NETWORKERS 2004

MULTI-TOPOLOGY ROUTING (MTR)

SESSION RST- 4313

Scott Sturgess (scsturge@cisco.com)
Chetan Khetani (cpk@cisco.com)

Routing DNA

MTR Agenda

Cisco.com

- Overview and Operation
- Protocol Extensions
- Comparison with Other Technologies
- Deployment Scenarios
- Management and Troubleshooting
- Platforms, Cisco IOS® Trains and Timelines

OVERVIEW AND OPERATION

MTR: TERMINOLOGY

Topology	Topology Means a Logical Path that the Traffic Will Take across the Given Network
Multi-Topology	Multi-Topology Means that Each Topology will Route/Forward a Subset of the Traffic as Defined by the Classification Criteria
Global/Base Topology	This Topology is the Same as the Default Global Routing Table that Exists Today—Without MTR Being Used; this is the Same as The Base Topology
Color Topology	New Topologies that Are Defined over and above the Existing Base Topology; Each Colored Topology is Represented by Its Own RIB and FIB
Classification	Selection/Matching of Traffic that Needs to be Marked/Provided with a Different Treatment
Marking	Setting a Value in the Packet or the Frame
Fallback	Use the base topology when no route exists in the colored topology and the router is explicitly configured to do so (incremental deployment)
CoS (L2)	Classification at Layer 2 Based on 802.1p Bits
ToS (L3)	1 Byte Field in IP Header
DSCP	DiffServ Code Point—6 Bits in the ToS (2 Bits Are Now Used for Explicit Congestion Notification)
IP Precedence	3 Bits in the ToS Field
PHB	Per Hop Behavior—DiffServ QoS Treatment on a Per Hop Basis
EFT	Early Field Trial

The Need for MTR

Cisco.com

- **Today, we have two main routing paradigms:**
  - **Destination-based routing:**
    - **Path followed by packet is based on destination address**
  - **Policy-Based Routing (PBR):**
    - **Statically configured path followed based on attributes such as DSCP, port numbers, etc.**
- **Problem Statement: How to dynamically use multiple paths to a given destination based on traffic types?**
- **MULTI TOPOLOGY ROUTING addresses this need**
  - **Initial focus is on the enterprise market**

What Is MTR?

Cisco.com

- **MTR is the third routing paradigm, combining:**
  - **Flexible, dynamic path selection criteria and traffic types**
- **Adding another dimension to destination based routing**
- **Creation of multiple topologies across the same underlying physical infrastructure using existing routing protocols:**
  - **EIGRP, OSPF, IS-IS and BGP**
- **Ability to map certain colored traffic to a topology based on attributes such as:**
  - **DSCP, NBAR, etc.**
- **Color-aware next-hops: e.g. Class of Service routing**
- **End Goal: To influence the path that certain types of traffic would take (to reach to a given destination) based on attributes such as DSCP, application type, etc.**
  - **Traffic separation across network infrastructure**

Conceptual View of MTR

Cisco.com

- Creation of multiple topologies
  - Logical path that traffic will take across the given network
  - MTR means that each topology will route/forward a subset of the traffic as defined by the classification criteria
- Mapping of traffic to a topology—topology selection
  - Determine which traffic (based on a classification criteria) is subject to topology specific forwarding
- Whereas QoS provides per-hop service differentiation within a single path, MTR provides PATH-BASED service differentiation within a single domain

RST-4313
9810_05_2004_c3 7

Design Principles of MTR

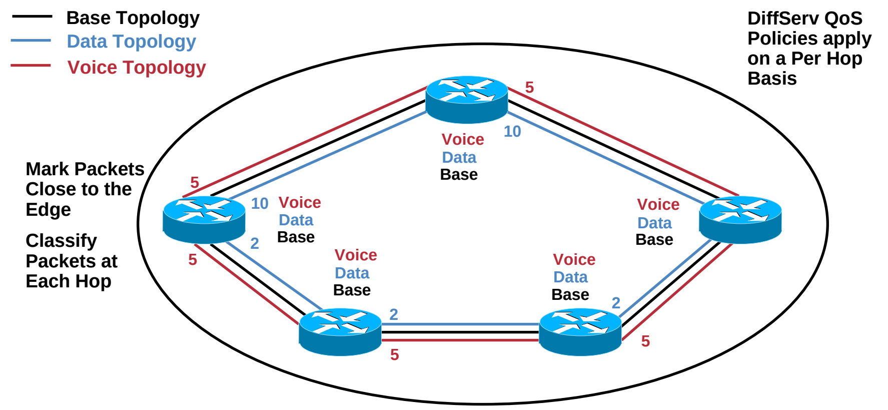
Cisco.com

- Topology is an independent routing entity in itself; all the rules that apply to the global routing table today, will also apply to the topology
  - i.e. Redistribution, summarization, default routes, gateway of last resort are no different for the colored topology and needs to be explicitly configured
- The view of a colored topology is based on the information that exists in that topology
- To attain end to end behavior the colored topology has to be contiguously defined

RST-4313
9810_05_2004_c3 8

Multi-Topology Routing with QoS

Cisco.com



- Goal: Destination-based routing based on traffic attributes…
  - Topologies are independent entities in themselves
  - Each link can have multiple topologies
  - Multiple DSCP Values can be mapped to a single topology

