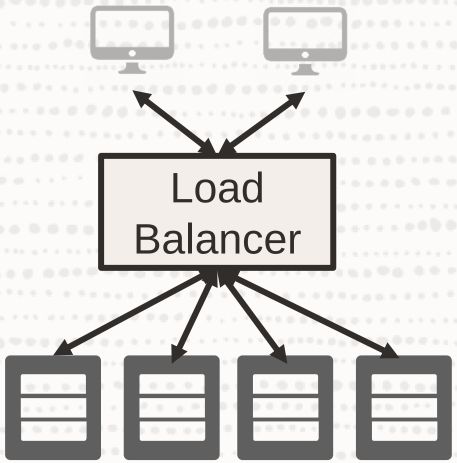
OCI Load Balancer

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

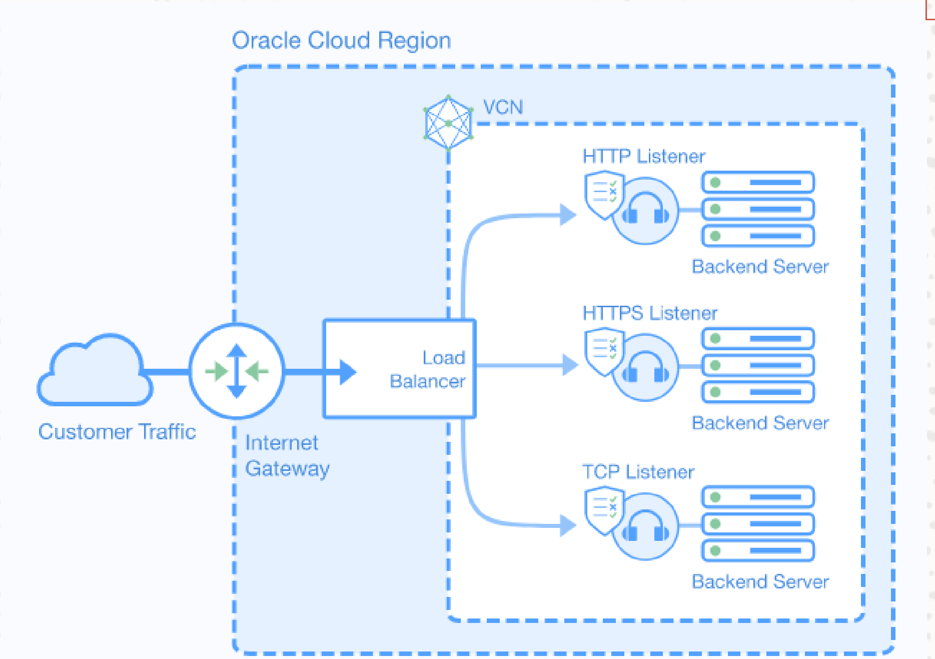
Load Balancer

- Resides between end users and backend servers.
- Load Balancer Improve Resource Utilization, Scalability and ensure High Availability.



Two Type of Load Balancer:

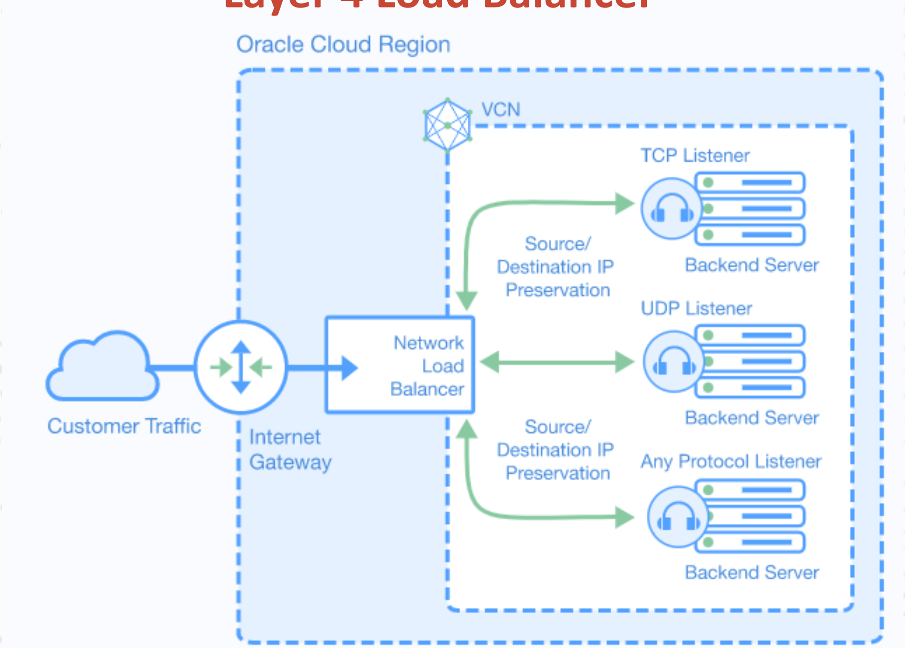
Layer 7 Load Balancer



Both LBs
Can be

Public or Private

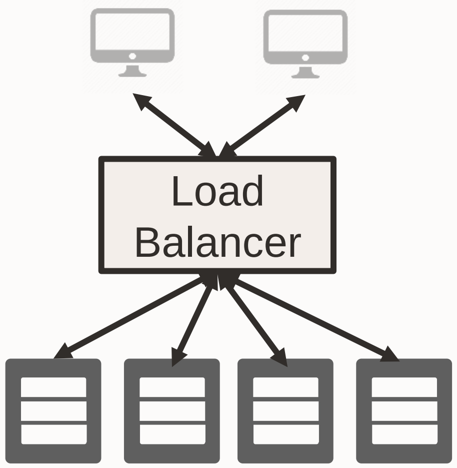
NLB: Network Load Balancer
Layer 4 Load Balancer



Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Flexible Load Balancer Benefits (Layer 7)

