# BLT
# BGP Label based Tunneling

Gunter Van de Velde
Sr Technical Leader
NOSTG, Cisco Systems

December 2012

# What is it?

- It is a new “IP” network overlay technology
- Can be seen as alternative for MPLS based upon BGP as scalable and proven control plane
- Within the Core ISP network, no more need for full routing table
- The network overlay mechanism is tunnel technology agnostic (LISP, GRE, L2TP, IPinIP, VxLAN, etc)
- BGP based Dynamic tunnelling works Intra- and Inter-domain
- Expected convergence time: <100msec

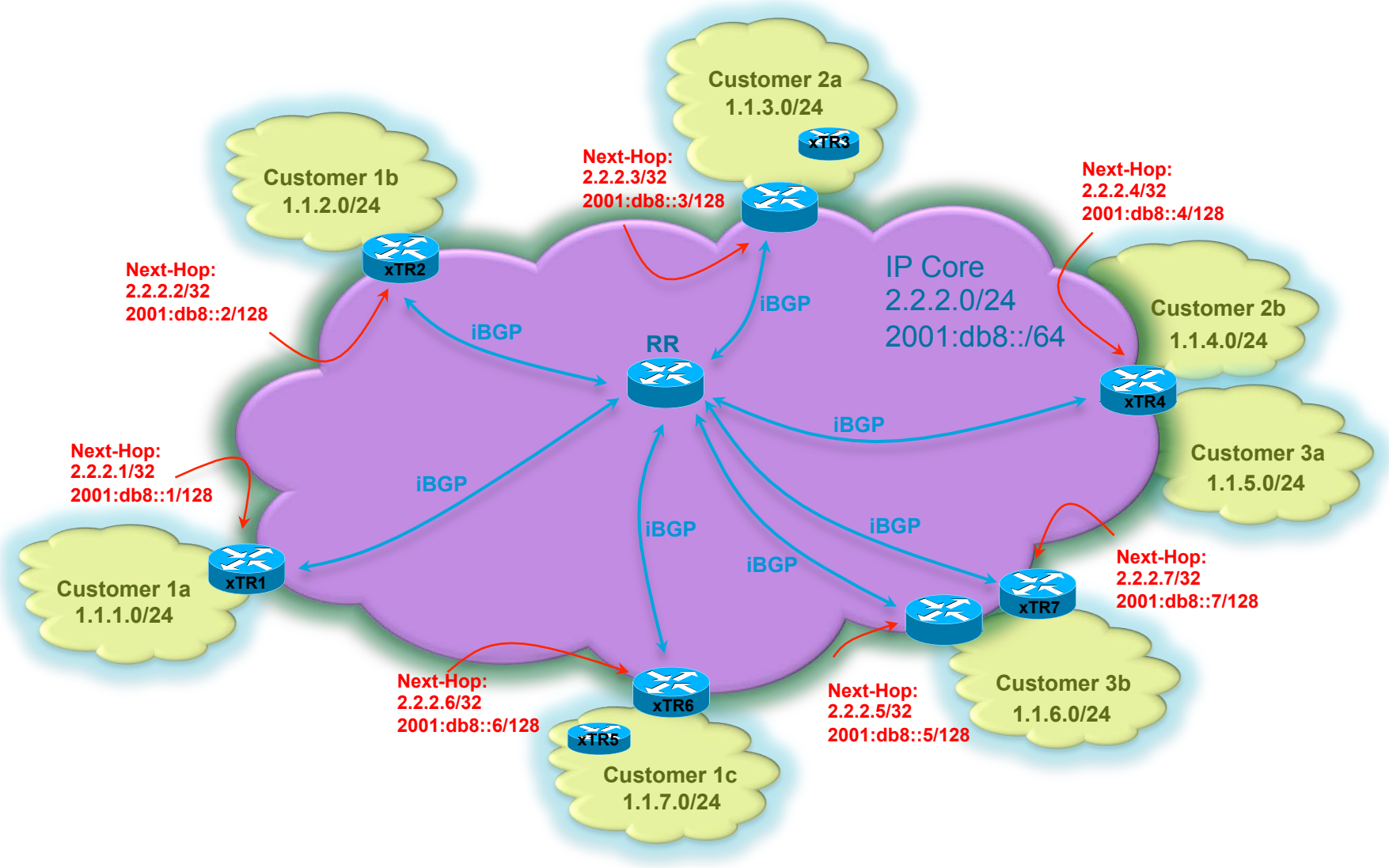
# Motivation?

- Address family (IPv4, IPv6, VPNv4, VPNv6, IP+Label) agnostic
- Usage of proven and highly scalable Internet technologies (BGP, PIC, LFA, etc…)
- BGP database carries tunnel end-points and identifiers as ships in the night to the traditional BGP routing BGP table
- Cost optimization by getting rid of:
  - Core MPLS control plane
  - Internet and customer prefixes from core
- Usage of BGP technology:
  - Fast Convergence
  - High scalability
  - High availability
  - VPN Support
  - Highly secure by utilisation of BGP security technologies (RPKI Origin Authentication, TCP-AO, etc..)
  - BGP Remote-Next-Hop (http://datatracker.ietf.org/doc/draft-vandevelde-idr-remote-next-hop/)
- Usage of VxLAN, GRE, LISP, IP-in-IP tunnels
  - Utilization of scalable and existing tunnel technology
  - Connect IPv6/IPv4 islands over an IPv6/IPv4 infrastructure for both Inter- and Intra-AS networks
  - Utilization of existing tunnel policy and RIB population mechanisms
  - Service differentiation: enable premium exit vs best-effort exit to Internet by Network Policy
- Backward compatible and support for gradual implementation

# Toolset for BGP based Dynamic Tunnelling

- BGP Remote-Next-Hop ( http://tools.ietf.org/html/draft-vandevelde-idr-remote-next-hop)
- LISP - http://tools.ietf.org/doc/draft-ietf-lisp/
- Other tunnel technologies: GRE, VxLAN, IP-in-IP, etc…
- BGP Route-Reflection (RFC4456)
- Cisco Prefix Independent Convergence
- BGP Diverse Path (RFC6774)
- BGP Add-Path (http://tools.ietf.org/html/draft-ietf-idr-add-paths)
- BGP/MPLS VPN (RFC4364)

