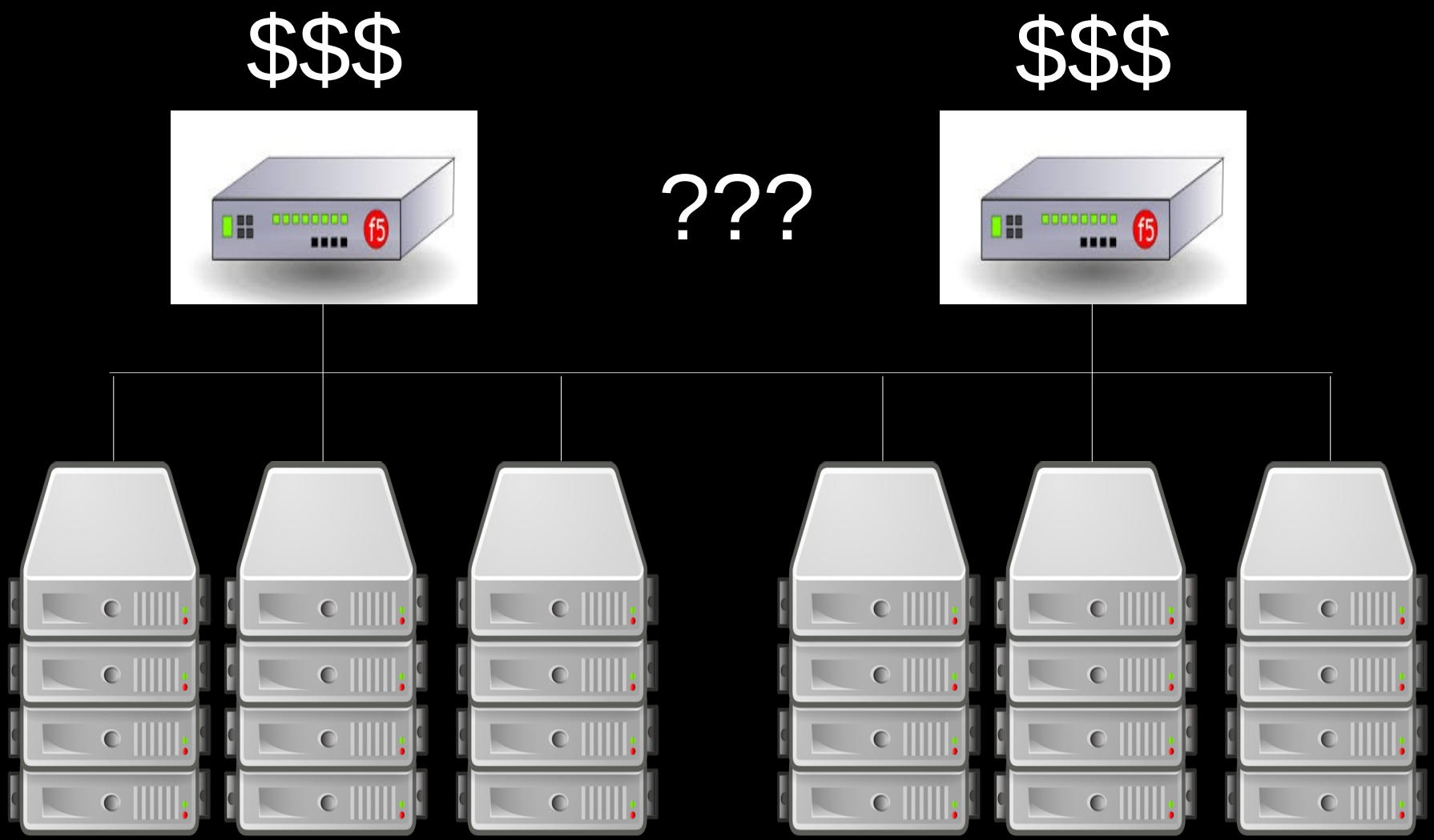
Softwarový LB

$$$


f5



$vendor

- 1 box nahradil LVS cluster
- Centralizovaný (SPoF)
- Nutnost použít víc IP (nepodporují a/a)
- Bugy
- Děsivý support
- Cena

$vendor

- Centralizovaný (SPoF)
- Nutnost použít víc IP (nepodporují a/a)
- Bugy
- Děsivý support
- Cena

- Konkurence ještě horší :(

$vendor

- Centralizovaný (SPoF)
- Nutnost použít víc IP (nepodporují a/a)
- Bugy
- Děsivý support
- Cena

