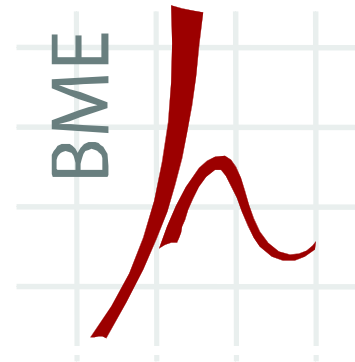


Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar

Mérnök informatikus szak, mesterképzés – Hírközlő rendszerek biztonsága szakirány
Villamosmérnöki szak, mesterképzés - Újgenerációs hálózatok szakirány



BMEVIHIM134 Hálózati architektúrák
3. Az újgenerációs hálózati (NGN) koncepció:
Követelmények – gerinchálózati technológiák I.
IP, IP/MPLS, IP/MPLS TE

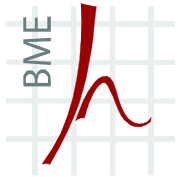
*Jakab Tivadar BME Híradástechnikai tanszék
2015*

IP és IP/MPLS áttekintés

- IP hálózati alapok (néhány nyilvános hálózati vonatkozás)
 - routing
 - CIDR
 - IGP,EGP
 - peering, tranzitálás
 - QoS
 - IntServ
 - DiffServ
- IP/MPLS
 - címke alapú továbbítás
 - IP/MPLS TE
 - tunnelek
 - védelmek
 - DiffServ-aware IP/MPLS TE

BGP – AS path

- az egyes prefixekhez vezető útvonal menti AS-ek számának listája
- az első AS (originator) által hirdetett útvonalban csak a saját AS-e szerepel
- egy BGP router (BGP speaker) az útvonal hirdetésekor a kapott útvonal elé beilleszti saját AS számát
- több útvonal is előállhat
- az AS path hossza fontos jellemző, a rövidebb a jobb
- az útvonalak hirdetése minden irányba történik
- a hurkok kialakulásának megakadályozása érdekében egy BGP router eldobja azokat az útvonalakat, amiben saját AS-e szerepel
- egyes prefixekhez több AS path is tartozhat



KAPCSOLATÁLLAPOT ALAPÚ IGP



Útvonalválasztó algoritmusok

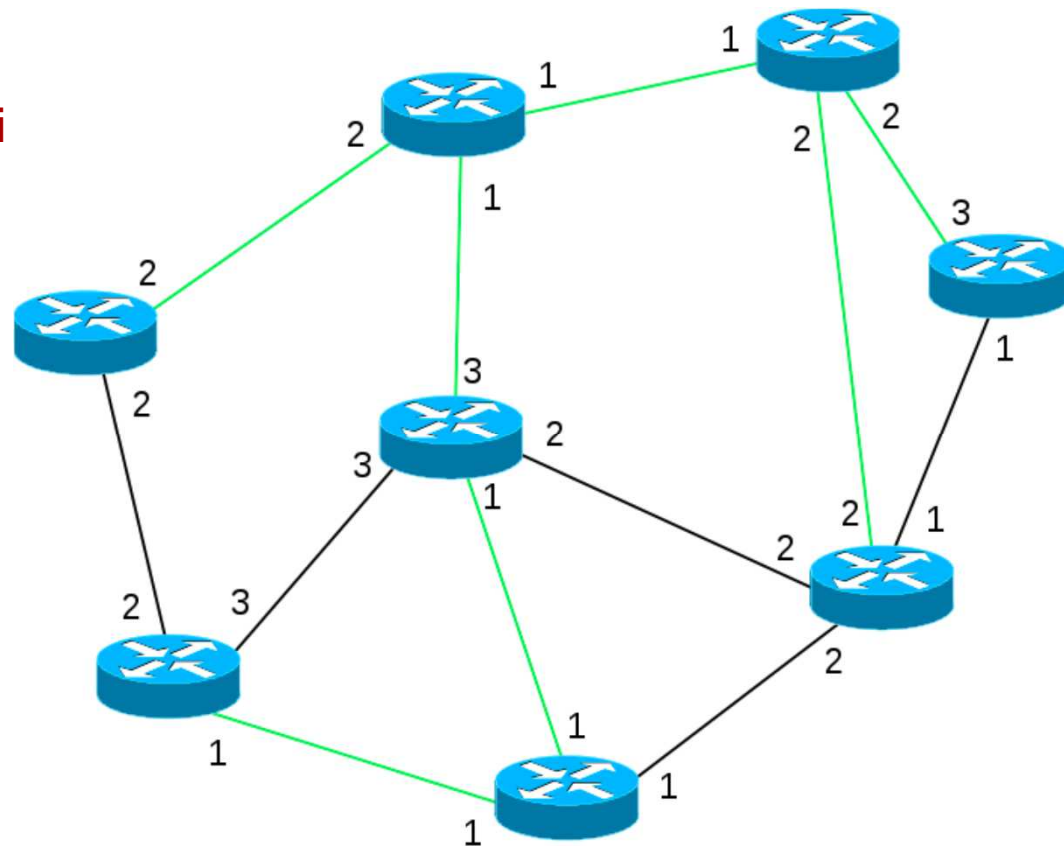
- **routing tábla**
 - legrövidebb út egy adott végponthoz (*IP prefix*)
 - nem a teljes út, csak a következő szakasz (*next hop*)
 - FIB (*Forwarding Information Base*)
- **két nagyobb routing protokoll csoport**
 - távolságvektor alapú (*distance vector*)
 - egy router nem lát a szomszédok „mögé”, a szomszédok FIB-jei alapján „látja” a hálózatot
 - RIP, EIGRP
 - kapcsolatállapot alapú
 - hálózat szerkezetére vonatkozó információk hirdetésére alapoz
 - OSPF, IS-IS

Kapcsolatállapot alapú routing

- minden router ismeri a hálózat szerkezetét, a linkek, csomópontok elrendezését
 - lényeges, hogy a ugyanaz a hálózati kép alakuljon ki az összes routerben, enélkül hurok vagy más irányítási inkonzisztencia jöhet létre
- a hálózat szerkezetének ismeretében minden router optimális utat tud
- számolni a célhálózatok felé
 - az adott router mint gyökér szerepel az elérési fában
- a hálózat állapotának változásakor a változás ténye terjed a hálózatban
 - link vagy router állapotának változását okozhatja például egy interfész állapotának változása, vagy időzítés lejárta
 - a változott linkkészleten minden router újraszámolja az egyes célhálózatok felé vezető utakat
 - a gyors konvergencia érdekében előbb állapotváltozási információ küldés, majd útvonalak újraszámolása a routerkeben

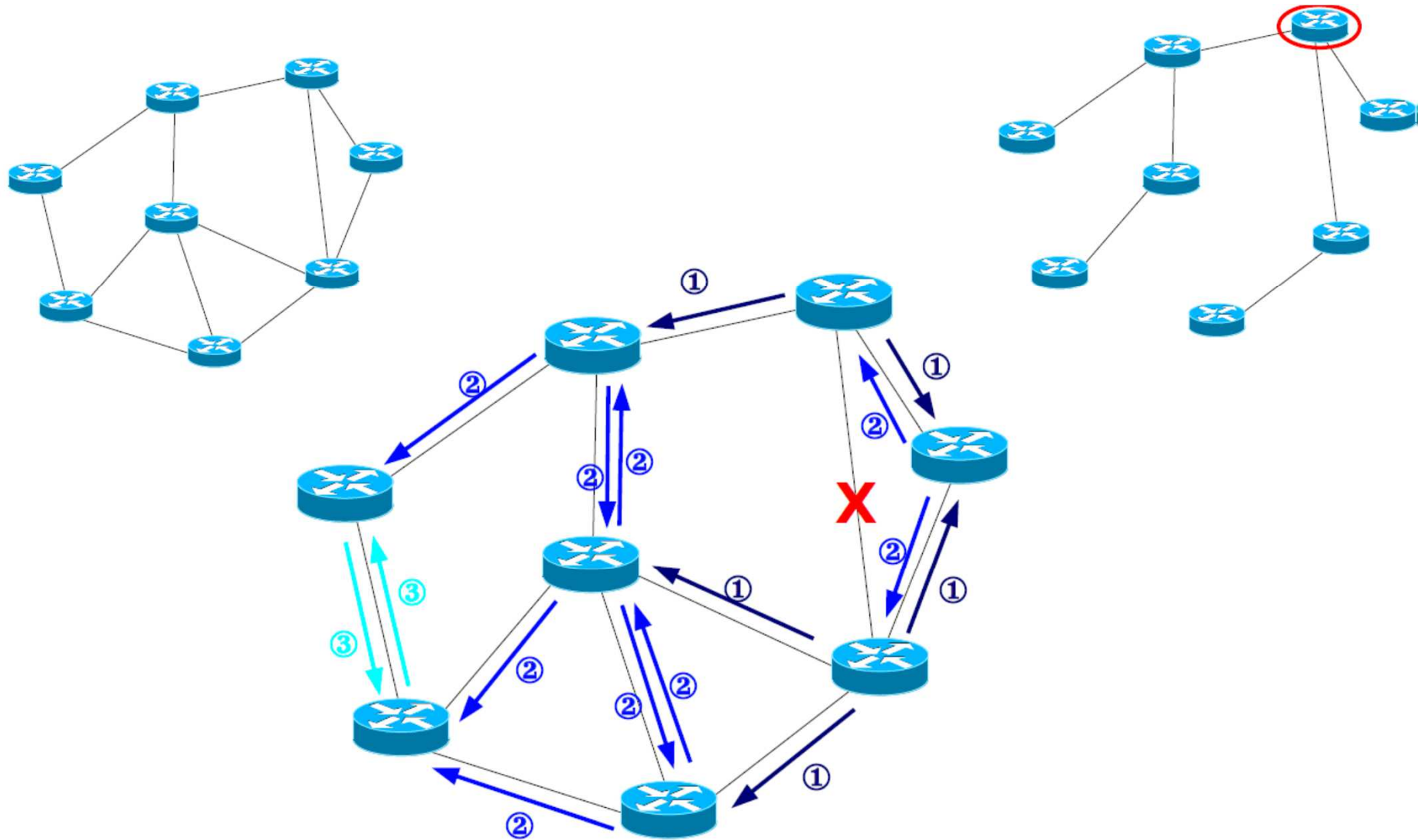
Linksúlyok, minimálutak

- az egyes linkek nem azonos tulajdonságúak, ezt figyelembe kell venni az utak számolásánál
 - a linkeknek "költsége" van
 - a költség lehet konfigurált, vagy származtatott érték
 - az utak kiszámolásánál a célfüggvény az út költségének minimalizálása

