Disaster Recovery Myths and Reality

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, open-source developer

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

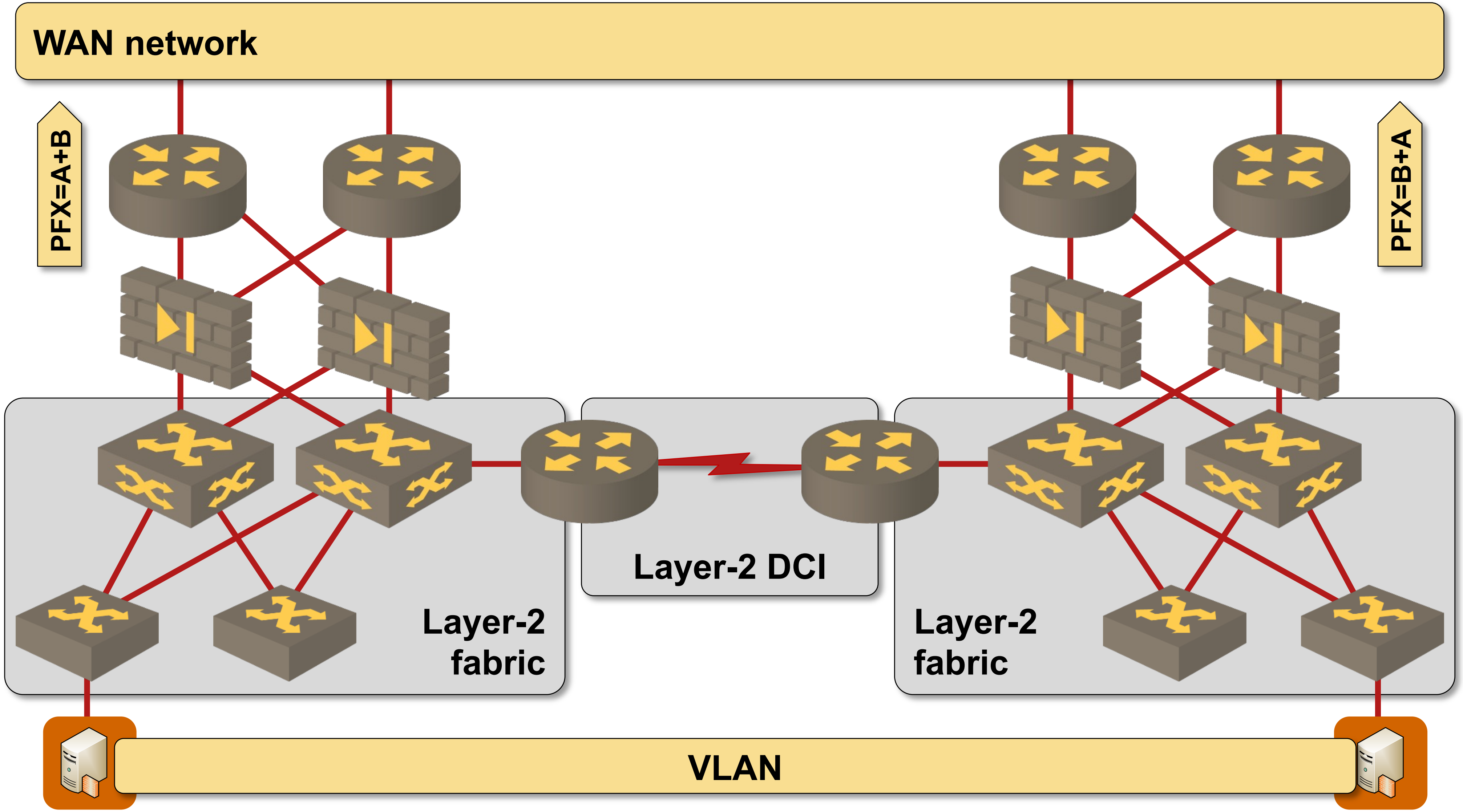
WE MUST HAVE

DISASTER RECOVERY

Vendor Consultants

Always have a solution

Typical Vendor-Recommended Solution

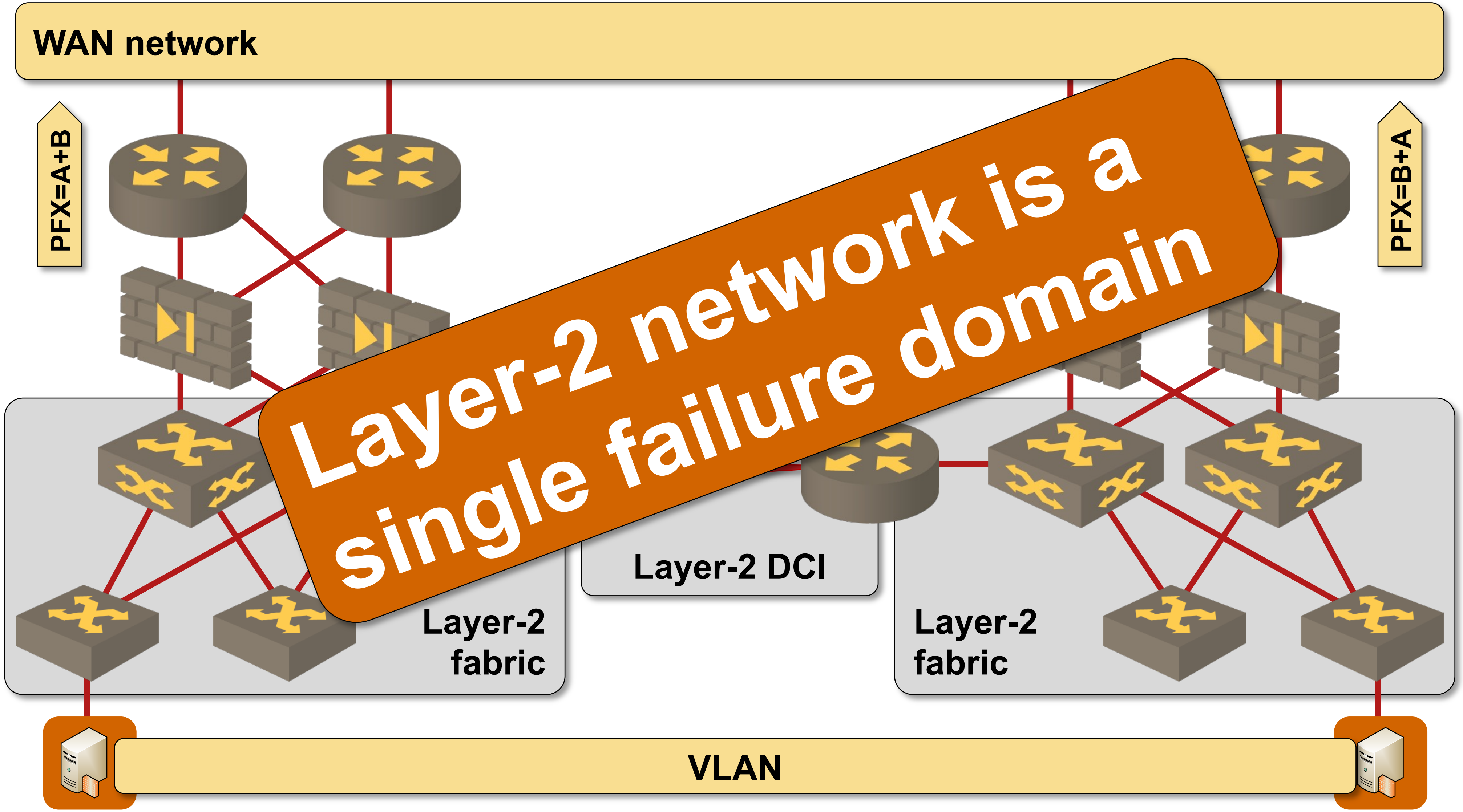


- Layer-2 fabric to support VM moves
- VLANs stretched across multiple sites
- Active/standby or distributed storage

