SDC
STORAGE DEVELOPER CONFERENCE
SNIA ■ SANTA CLARA, 2016

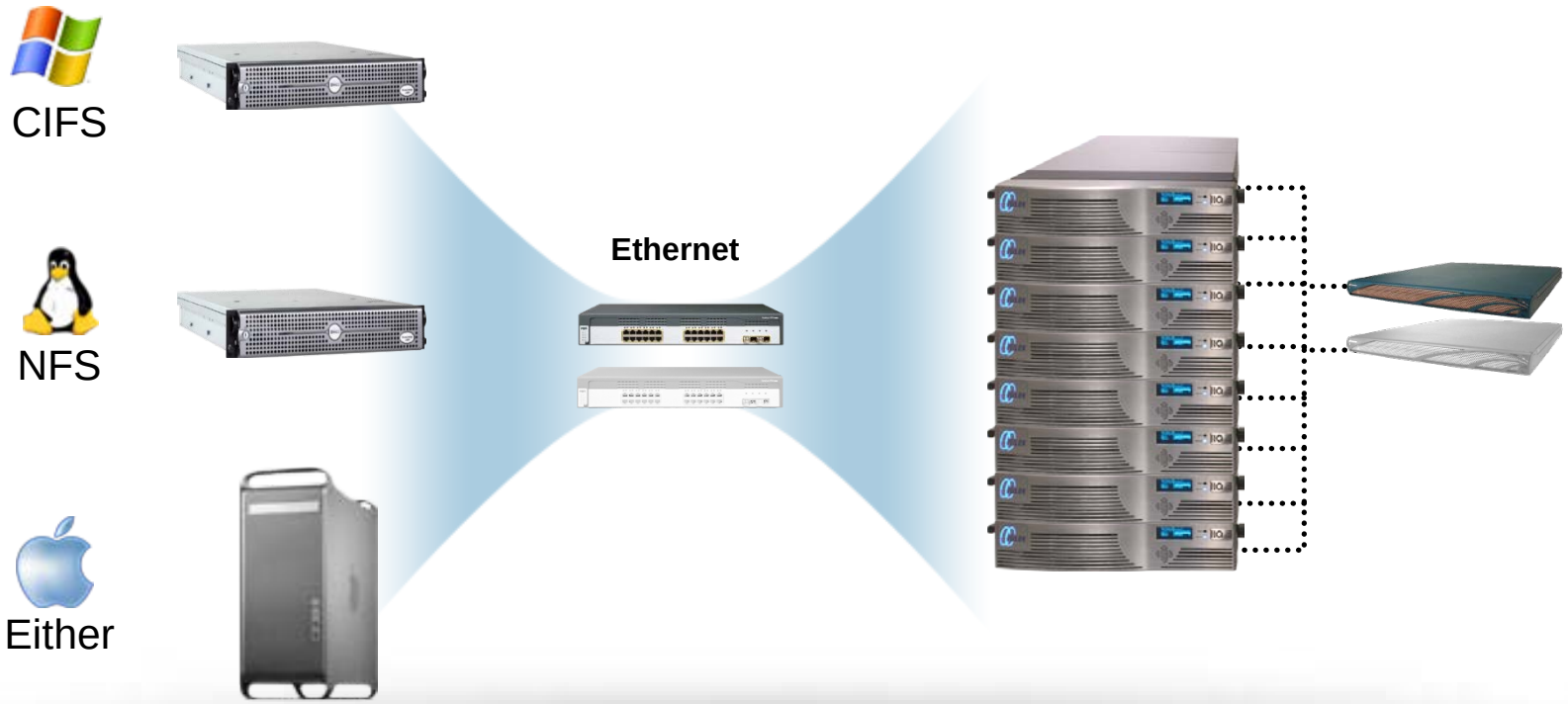
Active Directory Client Scaling Challenges

Marc VanHeyningen

Dell EMC Isilon

What is OneFS?