Some Benefits of MTR

Cisco.com

- **Delay vs. throughput**
  - Voice to follow paths that are delay sensitive, whereas data can follow paths that have good throughput, but propagation delay/jitter is not that important
- **Backup links**
  - Using under utilized (backup) links for batch traffic
- **Traffic separation**
  - Using certain links for certain traffic types only
- **Black hole**
  - Forwards all "suspicious traffic" on a separate topology that has security devices and/or to dump it in a "bit bucket"

MTR Topology Characteristics

Cisco.com

- **Maximum topologies**
  - Support for a maximum of 32 topologies per address family
- **Packet classification**
  - Components determine which topology chosen for route look-up
- **Topology fallback**
  - Only allowed to fallback to the base topology if explicitly configured to do so
- **Topology behavior**
  - No different within colored topology than that of the base
- **Topology features**
  - Some of the RP independent features such as max-paths also independently apply to colored topology

MTR Routing Protocol Characteristics

Cisco.com

- **Metrics**
  - Configure metrics on a per topology, per protocol basis
- **Optimal paths**
  - Calculate optimal paths (DUAL/SPF) on a per topology basis
- **Redistribution, defaults and summarization**
  - Apply as today on a per topology basis
- **Load balancing**
  - Across multiple paths within the same topology
  - Equal and un-equal (EIGRP) load balancing
- **Legacy support**
  - Coexistence of MTR-enabled and non-MTR-enabled routers during a transition phase

MTR Routing Protocol Requirements (Cont.)

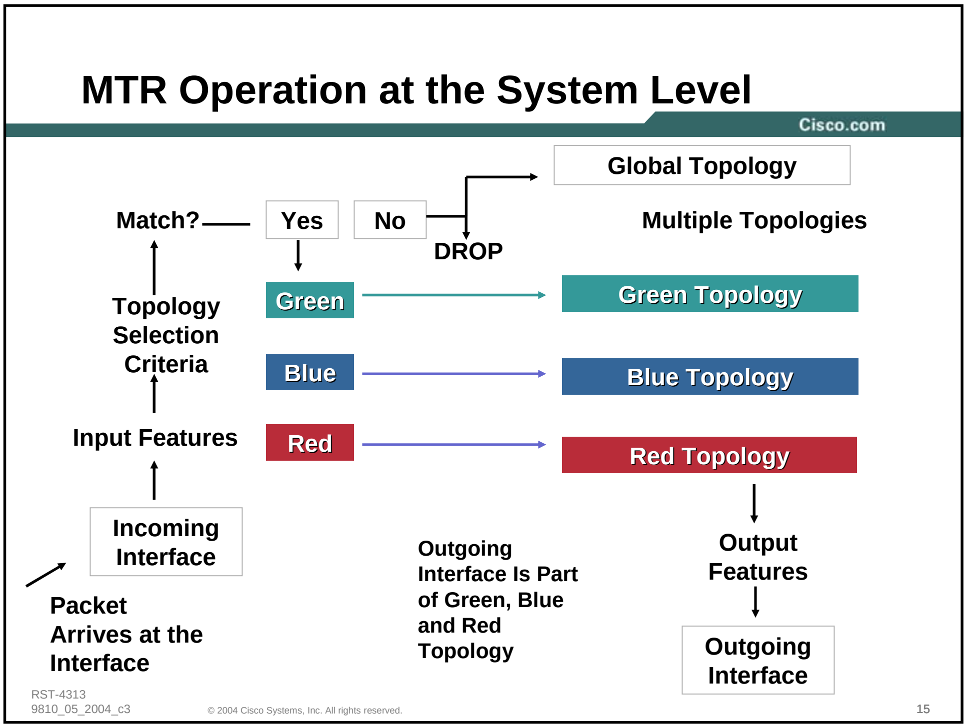
Cisco.com

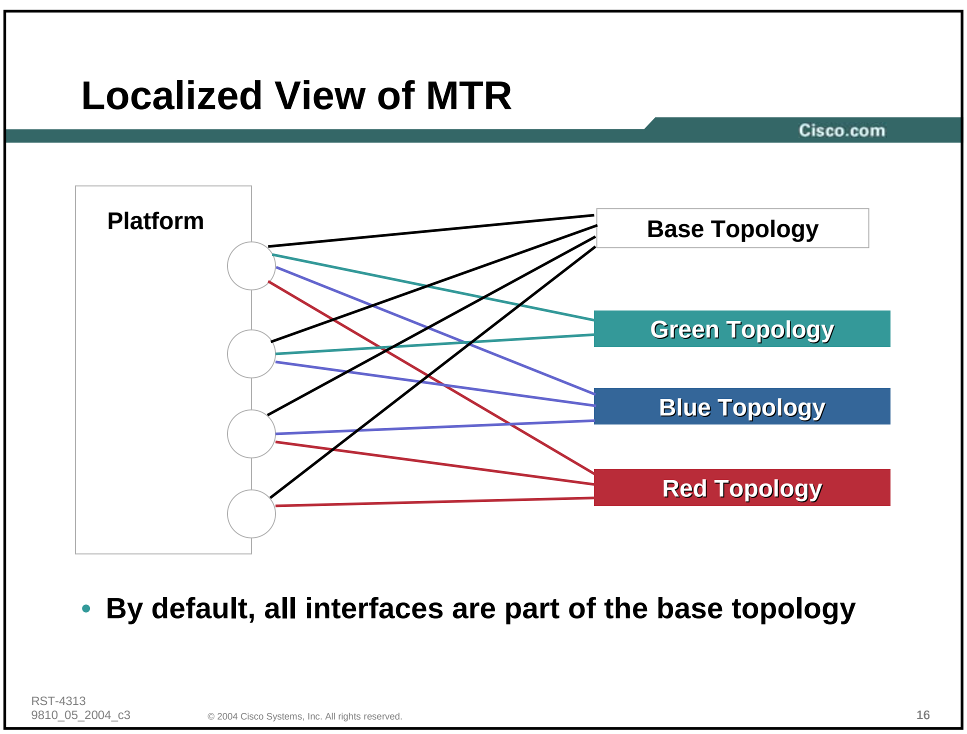
- Fast convergence features
  - Apply on a per topology basis where applicable
- NSF/SSO
  - MTR co-exists with NSF/SSO
- Static routes
  - Ability to define static routes within each topology