Handwaving

- Stateful services
- Ingress traffic
- Bandwidth requirements
- Latency

The Elephant in the Room



Layer-2 network is a single failure domain

Based on a True Story

Regional cloud (outsourcing) provider

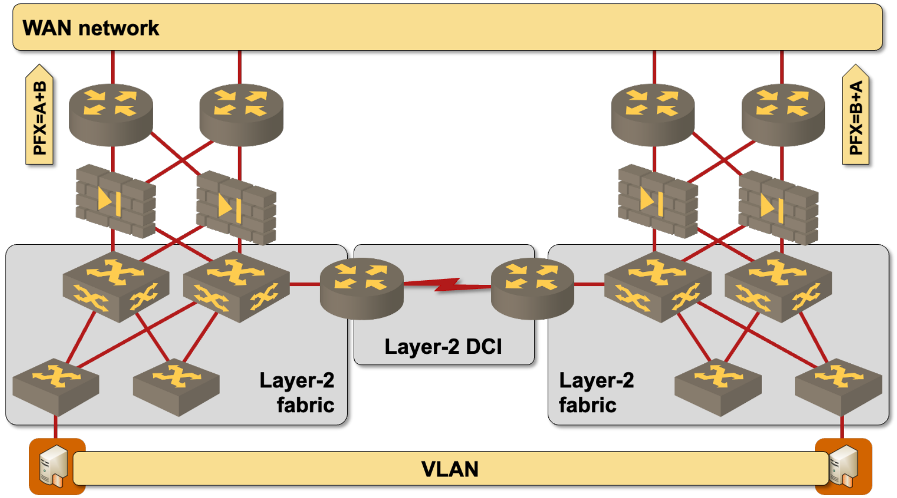
- Picture-perfect physical design: two sites, six fire zones per site, fully redundant infrastructure
- Two core switches per site (dubious) in an MLAG cluster (bad)
- Edge switches in MLAG clusters

Vendor-recommended solution

- VMware ESXi cluster stretched across both sites
- Automated VM restart on cluster node failures
- VLANs stretched across both sites to facilitate faster recovery (really bad)

What happened

- A software bug in fabric vendor MLAG code
- Forwarding loop between two edge switches
- Both data center fabrics became overloaded
- They lost the whole infrastructure for a few hours
- Vendor-proposed DR “solution” caused the disaster



The sad part: their DR ideas wouldn’t work anyway

Aside: Disaster Avoidance Is a Myth

The Promises

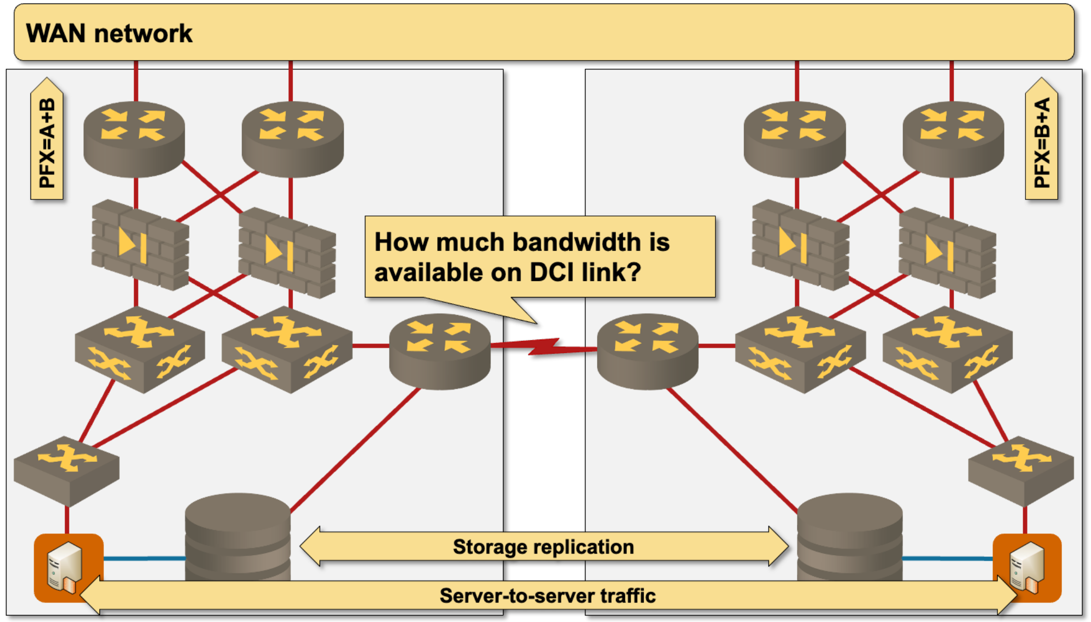
- Migrate virtual machines to the backup site before shutting down the primary site
- Seamless transition (according to vendors) with no downtime

The Reality

- How much bandwidth do you have available for migration?
- A lot more traffic will flow over the WAN link
- Non-trivial performance degradation during the whole migration process due to WAN link overload and latency

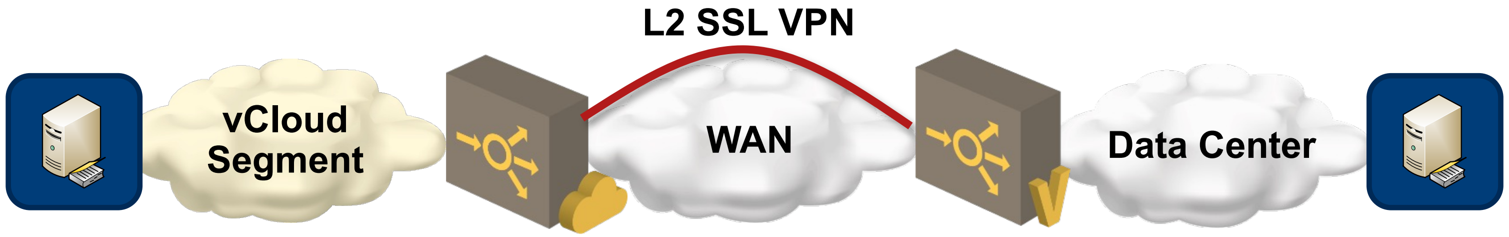
The Deal Breaker

- How long will it take to migrate the critical parts of the data center?



Much better: Shut down whole systems and restart them on the other site

Aside: Cloudbursting Works Best in PowerPoint

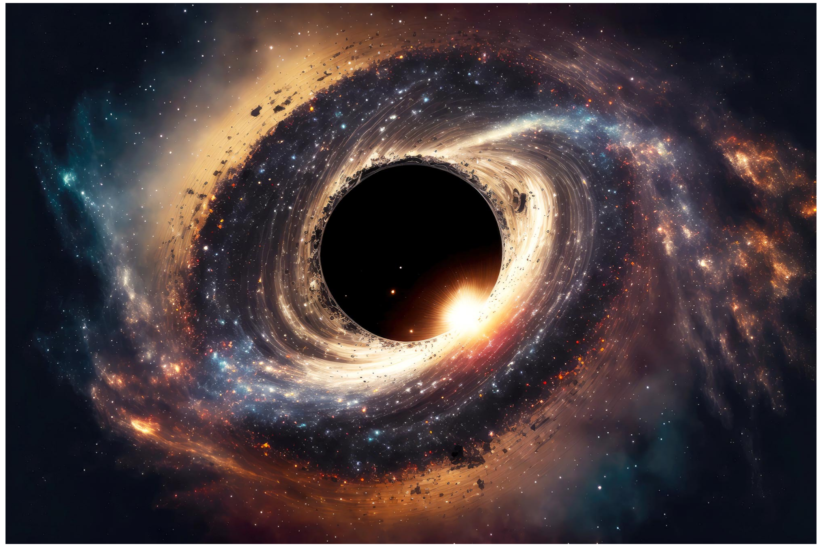


What vendors are promising

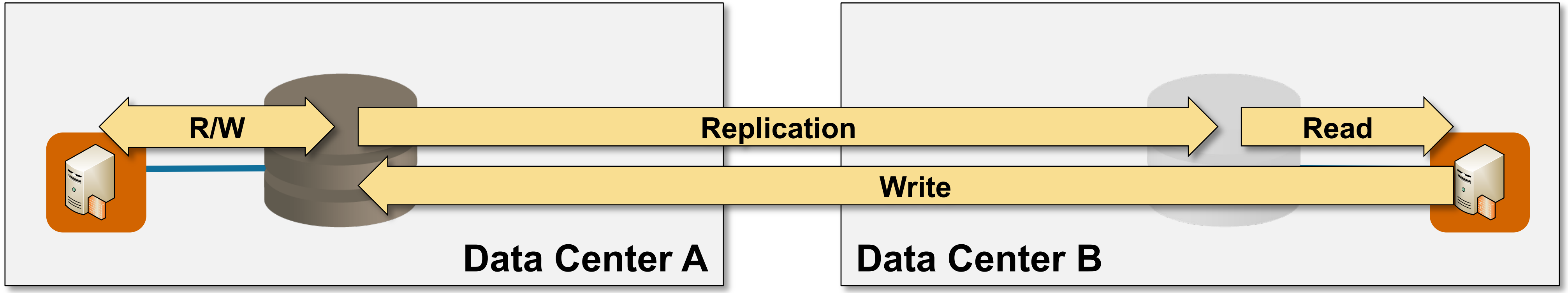
- Seamless migration into the public cloud
- Live VM migration with no downtime

