



Universidade Presbiteriana
Mackenzie

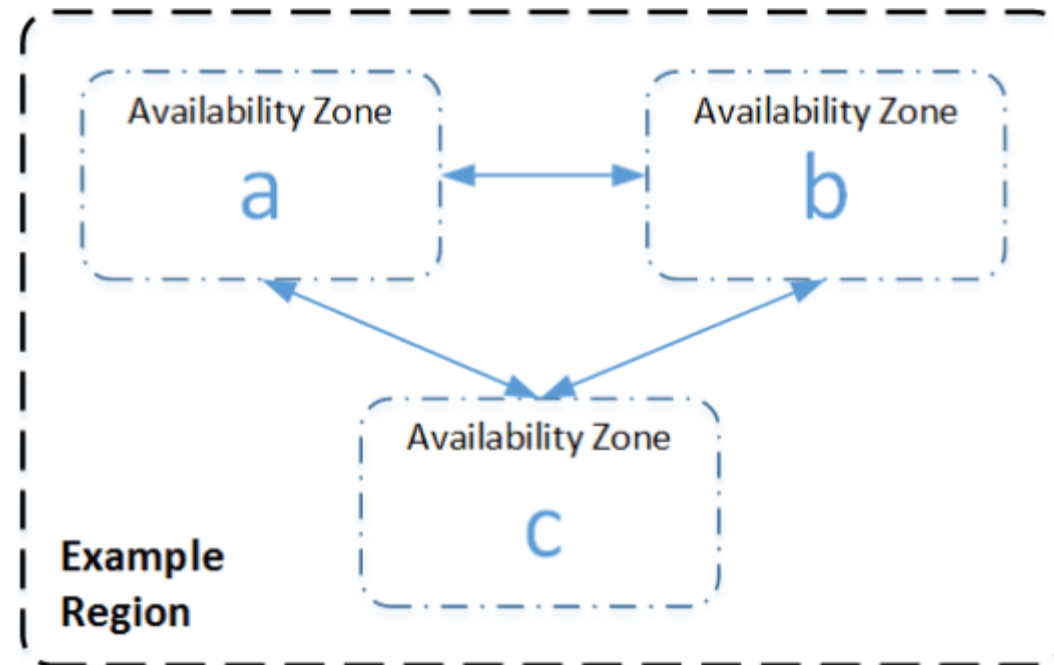
Visão geral da infraestrutura global de provedores de nuvem

Princípios da nuvem

Conceitos

- **Regiões**
- **Zonas de Disponibilidade (AZ)**
- **Benefícios da Virtualização**
- **High Availability (HA) e Disaster Recovery (DR)**
- **Gerenciamento de Ciclo de Infra Virtualizada**

Regiões e Zonas de Disponibilidade (AZ)



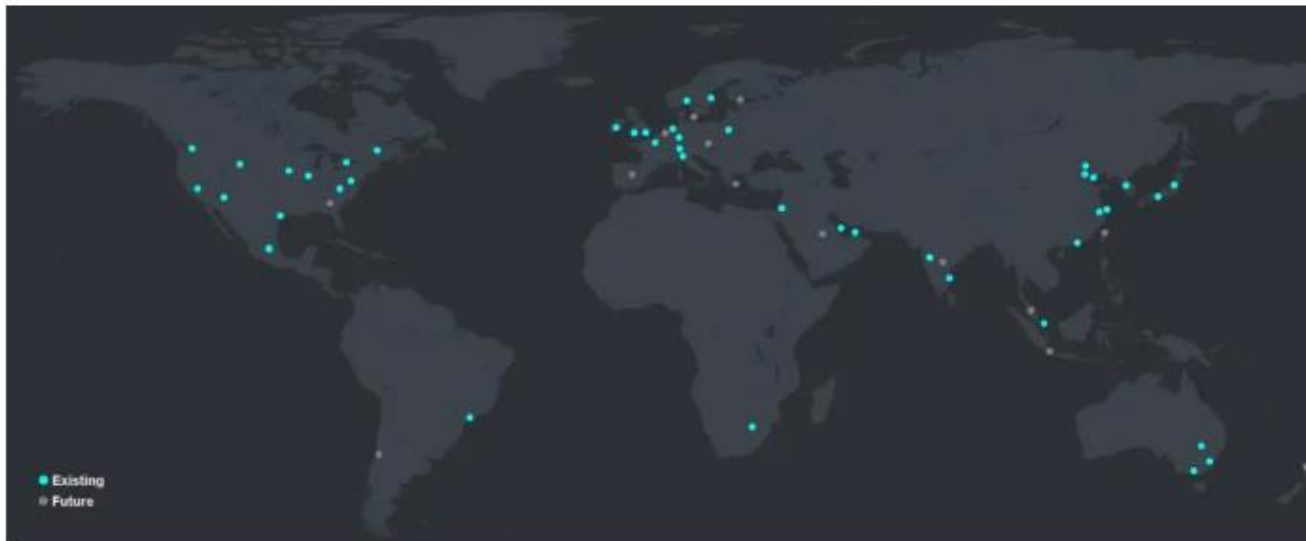
Alguns provedores podem ainda implementar data centers de **Locais de borda** e **caches de borda regionais** melhoram o desempenho **armazenando conteúdo em cache** em locais mais próximos dos usuários

Regiões

- Uma **Região** é uma área geográfica.
 - A replicação de dados entre Regiões é controlada pelo usuário
 - A comunicação entre regiões usa a infraestrutura de rede backbone
- Cada região fornece redundância total e conectividade com a rede.
- Uma Região normalmente consiste em duas ou mais **Zonas de Disponibilidade**.

Azure

In total, Microsoft Azure operates over 300 physical data centers in various global locations, with each distinct building home to a group of networked computer servers. These data centers are arranged into 64 regions – *Microsoft Azure's term for a set of data centers* – that are deployed within a latency-defined perimeter and linked by over 175,000 miles of terrestrial and subsea fiber-optic networks.



Amazon (AWS) Data Center Location in Eastern Oregon.

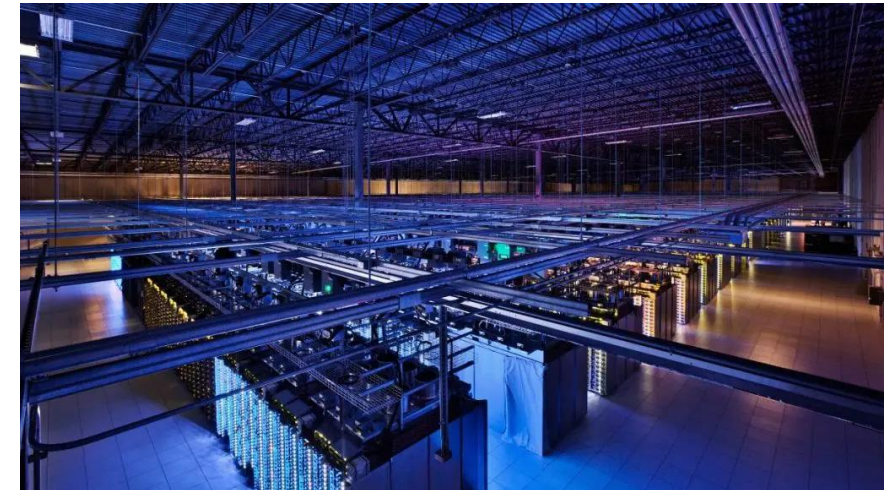
Americas – Regions and Availability Zones

Regions	Code	Zones	Location	Country	Opened
Brazil South	brazilsouth	3	Campinas	Brazil	2014
Brazil Southeast*	brazilsoutheast	1	Rio de Janeiro	Brazil	2020
Canada Central	canadacentral	3	Toronto	Canada	2016
Canada East	canadaeast	1	Quebec City	Canada	2016
Chile Central	chilecentral	3	Santiago	Chile	<i>Future</i>
Mexico Central	mexicocentral	3	Querétaro	Mexico	2024
Total	—	14	—	—	—

• • •

GPC

In total, Google operates or is developing 37 data centers around the world. These data centers support Google's current and planned 49 cloud regions – a *physical location where it clusters data centers* – that are designed to be highly available, fault-tolerant, and concurrently maintainable.



