Networking in Public Clouds

A Parallel Universe with a Different Geometry

Ivan Pepelnjak (ip@ipSpace.net)
Network Architect

ipSpace.net AG

Who is Ivan Pepelnjak (@ioshints)

Past

- Kernel programmer, network OS and web developer
- Sysadmin, database admin, network engineer, CCIE
- Trainer, course developer, curriculum architect
- Team lead, CTO, business owner

Present

- Network architect, consultant, blogger, webinar and book author

Focus

- SDN and network automation
- Large-scale data centers, clouds and network virtualization
- Scalable application design
- Core IP routing/MPLS, IPv6, VPN

CISCO CCIE EMERITUS

vmware vEXPERT

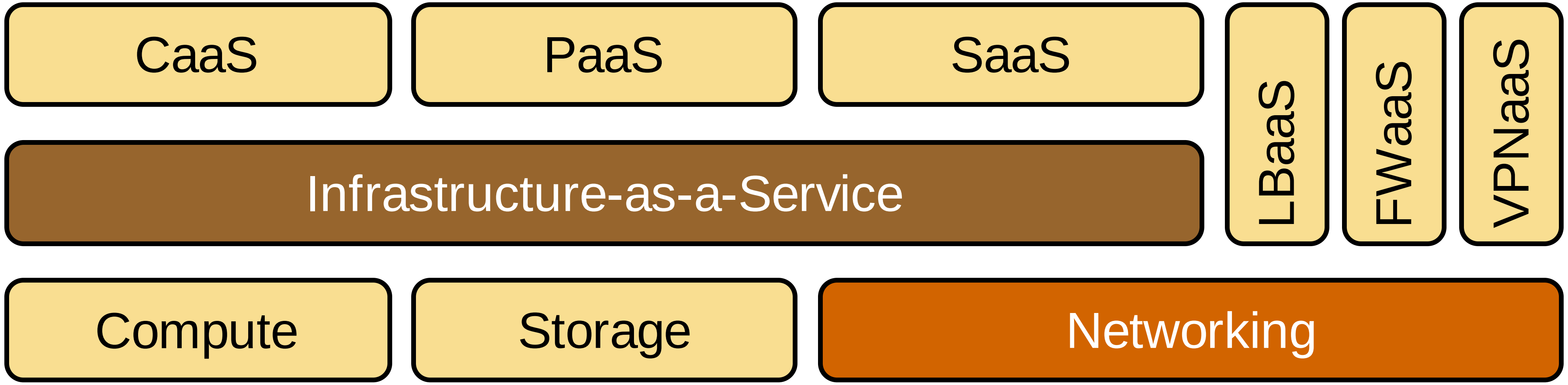
The only marketing slide in this talk

Clouds Need No Networking

Back to Reality

What abstraction are you working with?

- **Software-as-a-service**: it's just a web site, use Internet access or direct connection (warning: BGP ahead)
- **Platform-as-a-service** (aka serverless): most plumbing implemented by the cloud provider
- **Infrastructure-as-a-service**: where do you think you'll connect your VMs to? And how will you connect them to the outside world?
- **Containers-as-a-service**: welcome to (somewhat abstracted) NAT madness



CaaS

PaaS

SaaS

LBaaS

FWaaS

VPNaaS

Infrastructure-as-a-Service

Compute

Storage

Networking

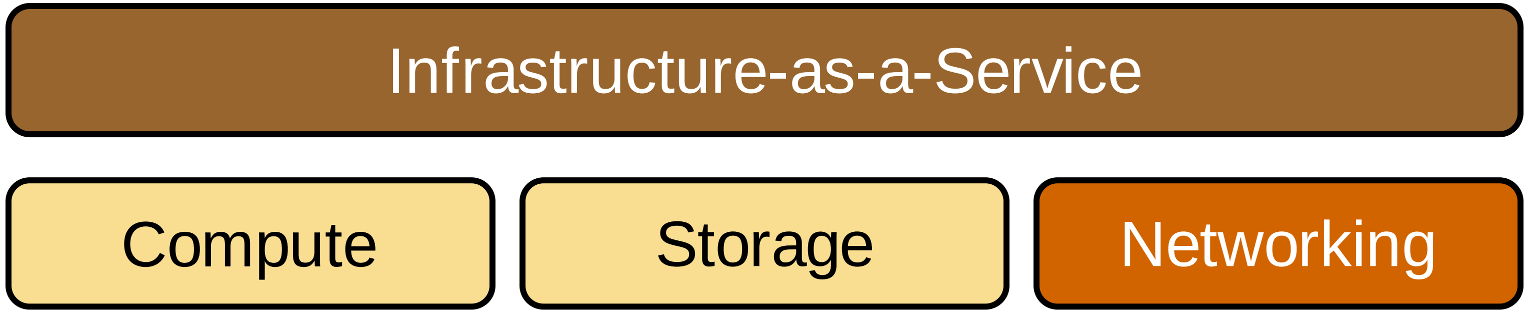
Agenda

Public cloud networking is…

- Different
- Nothing special
- Crazily complex

The Bare Minimum

- Create a tenant network
- Create one or more subnets in the tenant network
- Create VM NIC, assign an IP to the NIC
- Start a VM, attach VM NIC to a subnet

