CXS1822BCN

vmware EXPLORE

Demystifying Multi-Tier Gateways in VMware NSX

Tim Burkard (He/Him/His)

Staff Technical Instructor, VMware Learning

#vmwareexplore #CXS1823BCN

Agenda

1

Logical Routing in VMware NSX-T

2

Where do the packets go?

3

Investigating more deeply

Presenter

Tim Burkard

Staff Technical Learning Engineer

Tim is a VMware employee and has been a technical instructor for 20+ years. He has trained IT professionals in the areas of operating systems, security and networking. Tim specializes in teaching VMware NSX and has previously taught courses covering the entire vSphere product line.

Logical Routing in NSX-T

What's so special about this anyway?

Why should we care about Logical Routing?

Logical routing is provided in NSX-T to move packets between segments. It's flexible, in that we can run a single-tier deployment to provide a simple routing and services architecture, or we can provide more flexible design with multi-tier routing.

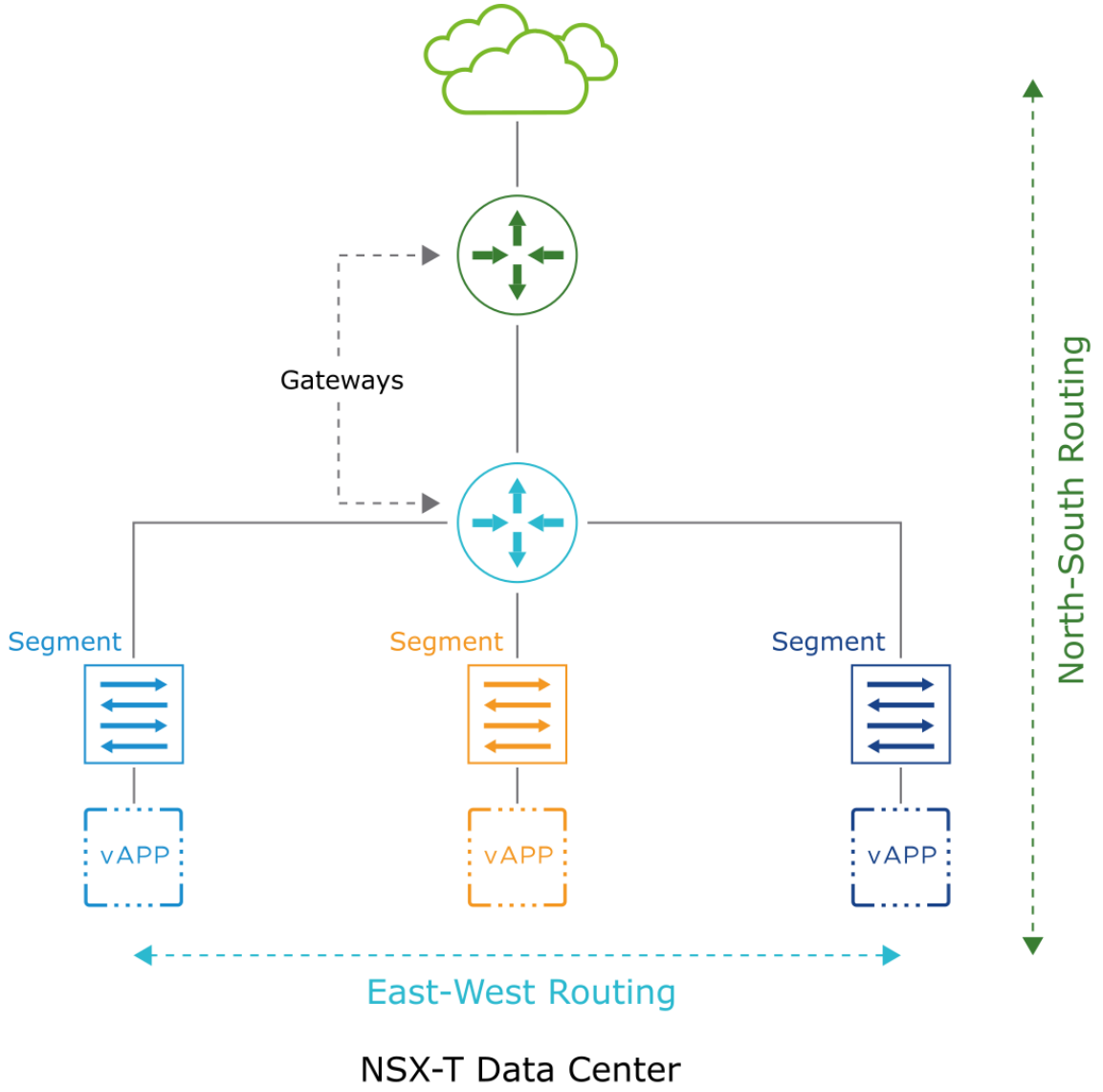
A logical router in NSX-T is a multi-component function, consisting of a DR, and optionally a SR. The DR, as the name implies, is distributed across all transport nodes, where the SR is centralized and realized on an NSX Edge Cluster.

Speaking of an NSX Edge Cluster, what is that? It's simply a logical grouping of NSX Edge transport nodes that provides enhanced scalability and high availability.

Logical Routing in NSX-T Data Center

NSX-T Data Center gateways provide:

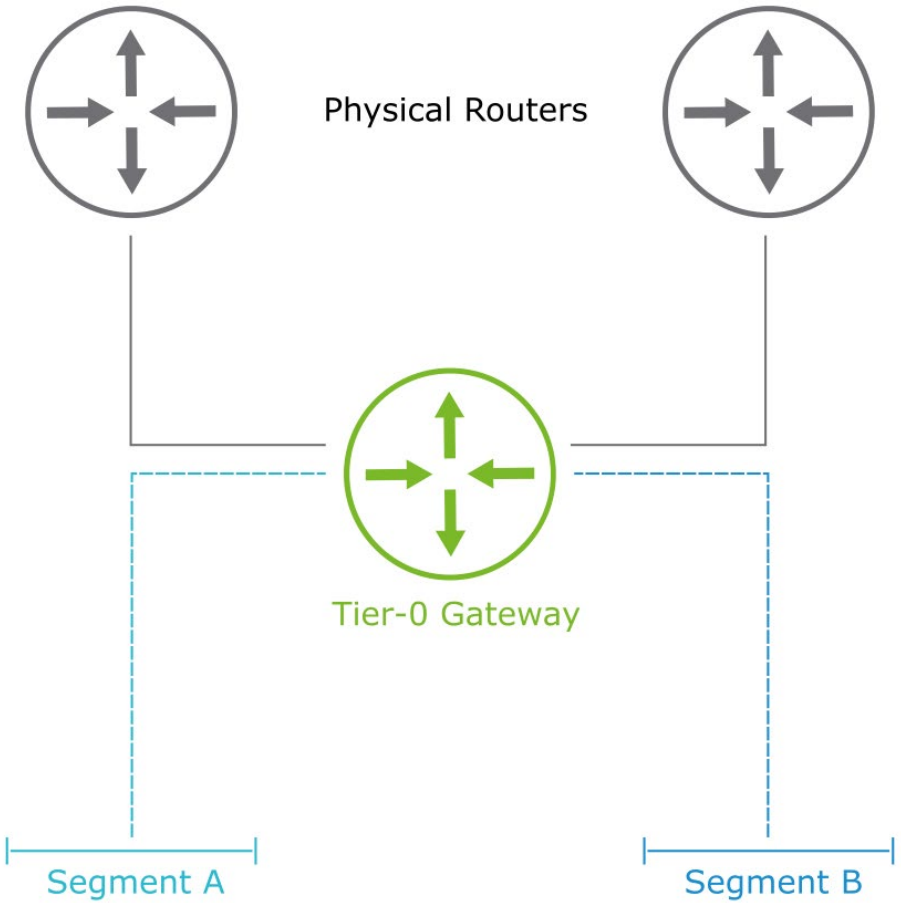
- Centralized north-south routing
- Distributed east-west routing
- Multitenant support
- Centralized stateful services, such as NAT or load balancing



Routing Topologies: Single-Tier

In a single-tier topology:

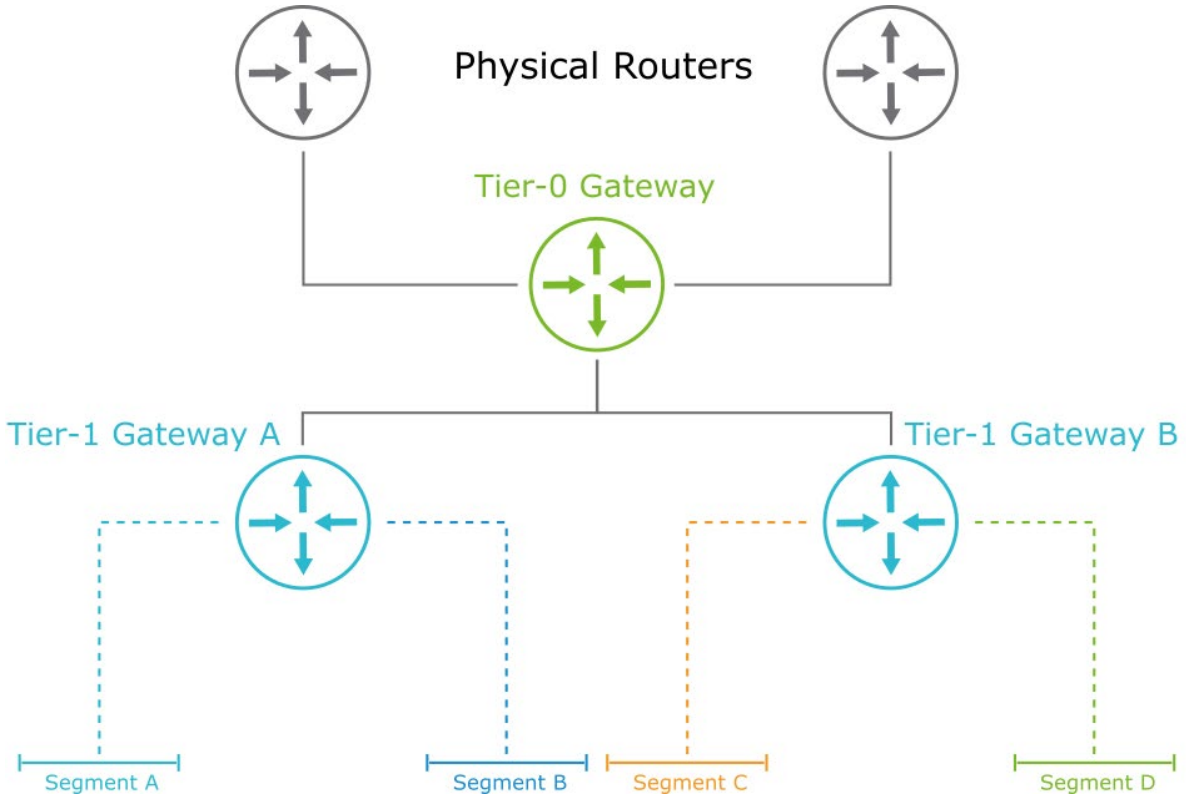
- Only Tier-0 gateways are included.
- Segments are connected directly to the Tier-0 gateway.



Routing Topologies: Multitier

In a multitier topology:

- Tier-0 and Tier-1 gateways are included.
- Tier-1 gateways are connected to the Tier-0 gateways.
- Segments are connected to the Tier-1 gateways.



Gateway Components: Distributed Router and Service Router (1)

A distributed router (DR) has the following features:

- Provides basic packet-forwarding functionalities
- Spans all transport nodes (host and edge transport nodes)
- Runs as a kernel module in the ESXi hypervisor and as an OVS file in the KVM
- Provides distributed routing functionality
- Provides first-hop routing for workloads

A service router (SR) has the following features:

- Provides north-south routing
- Provides centralized services, such as NAT and load balancing
- Required for the uplinks to external networks
- Deployed in edge transport nodes

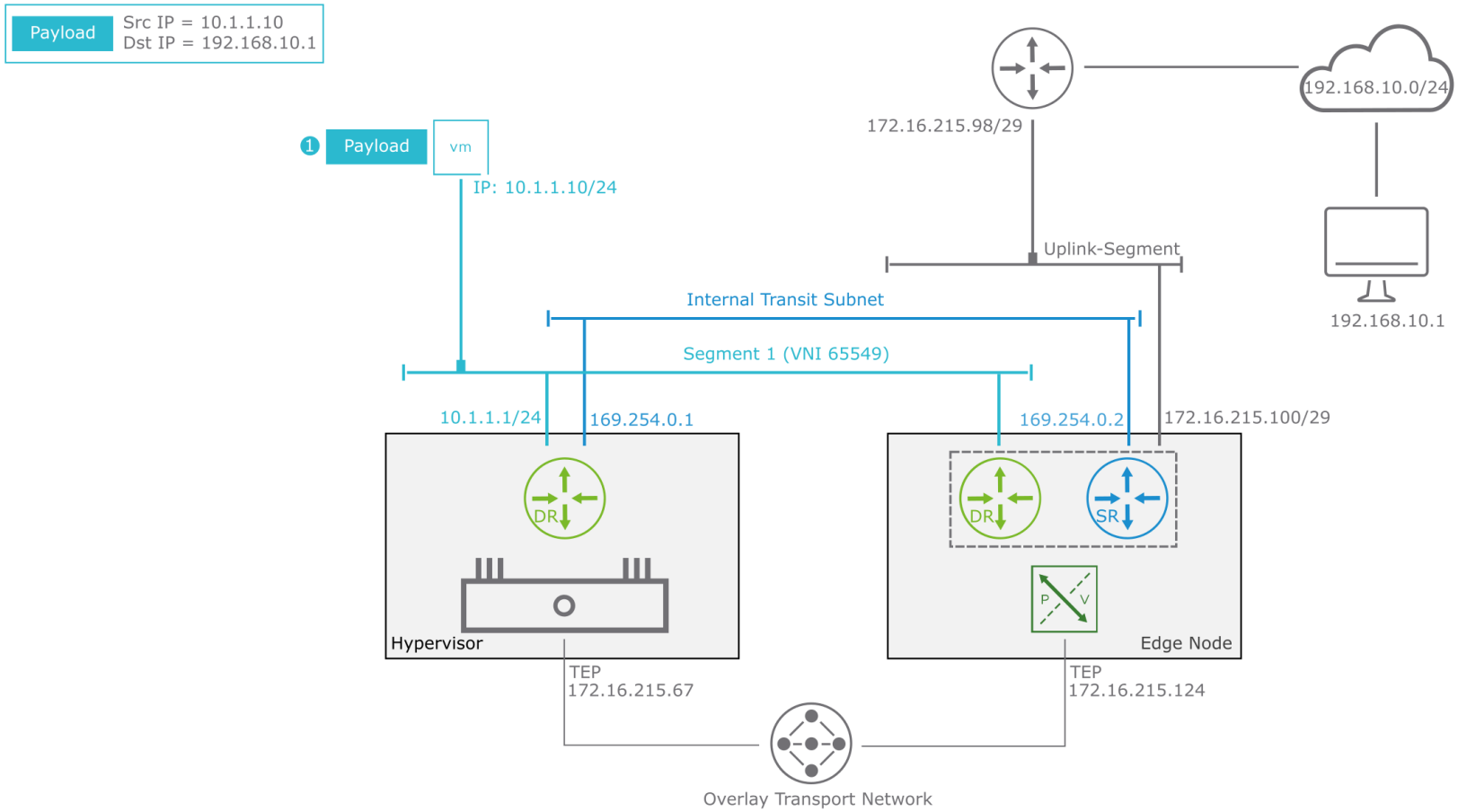
Where do the packets go?