MTR Classification Characteristics

Cisco.com

- Classification and topology selection
  - Should be independent of defining and building topologies
- Classification match criteria
  - Initially DSCP only
  - Multiple DSCP values can be mapped to a single topology
- QoS functionality
  - QoS and MTR policies will co-exist without any change in the existing QoS functionality
- Use existing MQC (Modular QoS Cli) to classify and select topologies


MTR Operation at the System Level
Cisco.com
Global Topology
Match?
Yes
No
DROP
Multiple Topologies
Topology
Selection
Criteria
Green
Blue
Red
Green Topology
Blue Topology
Red Topology
Input Features
Incoming
Interface
Packet
Arrives at the
Interface
Outgoing
Interface Is Part
of Green, Blue
and Red
Topology
Output
Features
Outgoing
Interface
RST-4313
9810_05_2004_c3
© 2004 Cisco Systems, Inc. All rights reserved.
15


Localized View of MTR
Cisco.com
Platform
Base Topology
Green Topology
Blue Topology
Red Topology
• By default, all interfaces are part of the base topology
RST-4313
9810_05_2004_c3
© 2004 Cisco Systems, Inc. All rights reserved.
16

Basic Forwarding Model/Behavior

Cisco.com

- **Forwarding path**
  1. Classifies packet
  2. Determines the corresponding class table
  3. Looks up the destination address in that table
  4. Forwarding entry is found for that destination
  5. Forwards the packet to the next hop
- **If no forwarding entry, packet is dropped**
- **If packet does not match any classifier, it is forwarded on the base topology**

RST-4313
9810_05_2004_c3

17

Basic Forwarding Model/Behavior (Cont.)

Cisco.com

- **Recommendation is that all packets are marked at "the edge"—as close to the source as possible before the packets enter into the MTR domain**
- **Re-mark within the MTR domain at your own risk**
- **Due to the risk of loops, there is no "fallback" between class-specific topologies or to the base topology**
  - Unless incremental mode is configured

RST-4313
9810_05_2004_c3

18

The Base Topology

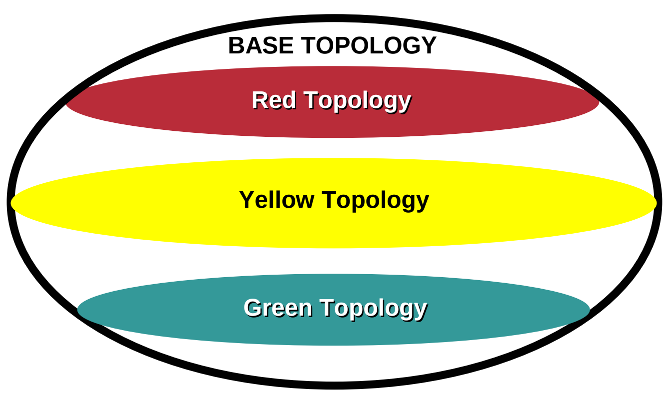
Cisco.com

- The base topology exists everywhere
  - It is also known as the default topology
- Interface metrics configured on the base topology can be inherited by the colored topology
  - Inheritance occurs if no explicit interface metric is configured in the colored topology
- Base topology can be used for fallback if explicitly configured
  - Known as incremental mode

Base Topology – An Example

Cisco.com

- Base topology is a superset of red, yellow, and green



MTR Quirks

Cisco.com

- Addressing is fixed
  - No address re-use
    (MTR is not intended to be a VPN solution)
  - Destination has the same address in the domain regardless of the topology used to reach it
- No tunneling
  - MTR is a Layer 3 technology
  - Fully exploits the characteristic dynamic adaptation of routing protocols

Incongruent IPv4 Ucast and Mcast Topologies

Cisco.com

- MTR provides incongruent unicast and multicast topologies
  - Metrics can be different for each on the same link
- Restrict traffic
  - Restrict multicast only to designated areas of the network
- No fallback from mcast to ucast for RPF
  - RPF checks based on mcast specific ucast table
- Multicast specific protocols such as PIM are not topology specific

High Level Steps in Configuring MTR

Cisco.com

- Define the topology:
  - Router commands (timers, etc.)
  - Interface commands (such as the topology cost)
  - Global topology instantiation (when the RIB/FIB are built)
- Map classification (DSCP values) to topology:
  - Classified packets are forwarded based on the colored topology