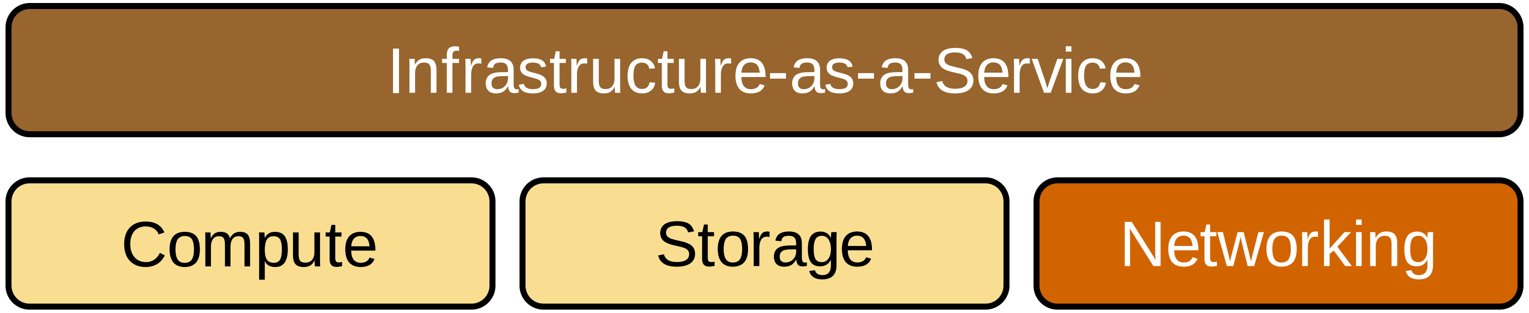


One Does Not Simply

Expose Servers to the Internet

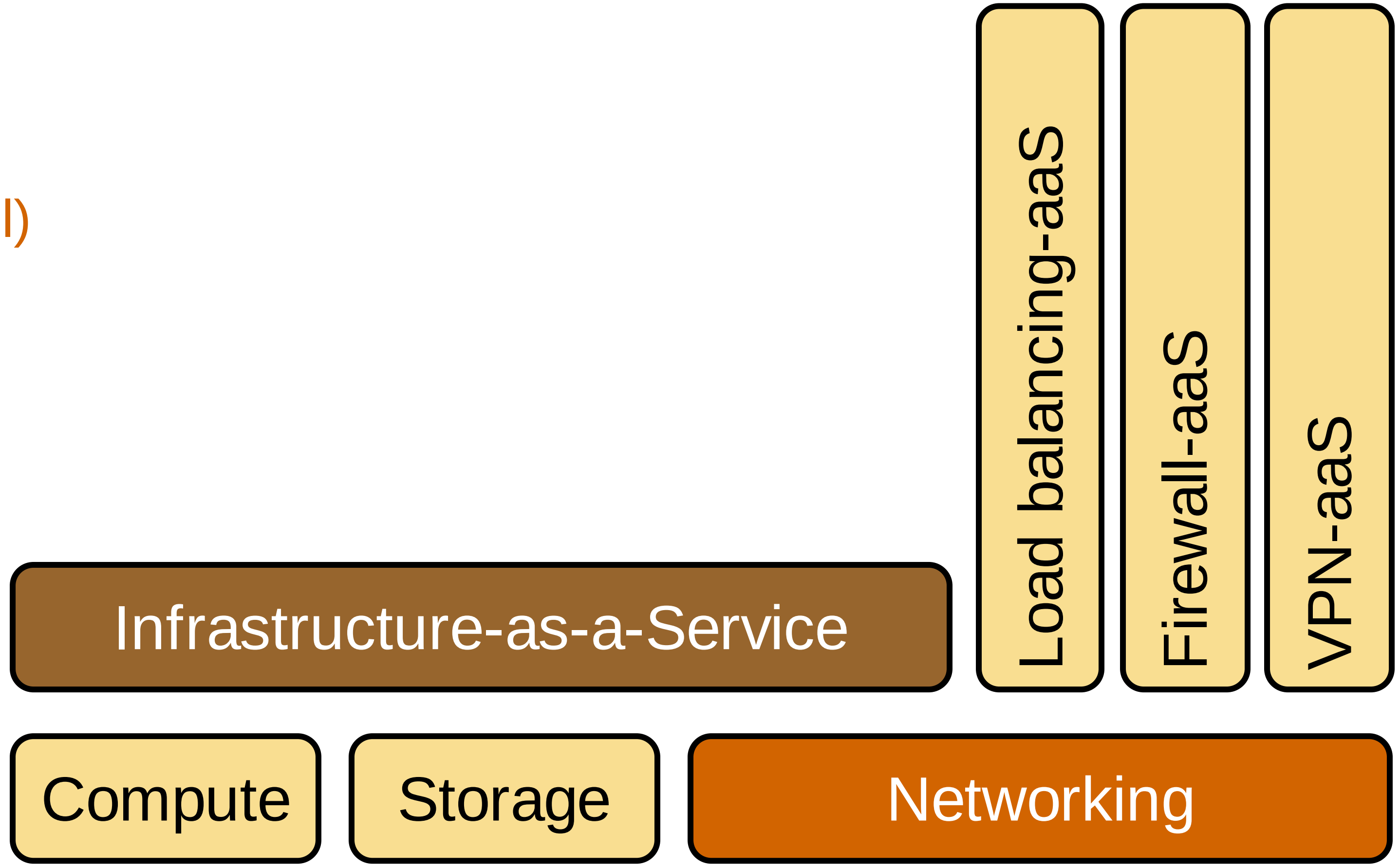
IaaS: Add Security

- Create a tenant network
- Create one or more subnets in the tenant network
- Create VM NIC, assign an IP to the NIC
- Start a VM, attach VM NIC to a subnet
- **Protect VMs (security groups or firewalls)**



IaaS: Welcome to Real World

- Create a tenant network
- Create one or more subnets in the tenant network
- Create VM NIC, assign an IP to the NIC
- Start a VM, attach VM NIC to a subnet
- Protect VMs (security groups or firewalls)
- Add load balancing (network- or application level)
- Availability zones and regions
- Protected links (VPN – IPsec, BGP)
- Direct connection to the cloud (BGP)
- Inter-cloud connectivity (have fun)



Who Do You Think Will Make That Work?