I'll give you a hint: They don't hide under the sofa

Single-Tier Routing: Egress to Physical Network (1)

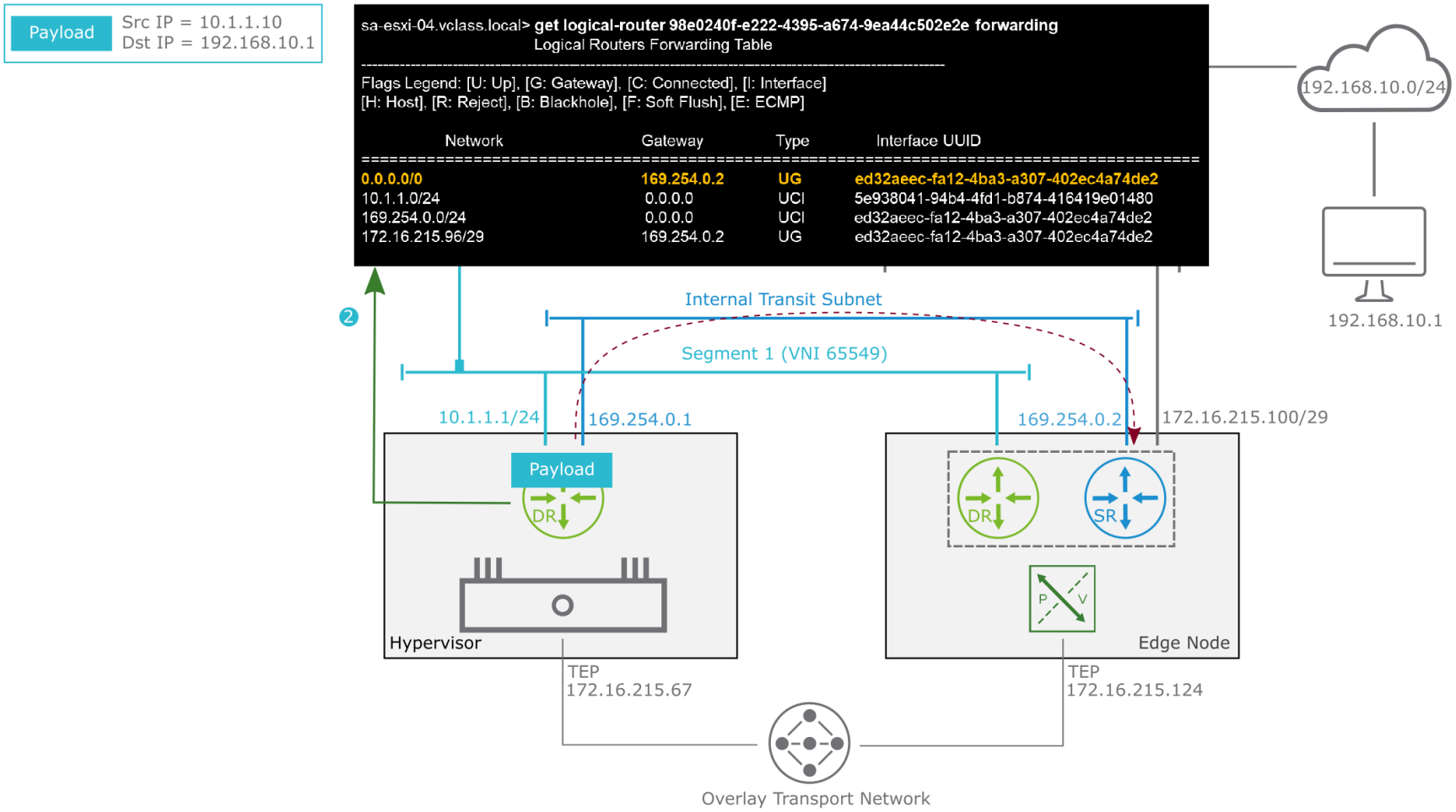
A packet is sent from the source VM 10.1.1.10 to the destination VM 192.168.10.1:

1. The packet is forwarded to its default 10.1.1.1 gateway.



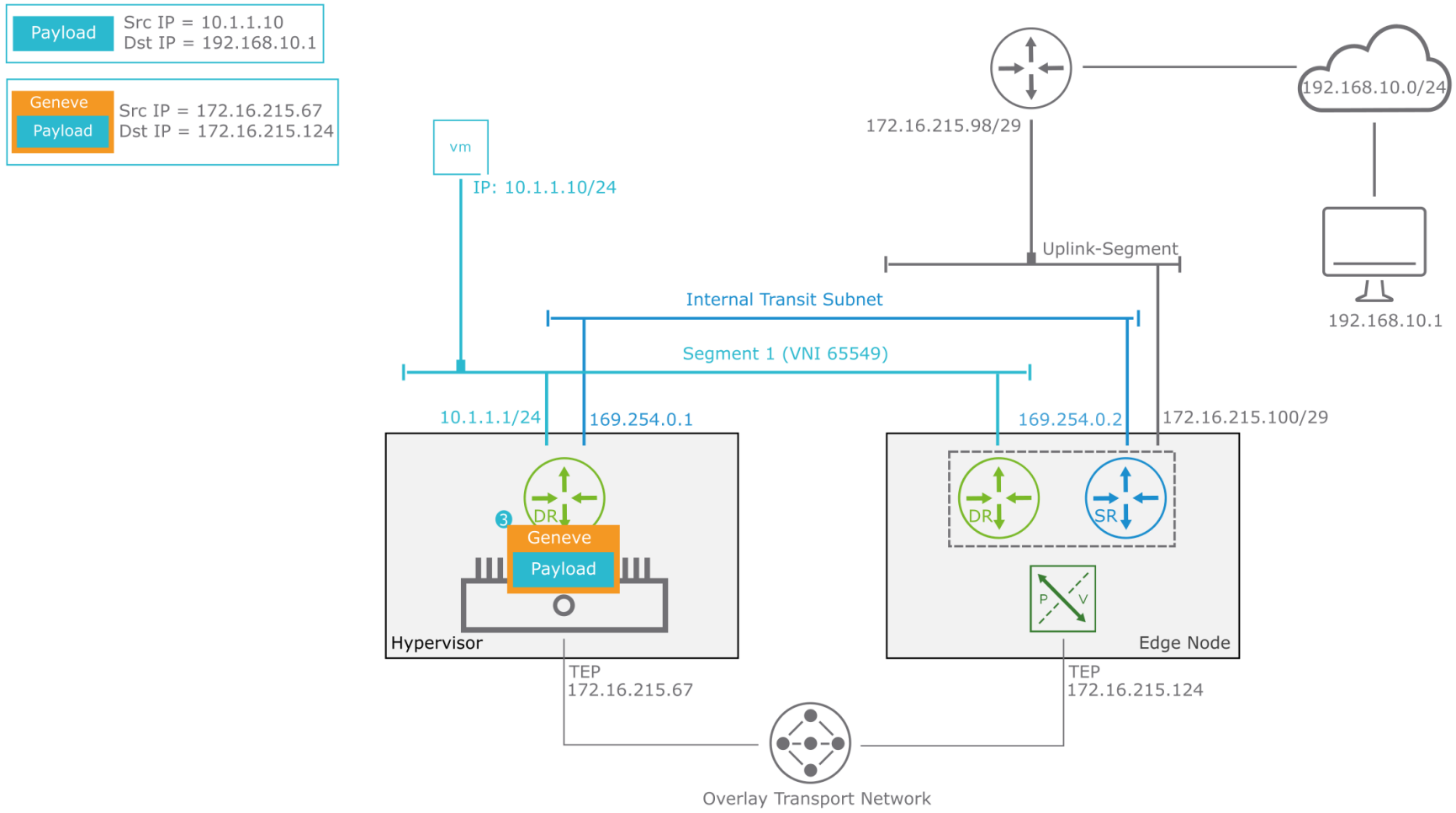
Single-Tier Routing: Egress to Physical Network (2)

2.The gateway (DR) checks its forwarding table. Because a specific route does not exist for the 192.168.10.0/24 network, the packet is sent to the default 169.254.0.2 gateway, which is the SR component on the edge node.



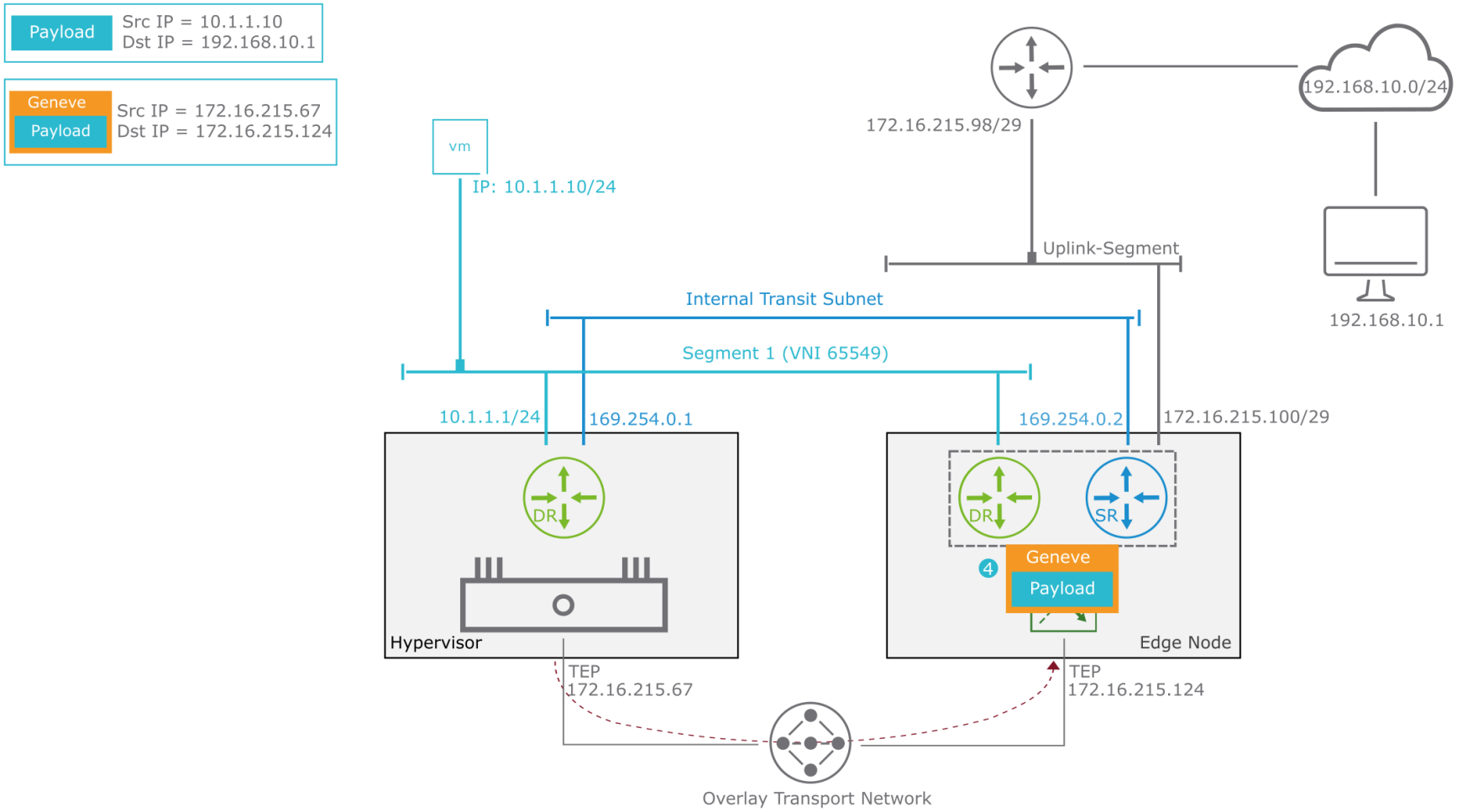
Single-Tier Routing: Egress to Physical Network (3)

3. To send the packet from the hypervisor to the edge node, the packet is encapsulated with a Geneve header.



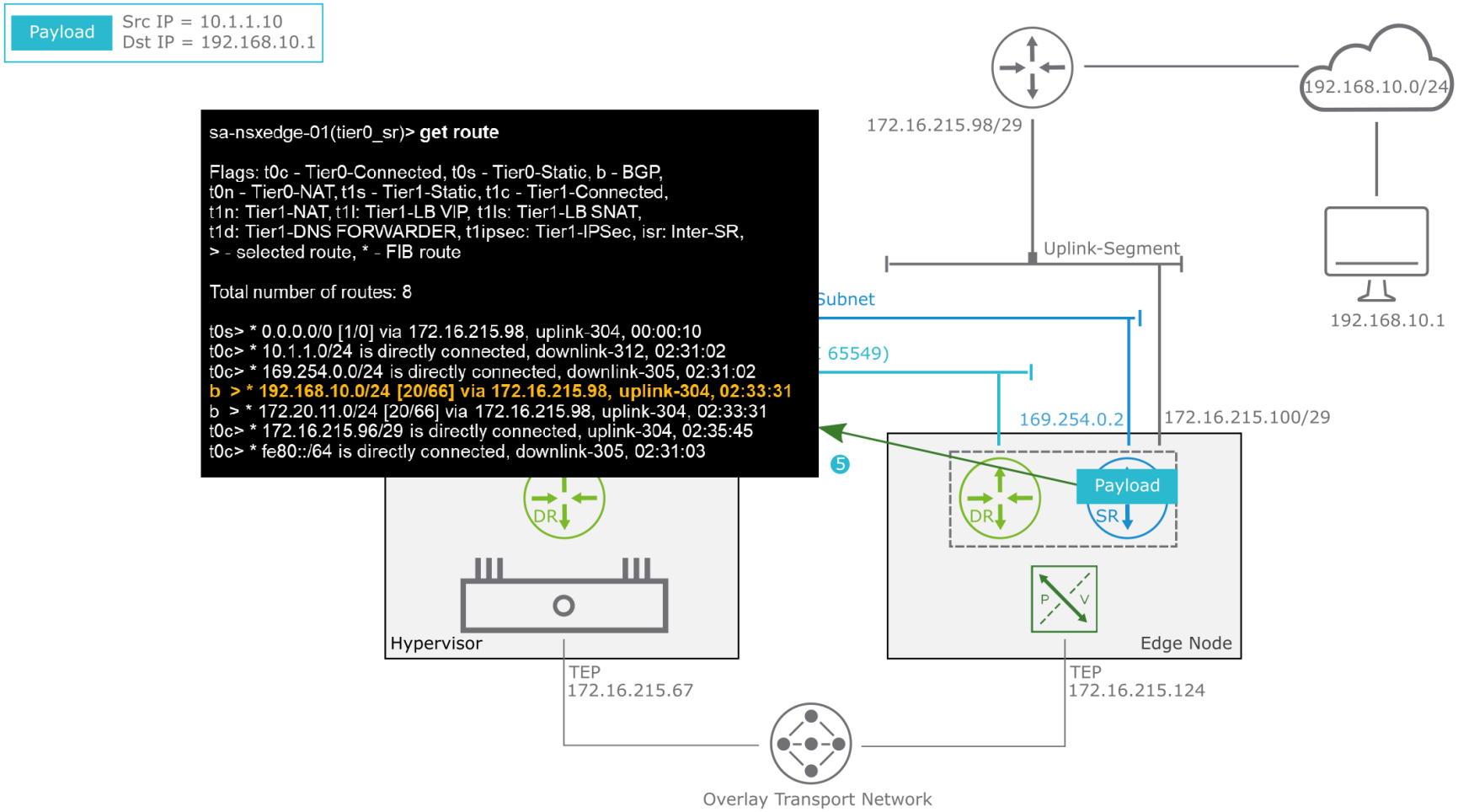
Single-Tier Routing: Egress to Physical Network (4)

4. The encapsulated packet is sent to the edge node across the overlay tunnel.



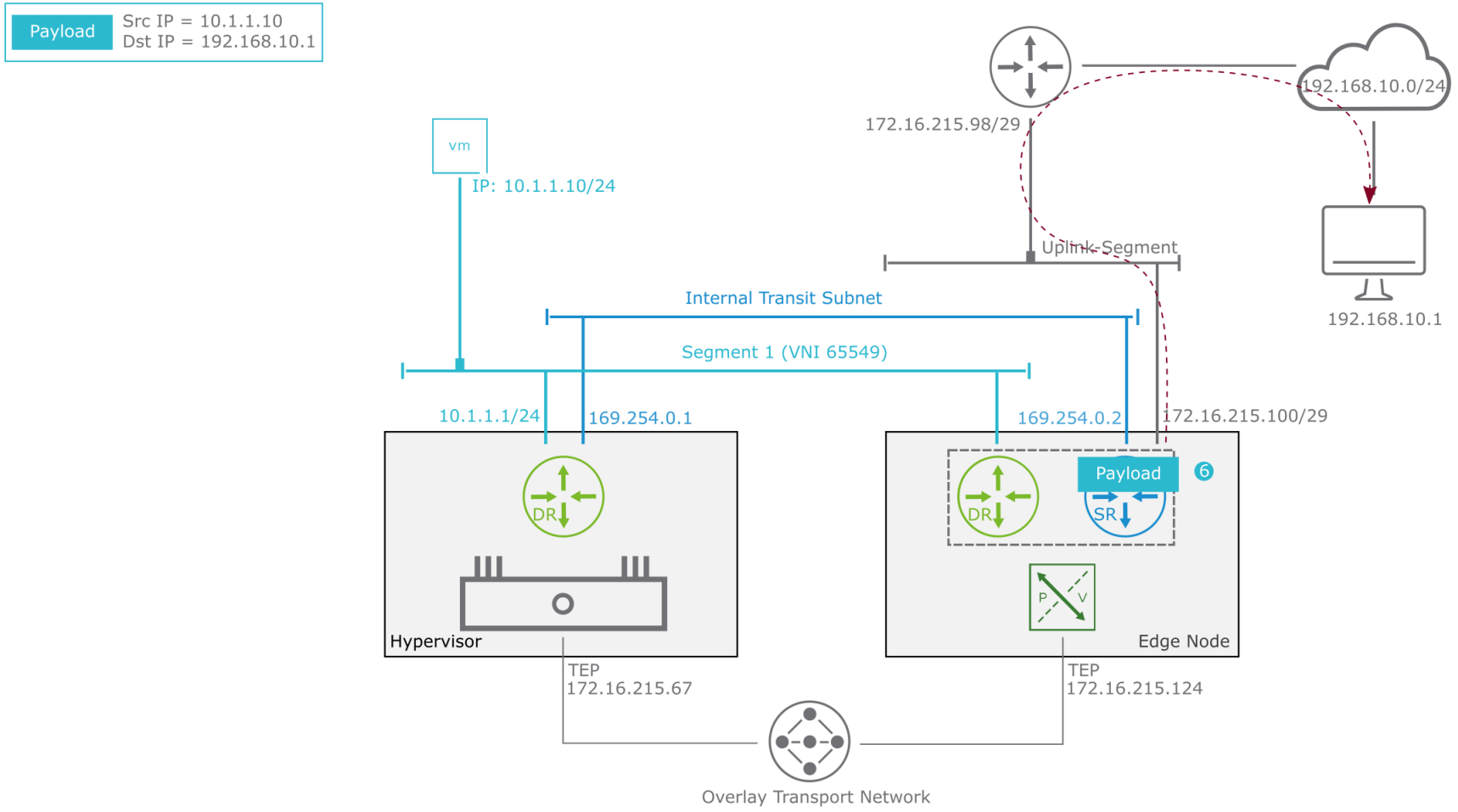
Single-Tier Routing: Egress to Physical Network (5)

5. The edge node decapsulates the packet and sends it to its SR component. The gateway (SR) routing table shows a route for the 192.168.10.0/24 network over the uplink segment.



