Booking.com

# Distributed load balancing

## Real case example using open source on commodity hardware

Pavlos Parissis | LinuxConf Berlin 2016


# Pavlos Parissis

Senior UNIX System Administrator

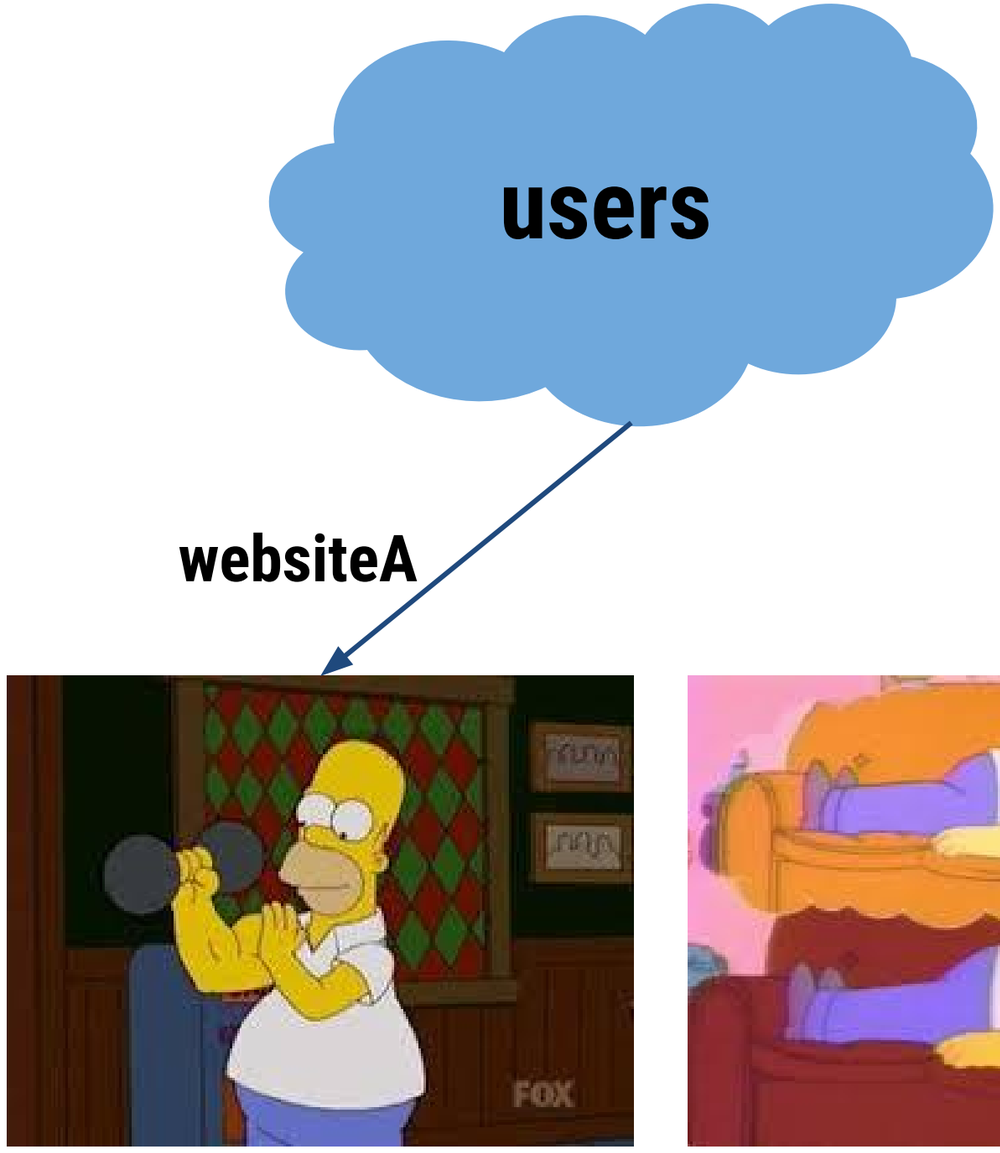
Global Traffic Distribution

pavlos.parissis@booking.com

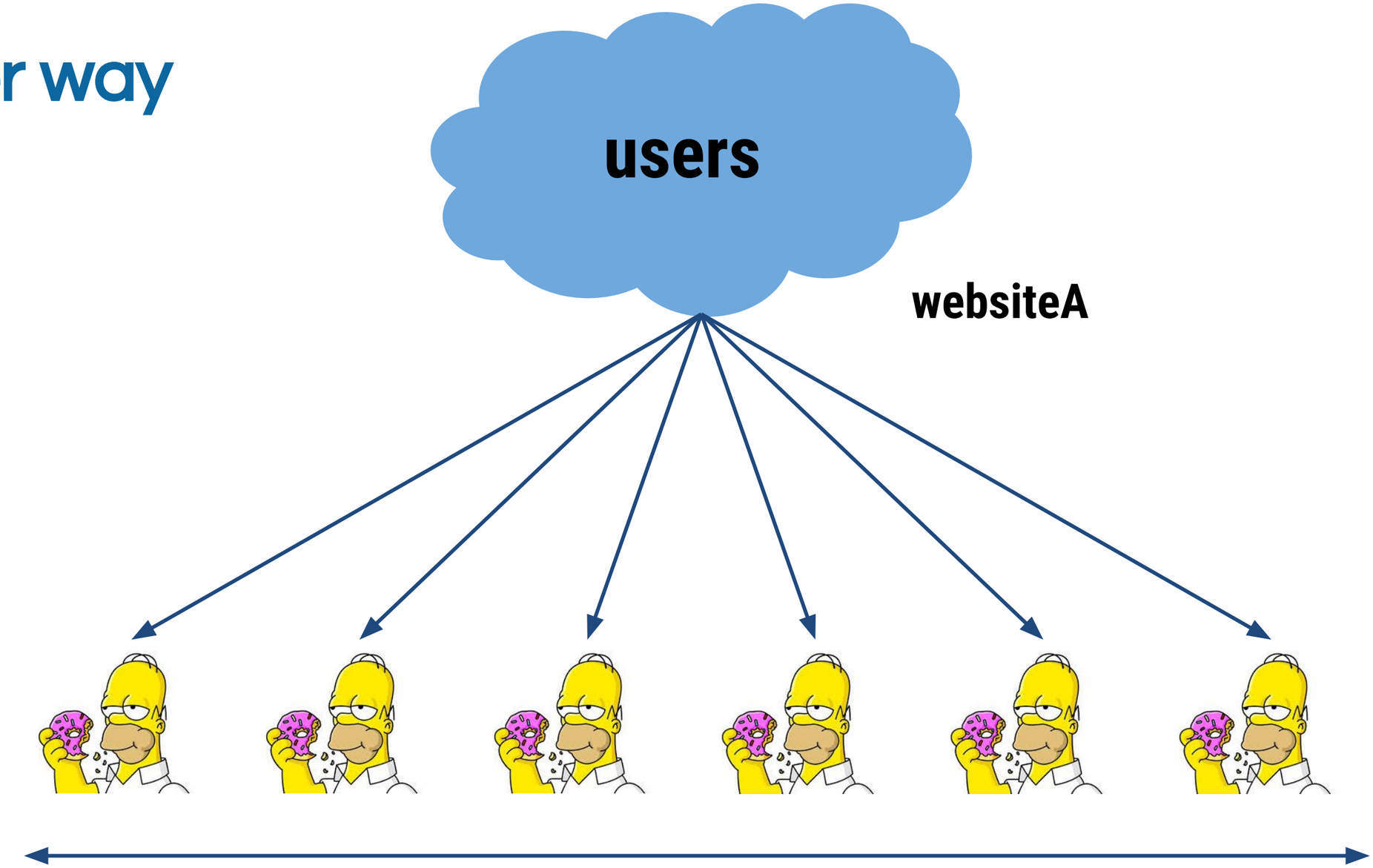
Booking.com