# Address Distribution

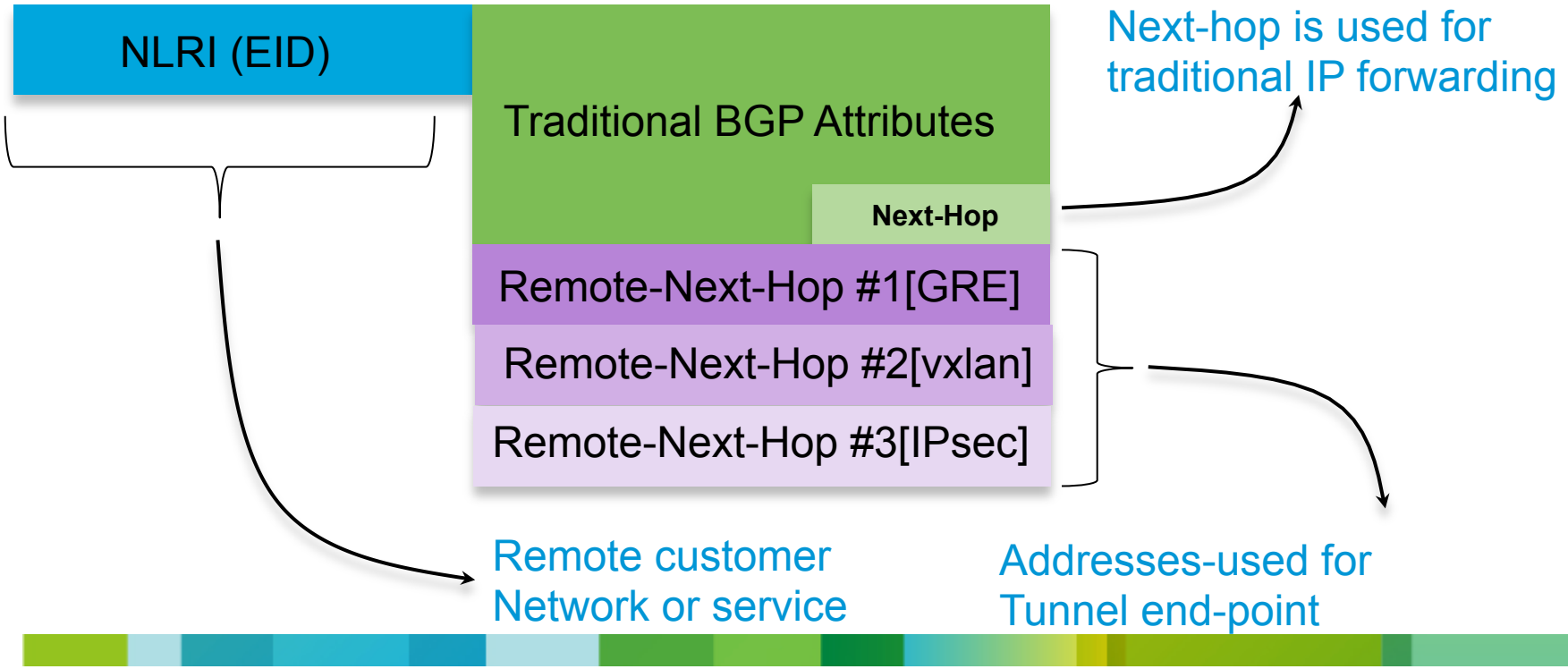


# Address Distribution

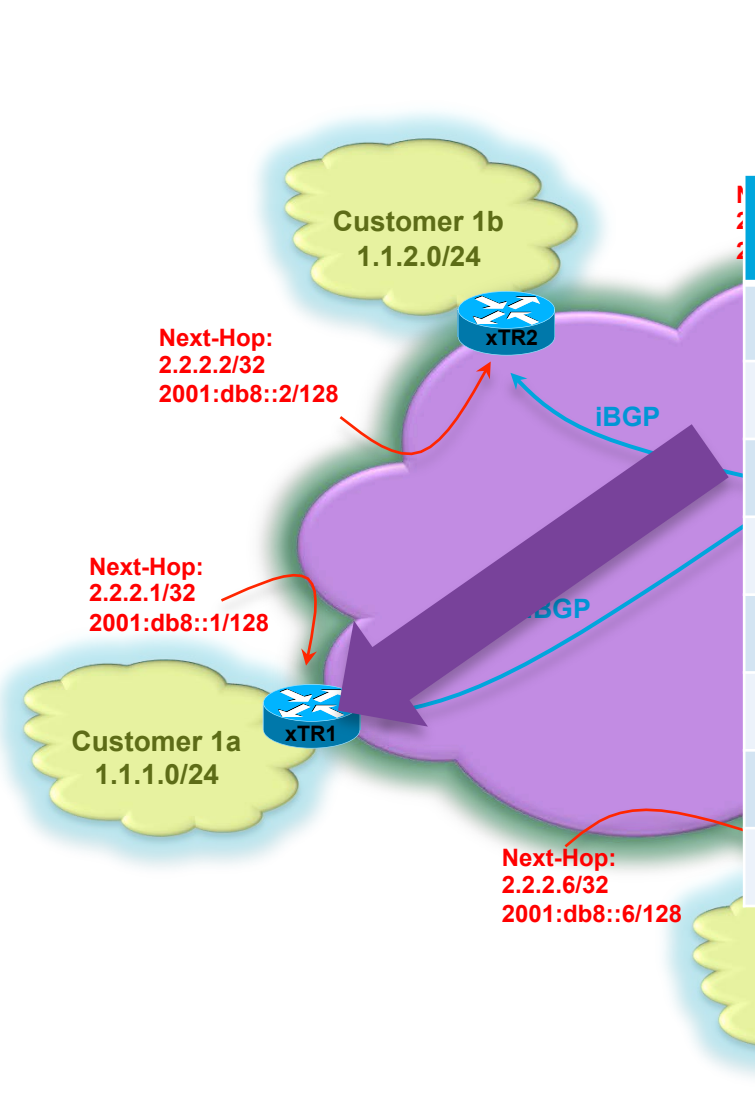
- Core
  - IGP: OSPF, EIGRP, ISIS
  - MPLS Free Core
  - BGP only is run only on the core edge and BGP RR
  - support of IGP LFA
- Edge
  - Location of the Tunnel in-/egress router
  - BGP NLRI is used as remote network identifier and the attached BGP Remote-Next-Hop as Locator
  - Forwarding in-/egress policy enforcement
  - Multi-tunnel loadsharing
- Customer Networks
  - Autonomous networks
    - DC, finance, IT department, engineering, customers, etc…
    - Independent address family agnostic address space
  - Customer networks and services are network identifiers

# BGP Remote-Next-Hop Attribute

- NLRI (Network Layer Reachability Information) is the customer network
- Next-hop is the traditional BGP Next-Hop used for traditional IP forwarding
- Remote-Next-Hop is the Tunnel End-Point used for dynamic tunnel based forwarding (Optional BGP transitive attribute)
- Multiple NLRI can point to identical Remote-Next-Hop