The reality

- Most (sane) public clouds don't support L2 extensions
- VMware can make connectivity work with NSX (but it's still useless)
- Latency will kill the application performance (data has gravity)



Back to Reality: Storage



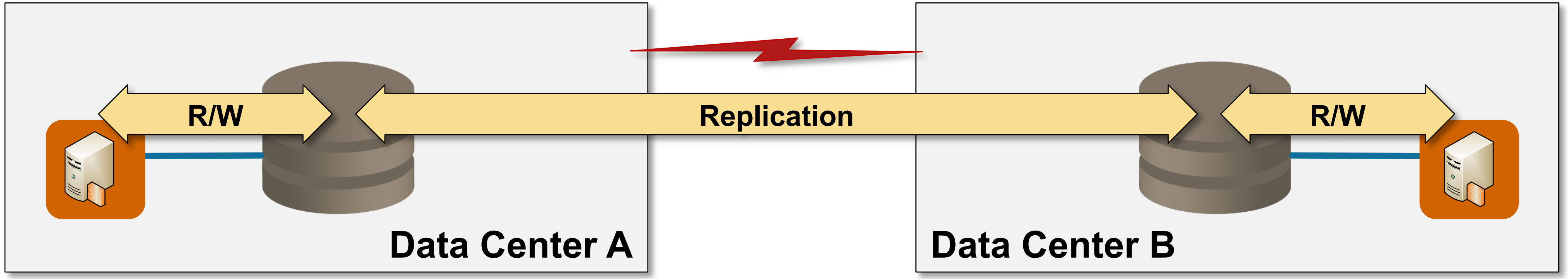
Storage and errors are permanent

- Synchronous replication between data centers (kills performance with SSD disks)
- One site is active; the other one is read-only
- Active-active storage arrays are dangerous

Alternate approach

- Asynchronous replication (could result in data loss)
- Database log shipping (non-zero RPO)
- Periodic backups (acceptable for VM disks)

It Gets Better: Distributed File Systems



- A cluster of servers replaces disk arrays
- NFS or iSCSI access to local target
- Works great within a site
  (used by all large cloud providers)

Vendors don’t stop there

- Stretch the active/active cluster across sites
- Continuous replication between sites

Minor annoyances

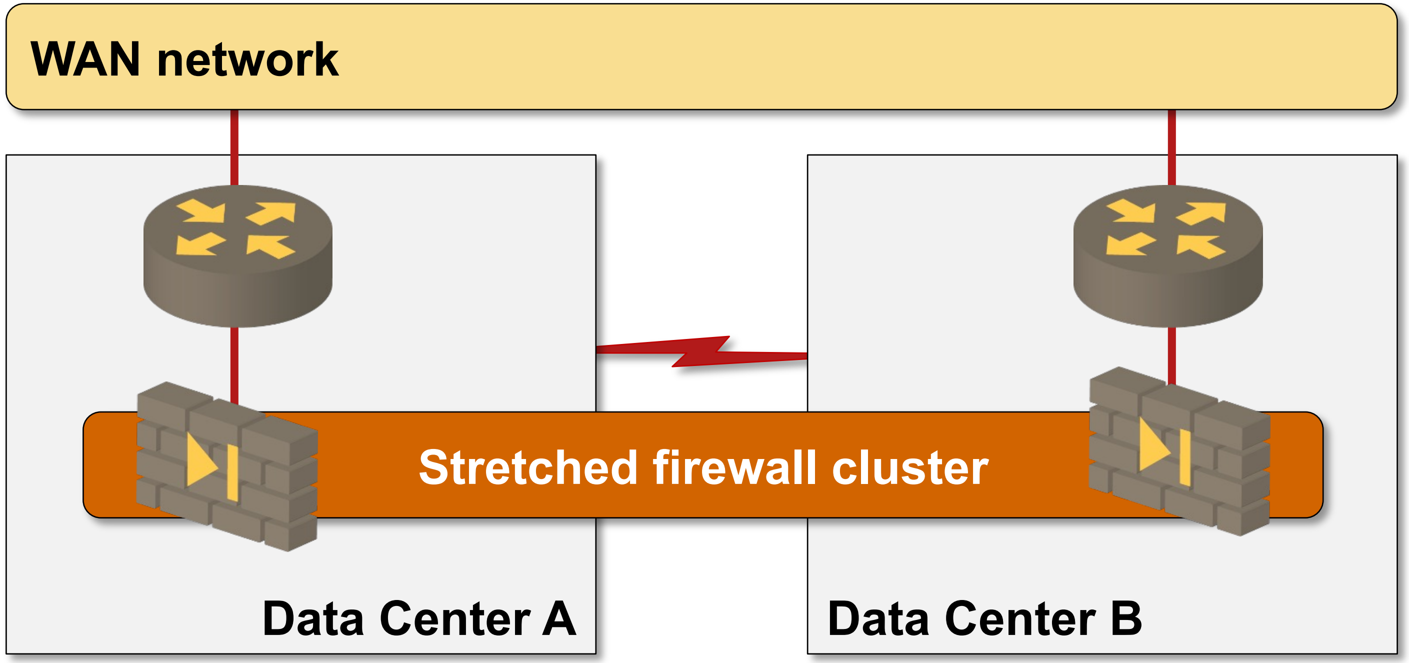
- Synchronous replication kills SSD performance
- Reduced performance for cross-site reads

Catastrophic scenario

- What happens when the WAN link fails?

Distributed Clusters Are Hard

- Sane clusters have an odd number of components
- How do you do that with two storage arrays or two firewalls?
- **Solution:** witness node