# The traditional way

- Scales only vertically
- Single point of failure
- Choke point for (D)DOS
- Very expensive



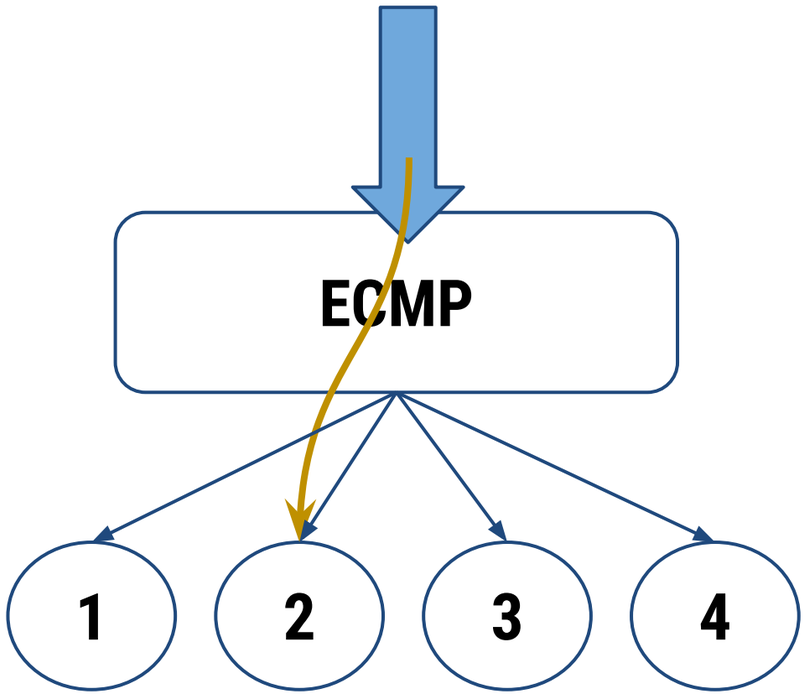
# A better way



# How to get there

- Equal-Cost Multi-Pathing routing
- Anycast network address scheme
- Bird Internet Routing Daemon
- A healthchecker for Anycasted services
- HAProxy Layer4-7 load balancer

