

A man in a white shirt and red tie is holding a large red pipe that arches over a colorful, abstract landscape. The landscape features a blue sky, a yellow wall, and a green and blue ground. The man is standing on a blue surface, and the pipe is held high above his head.

Principy technologie MPLS a její aplikace





Jaromír Pilař, CCIE #2910

E-mail: jpilar@cisco.com

IP telefon: 02/2143 5029



Agenda

- **Architektura MPLS**
- **Frame-mode MPLS - Přiřazení značek a jejich distribuce**
- **Cell-mode MPLS - Přiřazení značek a jejich distribuce**
- **Aplikace MPLS - přehled**
- **IP+ATM integrace**
- **MPLS Traffic Engineering**
- **MPLS VPN**



MPLS - architektura

MPLS - Multi-Protocol Label Switching

“The primary goal of the MPLS working group is to standardise a base technology that integrates the **label swapping** forwarding paradigm with **network layer routing**. This base technology (label swapping) is **expected to improve the price/performance** of network layer routing, **improve the scalability** of the network layer, and **provide greater flexibility** in the delivery of (new) routing services (by allowing new routing services to be added without a change to the forwarding paradigm)”

draft-ietf-mpls-framework

Co MPLS není

- MPLS není pouze metoda integrace IP a ATM, **ALE**
- Integrace IP a ATM je pouze jedna z významných aplikací MPLS

Co MPLS není

- MPLS není metoda jak učinit routery (mnohem) rychlejší, **ALE**
- Přepínací algoritmus MPLS je jednodušší než přepínací algoritmus IP, **A** umožňuje bohatší funkčnost

MPLS a referenční model OSI

- **MPLS není technologií síťové vrstvy**

nemá vlastní směrování a adresaci - využívá IP adresaci + IP směrování (s rozšířeními)

- **MPLS není technologií linkové vrstvy**

protože MPLS pracuje přes různé technologie linkové vrstvy (např., SONET, Ethernet, ATM, atd...)

- **MPLS není vrstvou ve smyslu OSI RM**

nemá jednotný formát pro převzetí dat z vyšší vrstvy

“shim” pro SONET, VCI/VPI pro ATM, lambda pro OXC, atd.

MPLS nezapadá přesně do referenčního modelu OSI

MPLS - základy

- Přepínání (**forwarding**) paketů je založeno na značkách (**labels**) (ne na informaci IP hlavičky)

přepínání je založeno na koncepci '**label swapping**'
podobné jako některé L2 přepínací mechanismy (dlci, vpi/vci)

MPLS umožňuje mít několik značek (label stack)

jakmile jsou pakety označeny už se znovu neklasifikují

pro klasifikaci paketů mohou být použita různá pravidla

destination-based unicast, TE, QoS, VPN

MPLS - základy

- **Značky jsou přiřazeny na vstupu do MPLS domény**

**vstupní klasifikace a přidělení značky -
label imposition**

umístění značky v paketu/buňce záleží na režimu, ve kterém MPLS pracuje

Label imposition používá CEF FIB => CEF je nutná implementační (nikoli principiální) podmínka pro MPLS

MPLS - základní komponenty

- **MPLS architektura je rozdělena do dvou oddělených komponent**
 - Přepínací komponenta (Forwarding Component)**
 - Řídící komponenta (Control Component)**
- **Poskytuje oddělení přepínacích a řídicích funkcí, které mají odlišnou podstatu**

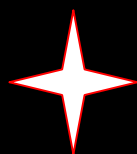
MPLS - přepínací komponenta

Přepínací komponenta

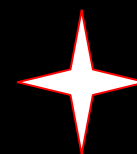
často se také označuje jako **data plane**

zodpovědná za přepínání paketů/buněk v závislosti na značkách (labelech)

využívá 'label forwarding database'



Jednoduché přepínání značek



'Hrubá síla'

MPLS - řídicí komponenta

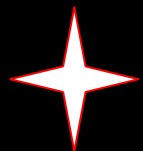
Řídicí komponenta

často se také označuje jako **control plane**

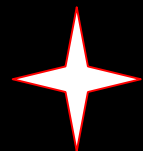
zodpovědná za tvorbu a udržování tabulek se značkami, kterými se řídí přepínací komponenta (**label bindings**)

využívá FIB (CEF)

přiřazení značek je distribuováno pomocí **label distribution protokolu**



Tvorba informace, podle níž se rozhoduje



'Rozum'

MPLS základy - řídicí komponenta

- Každý MPLS uzel musí provozovat IGP směrovací protokol (nebo statické směrování)

výměna informace o IP prefixech s ostatními MPLS uzly

Ize použít libovolný směrovací protokol
doporučuje se link state (OSPF, ISIS) kvůli TE

každý MPLS uzel realizuje řídicí komponentu IP směrovače (nerealizuje přepínací komponentu)

včetně ATM přepínačů v režimu cell-mode MPLS

MPLS základy - řídicí komponenta

- Každý MPLS uzel musí používat label distribution protocol

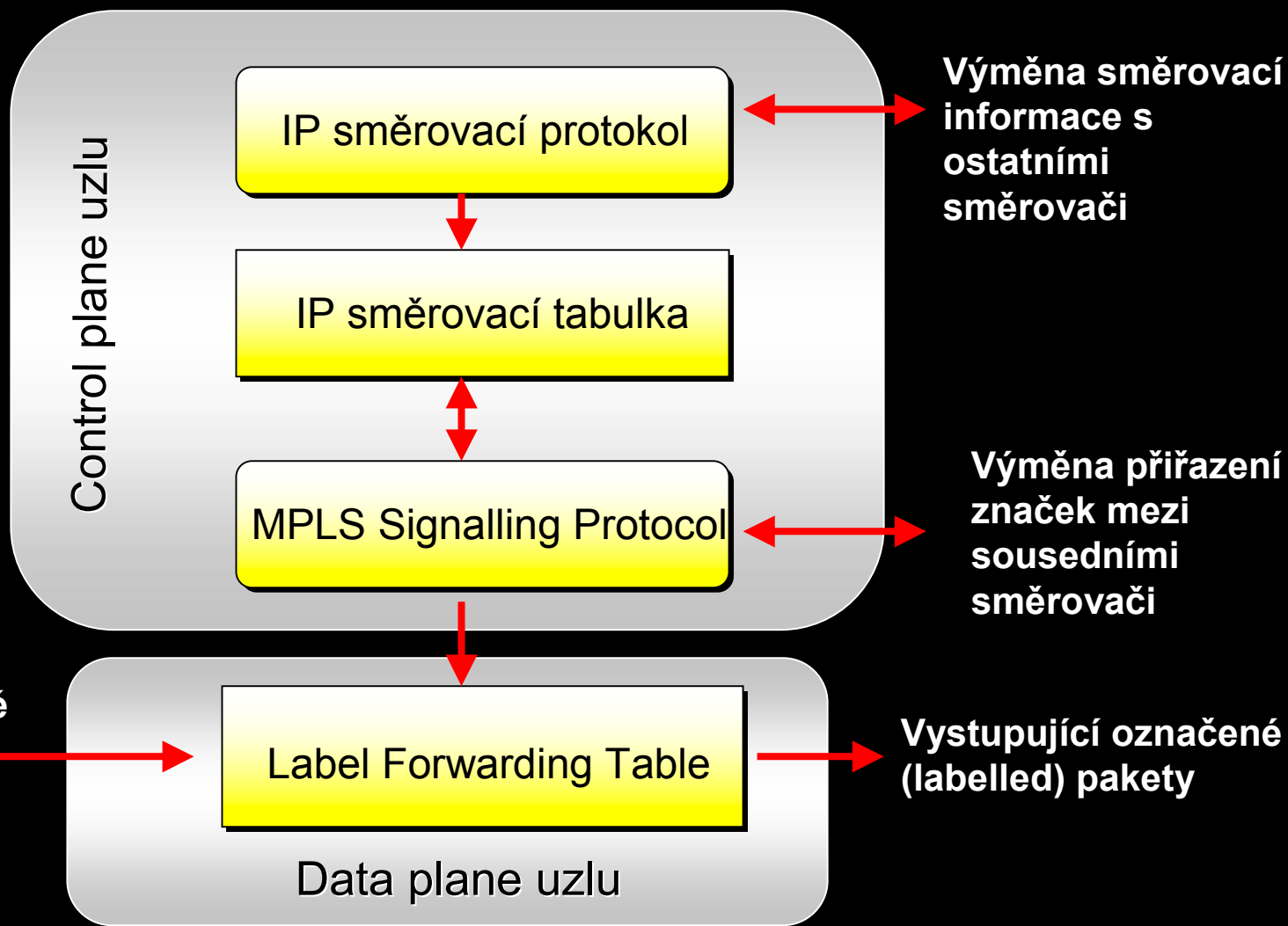
pro výměnu informací o přiřazení značek

pro naplnění přepínací tabulky MPLS

může mít více sousedů

next-hop z IP přepínací tabulky se používá pro výběr značky do přepínací tabulky MPLS

MPLS - základní komponenty



Label Switch Router

- **Jakýkoli směrovač nebo přepínač, který se podílí na přiřazování a distribuci značek a umí přepínat pakety/buňky v závislosti na značkách**
- **Může mít rozličnou funkcionálnítu podle toho, kde se v MPLS doméně nachází**
rozdíly mezi různými typy LSR jsou čistě architektonické, jedno zařízení může vystupovat v několika rolích