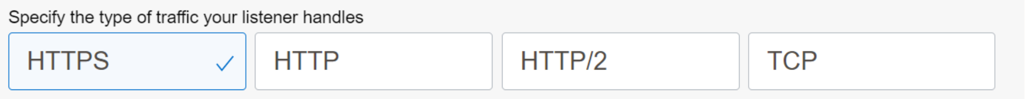
- **Fault tolerance and HA**: using health check + LB algorithms, a LB can effectively route around a bad or overloaded backend.
- **Scalability**: LB maximizes throughput, minimizes response time, and avoids overload of any single resource
- **Naming abstraction**: name resolution can be delegated to the LB; backends don’t need public IP addresses



Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Load Balancer Components

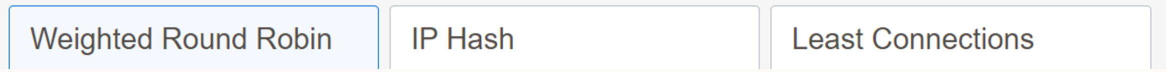
- **Listener** : Checks for incoming traffic on the load balancer's IP address, listener should configure for type of traffic.



Specify the type of traffic your listener handles

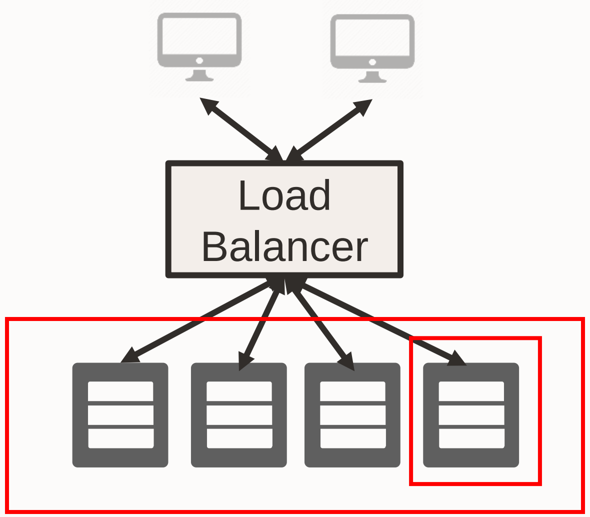
HTTPS ✓ | HTTP | HTTP/2 | TCP

- **Load Balancing Policy:** Distribution algorithm of incoming traffic.



Weighted Round Robin | IP Hash | Least Connections

- **Health Check**: What backends are currently healthy and available to accept requests?



- **Backend set** is a logical entity defined by a load balancing policy, a health check policy, and a list of backend servers.

- **Backend Server** – application server responsible for generating content in reply to the incoming TCP or HTTP traffic

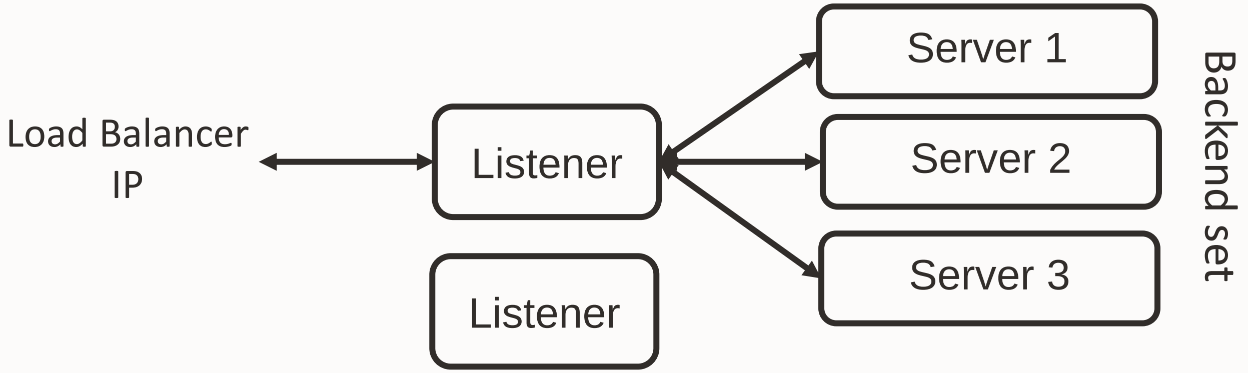
Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Health Check