# Equal-Cost Multi-Pathing routing



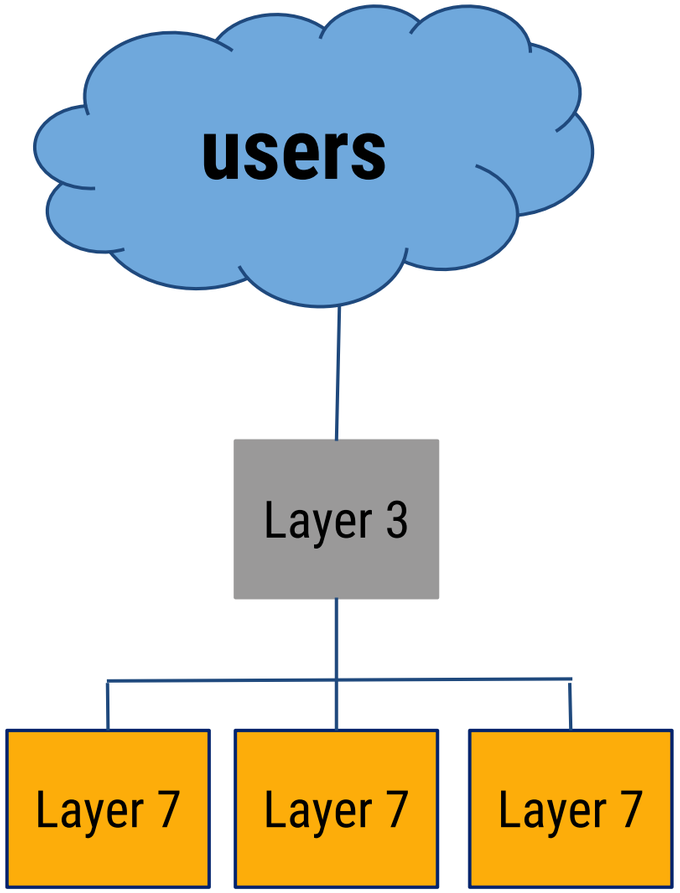

| Destination IP | Next hop |
|---|---|
| 5.56.17.220/32 | node1 |
| 5.56.17.220/32 | node2 |
| 5.56.17.220/32 | node3 |
| 5.56.17.220/32 | node4 |

- Nodes are distributed across multiple networks
- Preserves source and destination addresses
- Cheapest form of balancing
- Load balancing at wire-speed
- Adding/removing a path reshuffles flows

