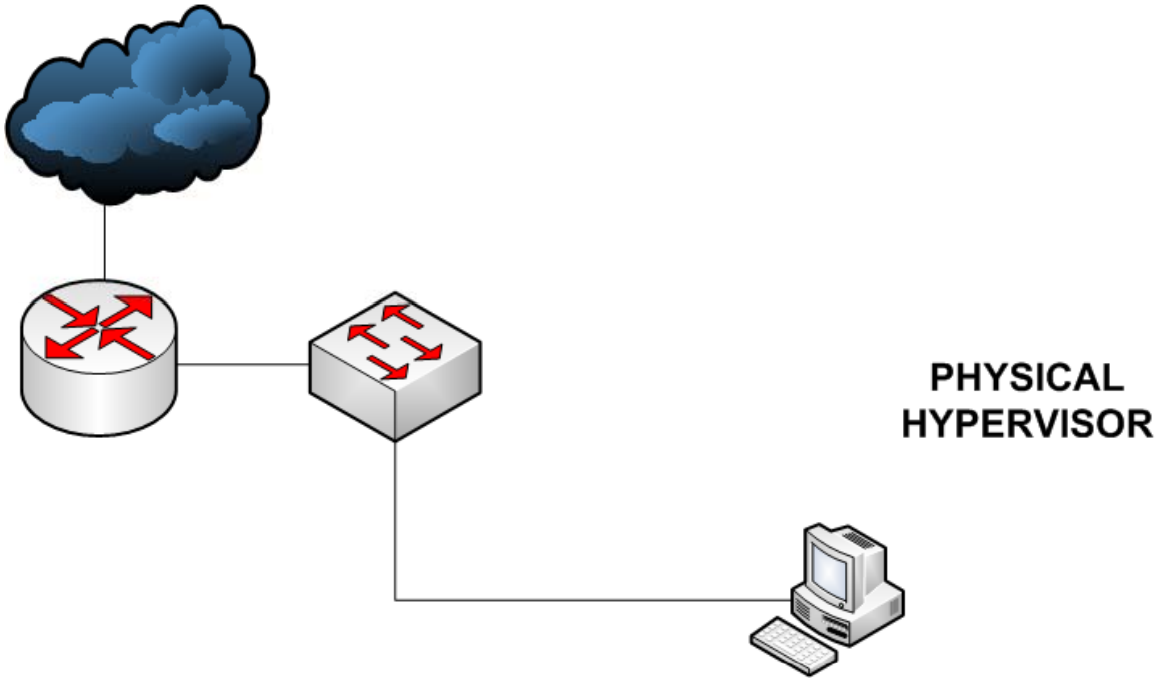
Virtual Machine hypervisors offer different connectivity solutions:

- Bridged
  - Virtual machine has its own address (IP and MAC).
  - Seen from the network as a different machine.
  - Needed when incoming IP connectivity is necessary.
- NAT
  - Uses the address of the HOST system (invisible for us).
  - Provides offsite connectivity using the IP of the hypervisor.
  - NAT is currently not allowed at CERN (for debugging and traceability reasons).
- Host-Only
  - VM has no connectivity with the outside world

