Horizontally scalable firewalls

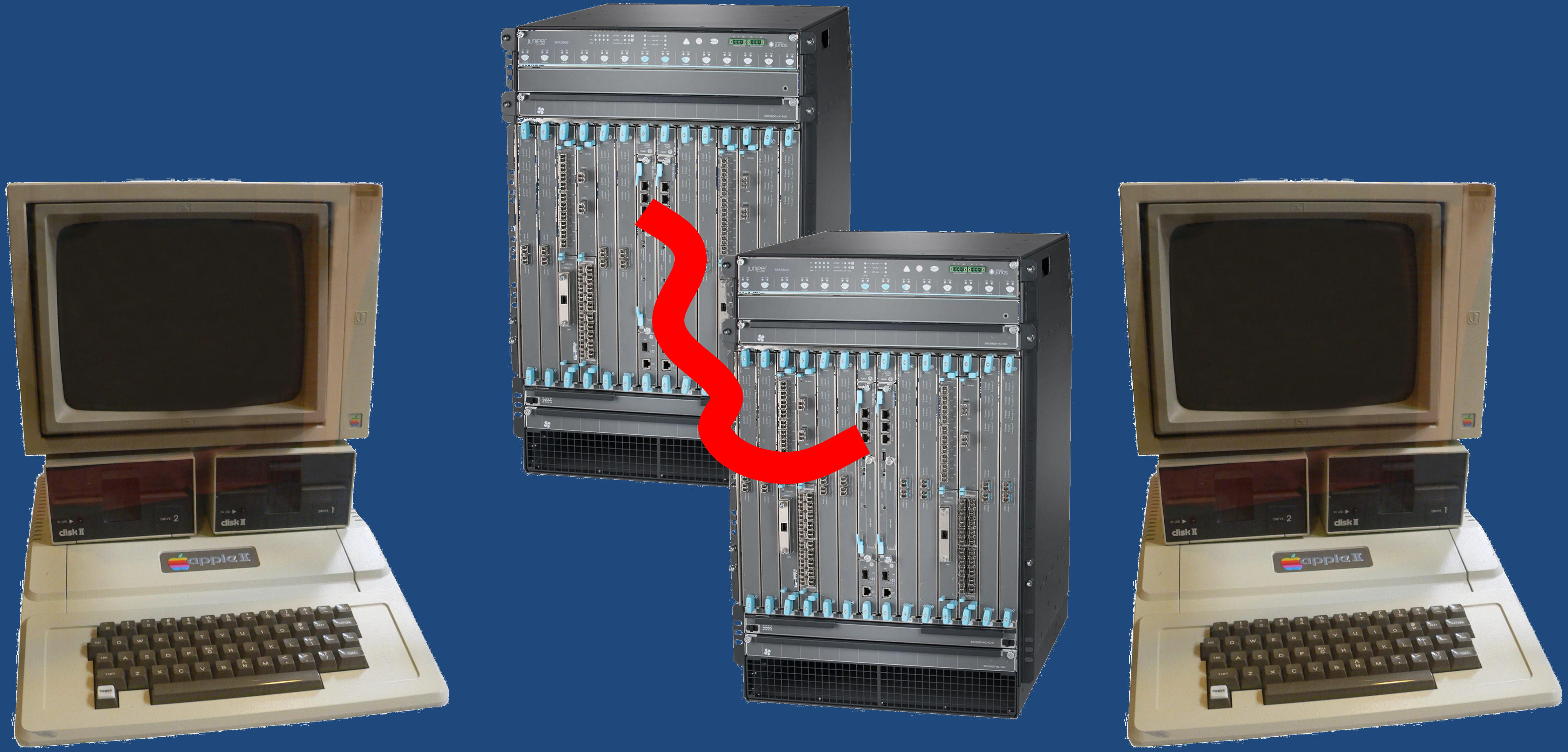
Ralf Ertzinger | 10.11.2020

Booking.com

The Problem

disk II
disk II
apple II

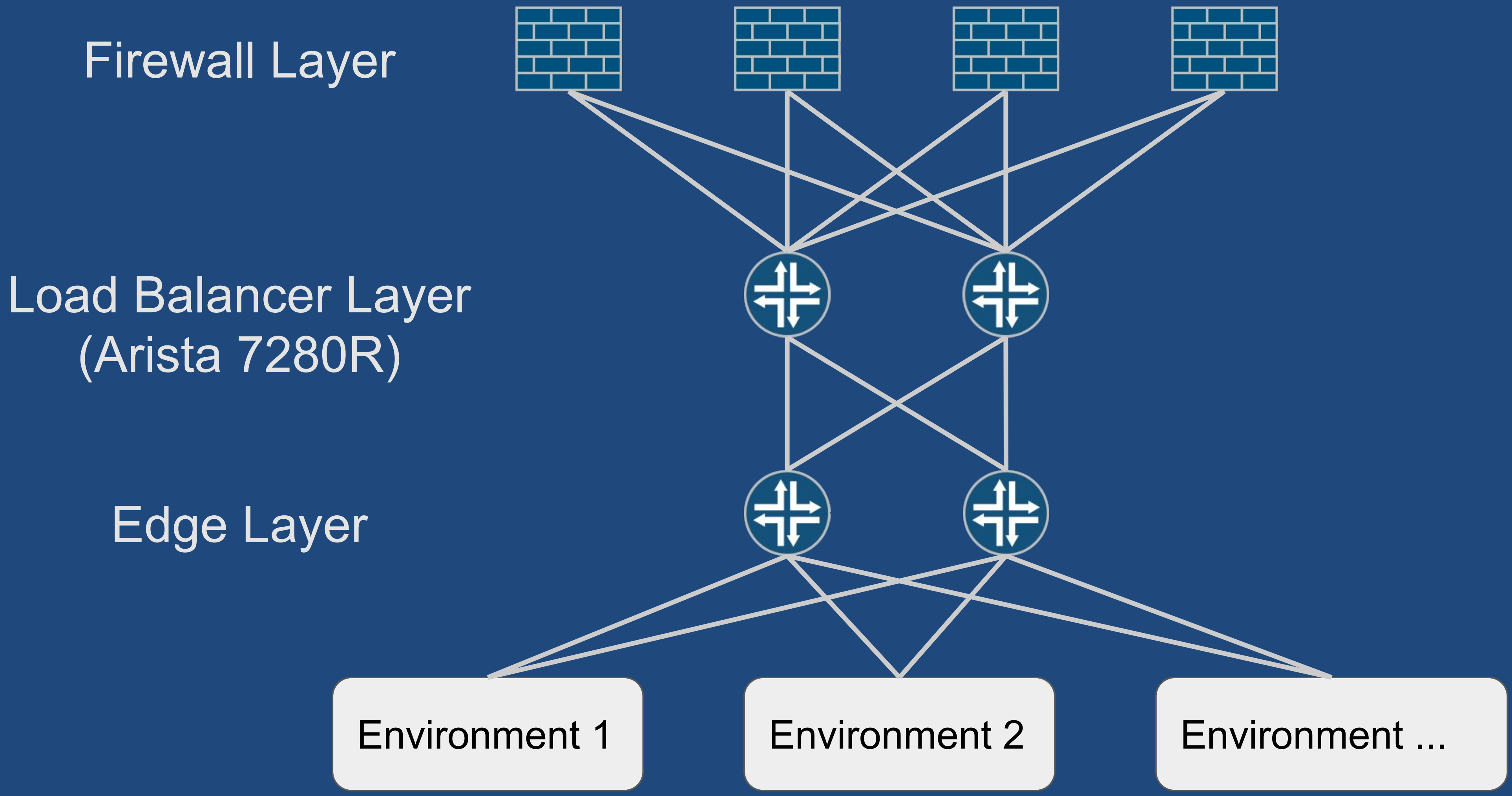
disk II
disk II
apple II



disk II
disk II
apple II

disk II
disk II
apple II

- Occasional disconnects are OK
- No NAT
- No VPN
- Have automatic firewall rule management



Firewall Layer

Load Balancer Layer
(Arista 7280R)