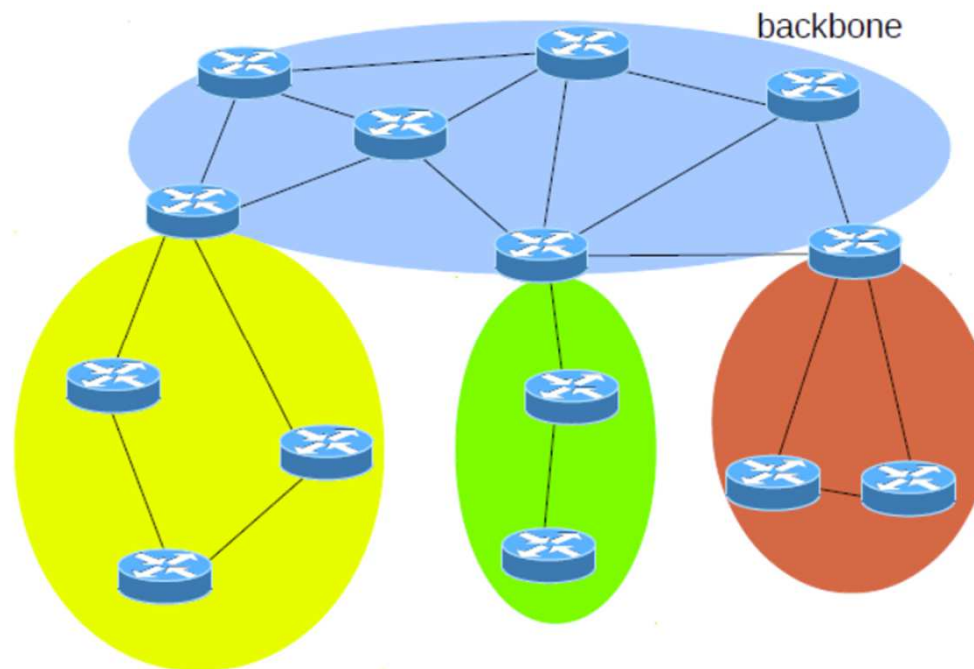


- a linkek és azok jellemzőinek halmazát az egyes routerek a linkállapot-adatbázisban tárolják (Link State Database)
 - ennek memóriaigénye nagyobb, mint a distance-vector protokolloké
 - az adatbázis mérete a hálózat méretének növekedésével nő
 - kb. a linkek mennyiségének növekedésével arányosan
 - az adatbázis méretének növekedésével az utak számolása is több CPU-t igényel
- az adatbázis méretének korlátozása a hálózat részekre bontásával érhető el

Állapotváltási információ terjesztése



Hálózat részekre bontása



- Hierarchikus szerkezet (csak hierarchikus tranzit)
- Elemi állapotinformációk hálózatrészeken belül
- Hálózatrészek között koncentrált információcsere (hálózatrészen belüli és hálózatrészek közötti rotuing funkciók megkülönböztetése)

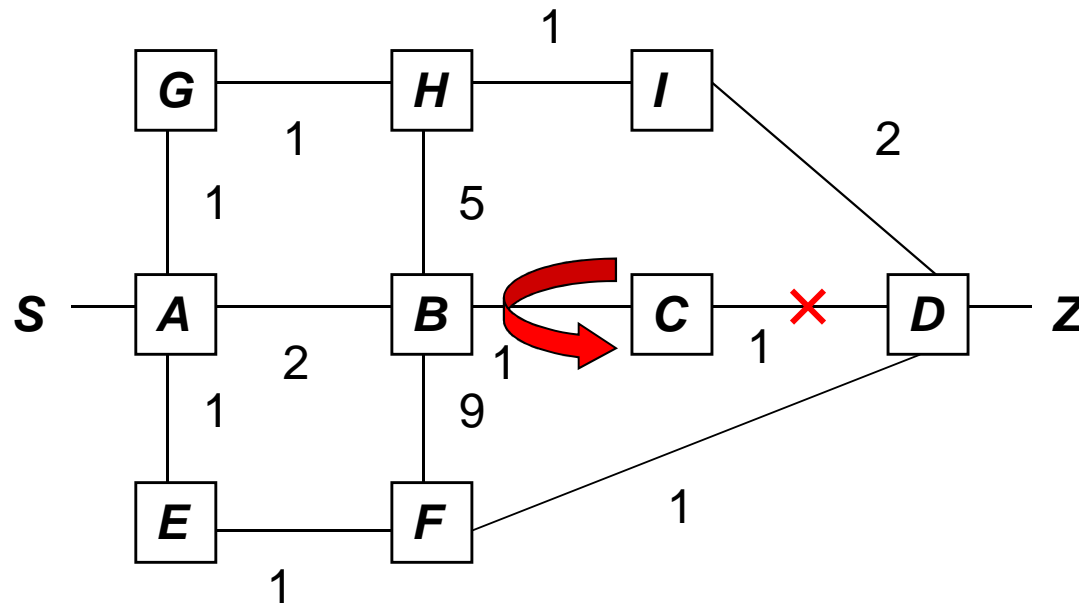
OSPF

- több másolatban létező elosztott adatbázismodellen alapul, lényeges, hogy ugyanaz a hálózati kép legyen mindegyik routerben (állapotváltozások terjedése!, ha nem akkor hurok vagy más inkonzisztencia)
- minden router elküldi környezetének jellemzőit (működő IF-ek, kapcsolódó IF-ek, linkek költségjellemzői, stb.) az összes többi routernek (LSA) egy elárasztásos mechanizmussal (*flooding*) – sorszámozott LSA-k, nyugtázott vétel
- ebből minden csomópont adatbázist épít (LSDB – Link State DataBase), és ennek alapján kalkulál minimálutat minden elérendő végponthoz (SPT - Shortest Path Tree)
- Ennek alapján FIB
- pl. OSPF, IS-IS

Kapcsolatállapot alapú

- Állapotváltozás detektálásakor a detektáló csomópont PDU-t generál és küld
- A PDU-t vevő csomópont
 - Nyugtázza a vételt (megbízható kommunikáció)
 - Ellenőrzi, hogy új információ érkezett-e
 - Ha igen továbbküldi, a saját LSDB-jét aktualizálja és a minimálút-meghatározást megismétli
 - Ha nem, akkor elbobja (nem küldi tovább)
- Útmeghatározás: Dijkstra algoritmus
- Skálázás: hierarchia (areak, ABR-ek)
- Gyorsítás: előbb küldés azután aktualizálás és számolás, hogy gyorsabb legyen a konvergencia, inkrementális SPT aktualizálás
- Terhelés szétosztás ECMP (Equal Cost Multipath) alapon

Kapcsolatállapot alapú nem konvergált állapot



$NH(A,Z) = B$

$NH(G,Z) = H$

$NH(B,Z) = C$

1. CD link meghibásodik

2. C-től LSA

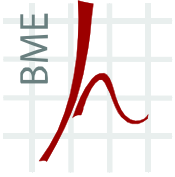
3. C újraszámol: $NH(C,Z) = B$

4. időszakos HUROK,
míg C nem konvergál

...