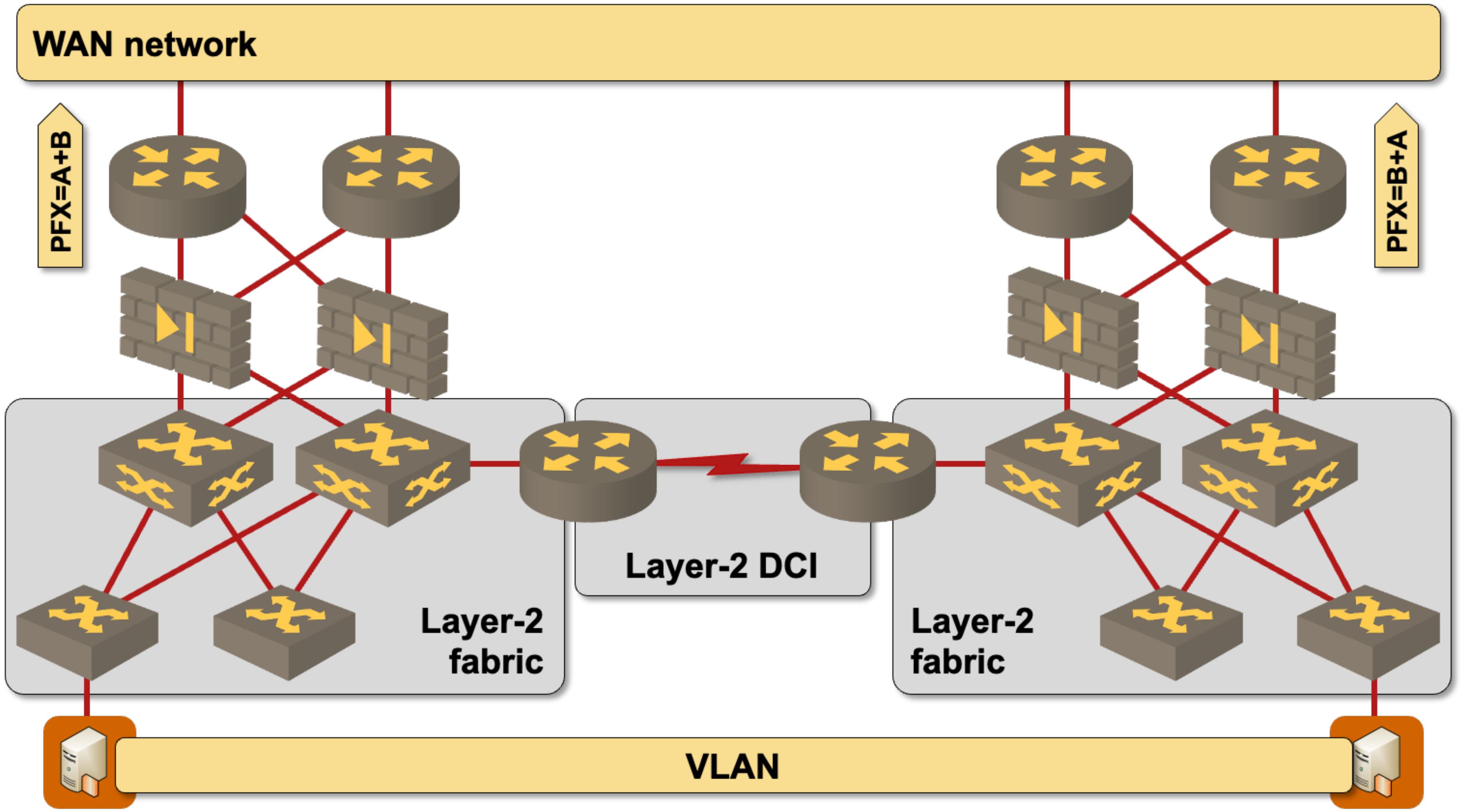


Minor details

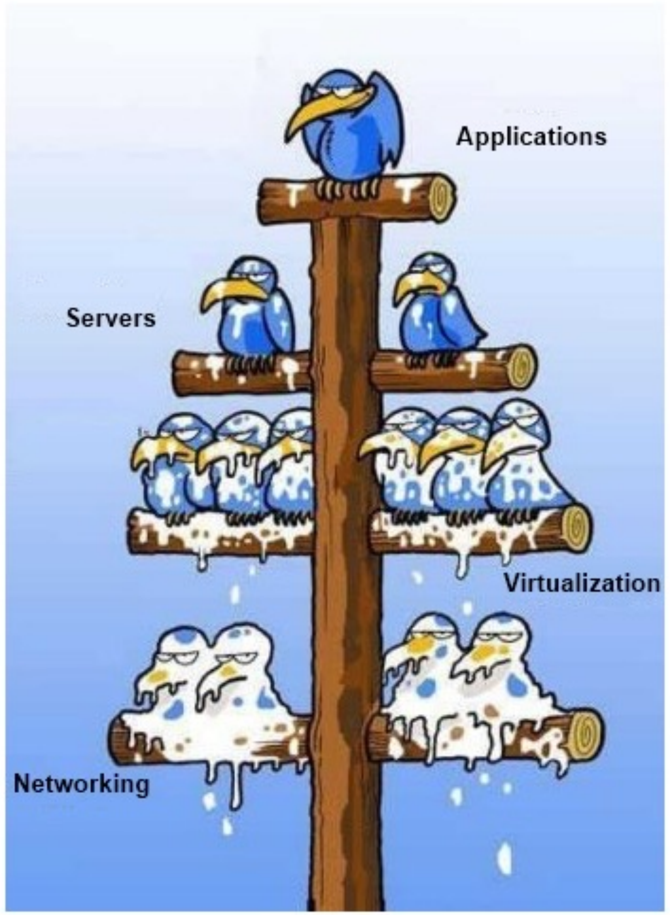
- What is a witness node for network components (switches or firewalls)?
- What happens when the WAN link fails?
- What happens when the storage array is a witness node for an application cluster?
- What happens when the primary site fails?

What Can We Do?

Pushing Application Problems Into the Infrastructure Does Not Work

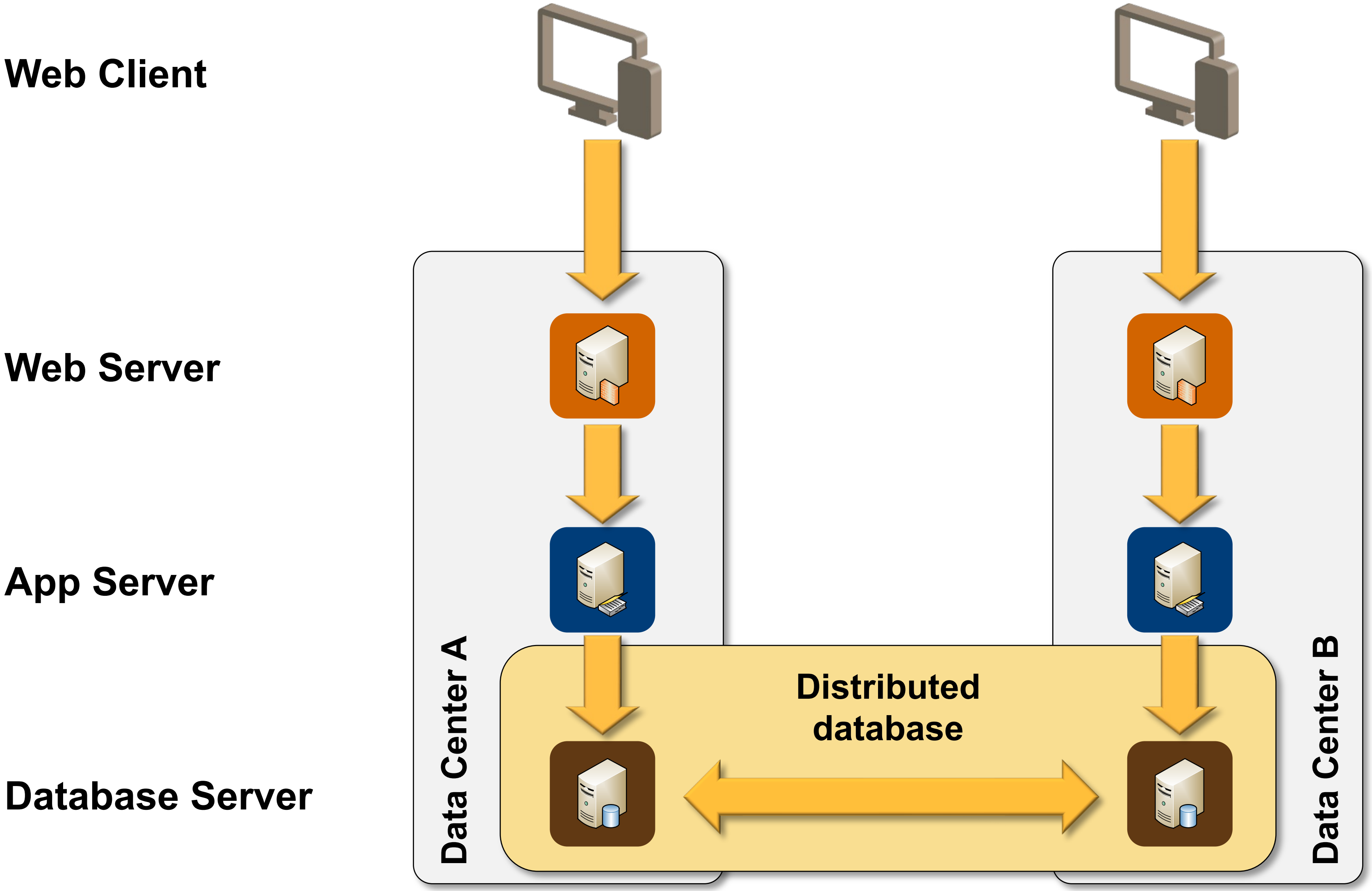


This is what makes networking so complex



There is no silver bullet (or infrastructure solution) for true high availability

Deploying Multiple Application Stacks (Swimlanes)