Google Cloud Data Center Location in Council Bluffs, Iowa.

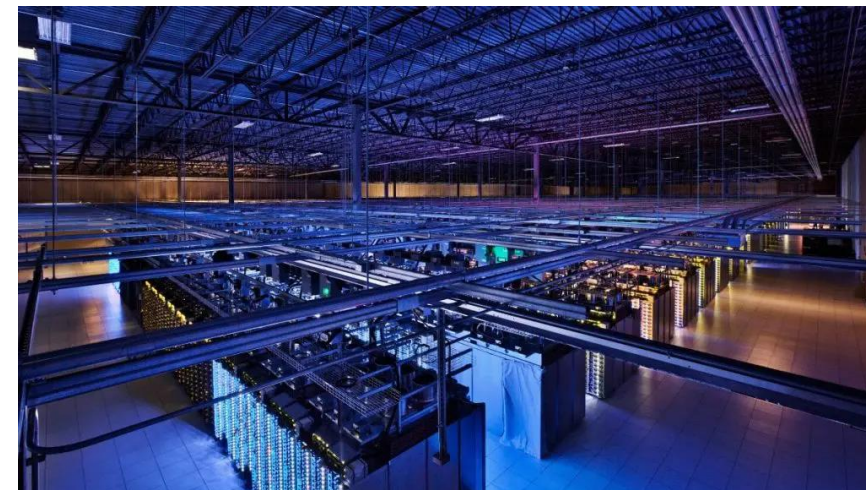
Europe – Regions and Availability Zones

Regions	Code	Zones	Location	Country	Opened
Warsaw	europe-central2	3	Warsaw	Poland	2021
Finland	europe-north1	3	Hamina	Finland	2018
Madrid	europe-southwest1	3	Madrid	Spain	2022
Belgium	europe-west1	3	St. Ghislain	Belgium	2013
London	europe-west2	3	London	United Kingdom	2017
Frankfurt	europe-west3	3	Frankfurt	Germany	2017
Netherlands	europe-west4	3	Eemshaven	Netherlands	2018
Zürich	europe-west6	3	Zürich	Switzerland	2019
Milan	europe-west8	3	Milan	Italy	2022
Paris	europe-west9	3	Paris	France	2022
Berlin	europe-west10	3	Berlin	Germany	2023
Turin	europe-west12	3	Turin	Italy	2023
Austria	TBD*	3	Vienna	Austria	Future

...

AWS

In total, Amazon Web Services (AWS) operates over 125 physical data centers in various global locations, with these facilities comprising over 38 million square feet. These data centers are or will be arranged into 39 regions – *AWS' term for a physical location where it clusters data centers* – that are deployed within a highly available, fault tolerant, and scalable environment.



Google Cloud Data Center Location in Council Bluffs, Iowa.

Asia Pacific – Regions and Availability Zones

Regions	Code	Zones	Location	Country	Opened
Asia Pacific (Auckland)	ap-southeast-5	3	Auckland	New Zealand	2024
Asia Pacific (Bangkok)	ap-southeast-6	3	Bangkok	Thailand	2024
Asia Pacific (Hong Kong)	ap-east-1	3	Hong Kong	SAR	2019
Asia Pacific (Hyderabad)	ap-south-2	3	Hyderabad	India	2022
Asia Pacific (Jakarta)	ap-southeast-3	3	Jakarta	Indonesia	2021
Asia Pacific (Malaysia)	ap-southeast-7	3	Kuala Lumpur	Malaysia	2024
Asia Pacific (Melbourne)	ap-southeast-4	3	Melbourne	Australia	2023
Asia Pacific (Mumbai)	ap-south-1	3	Mumbai	India	2016
Asia Pacific (Osaka)	ap-northeast-3	3	Osaka	Japan	2021
Asia Pacific (Seoul)	ap-northeast-2	4	Seoul	South Korea	2016
Asia Pacific (Singapore)	ap-southeast-1	3	Singapore	Singapore	2010
Asia Pacific (Sydney)	ap-southeast-2	3	Sydney	Australia	2012
Asia Pacific (Tokyo)	ap-northeast-1	4	Tokyo	Japan	2011
China (Beijing)	cn-north-1	3	Beijing	China	2014
China (Ningxia)	cn-northwest-1	3	Changsha	China	2017

• • •

Seleção de uma Região



Governança de dados,
requisitos legais



Proximidade com
os clientes (latência)



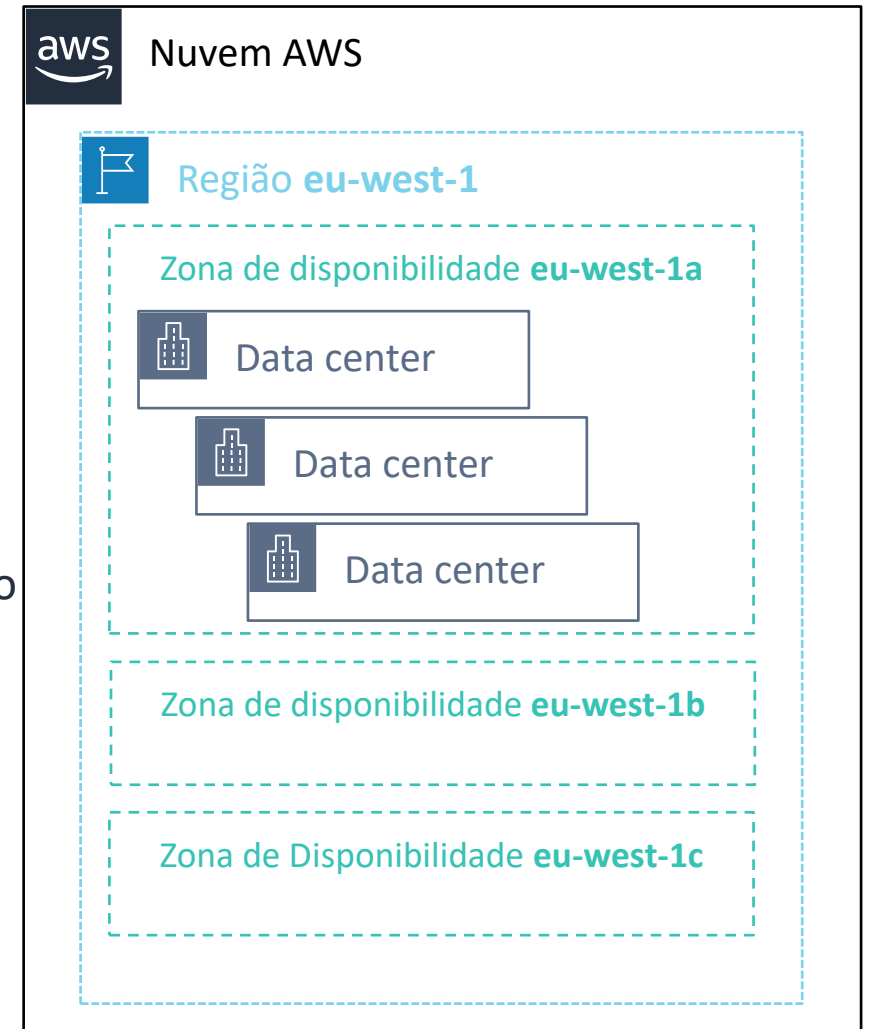
Serviços disponíveis na
região



Custos (variam por Região)