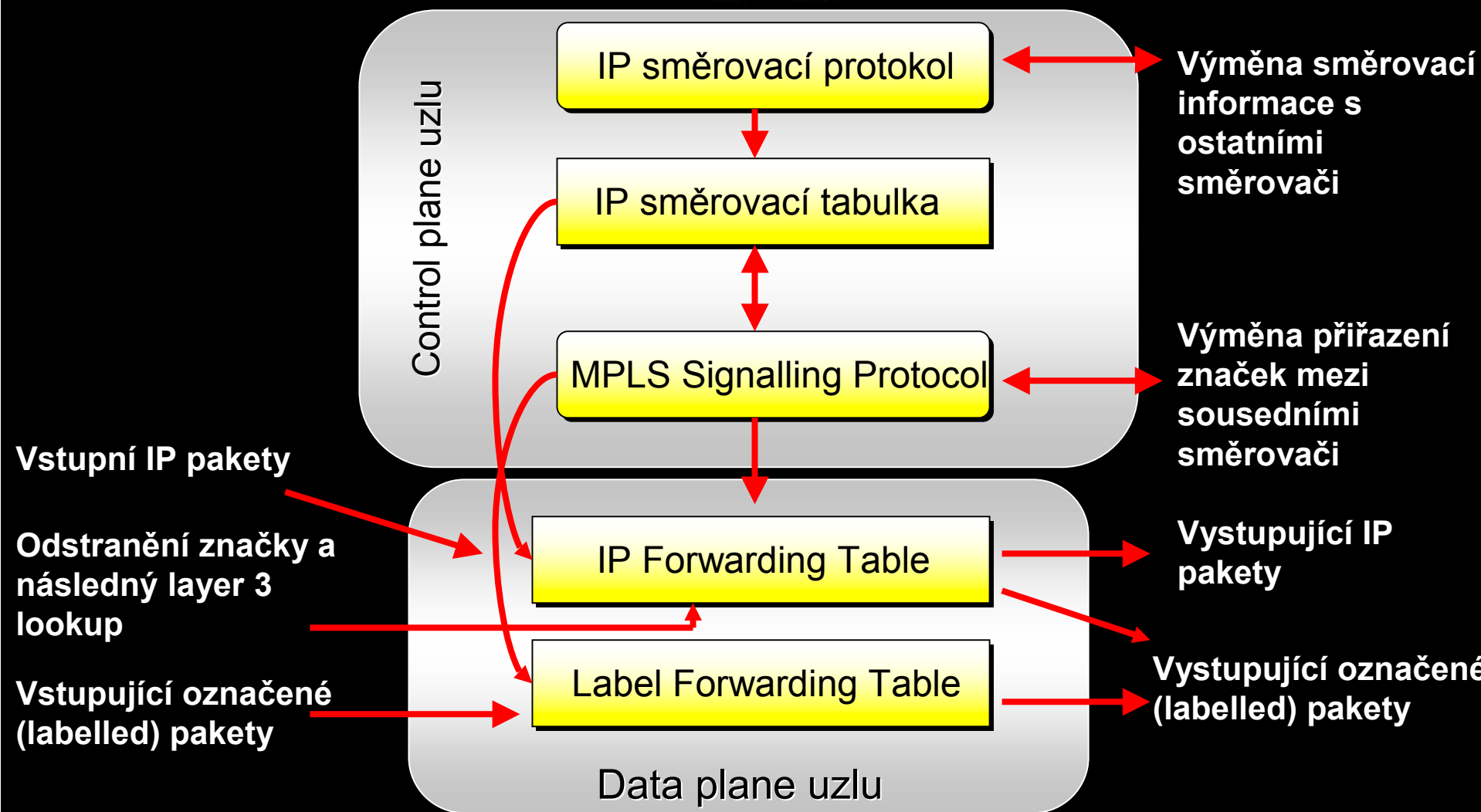
Edge-LSR

- LSR, který provádá vkládání a výběr značek na hranici MPLS domény (**label imposition nebo label disposition**)

každý LSR, který má sousedy neprovozující MPLS je hraniční-LSR

vyjimku tvoří ATM hraniční-LSR, který může mít MPLS sousedy, kteří nejsou ATM-LSR

MPLS hraniční-LSR

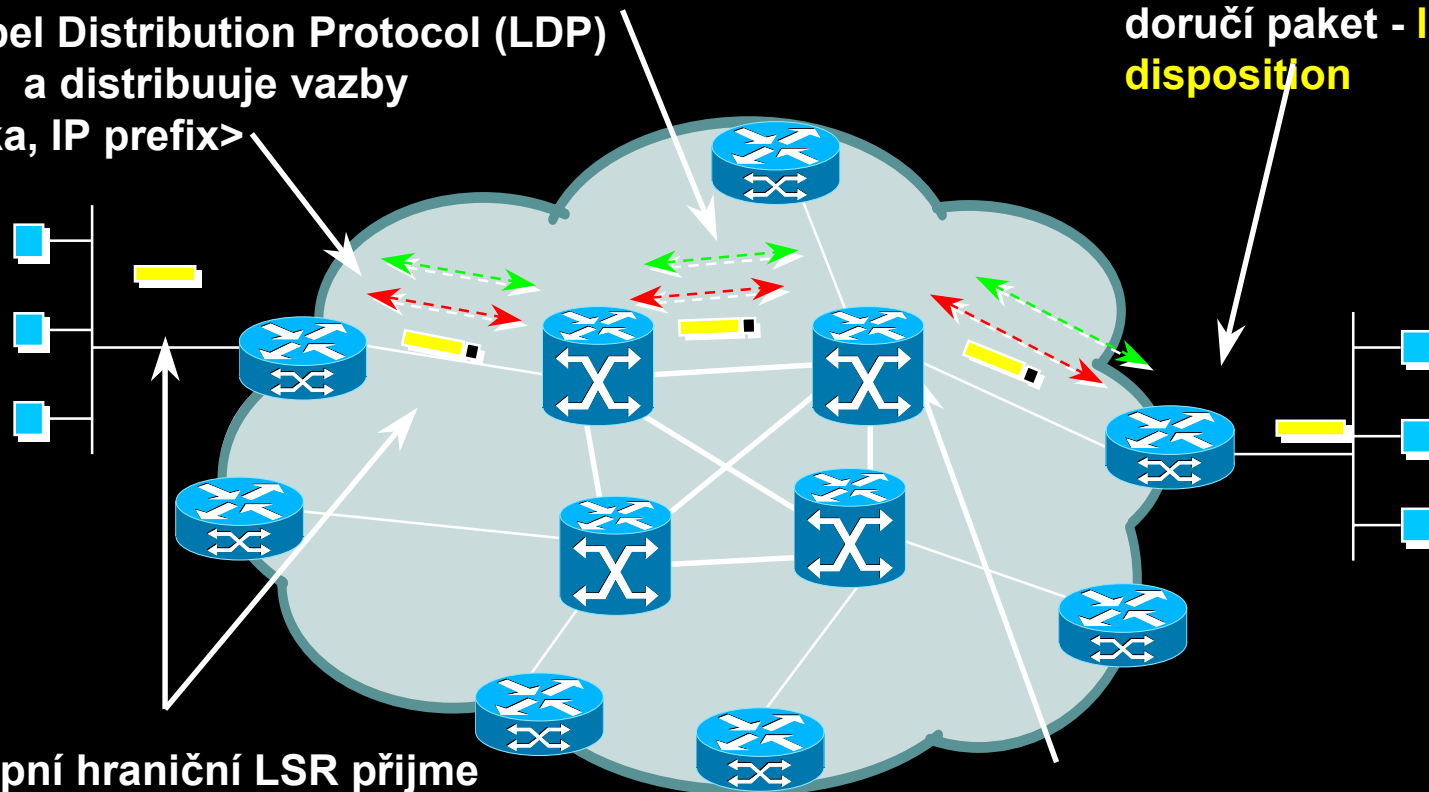


MPLS - princip funkce (shrnutí)

1a. Existující směrovací protokol (např. OSPF, IS-IS) vytvoří směrovací tabulky

1b. Label Distribution Protocol (LDP) vytvoří a distribuuje vazby <značka, IP prefix>

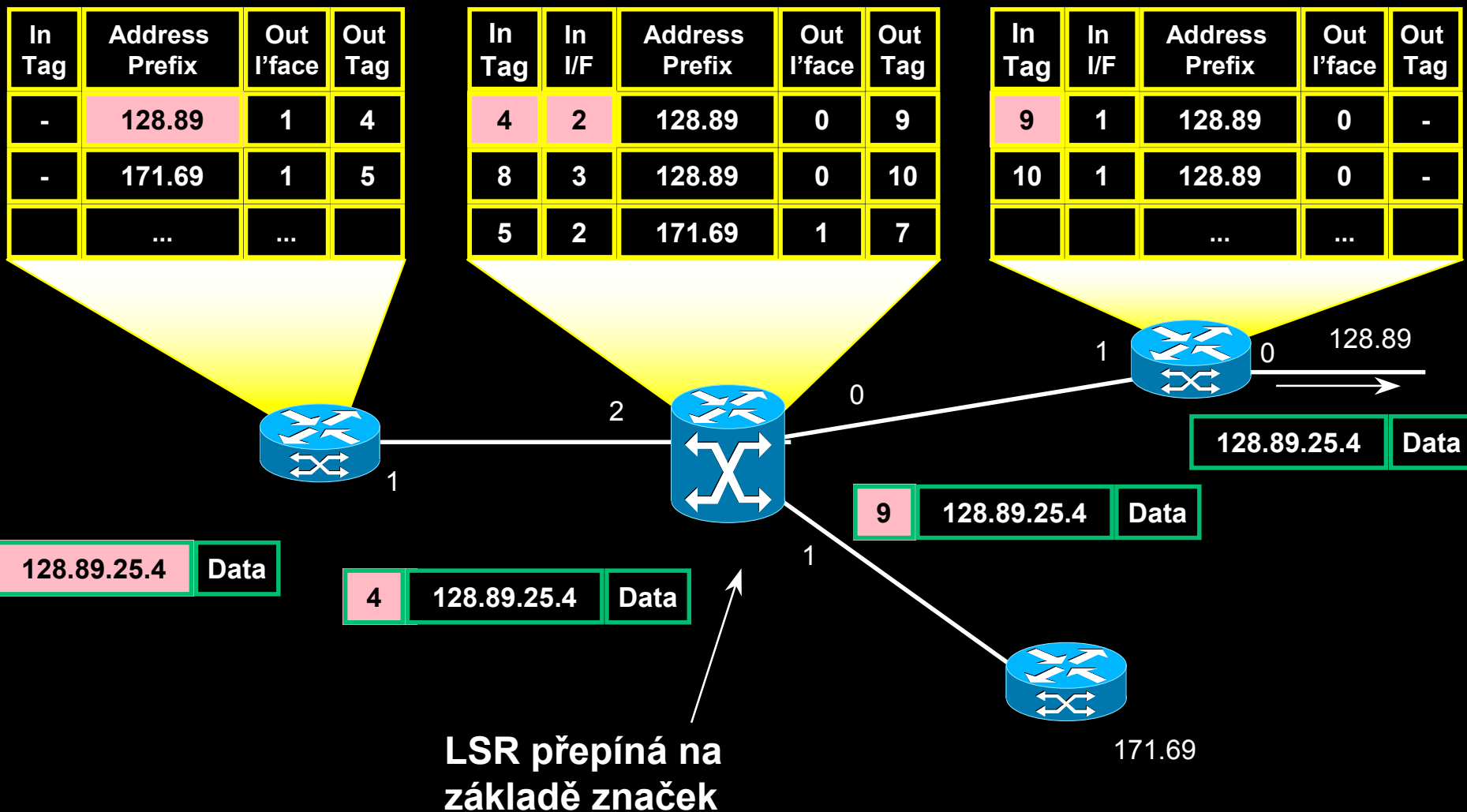
4. Výstupní hraniční LSR vyjme značku a doručí paket - **label disposition**



2. Vstupní hraniční LSR přijme paket, klasifikuje ho a označí značkou - **label imposition**

3. LSR přepíná pakety na základě značek - **label swapping**

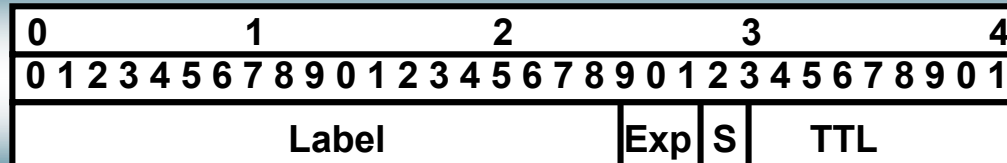
MPLS přepínání - příklad



Formát značky

- Formát značky a jeho délka závisí na typu enkapsulace
- Každý paket může mít více značek
koncept zásobníku značek
's' bit indikuje dno zásobníku
- MPLS LSR **vždy** přepíná paket podle značky, která je na vrcholu zásobníku

Formát značky



- **4 oktety**

Hodnota značky - 20 bitů

Experimentální bity (QoS) - 3 bity

S (bottom of stack) - 1 bit

TTL (Time to live) - 8 bitů

Umístění značky

Packet-over-SONET/SDH



Ethernet



Frame Relay PVC



ATM PVC



(další buňky)