Web Client

Web Server

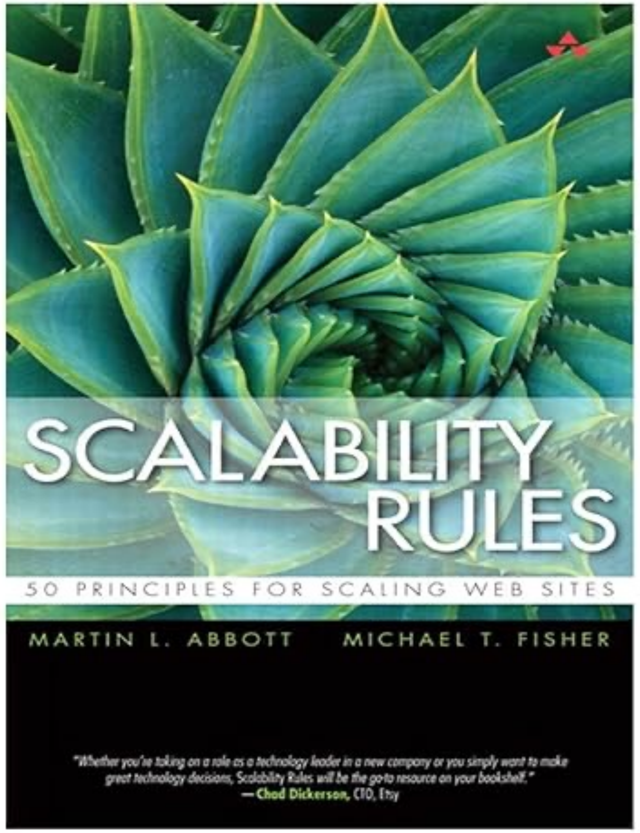
App Server

Database Server

Data Center A

Data Center B

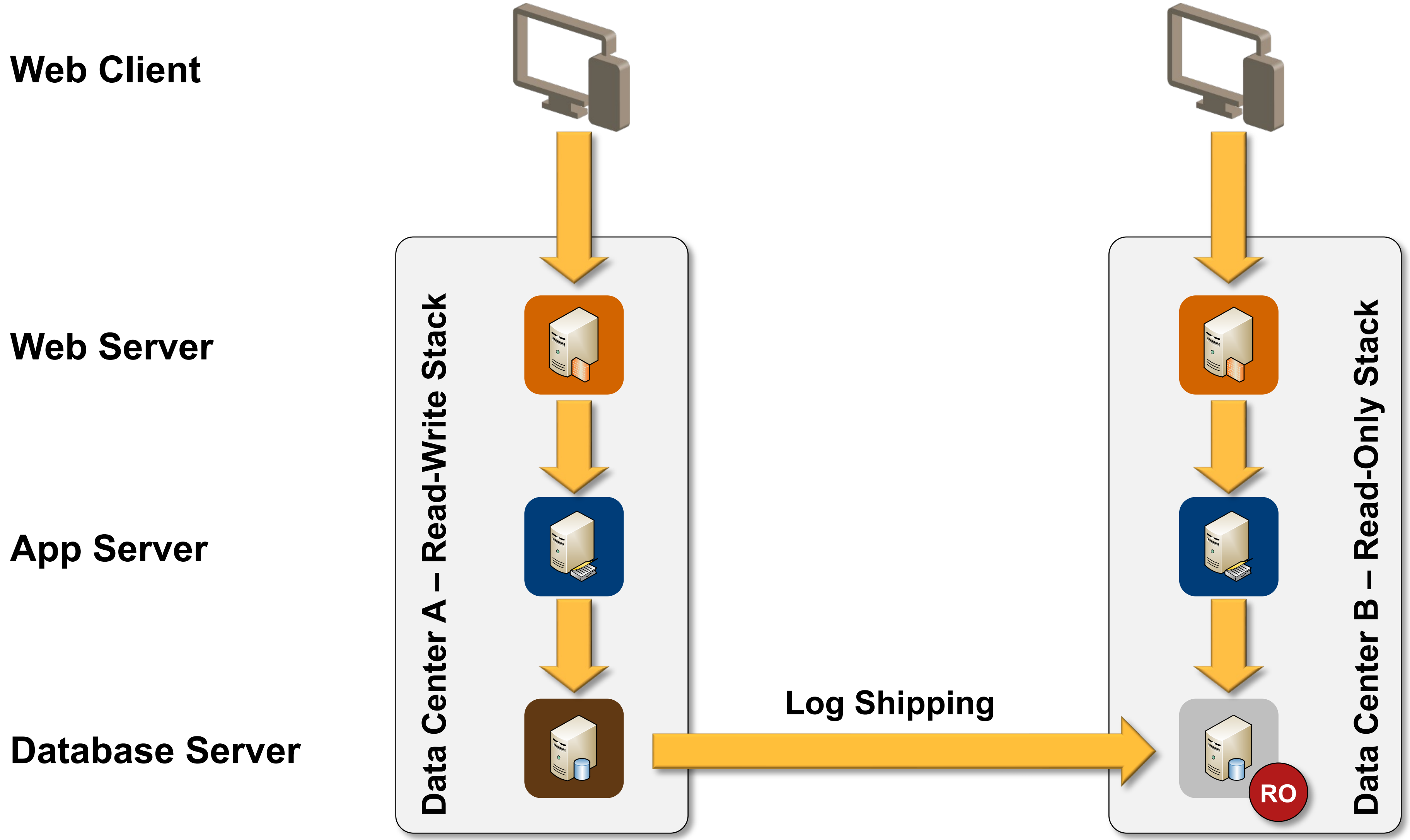
Distributed database



SCALABILITY RULES
50 PRINCIPLES FOR SCALING WEB SITES
MARTIN L. ABBOTT
MICHAEL T. FISHER
"Whether you're taking on a role as a technology leader in a new company or you simply want to make great technology decisions, Scalability Rules will be the go-to resource on your bookshelf."
—Chad Dickerson, CTO, Etsy

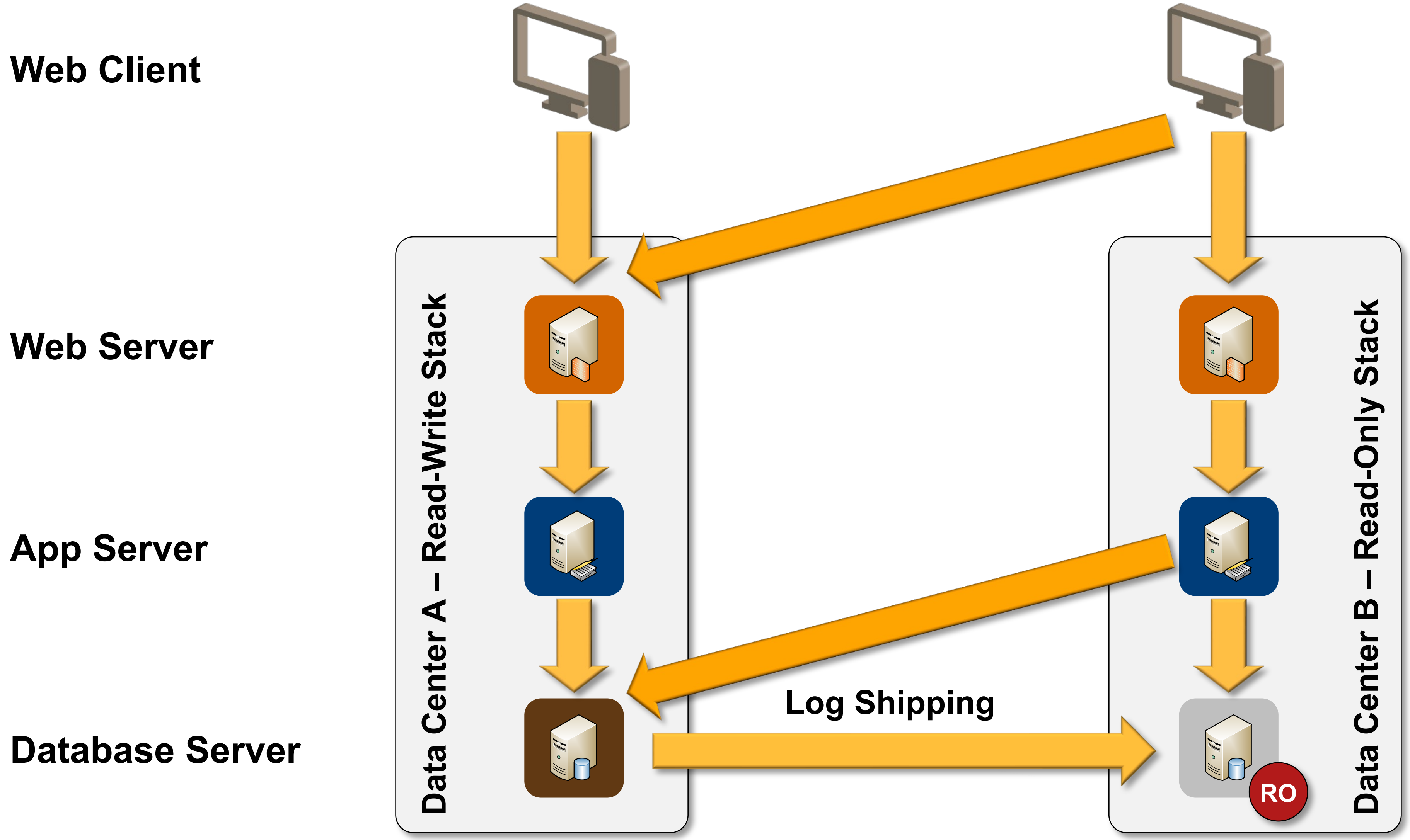
Scalability Rules: 50 Principles for Scaling Web Sites (akfpartners.com)

