Multivendor Datacenters

Disaggregating network services
Who’s that dude?

• 2 years at Criteo
• Network Infrastructure Team Lead
Who’s that dude?

- 3 Missions:
  - Network architecture & design
  - Throwing out entrenched vendors
  - Automate myself out of a job
Transforming digital advertising into a personal experience.
About Criteo
About Criteo

8 Datacenters
7 PoPs
About Criteo

2,000 Network devices
10+ Hardware vendors
2.5 m reqs/s
1.1Bn internet users
What usually happens when you have one vendor

- Build RFP
- Consult
- Choose
What usually happens when you have one vendor

Choose your dream Hardware  Send money to the Vendor  Sales moves to the Bahamas
Meanwhile…

- Please upgrade
- Known bug, goto 1
- Yeah, that’s a feature
How do you solve this problem

• Story time! How we built PA4
  • Criteo’s 2\textsuperscript{nd} HPC cluster
  • 40G non-blocking L3 Clos fabric
How do you solve this problem

- Design:
  - Proprietary features do not save lives
  - Find the smallest set of features possible
    - IEEE and IETF are your friends
How do you solve this problem

• L3 Clos Fabrics
• One feature: BGP
• All the rest is useless
How do you solve this problem

• During the RFP process try to qualify at least 2 vendors
  • People will say you’re crazy:
  • Twice the work, twice the problems
How do you solve this problem

• No vendor is used to this
• Most will work with you
• Some will spread rumours of unfairness
How do you solve this problem

• Be Transparent
• Brief your vendors on the qualification results
• Share anonymized scores
How do you solve this problem

<table>
<thead>
<tr>
<th>feature</th>
<th>sc</th>
<th>score</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP</td>
<td>1</td>
<td>2</td>
<td>-1 (green)</td>
</tr>
<tr>
<td>LLDP</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IPV4</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IPV6</td>
<td>2</td>
<td>2</td>
<td>-2</td>
</tr>
<tr>
<td>SPAN (Span)</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Flow</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>L2 overlay</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>API</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Code Stability</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Code maturity</td>
<td>1</td>
<td>2</td>
<td>-1</td>
</tr>
<tr>
<td>Auto Install</td>
<td>2</td>
<td>2</td>
<td>-1 (drop)</td>
</tr>
<tr>
<td>Debugs</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Documentation</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>User Base</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>User Community</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>CPU (x86)</td>
<td>2</td>
<td>2</td>
<td>-1 (PPC)</td>
</tr>
<tr>
<td>Redundant PS</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1U</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Port capacity</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Strong points</td>
<td>Deep Buffers</td>
<td>support code maturity</td>
<td>Price</td>
</tr>
<tr>
<td>Weak points</td>
<td>Price Documentation user base only China BGP code 09/2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>21</td>
<td>36</td>
</tr>
</tbody>
</table>
How do you solve this problem

• Chose the 2 best that fit your requirements
  • When you have time qualify more

• Give the power back to your purchasing department
Automation is key

- The only way this can work is through automation
- If you use the CLI you will never scale
- CLI is the last resort when stuff is broken
• Find an abstraction layer that fits you
  • Ansible var files

• Build a DC factory
Automation is key

- dc.yml
- dcgen.py
- BOM.csv
- Asset Mgmt
- Cablers
- Vendor
- IPAM
- CMDB
- Cabling.csv
Automation is key

IPAM
CMDB
Asset Mgmt

Play
Ansible

Variable file
Automation is key

Play → Ansible → Vendor specific template → Config file → Variable file
Automation is key

This is the only thing to develop to add a new vendor
Automation is key

Netcompare
https://github.com/criteo/netcompare
The Future

- Disaggregate, Distribute, Decentralize!
- LBs, FWs, Switches
- One OS – Multiple HW vendors
- One template
The Future

• Onboard code:
  • Counter streamers (kill SNMP)
    • Self healing devices
  • Kill Centralization (LB/FW)
    • On demand WAN TE
Join Us!

• Criteo is recruiting!
• Want to work with us on the future of DC networking?
  • [http://labs.criteo.com](http://labs.criteo.com)
  • [rndrecruitment@criteo.com](mailto:rndrecruitment@criteo.com)
• Come see us at the break!
Thank you.

Questions:

m.moriniaux@criteo.com