ATM label switching



(další buňky)





Frame-mode MPLS Přiřazení značek a jejich distribuce

Frame-mode MPLS

- Označené pakety se vyměňují jako rámce na layer-2

Frame-relay, Ethernet, ATM PVC

- **Shim** záhlaví mezi layer-2 a layer-3 záhlavím
- MPLS je identifikován v layer-2 záhlaví pomocí Ethertype, PPP protocol field, frame-relay NLPID

Frame-mode - přiřazení značek

- Frame-mode LSR používá pro přiřazení značky **independent control**

jakmile se FEC objeví v IP směrovací tabulce (resp. FIB) je mu okamžitě přiřazena značka a mapování je uloženo do LIB

LSR může přepínat pakety na next-hop, který ještě nemá značku pro FEC

Frame-mode - distribuce značek

- **Přiřazené značky jsou distribuovány sousedním uzlům**

pouze značka od souseda, který je zároveň next-hop se umístí do LFIB (Label Forwarding Information Base)

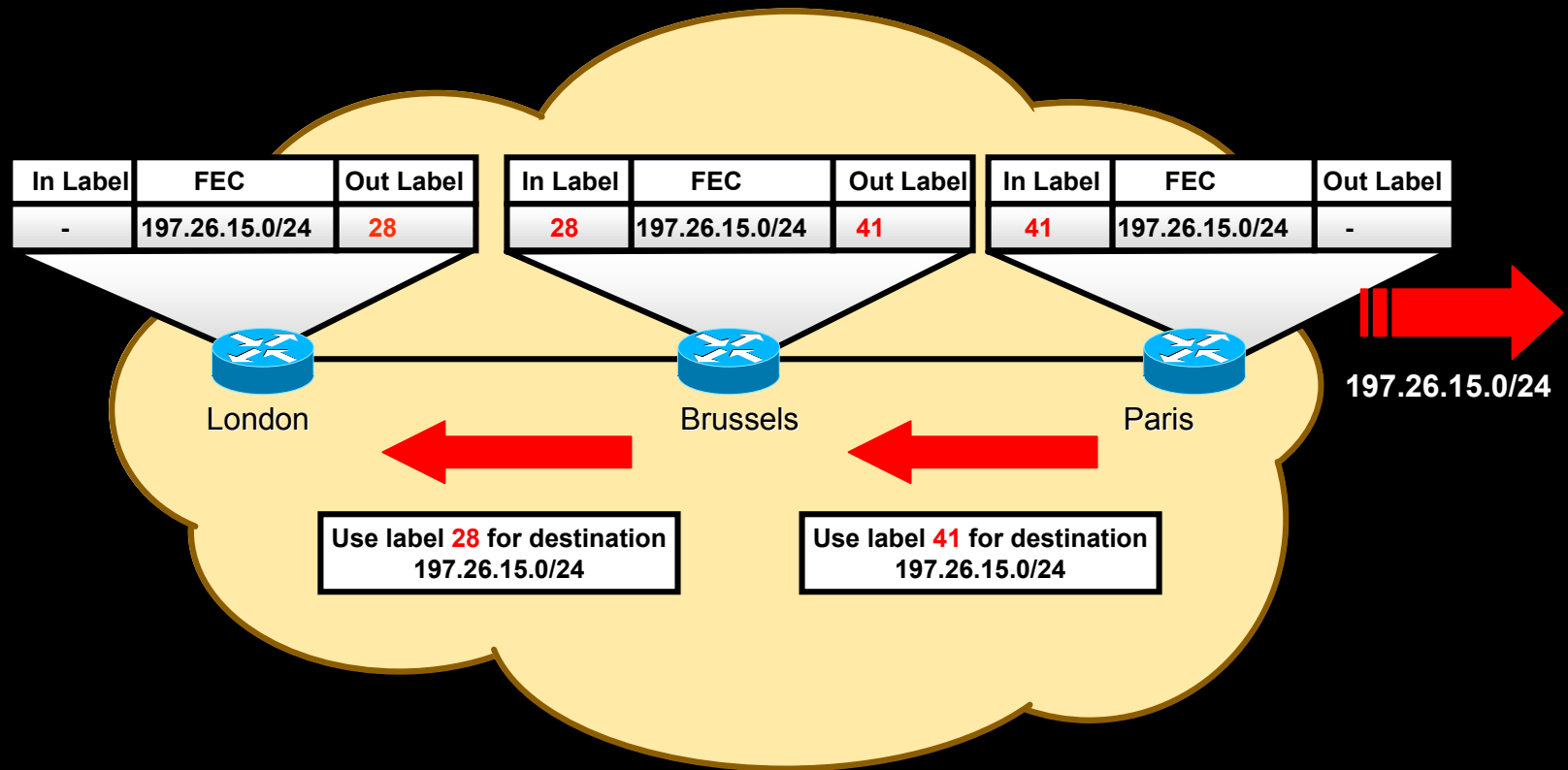
všechny značky jsou v LIB (Label Information Base)

- **Unsolicited downstream** distribuce značek

unsolicited - značky jsou distribuovány automaticky

downstream - LSR přiřazuje značky, které upstream soused používá pro přepínání paketů na downstream
FEC

Frame-mode - distribuce značek



Distribuce značek - protokoly

- **MPLS architektura nepředpokládá, že bude existovat pouze jeden protokol**

TDP/LDP

IP destination unicast

RSVP

traffic engineering

BGP

MPLS VPN


Udržování značek - frame mode

- Frame-mode LSR používá **liberal retention mode**

LSR udržuje pro danou FEC značky od všech sousedů

větší nároky na paměť a počet značek (label space)