Read-Write and Read-Only Application Stacks

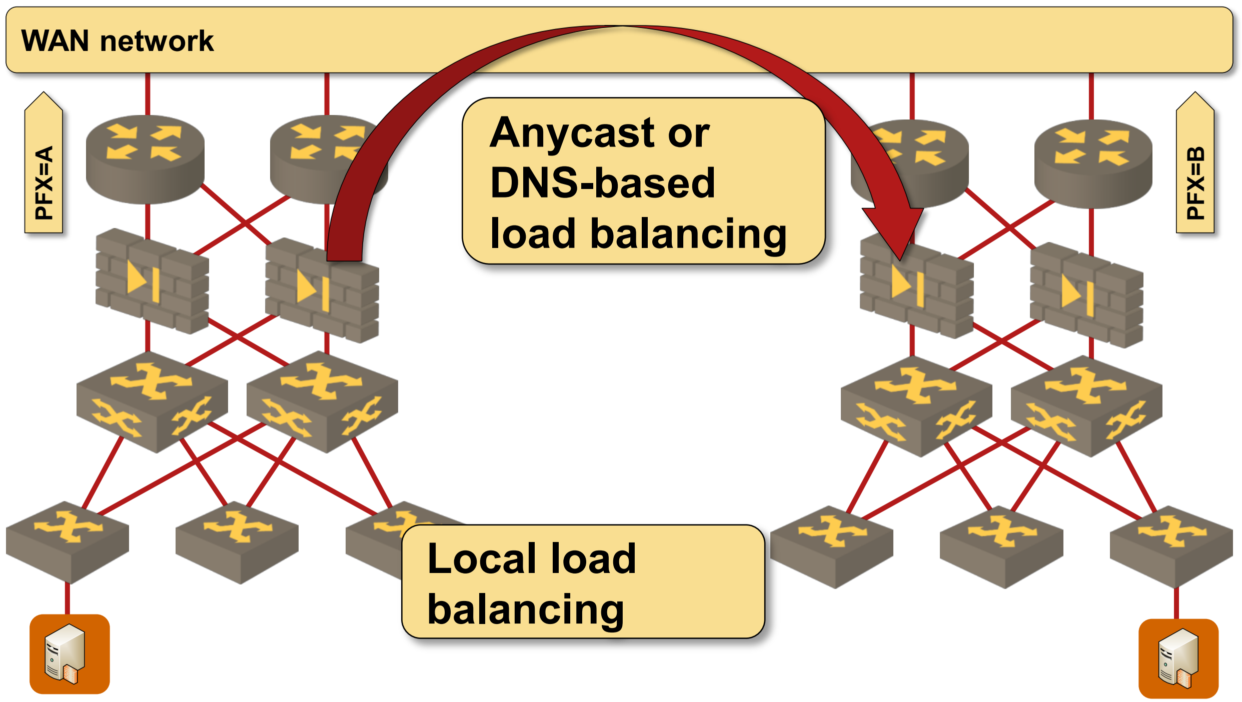


Web Client
Web Server
App Server
Database Server
Data Center A – Read-Write Stack
Log Shipping
RO
Data Center B – Read-Only Stack

Potential Request Redirection Paths

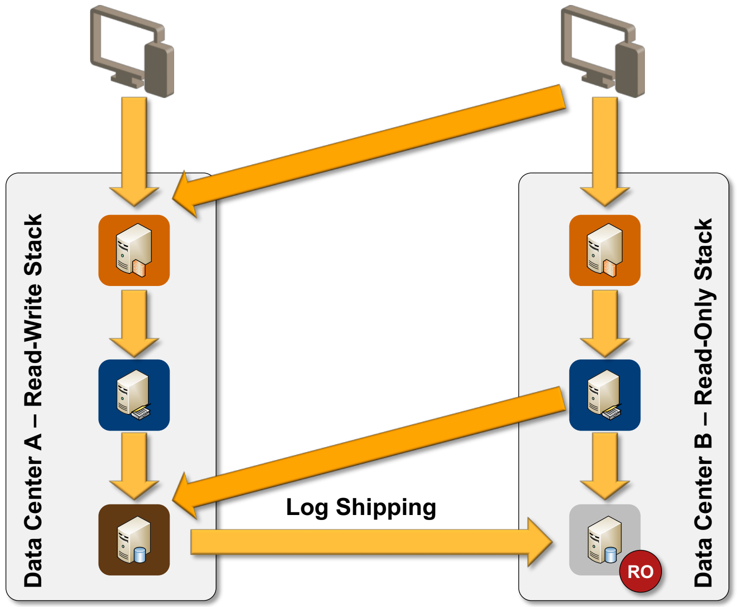

Web Client
Web Server
App Server
Database Server
Data Center A – Read-Write Stack
Data Center B – Read-Only Stack
Log Shipping
RO

Distributed Data Centers with Scale-Out Applications



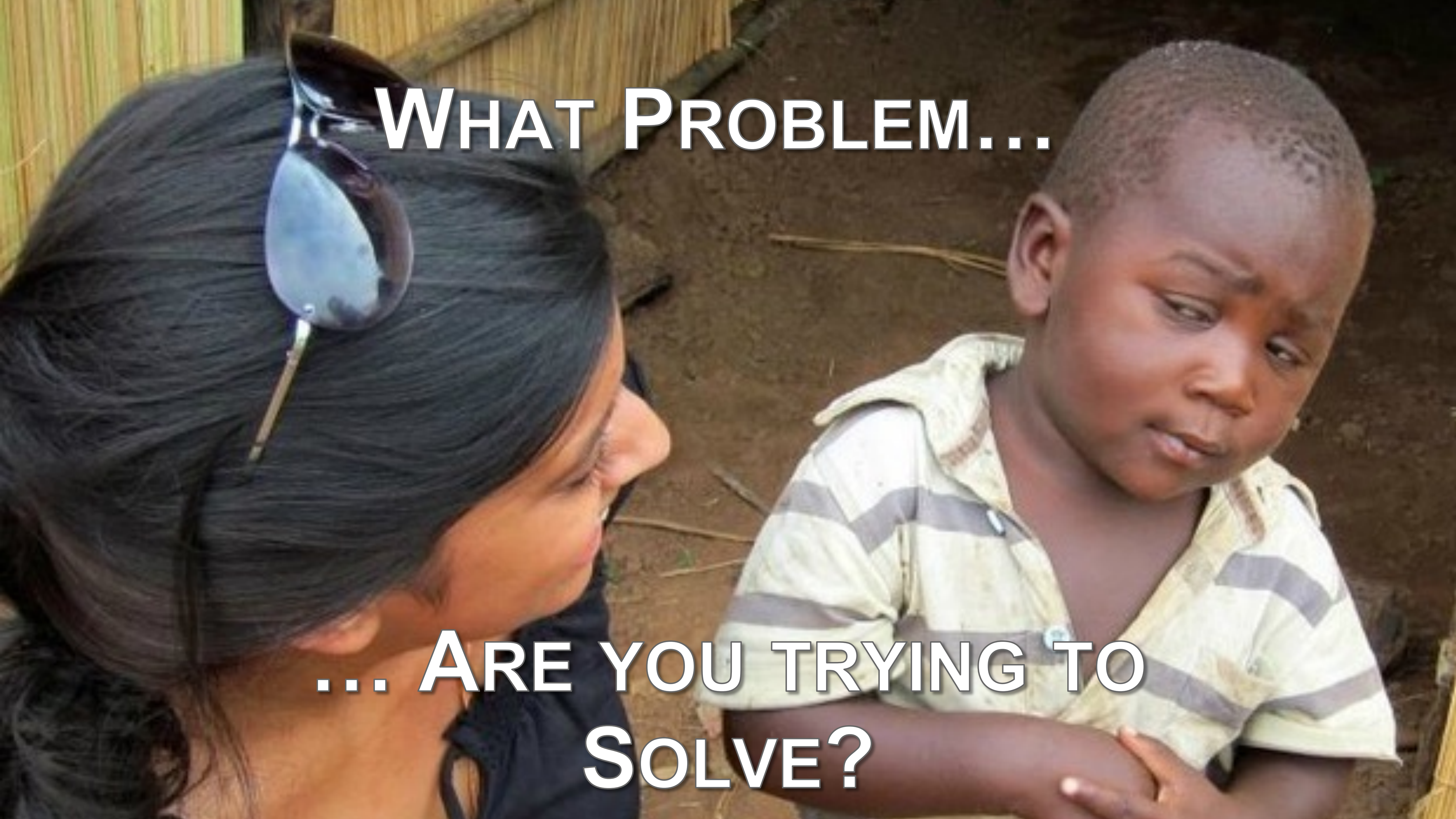
Anycast or DNS-based load balancing

Local load balancing


Data Center A – Read-Write Stack
Log Shipping
RO
Data Center B – Read-Only Stack

- Simple network design
- Easy-to-implement N+1 redundancy (particularly with eventual consistency)
- Works great in many large-scale deployments

Time to Wrap Up



What Problem…

… Are you trying to Solve?