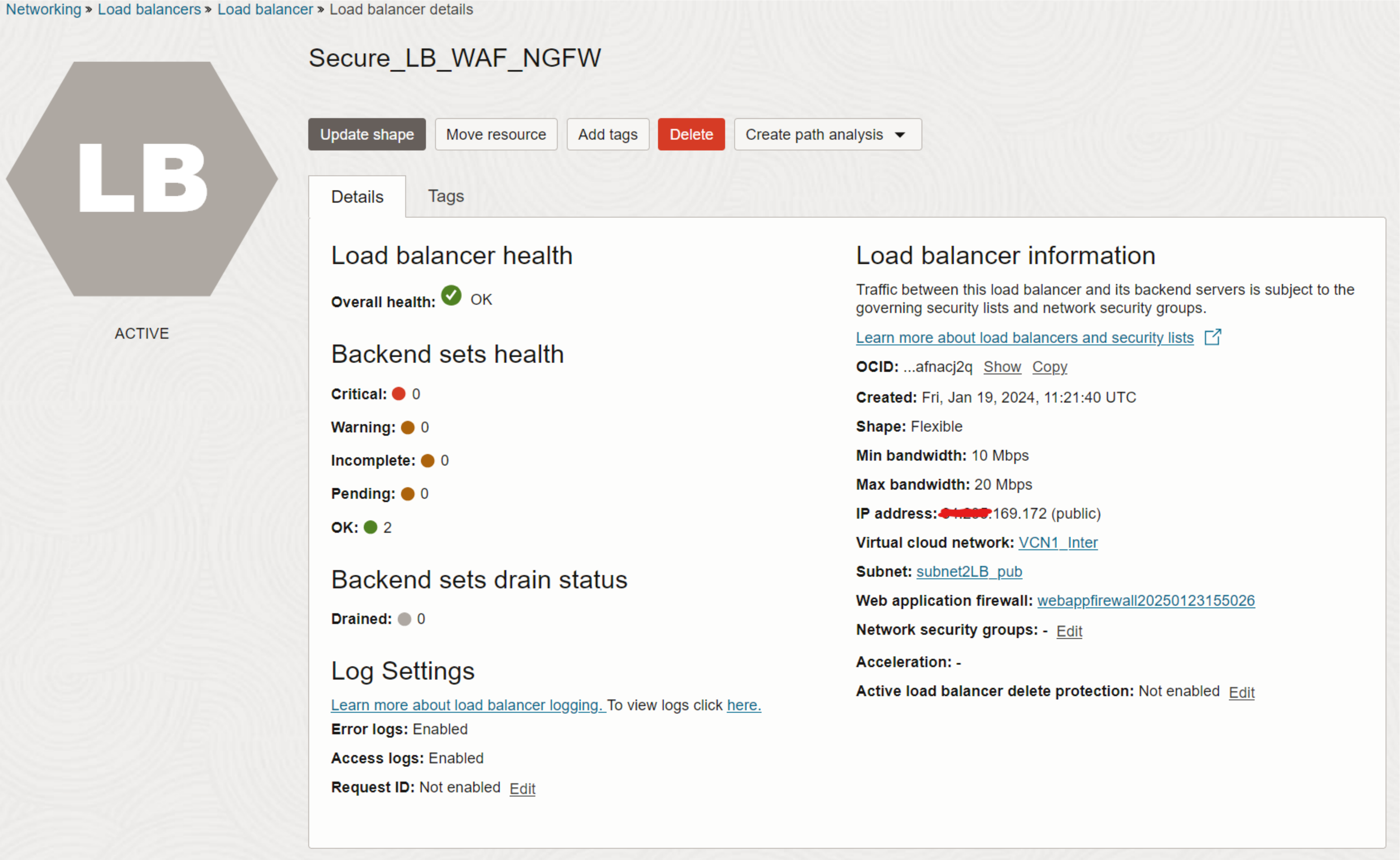
- Health check is a test to confirm the availability of backend servers.
- Health API provides a 5-state health status (**ok**, **pending**, **incomplete**, **warning**, **critical**)
- Health status is updated **every three minutes**

Listeners

- A Load Balancer IP can have **up to 16 listeners** (port numbers). Each listener has a backend set that can have 1 to N backend servers



Copyright © 2025, Oracle and/or its affiliates. All rights reserved.



Secure_LB_WAF_NGFW

ACTIVE

Load balancer health

Overall health: ✓ OK

Backend sets health

Critical: ● 0

Warning: ● 0

Incomplete: ● 0

Pending: ● 0

OK: ● 2

Backend sets drain status

Drained: ● 0

Log Settings

Learn more about load balancer logging. To view logs click here.

Error logs: Enabled

Access logs: Enabled

Request ID: Not enabled Edit

Load balancer information

Traffic between this load balancer and its backend servers is subject to the governing security lists and network security groups.

Learn more about load balancers and security lists

OCID: ...afnacj2q Show Copy

Created: Fri, Jan 19, 2024, 11:21:40 UTC

Shape: Flexible

Min bandwidth: 10 Mbps

Max bandwidth: 20 Mbps

IP address: [redacted].169.172 (public)