- A clustered scale out file system
  - Each node has a full view of the file system



OneFS and Active Directory

- OneFS cluster can join an AD domain
- We act as a client - a really big client distributed across hundreds of nodes
- Scaling problems in server operations are well-understood, but client scaling is less mature

Talking to AD: Naïve Approach

Have each OneFS node:

1. Find all domain controllers at your site
2. CLDAP ping them all, see who answers first
3. Use that DC for everything (“affinitize”)

Problems with this approach:

- All nodes pick the same DC, bullying it
- Ping time poor measure of server load (or even availability)

Talking to AD: DCLocator Option

- Choose a DC based on admin-defined weights
- Weights determine DC frequency linearly
  - DC with 2x the weight chosen 2x as often

Why we didn't use this approach:

- We're not Windows, don't have functions
- Weights static, don't respond to changing loads
- Weights outside control of storage admin

Talking to AD: What we do Today

- Maintain statistics on interactions with DCs
- Measure transaction latency (not ping time) as *t*
- Average latency for DC *i* is $\overline{t_i}$
- Create our own balancer:
  - $\forall i \ (w_i = \frac{1}{(\overline{t_i})^2})$
- A server 2x as fast will be selected 4x as often
- Each node chooses independently

Talking to AD: Configured Blacklists?

- Request: let storage admin set lists of “good” and/or “bad” domain controllers
- We have resisted this
  - Ugly hack
  - Hard to test
  - List will inevitably become stale
  - Nobody will check list