Zonas de Disponibilidade (AZ)

- Cada **Região** tem várias Zonas de Disponibilidade.
- Cada **Zona de Disponibilidade** é uma partição totalmente isolada de infraestrutura
 - As Zonas de Disponibilidade consistem em **data centers** distintos
 - Elas são projetadas para isolamento de falhas
 - Elas são interconectadas a outras zonas de disponibilidade usando redes privadas de alta velocidade
 - Você escolhe suas zonas de disponibilidade
 - **Recomenda-se a replicação de dados e recursos entre Zonas de Disponibilidade** para fins de resiliência



Data centers

- São projetados para segurança.
- Os DC data centers (new) = CPD, centros de processamento de dados (old)
- Cada data center tem energia, redes e conectividade redundantes e está hospedado em uma instalação separada.
- ~ 50.000 a 80.000 servidores físicos.

Recursos de infraestrutura

- **Elasticidade e dimensionamento**

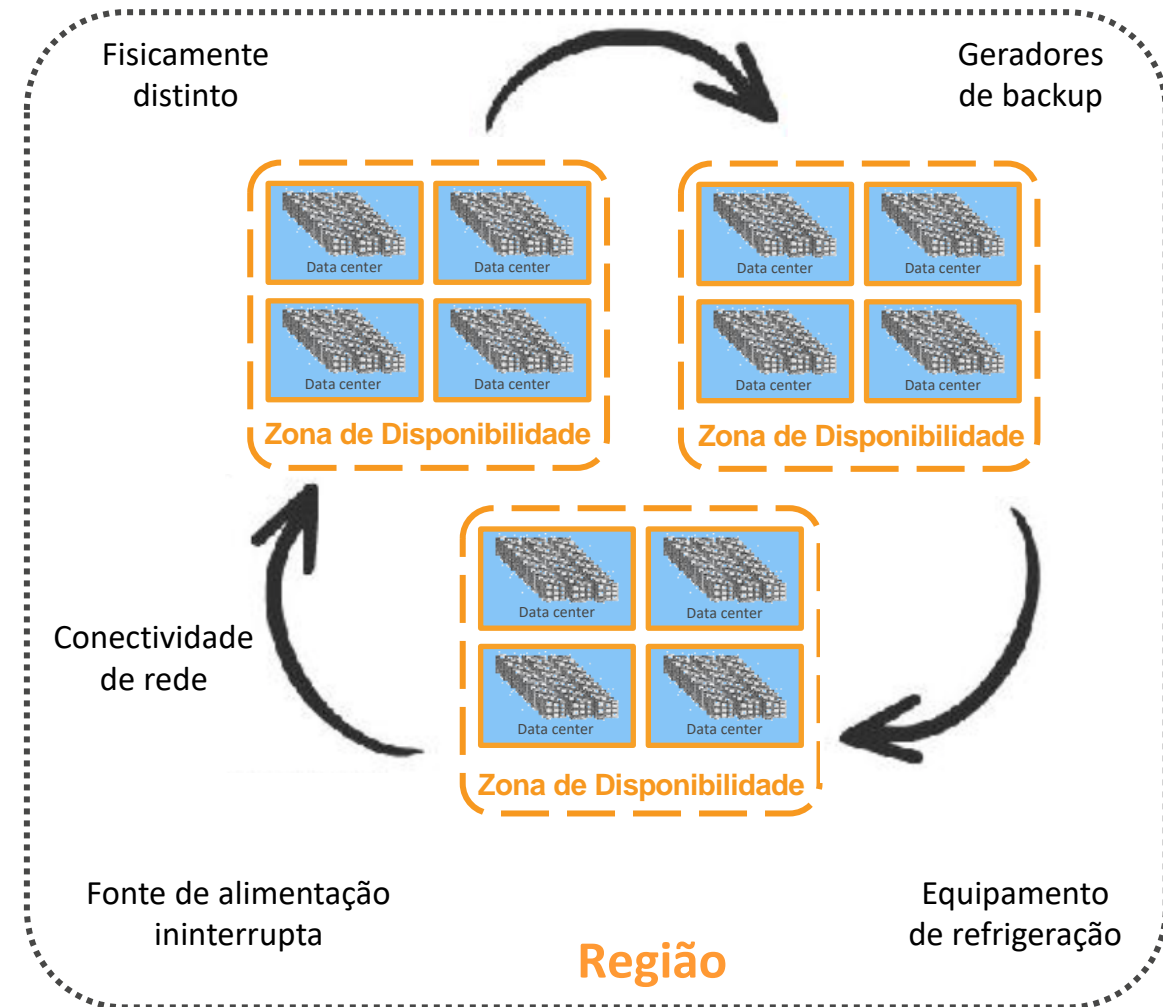
- Infraestrutura elástica; adaptação dinâmica da capacidade
- Infraestrutura escalável; adapta-se para acomodar o crescimento

- **Tolerância a falhas**

- Continua funcionando corretamente na presença de uma falha
- Redundância de componentes

- **Alta disponibilidade**

- Alto nível de desempenho operacional
- Tempo de inatividade mínimo
- Sem intervenção humana



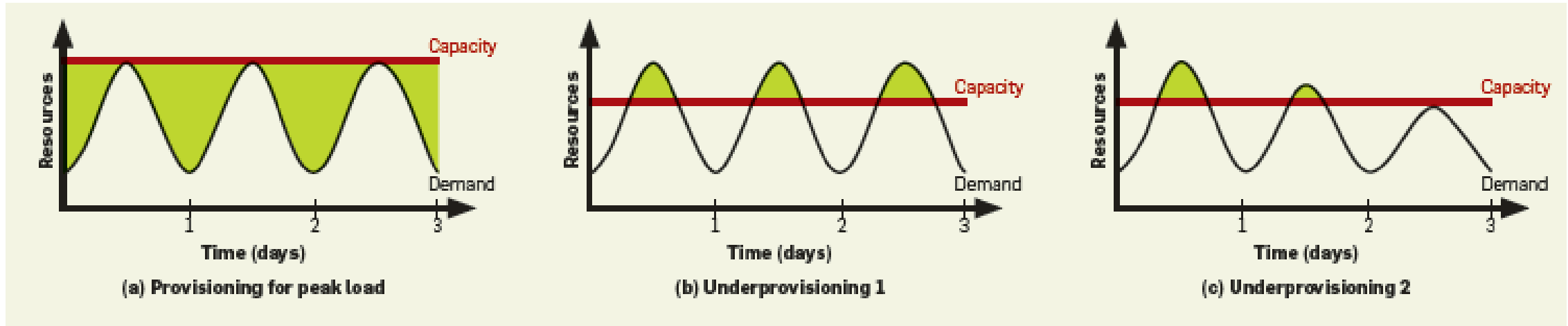
Benefícios do Uso da Virtualização

- *Tabela 2 – Problemas comuns dos serviços de TI encontram na virtualização uma alternativa.*

Alternativa da virtualização	Problemas comuns dos serviços de TI
Otimização e centralização de recursos	Baixa utilização dos recursos de <i>hardware</i>
Isolamento de aplicações	Segurança
Flexibilidade na alocação de recursos	Disponibilidade e confiabilidade dos serviços
Simplificação da infraestrutura	Processos laboriosos de operação

Otimização e Centralização de Recursos

- Menor capacidade ociosa
- Recursos mais centralizados
- Por exemplo, a centralização traz ganhos de eficiência em termos de espaço, gastos com energia, segurança física e administração do ambiente.