Virtual cloud network: VCN1_Inter

Subnet: subnet2LB_pub

Web application firewall: webappfirewall20250123155026

Network security groups: - Edit

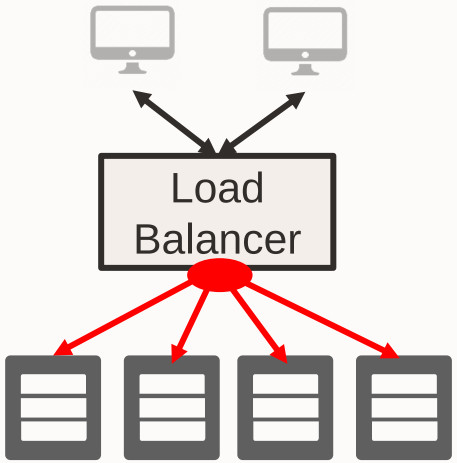
Acceleration: -

Active load balancer delete protection: Not enabled Edit

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

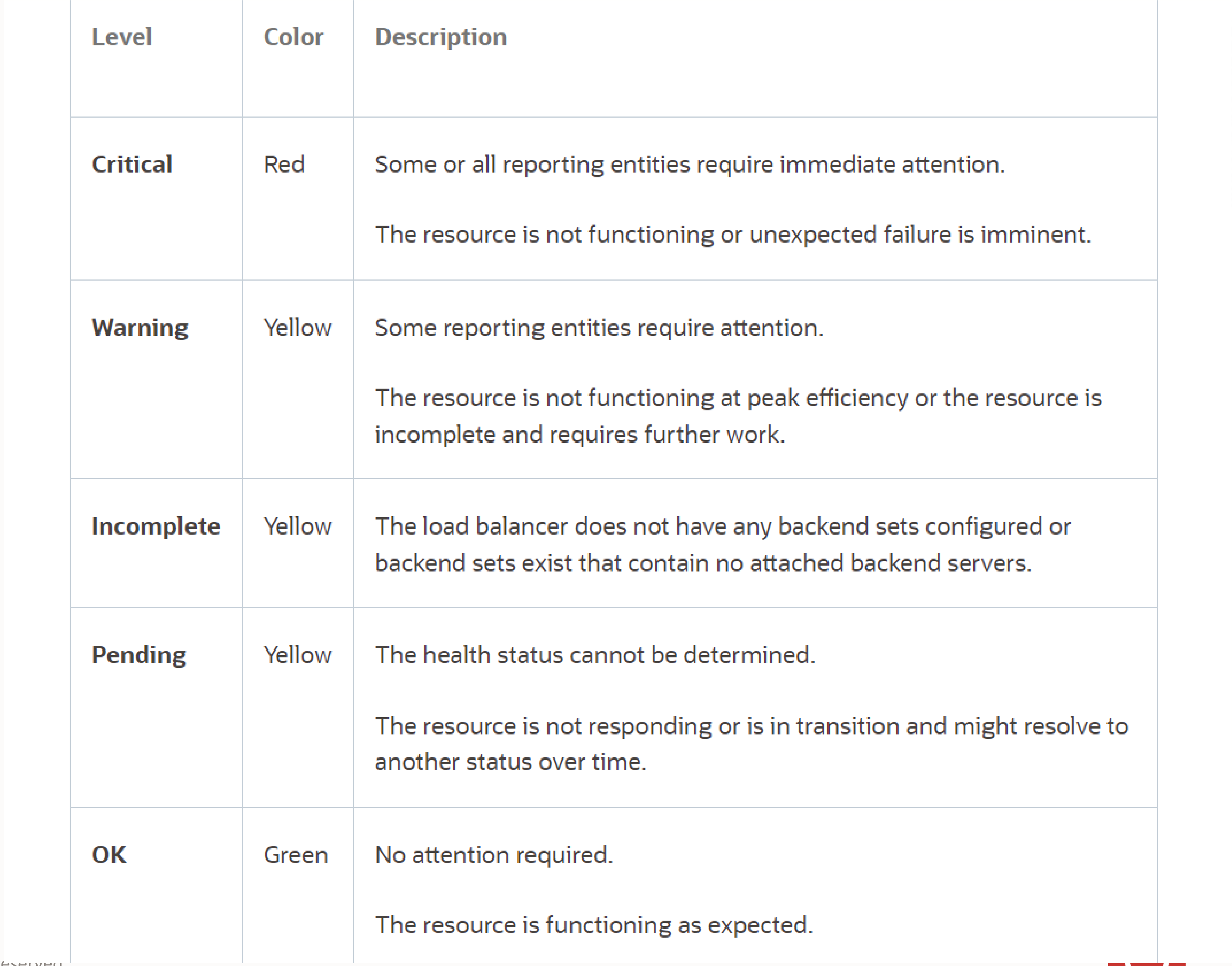
Load Balancing Policies

- **Round Robin**: default policy, distributes incoming traffic sequentially to each server in a backend set. After each server has received a connection, the load balancer repeats the list in the same order.
- **IP Hash**: uses an incoming request's source IP address as a hashing key to route non-sticky traffic to the same backend server.
- **Least Connection**: routes incoming non-sticky request traffic to the backend server with the fewest active connections.
- Load balancer policy decisions apply differently to TCP load balancer, cookie-based session persistent HTTP requests (sticky requests), and non-sticky HTTP requests
  - A TCP load balancer considers policy and weight criteria
  - An HTTP load balancer w/ cookie-based session persistence forwards requests using cookie's session info
  - For non-sticky HTTP requests, the load balancer applies policy and weight criteria



Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Backends Health Status



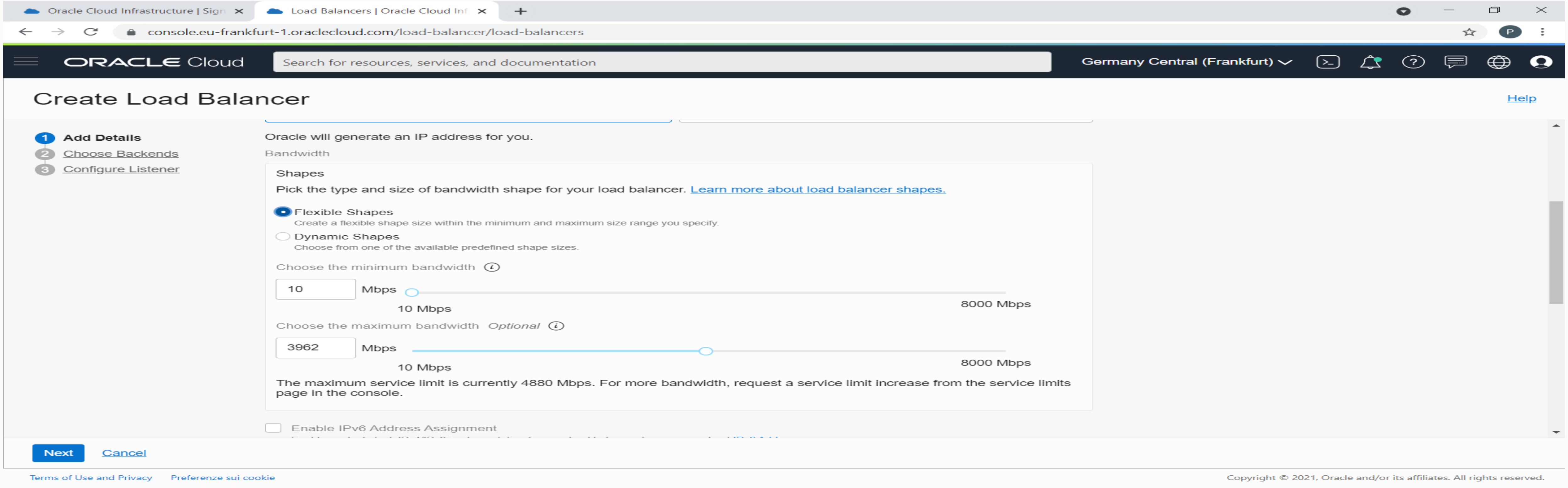
Level	Color	Description
Critical	Red	Some or all reporting entities require immediate attention. The resource is not functioning or unexpected failure is imminent.
Warning	Yellow	Some reporting entities require attention. The resource is not functioning at peak efficiency or the resource is incomplete and requires further work.
Incomplete	Yellow	The load balancer does not have any backend sets configured or backend sets exist that contain no attached backend servers.
Pending	Yellow	The health status cannot be determined. The resource is not responding or is in transition and might resolve to another status over time.
OK	Green	No attention required. The resource is functioning as expected.