Konvergált hálózat

- minél tovább tart az LSA flooding, annál nagyobb a valószínűsége hurkok kialakulásának
- a hálózatban egységes időzítési beállítások célszerűek
- linksúlyok megváltoztatása is vezethet hurok kialakulásához
- hiba elhárítása után is a példához hasonlóan alakulhatnak ki hurkok

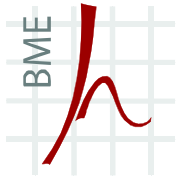


Terhelés szétosztás ECMP alapon

- Azonos súlyú utak egy célhoz
- Csomag alapú: round robin (sorrend, késleltetés, stb. különbözhet)
- Session alapú: *hashing* forrás- és célcím alapján (linkenként eltérő hash-függvény)
- Csökkenti a hiba hatását
- Gyorsítja a konvergenciát

Hierachikus szerkezet

- Areak, Area-0 a tranzit
- Area Border Router
- útvonal preferenciák (intra-area, inter-area, external type 1, external type 2)



EXTERIOR GATEWAY PROTOCOL: BGP



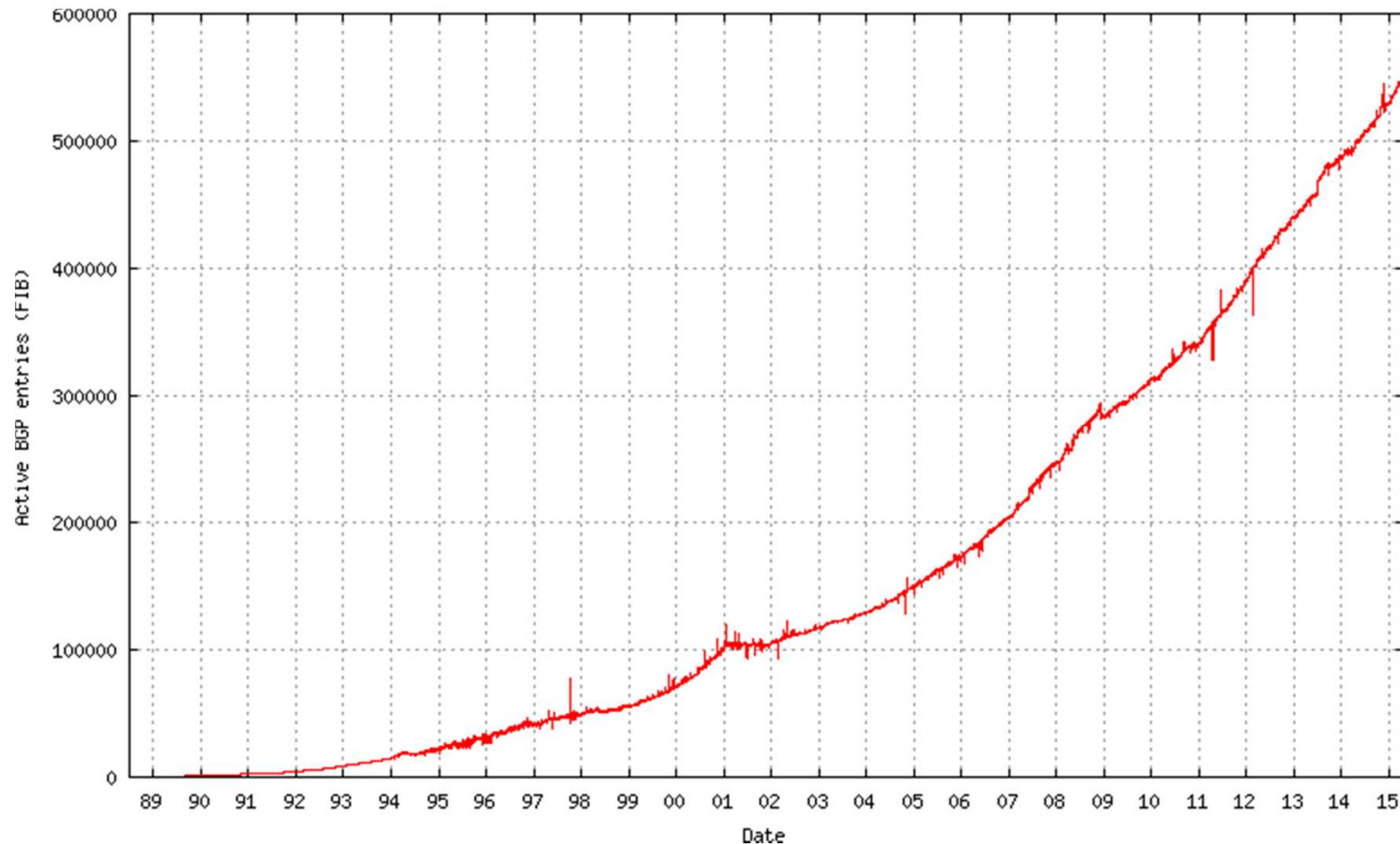
Exterior Gateway Protocols

- **Külső forgalomirányítási protokollok**
 - a "külső" jelző arra utal, hogy nem egy hálózaton belüli, hanem hálózatok közti forgalomirányítás a cél
 - az egyes hálózatok belső szerkezetével alapvetően nem foglalkozik, azok külső kapcsolatainak halmazával dolgozik
 - alapegysége az autonóm rendszer
- **Autonomous System, AS**
 - egy olyan hálózat, ami egységes, egyértelmű belső routing policy-val rendelkezik
 - a külvilág számára egységként kezelhető hálózatrész (RFC4271, page 4)
 - legtöbbször egy vállalat, intézmény, ISP hálózata
 - BMENET: AS2547 , HBONE: AS1955
 - AS number: az autonóm rendszer globálisan egyedi azonosítója

EGP vs. IGP

- **IGP: autonóm rendszeren belüli optimális forgalomirányítás**
 - gyors konvergencia
 - nagy méretű hálózatokra nem skálázódik
- **EGP: autonóm rendszerek közti forgalomirányítás**
 - világméretű hálózatra skálázható
 - jelenleg ~540000 IPv4, ~22000 IPv6 prefix a globális táblában
 - az AS-ek közti konnektivitásra koncentrálnak
 - az AS belsejének "finomságait" nem kezeli (pl. prefix aggregáció)
 - lehetővé teszi routing policy megadását

Aktív BGP entry-k számának növekedése



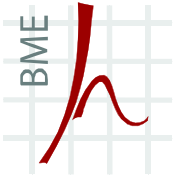
<http://bgp.potaroo.net/as2.0/bgp-active.html>

BGP data obtained from AS65000

Report last updated at Tue Apr 21 06:10:21 2015 (UTC+1000)

BGP

- egyetlen EGP implementáció: BGP
- jelenleg 4-es verzió, RFC1771 -> RFC4271
- AS-AS útvonalakat kezel
 - a prefix felé vezető útvonal alapján állítja be a next hop-ot
 - path-vector protokollként is említik
 - az útvonal-információk változásakor az útvonalat
 - újraszámítja, és az általa használtat hirdeti tovább
 - az útvonalakhoz nem egy metrika tartozik, hanem több attribútum
 - ezek összetett módon befolyásolják a választott útvonalat

