Infrastructure as a Service

Infrastructure as a Service, or IaaS is a service offering by most cloud providers that provides *virtual mac*

Virtual machines

Virtualization is a surprisingly old technology. The first virtualized system was the IBM System/370 mainfr

When you think about mainframes you have to consider that these machines were *very* expensive and m

Initially CPUs in personal computers did not have application separation. The x86 line of Intel CPUs only
DOS.

Protected mode introduced the concept of *rings* in the CPU. The operating system *kernel* would run in rin

> **Note**
>
> Device drivers today typically run on ring 0 instead of 1 or 2.

This ring system allowed the operating system to restrict the higher ring numbers from accessing certain

> **Note**
>
> If you try and set up a really old computer game like Commander Keen in DOSBox you will realize that you have to provide the g

To work around the problems with protected mode the 80386 successor introduced virtual mode. The nev

For instance the CPU would create a simulated *virtual* memory space the program could write to and tran

> **Note**
>
> VM86 does not capture *every* instruction the application runs in virtual mode, only the sensitive CPU instructions. This enables le

In the mid 2000's CPUs became so powerful that it made sense to not only virtualize applications but wh
techniques, such as Xen required the guest operating system to run a modified kernel to facilitate them ru

Hardware vendors, of course, followed suit. In 2005 Intel added the VT-x (Vanderpool) feature to its new

VT-x and AMD-V added new ring -1 to accommodate *hypervisors*. This new ring allowed for separation b
machines.

> **Note**
>
> Intel also introduced a ring -2 for the Intel Management Engine, a chip that functions as an OOBM in modern Intel chips. The ME

Virtualization also gave rise to Infrastructure as a Service. AWS was the first service that offered virtual m

This allowed customers to create virtual machines as they needed it and they were billed for it on an hou

The presence of an API makes the difference between IaaS and plain old virtual machines as a service. I

Typical instance types

When the cloud became popular in the late 2000s several providers attempted to offer a service that was

Instead most cloud providers nowadays opt to offer fixed machine sizes. To accommodate high-CPU and

- **Shared CPU:** These are small instances where a single CPU core is shared between multiple virtual
- **Standard, dedicated core CPU:** These instance types receive one or more physical cores leading t
- **High CPU:** These instance types are usually hosted on physical servers that have a very high CPU t
- **High RAM:** This offering is the exact opposite of the high CPU offering. The machines on offer here i
- **Storage:** These instance types contain large amounts of local storage (see below in the storage sect
- **Hardware-specific:** These instance types offer access to dedicated hardware such as graphics card

Automation

As discussed before, that makes an IaaS cloud provider a cloud provider is the fact that they offer an API

Initially this problem would be solved by creating *templates* for the operating system that launches. In larg

Thankfully in the last decade a lot has happened and Cloud Init has established itself as a defacto standa

A DevOps engineer can simply inject a script that runs at the first start that takes care of all the installatio

Tools like Terraform or Ansible assist with managing the whole process of provisioning the virtual machin

Virtual machine pools

One other use of user data are virtual machine pools. Each cloud provider adopts a different name for the
health check the cloud deletes the machine and creates a new one.

The number of machines in a pool can, of course, be changed either manually or in some cases automat

Combined with the aforementioned user data this can be a very powerful tool to create a dynamically size

These pools are often integrated with the various load-balancer offerings cloud providers have in their po

Storage

When it comes to data storage virtual machines work exactly like your physical machine would: there is a machine will not cause a data loss.