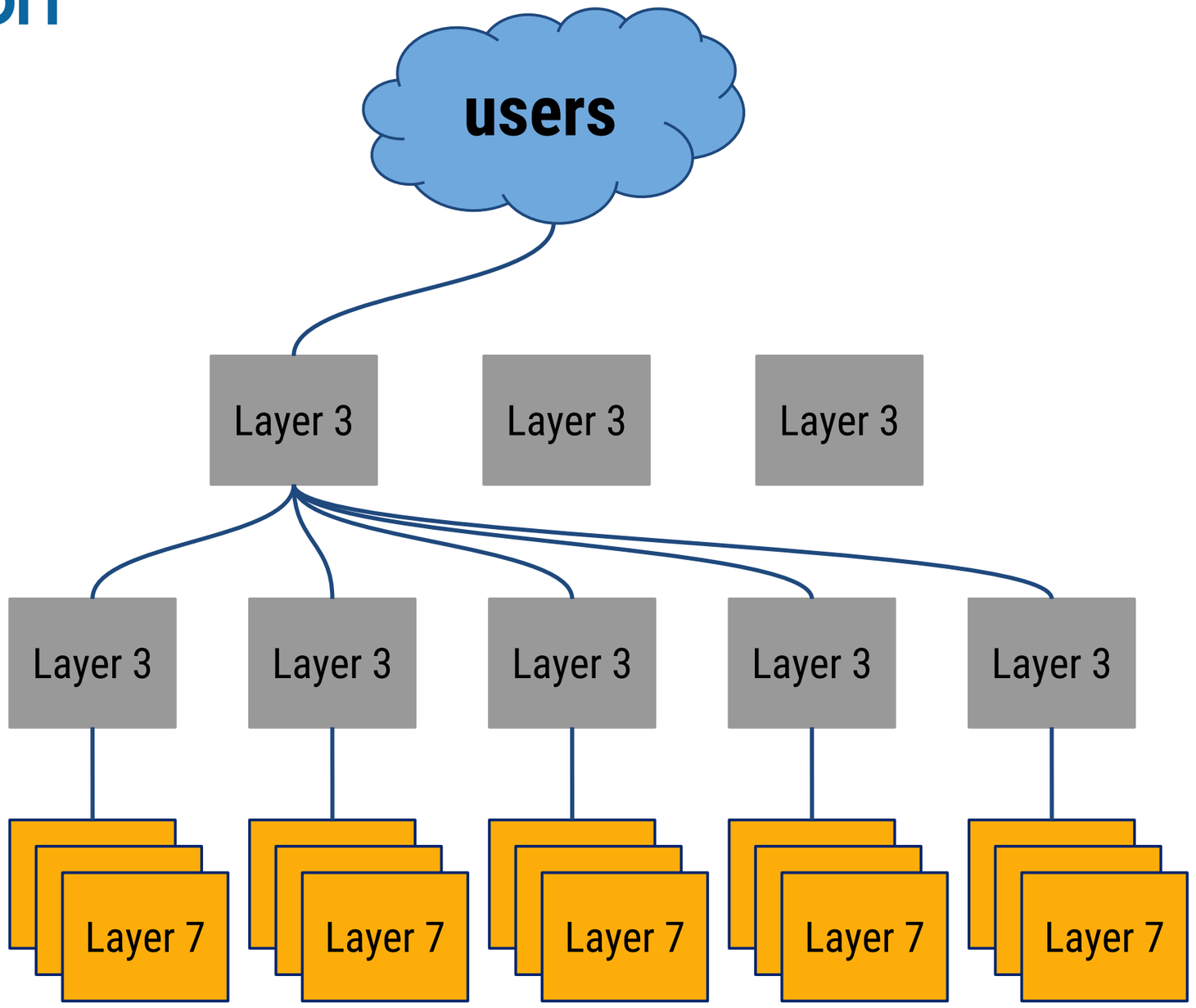
# Equal-Cost Multi-Pathing

Tier 1 Load balancer

Tier 2 Load balancer



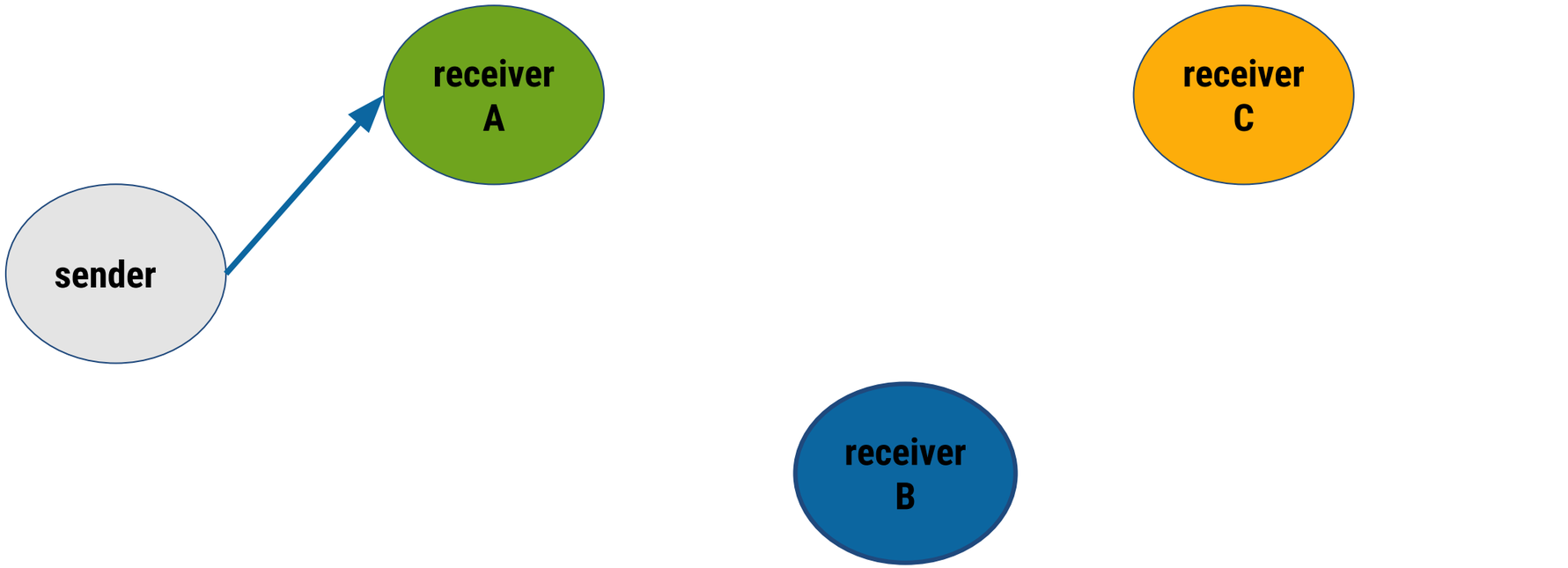
# 2-Tier setup in production



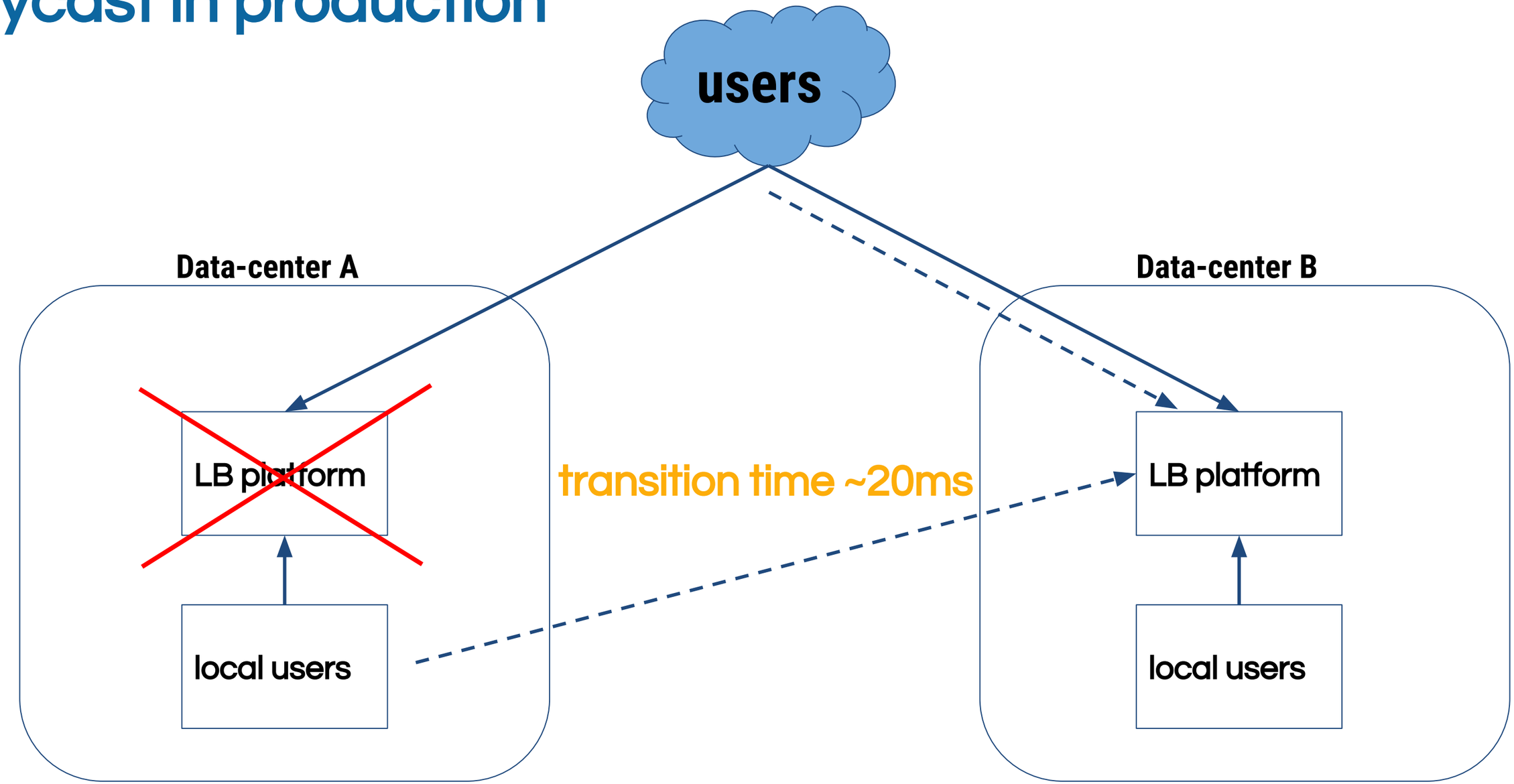
# Benefits of 2-Tier setup

- Horizontally scalable
- Scaling and managing each tier independently
- Single device becomes less critical