Networking in Public Clouds Is Different

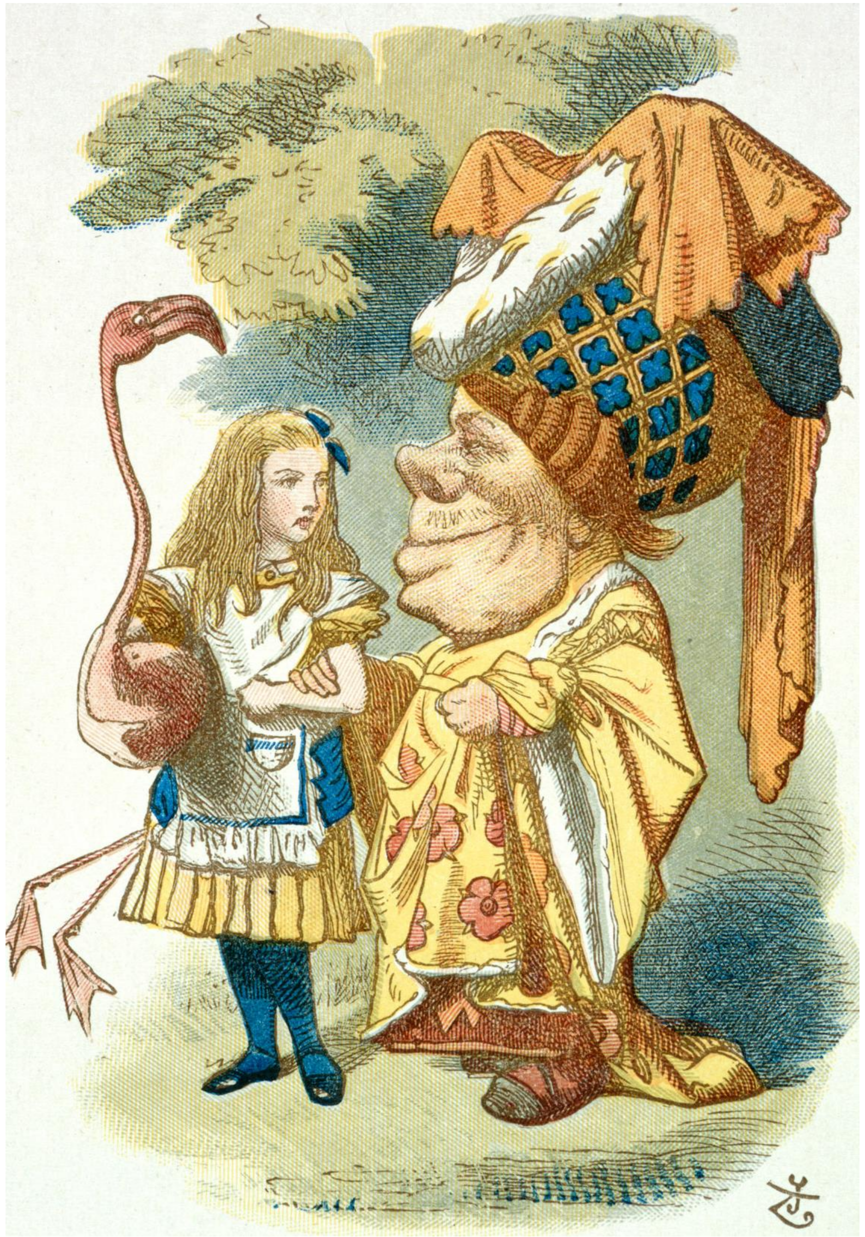
What a Weird Land We’re Entering

There’s no layer-2 in (sane) public cloud

- VMware-based approximations don’t count
- We’re talking about stuff that scales beyond 1000 hosts

How am I supposed to:

- Move virtual machines
- Implement high-availability clusters
- Deploy firewall clusters
- Migrate workloads from on-premises data center



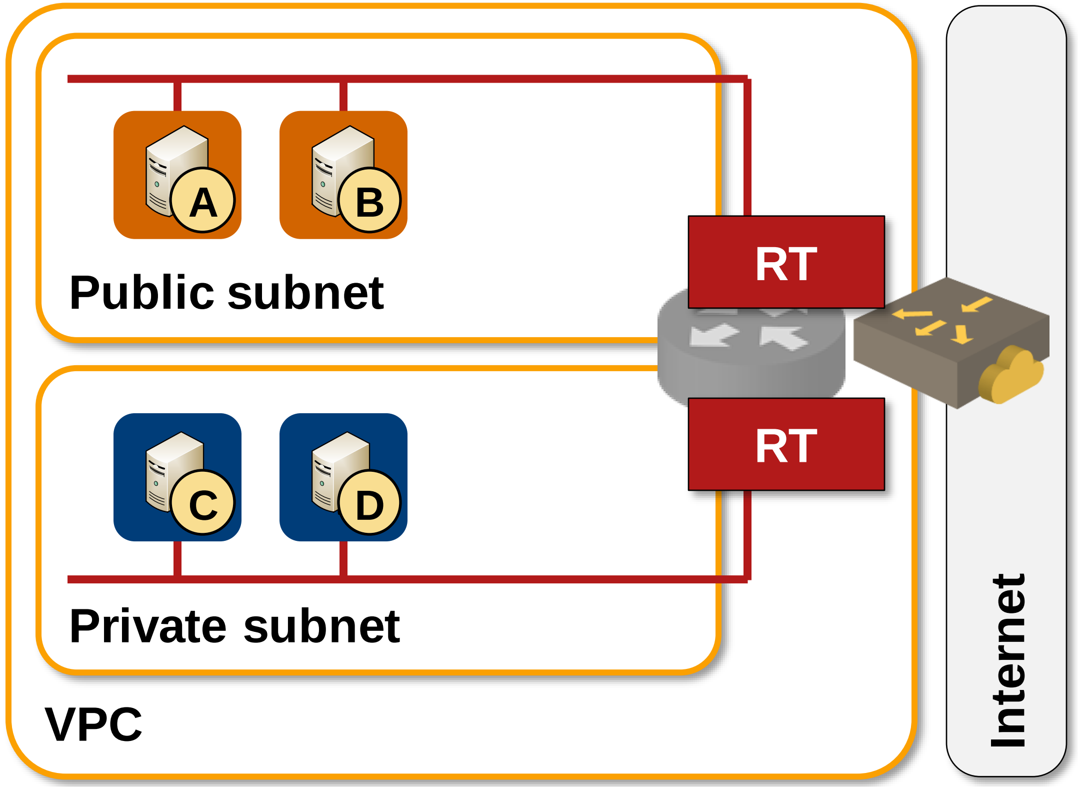
Each cloud provider does things a bit different

AWS versus Azure: Common Concepts

- Isolated tenant routing domains (VPC, VNet...)
- Multiple subnets within a tenant routing domain
- IP and MAC addresses assigned by the orchestration system
- Strict IP+MAC RPF checks (can be disabled)
- Routing controlled by the orchestration system

Consequences

- You cannot change VM IP or MAC address without an orchestration system API call
- You cannot use FHRP
- Most MAC+IP high availability hacks don’t work
- You cannot run a routing protocol within the cloud to influence forwarding decisions