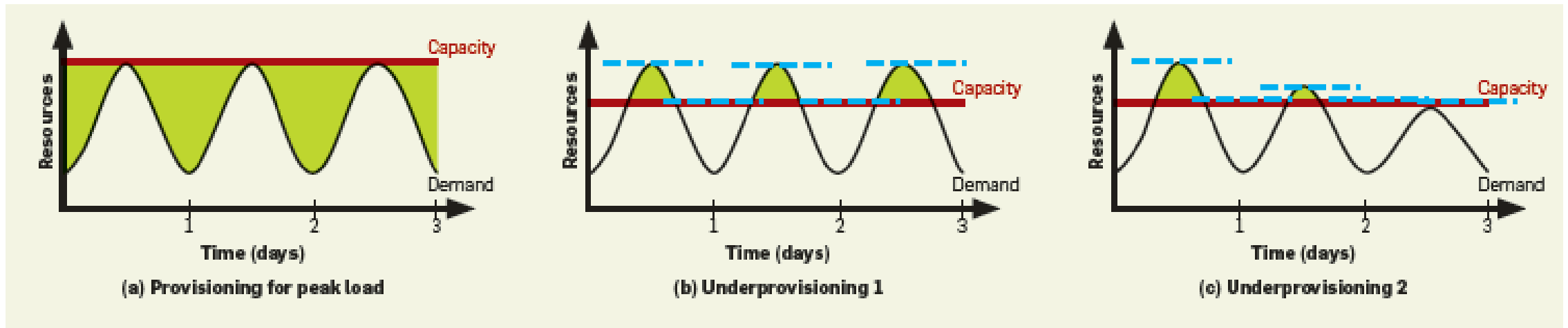


Isolamento de Aplicações (+Segurança)

- Encapsulamento – Embora não faça dos requisitos de virtualização, uma funcionalidade adicional dos sistemas recentes de virtualização, é a capacidade do MMV encapsular por completo o estado de uma máquina virtual em execução, incluído seus dados, SO, aplicações e processos [Rosenblum e Garfinkel 2005].
- VPC/VPN – Flexibilidade para Isolamento de redes e comunicação.
- Controle de Recursos, é feito pelo MMV.
- Por exemplo, uma máquina comprometida por um vírus, pode estar em uma rede isolada ou seus acessos alterados sem intervenção física e/ou pode ser recriada a partir de uma imagem original não comprometida.

Flexibilidade na Alocação de Recursos

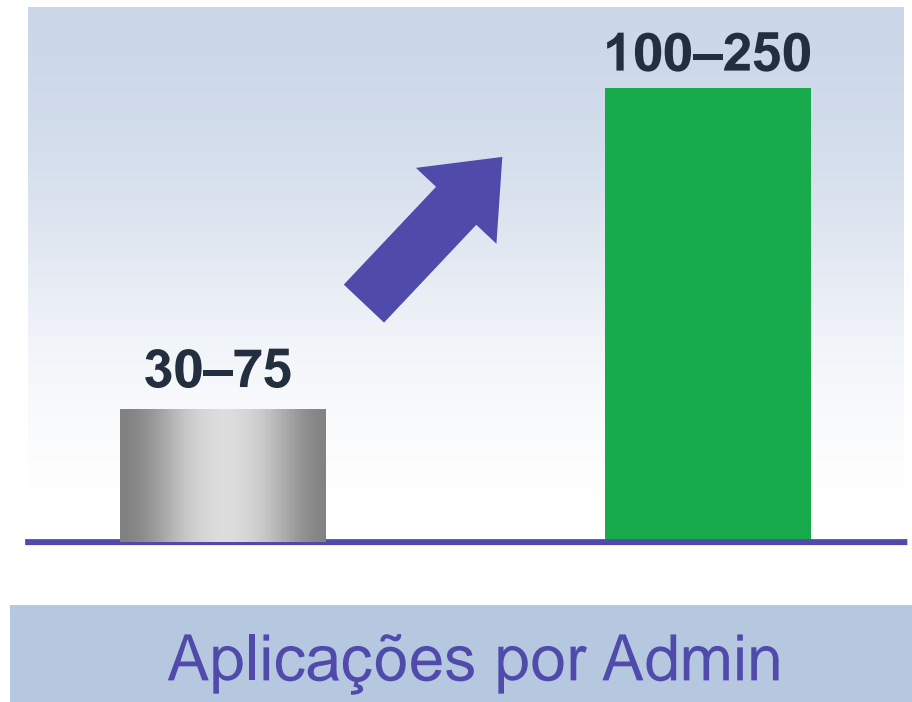
- Gestão de Workload. Aumento e redução de capacidade por dia, hora ou menos.
- Escalonamento automático.
- Micropartições.
- Por exemplo, o escalonamento permite o aumento de um recurso como memória ou cpu quando uma máquina atinge um threshold de 80%. Sistemas modernos permitem fazer isso e reiniciar a máquina (cpu).



Simplificação da Infraestrutura

- Software Defined Network, Storage...
- Ferramentas de Automação
 - Boto3 (AWS)
 - Azure SDK (Azure)
 - OpenStack SDK
 - Terraform
 - GPC Client (Google)

2–3x em Ganhos de Produtividade



Fonte IDC

Simplificação da Infraestrutura (Boto3)

```
import boto3

# Definir as credenciais da AWS
aws_access_key = 'SEU_ACCESS_KEY'
aws_secret_key = 'SEU_SECRET_KEY'
region = 'us-west-2'

# Criar um cliente EC2
ec2_client = boto3.client('ec2', region_name=region, aws_access_key_id=aws_access_key, aws_secret_key=aws_secret_key)

# Definir parâmetros para a instância EC2
instance_params = {
    'ImageId': 'ami-0c55b159cbfafa1f0',
    'InstanceType': 't2.micro',
    'KeyName': 'seu-key-pair',
    'SecurityGroups': ['seu-security-group']
}

# Criar a instância EC2
response = ec2_client.run_instances(**instance_params)
```