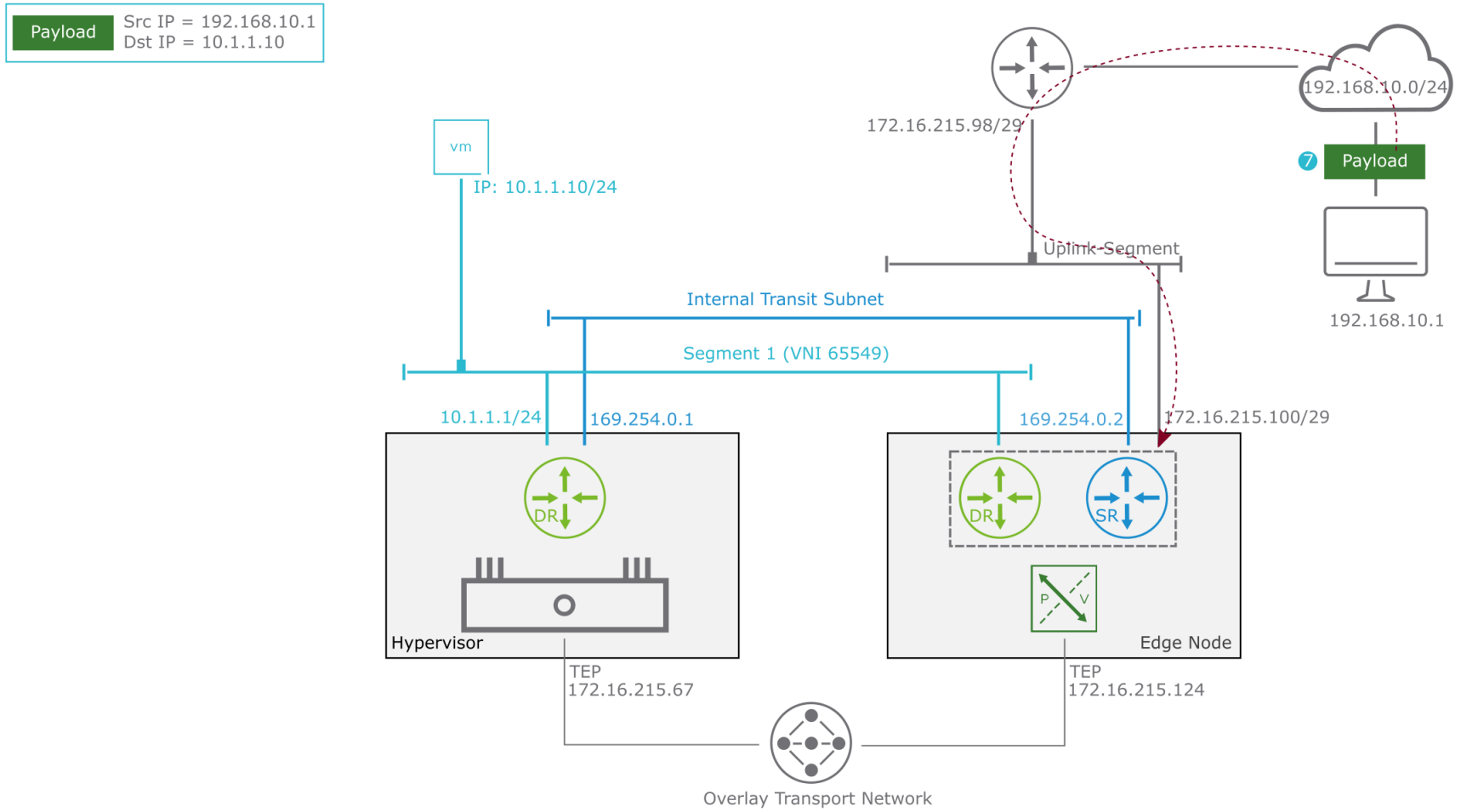
Single-Tier Routing: Egress to Physical Network (6)

6. The edge node sends the packet to its upstream physical gateway, which routes the packet to its destination 192.168.10.1.



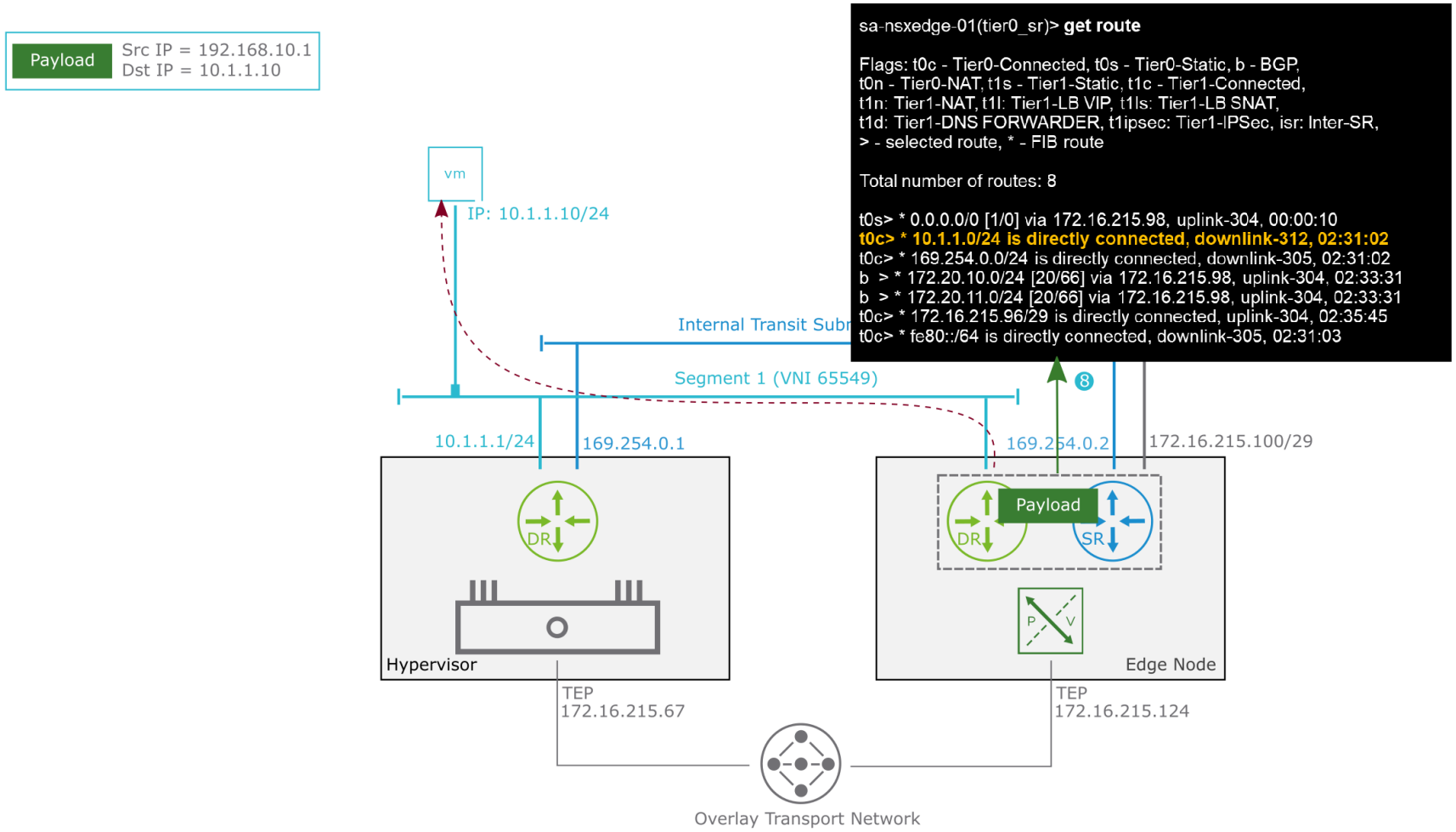
Single-Tier Routing: Ingress from Physical Network (7)

7. For the return packet, the source VM 192.168.10.1 sends the packet to its default gateway, which routes the packet to the edge node.



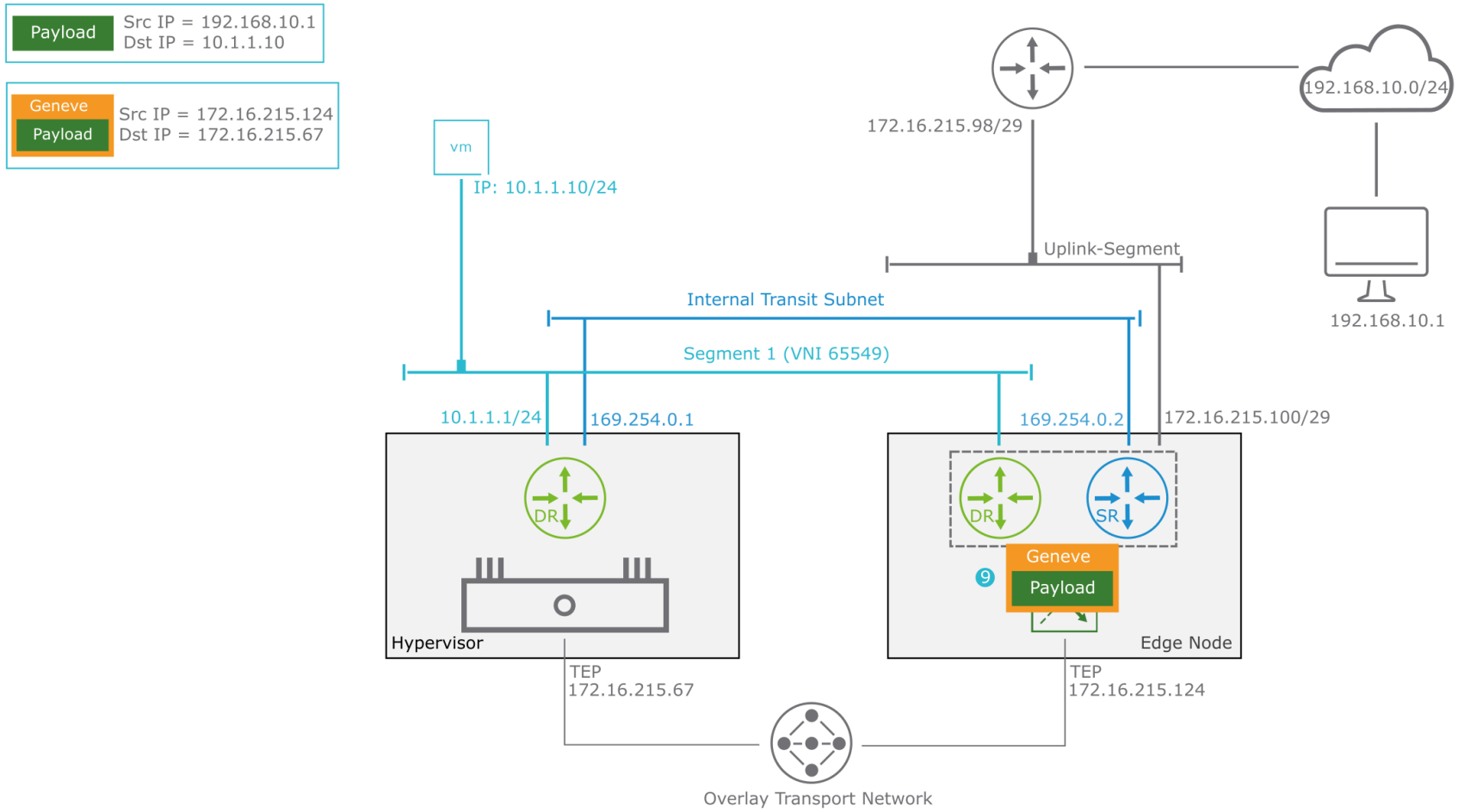
Single-Tier Routing: Ingress from Physical Network (8)

8. The SR and the DR components on an edge node share their routing table. A route is directly connected to the 10.1.1.0/24 network over Segment 1. The packet is sent to the remote host by using the DR interface.



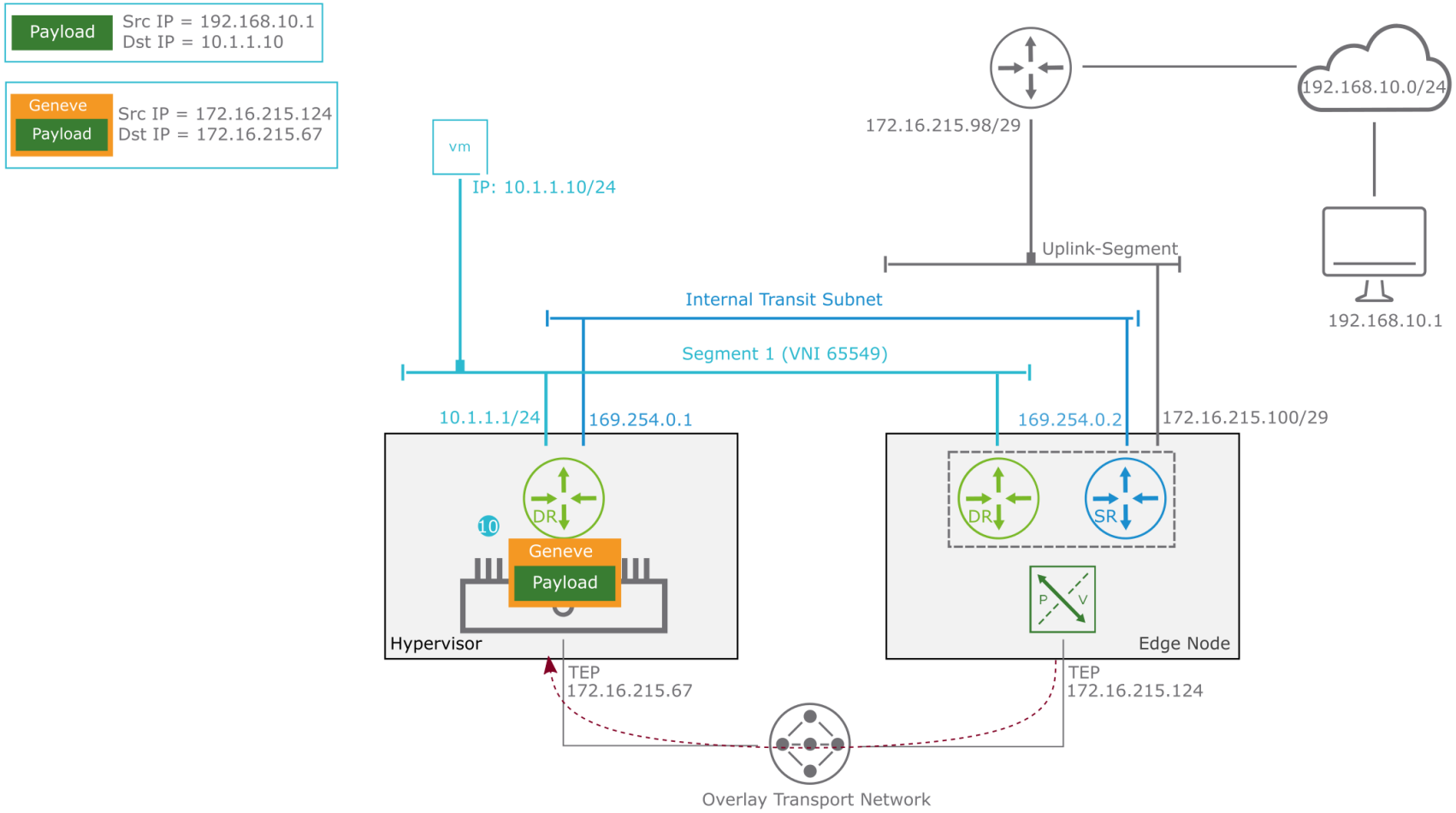
Single-Tier Routing: Ingress from Physical Network (9)