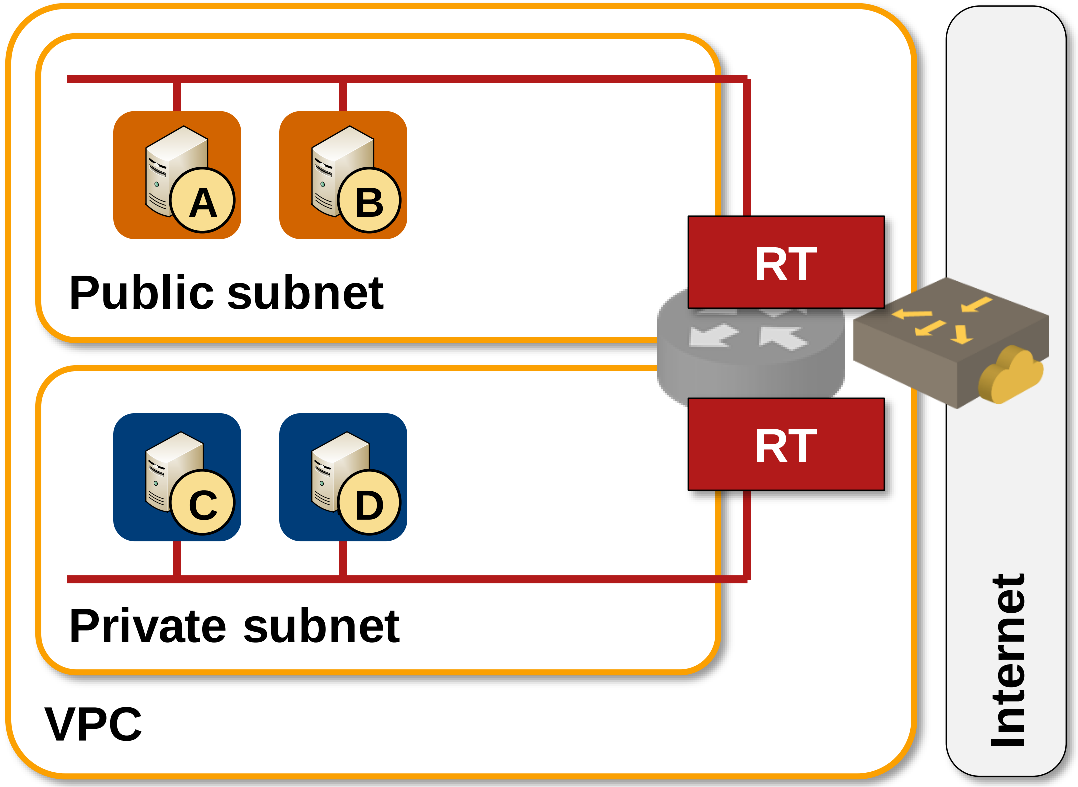


AWS Networking 101

- Each subnet is limited to a single availability zone
- Unicast IPv4 + IPv6 forwarding
- Limited IPv4 multicast support
- Unicast MAC forwarding within the subnet
- No L2 flooding
- Each subnet can have a different route table
- There’s no way to influence intra-VPC packet forwarding

Consequences

- Service insertion is interesting
- Intra-VPC service insertion is **really** hard
  (and usually involves a lot of NAT duct tape)
- You could do routing tricks within a subnet but
  not across subnets

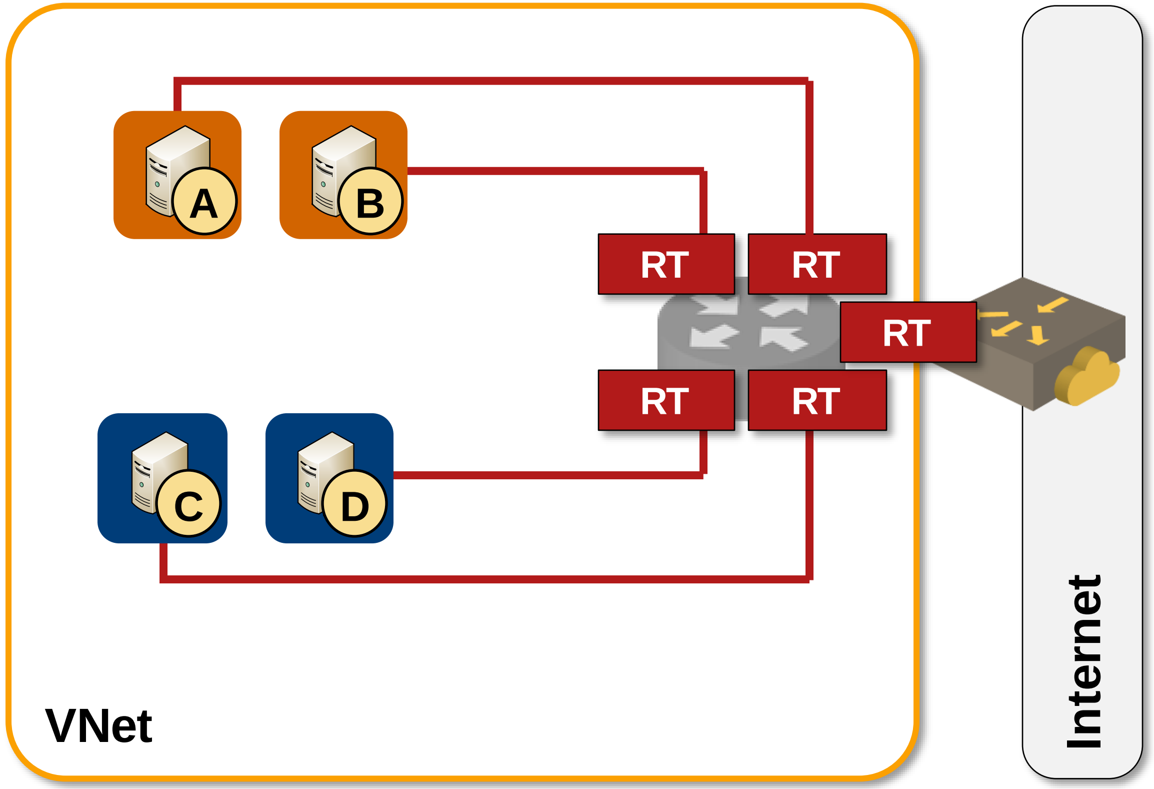


Azure Networking 101

- Subnets span availability zones
- Unicast IPv4 + IPv6 forwarding
- No L2 forwarding (every instance connected directly to a router)
- ARP always returns the MAC address of Azure router
- Each subnet could have a different route table
- Route tables can contain intra-VNet prefixes

Consequences

- Service insertion is relatively easy (but messy)
- Building application swimlanes tied to availability zones is hard(er)