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Bridged and IPv4

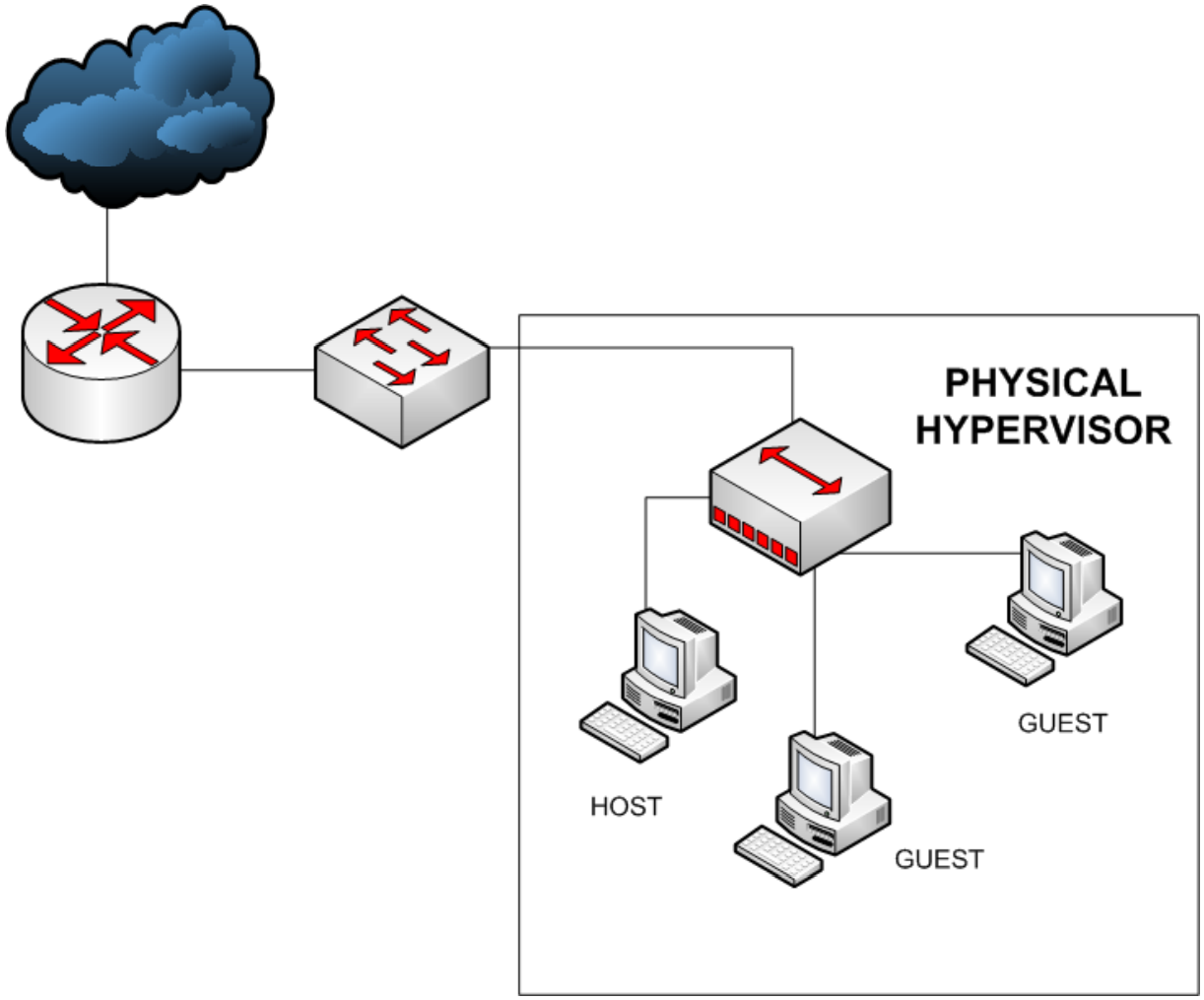
- For bridged reality is this:



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Bridged and IPv4

- Observed by us as this:



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Bridged and IPv4 (II)

- It's just the same as a physical machine, therefore should be considered as such!
- Two possibilities for addressing:
  - Private addressing
    - Only on-site connectivity
    - No direct off-site (NO INTERNET) connectivity
  - Public addressing: best option, but…
    - Needs a public IPv4 address
    - IPv4 address is limited.
    - IPv4 address allocation: IPv4 address are given in form of subnets (no single IPv4 addresses around the infrastructure)-> Fragmentation -> Use wisely and fully.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Why not IPv6?

- No address space problem, but:
  - ALL computers that the guest wants to contact would have to use IPv6 to have connectivity.
  - IPv6 “island” would not solve the problem
    - If these machines need IPv4 connectivity IPv6 to IPv4 conversion is necessary.
    - If you have to map each IPv6 address to one IPv4 address we are hitting the same limitations as IPv4.
  - All applications running in the VM should be IPv6 compatible.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Private Addressing

- Go for it whenever possible! (space not as limited as if we use public addresses).
- But… no direct off-site connectivity (perfect for the hypervisors!)
- Depends on the use case for the VM

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

NAT

- Currently not allowed at CERN:traceability...
- NAT where?
  - In the Hypervisor
    - No network ports in the VM would be reachable from outside.
    - Debugging network problems for VMs impossible
  - Private addressing in the VM and NAT in the Internet Gate:
    - Would allow incoming in-site connectivity
    - No box capable of handling 10Gb+ bandwidth
  - Distribution Layer (access to the core)
    - Same as above plus more number of high speed NAT engines required.
- No path redundancy possible with NAT!