Simplificação da Infraestrutura (Azure Client)

```
# Credenciais do Azure
credential = AzureCliCredential()

# Cliente de gerenciamento de computação
compute_client = ComputeManagementClient(credential, 'SUA_SUBSCRIPTION_ID')

# Definir parâmetros para a máquina virtual
vm_params = {
    'location': 'SUA_LOCALITY',
    'hardware_profile': HardwareProfile(vm_size='Standard_B1s'),
    'os_profile': OSProfile(computer_name='vm-test', admin_username='azureuser',
    'storage_profile': StorageProfile(image_reference=ImageReference(publisher='
    'network_profile': NetworkProfile(network_interfaces=[{'id': '/subscriptions

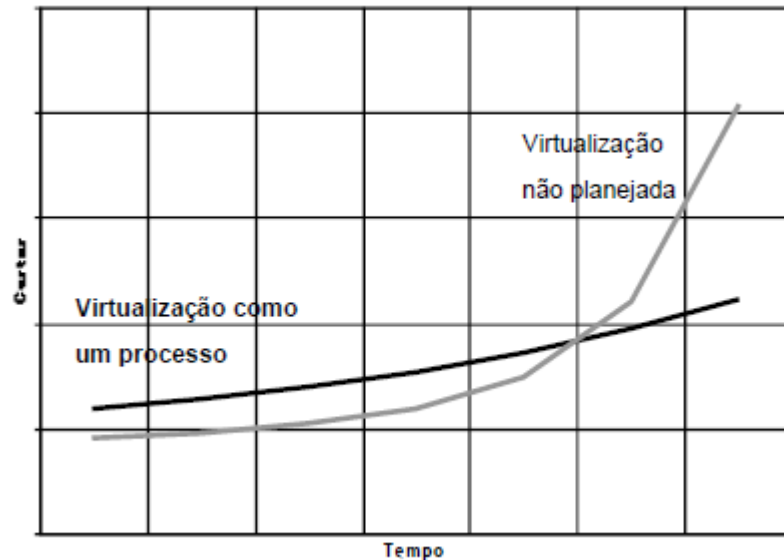
}

# Criar a máquina virtual
async_vm_creation = compute_client.virtual_machines.begin_create_or_update('SUA_

# Aguardar a conclusão da criação ↓
async_vm_creation.wait()
```

Virtualização SEM Processos

- Figura 4 – A adoção da virtualização com foco unicamente na redução de máquinas físicas (virtualização não planejada) apresenta um custo inicial menor que a criação de um processo permanente com vistas a virtualização. Este último requer um investimento inicial maior em processos, ferramentas e automação. Entretanto, a medida que o ambiente torna-se mais complexo, prolifera-se o número de imagens e alocações dinâmicas, torna-se mais difícil o controle do uso de recursos. Aumenta-se a necessidade de provisionamento de recursos ociosos e o tempo de alocação de recursos para as aplicações.

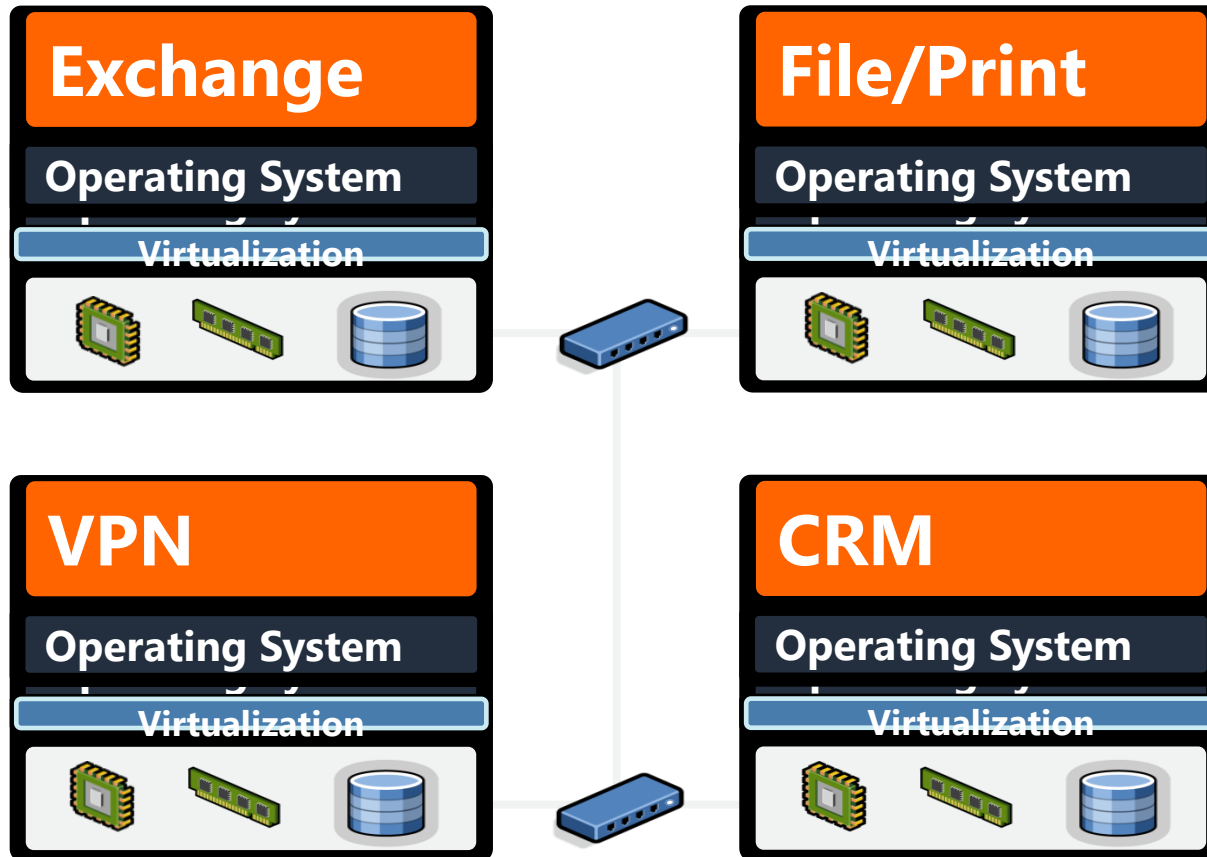


Como evitar essa deterioração?