However, a distributed storage system is generally either slower or more expensive for the same perform

When we talk about storage systems we are talking about two types: block devices and filesystems. On t blocks of a single file may be distributed all over the whole disk randomly so that's something the filesyste

Therefore we traditionally call raw disk devices *block devices*. Block devices are (with very few exception enforce a single-VM access policy. In other words, one block storage device can only ever be used by a s

Local Storage

As described above the simplest and most widely supported option to store data from your virtual machin

Some cloud providers offer disk redundancy (RAID) while others don't. At any rate a hardware failure on t

It is therefore very advisable to solve redundancy on top of the virtual machine, e.g. by building a replicat

Network Block Storage

Network block storage means a block storage that is delivered over the network. The network here can m

As described before block storage is, in general, single-VM only. You can't access the files stored on a bl

Also note that Network Block Storage does not automatically come with redundancy. Some solutions, suc

At any rate, using Network Block Storage does not absolve you from the duty to make backups and have

Network File Systems

In contrast to network block storage network file systems offer access to data not on a block level, but on

The filesystem has to keep track of which machine has which file open, or has locks on which file. When

Object storage

Object storage systems are similar to network file systems in that they deal with files rather than blocks. H

While object storages technically *can* be used as a filesystem on an operating system level for example b

Operating system level integration should only be used as a last resort and object storages should be inte

Network

The next big topic concerning IaaS services is networks. Before we go into the cloud-aspect let's look at h

How cloud networks are built

So, let's get started. Imagine a data center from the first lecture. Your task is to build an IaaS cloud provic
s.

This sounds like a lot of bandwidth available but keep in mind that your virtual machines get assigned to t
between two distinct virtual machines.

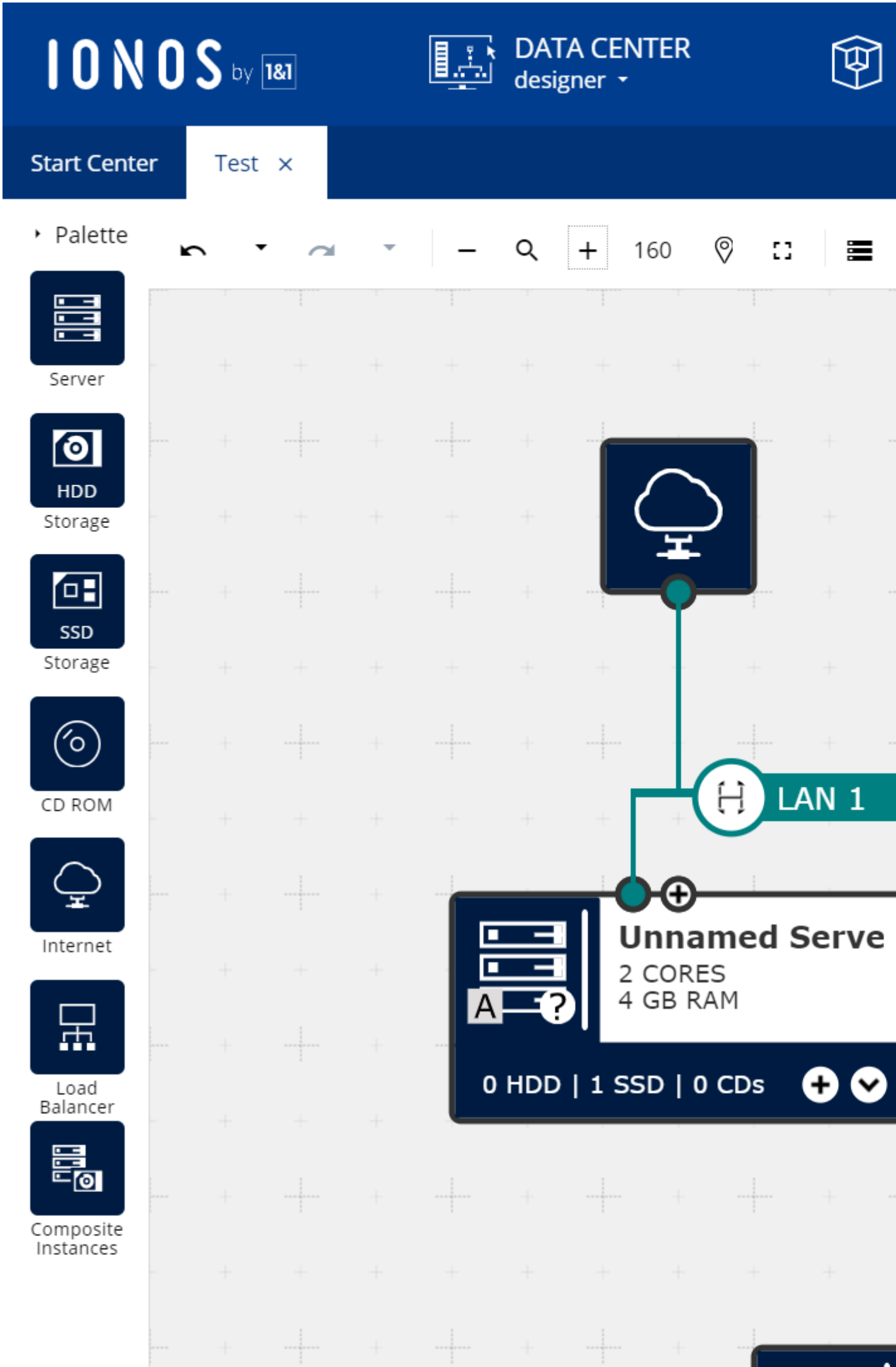
This is part of the reason why in the cloud scaling horizontally (adding more machines) is preferred rather