# Anycast network address scheme



# Anycast in production



# Benefits of Anycast in production

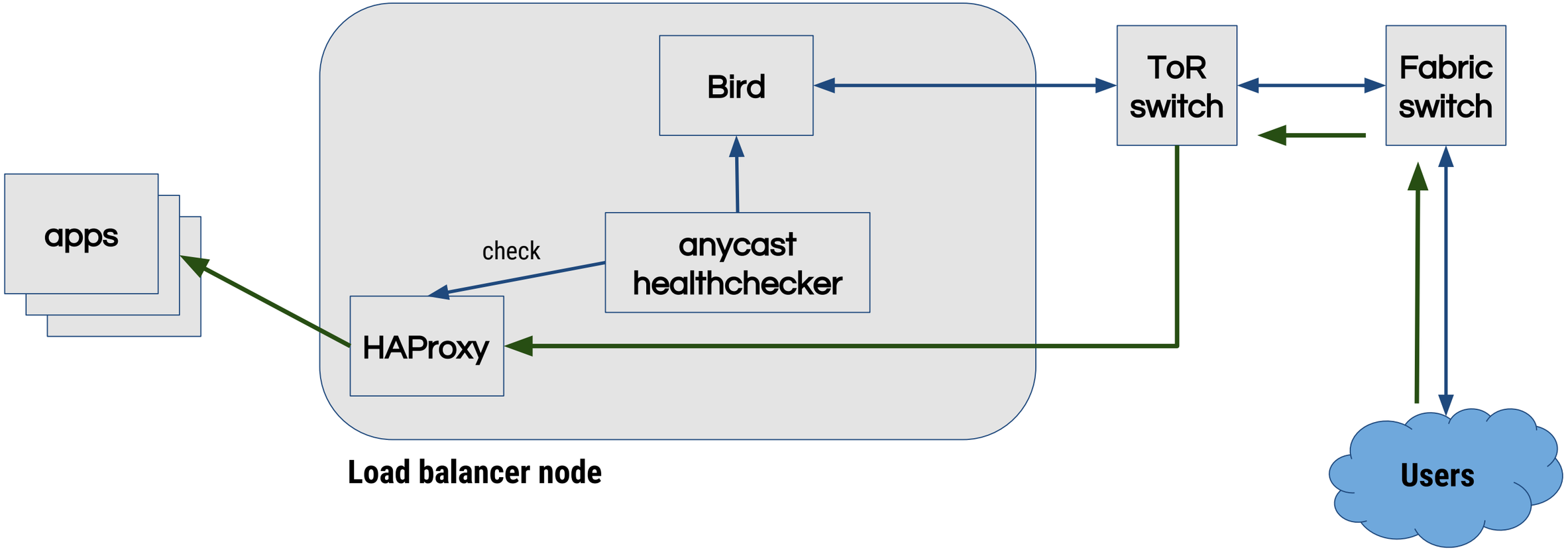
- Network detect failures within 1.2secs ( BFD protocol helps a lot)
- Switches traffic to other location within 1sec
- Reduces network distance which lowers response time
- Provides a very fast and without manual intervention fail-over which improves service reliability
- **Works** for **TCP** protocol