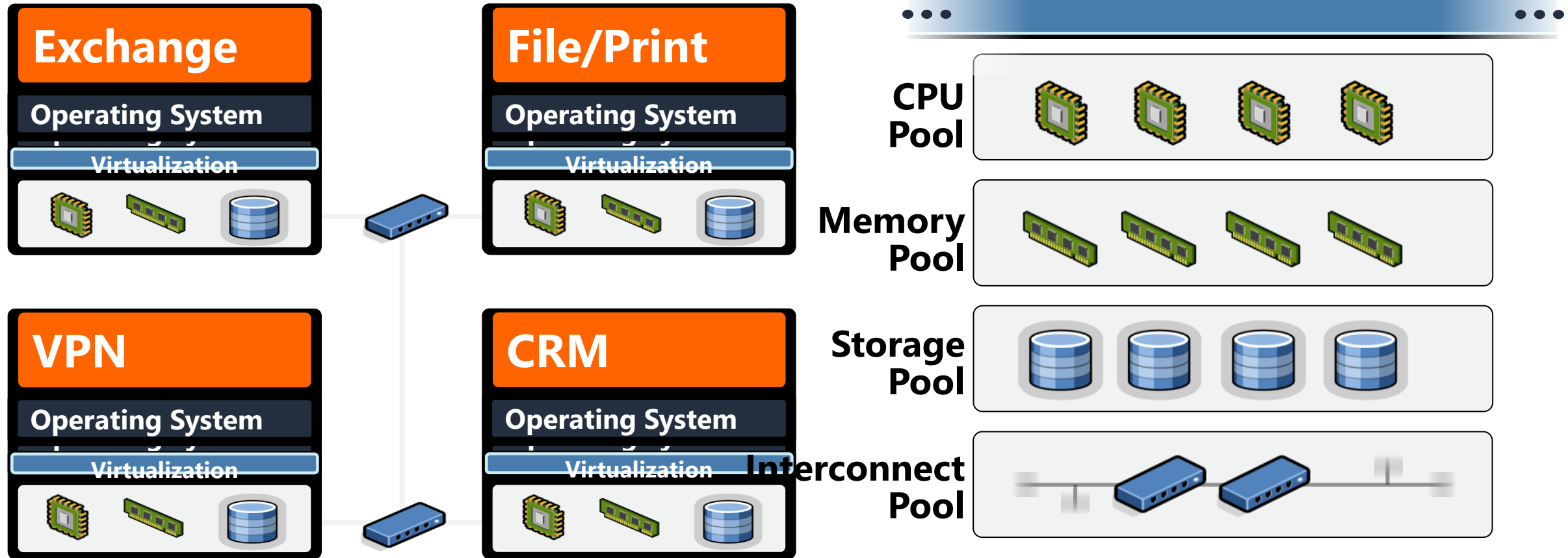
Questões da Virtualização

- *Tabela 3 – Questões e riscos envolvidos no emprego da virtualização e possíveis abordagens e ações.*
- Custos de Entrada na Tecnologia
- Ausência de Dados e Métricas para Lidar com Recursos Compartilhados
- Licenças de Software
- Gerenciamento
- Impactos Humanos
- Diferente Complexidade do ambiente
- Questões de Segurança
- Compartilhamento de Ambientes com Diferentes Níveis de Serviço
- Desempenho
- Definição de Hardware e Capacidade