9. To send the packet from the edge node to the hypervisor, the packet is encapsulated with a Geneve header.



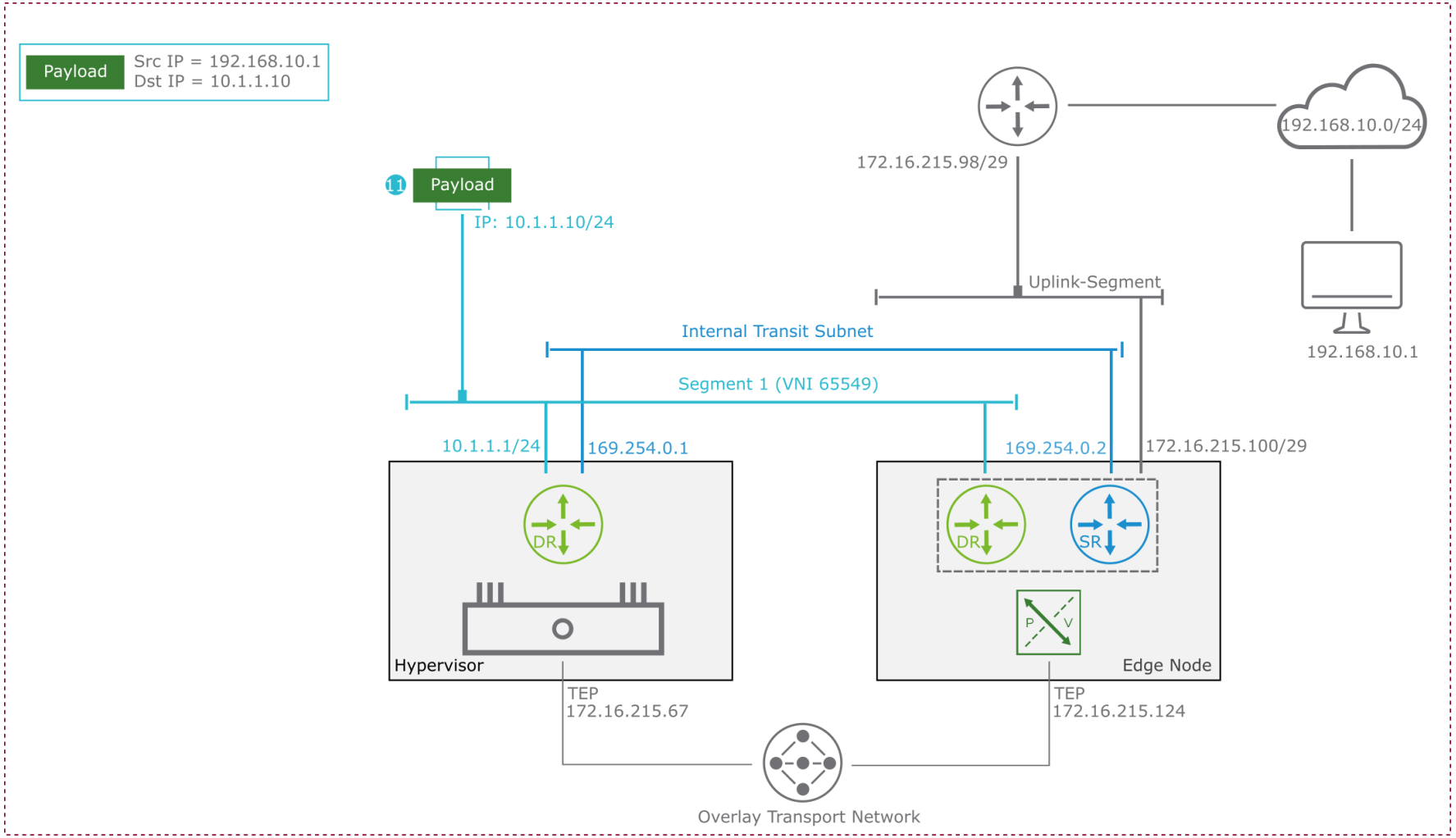
Single-Tier Routing: Ingress from Physical Network (10)

10. The encapsulated packet is sent across the overlay tunnel.



Single-Tier Routing: Ingress from Physical Network (11)

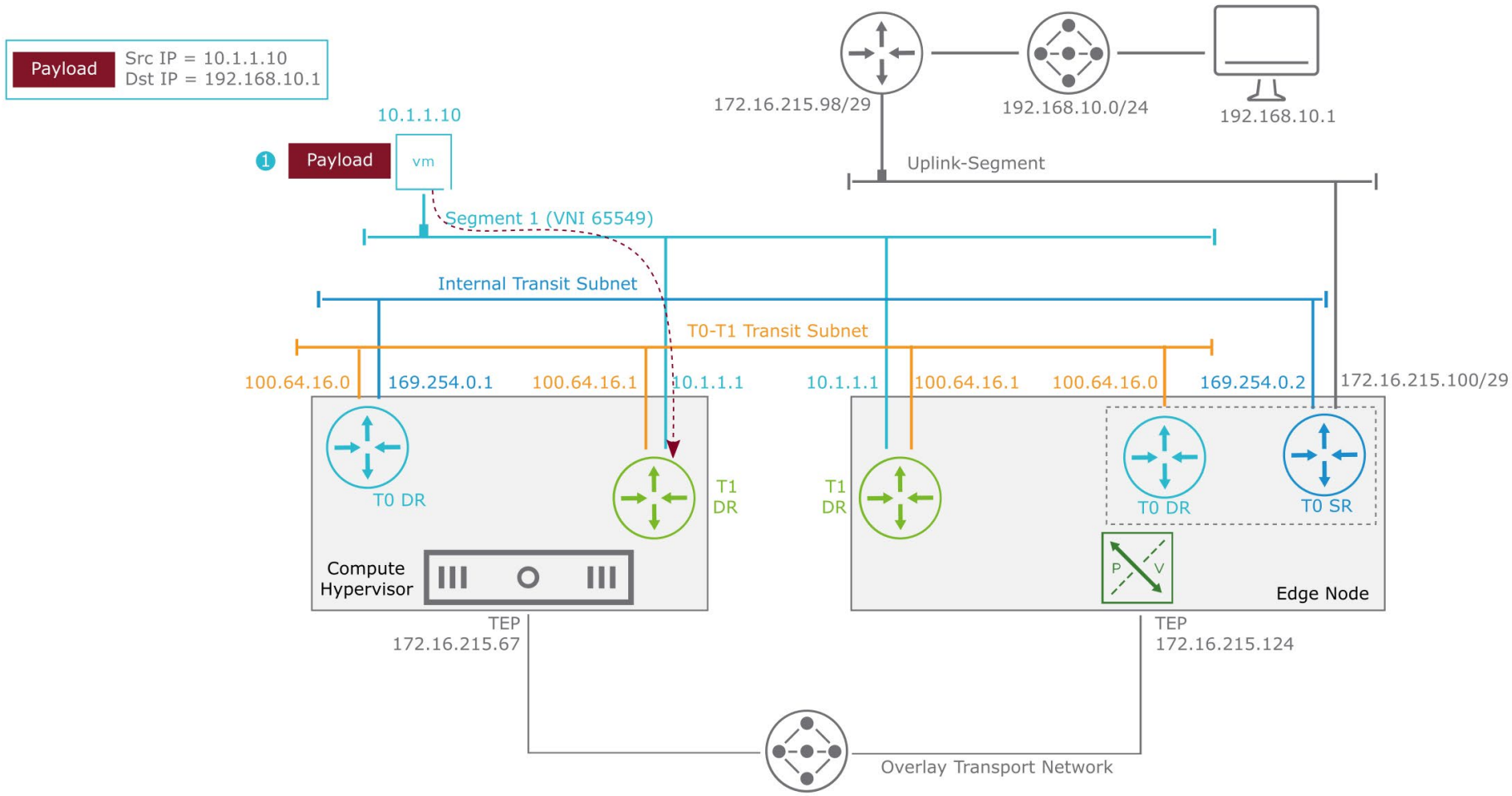
11. The receiving host decapsulates the packet and routes it to its destination (VM 10.1.1.10).



Multitier Routing: Egress to Physical Network (1)

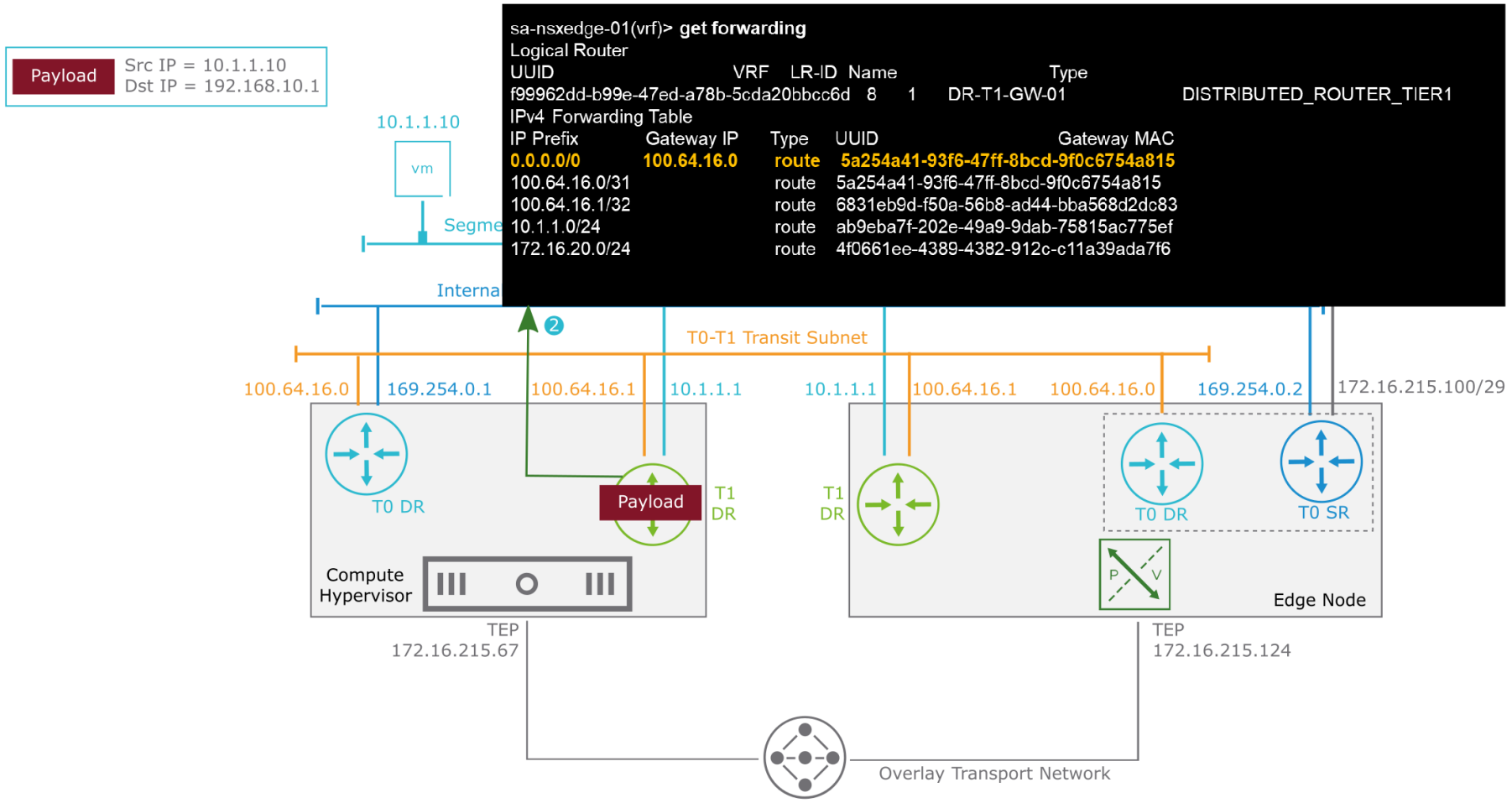
A packet needs to be sent from the source VM 10.1.1.10 to the destination VM 192.168.10.1:

1. The packet is forwarded to its default 10.1.1.1 gateway.



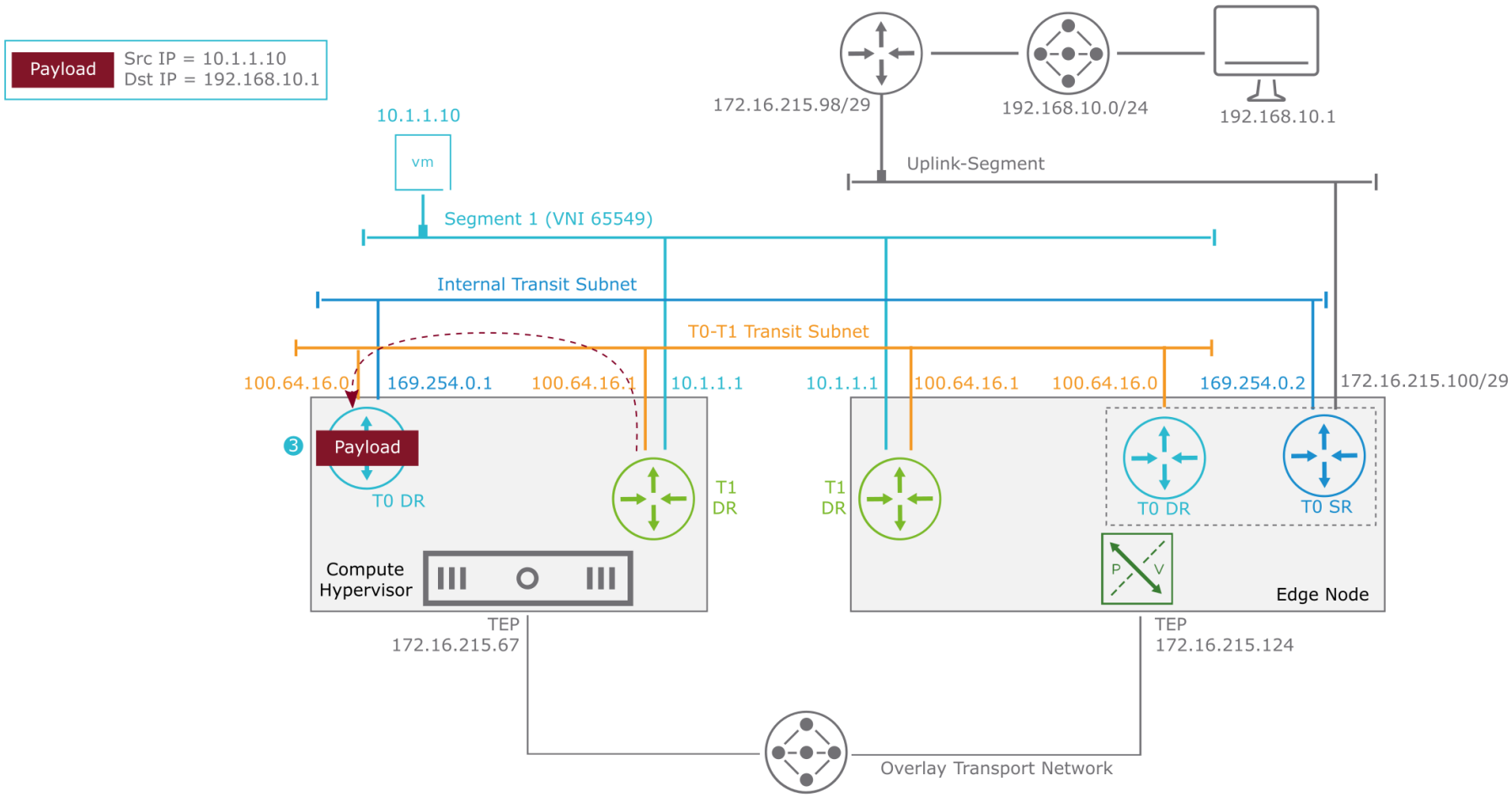
Multitier Routing: Egress to Physical Network (2)

2.The gateway (T1 DR) checks its forwarding table to make a routing decision. Because no specific route exists for the 192.168.10.0/24 network, the packet is sent to the default 100.64.16.0 gateway, which is the DR instance of Tier-0 on the same hypervisor.