# Dive into details

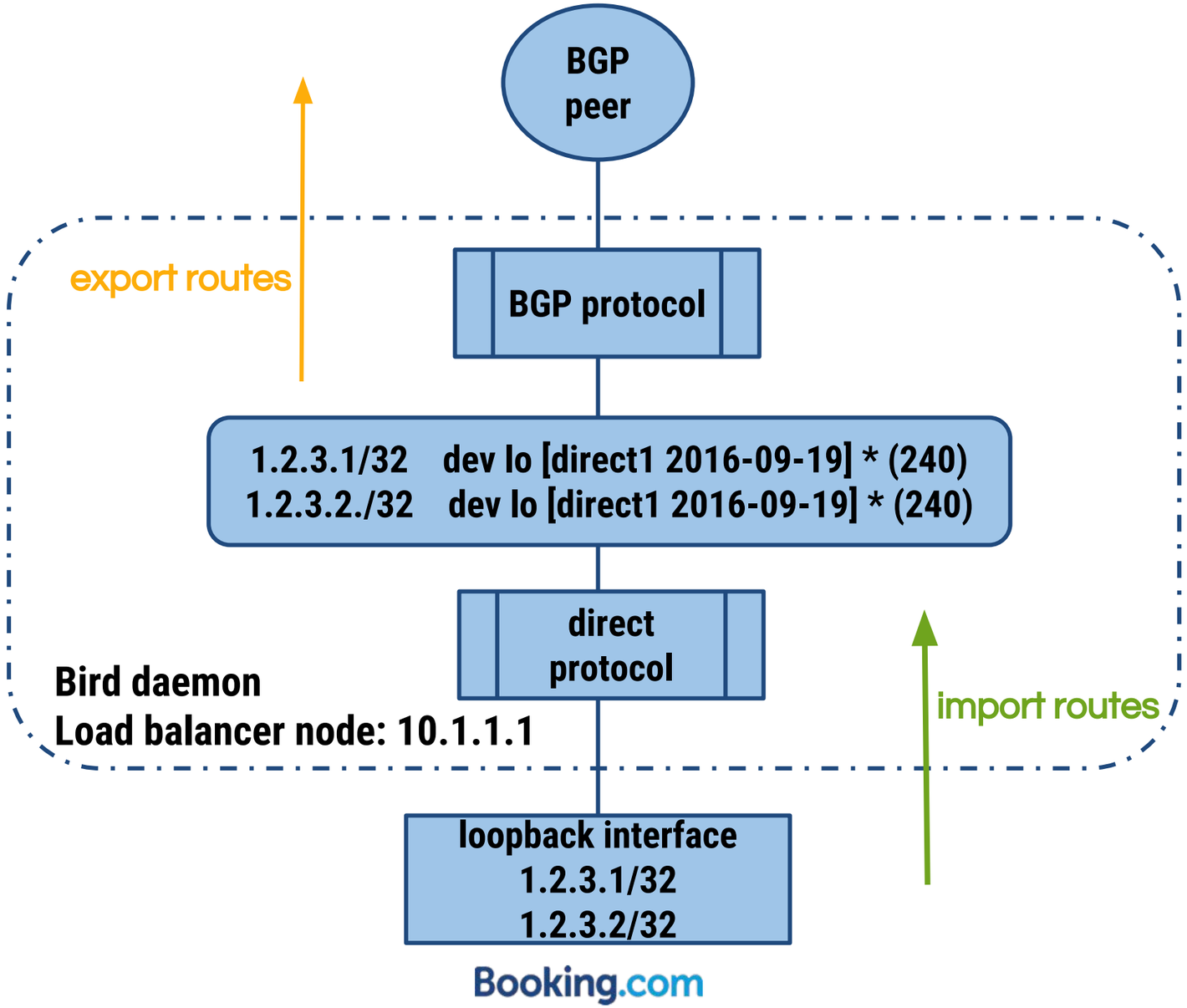
- Bird Internet Routing daemon
- A healthchecker for anycasted services
- HAProxy Layer4-7 load balancer