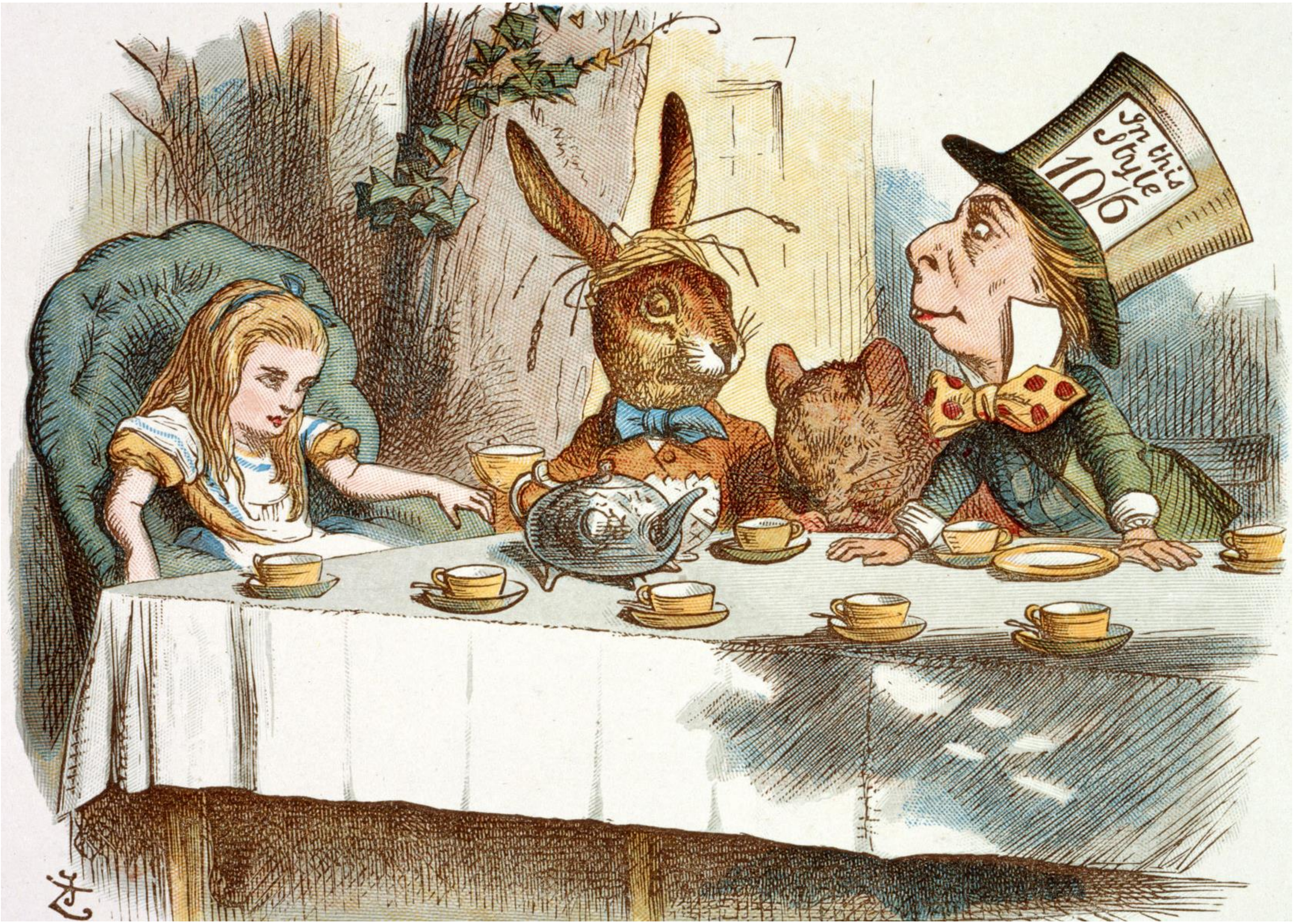


Why Couldn’t They Be The Same?

- Convergent evolution
- Different audiences
- Different scalability goals?
- Fixing different problems
- Solving the same problem in different ways (aka leverage the investment)

Consequences

- Nobody wants to be limited to the least common denominator
- Real-life tools have cloud-specific plugins or modules (Terraform, Ansible)
- Multi-cloud works best in PowerPoint



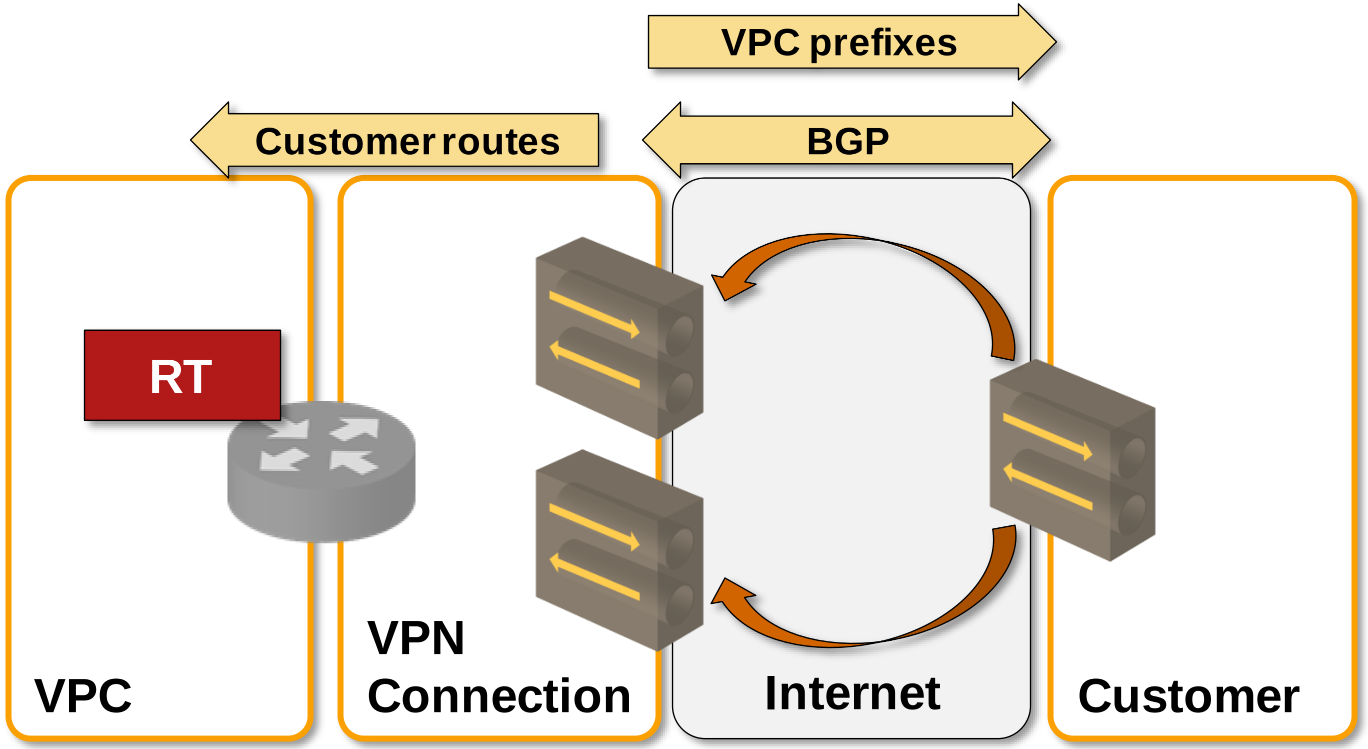
Networking in Public Cloud Is Nothing Special