rychlejší konvergence



Cell-mode MPLS

Přiřazení značek a jejich distribuce

Cell-mode MPLS

- **ATM přepínače nemohou provádět MPLS label ani IP lookup**

Pakety jsou přenášeny rozložené do buňek na layer-2

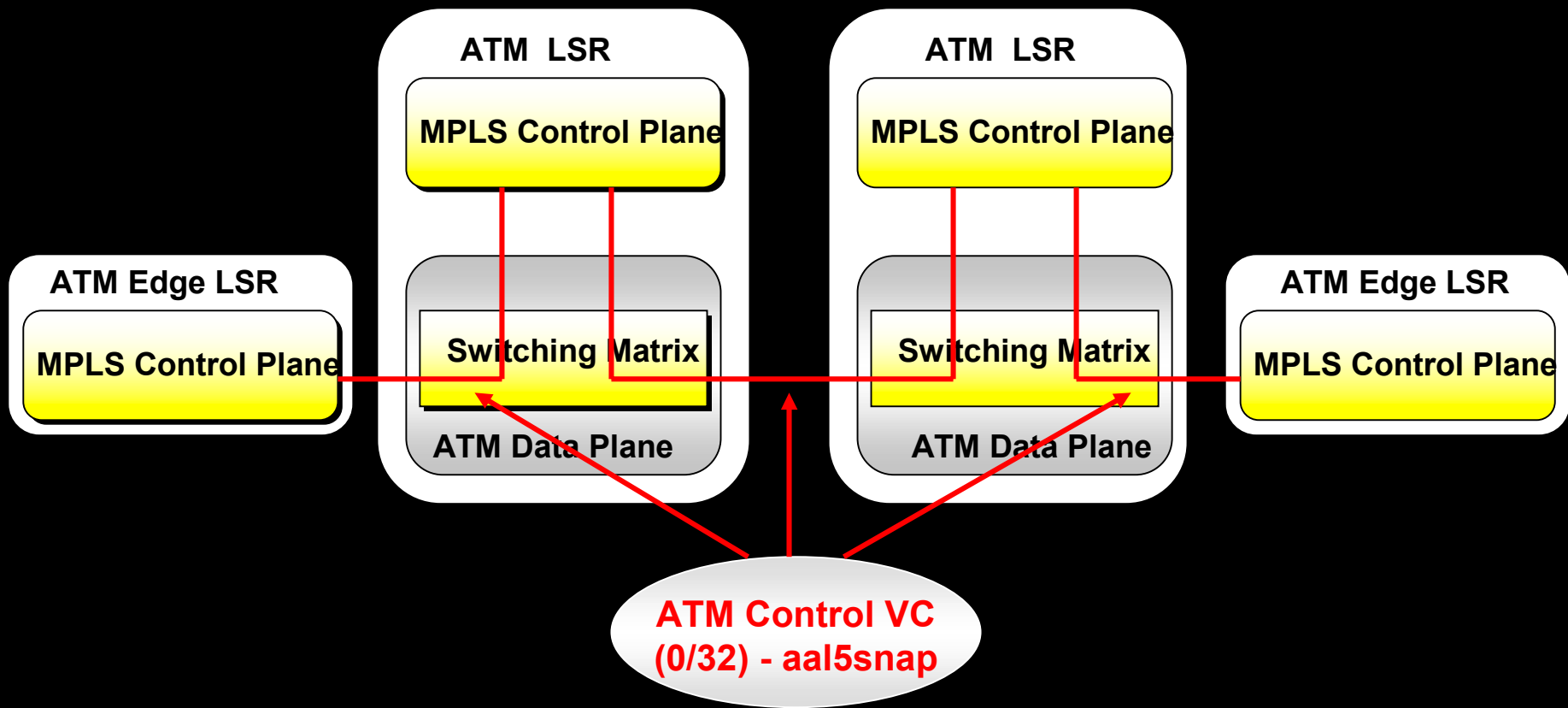
VPI/VCI pár je použit jako značka

- **ATM přepínače nemohou přímo přepínat IP pakety**

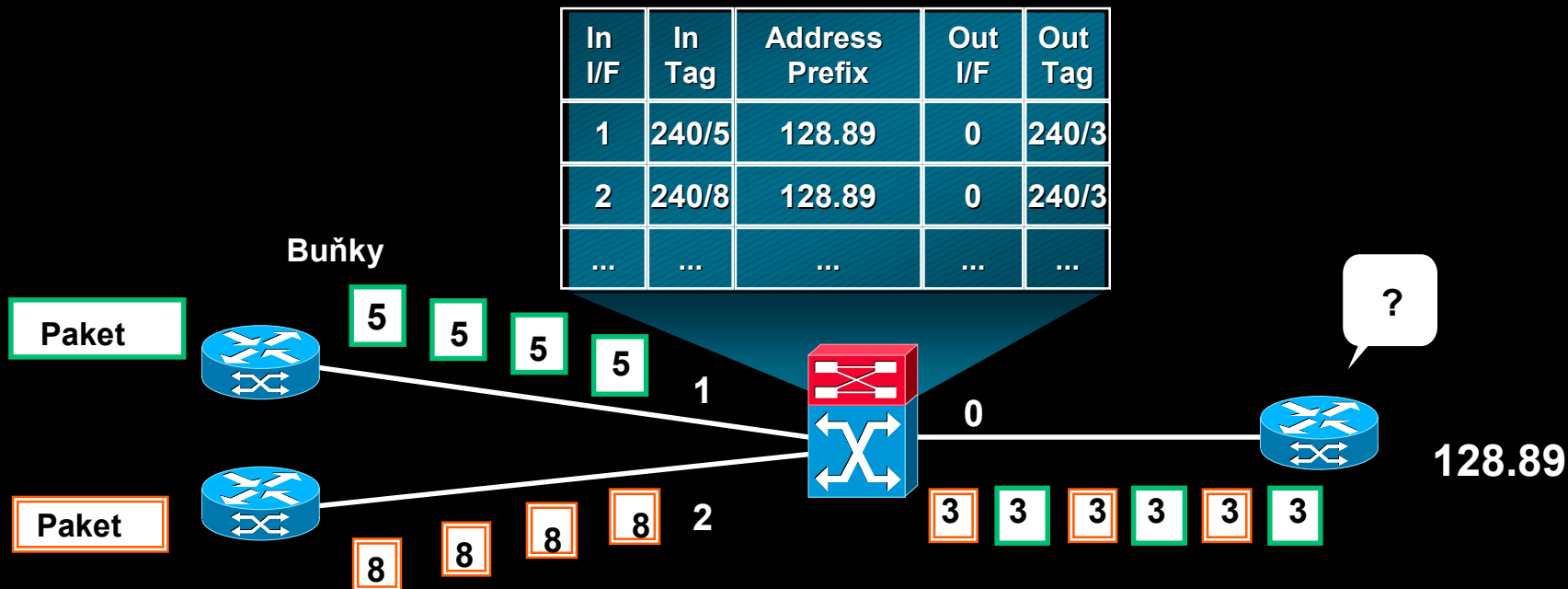
jsou vyžadovány virtuální okruhy

pro výměnu režijní informace mezi sousedními ATM LSR se používá speciální VC

MPLS - řídicí VC



Cell-mode - distribuce značek

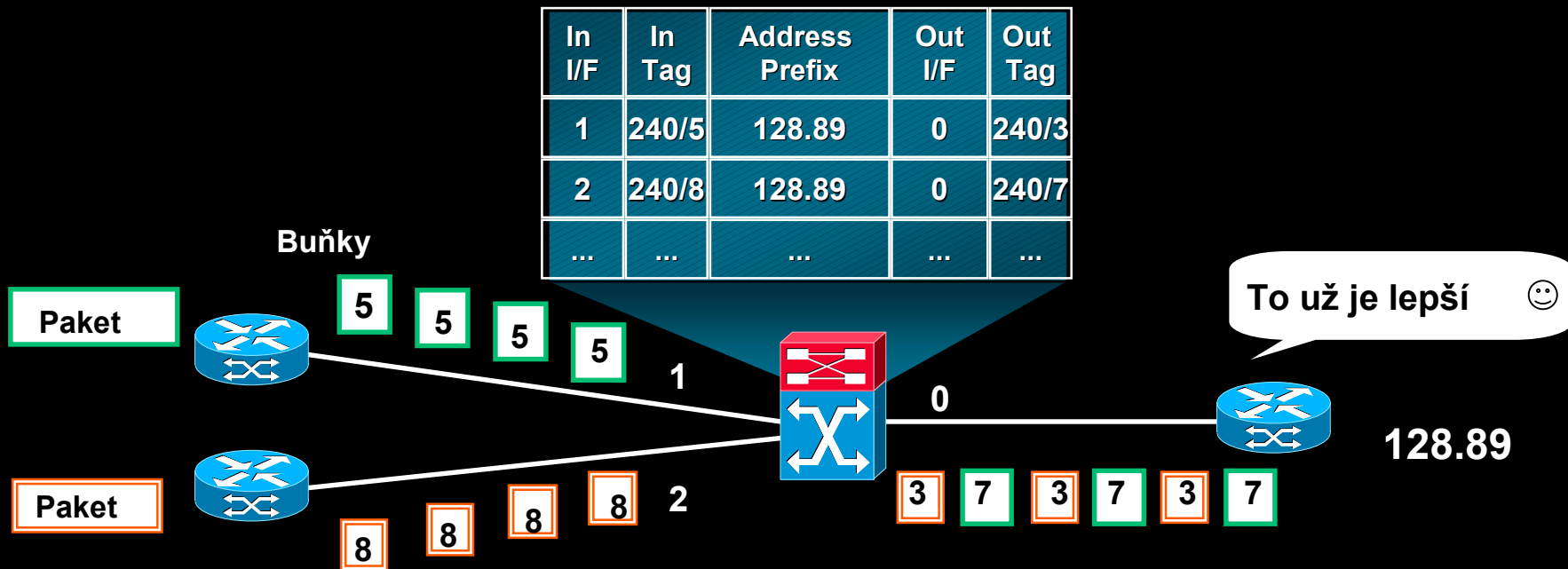


Unsolicited Downstream metoda pro cell-mode nefunguje !

Cell-mode - distribuce značek

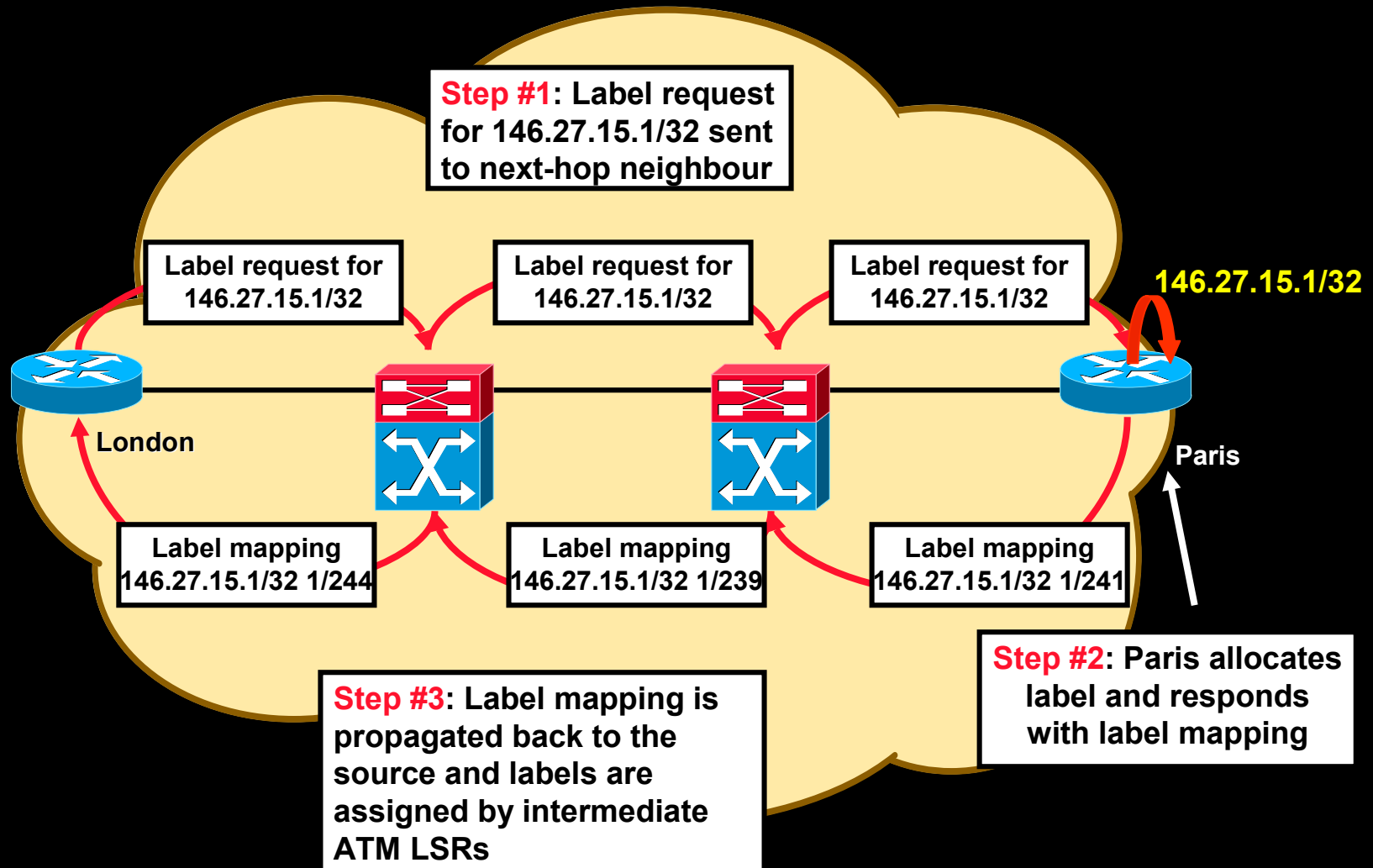
- Cell-mode používá přidělování a distribuci značek **Downstream on Demand** a **Ordered Control**
 - unikátní značka** pro každého upstream souseda
 - ordered control** existuje jasná posloupnost

Cell-mode - distribuce značek



Unikátní značka (VPI/VCI) pro každého 'upstream' souseda

Cell-mode - distribuce značek



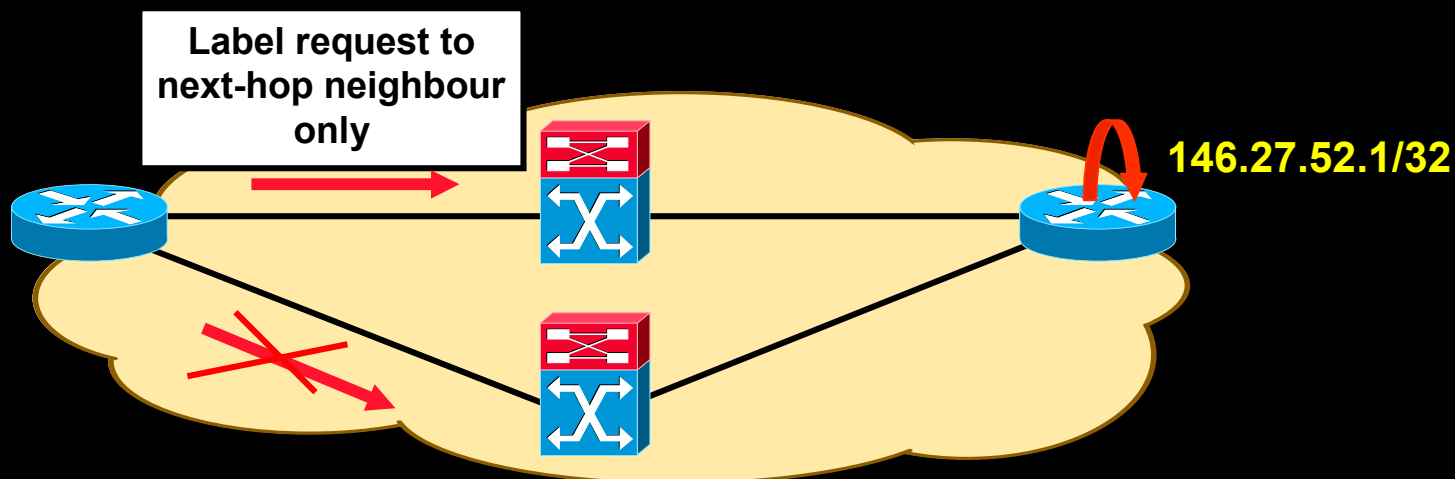
Udržování značek - cell mode

- Cell-mode používá **Conservative** label retention mode

udržují se pouze značky aktuálně používané k přepínání

dosaženo tím, že explicitní dotaz jde jen na next-hop

Dvě cesty do
146.27.52.1/32

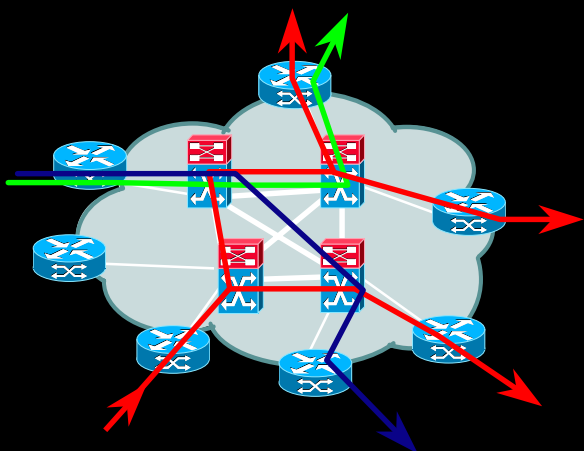


A man in a white shirt and red tie is holding a large red pipe that arches across the top of the frame. The background is a colorful, abstract landscape with yellow, blue, and green areas. The title 'Aplikace MPLS' is centered in the lower half of the image.