Booking.com

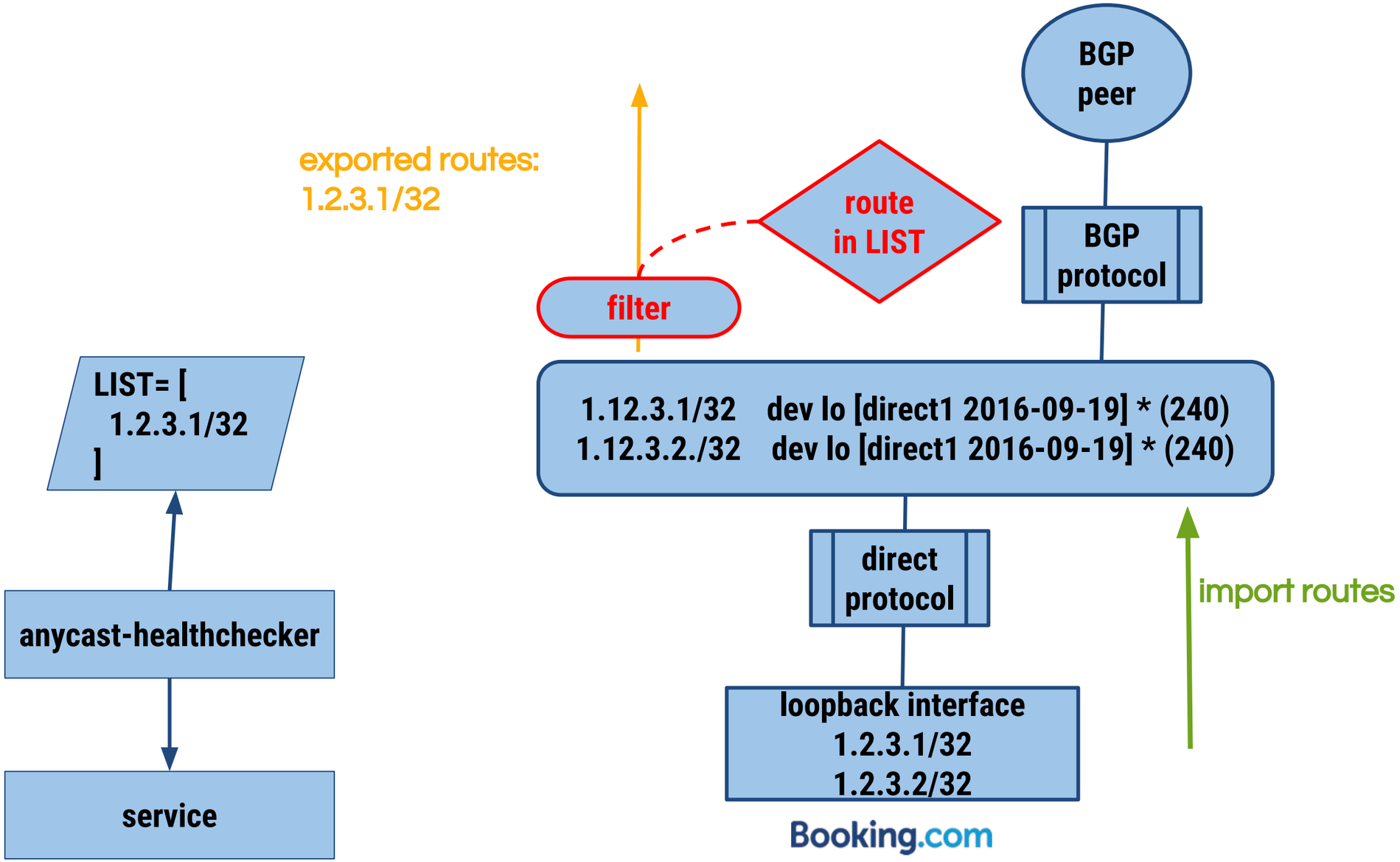
# How it works


Bird
ToR switch
Fabric switch
apps
check
anycast healthchecker
HAProxy
Load balancer node
Users