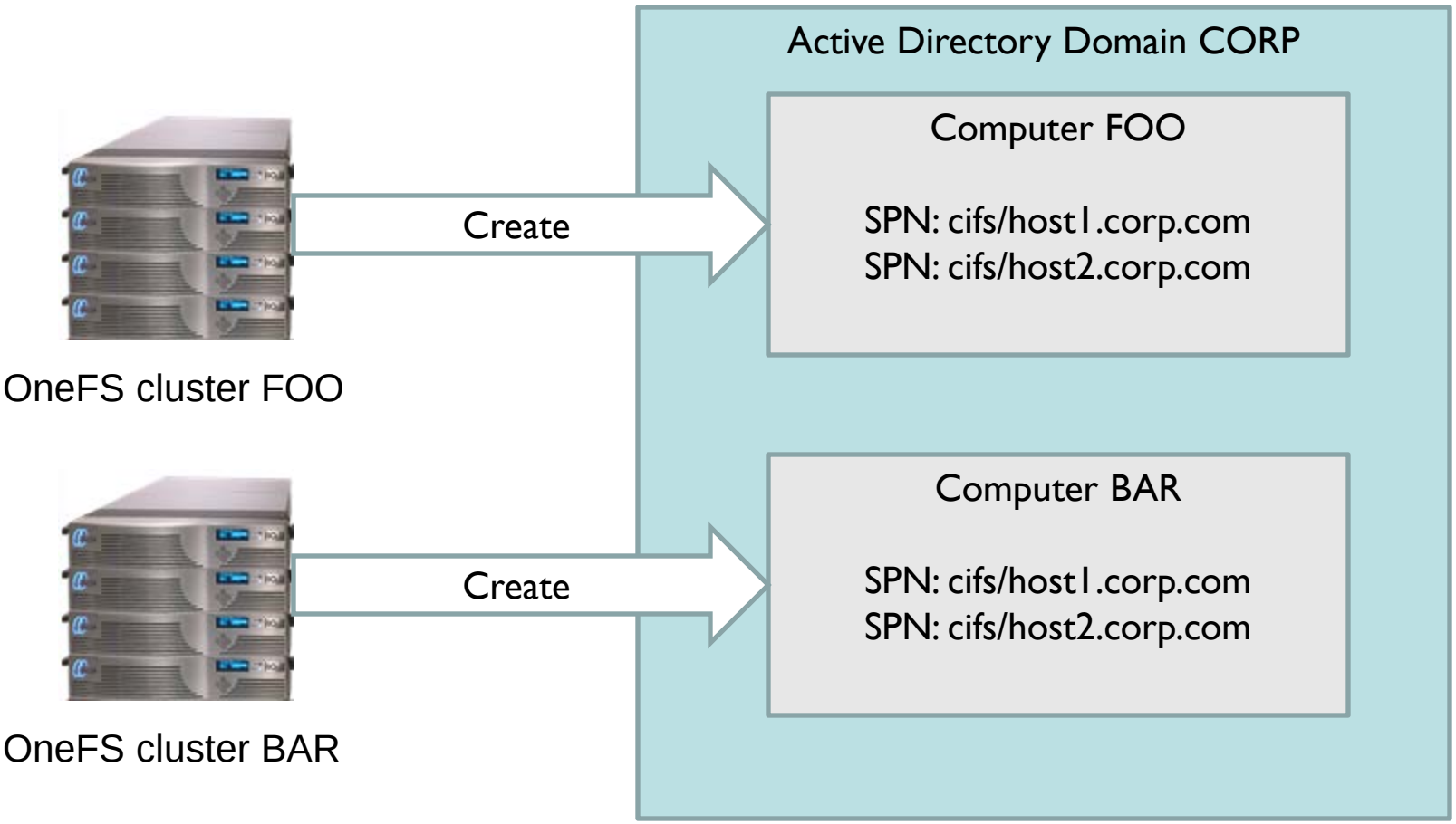
AD Machine Account Password Change

- Updating shared state across two different distributed systems with transactional semantics
- AD can take 15 minutes to update all DCs
- Old behavior: all nodes affinitize to a single DC for a while after a password change
- New behavior: if new password fails, try the old one (so we don't need transactional semantics)
- Plurality of AD issues involve password change

AD Machine Account: Testing

- Verification problems unlikely for customers
- Same cluster joins and un-joins domain hundreds of times for unit tests
  - Sometimes using same names
  - Sometimes using same names for different domains in same forest
  - Sometimes un-joining cleanly, sometimes not
  - Sometimes all this within <15 minutes

AD Machine Account: SPN Collision



How to Address Testing Challenges?