MTR: Configuration

Cisco.com

- Configuration is broken down into steps:
  - Define topology—Red, Green, Blue, etc.
  - Define protocol specific topology metrics
  - Define topology mapping
- Topology submodes under different configuration levels for topology specific values
- CLI aspects including configuration, show commands, debug commands that exist for the base topology also apply on a per topology basis

Topology Instantiation

Cisco.com

- **Topologies are instantiated globally**
  - **AFI-typed**
  - **Empty until configured on interfaces—installs connected routes**
  - **Configured interface is enabled for an IGP—installs calculated route**
- **Classifiers are defined globally**
  - **Once enabled, the classifier sequence applies on all interfaces**

Global CLI

Cisco.com

- **Topology Instantiation**

```
address-family ipv4
   topology RED
      [no] shutdown
      [no] forward-base
   topology YELLOW
   topology GREEN
```

Creates topology and data structures

Note: CLI Is Under Implementation and Can Change

RST-4313
9810_05_2004_c3
© 2004 Cisco Systems, Inc. All rights reserved.

Global CLI

Cisco.com

- **Classifier Definition**

```
class-map VIDEO_CLASS
  match {VIDEO DSCP value}
class-map VOICE_CLASS
  match {VOICE DSCP value}
class-map DATA_CLASS
  match {DATA DSCP value}
```

Determines the class of a particular packet

Rest of the packets match in the default class

Note: CLI Is Under Implementation and Can Change

Global CLI

Cisco.com

- **Classifier Sequencing**

```
policy-map class-routing CLASS_ROUTE_POLICY
  class VIDEO_CLASS
    select-topology RED
  class VOICE_CLASS
    select-topology YELLOW
  class DATA_CLASS
    select-topology GREEN
```

Examines sequence until match occurs

Default class packets are routed through the base topology

Note: CLI Is Under Implementation and Can Change

Global CLI

- **Classifier Enabling**

```
service-policy input CLASS_ROUTE_POLICY
```

Enables policy-map

This command invokes packet forwarding on all applicable colored topologies

Note: CLI Is Under Implementation and Can Change

Router CLI

- **Configure MTR under the routing process**

```
router ospf 1
  topology router {topology name} tid {mtid value}
```

Note: CLI Is Under Implementation and Can Change

Interface CLI

Cisco.com

- **Configure MTR on the interface**

```
interface ethernet1/0
  ip ospf cost 20
   ip topology {topology name}
         ip ospf cost 10
```

Note: CLI Is Under Implementation and Can Change

RST-4313
9810_05_2004_c3

Incremental Deployment

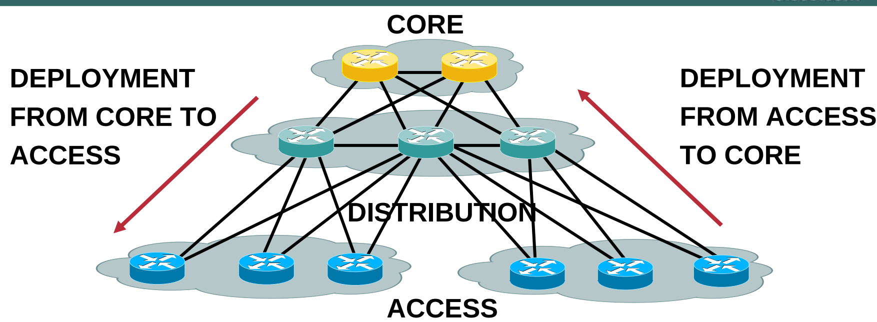
Cisco.com

- **Mode used when deploying within an existing non-MTR environment**
  - **Used temporarily for a transition phase to MTR**
- **Used when "Islands" of the same color exist but have no colored connectivity**
- **Incremental deployment used to maintain connectivity**
- **Colored prefixes are reached via the base topology**
  - **Does the route exist in a colored topology?**
  - **If not, does the route exist in the base topology?**

RST-4313
9810_05_2004_c3

Core, Distribution and Access Deployment

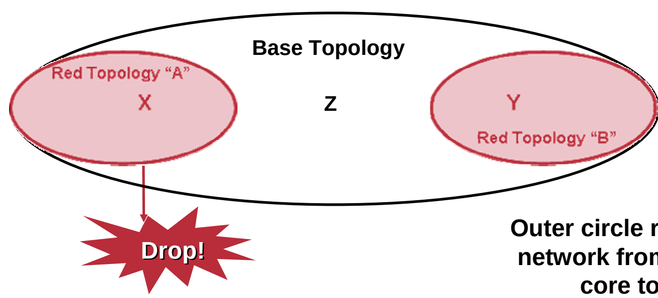
Cisco.com



- Deployment in core, distribution and access nodes
- MTR can be deployed from the core out towards the access nodes or from the access in towards the core
- Choose deployment model that suits best…