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Recommendations

CERN IT Department

- Everything depends on the behavior of the VM and its intended usage.
- Public addresses are a scarce resource. Can be provided if limited in number.
- Use private addressing if there is no other special need besides the use of local on-site resources.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Part III

VM machine network management.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CS proposed solutions

CERN IT Department

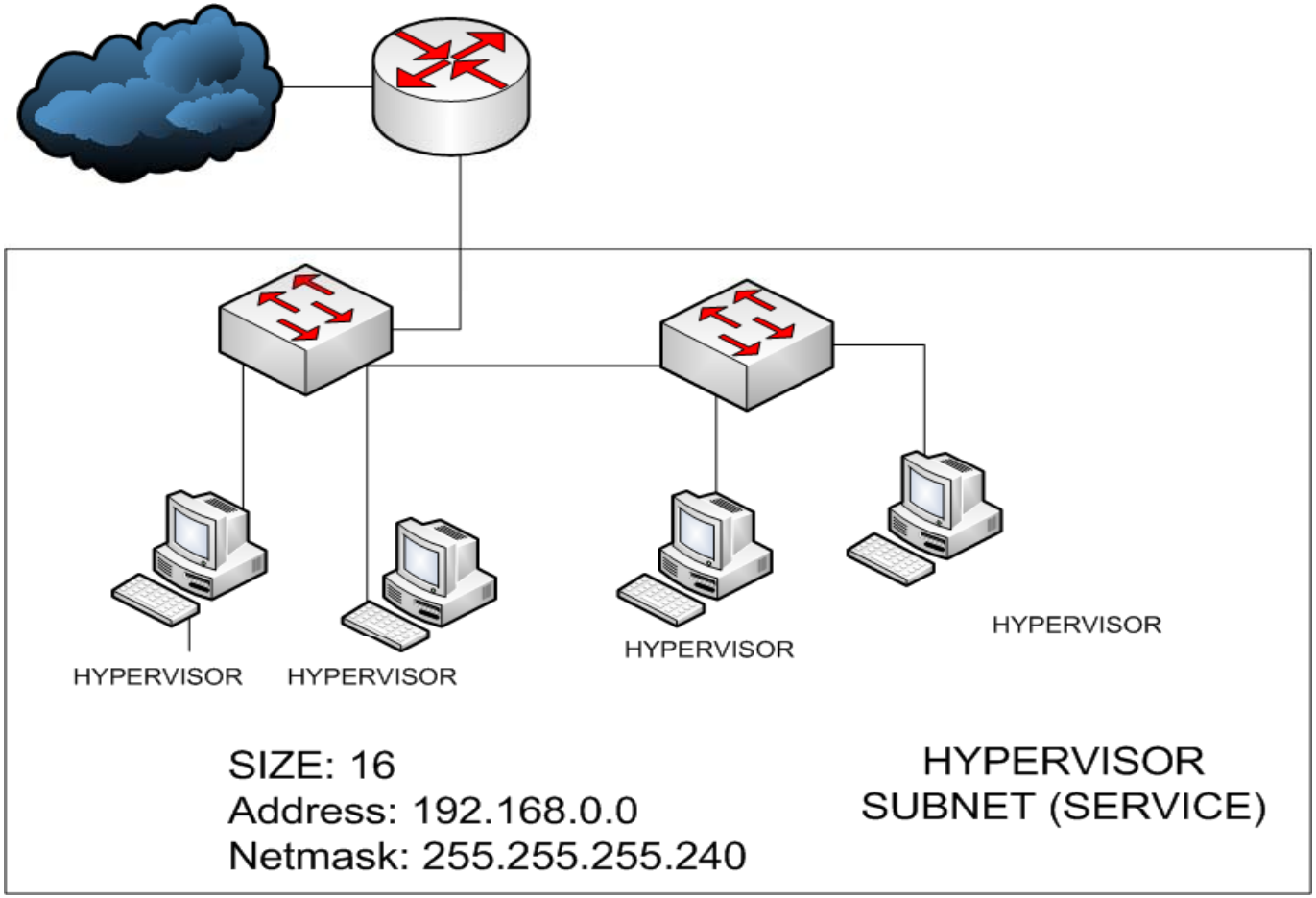
- For desktops:
  - Desktops are not servers, therefore…
  - NAT in the hypervisor proposed:
    - Responsible of the hypervisor is the same as responsible of VMs
- VM as a service (servers, batch, etc…):
  - For large number of VMs (farms)
  - Private addressing preferred
  - VMs should not be scattered around the physical infrastructure.
  - Creation of the “VM Cluster” concept.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