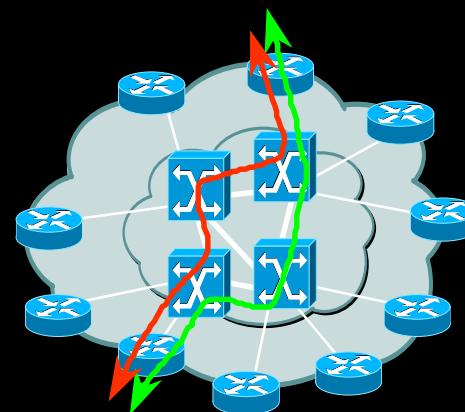
Aplikace MPLS



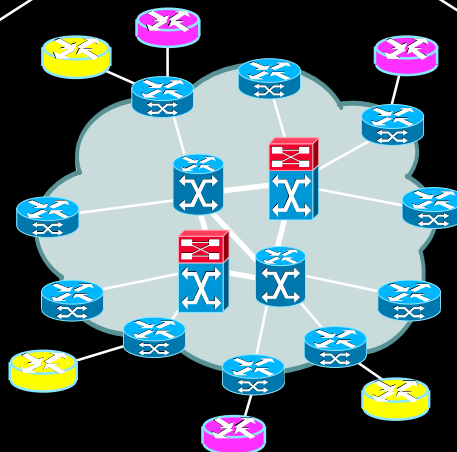
Základní MPLS aplikace



IP+ATM intergrace



Efektivní enkapsulace

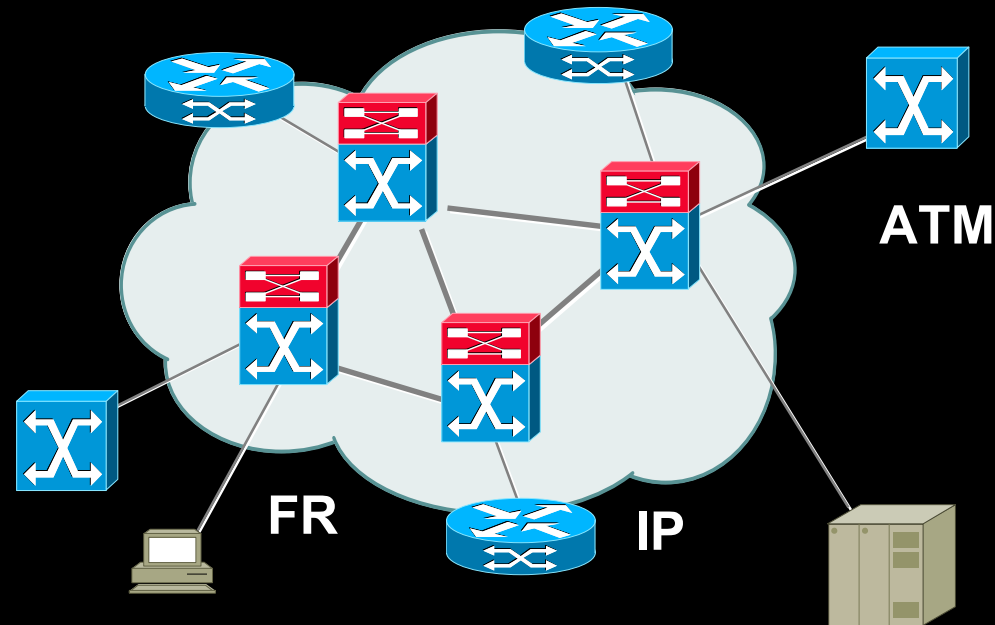


MPLS VPN

www.cisco.com

IP+ATM integrace

- Multifunkční ATM páteř s vazbou na IP
- ATM, FR, pevné okruhy, MPLS, IP VPN
- Dostupné platformy: BPX, MGX, Catalyst 8500 MSR, LS1010, směrovače, IOS & MPLS



MPLS Traffic Engineering

- **Vytváření 'traffic trunks'**
Provoz směřován stejnou cestou se stejnou CoS
- **Cesta je sestavena pomocí RSVP**
Routing with Resource Reservation (R³)
- **Rozšířené vlastnosti IGP (IETF drafty)**
propagují informace o zdrojích sítě
- **Dynamické přesměrování v případě výpadku**
Pružné řízení politiky, rozdělování zátěže



Virtuální privátní sítě

Virtuální privátní sítě

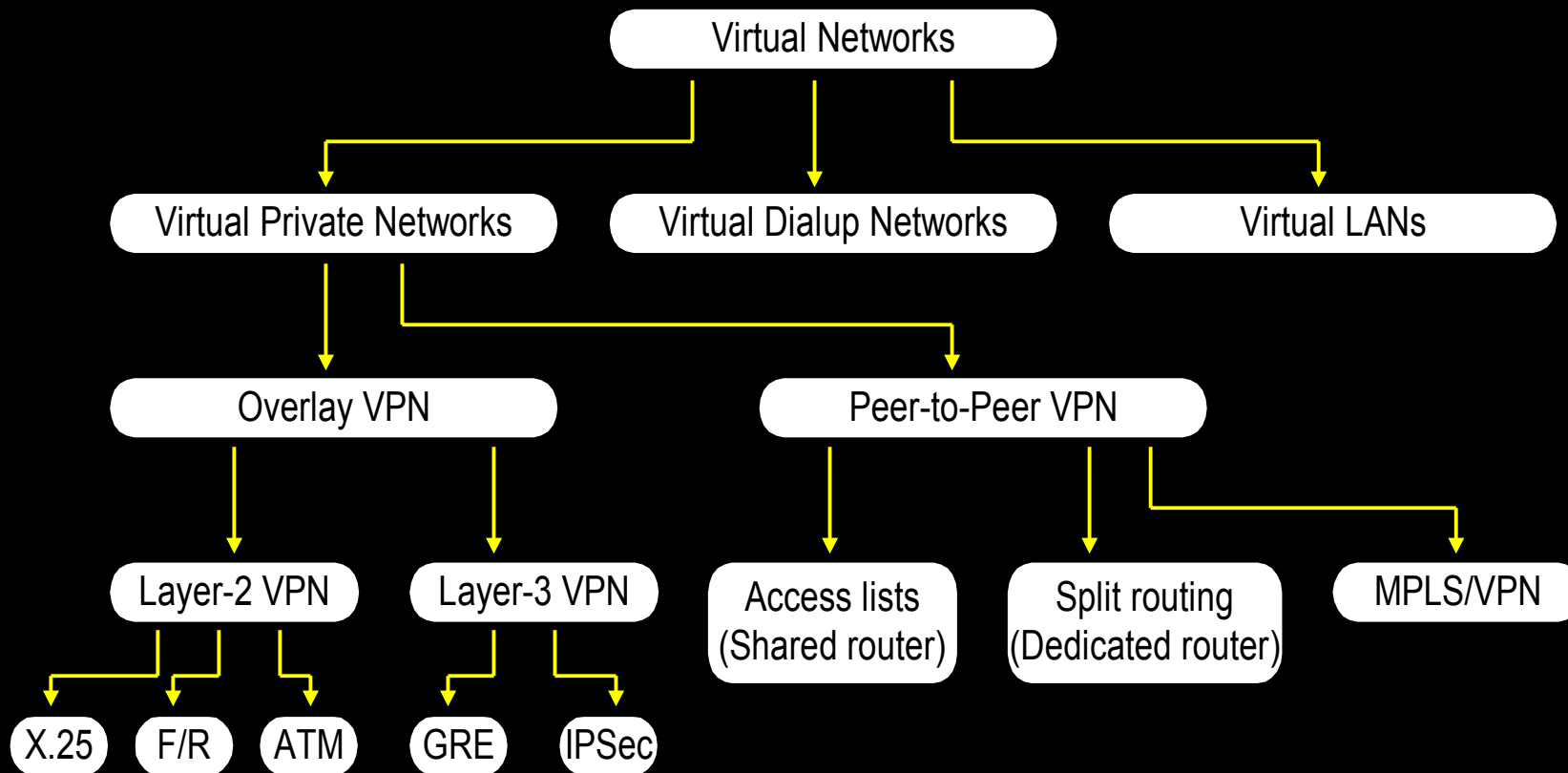
- Infrastruktura poskytující privátních síťové služby založené na veřejné infrastruktuře

Nejedná se o nový koncept

Layer-1, layer-2, layer-3

Nutnost řešit překrývající se adresní prostory

Virtuální privátní sítě - členění



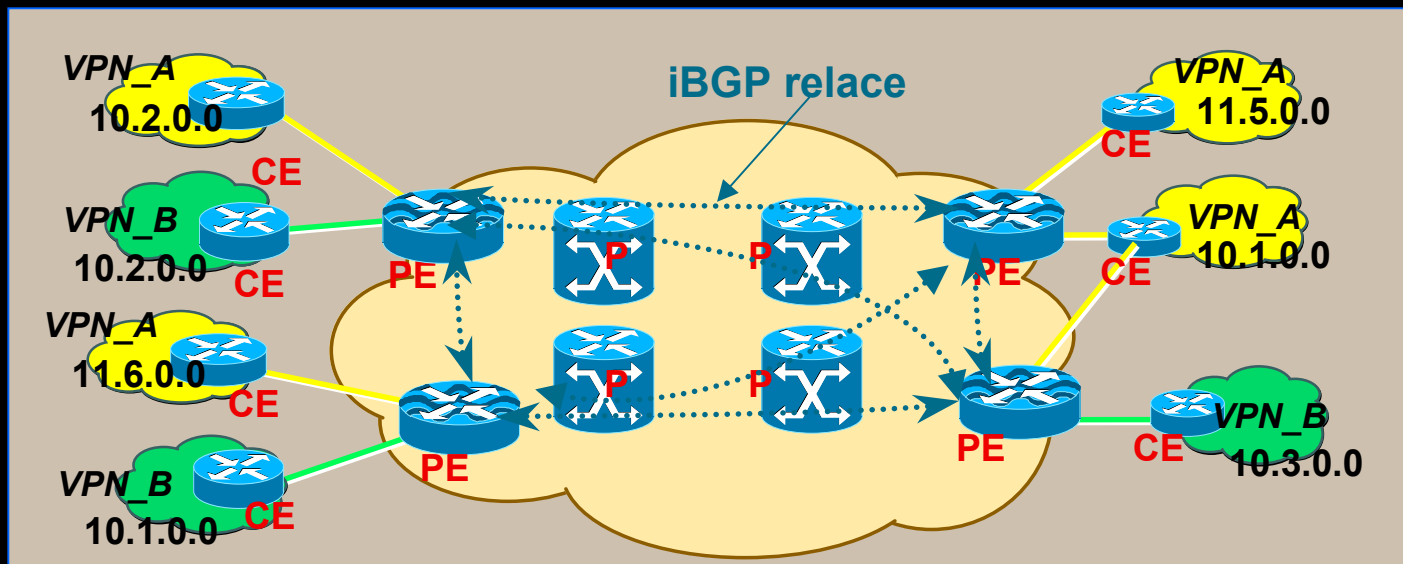
VPN - MPLS/VPN model

- **Spojuje výhody překryvného a peer-to-peer modelu**
 - překryvný (bezpečnost a izolace jednotlivých zákazníků (VPN))
 - peer-to-peer (jednoduché směrování, přidávání atd.)
- **PE směrovače drží informace pouze pro připojené VPN**
 - snižuje nároky na PE
 - zvyšuje škálovatelnost
- **MPLS je použito pro přepínání v páteři**
 - plný routing v páteři není potřeba