Without Incremental Deployment

Cisco.com

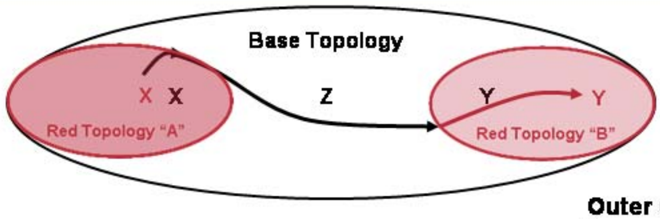


Outer circle represents a network from edge thru core to edge

- Base topology exists everywhere
  - Prefixes X, Y and Z are present in the base topology
- Red topologies A and B are separated "Islands"
  - Prefixes X and Y are present in the red topology as shown
- Red packets from A to B are dropped [by default] as destinations in B are not known in A

With Incremental Deployment

Cisco.com



Outer Circle Represents a Network from Edge thru Core to Edge

- **Base topology exists everywhere**

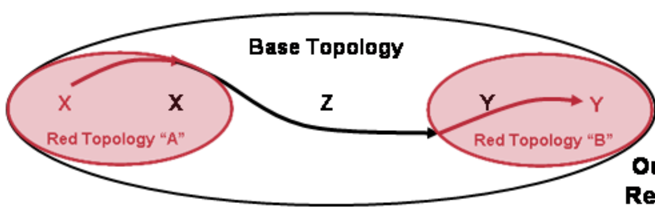
  Prefixes X, Y and Z are present in the base topology
- **Red topologies A and B are separated "Islands"**

  Prefixes X and Y are present in the red topology as shown
- **Red packets from A to B follow the base topology and then follow the red route within B to prefix Y**

RST-4313
9810_05_2004_c3

Incremental Deployment and Colored Default

Cisco.com



Outer Circle Represents a Network from Edge thru Core to Edge

- **Base topology exists everywhere**

  Prefixes X, Y and Z are present in the base topology
- **Red topologies A and B are separated "Islands"**

  Prefixes X and Y are present in the red topology as shown
- **Red packets from A to B follow the red default route, then the base topology and finally follow the red route within B to reach prefix Y**

RST-4313
9810_05_2004_c3

Presentation_ID.scr

Incremental Deployment Scenarios

- **MTR at the edge**

  Multiple links, but can't upgrade all routers in the core yet

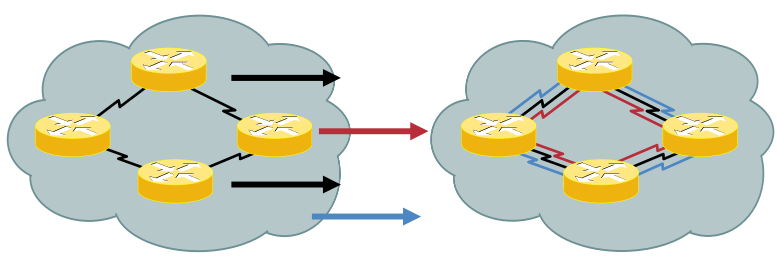
- **Services from MPLS-VPN provider:**

  Run MTR within sites, but PE-CE protocols not MTR aware yet

- **eBGP Peering**

  Run colored topology within network, but not on eBGP peering with core backbone or eBGP peering with service providers

Flash Cut-Over Deployment



- **MTR flash cut-over is also possible…run in parallel**
- **Verify all colored links and routes before cut-over**
- **Will require a maintenance window**

Deployment Caveats

Cisco.com

- **Incremental deployment**
  - Design consideration to ensure no conflicts between topologies
- **Border controls**
  - Border between colored/non-colored network must be clearly defined; this becomes a network design criteria and nothing "special" will be done to avoid loops between colored/non-colored networks
- **Consistency in topology selection**
  - Consistent mapping of traffic types to colored topologies is required
- **Packet remarking**
  - Re-marking of packets within the network is not recommended and if performed is done so at a risk of looping

PROTOCOL EXTENSIONS

MTR EIGRP

Cisco.com

- Single EIGRP topology table per color
  - Isolate topologies
  - Minimize scalability impact on uni-topology
- DUAL per topology
- Single peering relationship
  - Maximize peer scaling
  - Minimize network traffic
  - MTR aware and non-aware routers will peer with each other
- Topology identifier will be encoded in new TLVs

MTR OSPF

Cisco.com

- Single OSPF instance to compute routes for multiple independent topologies
- Single adjacency
- OSPFv2 is extended to:
  - Advertise multiple metrics for each link/prefix
  - Compute SPT for each topology independently
  - Calculate prefixes belonging to each topology independently and update corresponding routing/forwarding tables
  - Interact with multiple routing tables, one per topology
- OSPF: draft-psenak-mt-ospf-00.txt
  - Redefine the TOS-based metric fields in LSAs as specified by OSPFv2 (RFC 2328) to be MTID and MTID metric
  - Selectively include/exclude metric for each topology in the link advertisement
  - MTID range is 127, MTID #0 is reserved for the base topology

MTR IS-IS

Cisco.com

- **Addition of 4 new TLVs**
  - TLV-229: Multi-Topologies Identifier
  - TLV-222: Multi-Topologies intermediate system
  - TLV-235: Multi-Topologies Reachable IPv4 address
  - TLV-237: Multi-Topologies Reachable IPv6 address
- **Some well known MT ID values will be used…**
  - MT ID #0 Base topology
  - MT ID #3 IPv4 multicast topology
- **ISIS—draft-ietf-isis-wg-multi-topology-06**
- **Backward compatible with the existing "Multi-Topology IS-IS"**