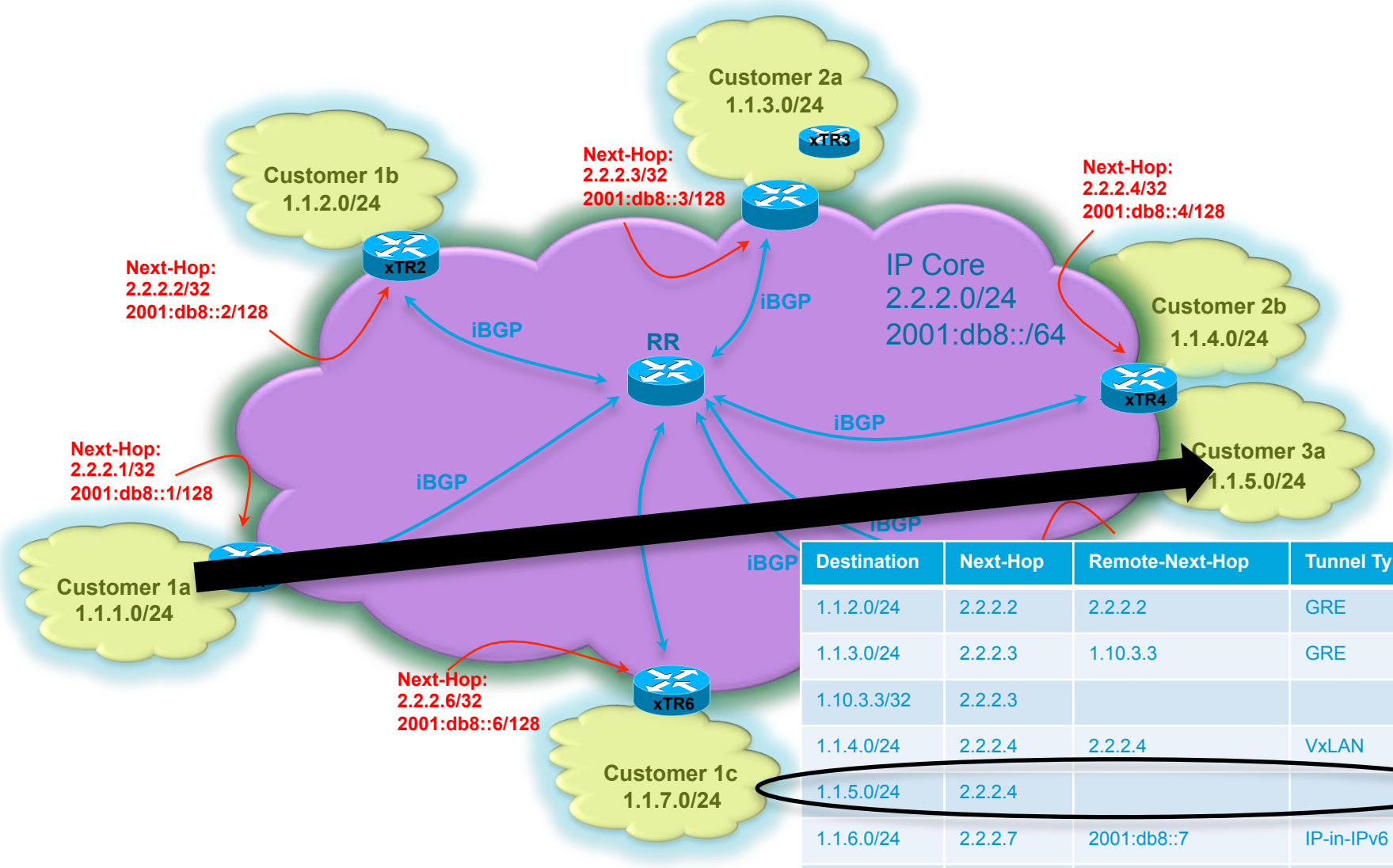
# Address Distribution: BGP Table at xTR1




## BGP Table

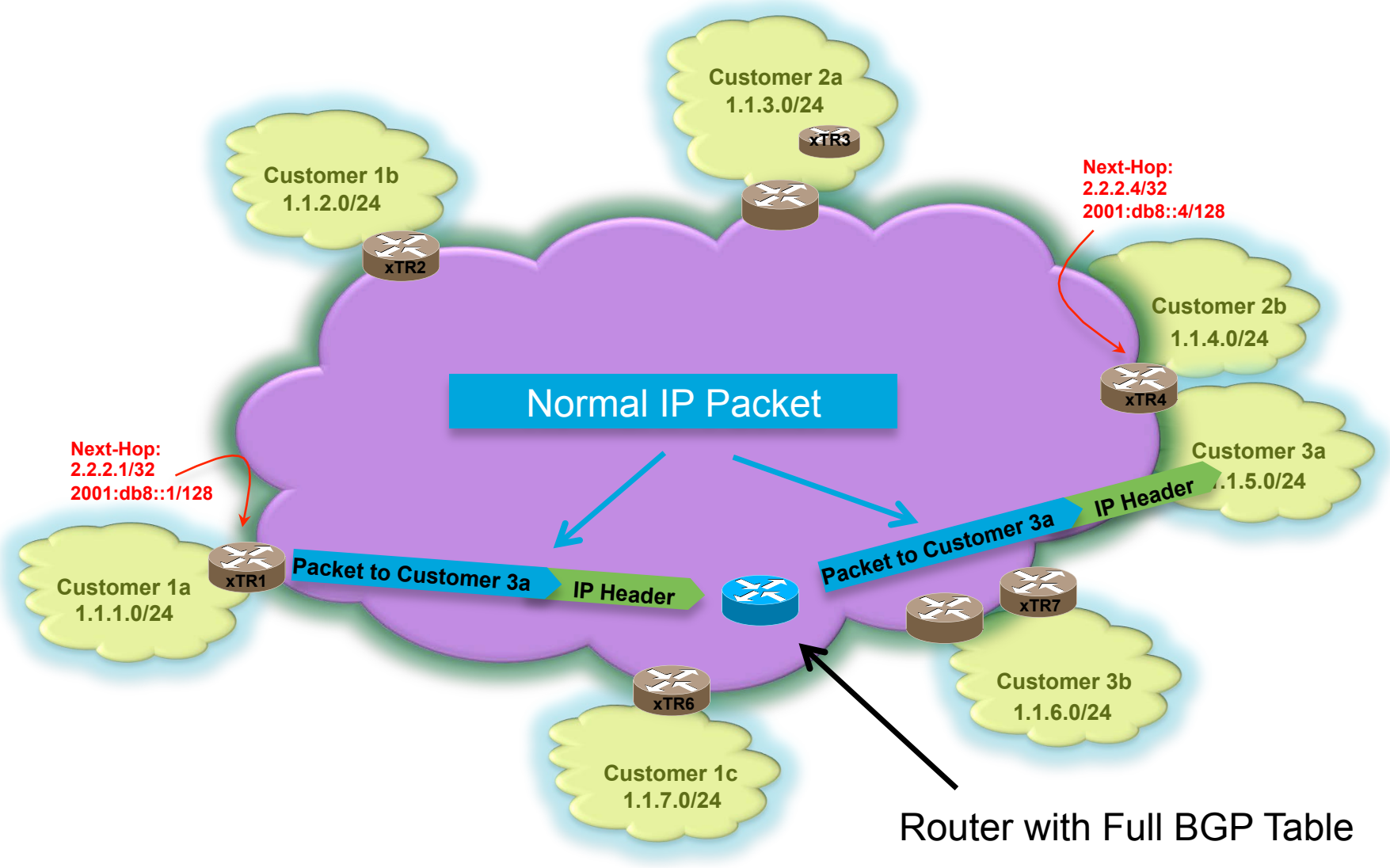
| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