MPLS/VPN - model propojení

MPLS VPN - model a komponenty



- P routery (LSR) jsou v páteři MPLS sítě
- PE routery (Hraniční LSR) používají MPLS směrem do páteře a IP směrem k CE routerům
- P a PE routery sdílí společný (globální) IGP
- PE routery mezi sebou používají MP-iBGP

MPLS/VPN - model propojení

- **VPN je množina lokalit sdílejících stejnou směrovací informaci**
- **Lokalita může patřit do více než jedné VPN**
- **Analogie Closed User Group (CUG)**

VPN Routing & Forwarding (VRF)

- **PE router obsluhují několik oddělených směrovacích tabulek**

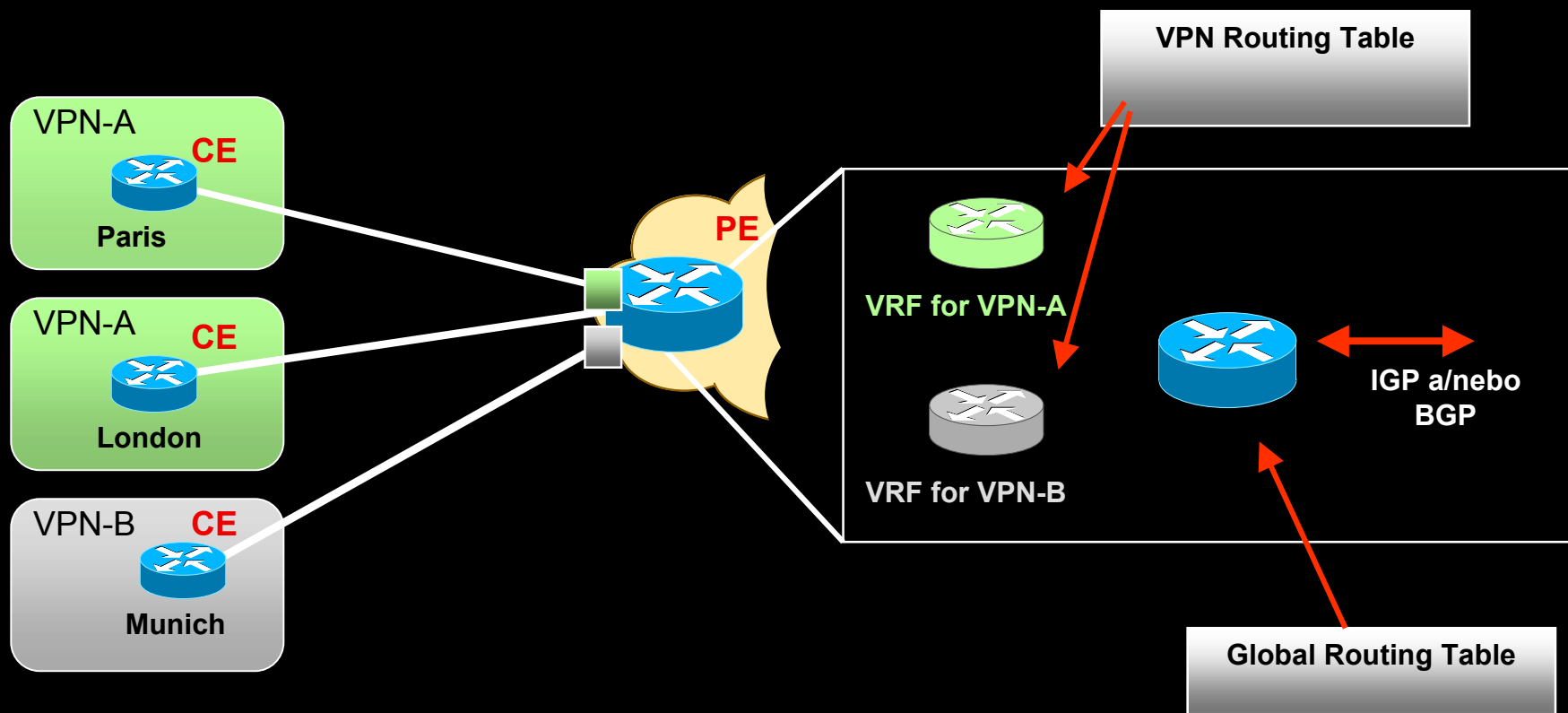
Globální směrovací tabulka

- obsahuje všechny PE a P routes (možná BGP)
- naplňovaná IGP, který běží v páteři
- neobsahuje VPN směrovací informaci

VRF (VPN routing & forwarding)

- směrovací tabulky příslušných VPN
- každé rozhraní může být přiřazeno pouze do jednoho VRF
- **Velká změna proti původním implementacím**

VPN Routing & Forwarding (VRF)

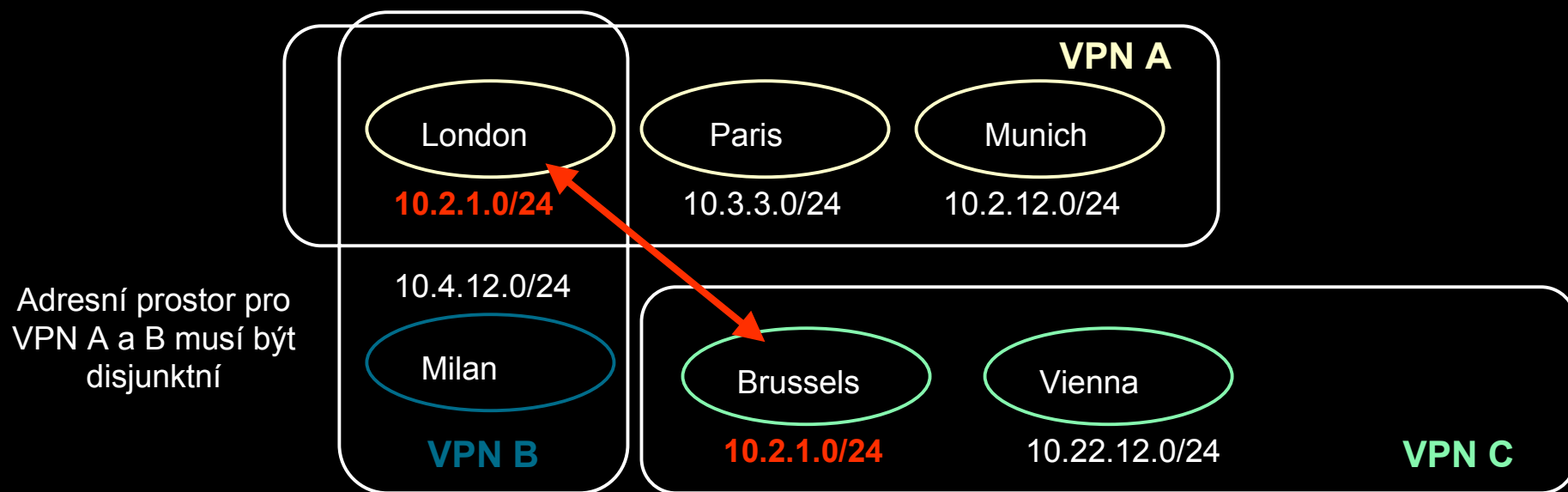


Na VRF je možno pohližet jako na 'virtuální router'

MPLS/VPN - model propojení

- **Privátní adresace a překrývající se adresní rozsahy nejsou problém**

v případě lokalit patřících do více lokalit je třeba dodržet určitá pravidla



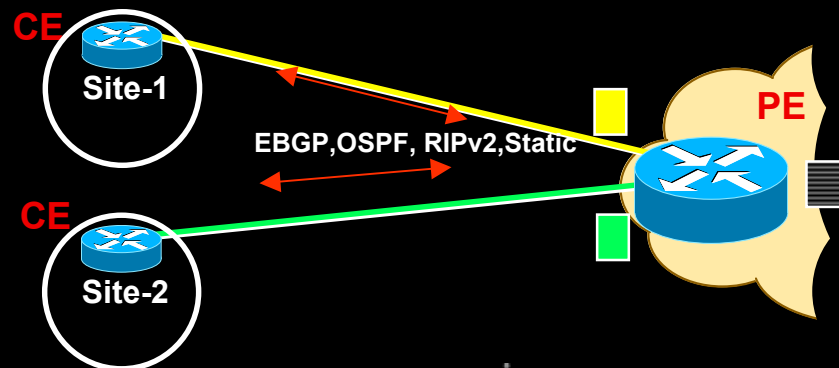
Interakce mezi PE a CE

- **Podporované směrovací protokoly mezi PE a CE**

RIP V2, OSPF, BGP-4 (externí), statické směrování, connected

- **Oddělený směrovací kontext pro každé VRF**

oddělený kontext v jednom procesu (BGP-4, RIP V2)
oddělené procesy (OSPF)

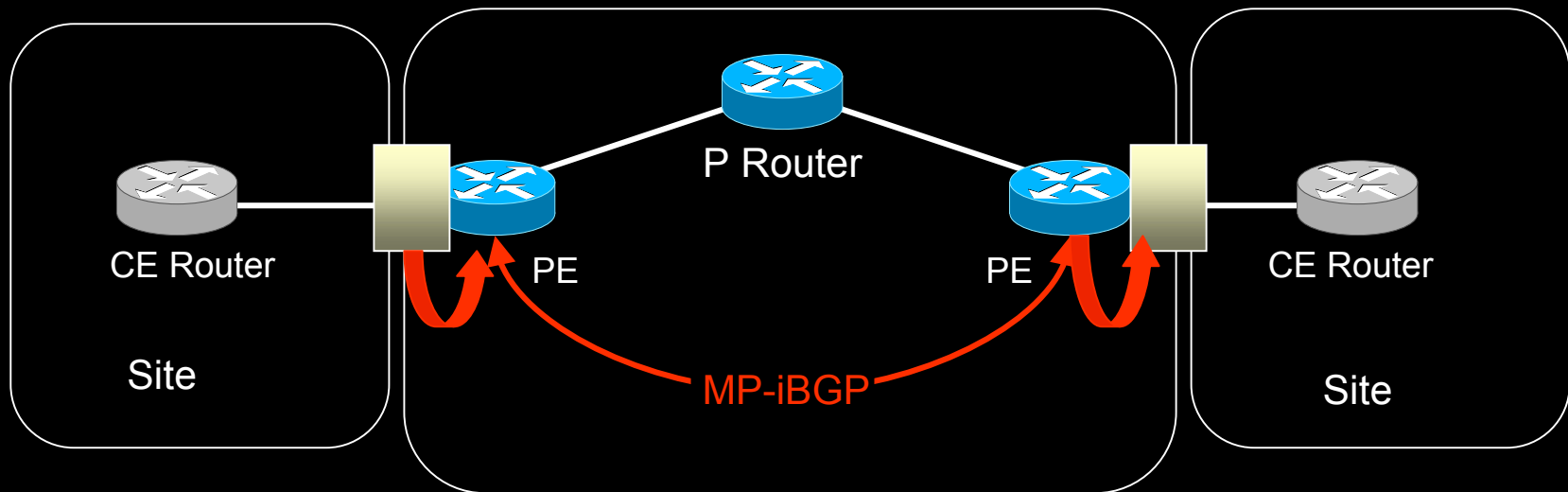


VRF - distribuce směrovací informace

- PE routery distribuují lokální VPN informace skrz MPLS/VPN páteř

využití MP-iBGP a redistribuce z VRF

přijímající PE **importuje** směrovací informaci do VRF



VRF - přenos směrovací informace pomocí MP-iBGP

- **Přijímající PE router musí znát:**
 - odkud daný route pochází;
 - do kterého (kterých) VRF má být route umístěn
 - jak rozlišit mezi duplicitními adresami
- **Unikátnost IPv4 prefixů je dosažena pomocí parametru **Route Distinguisher****
 - RD (64 bit) identifikátor
 - nový typ adresy VPN-V4 prefix - RD + IPv4 Prefix