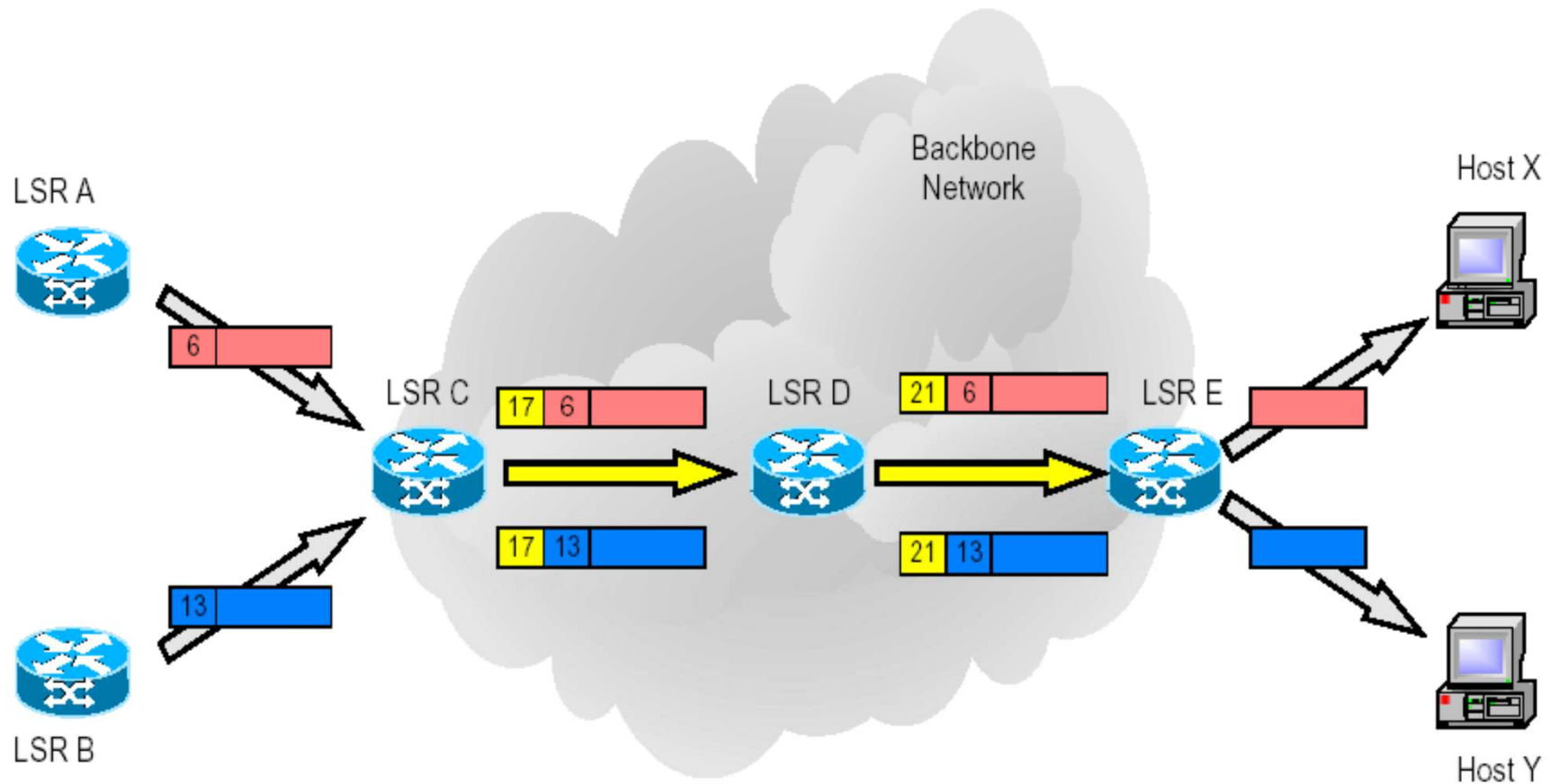


IP/MPLS

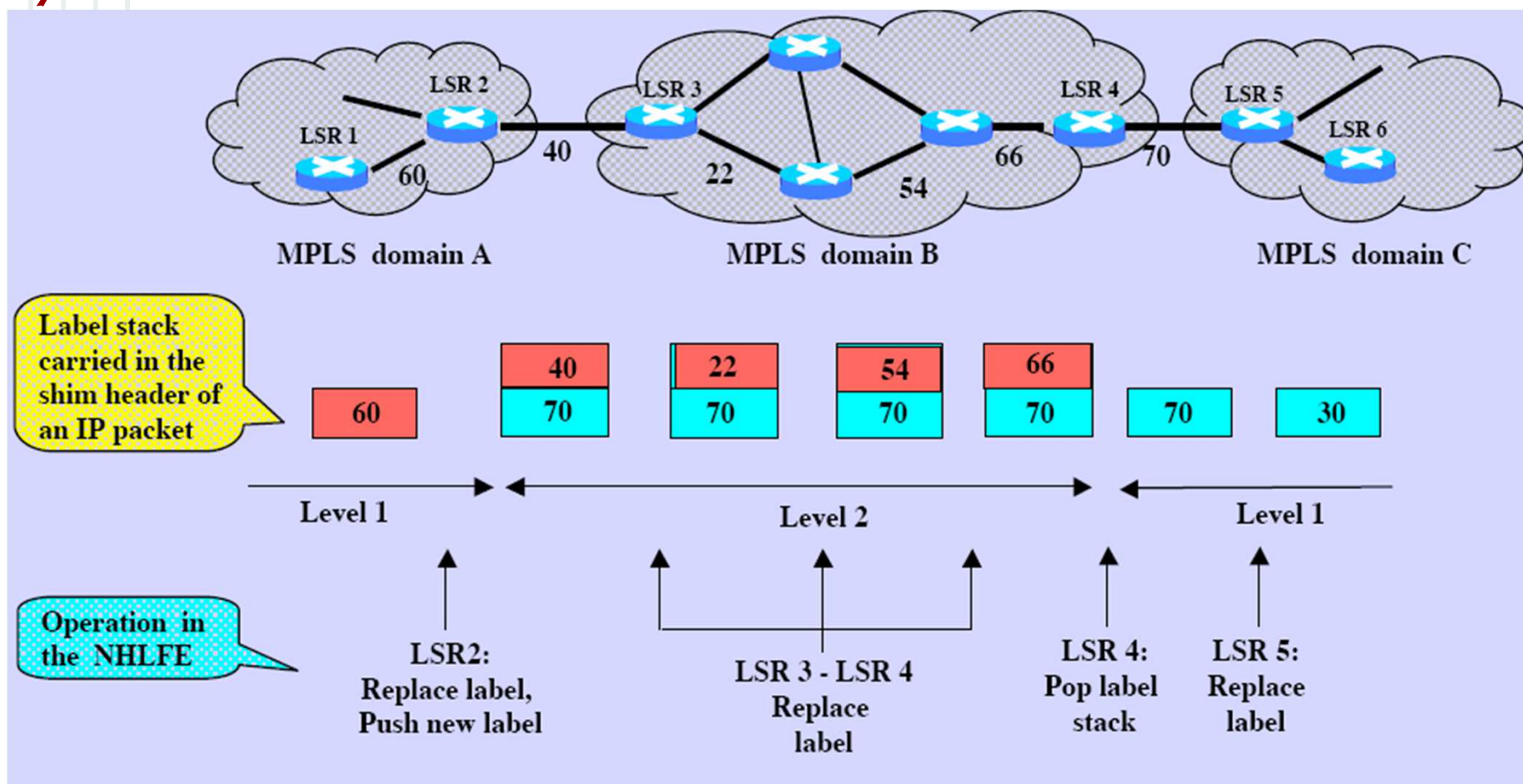


- címke alapú továbbítás Core-ban
 - nagy mennyiségű, nagy aggregáltságú forgalom
 - hierarchikus címkék
- Label Edge Router (Provider Edge Router)
 - encapsulation, decapsulation, service selection
 - label push, pop
- Label Swithc Router (Provider Router)
 - label based forwarding,
 - label swap, push, pop
- Címkekezelés: LDP
 - megállapítás, fenntartás, visszavonás

LSP hierarchy: hierarchikus címkék

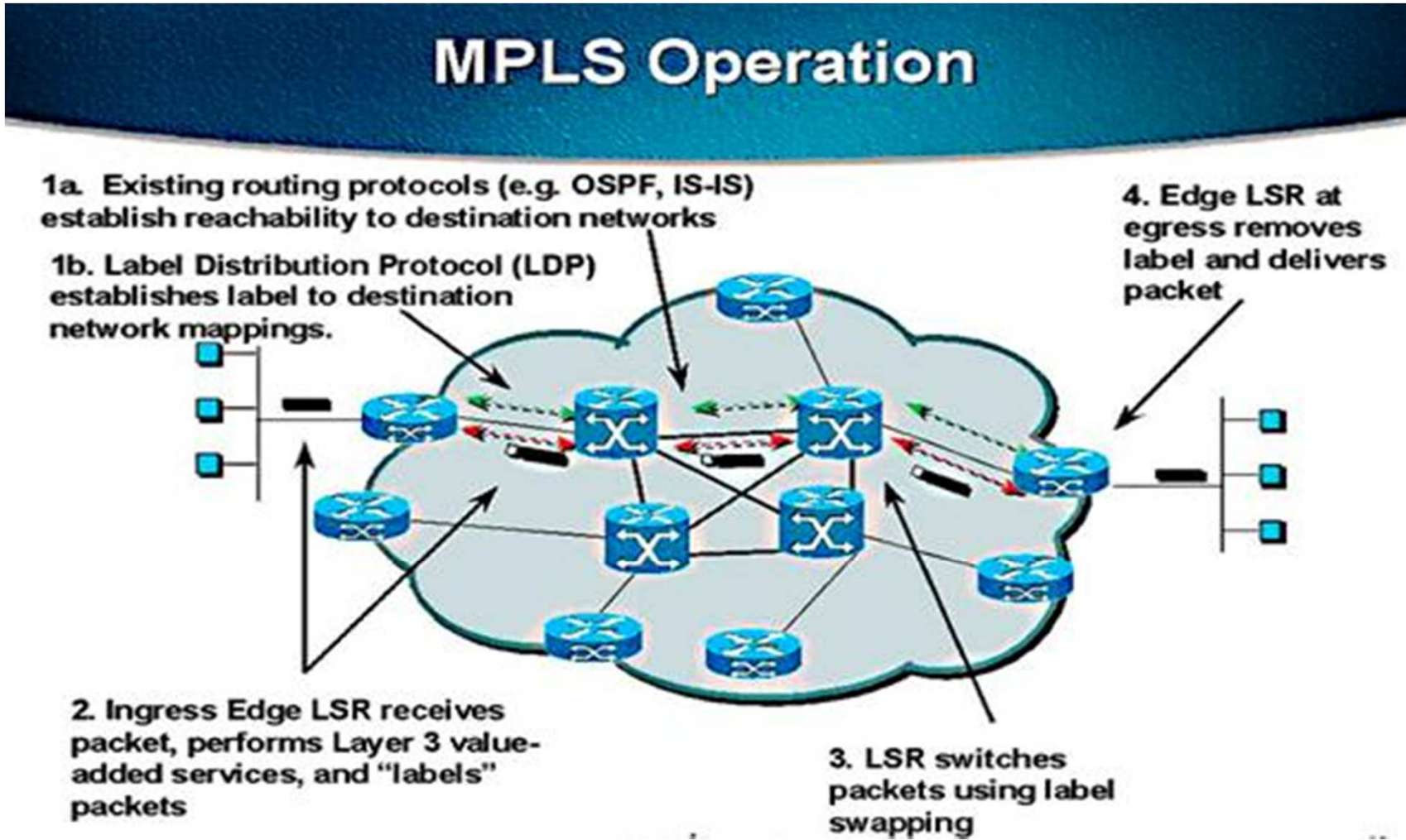


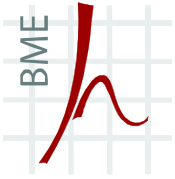
LSP hierarchia: címke stack



- A B MPLS tartományban a címkék egy alagutat alkotnak, az LSR-ek, az egress-t is beleértve, csak a tartományon történő áthaladás szempontjából ismerik a csomagok célját (NHLFE Next Hop Label Forwarding Entry)

IP/MPLS működés





MPLS TE



MPLS TE

- **Miért?**
 - gyors, szelektív hálózatvédelem
 - hálózati erőforrások hatékony kihasználása kontrollált szolgáltatásminőség mellett
- **Hogyan?**
 - kapcsolatorientált explicit utak: TE tunnel
 - off-line: feltételezett (forgalmi prognózis) követelmények alapján megtervezett, konfigurált (szemi-permanens állapotok, újratervezés valamilyen rendszerességgel)
 - on-line: folyamatosan kiértékelt állapotinformációk alapján (dinamikusan változó állapotok, újratervezés kritériumok teljesülésekor)