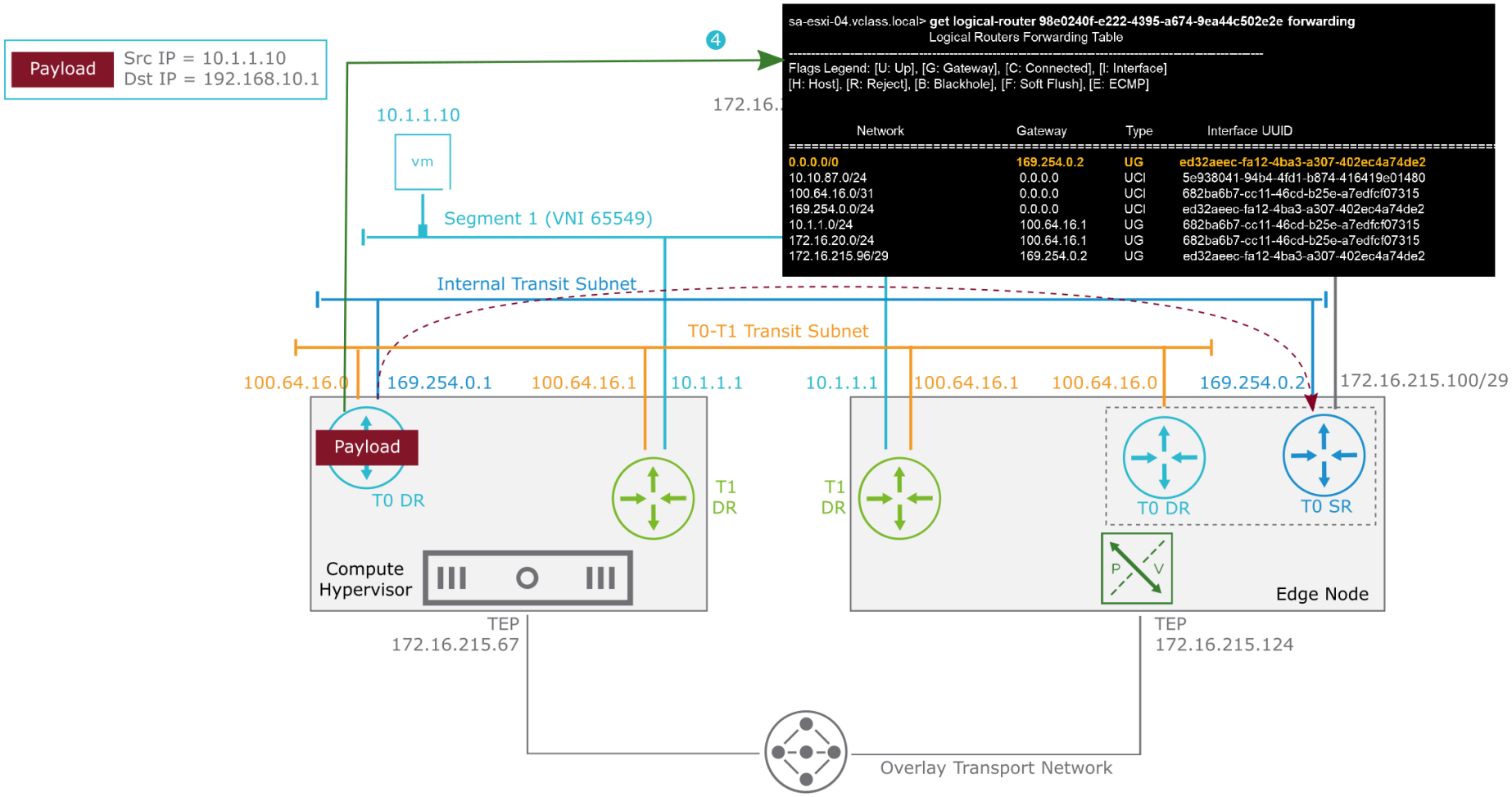
Multitier Routing: Egress to Physical Network (3)

3. The packet is sent to the T0 DR instance on the same hypervisor through T0-T1 Transit Subnet.



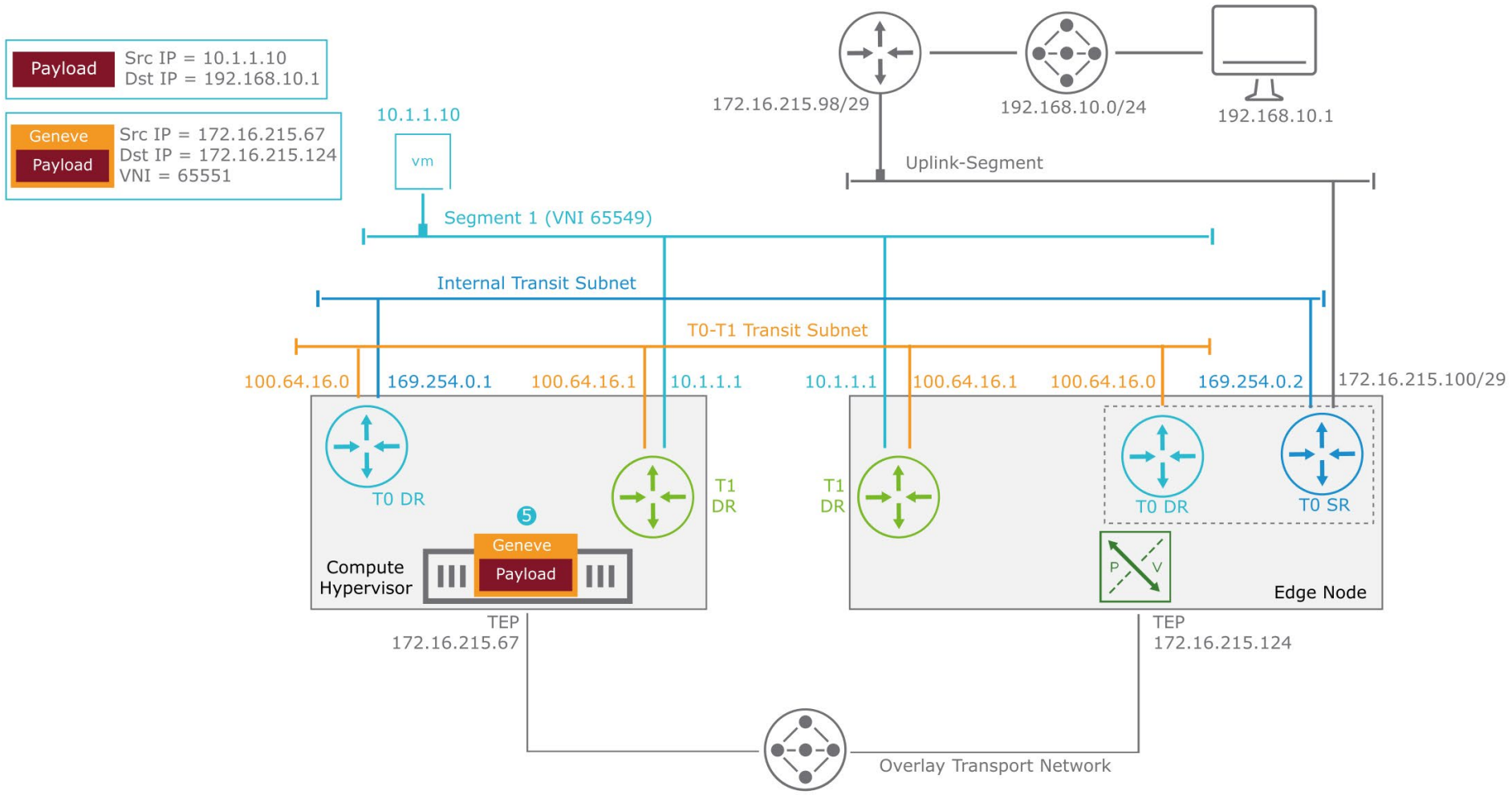
Multitier Routing: Egress to Physical Network (4)

4. The gateway (T0 DR) checks its forwarding table to make a routing decision. The packet is sent to the default 169.254.0.2 gateway, which is the T0 SR component on the edge node.



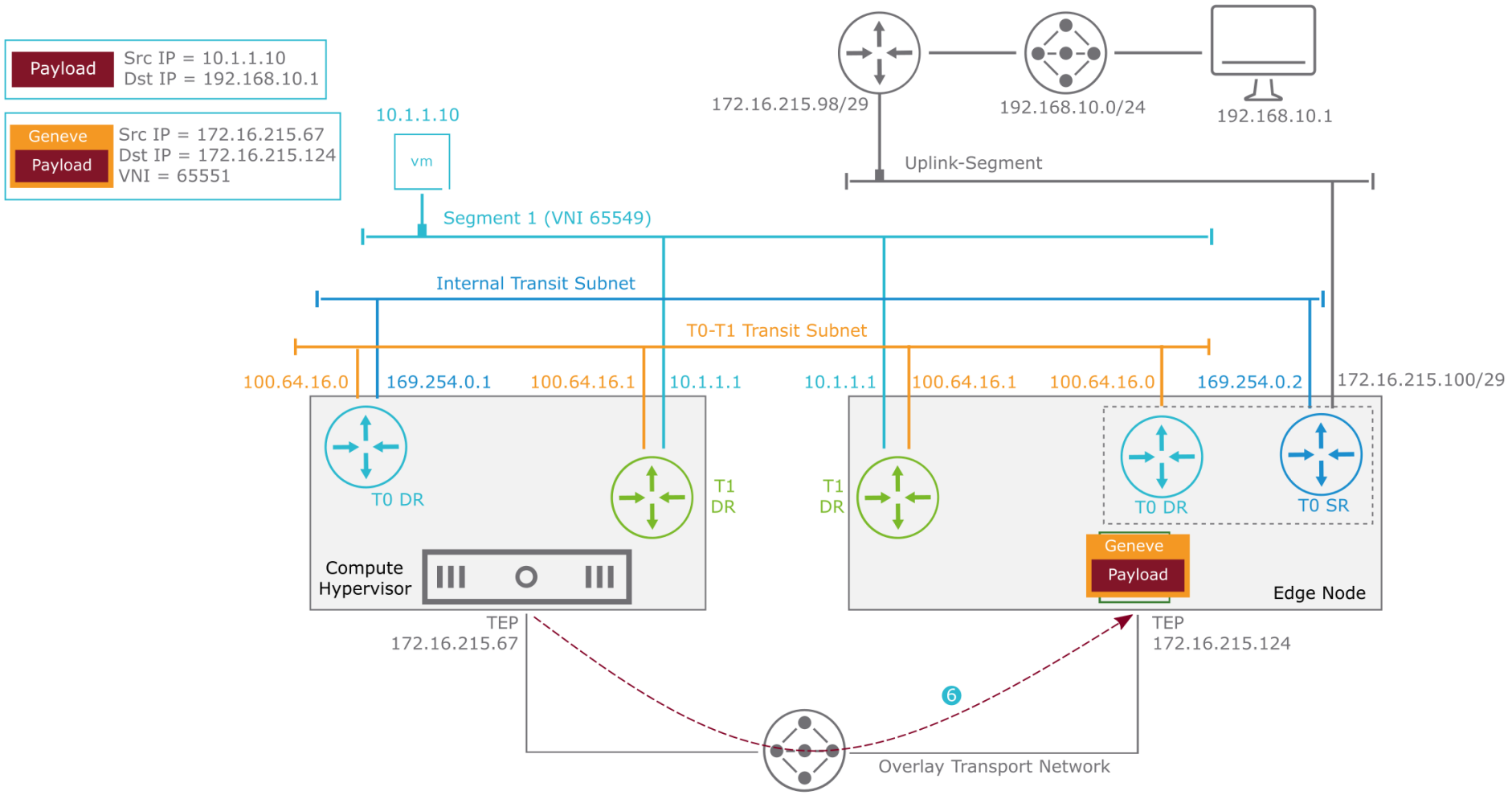
Multitier Routing: Egress to Physical Network (5)

5. To send the packet from the hypervisor to the edge node, the packet is encapsulated with a Geneve header.



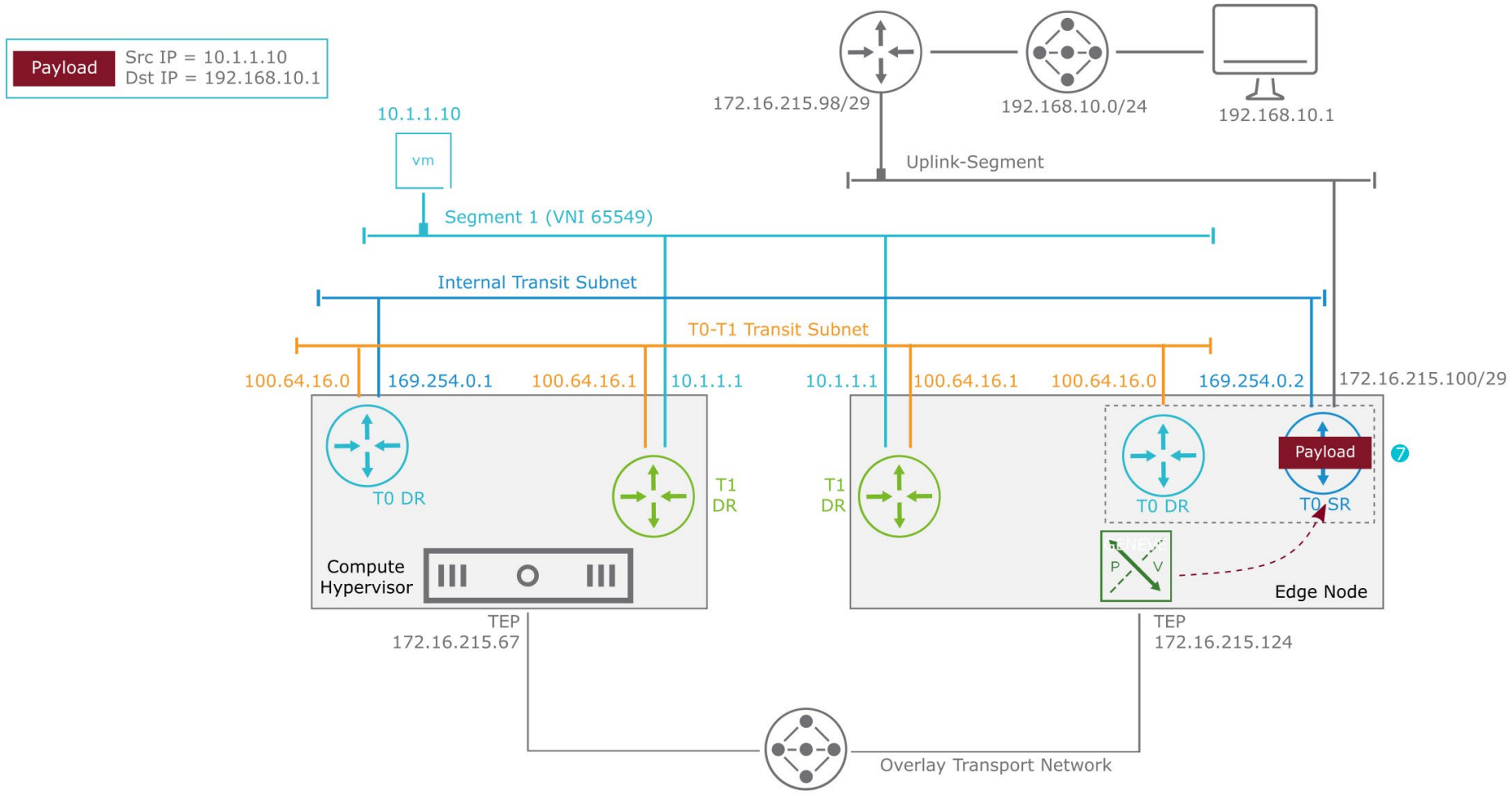
Multitier Routing: Egress to Physical Network (6)