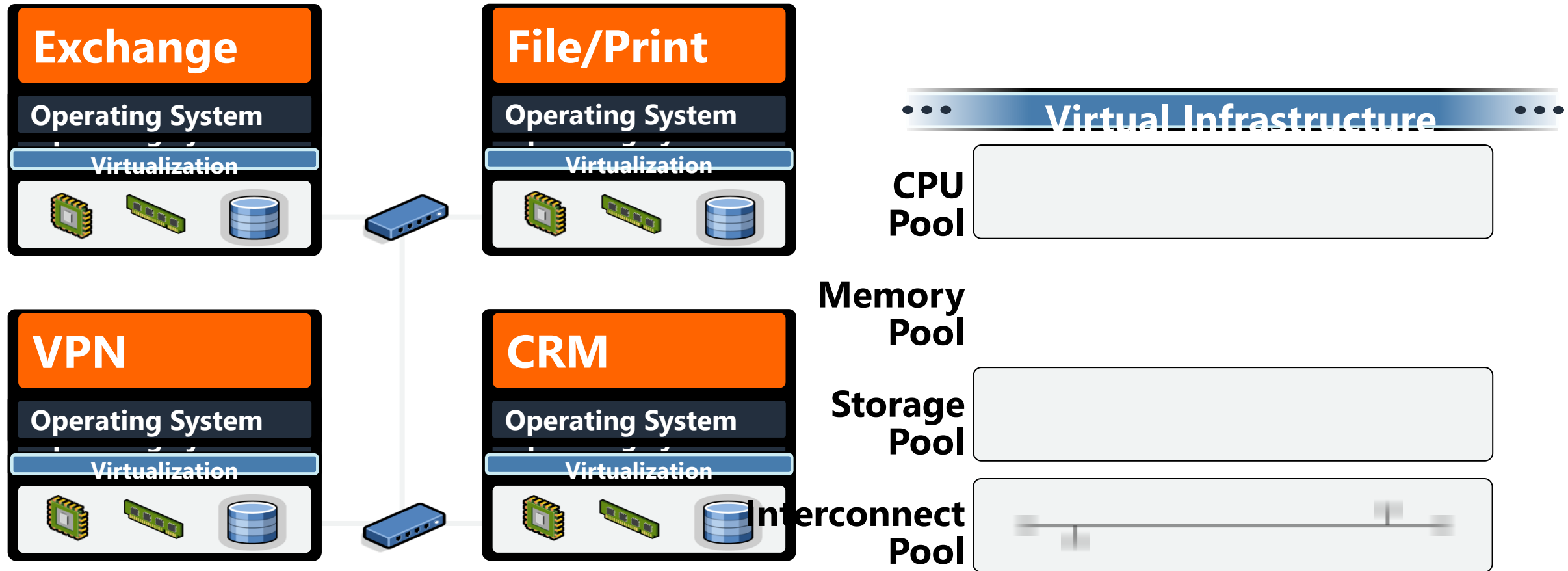
Visão Tradicional



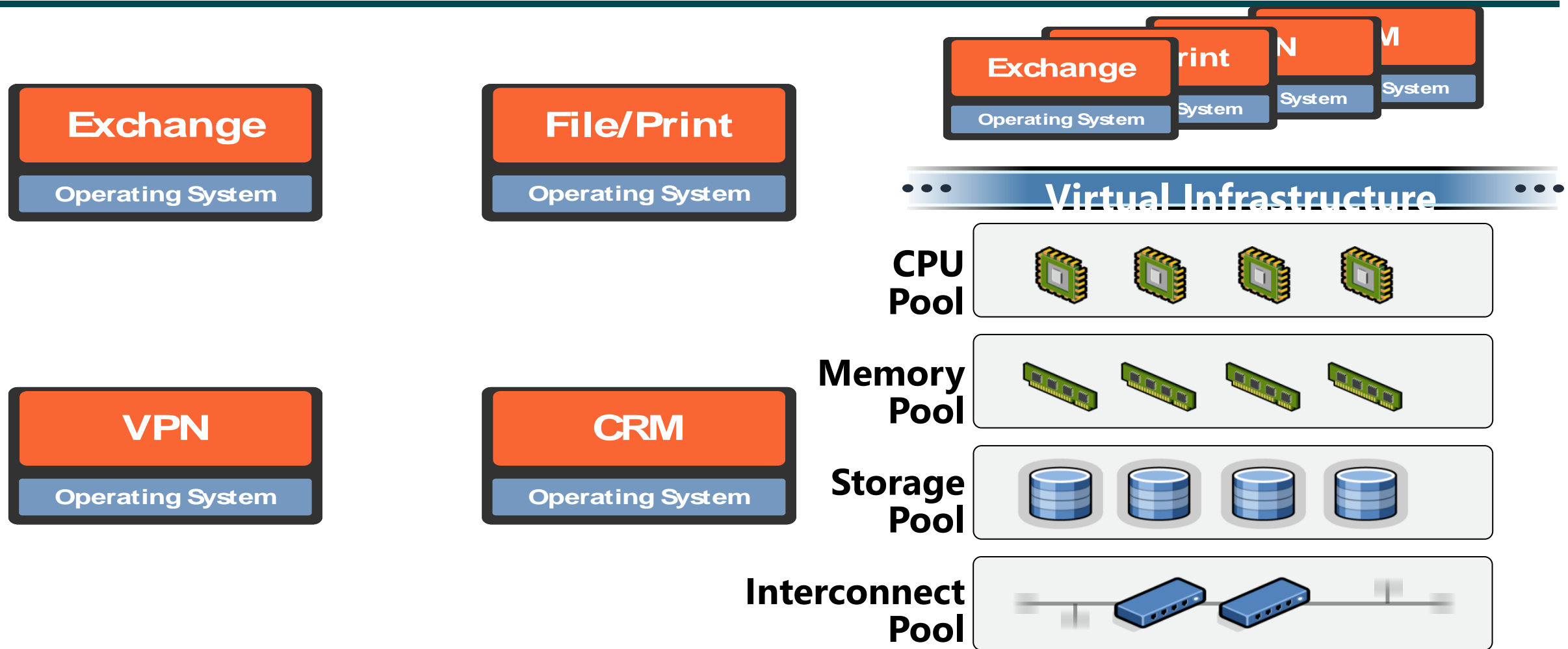
Virtual Data Center



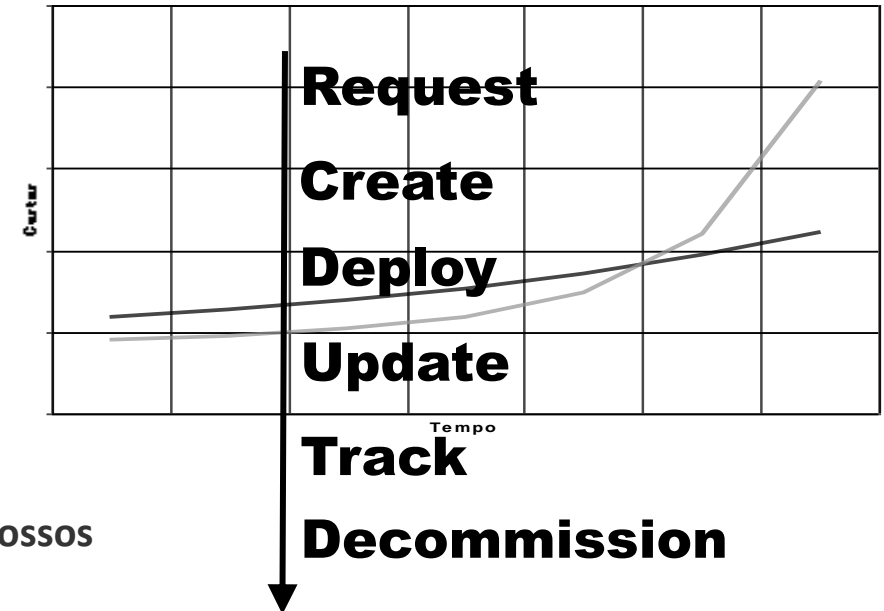
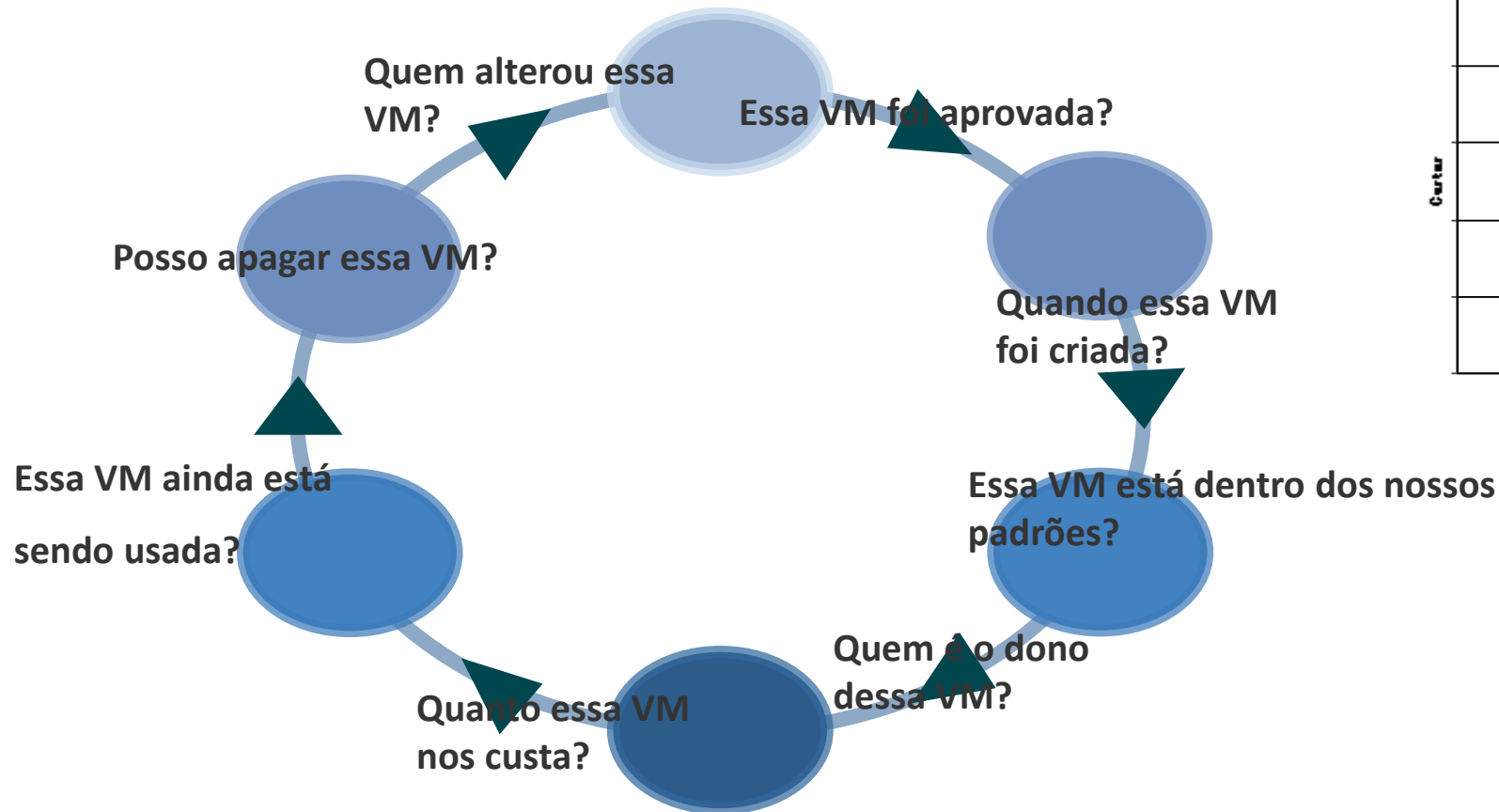
Virtual Data Center