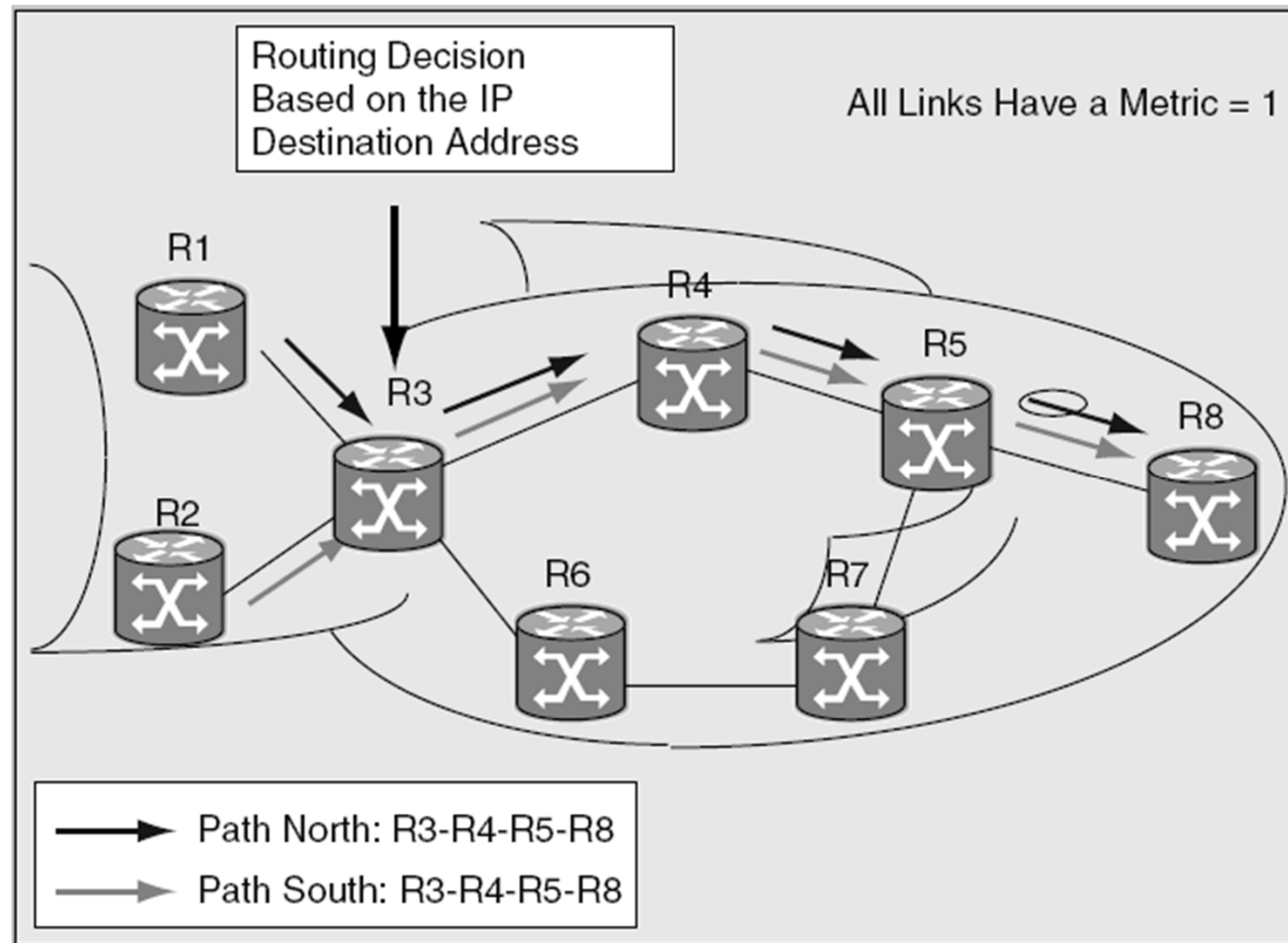
Milyen védelem van MPLS TE nélkül?

- **Védelmek**
 - IP adaptáció (pl. OSPF)
 - IP load balancing (ECMP)
 - kapcsolt L2
 - adaptáció (SPT, RSTP, ...)
 - portduplikálás
 - link aggregation
 - EAPS (Ethernet automatikus védelmi átkapcsolás – VLAN szintű átkapcsolás)
 - L1 (...)

Mi van MPLS TE nélkül?

- **Erőforrás kihasználás**
 - az alapprobléma (pl. OSPF)
 - mininmálutak hálózati szintű aggregáló jellege
 - a „hal”
 - lehetséges megoldások
 - „reverse engineering” – OSPF-súlyok a kívánt utakhoz <-> komplexitás, skálázhatóság, bővíthetőség, menedzselhetségé
 - kapcsolt L2 VLAN-ok <-> menedzselhetségé, QoS (direkt Ethernet szolgáltatások mellett)