Defining the Challenge

RTO (Recovery Time Objective)

- How long can it take to be back in business?

RPO (Recovery Point Objective)

- How much data can we lose?

Recovery type

- Recovery from backups
- Cold standby
- Warm / hot standby
- Active-active data centers (failover is automatic)

What failures are we protecting against?

- Hardware failures (power supplies, transceivers, links)
- System software failures (server/node crash)
- Application failures

Are things really mission-critical?

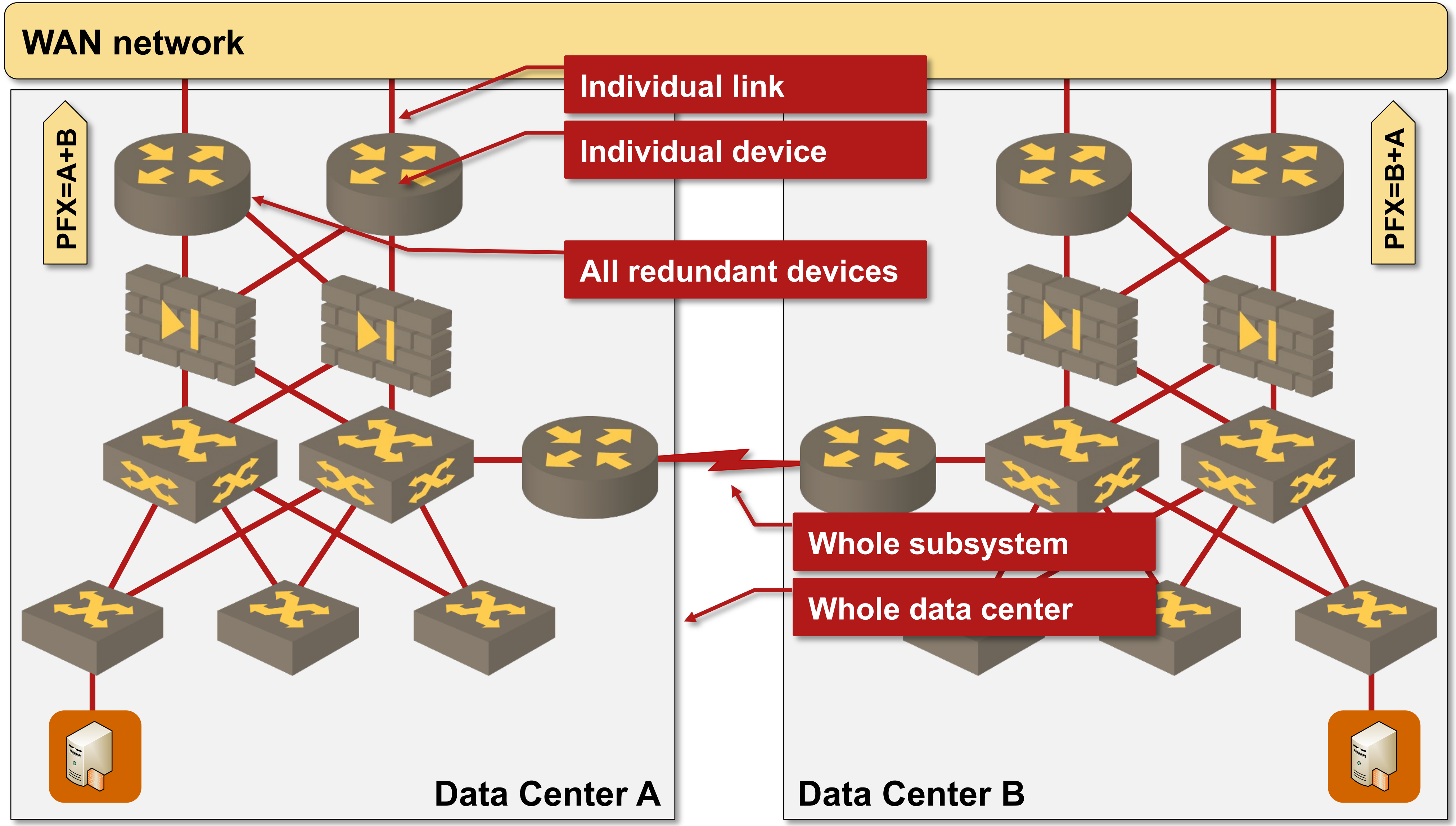
- Does maintenance count as downtime?
- Can you patch servers without downtime?

How much redundancy do we need?

- How often do we have failures?
- What is the impact of failures?
- Can we even get redundant infrastructure?

Active / Active architecture can only be implemented within the applications

What Can Fail?



Prerequisites

Infrastructure

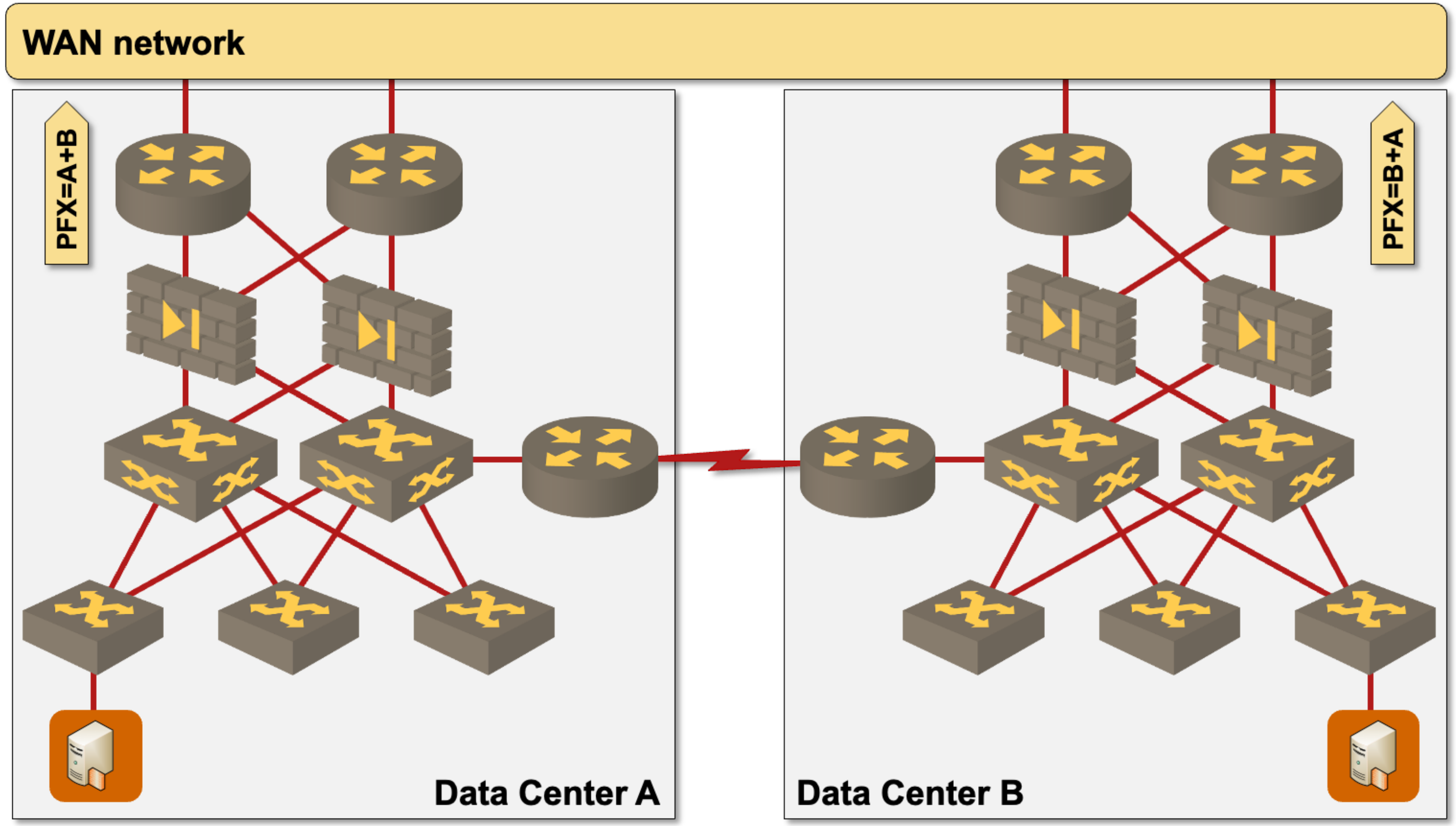
- Networking infrastructure
- Network services
- Storage (databases + virtual disks)
- Compute