COMPARISON WITH OTHER TECHNOLOGIES

MTR vs. PBR

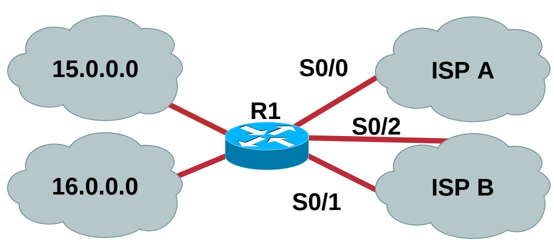
- MTR dynamically creates topologies and influences the path of the packet
- PBR has static influence, overriding destination based routing
- PBR selects next hop based on a number of attributes
- MTR allows selection of desired topology to reach destination

PBR Refresher

- Forwarding decision not based on destination prefix (overrides dynamic routing protocol next-hop decision)
- Selects next-hop based on attributes of user packet (source/destination IP address, application port, packet lengths, tracking criteria, and so forth)
- Set next-hop, interface, VRF, etc.

Functional Example:

If { match src 15/8
Then { set next-hop S0/0
If { match src 16/8
Then { set next-hop S0/1
Else { set next-hop S0/2



Key Differences Between MTR and PBR

Cisco.com

PBR	MTR
Manual Selection of Path	Automatic Selection of Path
Manual Configuration and Definition of Next Hops	Per Hop Configuration but Dynamic Definition of Next Hops
Manual Configuration and Definition of Dynamic Failover	Dynamic Routing Protocols Take Care of This
Manual Configuration and Definition of Load Balancing	Dynamic Routing Protocols Take Care of This
Manual Configuration for End to End Path Influencing	Dynamic Routing Protocols and Path Influenced via Metrics

RST-4313
9810_05_2004_c3

Key Differences Between MTR and Multiple Routing Processes with VRF

Cisco.com

	MRP with VRF	MTR
Relationship	PE-CE Type Scenario Support Multiple Topologies under a VPN VRF	Multiple Topologies within aVRF (End Goal)
Interface	Interface Isolation; multiple Routing Processes Cannot Run on the Same Interface	Same Interface Can Have Multiple Topologies; can Share the Same Link Across Multiple Colors
Processes	Multiple Routing Processes with VRF	Single Routing Process
Overlapping Addresses	Allowed across VPNs	Not the Goal (Not a VPN Solution)
Updates and Hellos	Separate Updates and Hellos per Process	Same Updates and Hellos Carry Color Specific Information

RST-4313
9810_05_2004_c3

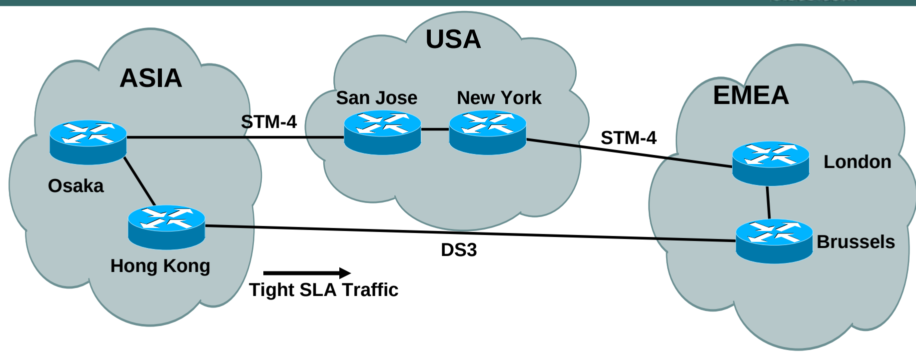
Key Differences Between MTR and MPLS-TE

- MTR primarily for IP
- MTR does Class of Service routing
  - Route IP packets over different topologies
- Traffic Engineering requires MPLS
- MPLS-TE requires tunnels, LDP
- MTR creates separate topologies
- MPLS-TE more suitable if the requirement is to have 100s to 1000s of TE tunnels

DEPLOYMENT SCENARIOS

Traffic Separation

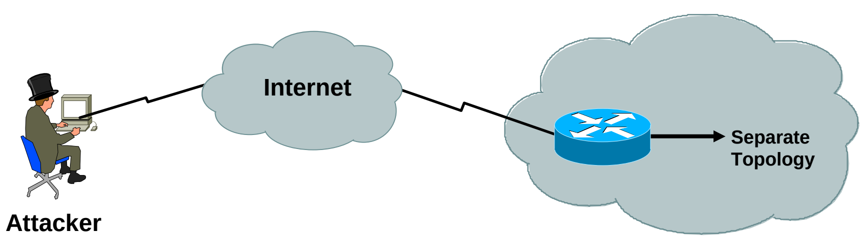
Cisco.com



- Tight SLA traffic from ASIA to EMEA MUST use DS3 and NOT transit USA as this would result in excessive cumulative propagation delay
- Other traffic MUST NOT use the DS3 as this does not offer sufficient bandwidth, so MUST transit USA over the high bandwidth link
- Must route one class of traffic via STM-4 and other via DS3