Edge Layer

Environment 1

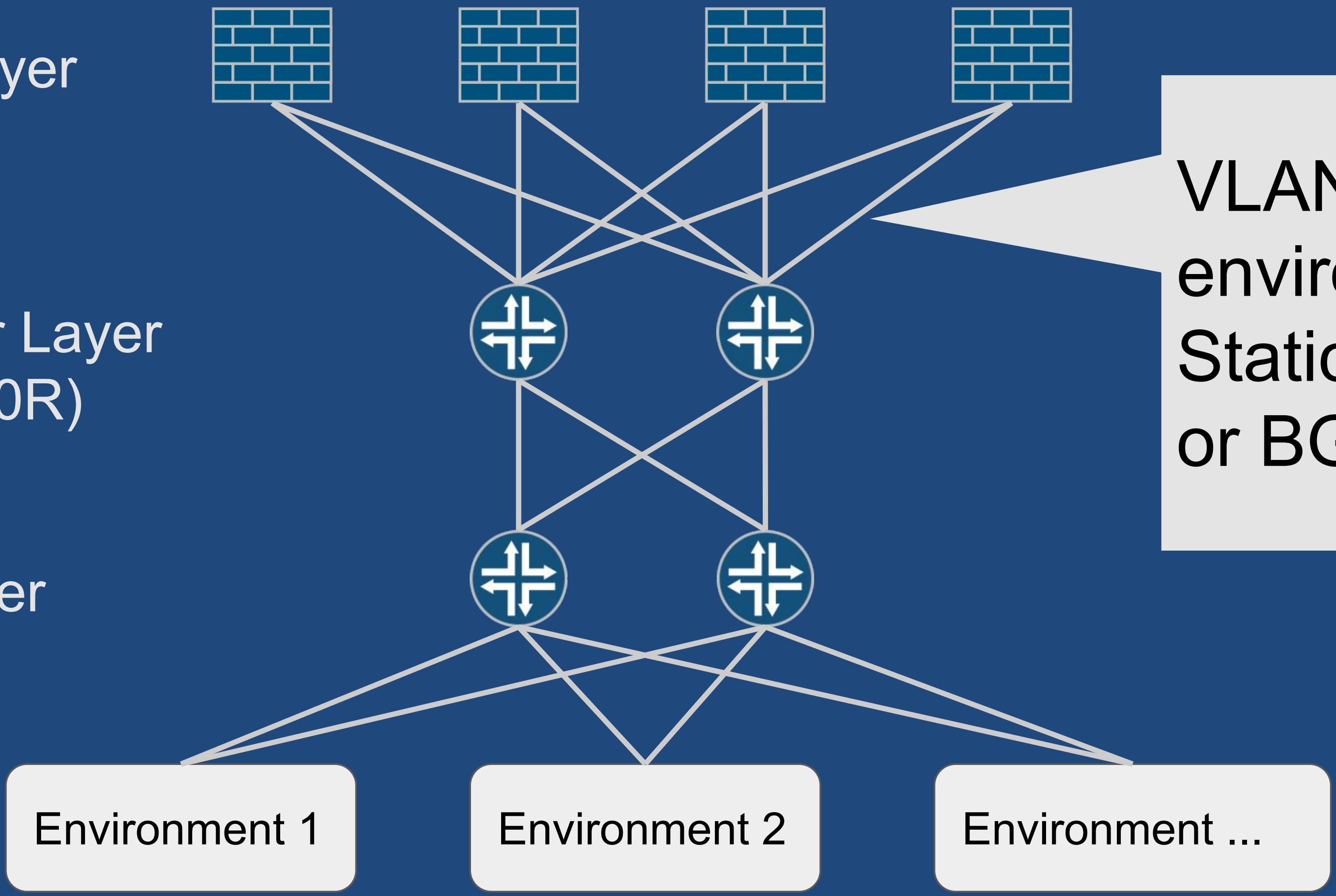
Environment 2

Environment ...