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Flexible Load Balancer

Customer just defines the minimum and maximum bandwidth

- Minimum bandwidth provides instant readiness for load
- Maximum bandwidth allows control of maximum cost
- Customer pays a minimal base cost for the load balancer and then pays a simple single rate for the larger of the reserved bandwidth or the maximum bandwidth actually used each minute



Oracle Cloud Infrastructure | Sign
Load Balancers | Oracle Cloud Inf
console.eu-frankfurt-1.oraclecloud.com/load-balancer/load-balancers
ORACLE Cloud
Search for resources, services, and documentation
Germany Central (Frankfurt)
Help
1 Add Details
2 Choose Backends
3 Configure Listener
Next
Cancel
Terms of Use and Privacy
Preferenze sui cookie

Create Load Balancer

Oracle will generate an IP address for you.

Bandwidth

Shapes

Pick the type and size of bandwidth shape for your load balancer. Learn more about load balancer shapes.

● Flexible Shapes
Create a flexible shape size within the minimum and maximum size range you specify.

○ Dynamic Shapes
Choose from one of the available predefined shape sizes.

Choose the minimum bandwidth: 10 Mbps (10 Mbps – 8000 Mbps)

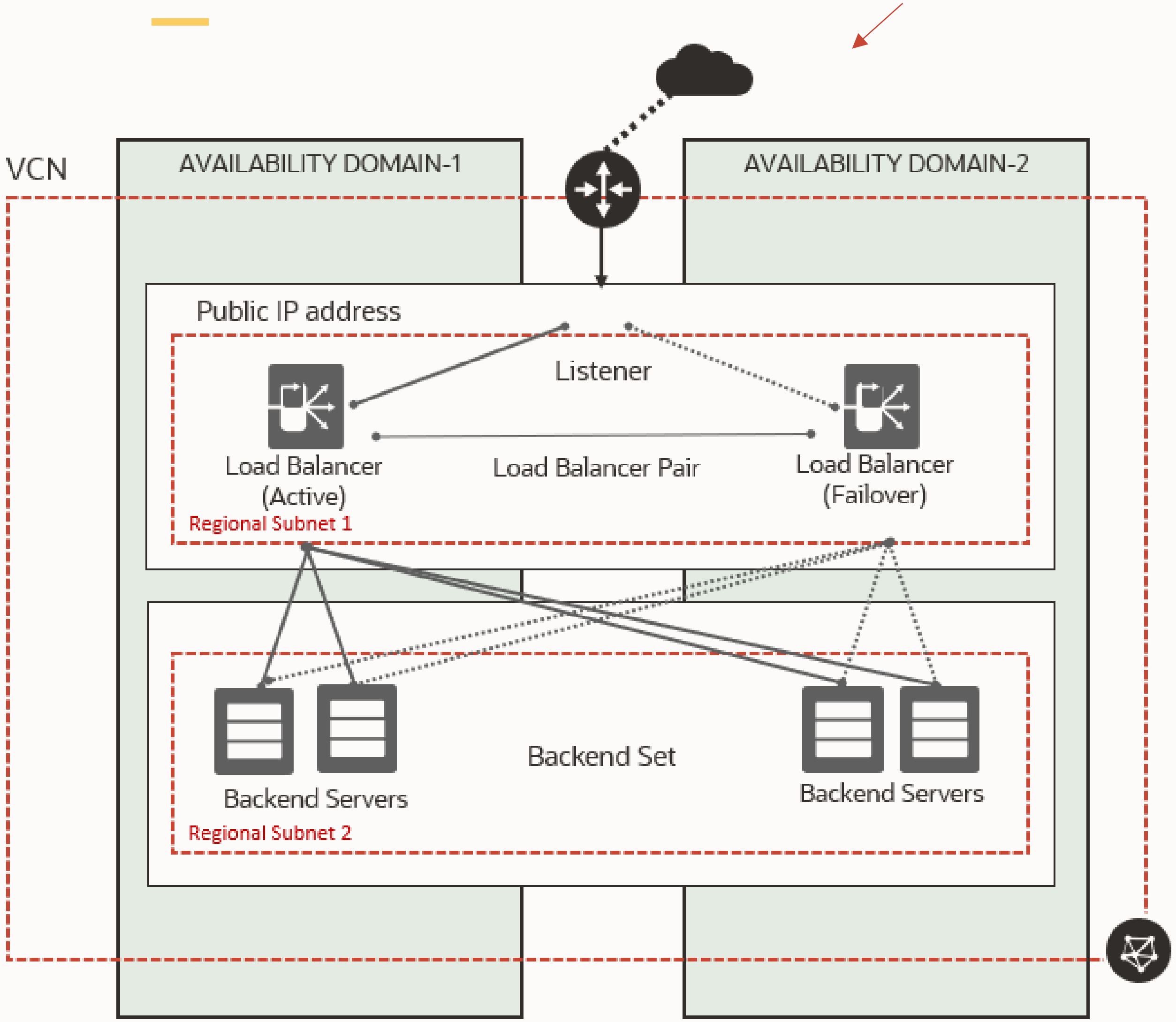
Choose the maximum bandwidth *Optional*: 3962 Mbps (10 Mbps – 8000 Mbps)