Network architectures offered by cloud providers

When we look at the network offerings by cloud providers there are three types:

1. **Private-only network with NAT**: This option is provided by the larger cloud providers such as AWS,
   Destination NAT. Multiple private networks (VPC's) can be assigned to a virtual machine and they ca
2. **Default public IP**: This option is provided by smaller IaaS providers such as DigitalOcean, Exoscale

3. **Fully customizable:** This setup allows the customer to design their network connectivity as they see


IONOS by 1&1
DATA CENTER
designer
Start Center
Test
Palette
160
Server
HDD
Storage
SSD
Storage
CD ROM
Internet
Load
Balancer
Composite
Instances
LAN 1
Unnamed Serve
2 CORES
4 GB RAM
0 HDD | 1 SSD | 0 CDs

> **Note**
>
> There are several other cloud providers which we have no information on, such as the Deutsche/Open Telekom Cloud, or the Ali

> **Note**
>
> Out of group 2 it is worth mentioning that the services that are available on the public network (firewalls, load balancers) are ofte

TODO: add illustration

Firewalling

IaaS providers often also offer network firewalls as a service, included in the platform. Firewalls generally

Firewall providers often employ the concept of *security groups*. The implementation varies greatly, but in

For most cloud providers you will need to create an explicit rule allowing traffic to flow between two mach

The advantage of security groups is that the rules can be made in such a way that they reference other s
application servers.

TODO: add screenshot of security group configuration

Network load balancers

Network load balancers are an option some cloud providers offer. In contrast to Application Load Balance

TODO add illustration

Depending on the cloud provider in question network load balancers may or may not offer terminating en

When designing an architecture it is worth considering if the real IP address of the connecting client will b
should, in general, make the client IP available to the backends.

When talking about load balancers an interesting question is the load balancing strategy. Most load balan

VPNs, private interconnects, and routing services

While it seems convenient at first to use only the public network several organizations have security mod

However, this presents a problem when moving data between several, geographically distributed location
an MPLS tunnel or VPN.

MPLS tunnels create a virtual connectivity that does not go on the Internet. While being expensive and sl

VPN's on the other hand create a virtual connectivity by sending data over the Internet in an encrypted fo

Larger cloud providers tend to offer both options. MPLS is supported by the larger cloud providers (AWS

VPN is also offered mostly by large providers (AWS VPC VPN, Azure VPN, or Google Cloud VPN). Howe
Azure's Point-to-Site VPN.

It is also worth noting that VPN's can be used to connect cloud providers together.

DNS

The Domain Name Service is one of the services that are all but required for building an infrastructure. It

There is a difference, however, between DNS services on offer. Some DNS services by cloud providers o

More advanced features may include automatic DNS failover. This involves running regular health checks
building a custom CDN.

Monitoring

Some cloud providers offer included basic monitoring, such as CPU or memory usage. Some providers a
next lecture.

Automation