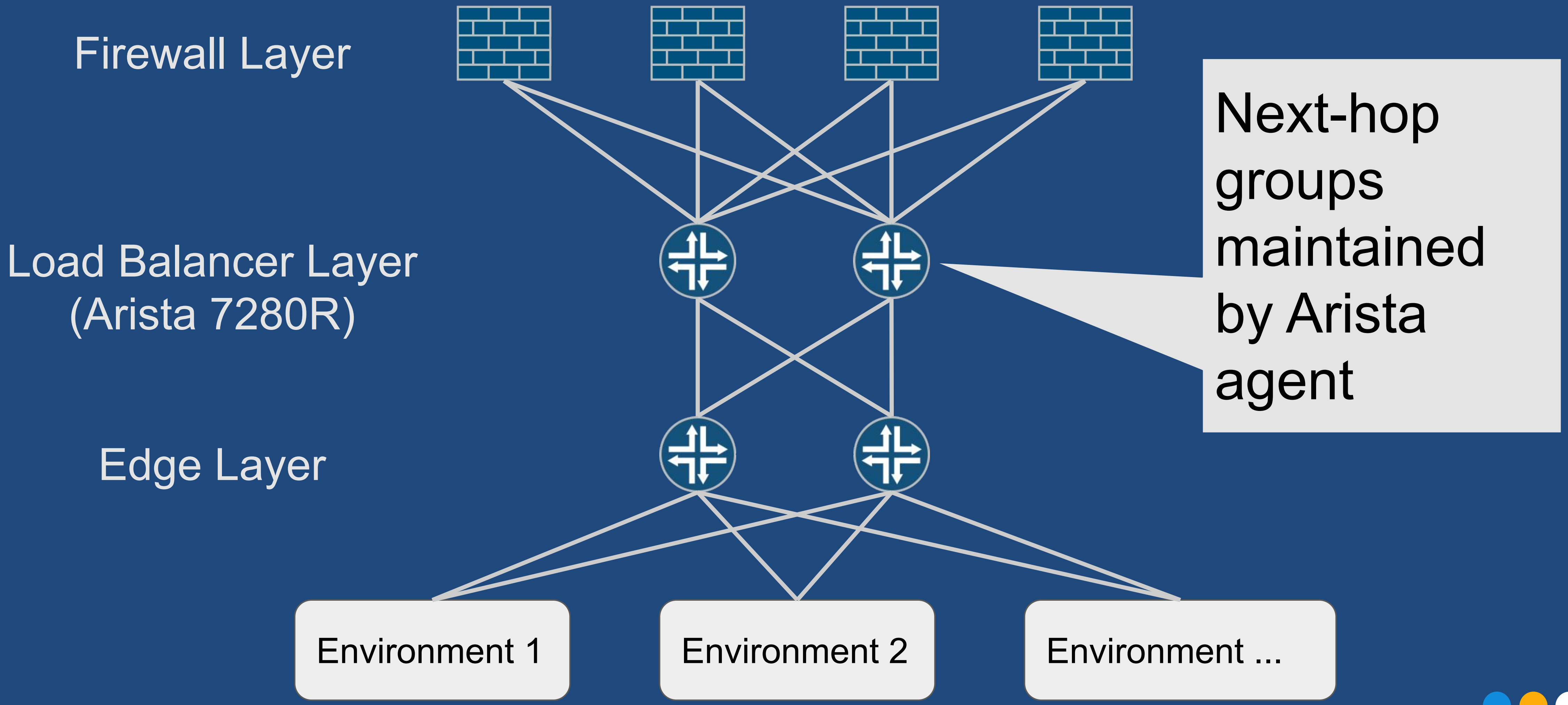
Firewall Layer

Load Balancer Layer
(Arista 7280R)

Edge Layer



VLAN per environment. Static routes or BGP



Firewall Layer

Load Balancer Layer
(Arista 7280R)