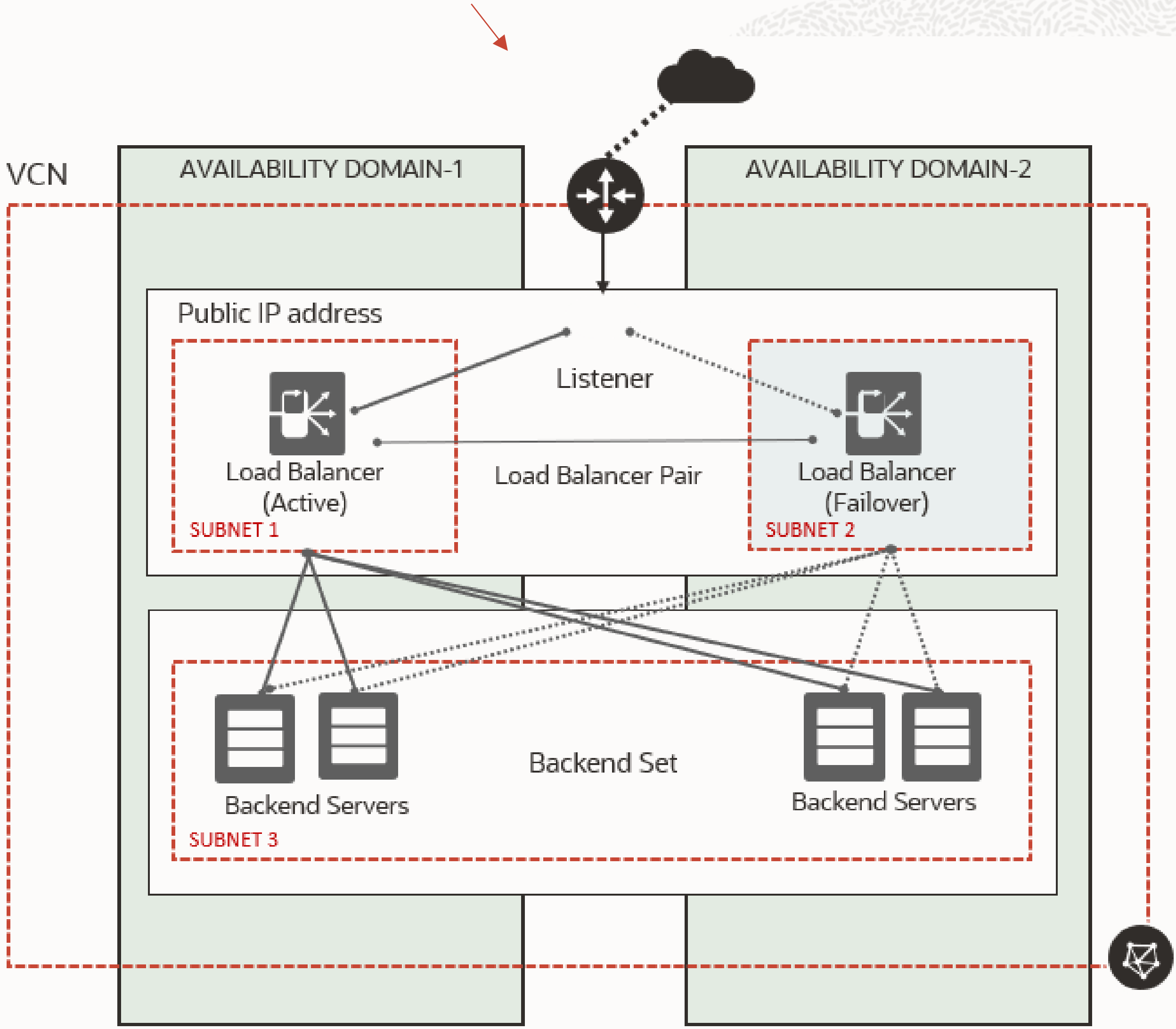
The maximum service limit is currently 4880 Mbps. For more bandwidth, request a service limit increase from the service limits page in the console.

□ Enable IPv6 Address Assignment

Copyright © 2021, Oracle and/or its affiliates. All rights reserved.

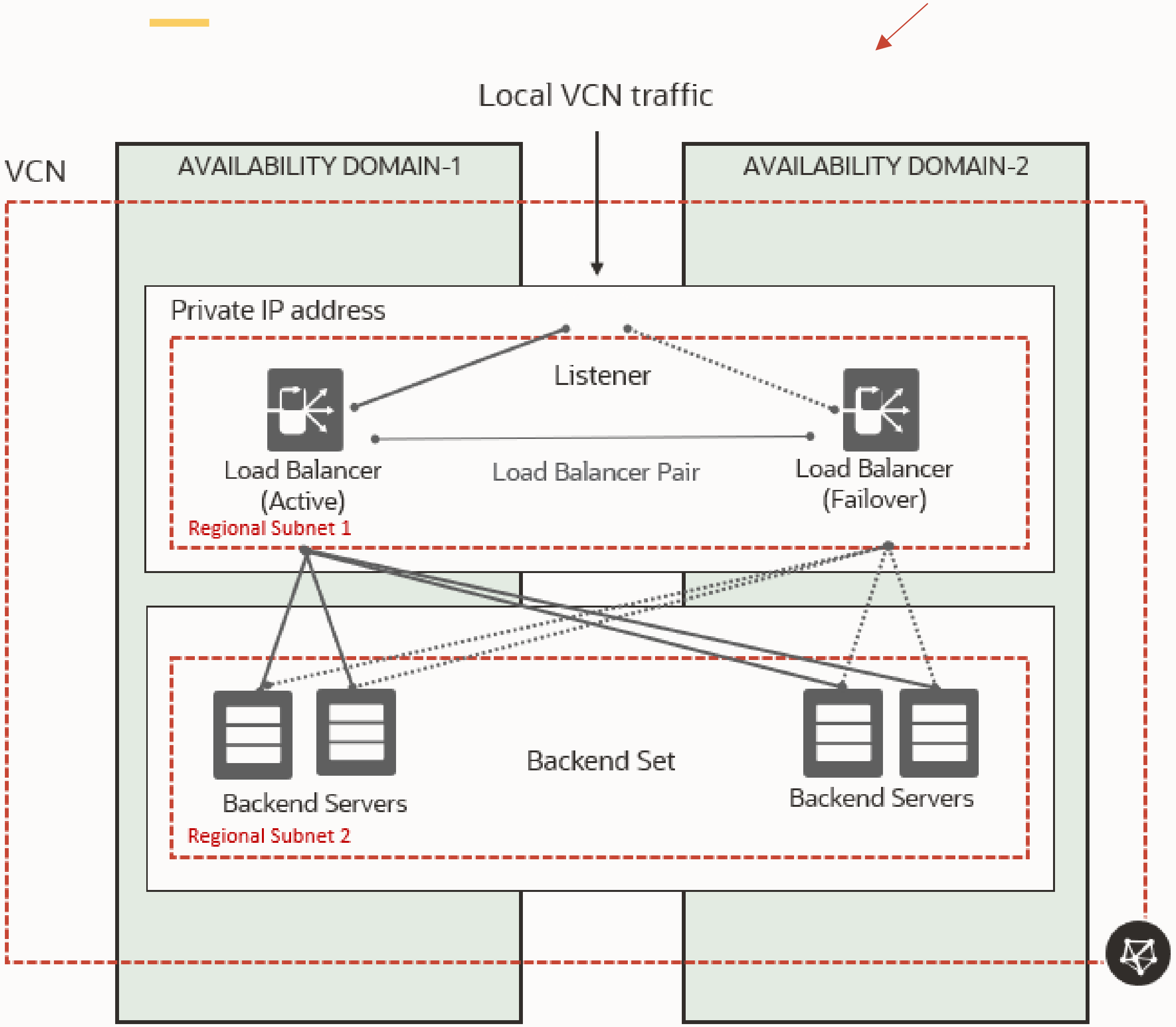
Public Load Balancer (Regional Subnets -Recommended- vs. AD Specific Subnets)