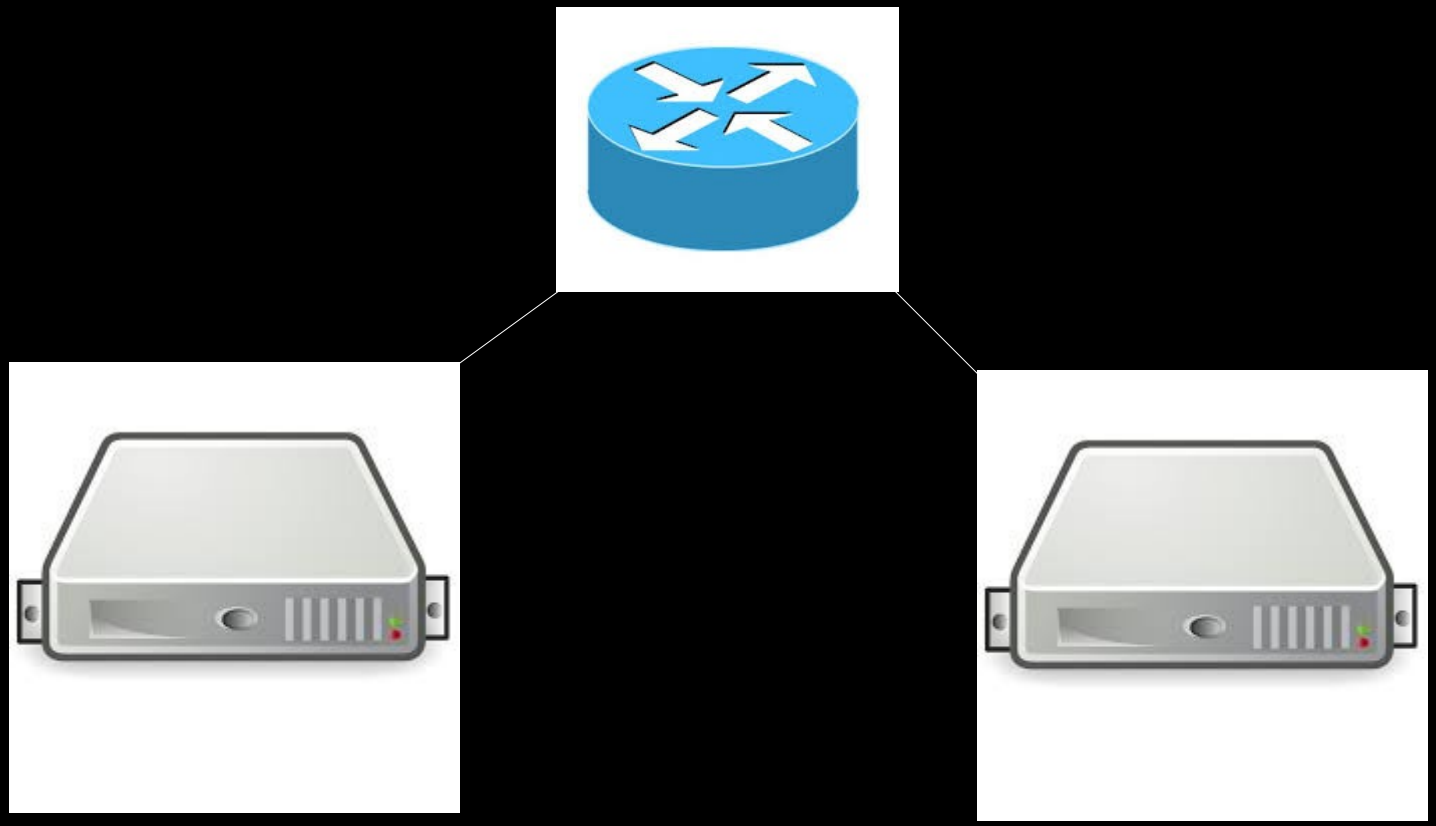
- Konkurence ještě horší :(
- Nová naděje – AviNetworks. Cena :(

Vlastní LB!

- Distribuovaný (bez SPoF)
- Snadné škálování
- Pouze základní features
- Levnější
- Využít stávající HW (servery, síť)
- Opensource
- IPv6

ECMP!

- Umožňuje anycast (stejná IP, víc strojů)
- Funguje nám na DNS!



ECMP! Ale...

- 1/10 gbps?
- Requests/second?
- Velikost routovacích tabulek?
- Routovací protokol (BGP/OSPF)?
- Počet sessions routovacího protokolu?

ECMP! Ale...

- 1/10 gbps?
- Requests/second?
- Velikost routovacích tabulek?
- Routovací protokol (BGP/OSPF)?
- Počet sessions routovacího protokolu?

- Hashovací algoritmus?!?

Hashovací algoritmus?

- hash(src_ip⊕dst_ip⊕src_port⊕dst_port)

Hashovací algoritmus?

- hash(src_ip⊕dst_ip⊕src_port⊕dst_port)
- Linka = Hash % počet_linek

Hashovací algoritmus?

- hash(src_ip⊕dst_ip⊕src_port⊕dst_port)
- Linka = Hash % počet_linek
- Consistent hashing! Drahé routery :(

Hashovací algoritmus?

- hash(src_ip⊕dst_ip⊕src_port⊕dst_port)
- Linka = Hash % počet_linek
- Consistent hashing! Drahé routery :(
- Resilient hashing! Moc nefunguje :(
  - IPv6 :(
  - Neřeší přidání node, pouze odebrání

Dvojvrstvý LB!

- První vrstva LVS + consistent hashing
- Druhá vrstva ukončí TCP

Dvojvrstvý LB!

- První vrstva LVS + consistent hashing
  - IPIP/GRE tunel
  - Neřeší návratový provoz
- Druhá vrstva ukončí TCP

- Synchronizace konfigurace?
- Health checky?

Aktuální stav u nás

- Cca 100mbit/s provozu
- Statické nody (bez škálování)
- OSPF
- Dynamická konfigurace obou vrstev
- Neřešíme výpadky 2. vrstvy

Závěr

- Levnější než HW LB
- Dynamické škálování
- Anycast + ECMP – žádný SPoF!
- Všechno na L3 - žádný NAT!
- Možnost opravovat chyby
- DDoS ochrana
- Šetření IP adres (resp dynamický failover)

Závěr (cont.)

- Založený na opensource
  - LVS, nginx, bird
  - Docker/LXC
- Navázané spojení na backend (proxy)
- Možná přepis LVS části do userspace?
- Možná zveřejníme ?

Ď za pozornost

- Toto PDF – https://safar.sk/lb.pdf
- Facebook LB - https://www.usenix.org/conference/srecon15europe/program/presentation/shuff
- Google LB - https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/44824.pdf