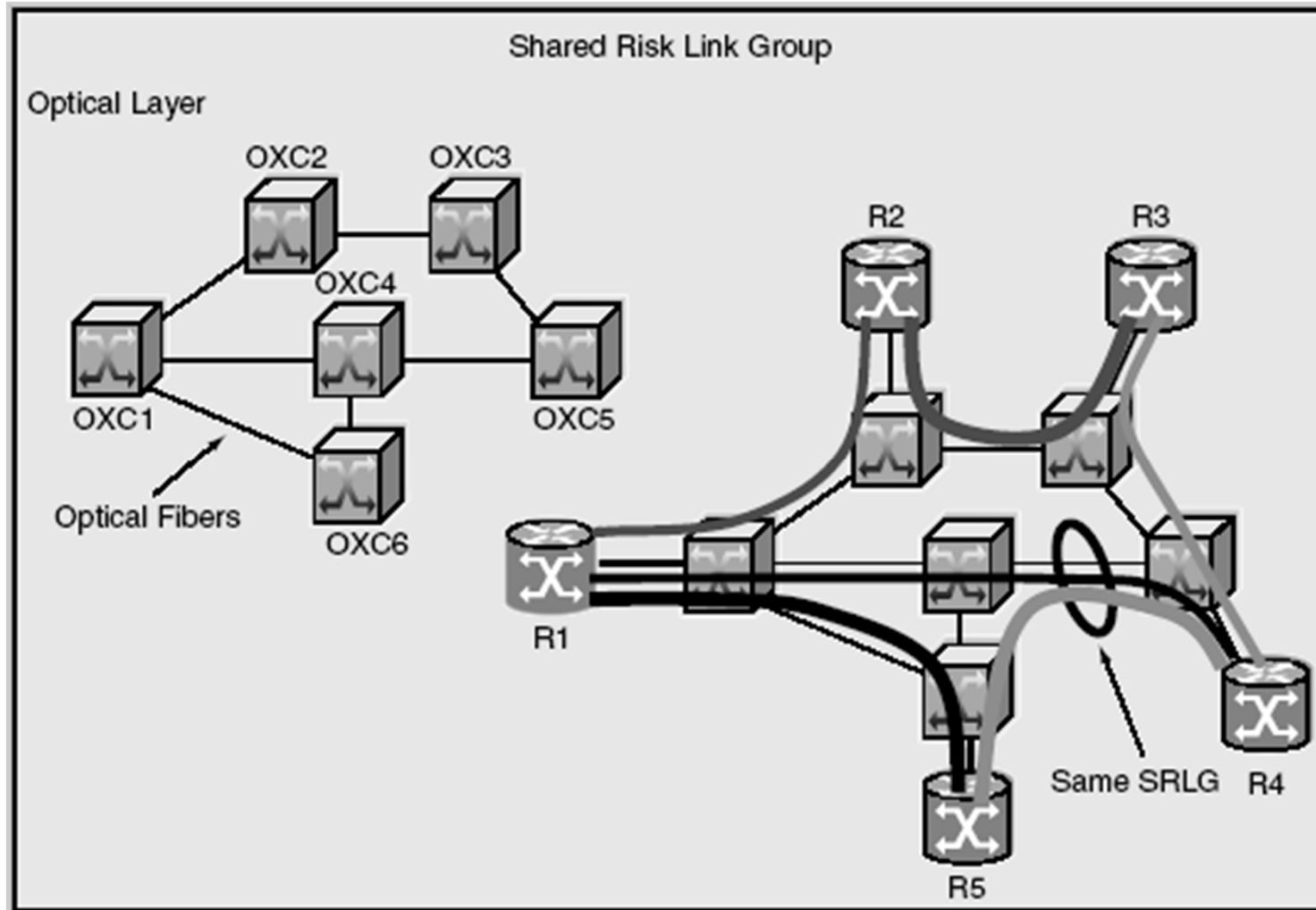
A klasszikus „hal” probléma



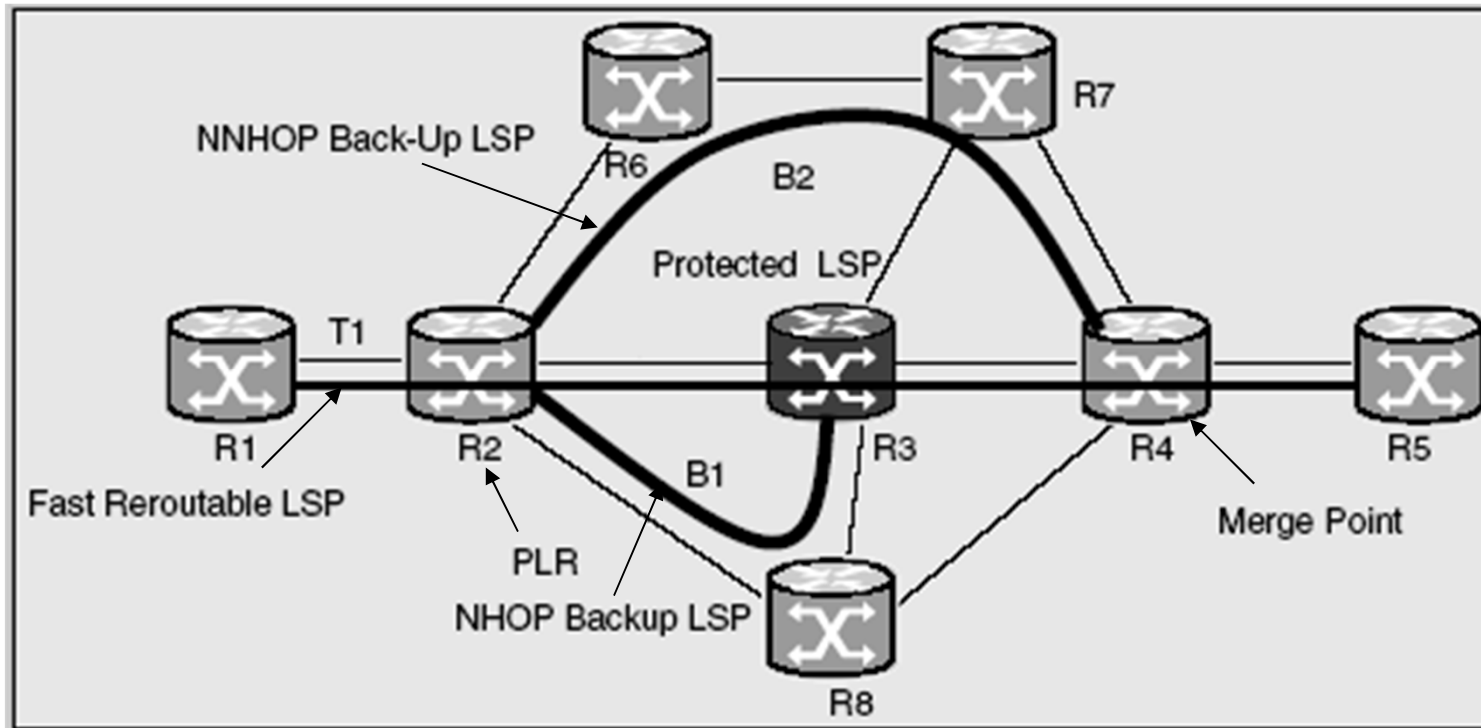
MPLS TE - védelmek

- **SRLG**
 - Shared Risk Link Group: azon linkek csoportja amelyek ugyanazon fizikai hiba hatására esnek ki
 - lazább értelmezés: L2 és L1 hibák
 - szigorúbb értelmezés: berendezés szintű L3 (pl. portkártya) is
- **Függetlenség (SRLG, csp., link)**
- **Védelmi mechanizmusok**
 - Fast Reroute/Path Protection
 - előre tervezett és konfigurált/on-line

SRLG



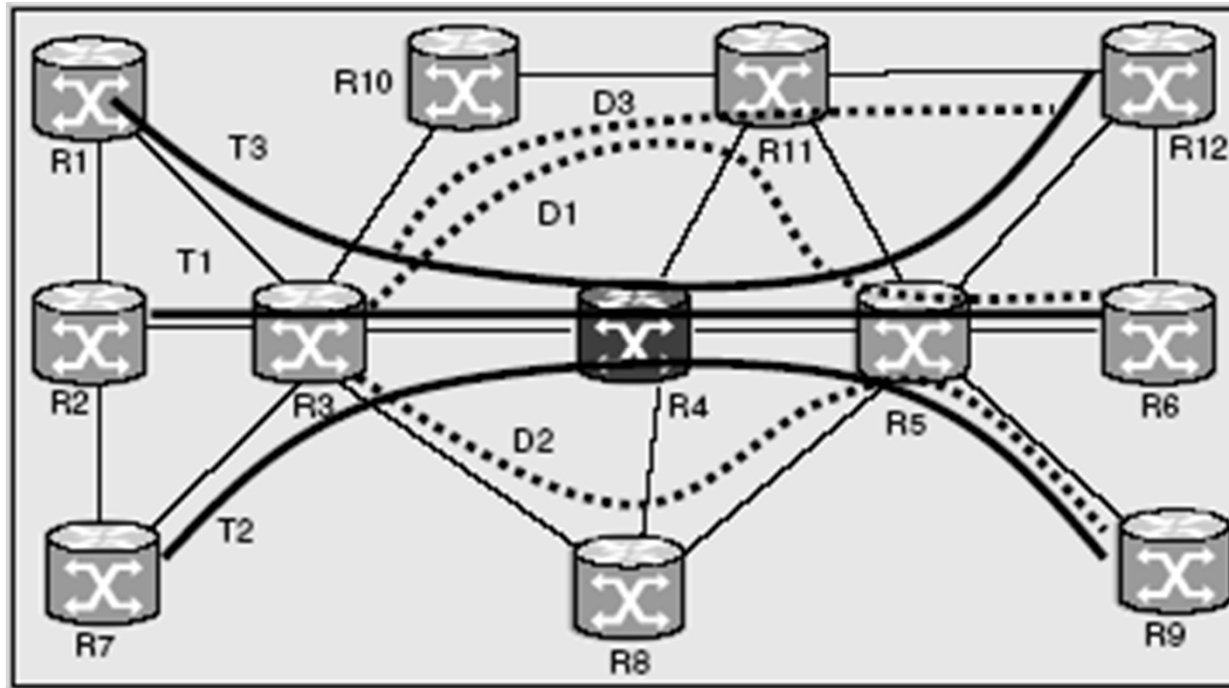
Alapfogalmak



- **PLR – Point of Local Recovery**
- **NHOP Recovery LSP – Next Hop Recovery LSP (végződés PLR-hez képest)**
- **NNHOP Recovery LSP – Non Next Hop Recovery LSP (végződés PLR-hez képest)**

Lokális védelem LSP-nként

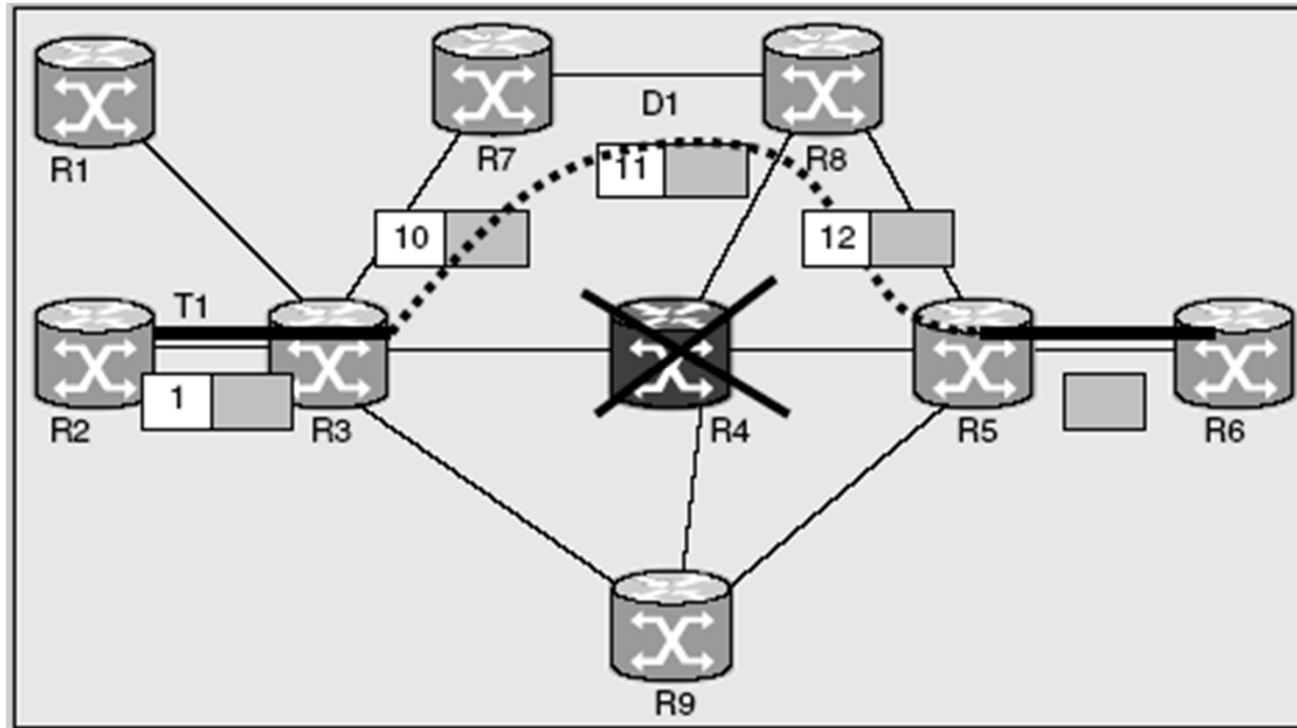
(Local Protection – One to One Backup)



- LRP: R3, Merge Point : R5
- minden védett LSP-nek saját védelmi LSP
- LSP merging a menedzselendő LSP-k számának csökkentésére (pl. D1 és D3 az R3-R10-R11 szakaszon)