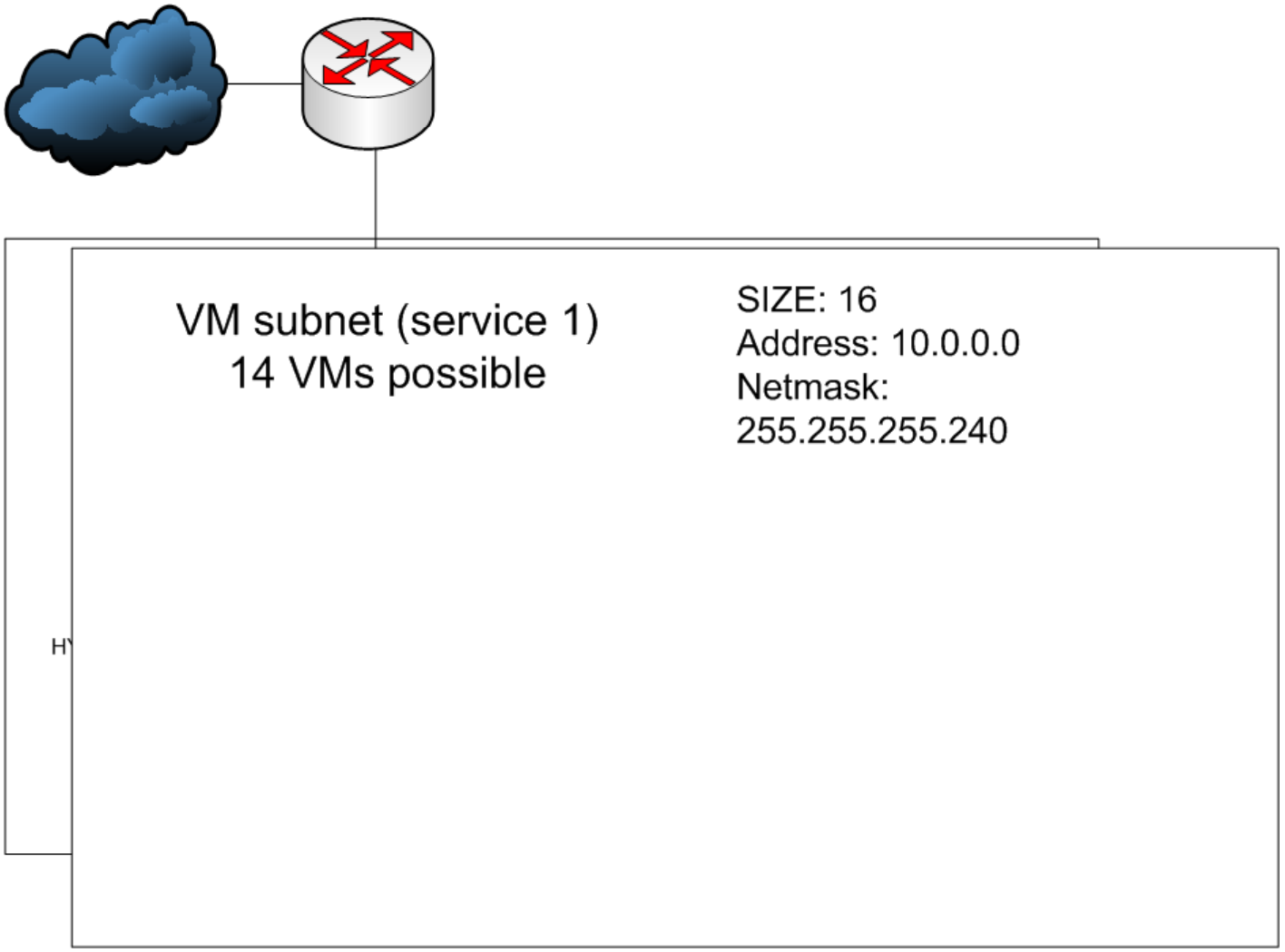
VM Clusters

- VM Cluster: separate set of subnets running in the SAME contiguous physical infrastructure:



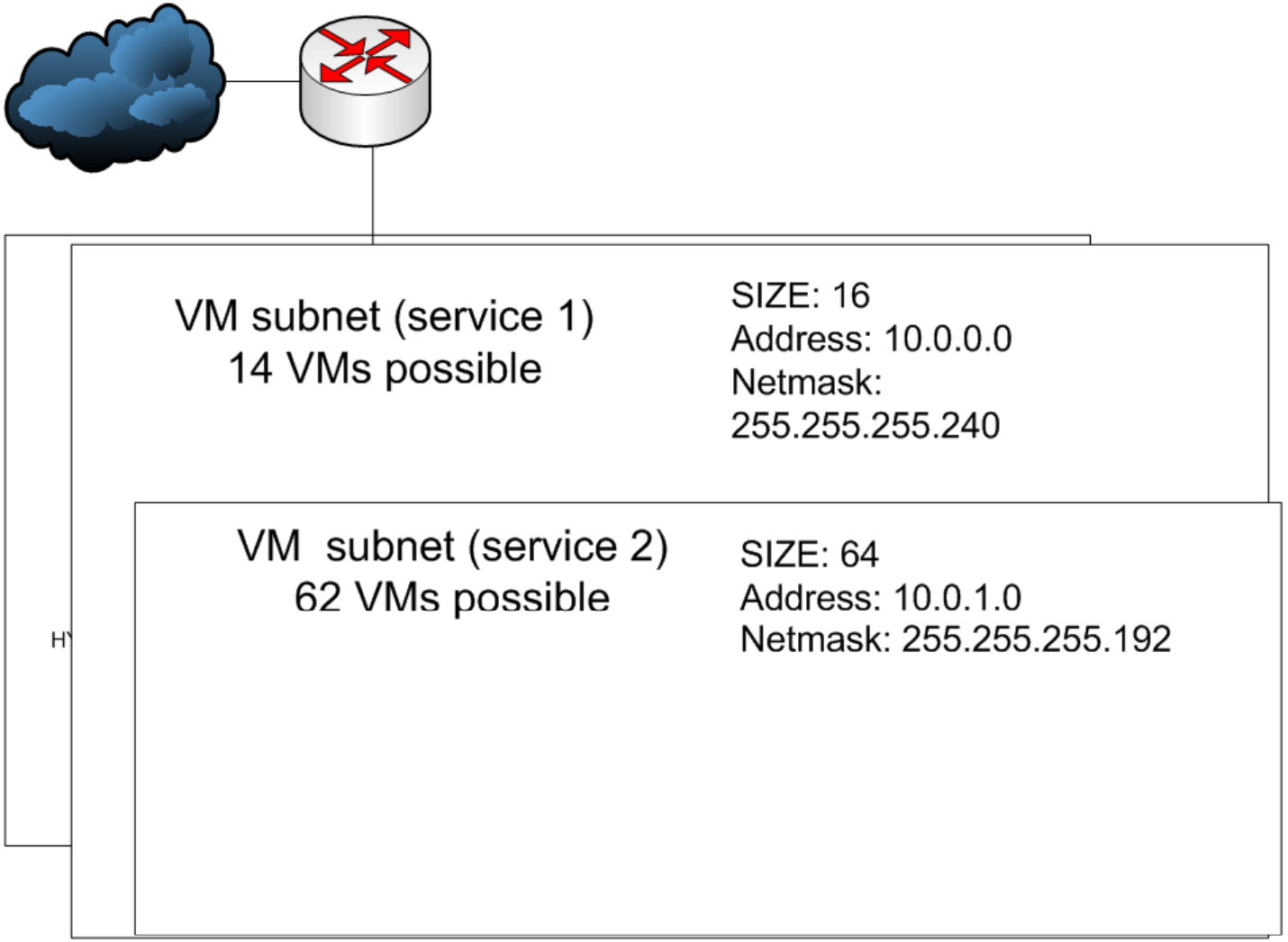
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

VM Clusters



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

VM Clusters



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

VM Cluster advantages

CERN IT Department

- Allows us to move the full virtualized infrastructure (without changing IP addresses for the VMs) in case of need.
- Delegate to the VM Cluster owner full allocation of network resources.
- All combinations possible:
  - Hypervisor in public address/private (preferred)
  - VM subnet1 public/private
  - VM subnet2 public/private
- Migration within the same VM subnet to any host in the same VM cluster possible.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

VM Clusters

- How this service is offered to service providers: SOAP
- Is flexible: can represent the actual VM or a VM Slot.
- VM Cluster is requested directly to us
  - Adding a VM subnet also has to be requested.
- What can be done programmatically?

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

VM representation in LANDB

CERN IT Department

- Several use cases for VMs: we need flexibility
- They are still machines, responsible may differ from hypervisor. Should be registered as such:
  - Added a flag that indicates this is a Virtual Machine.
  - Pointer to the HOST machine using it at this moment.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

CERN

Operations allowed for service providers in LANDB

- Allows to document the VM infrastructure in LANDB:
  - Create a VM (creates device, IP allocation in the cluster)
  - Destroy a VM
  - Migrate a VM (inside the same VM subnet)
  - Move a VM (inside the same cluster or other cluster -> VM will change IP)
  - Query information on Clusters, hypervisors, and VMs
    - What hypervisor is my VM-IP on?
    - What VM-IPs are running in this hypervisor?

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Conclusions

- Is not obvious how to manage virtualization on large networks.
- We are already exploring possible solutions
- When the requirements are defined we are confident to find the appropriate networking solutions.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Questions?

CERN IT Department

THANK YOU!

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it