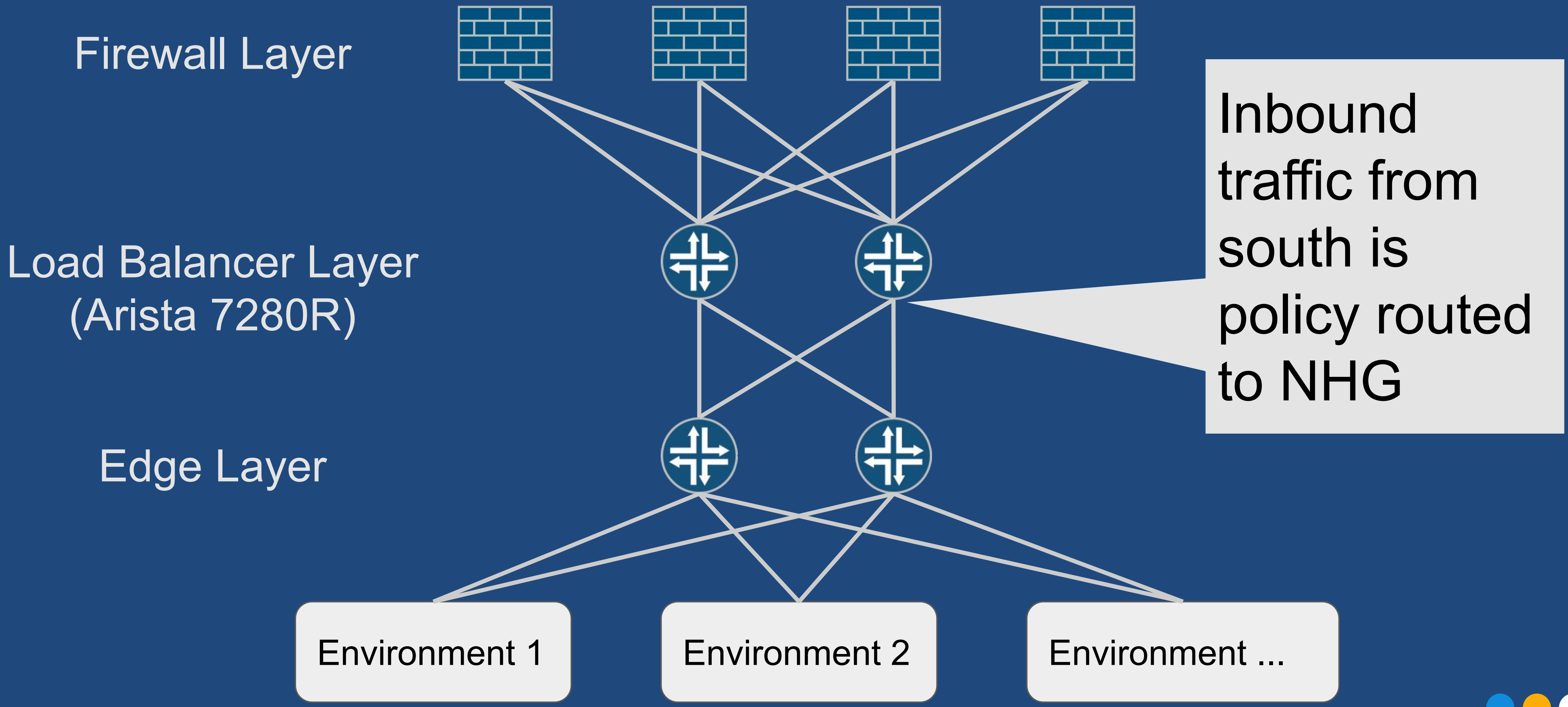
Edge Layer

Environment 1

Environment 2

Environment ...

Next-hop groups maintained by Arista agent



Firewall Layer

Load Balancer Layer
(Arista 7280R)