Example: VPN Connectivity (AWS and Azure)

- IPsec tunnel
- Unnumbered interfaces
- EBGP multihop
- Static host route to BGP next hop

Similar setup for direct connection

- VLAN trunk
- EBGP
- AS path prepending to influence route selection



Networking in Public Cloud Can Get Complex

Example: AWS Gateway Load Balancer

- Create a dedicated subnet for GWLB endpoint in each availability zone

Ingress routing

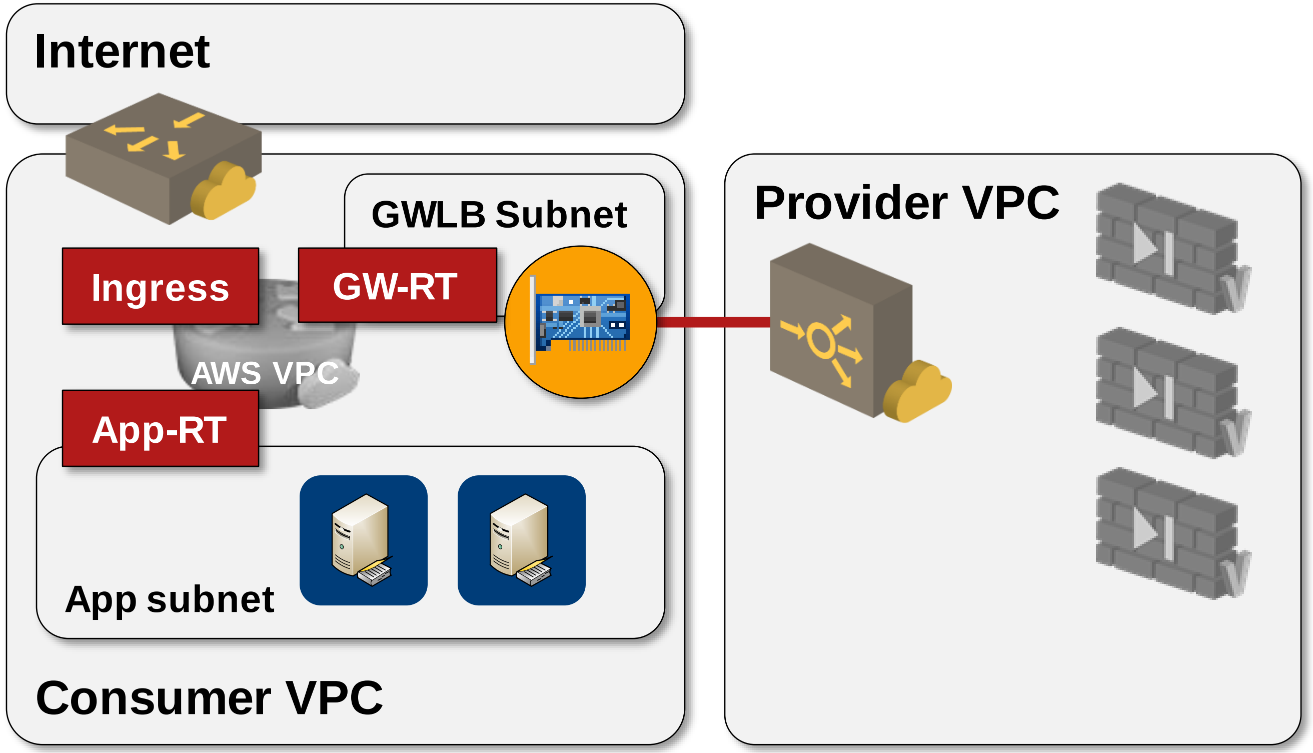
- Use Internet Gateway ingress routing
- GWLBE is next hop for consumer VPC CIDR block or specific subnets

Egress routing (App-RT)

- Use custom route table for App subnet
- GWLBE is next hop for default route

GWLB egress routing (GW-RT)

- Use default route table
- Internet Gateway is next hop for default route



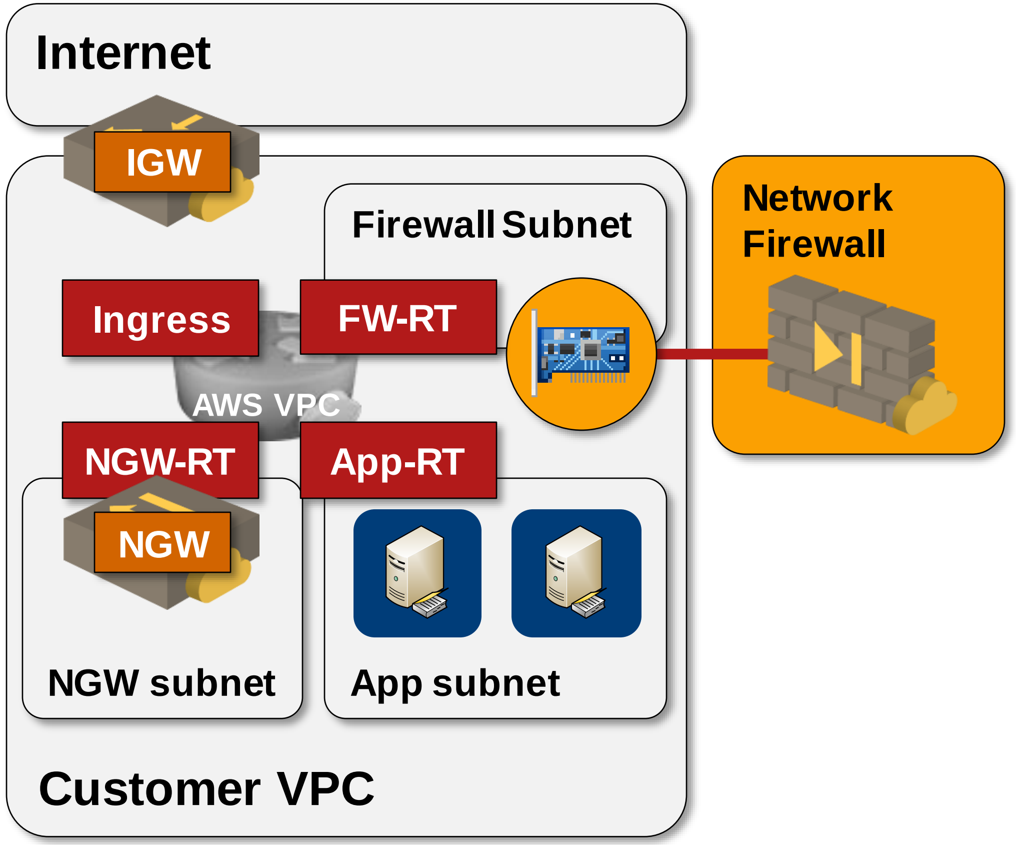
Example: Combining AWS NAT Gateway with Network Firewall