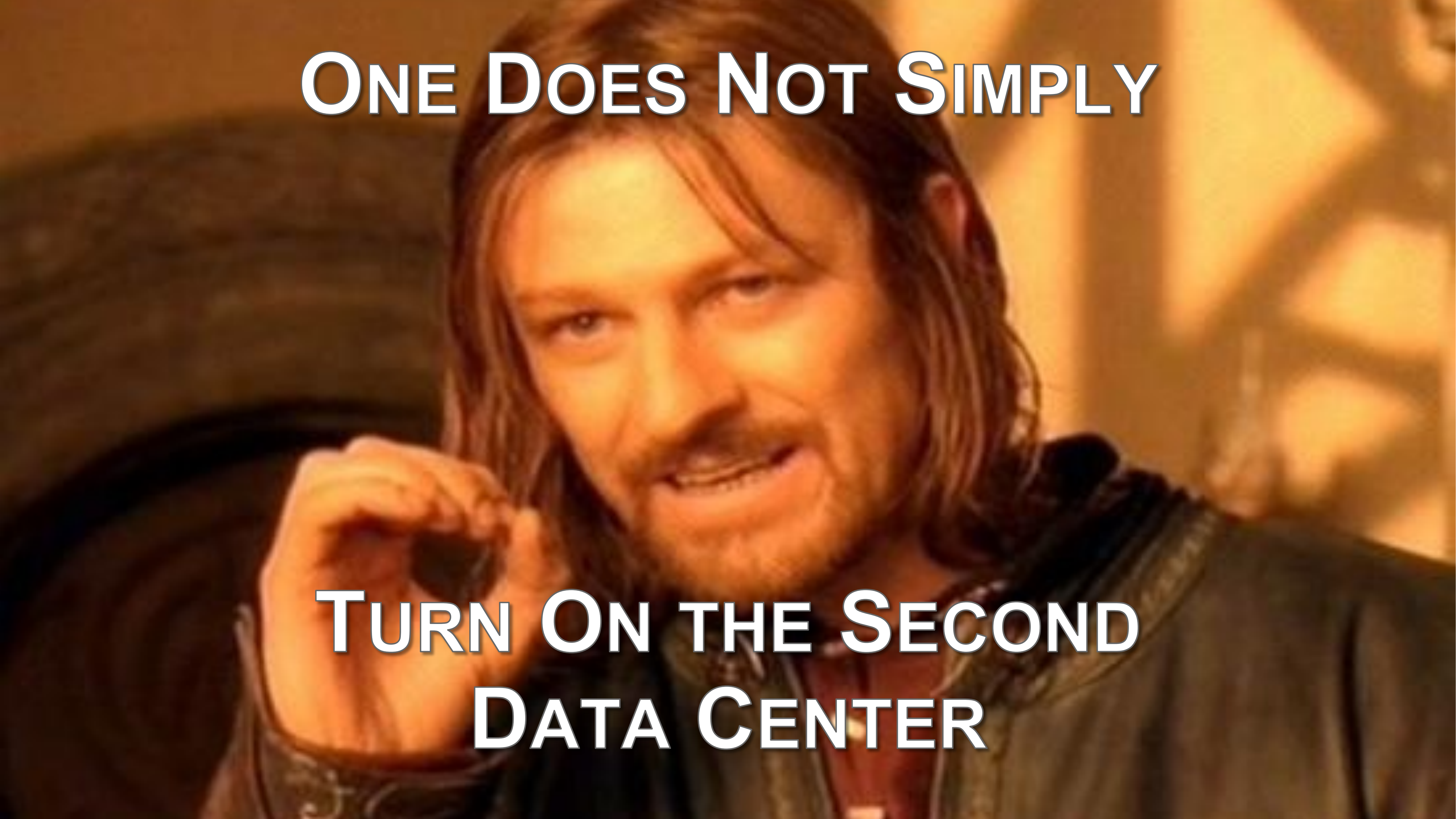
Provisioning the infrastructure

- Virtualize as much as you can
- Keep track of configuration changes
- Virtual appliances help

Questions to ask

- How are we doing storage failover?
- Is the failover manual or automatic?
- Have you tested it?





One Does Not Simply

Turn On the Second Data Center

Do You Know What Needs to Be Done?

Disaster Recover Plan

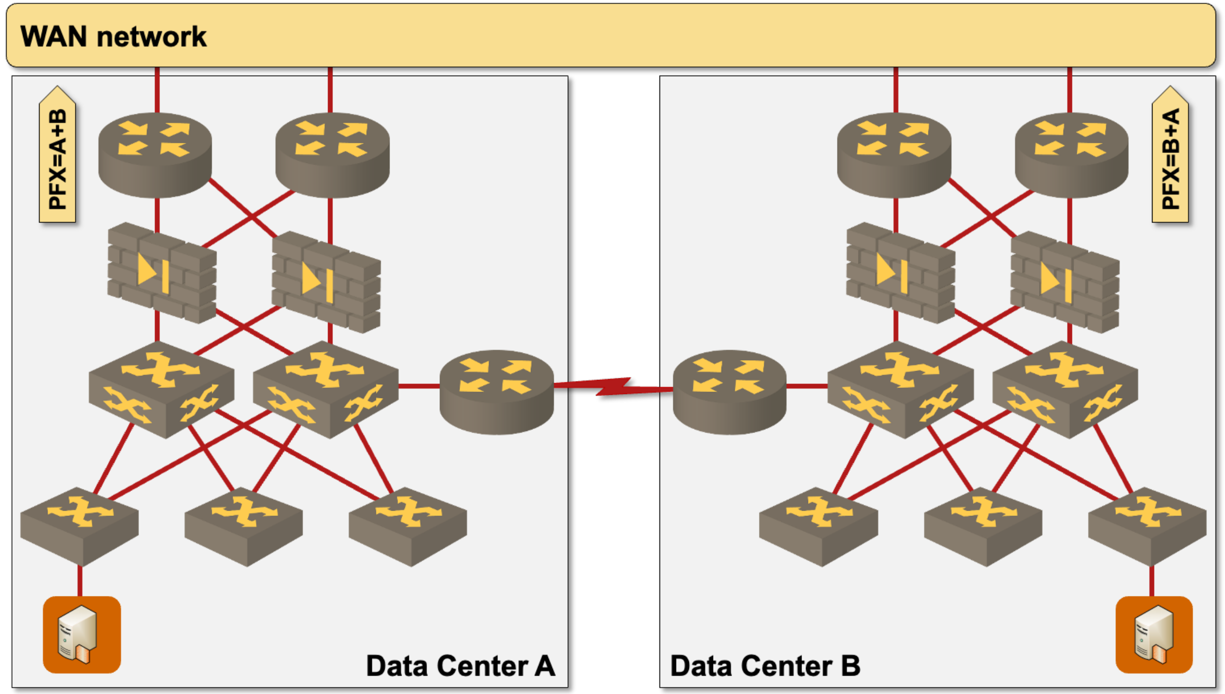
- You need it *before* the disaster strikes
- It is always out-of-date (but might still work)
- It probably won’t work if you haven’t tested it

Disaster Recovery Tests

- Hard to do without active/active infrastructure
- Pull the plug – migrating a few running VMs back and forth does not count

Don’t panic

- Follow the plan – no shortcuts
- It’s better to be a bit late than to mess up everything



What I Learned in the Last Fifteen Years

- Application workload migration is an orchestration problem
- Works best for properly designed applications
- SDDC and SDN are not a solution, but they make orchestration easier
- Disaster recovery is relatively easy to solve with orchestration tools
  ➔ Make network and services recovery part of the orchestrated process
- Disaster avoidance and workload load balancing are usually just a dream
- Stretched VLANs might be a great migration tool (disable after migration)
- An untested recovery procedure will likely fail when used for the first time

Questions?

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