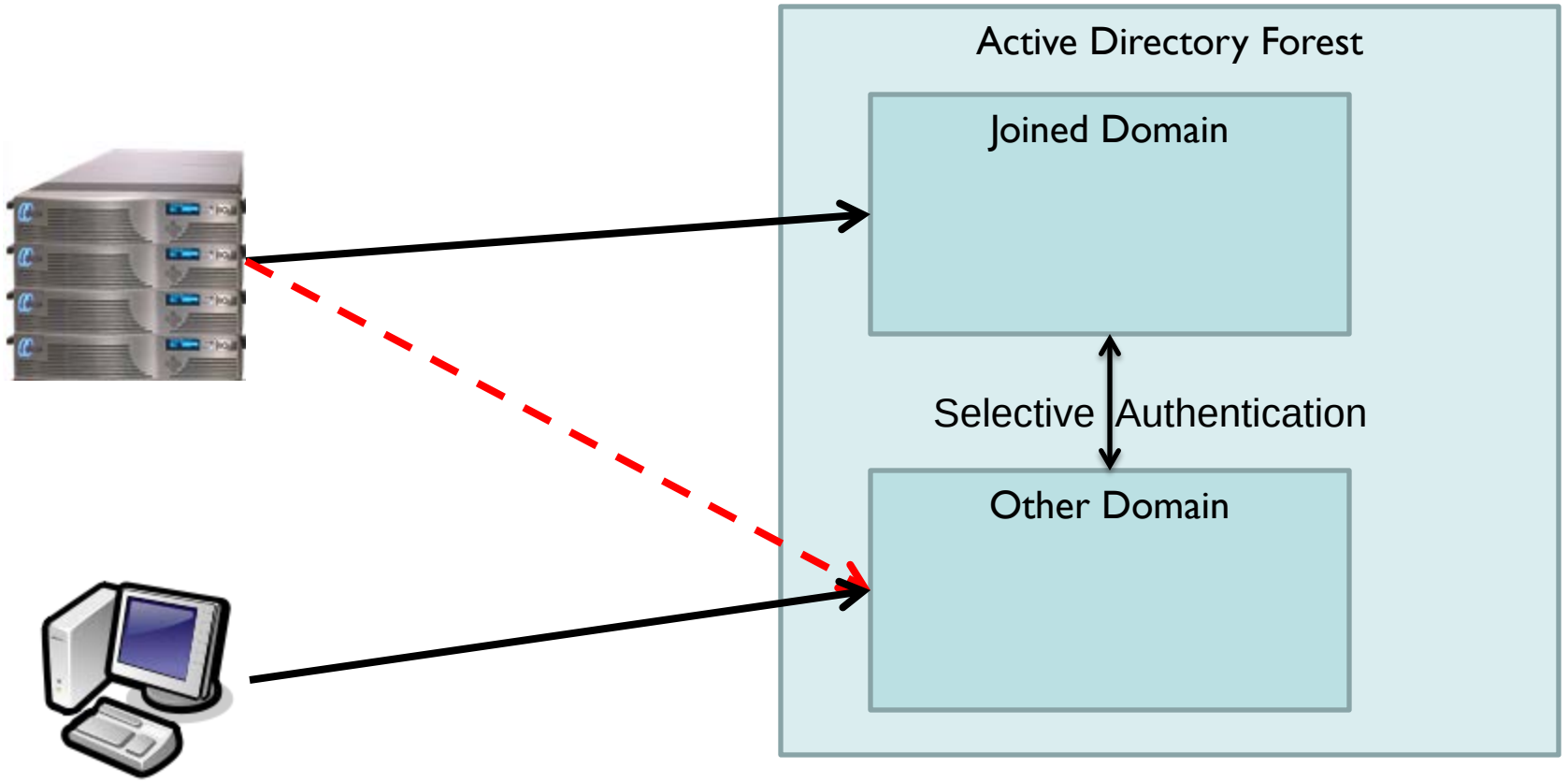
- Don’t re-use hostnames so much
- Write tools to detect things like SPN collisions
- Try to make sure tests clean up after themselves
- Active Directory as a Service (ADaaS)
  - Dynamically create a forest (on virtual hardware) for a test
  - Re-create it anew every time

AD: Read-only DC Challenges

- RODCs do not register in DNS for non-sited lookups
  - DCLocator falls back to NetBIOS, not us
- If only RODC is reachable we fail to join
  - We need a DC to learn our site
  - We need a site to find a RODC
- Crypto separation (different key, kvno, signed/unsigned bug in Kerberos)

https://technet.microsoft.com/en-us/library/ee522995(WS.10).aspx
https://support.microsoft.com/en-us/kb/2716037

AD: Selective Authentication



The client can authenticate to the other domain, but our cluster (the machine account) can not!

AD: Selective Authentication Challenges

- Cluster can’t look up group info
- PAC contains group info, but not all authentication methods include a PAC
- Workaround: get one (e.g. make PAM back-end to kinit so we get a PAC)
- Workaround: use LsaRpc calls instead of LDAP
- Preferred: add cluster account to selective authentication list (if customer can/will!)

Challenge: Technical

- Big clients, like big servers, must be smart and adaptable to perform well in a wide variety of situations
- The environment is constantly changing
  - Scale always increasing
  - Customer deployments can be unusual
  - Microsoft always inventing new wrinkles

Challenge: Customer Organization

- In large orgs, storage cluster likely administered by a different team than Active Directory
- Teams can face:
  - Understaffing
  - Conflicting priorities
  - Poor communication
  - Hate each other

Questions?