Rozšířené komunity

- Místo vzniku a identifikace 'místa určení' je dosažena pomocí BGP **Extended Community Attribute**

- BGP

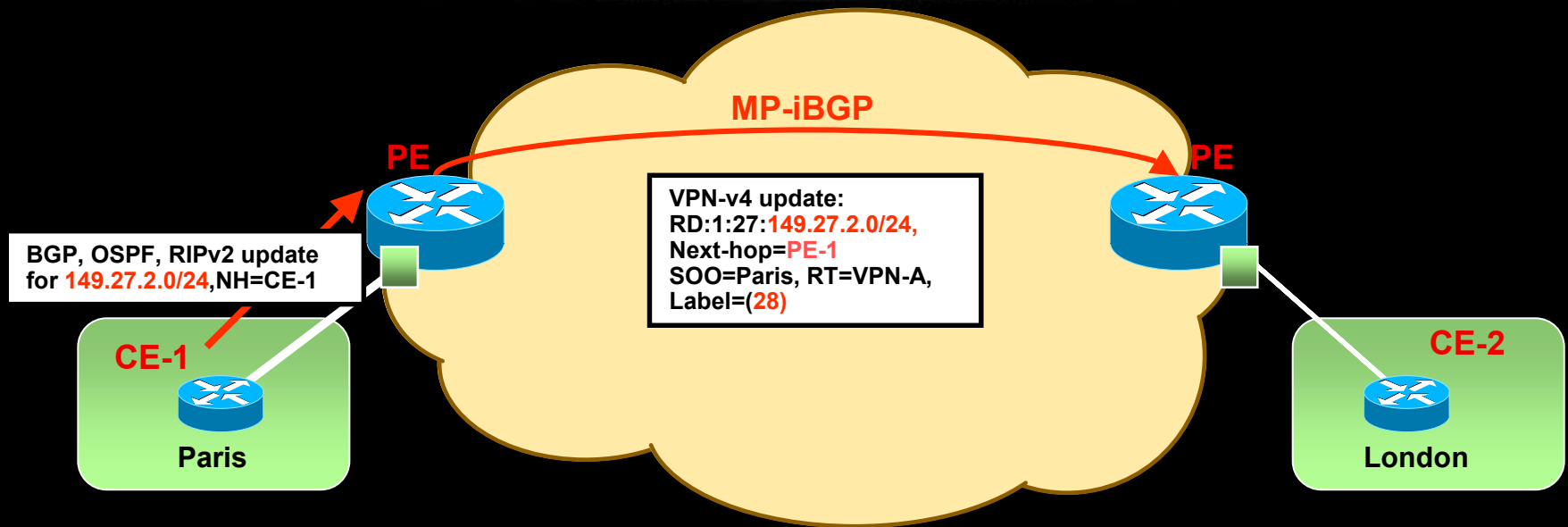
SOO (Site of Origin)

používá se pro identifikaci lokality odkud daný route pochází

Route Target

určuje, kam má být daný route exportován

Co dělá vysílající PE?



- **PE router přeloží IPv4 route do VPN-V4 routu**
 - Přiřadí RD, SOO a RT podle konfigurace
 - Přepíše Next-Hop atribut (na PE loopback)
 - Přiřadí VPN label
 - Pošle MP-iBGP update všem PE sousedům

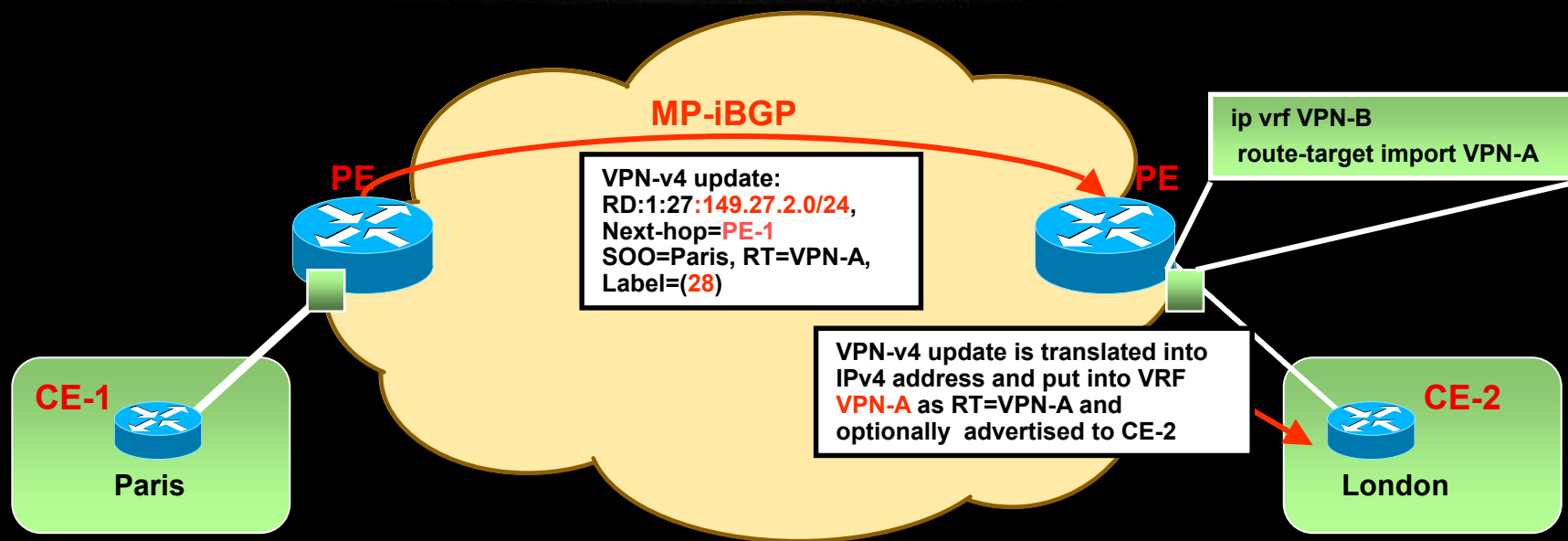
VPN značka

- **VPN značka identifikuje:**

Výstupní interface nebo VRF (tam, kde se musí udělat lookup, aggregate/connected)

BGP značka je druhá značka ve stacku, není používána pro přenos paketu v MPLS páteři, ale pouze na PE

Co dělá přijímající PE?



- **Překládá VPN-V4 adresy do IPv4 adres**

Importuje routy do příslušných VRF v závislosti na RT

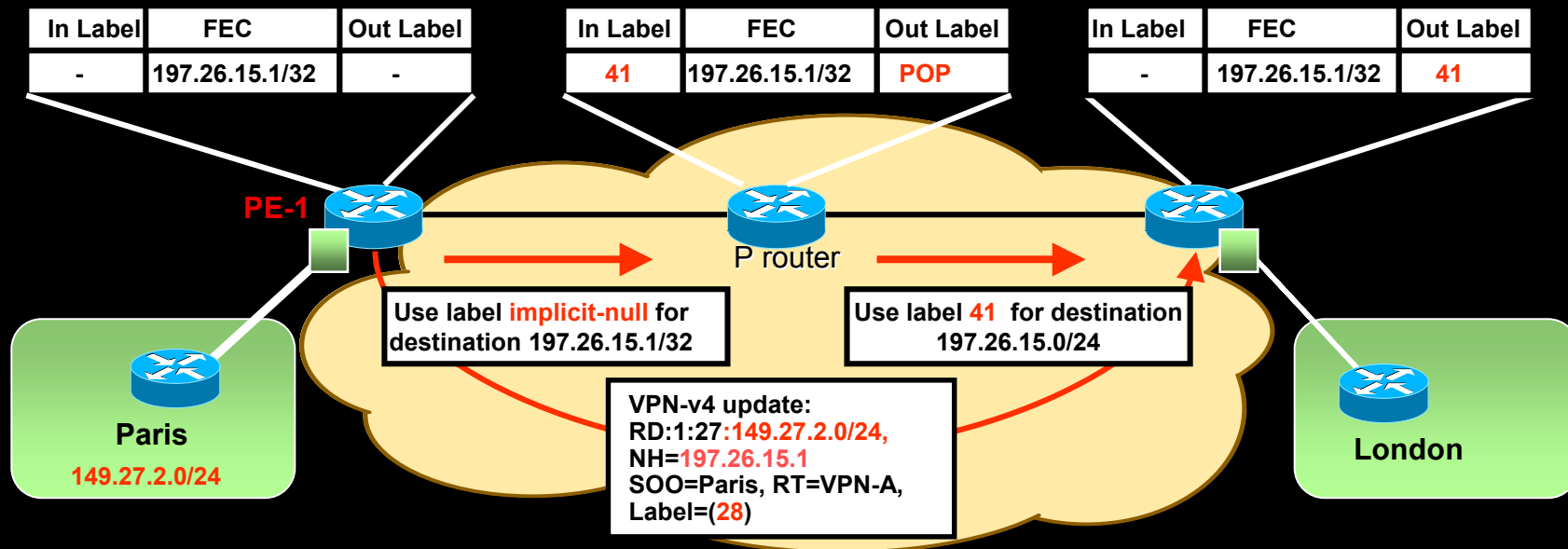
MPLS/VPN - doručení paketu

- **Pro přenos paketu se používá stack dvou značek**

Značka na vrcholu určuje BGP Next-Hop (interior label)

Druhá značka určuje výstupní rozhraní nebo VRF (exterior VPN label)