# Traditional BGP Forwarding



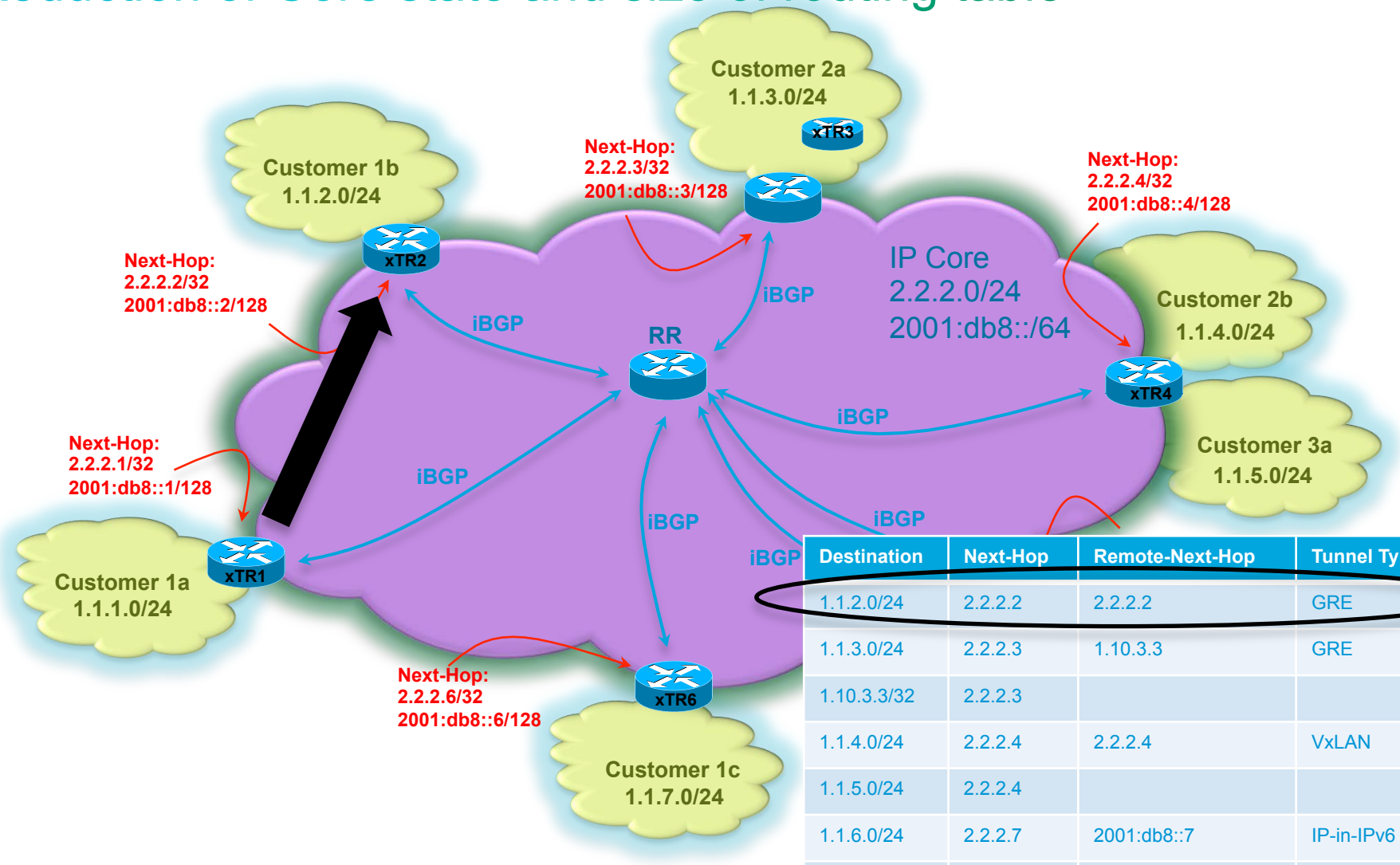

| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

# Traditional BGP Forwarding



# Tunnel Based Forwarding: Case 1

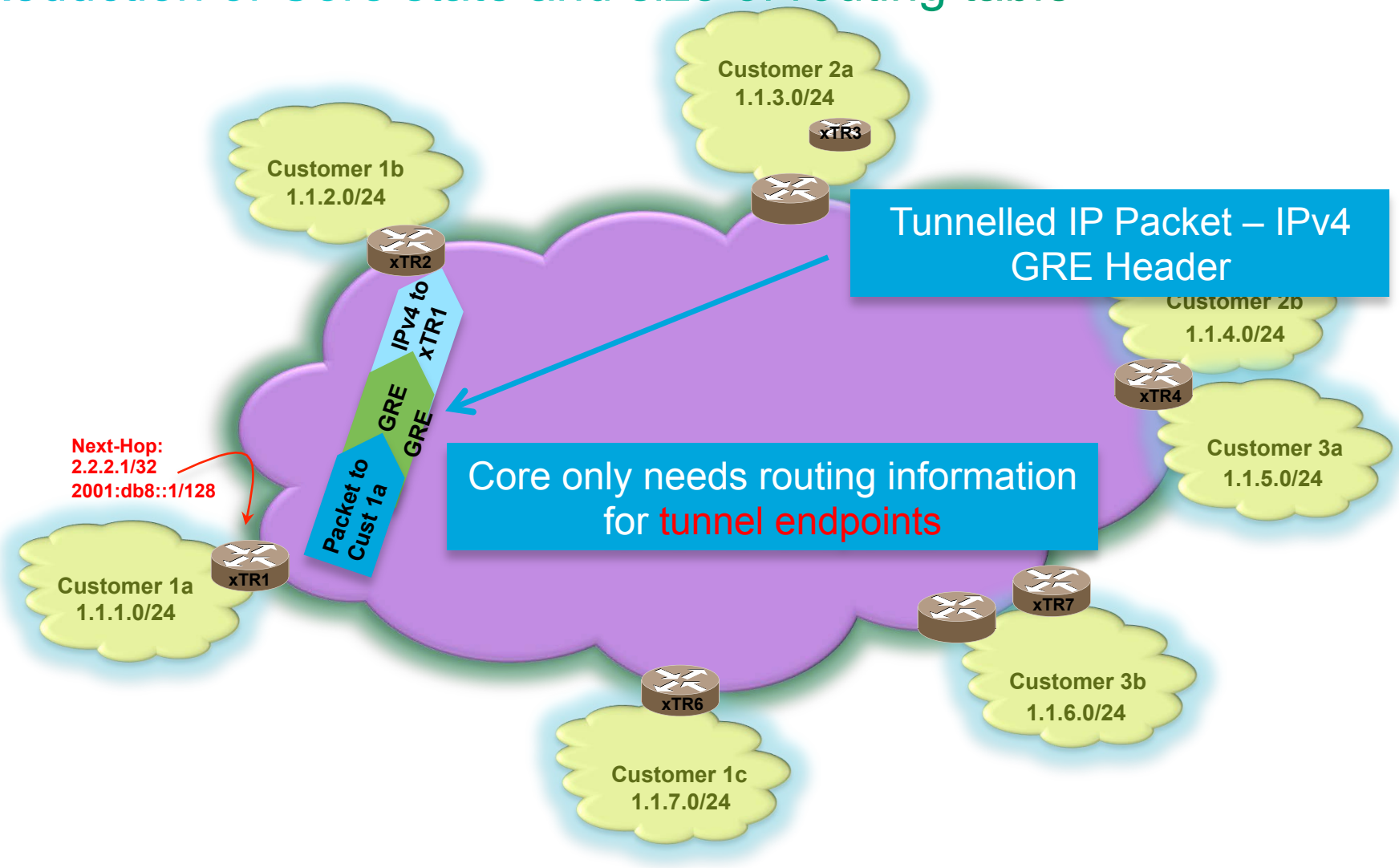
## Reduction of Core state and size of routing table




| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

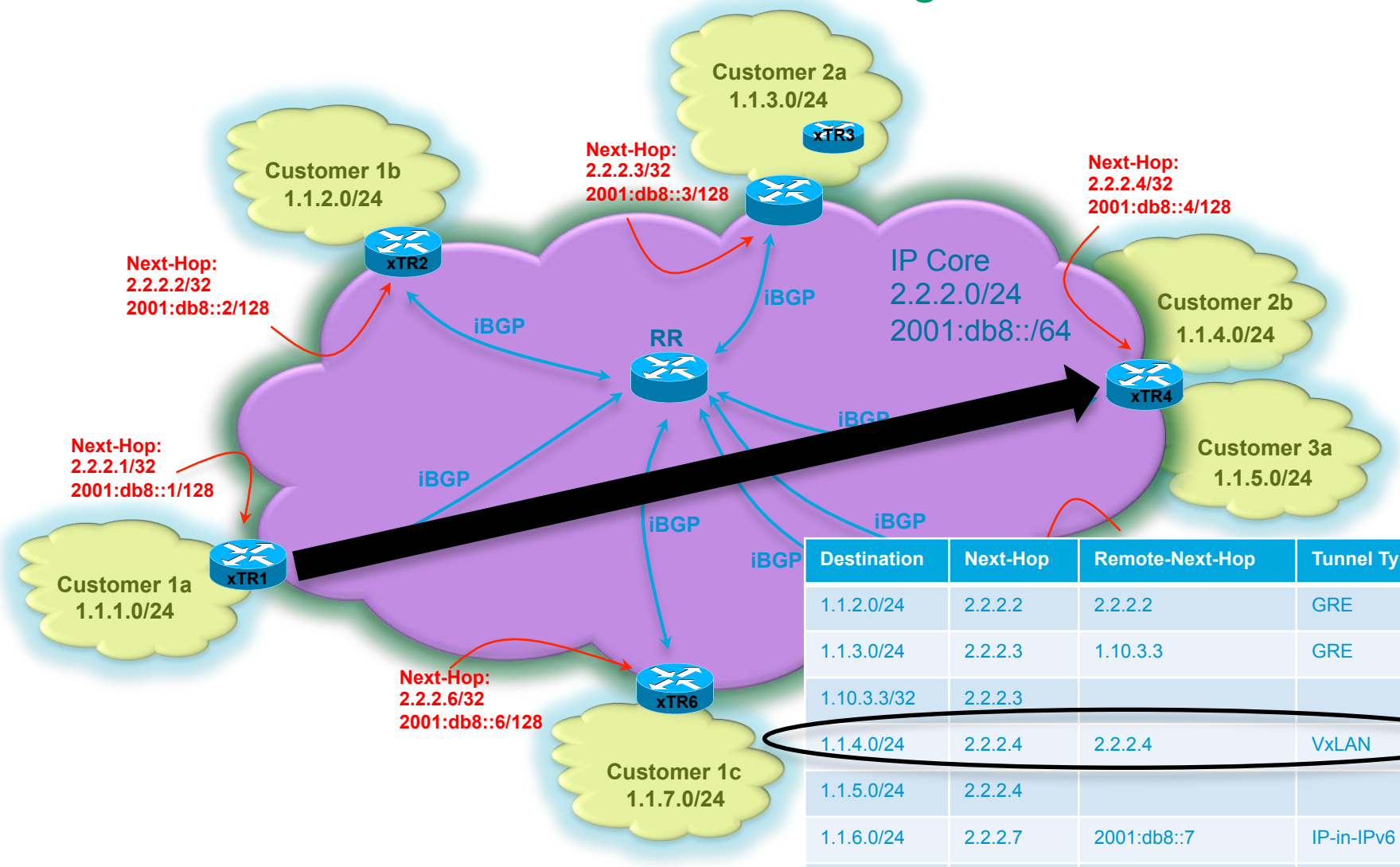
# Tunnel Based Forwarding: Case 1

## Reduction of Core state and size of routing table



# Tunnel Based Forwarding: Case 2

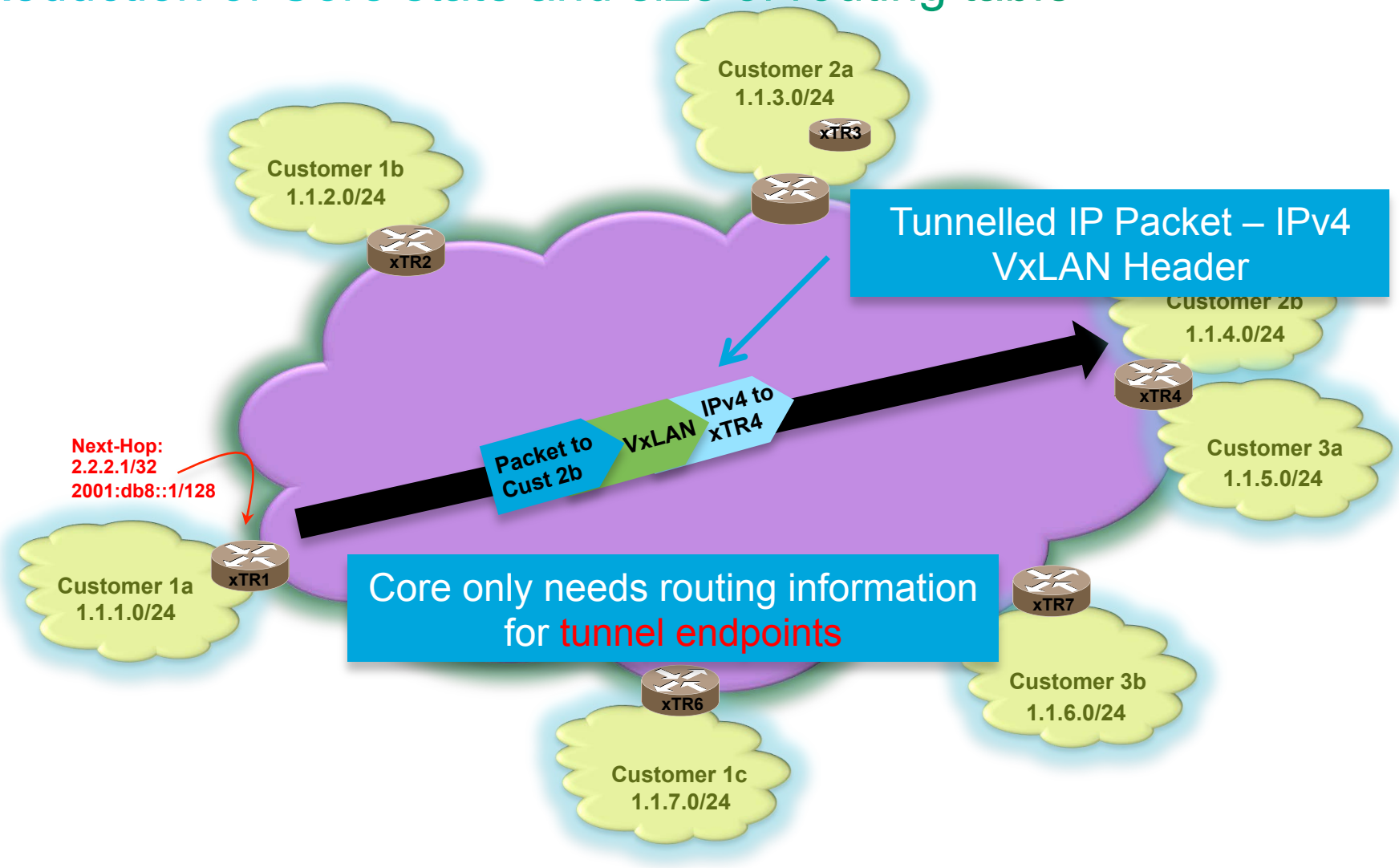
## Reduction of Core state and size of routing table




| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

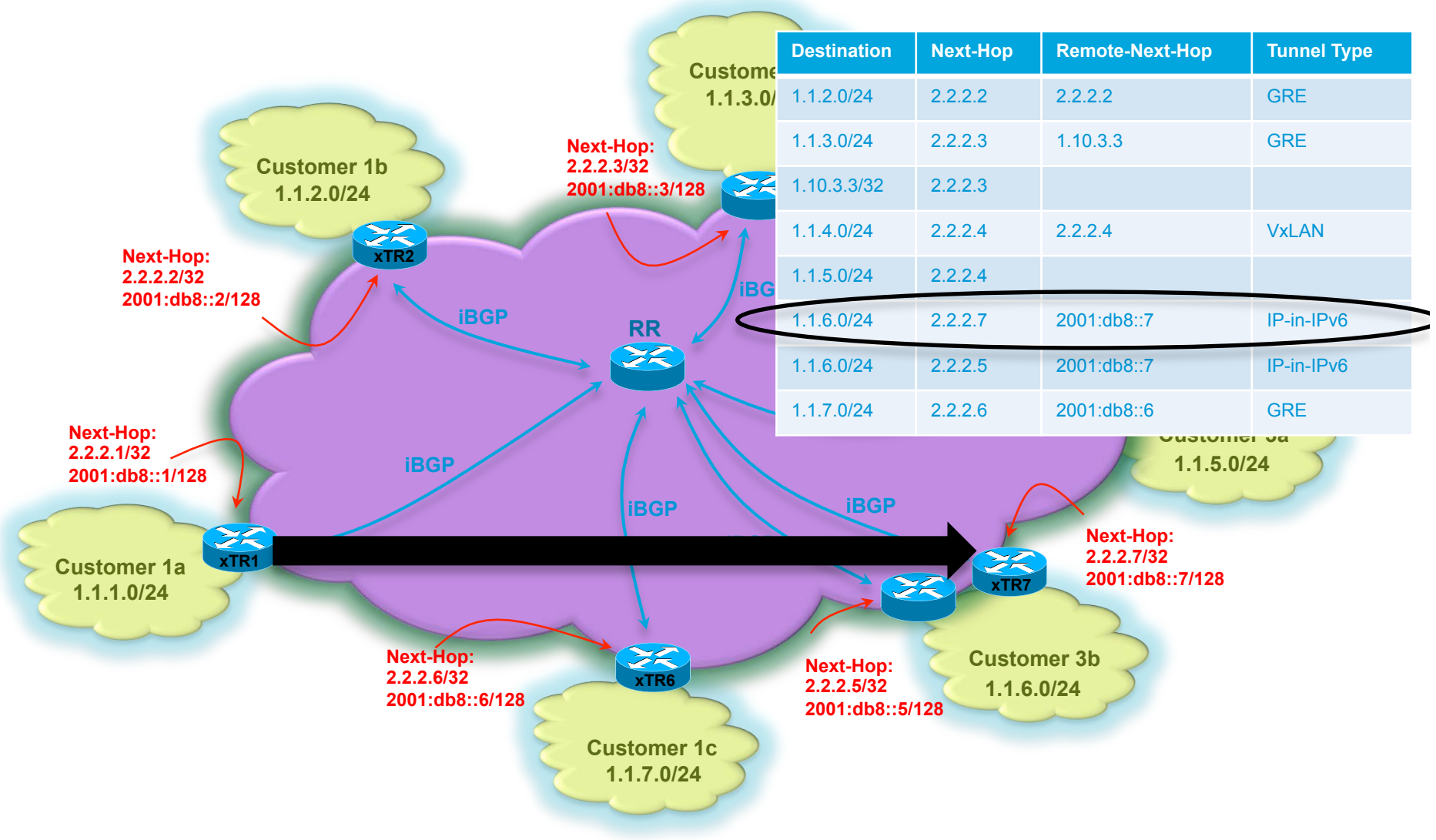
# Tunnel Based Forwarding: Case 2

## Reduction of Core state and size of routing table



# Tunnel Based Forwarding: Case 3

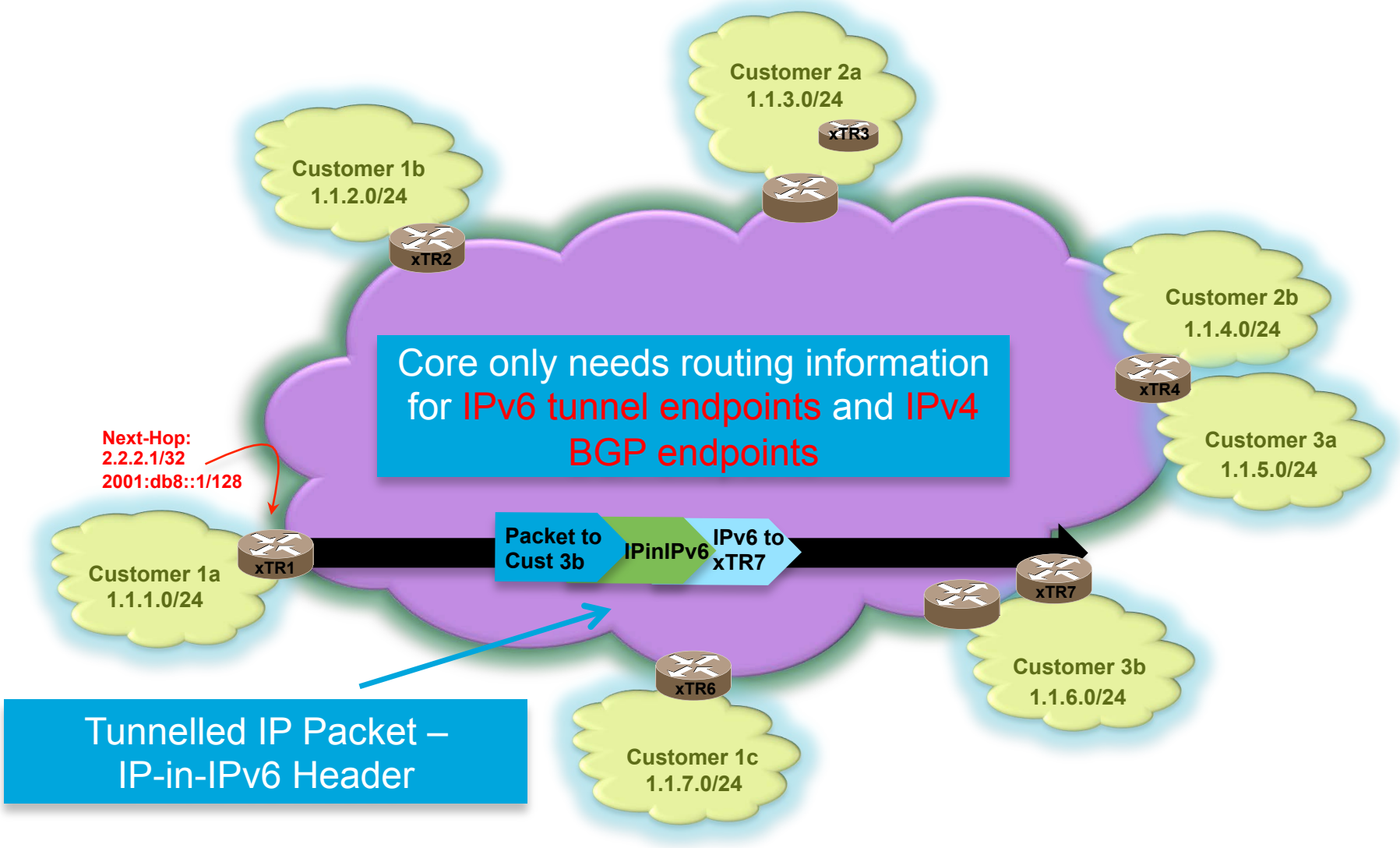
## IPv4 over IPv6 enabled core




| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

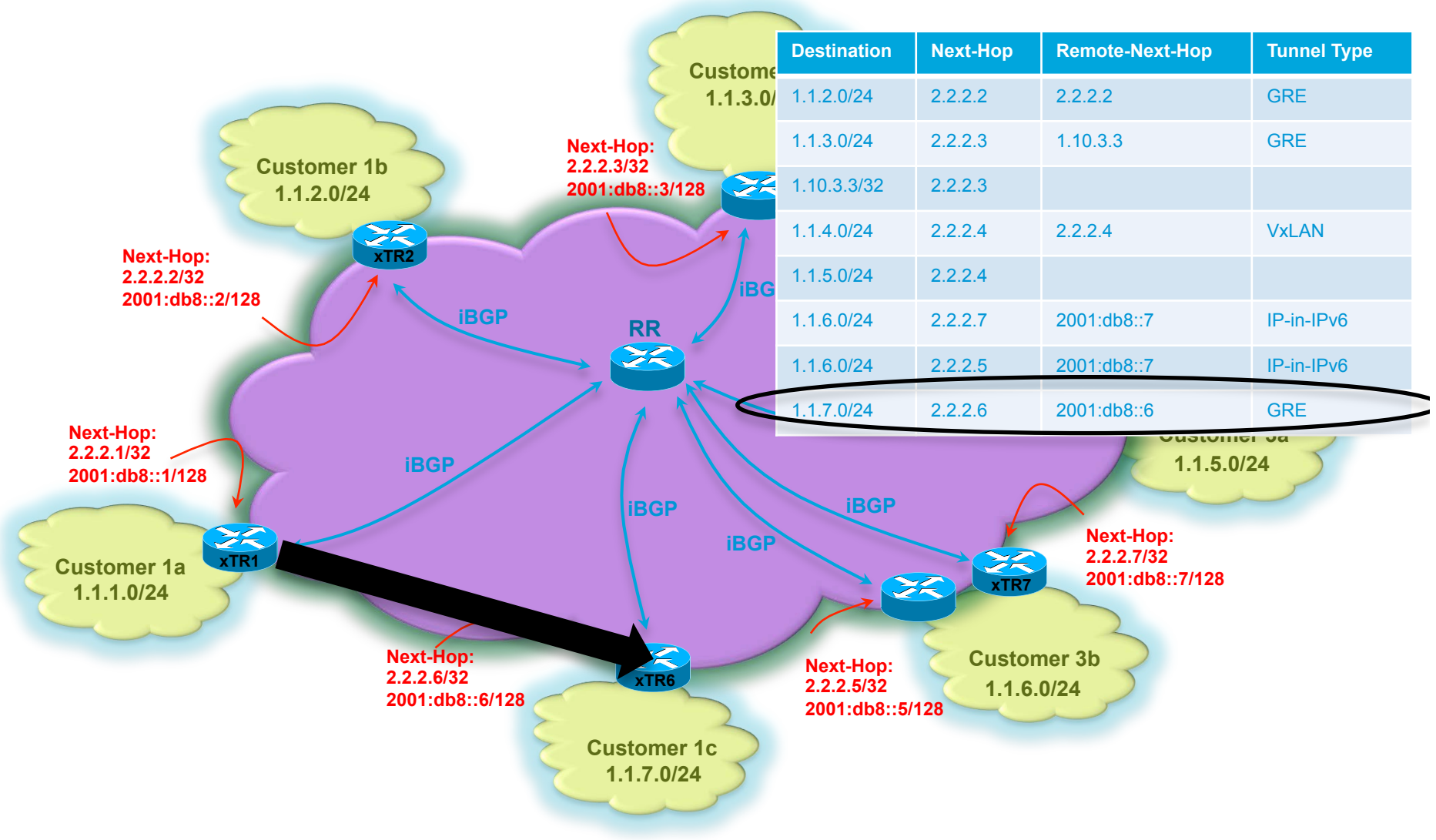
# Tunnel Based Forwarding: Case 3

## IPv4 over IPv6 enabled core



# Tunnel Based Forwarding: Case 4

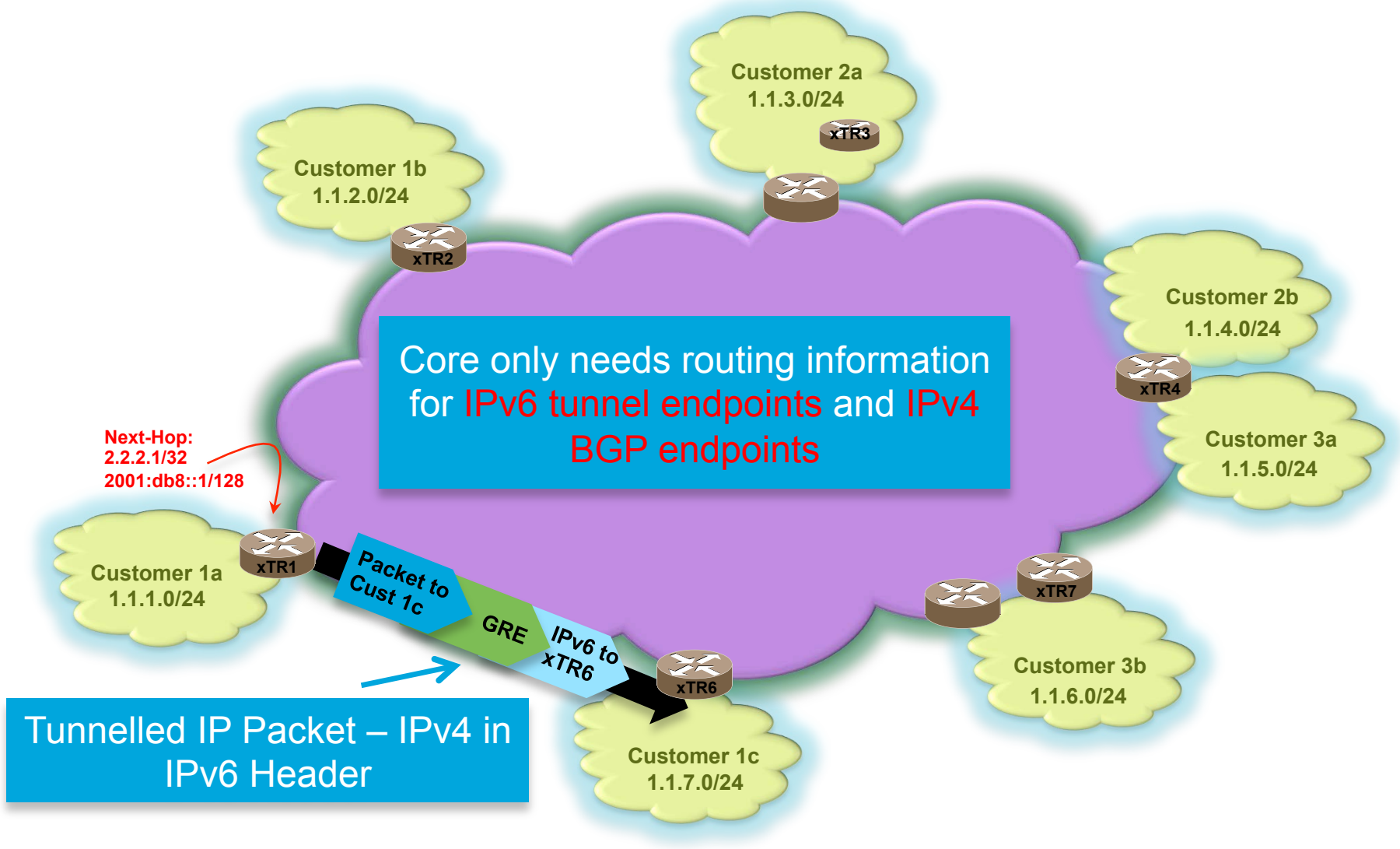