6. The encapsulated packet is sent to the edge node across the overlay tunnel.



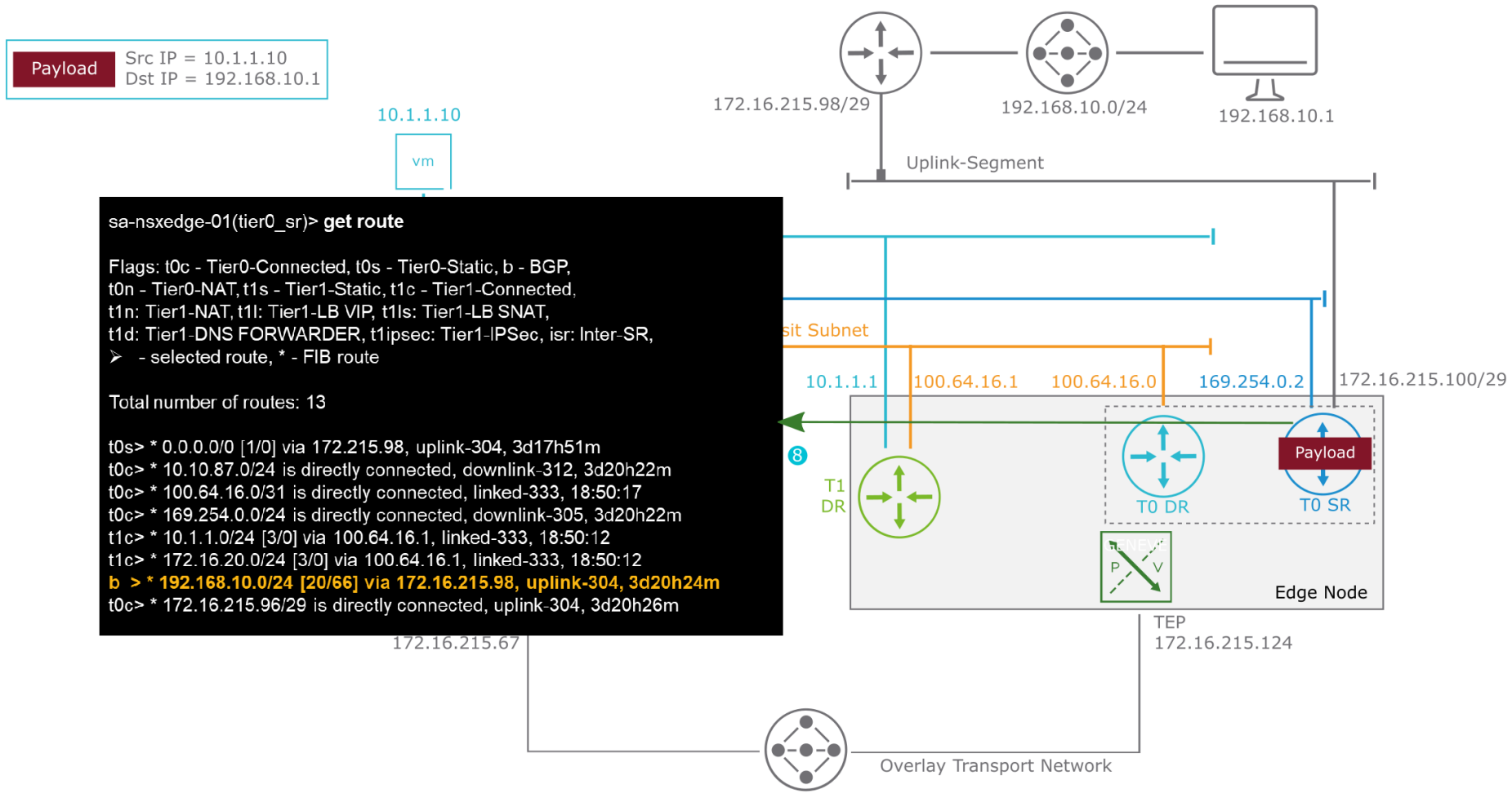
Multitier Routing: Egress to Physical Network (7)

7. The edge node decapsulates the packet and sends it to its T0 SR instance.



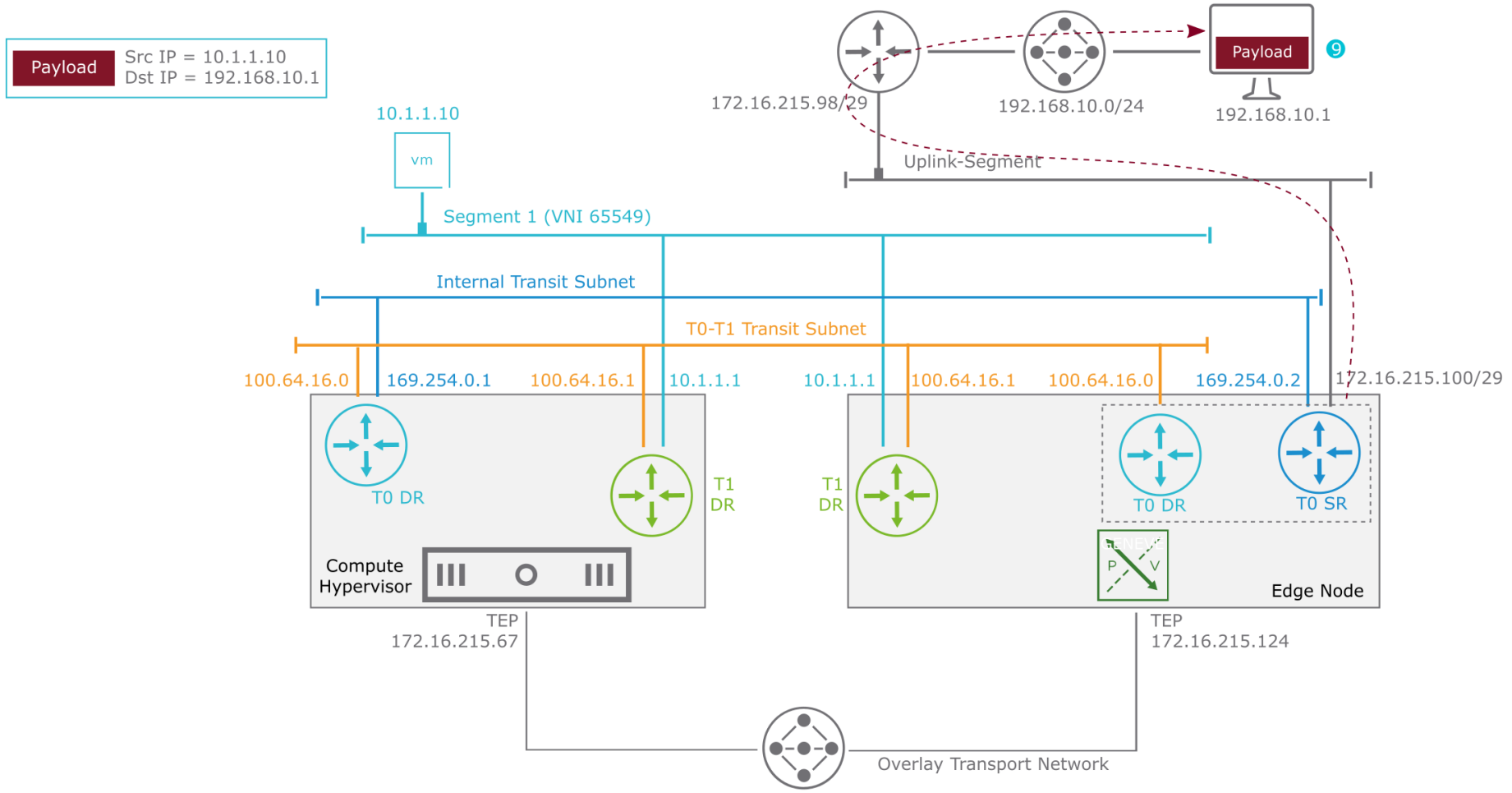
Multitier Routing: Egress to Physical Network (8)

8. The gateway (T0 SR) routing table shows a route for the 192.168.10.0/24 network over the uplink segment.



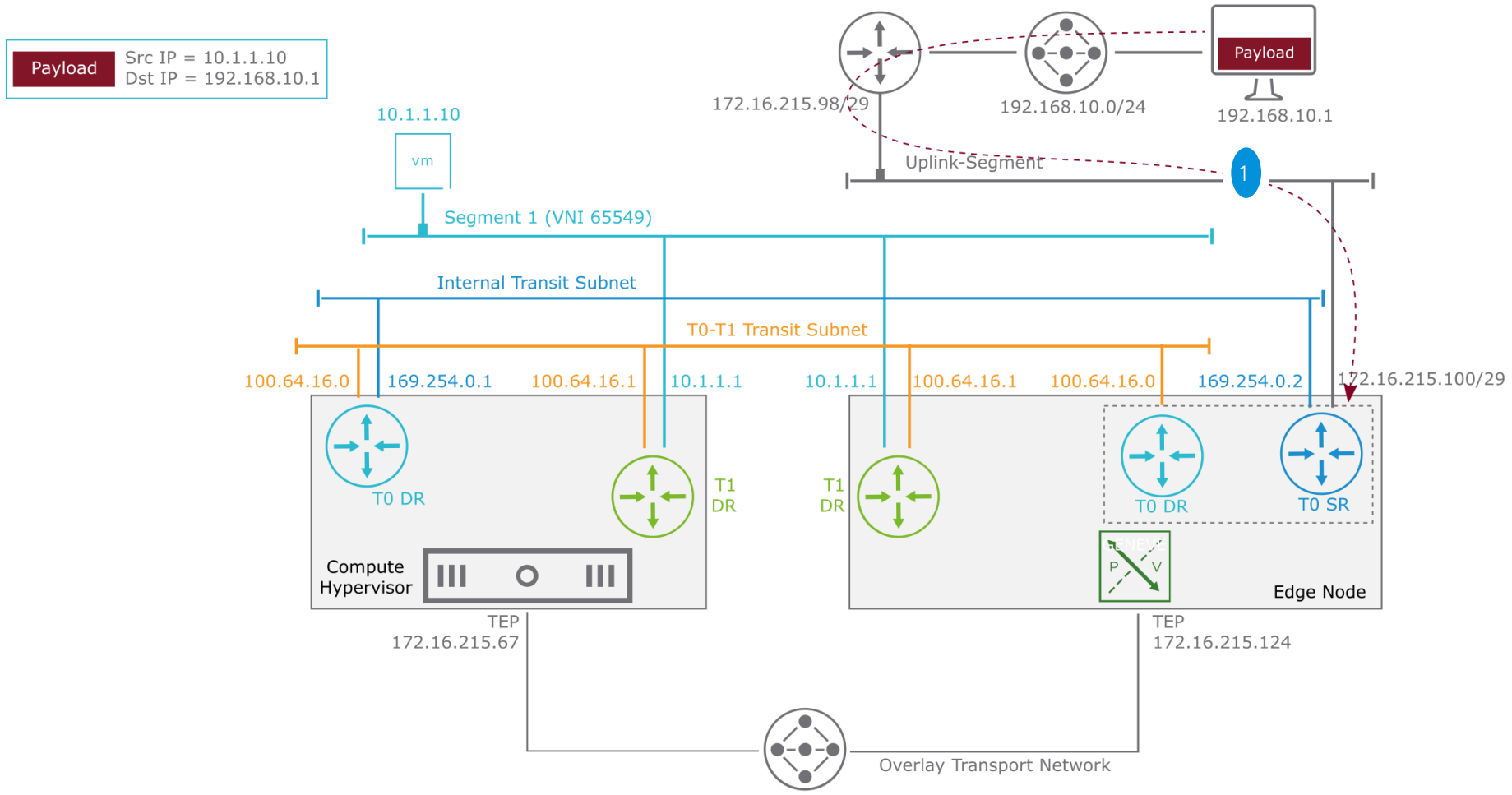
Multitier Routing: Egress to Physical Network (9)

9. The edge node sends the packet to its upstream physical gateway, which routes the packet to its destination, 192.168.10.1.



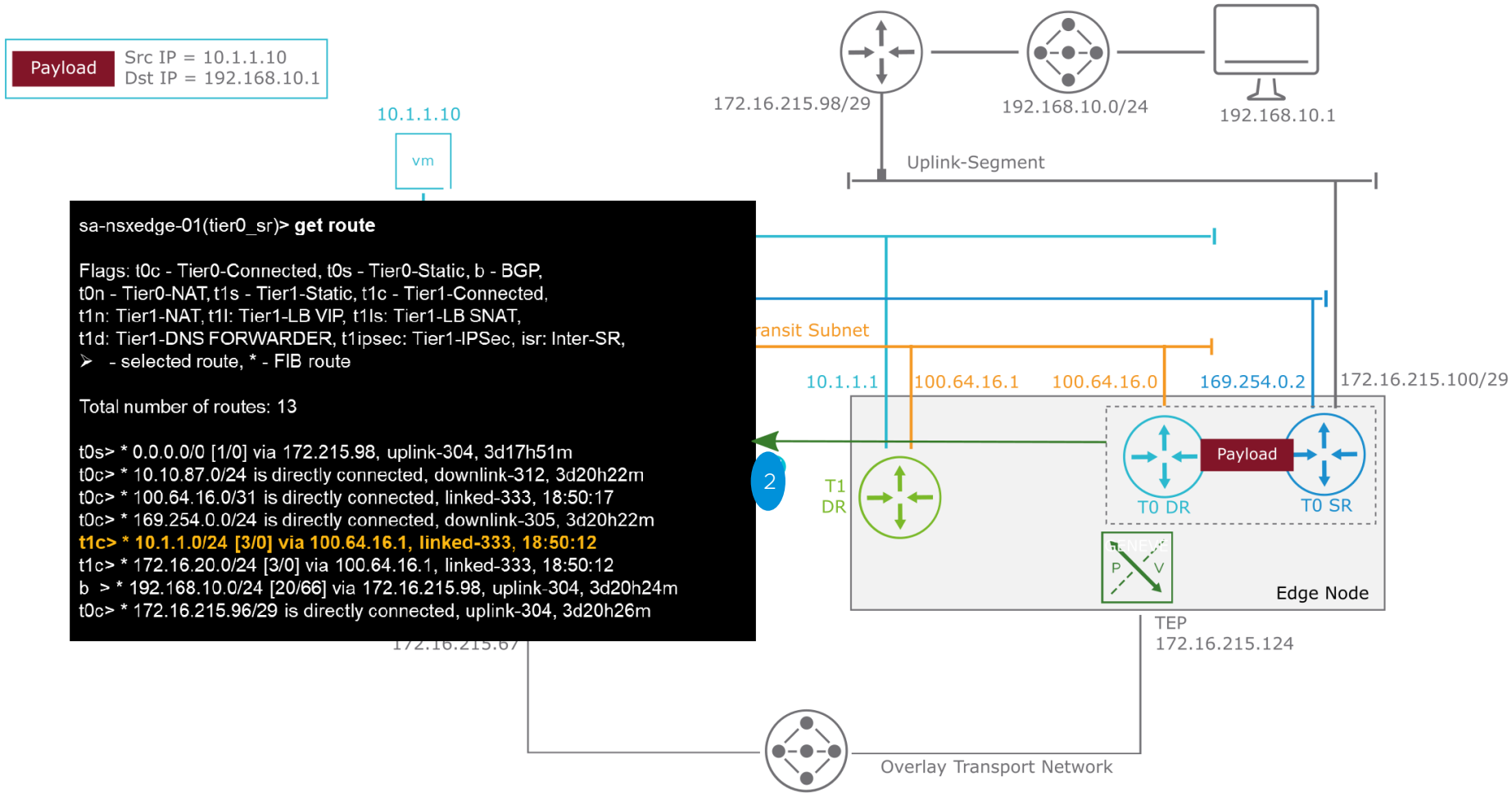
Multitier Routing: Ingress from Physical Network (1)

1. For the return packet, the source VM 192.168.10.1 sends the packet to its default gateway, which routes the packet to the edge node.