# How Bird advertise routes



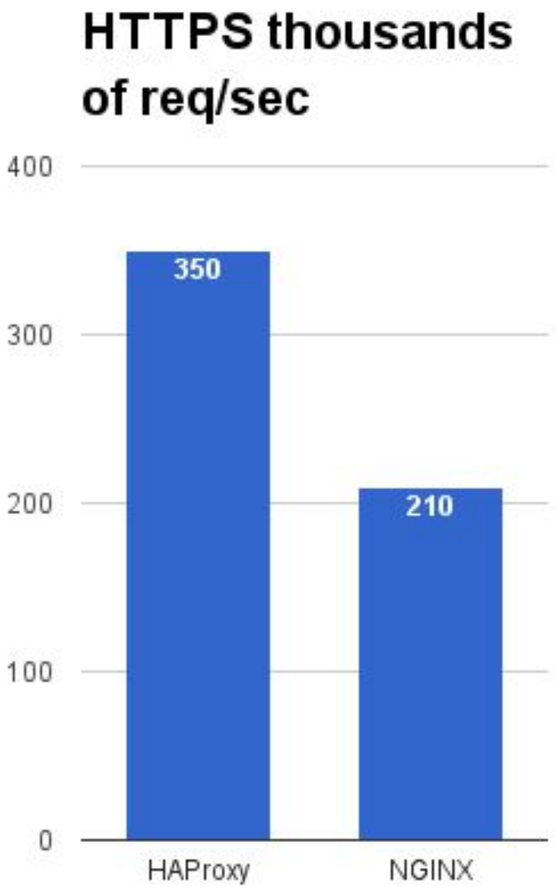
# Filtering routes for unhealthy services

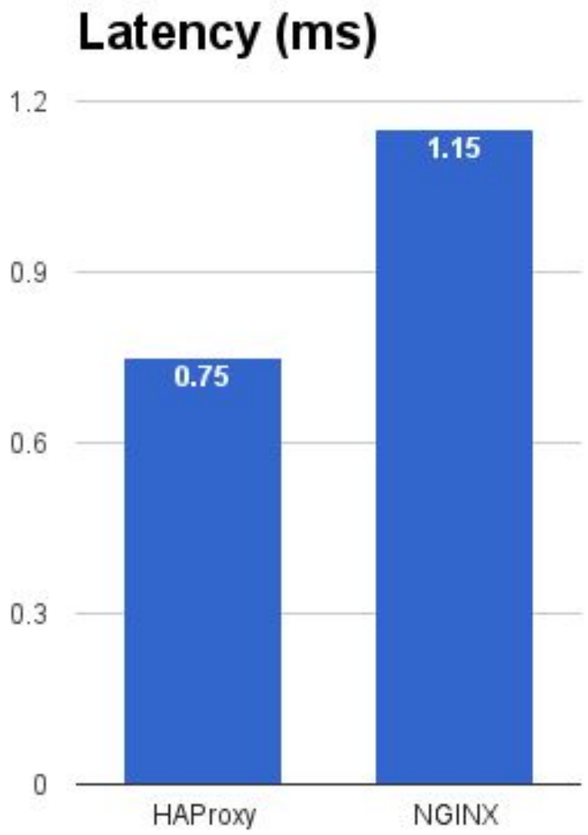


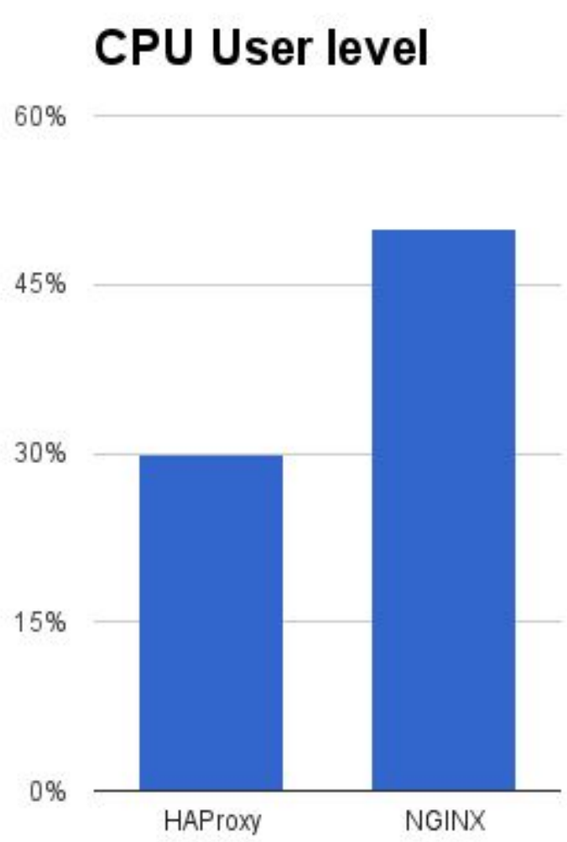
# HAProxy load balancer

- Highly configurable
- Rock solid
- Excellent support
- Supports Lua
- Faster than Nginx in our setup, benchmark yours

# HAProxy load balancer performance



# Software and Hardware we use

- Arista switches
- 2 x 10GbE interfaces on servers and 160GbE (4 x 40GbE) on switches
- Bird Internet Routing Daemon http://bird.network.cz
- HAProxy load balancer http://www.haproxy.org
- https://github.com/unixsurfer/anycast_healthchecker
- https://github.com/unixsurfer/haproxystats
- https://github.com/unixsurfer/haproxyadmin
- HP discrete/blade servers


We are hiring
Site Reliability Engineers
https://workingatbooking.com
Booking.com