## IPv4 over IPv6 enabled core




| Destination | Next-Hop | Remote-Next-Hop | Tunnel Type |
|---|---|---|---|
| 1.1.2.0/24 | 2.2.2.2 | 2.2.2.2 | GRE |
| 1.1.3.0/24 | 2.2.2.3 | 1.10.3.3 | GRE |
| 1.10.3.3/32 | 2.2.2.3 | | |
| 1.1.4.0/24 | 2.2.2.4 | 2.2.2.4 | VxLAN |
| 1.1.5.0/24 | 2.2.2.4 | | |
| 1.1.6.0/24 | 2.2.2.7 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.6.0/24 | 2.2.2.5 | 2001:db8::7 | IP-in-IPv6 |
| 1.1.7.0/24 | 2.2.2.6 | 2001:db8::6 | GRE |

# Tunnel Based Forwarding: Case 4

## IPv4 over IPv6 enabled core



# Conclusion

- BGP based Dynamic Tunnelling is allows a single IP based control base
- High scalability due to proven BGP technology
- Fast Convergence due to proven BGP and IGP tuning technology
- Network core devices enjoy reduction in the size of the BGP table
- BGP based Dynamic Tunnelling allows virtualisation based upon IP technology
- IPv4 and IPv6 agnostic
- Incremental implementation is supported
- BGP based Security is supported and scalable

# Thank You