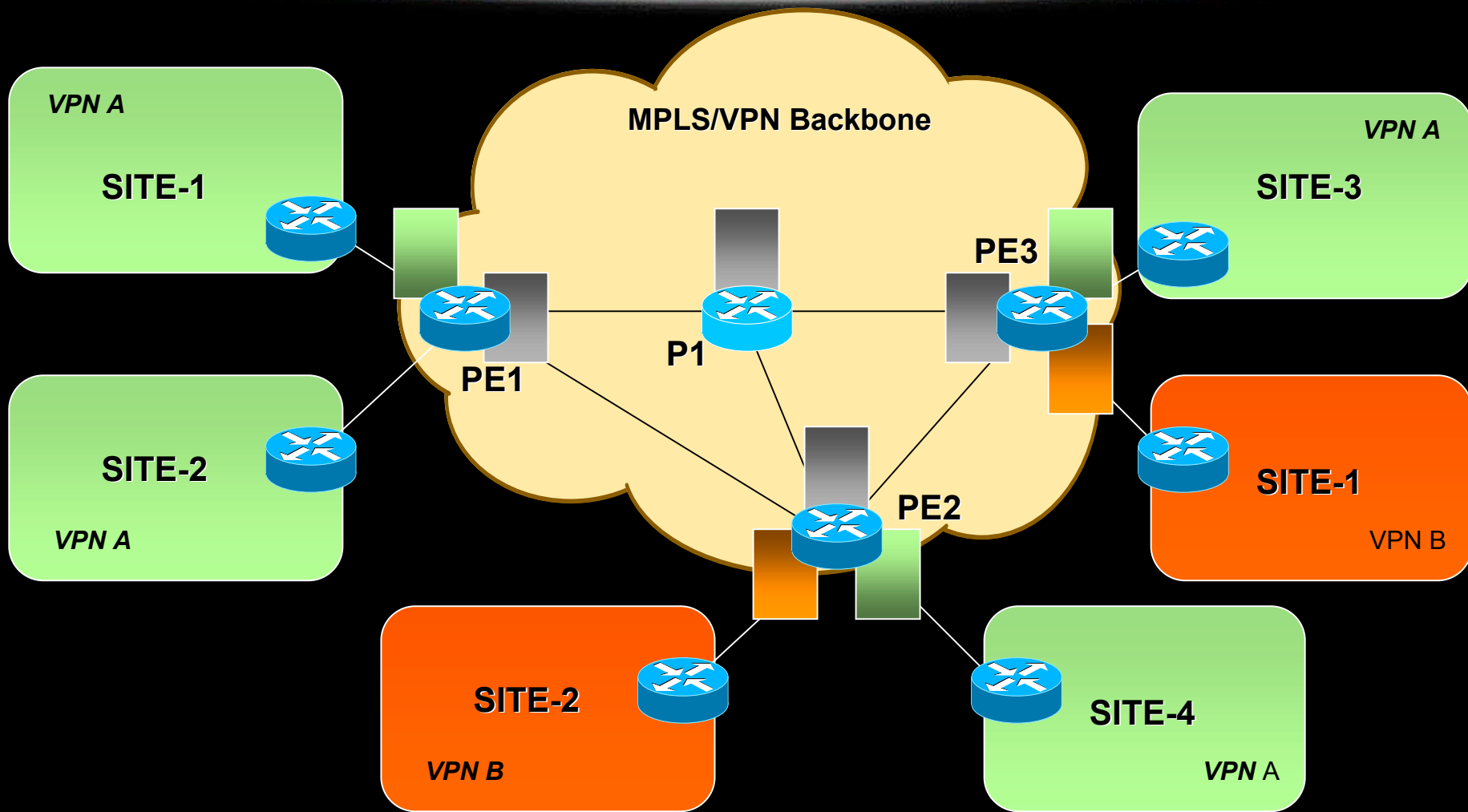
MPLS/VPN - doručení paketu



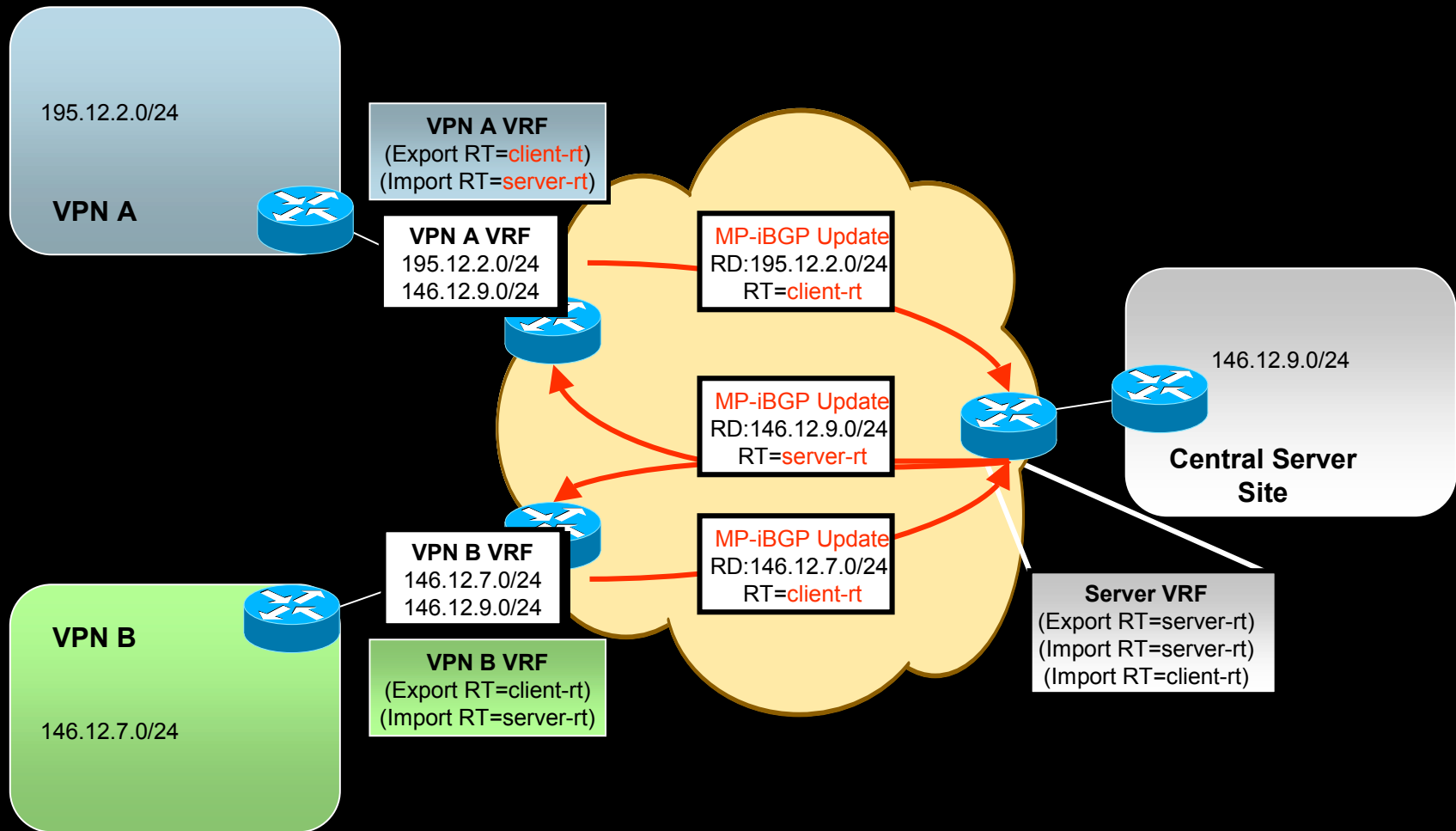


MPLS/VPN - příklady

Základní model VPN

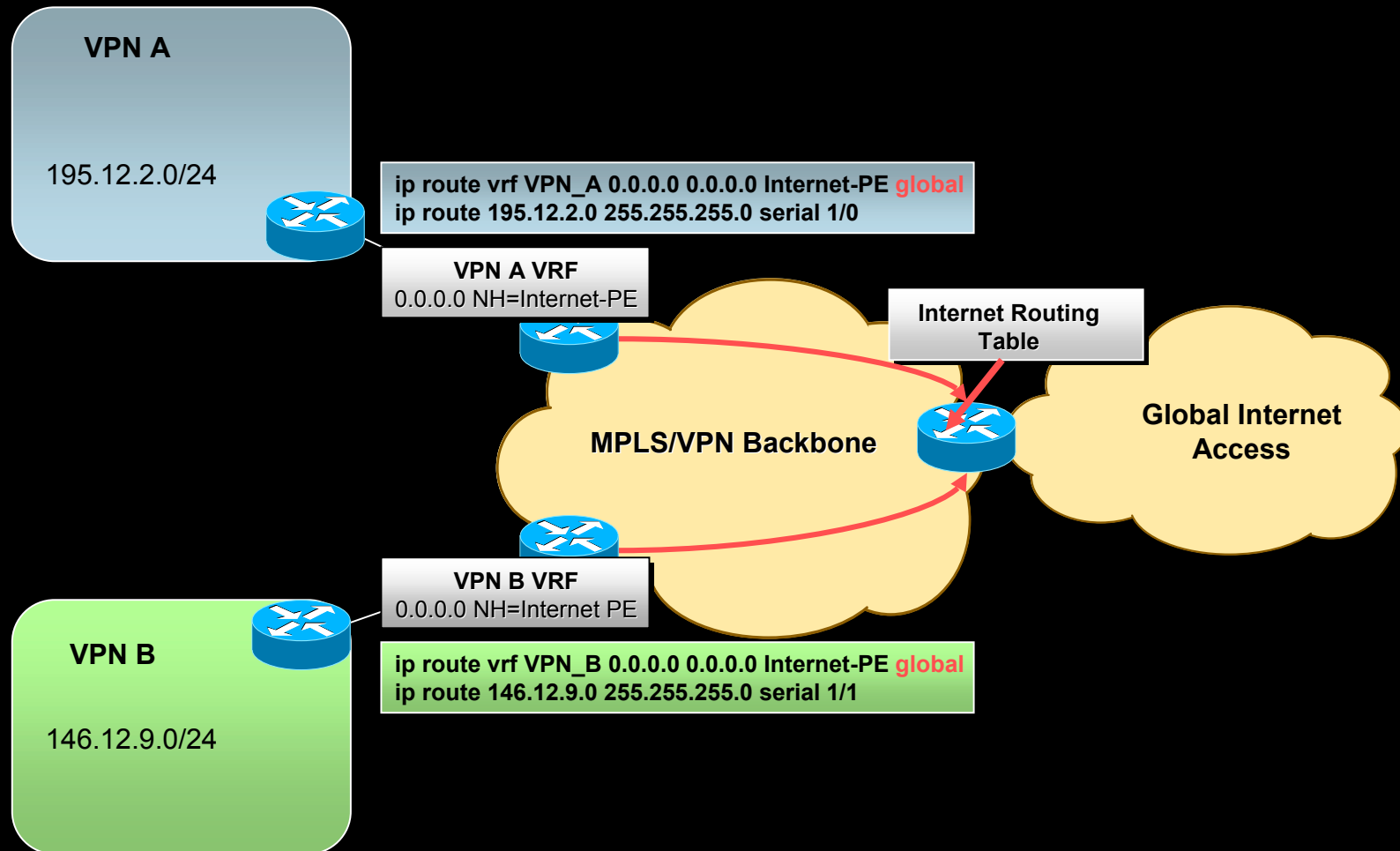


MPLS/VPN - centrální servery



MPLS/VPN - Internet konektivita

Static Default Route



CISCO SYSTEMS



EMPOWERING THE
INTERNET GENERATIONSM