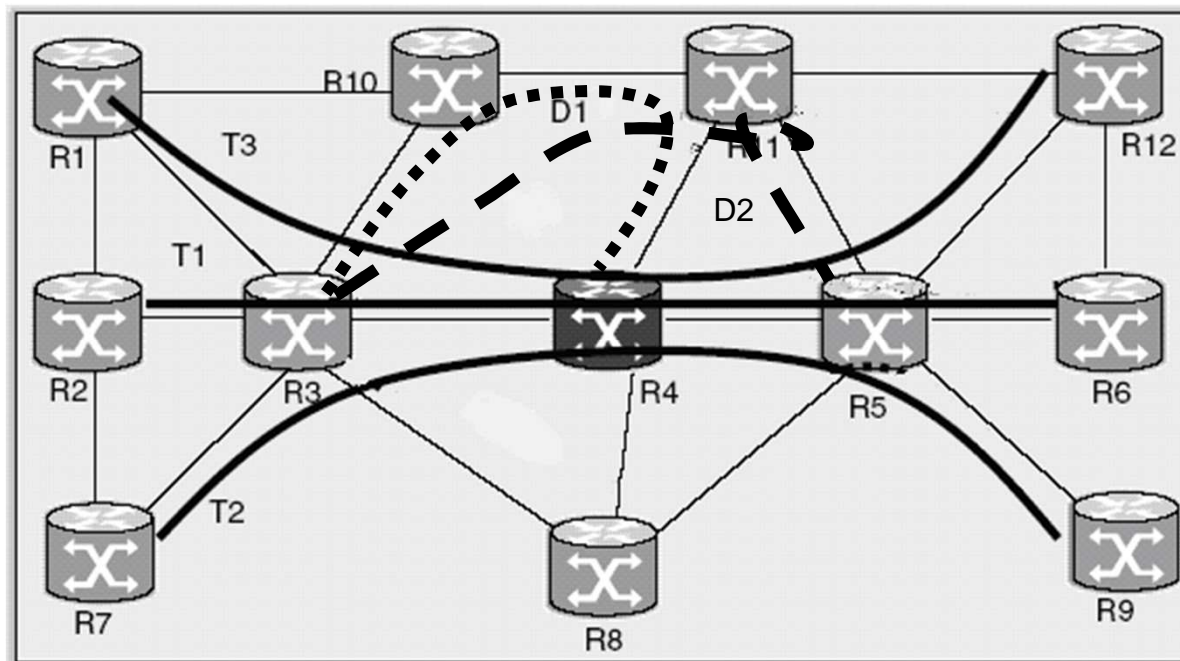
Local Protection – One to One Backup

Működés



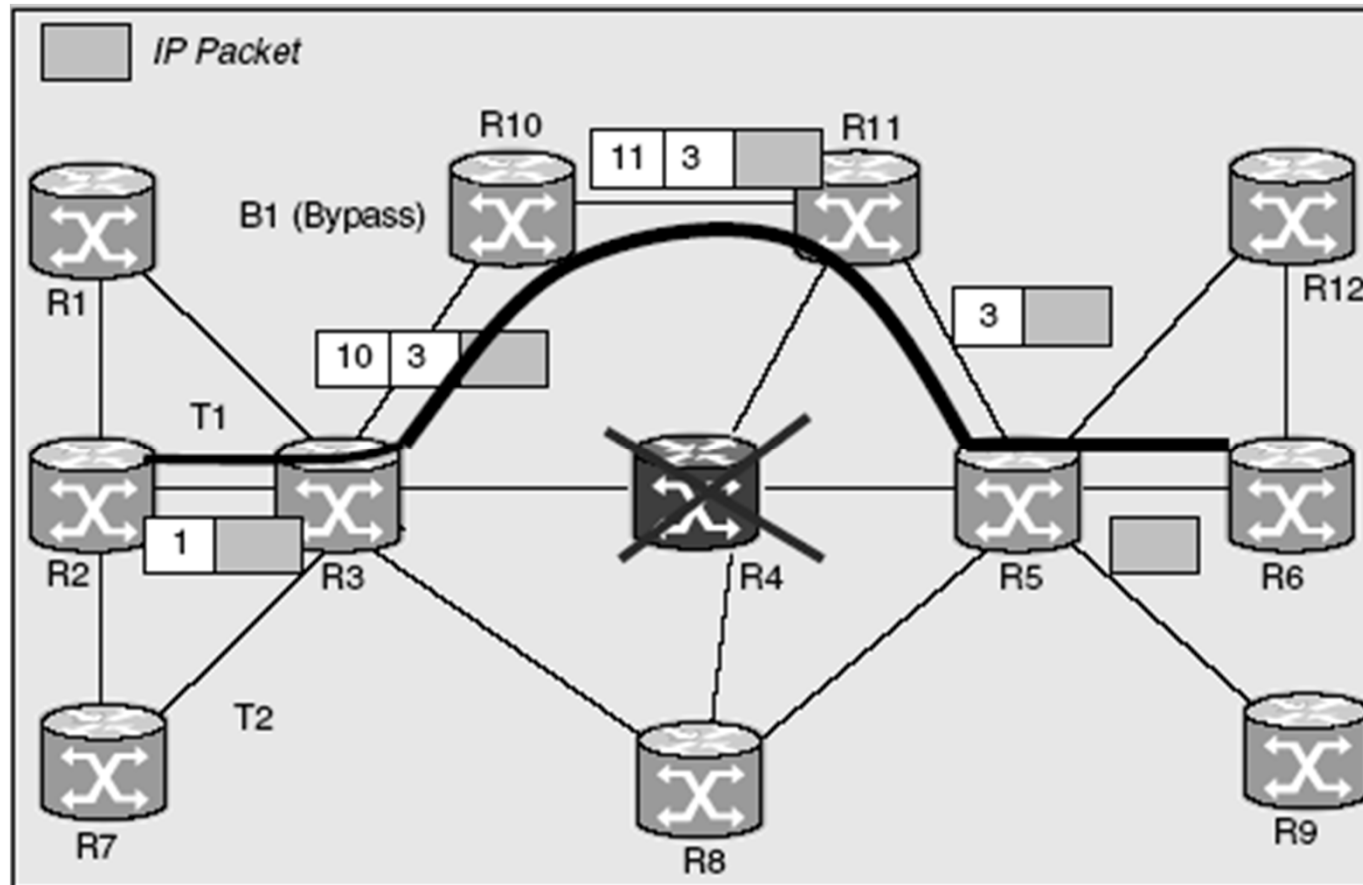
Lokális védelem erőforrásonként

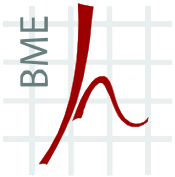
(Local Protection – Facility Backup)



- LRP: R3,
- egy NHOP BAckup LSP szakaszhiba ellen (D1), Merge Point : R4
- egy NNHOP BAckup LSP LSR hiba ellen (D2), Merge Point : R5
- védett LSP-nként, vagy összevontan

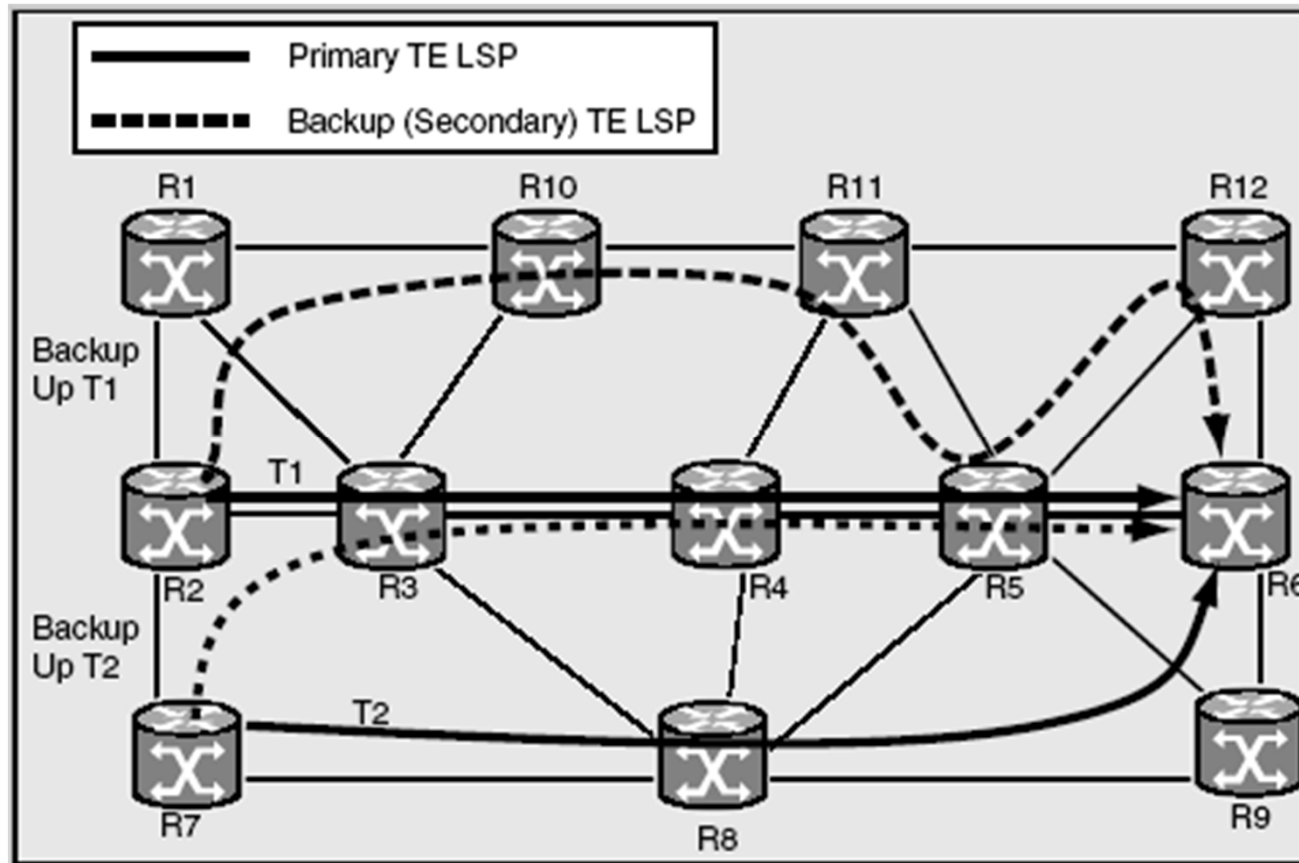
Local Protection – Facility Backup Működés