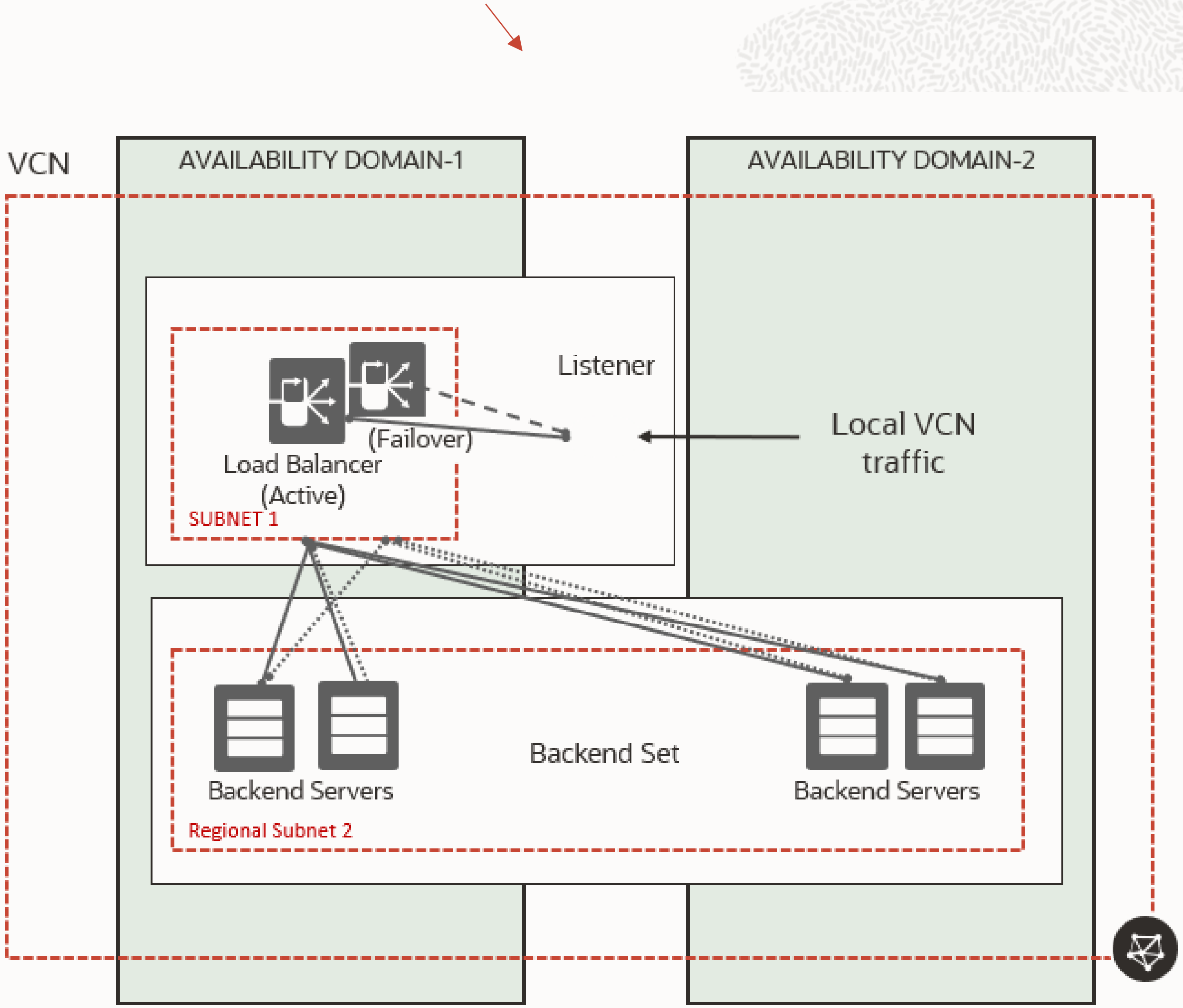




Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Private Load Balancer (Regional Subnets -Recommended- vs. AD Specific Subnets)





Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

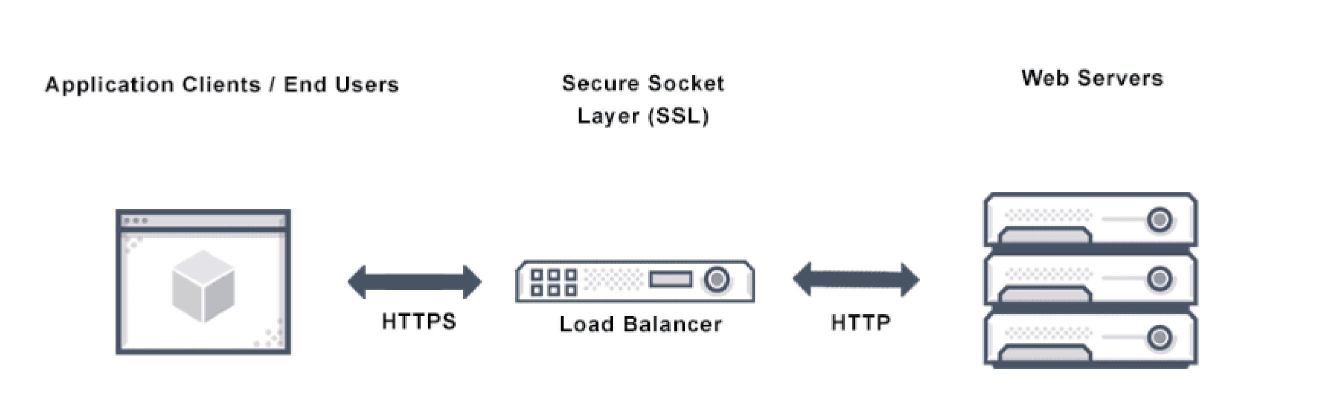
SSL Handling

Terminate SSL at the load balancer. This configuration is ***frontend SSL***. Your load balancer can accept encrypted traffic from a client. No encryption of traffic exists between the load balancer and the backend servers.

Implement SSL between the load balancer and your backend servers. This configuration is ***backend SSL***. Your load balancer does not accept encrypted traffic from client servers. Traffic between the load balancer and the backend servers is encrypted.

Implement point-to-point SSL. Your load balancer can accept SSL encrypted traffic from clients and encrypts traffic to the backend servers.

- To use SSL with your load balancer, you must add one or more certificate bundles to your system.
- Oracle Cloud Infrastructure accepts x.509 type certificates in PEM format only.



SSL Termination

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

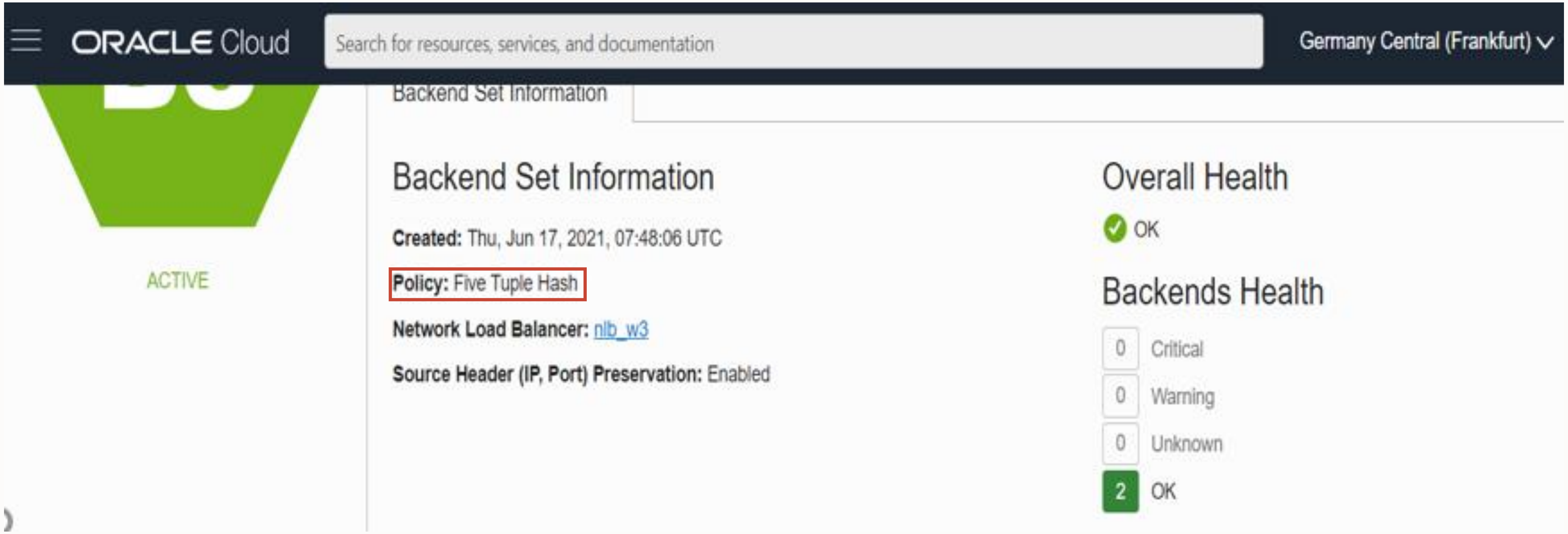
Network Load Balancer (Layer 4)

- The Network Load Balancer load balance **layer 3 and layer 4 (TCP/UDP/ICMP) workloads**
- It's designed to handle volatile traffic patterns and millions of flows, offering high throughput while maintaining **ultra-low latency**
- Ideal load balancing solution for **latency-sensitive workloads includes real-time streaming, VoIP, Internet of Things, HA scenarios and trading platforms**
- OCI Network Load Balancer can be public or private
- **Symmetric Hashing** support for Active/Active configurations
- Full NAT**,** Source/Destination Header (IP/Port) Preservation (Bump-in-the-wire), Source Header (IP/Port) Preservation

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Network Load Balancer - Policies

5-Tuple Hash: Routes incoming traffic based on 5-Tuple (source IP and port, destination IP and port, protocol) Hash. This is the default network load balancer policy.
3-Tuple Hash: Routs incoming traffic based on 3-Tuple (source IP, destination IP, protocol) Hash.
2-Tuple Hash: Routs incoming traffic based on 2-Tuple (source IP Destination, destination IP) Hash.


ORACLE Cloud
Search for resources, services, and documentation
Germany Central (Frankfurt)
Backend Set Information
ACTIVE

Backend Set Information

Created: Thu, Jun 17, 2021, 07:48:06 UTC

Policy: Five Tuple Hash

Network Load Balancer: nlb_w3

Source Header (IP, Port) Preservation: Enabled

Overall Health

✓ OK

Backends Health

0	Critical
0	Warning
0	Unknown
2	OK

Copyright © 2025, Oracle and/or its affiliates. All rights reserved.

Thanks