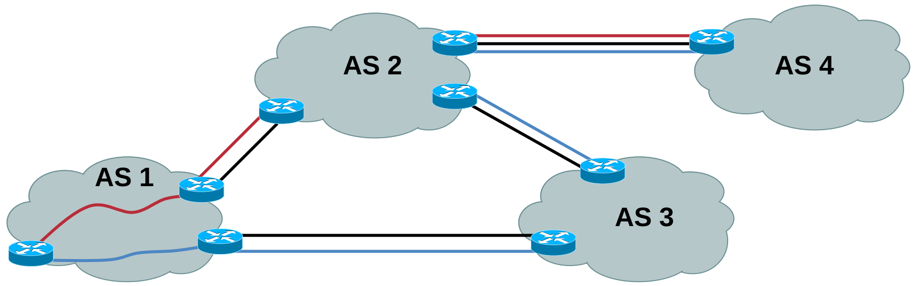
Directed DoS Attacks

Cisco.com



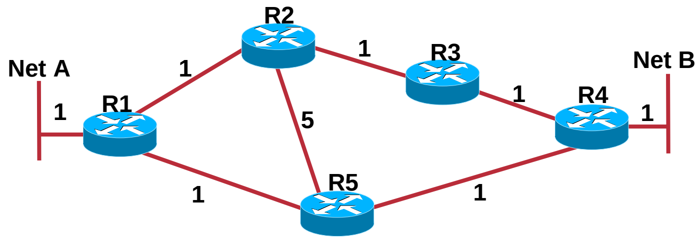
- Traffic that is identified as being suspicious and/or denial of service attacks can be routed onto a separate topology
- Traffic can then be controlled, monitored and/or sent to the bit bucket!

BGP MTR



- All the ASes have BGP MTR configured
- All the ASes send and receive colored updates; iBGP follows the IGP colored path
- Red packets follow the red path to AS 4 and blue packets follow the blue path to AS 4

Default Behavior (IPv4 Unicast)

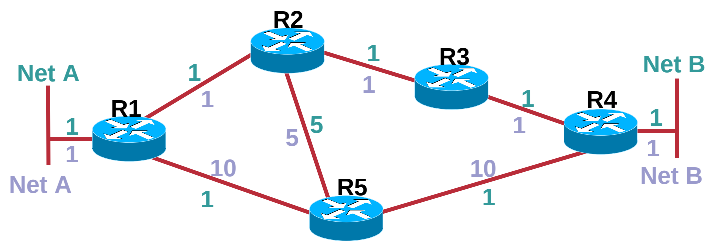


Routing IPv4 from Net A
✓ to Net B via (R1 -> R5 -> R4) metric 3

- Routing decisions are done per destination prefix
- This network shows the individual links and metrics that compose the single IPv4 unicast topology

Separate Topologies (Unicast/Multicast)

Cisco.com

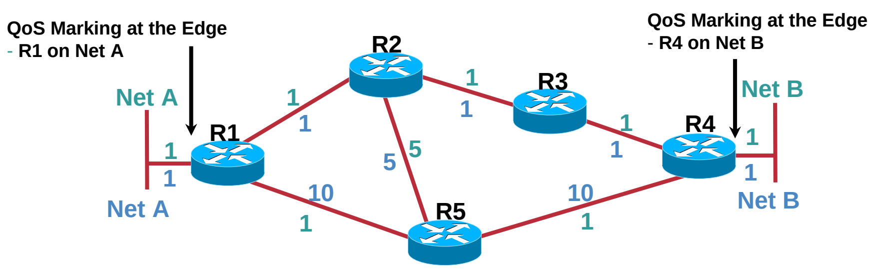


Routing IPv4 Unicast from Net A
✓ to Net B via (R1 -> R5 -> R4) metric 3
Routing IPv4 Multicast from Net A
✓ to Net B via (R1 -> R2 -> R3 -> R4) metric 4

- **Multicast can start populating separate RPF tables from different IGP topology within IPv4 or IPv6 for same source/destination prefixes**

Separate Topologies per Class of Service

Cisco.com

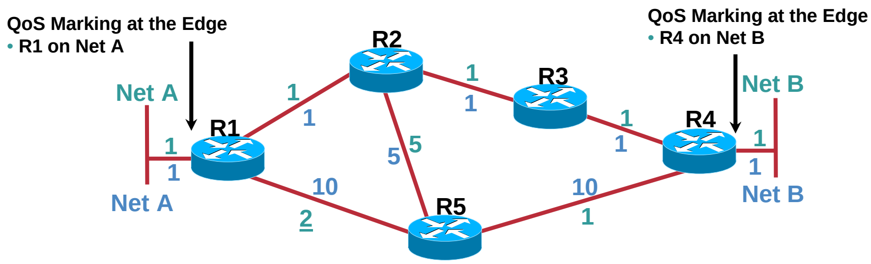


Routing IPv4 CoS:Data from Net A
✓ to Net B via (R1 -> R5 -> R4) metric 3
Routing IPv4 CoS:Voice from Net A
✓ to Net B via (R1 -> R2 -> R3 -> R4) metric 4

- **Separation of Classes of Service according to DSCP/precedence per topology definition (different metrics per topology per CoS)**
- **As many topologies as CoS definitions—Usually 4-6 classes**

Equal Cost Load Balancing per CoS

Cisco.com



Routing IPv4 CoS:Data from Net A
- ✓ to Net B via (R1 -> R5 -> R4) metric 4
- ✓ to Net B via (R1 -> R2 -> R3 -> R4) metric 4