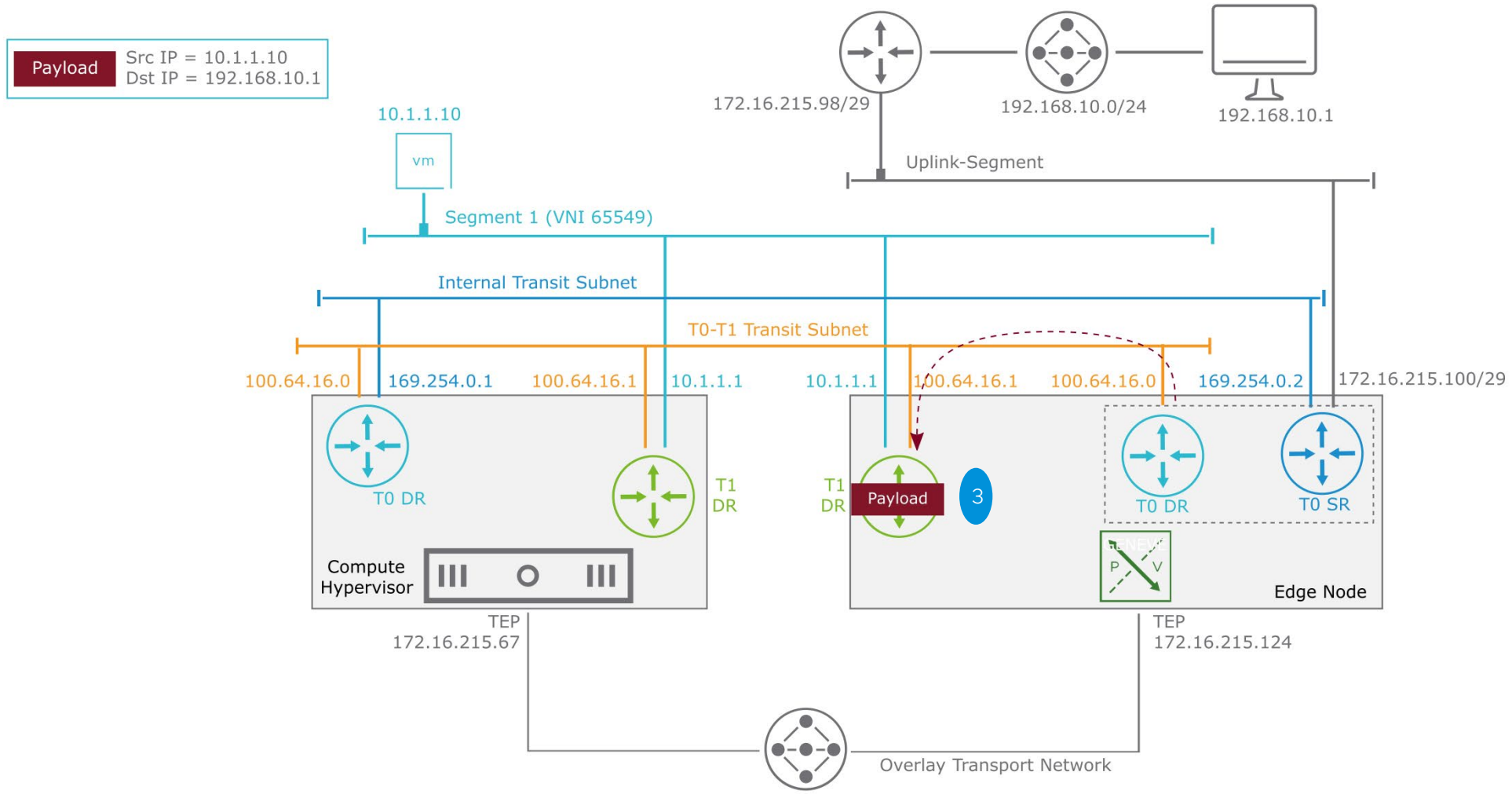
Multitier Routing: Ingress from Physical Network (2)

2. The SR and the DR components of the Tier-0 gateway share their routing table because they are both on the edge node. The routing decision is made to send the packet to the Tier-1 DR instance in the same edge node.



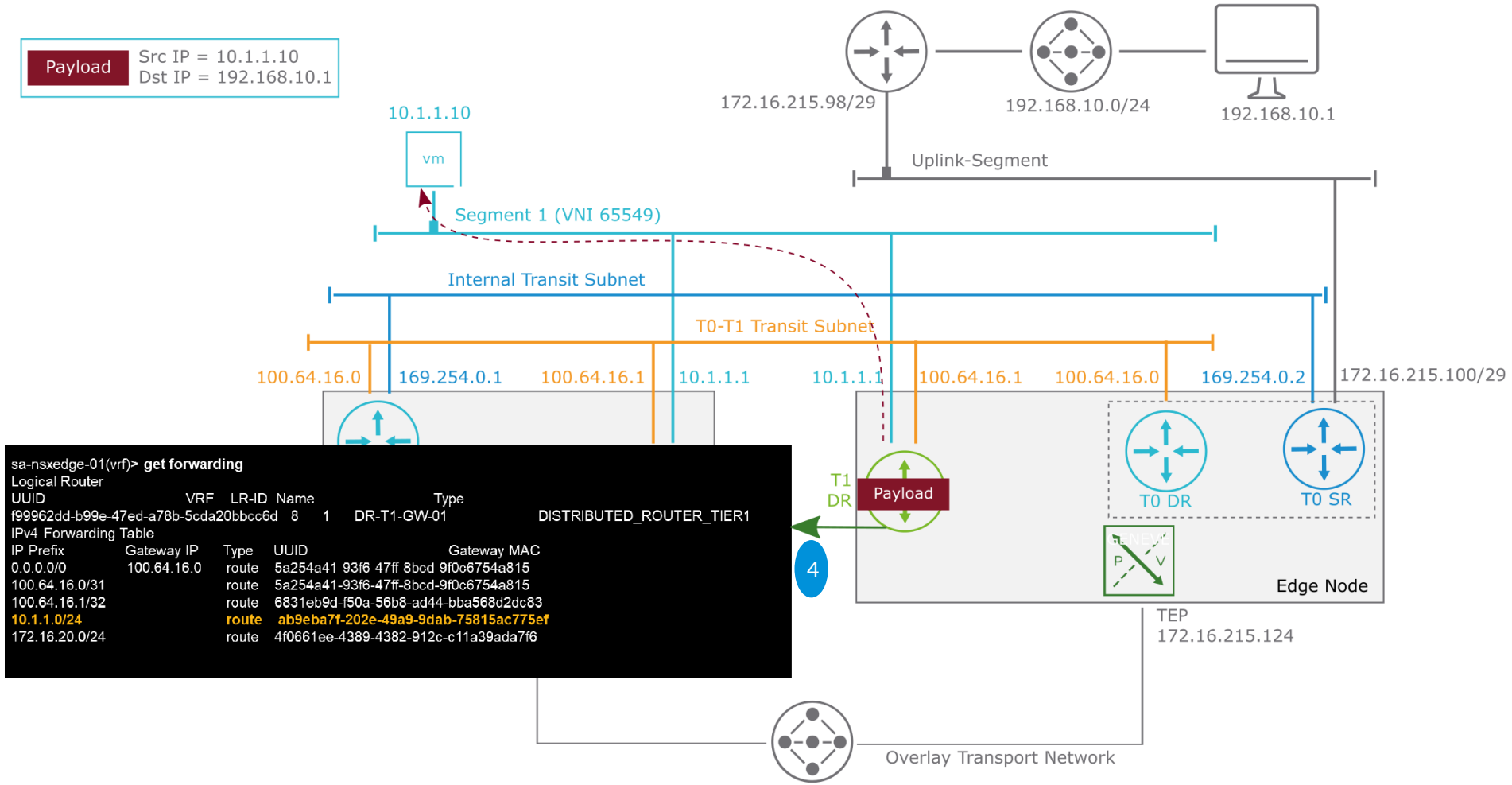
Multitier Routing: Ingress from Physical Network (3)

3. The packet is sent to the T1 DR instance on the edge node through T0-T1 Transit Subnet.



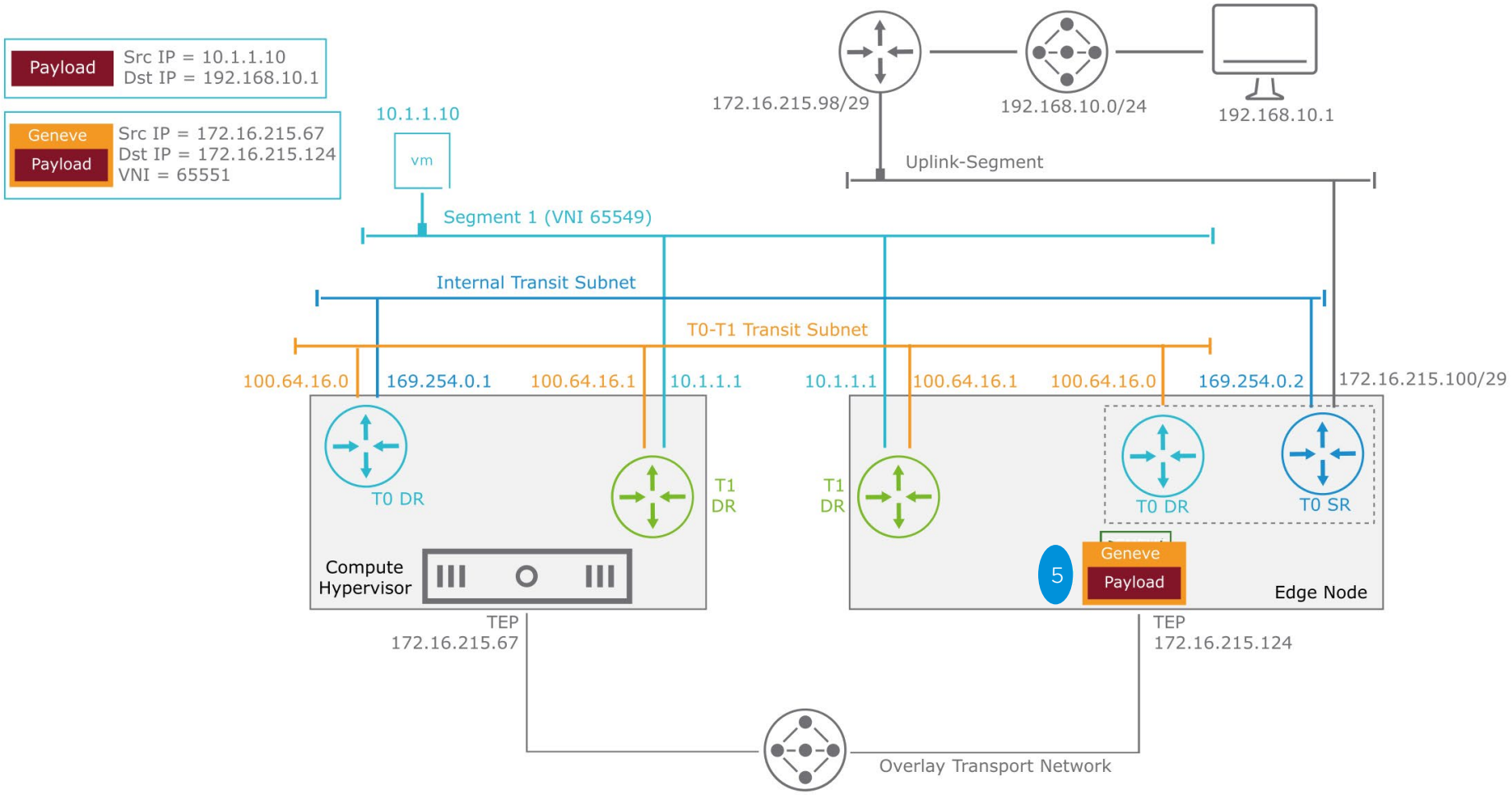
Multitier Routing: Ingress from Physical Network (4)

4. The gateway (T1 DR) checks its forwarding table to make a routing decision. A route is directly connected to the 10.1.1.0/24 network over Segment 1. The packet is sent to the remote host.



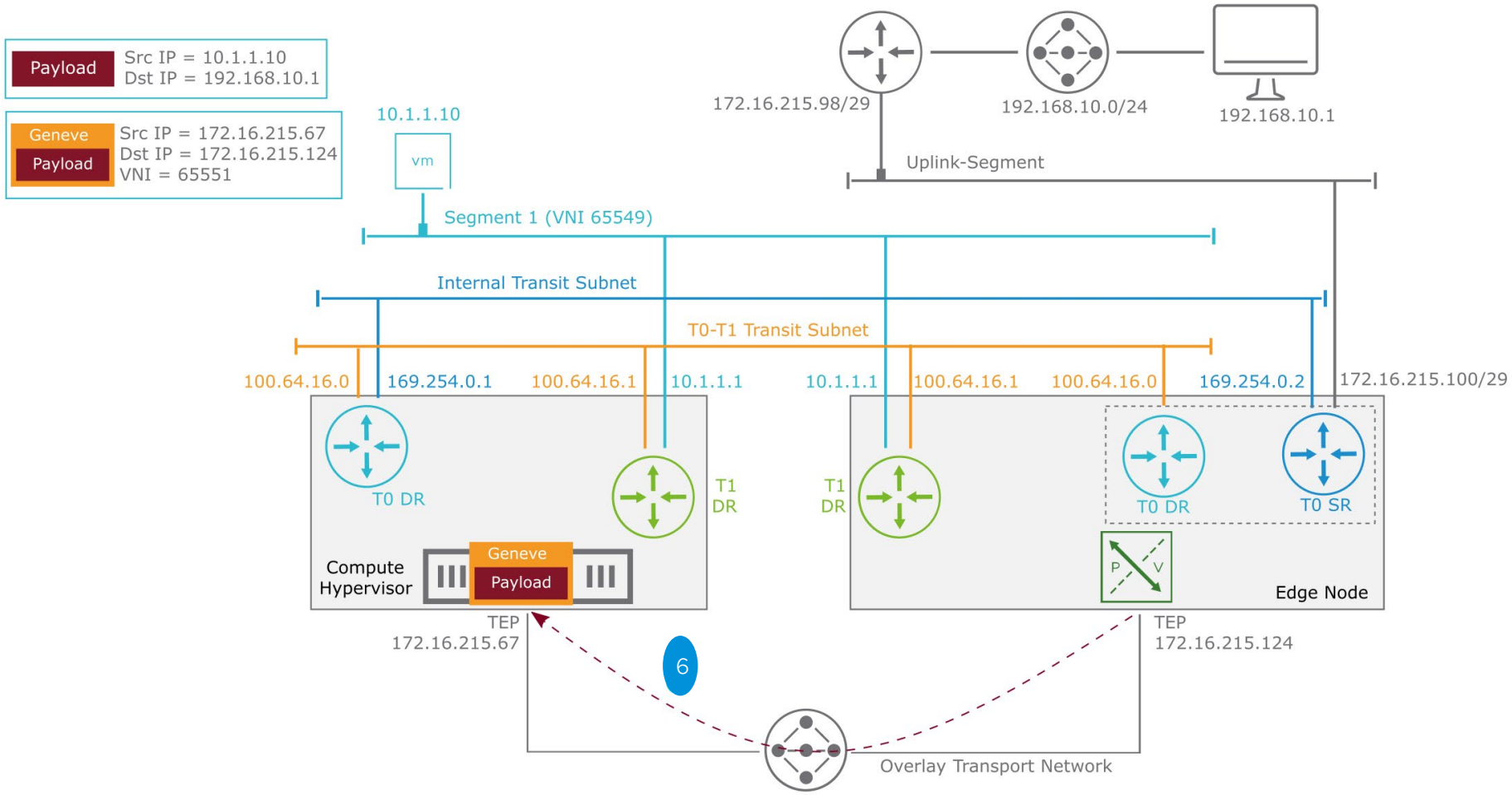
Multitier Routing: Ingress from Physical Network (5)