Virtual Data Center



Gerenciamento Ciclo de Vida das VMs

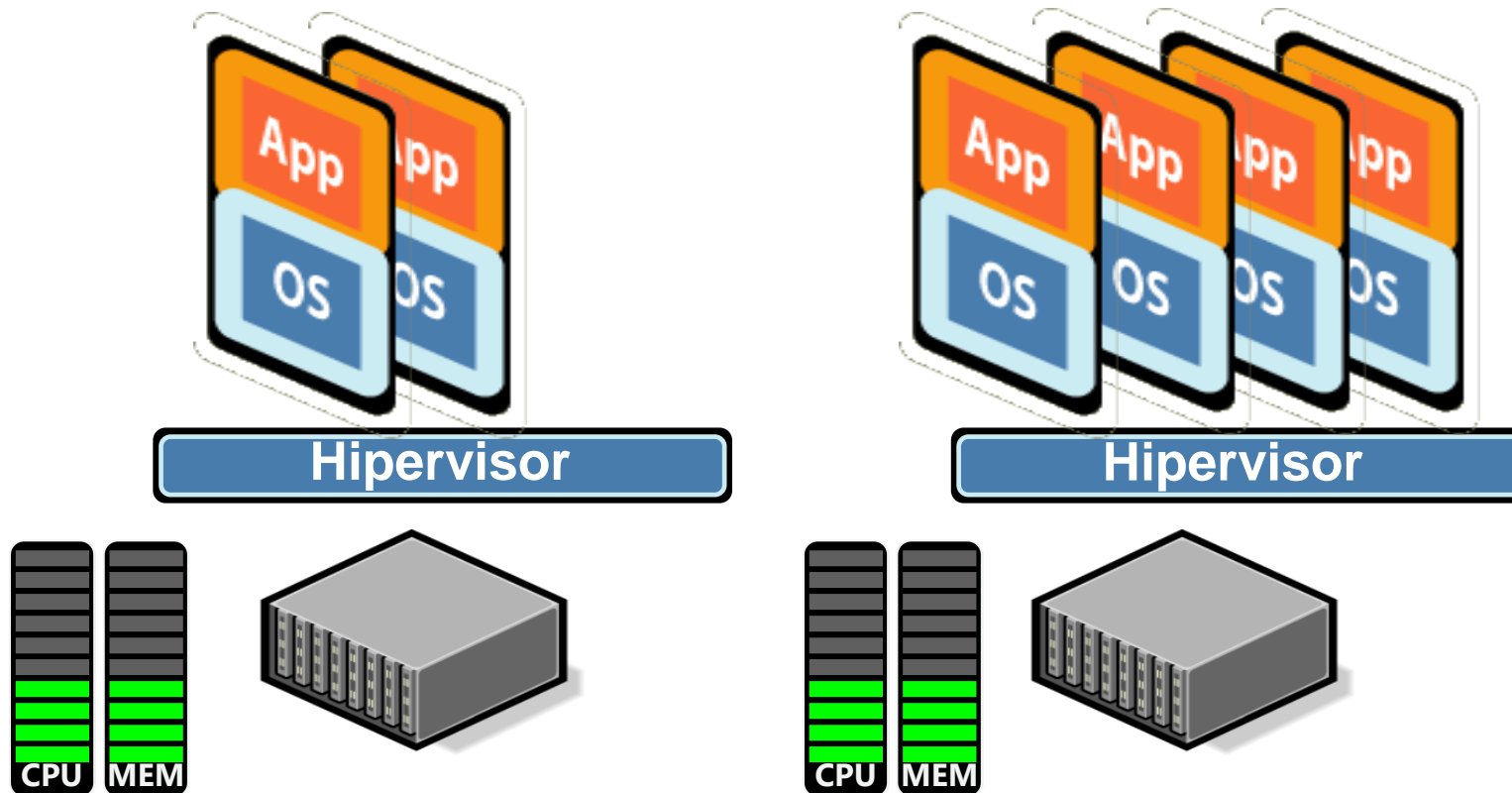


+Processos

Downtime Planejado



Scheduler Distribuição de Recursos

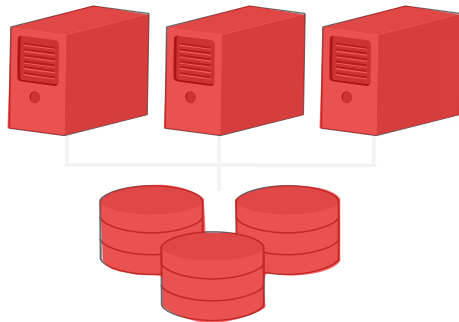


Disaster Recovery

Produção



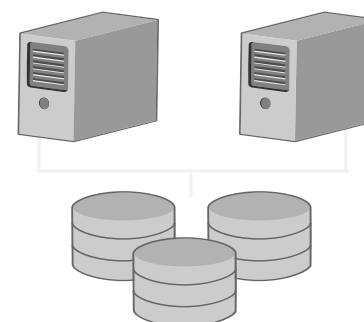
VM Infrastructure



Site Backup (AZ ou Região)

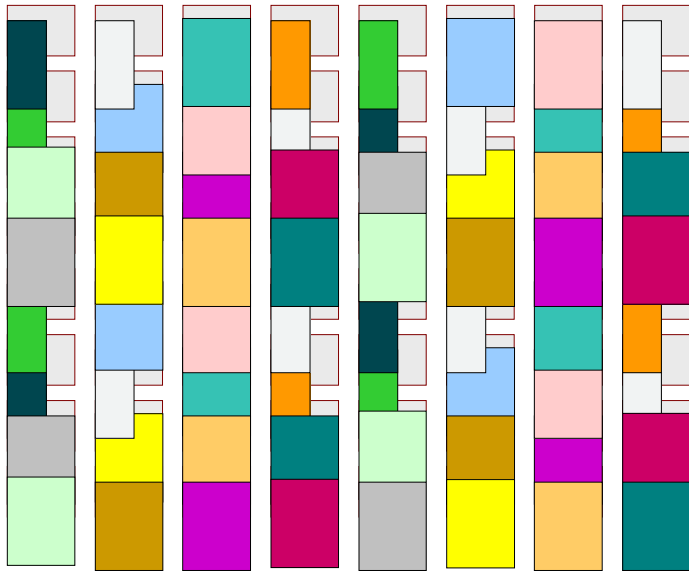


VM Infrastructure

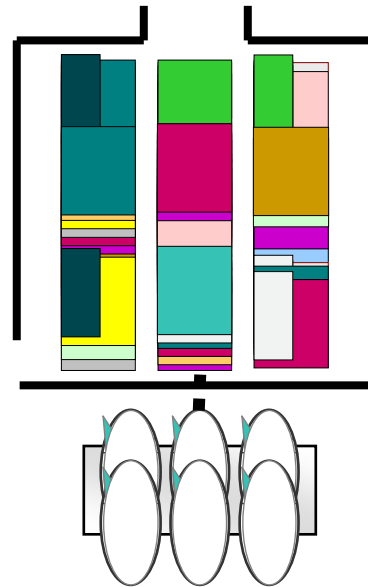


Virtual Data Center

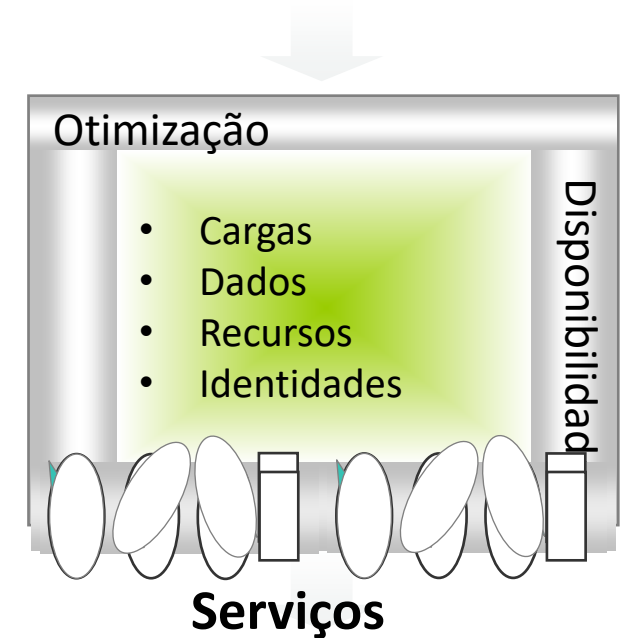
Física



Virtualizada



Automatizada



De "Virtualization Changes Virtually Everything," Thomas Bittman, Gartner, Novembro 2007.