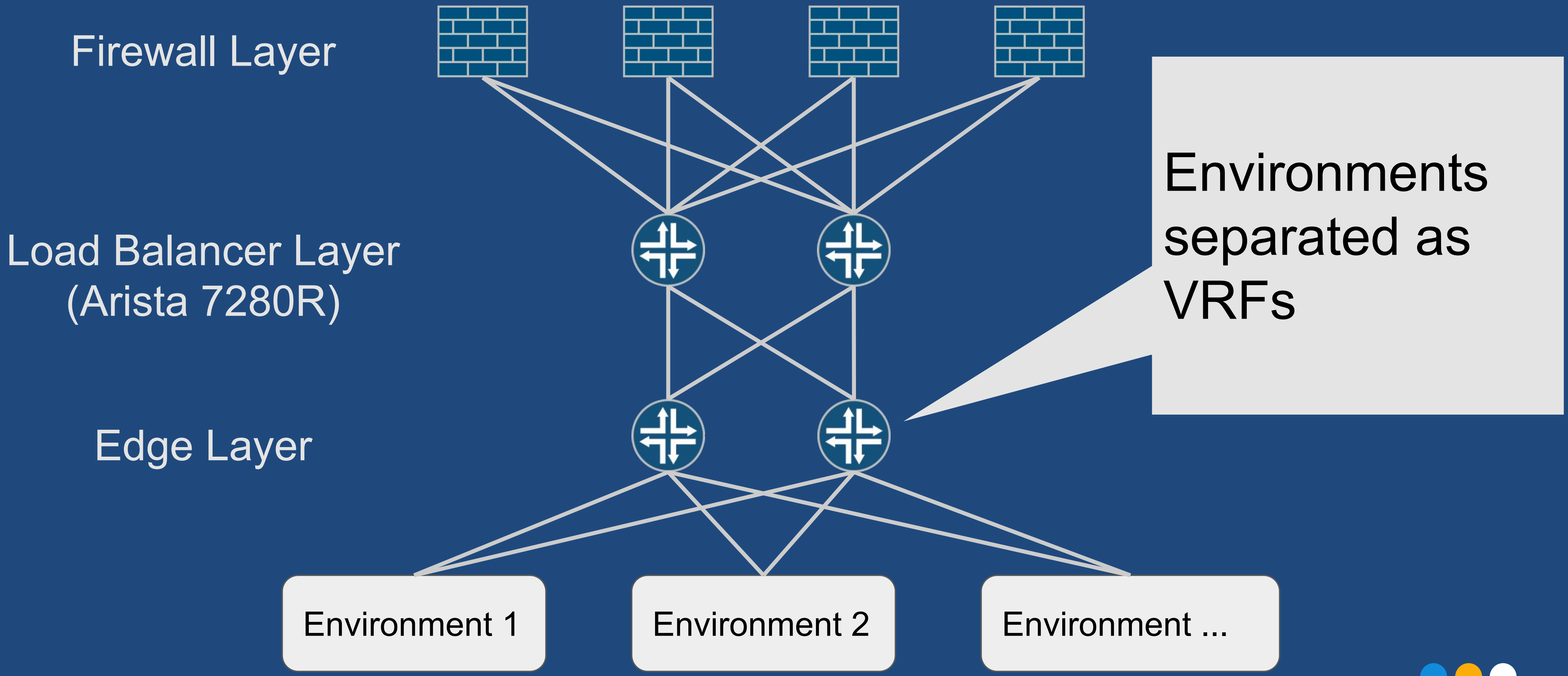
Edge Layer

Inbound traffic from south is policy routed to NHG

Environment 1

Environment 2

Environment ...



Firewall Layer

Load Balancer Layer
(Arista 7280R)

Edge Layer

Environment 1

Environment 2

Environment ...

Environments separated as VRFs

Consistent hashing

```
load-balance policies
   load-balance sand profile SYM
      no fields mac
      no fields mpls
      fields ipv4 symmetric-ip
      fields ipv6 symmetric-ip
      no fields l4

no ip load-sharing sand fields ingress-interface
ip load-sharing sand hash preset 1
port-channel load-balance sand profile SYM
```

Moar consistent hashing

```
/*
 * Disable Consistent Hashing for IPv4 Unicast packets
 * This is needed to disable consistent hashing between the multiple levels of FECs
 * as this effectively disables symmetric hashing for nexthop groups (in 4.21).
 * Arista promised a command to do this in configuration
 */
if( ARGC > 0 ) {
   int unitId = 0;
   int i;
   for( i=0; ARGV[0][i] != '\0'; ++i )
      unitId = unitId * 10 + ( ARGV[0][i] - '0' );
   bshell( unitId, "mod IHP_CONSISTENT_HASHING_PROGRAM_SEL_TCAM 0 4 VALID=0" );
}
```

Next-hop groups

```
daemon ResilientNexthopsCORP
   exec
/usr/local/bin/ResilientNexthops
   option 0 value 10.70.100.1
   option 1 value 10.70.100.3
   option 2 value 10.70.100.5
   option 3 value 10.70.100.7
   option GROUP_NAME value CORP
   no shutdown
```

```
daemon ResilientNexthopsPROD
   exec
/usr/local/bin/ResilientNexthops
   option 0 value 10.70.102.1
   option 1 value 10.70.102.3
   option 2 value 10.70.102.5
   option 3 value 10.70.102.7
   option GROUP_NAME value PROD
   no shutdown
```

Next-hop groups

```
All interfaces are UP

  Entries
    0  10.70.100.1
    1  10.70.100.3
    2  10.70.100.5
    3  10.70.100.7
    4  10.70.100.1
    5  10.70.100.3
    6  10.70.100.5
    7  10.70.100.7
    8  10.70.100.1
    9  10.70.100.3
    10  10.70.100.5
    11  10.70.100.7
```

```
Interface Et1 is DOWN

  Entries
    0  10.70.100.3
    1  10.70.100.3
    2  10.70.100.5
    3  10.70.100.7
    4  10.70.100.5
    5  10.70.100.3
    6  10.70.100.5
    7  10.70.100.7
    8  10.70.100.7
    9  10.70.100.3
    10  10.70.100.5
    11  10.70.100.7
```

```
Interface Et1, Et3 DOWN

  Entries
    0  10.70.100.3
    1  10.70.100.3
    2  10.70.100.7
    3  10.70.100.7
    4  10.70.100.3
    5  10.70.100.3
    6  10.70.100.7
    7  10.70.100.7
    8  10.70.100.3
    9  10.70.100.3
    10  10.70.100.7
    11  10.70.100.7
```

Failover times

Thank you!

We’re Hiring

careers.booking.com

All references to “Booking.com", including any mention of “us”, “we” and “our” refer to Booking.com BV, the company behind Booking.com™