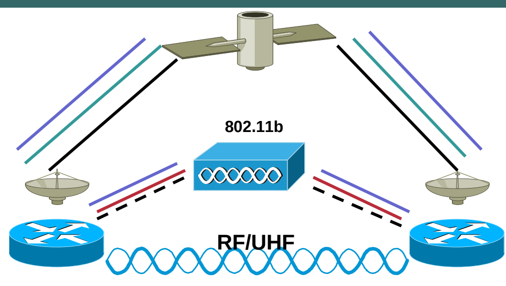
Routing IPv4 CoS:Voice from Net A
- ✓ to Net B via (R1 -> R2 -> R3 -> R4) metric 4

- CoS:Data metric between R1->R5 changed to 2
- CoS:Data traffic from Net A to Net B will load share across two paths
- Notice that QoS queuing still applies on path across R1 -> R2, etc.

RST-4313
9810_05_2004_c3

Wireless Scenario

Cisco.com

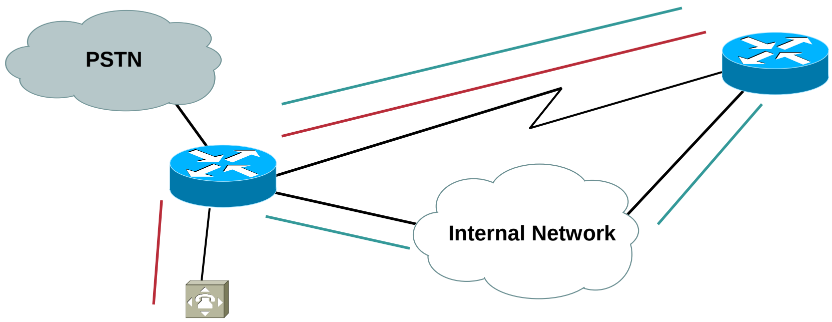


- Red prefers the RF/UHF link and uses the 802.11b link as backup; red never goes over the satellite link
- Blue prefers the 802.11b link and uses the satellite link as backup; blue never goes over the RF/UHF link
- Green prefers the satellite link and never uses the 802.11b link or the RF/UHF link
- Preferences above can all be achieved by applying different metrics

RST-4313
9810_05_2004_c3

Voice Policy Scenario

Cisco.com



- If Voice (Red) fails, use off-network connection to the PSTN; do not use other internal links
- Critical Data (Green) may use another path or load balance within the internal network

MANAGEMENT AND TROUBLESHOOTING

MTR Management

Cisco.com

- MTR MIBs for: EIGRP, OSPF, BGP
- Network wide configuration distribution tool
  - Mapping of DSCP values to colored topologies
  - Configuration of colored topologies
- GUI tools
  - Development of internal management tools
  - Possibility of partnership with external management vendors

MTR Troubleshooting

Cisco.com

- Topology specific show commands
  - show ip route topology <color>
  - etc.
- Topology specific debug commands and syslog messages
  - debug ip <protocol> topology <color>
  - etc.
- Generic troubleshooting commands that apply to a colored topology
  - MTR Ping, MTR trace route

PLATFORMS, CISCO IOS TRAINS AND TIMELINES

RST-4313
9810_05_2004_c3

First Target: Address Family IPv4

Cisco.com

- EIGRP, OSPF, IS-IS, BGP, static routes
- Classification: DSCP
- Incongruent multicast and unicast topologies
- Target platforms:
  - Cisco 7600/Catalyst 6500
  - Cisco 1700, 2600, 36/3700, 7200, 7500
  - Cisco 7300, Catalyst 3500, Catalyst 4500
- Future: IPv6 MTR—EIGRP, OSPF, IS-IS, BGP

For More Information, Contact the MTR Product Manager:
Chetan Khetani—cpk@cisco.com

RST-4313
9810_05_2004_c3

Platforms, Cisco IOS Trains and Timelines

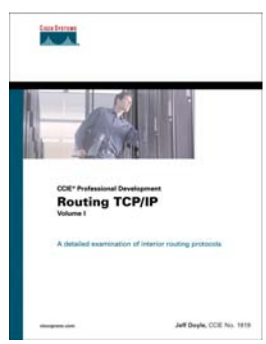
Cisco.com

Platforms	Release	Date
Cisco 7600 and Cat 6500	12.2S-Based Release	2H–CY 2005
Cisco 1700, 2600, 3700, 7200 (NPE), 7500	T-Based Release	2H–CY 2005
Cat 4500, Cisco 7300	12.2S-Based Release	2H—CY 2005
Generic T-Based Platforms Generic 12.2S-Based Platforms	T Release 12.2S	1H—CY 2006

Recommended Reading

Cisco.com

- Continue your Networkers learning experience with further reading for this session from Cisco Press.
- Check the Recommended Reading flyer for suggested books.



Available on-site at the Cisco Company Store

Complete Your Online Session Evaluation!

Cisco.com

WHAT: Complete an online session evaluation and your name will be entered into a daily drawing

WHY: Win fabulous prizes! Give us your feedback!

WHERE: Go to the Internet stations located throughout the Convention Center

HOW: Winners will be posted on the onsite Networkers Website; four winners per day

CISCO SYSTEMS