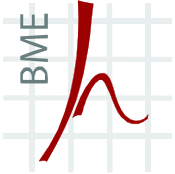




Global Protection

Path Protection

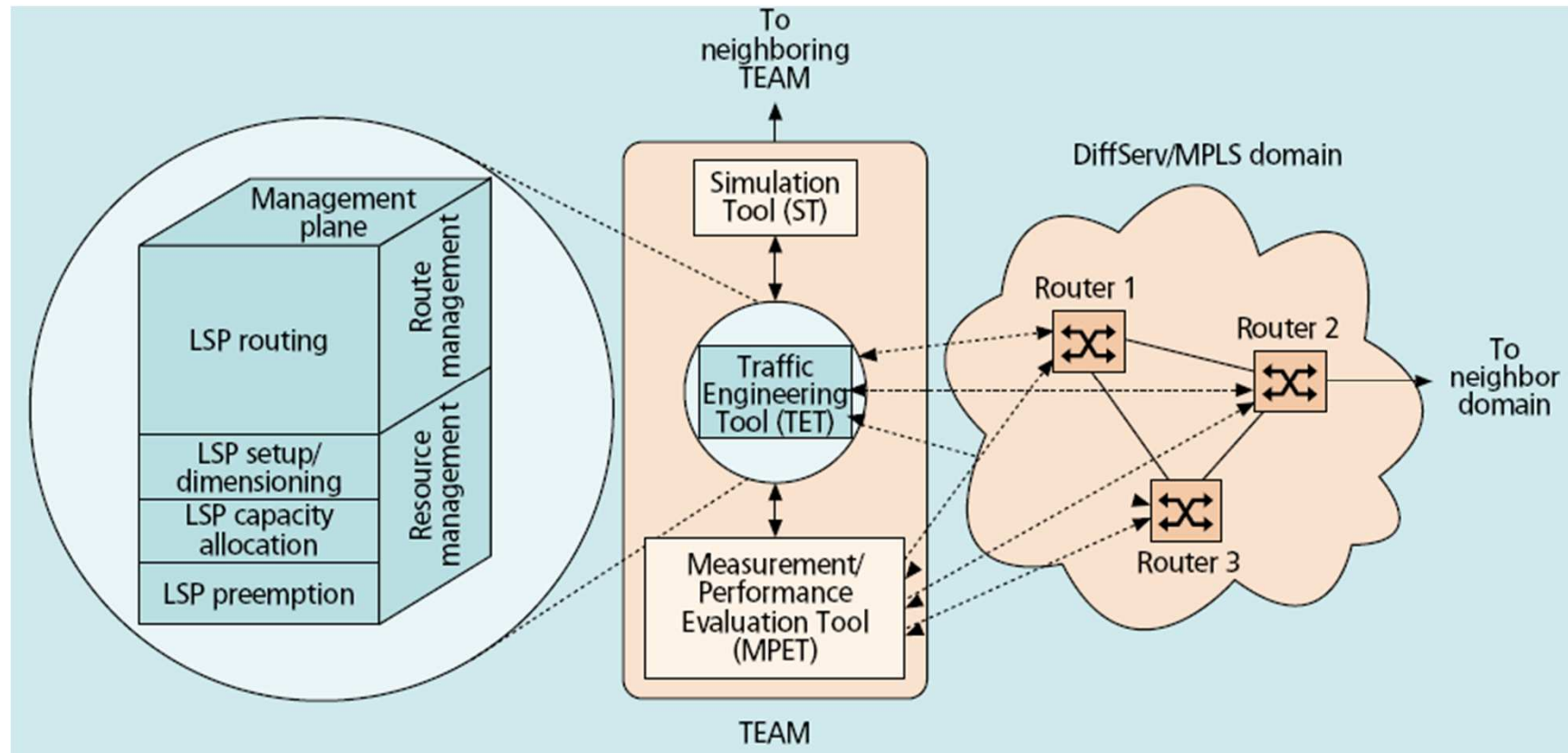




DiffServ Aware IP/MPLS TE

- Traffic Engineering Automated Manager IP MPLS/Diffserv hálózatokhoz
- IEEE Comm. Mag. 2004. október 134-145. oldal alapján

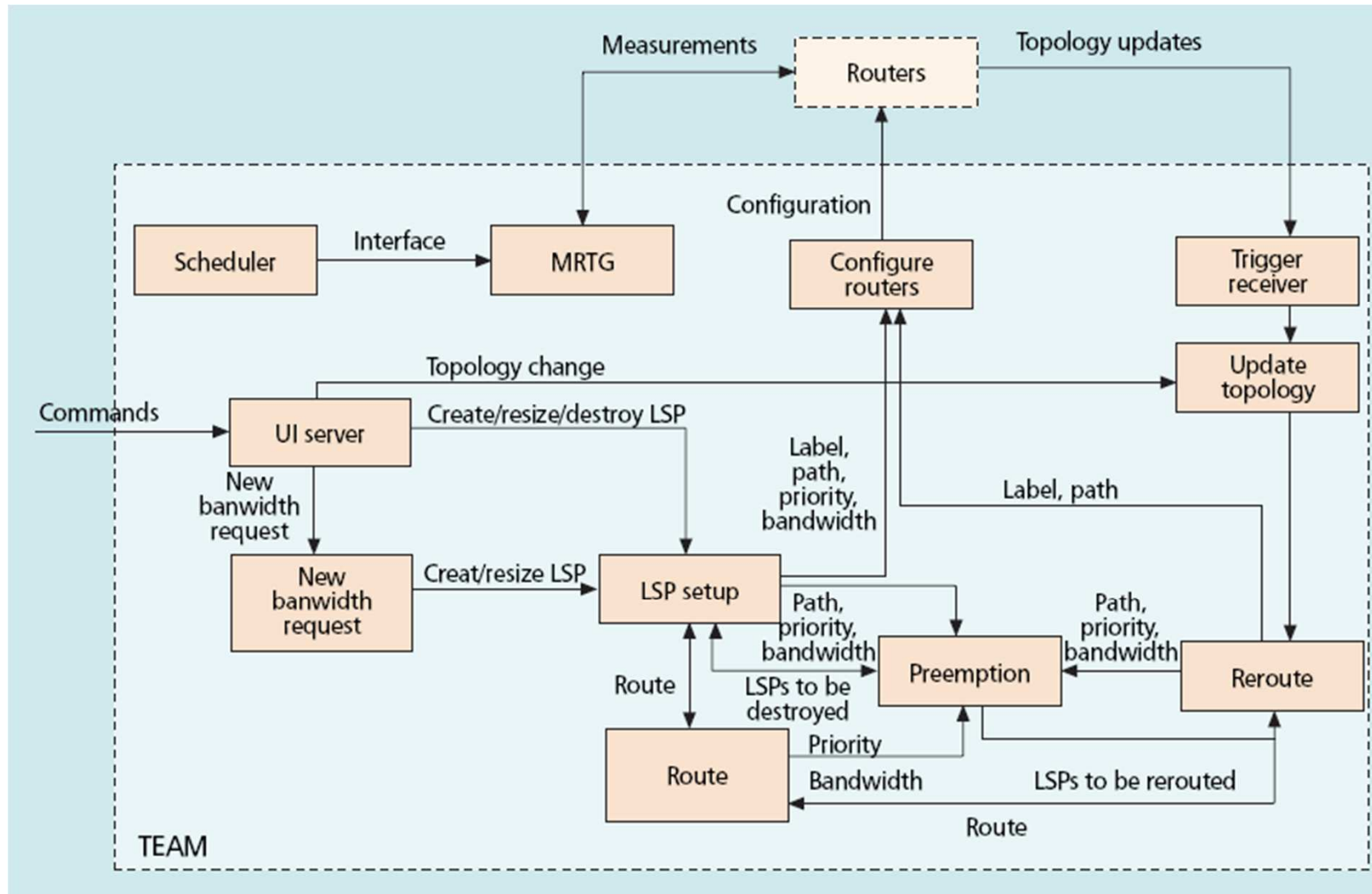
Keretrendszer és funkciók



Keretrendszer és funkciók

- **Traffic Engineering Tool (TET)**
 - LSP létrehozása, méretezése és erőforrás lefoglalás
 - LSP preemption (kikényszerítés) – annak eldöntése, hogy versenyhelyzetben melyik LSP melyik másik erőforrásait veheti el
 - ha ilyen helyzet áll elő, akkor Preemption Policy alapon alacsonyabb prioritású LSP keresése (ezt utóbb a rendszer megpróbálhatja új nyomvonalon elvezetni)
 - tipikus célok
 - a legalacsonyabb prioritású LSP-k megkeresése
 - a lehető legkevesebb LSP kikényszerítésével kiszolgálni a magasabb prioritásút
 - a lehető legkisebb sávszélesség kikényszerítésével kiszolgálni a magasabb prioritásút (single LSP)
 - LSP útvonalválasztás – útvonal a fizikai hálózaton, erőforrások (sávszélesség) az MPLS hálózaton

Részletes funkcionális felépítés



IP/MPLS TE

- menedzselt, szelektív védelmi funkciók
- az MPLS menedzsment-funkciókra alapozottan (mérés, vezérlés) on-line folyamatok a jobb hálózatkihasználás érdekében (a QoS garanciák fenntartása mellett)
- komplex SW, on-line mérések, kiértékelés, tervezési és konfigurációs akciók