5. To send the packet from the edge node to the hypervisor, the packet is encapsulated with a Geneve header.



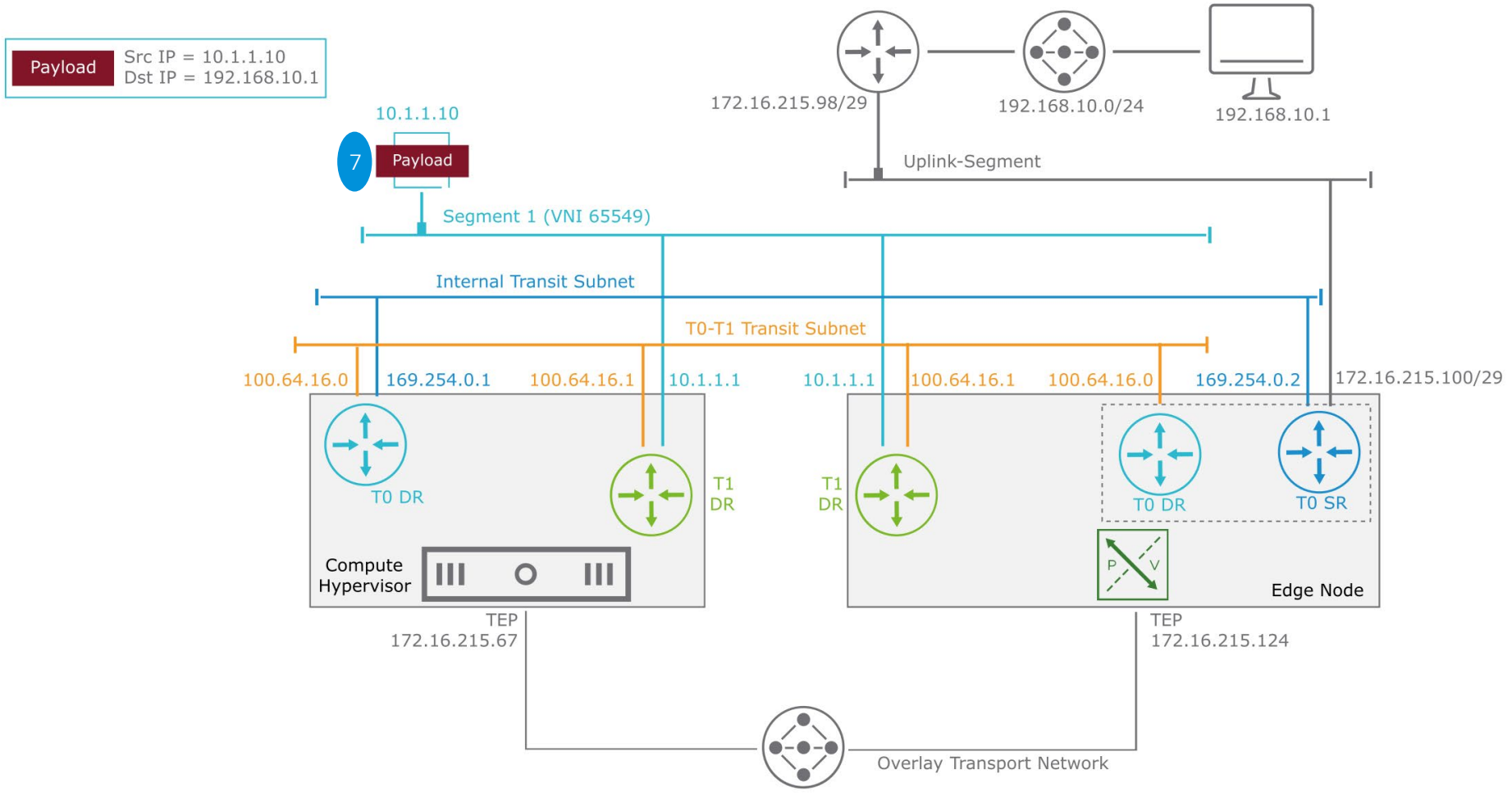
Multitier Routing: Ingress from Physical Network (6)

6. The encapsulated packet is sent to the edge node across the overlay tunnel.



Multitier Routing: Ingress from Physical Network (7)

7. The receiving host decapsulates the packet and routes it to its destination (VM 10.1.1.10).



Investigating More Deeply

What else can we see?

Uplink Configuration: Edge CLI Validation

Use the **`get interfaces`** command in the VRF context of the Tier-0 service router to get the uplink interface-related information.

```
sa-nsxedge-01> get logical-routers
Logical Router
UUID VRF LR-ID Name...
9ffdac61-d645-4b2d-957e-ef1e422767a7 14 11266 SR-Prod-T0-GW-01...

sa-nsxedge-01> vrf 14
sa-nsxedge-01(tier0_sr)> get interfaces
Interface :          685eab8a-b860-4411-9919-92b7f70e7048
Ifuid :              376
Name :               Uplink-01-Intf
Fwd-mode :           IPV4_ONLY
Internal name :      uplink-376
Mode :               lif
Port-type :          uplink
IP/Mask :            192.168.100.2/24
...
```

Tier-0 and Tier-1 Connection: Edge CLI Verification

Use the **`get interfaces`** command in the VRF context of the Tier-1 service router to get the uplink interface used to connect to the Tier-0 gateway.

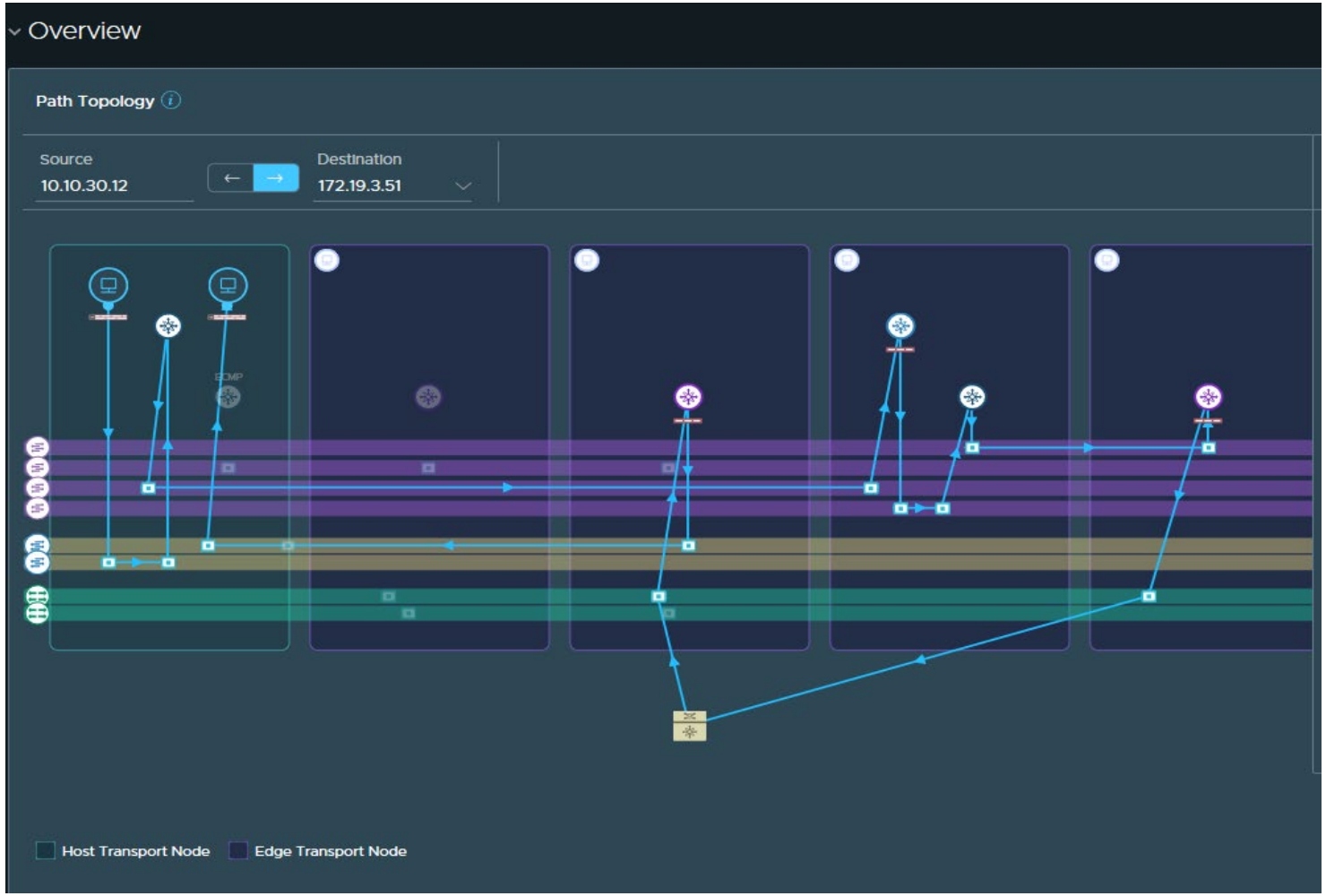
```
sa-nsxedge-01> get logical-routers
Logical Router
UUID VRF LR-ID Name...
a2f27e39-5b4c-4c4e-a40f-ea7bb4b3e434 12 11265 SR-Prod-T1-GW-01...

sa-nsxedge-01> vrf 12
sa-nsxedge-01(tier1_sr)> get interfaces
Interface :          afe905c7-9c8c-47cb-a213-2f79240d5b14
Ifuid :              382
Name :               Prod-T0-GW-01-Prod-T1-GW-01-t1_
Fwd-mode :           IPV4_ONLY
Mode :               lif
Port-type :          uplink
IP/Mask :            100.64.48.0/31
...
```

Optimizing and Troubleshooting Virtual and Physical Networks

Use Aria Operations for Networks to optimize and troubleshoot your networks:

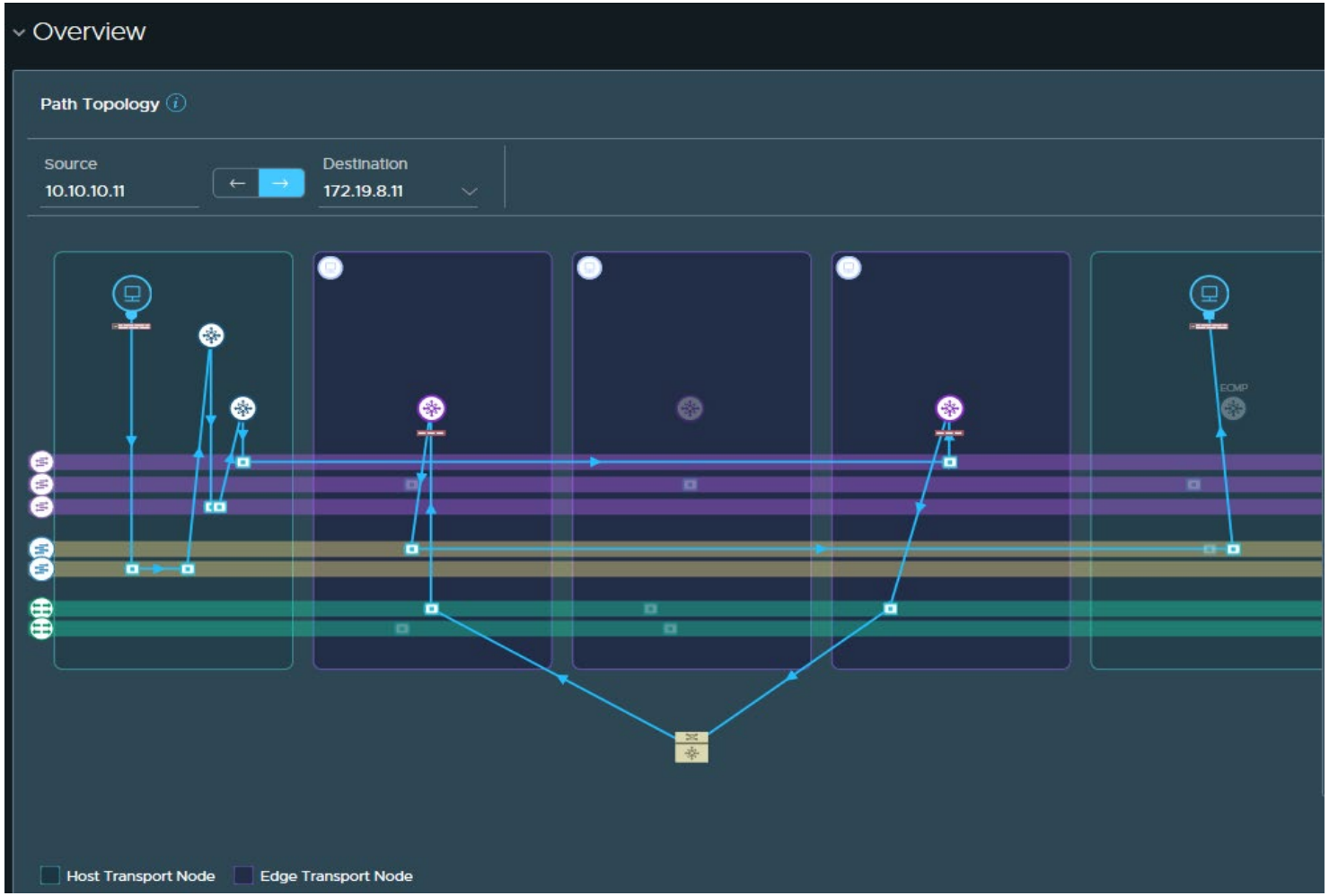
- Virtual and physical network topology mapping
- Performance optimization across overlay and underlay
- Correlated problems and performance metrics
- Firewall rules and security policies across NSX and third-party devices



Optimizing and Troubleshooting Virtual and Physical Networks

Use Aria Operations for Networks to optimize and troubleshoot your networks:

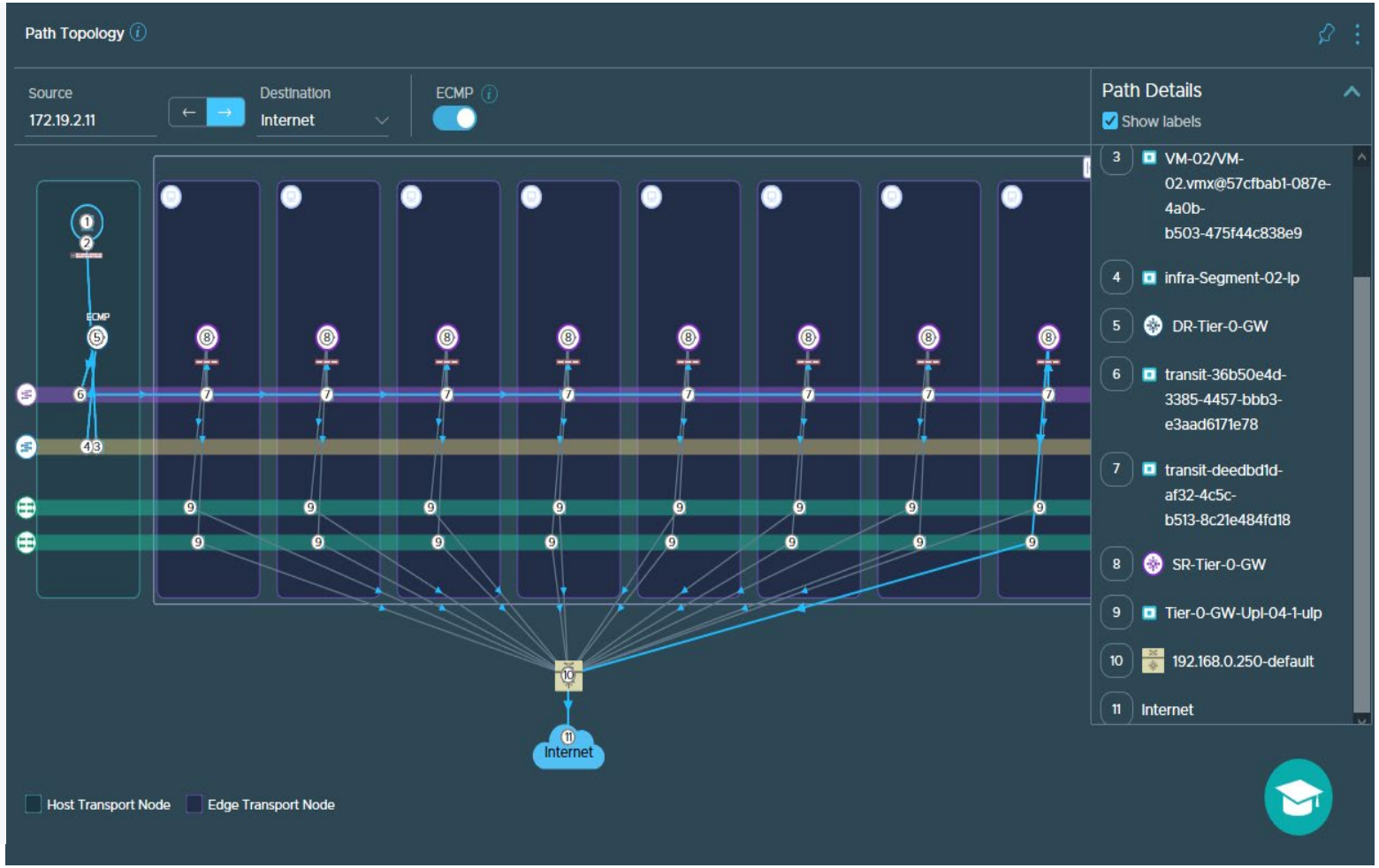
- Virtual and physical network topology mapping
- Performance optimization across overlay and underlay
- Correlated problems and performance metrics
- Firewall rules and security policies across NSX and third-party devices



Optimizing and Troubleshooting Virtual and Physical Networks

Use Aria Operations for Networks to optimize and troubleshoot your networks:

- Virtual and physical network topology mapping
- Performance optimization across overlay and underlay
- Correlated problems and performance metrics
- Firewall rules and security policies across NSX and third-party devices



vmware EXPLORE

Thank You