Egress routing

- App route table: Default route ➔ NAT gateway instance
- NAT gateway route table: Default route ➔ Firewall endpoint
- Firewall subnet route table: Default route ➔ Internet gateway

Ingress routing

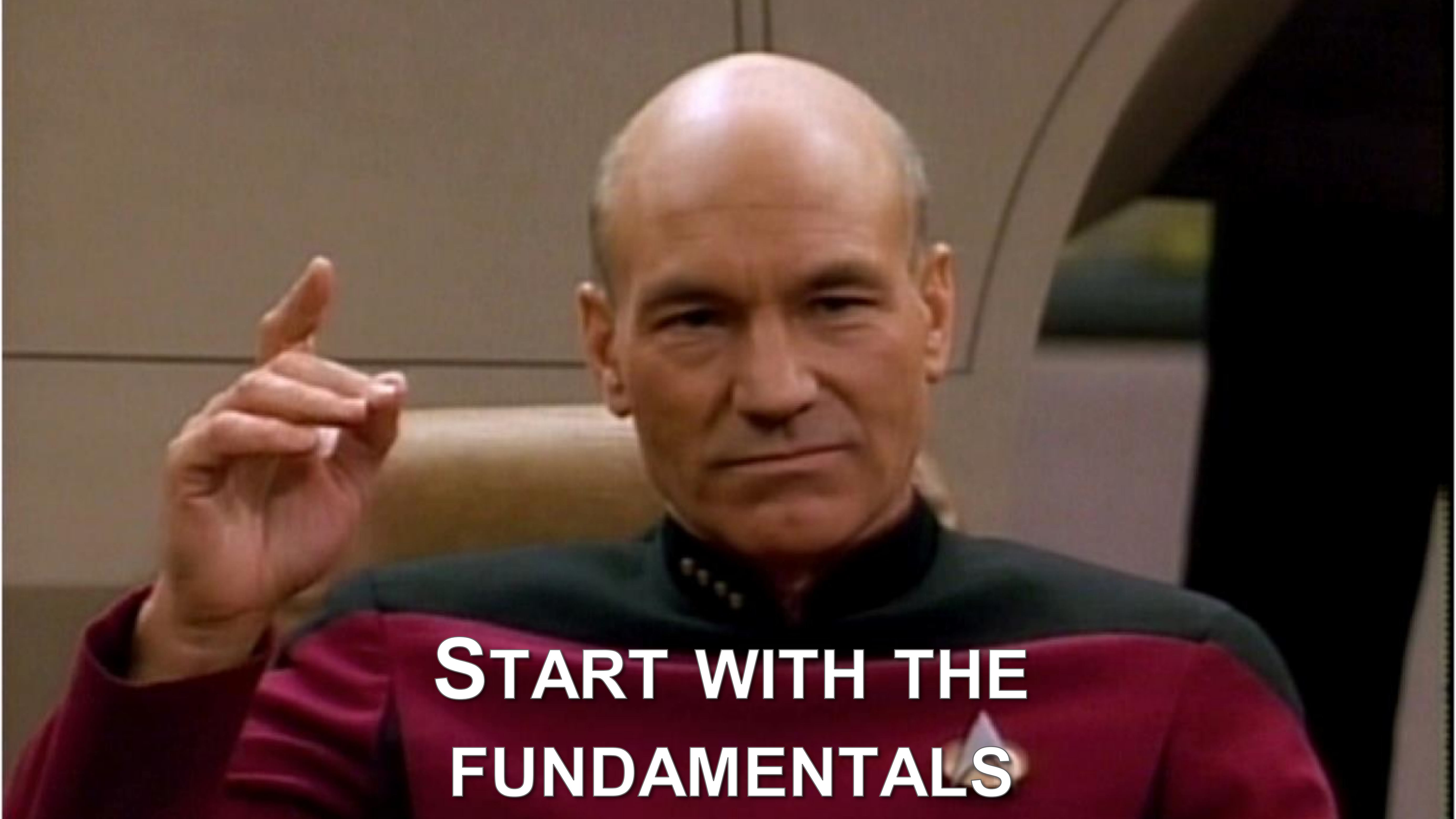
- Ingress route table: VPC CIDR prefix ➔ Firewall endpoint
- Traffic is sent from firewall endpoint to NAT gateway based on destination IP address
- NAT gateway sends translated traffic to VM instances



What Can You Do?

Forget PowerPoint

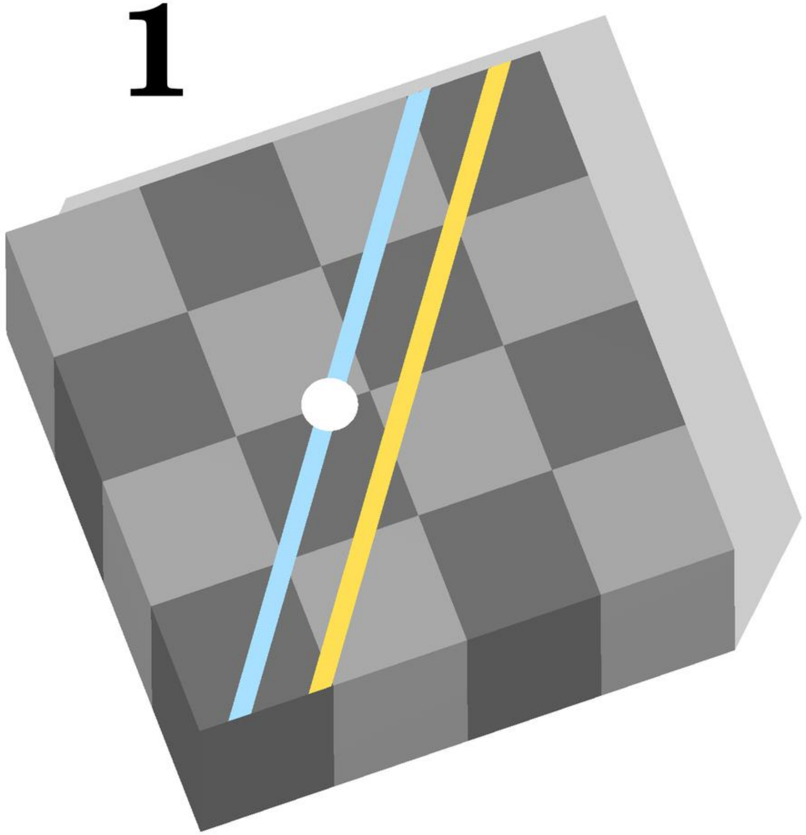
It's Time for the Red Pill



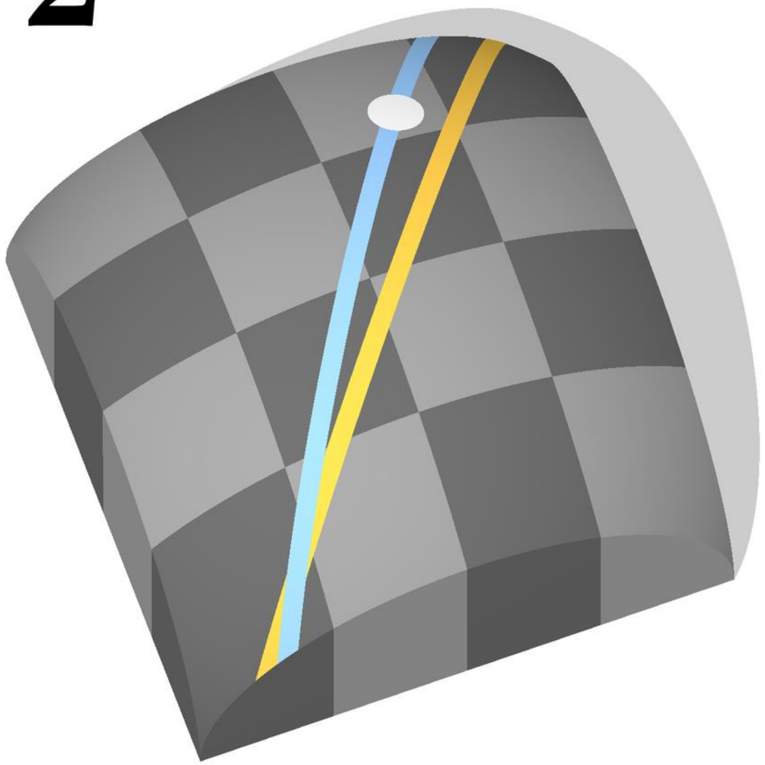
START WITH THE FUNDAMENTALS

It’s Just Another Case of Alternate Geometries

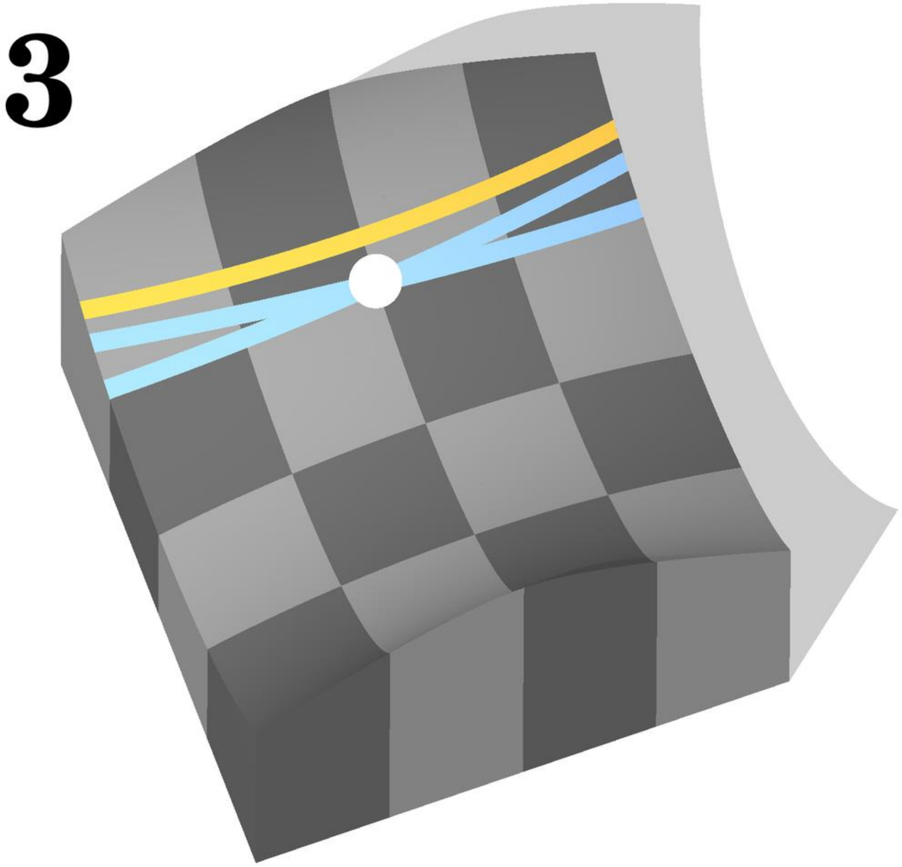
1



2



3



Master the cloud…

…You will

Questions?

Web:	ipSpace.net
Blog:	blog.ipSpace.net
Email:	ip@ipSpace.net
Twitter:	@ioshints

Public Cloud:	ipSpace.net/PubCloud
Automation:	ipSpace.net/NetAutSol
Webinars:	ipSpace.net/Webinars
Consulting:	ipSpace.net/Consulting