A távközlési infrastruktúra fejlődése a digitális átállás és a felhőalapú technológiák korában

KÁKONYI ISTVÁN
Cisco Systems Magyarország Kft.
ikakonyi@cisco.com

Kulcsszavak: SDN, „merchant silicon", automatizálás, virtualizáció, SDN-vezérlő, adatsík, IPoDWDM, konténeralapú virtualizáció

A távközlési szolgáltatóknak meg kell küzdeniük az egyre növekvő sávszélesség-felhasználással (amit a Wi-Fi, 5G és más szélessávú technológiák indukálnak) és a felhőalapú alkalmazások elterjedésével. Nemcsak a sávszélesség, hanem a forgalmi minták is változnak. Másrészt a hálózatfejlesztésnek a szűkülő CAPEX- és OPEX-feltételek ellenére is meg kell történnie. Egyszerűsítésre és automatizálásra van szükség a hálózatban.
A cikk bemutatja a szolgáltatói hálózatok fejlődésének legfontosabb szempontjait.

1. Bevezetés

A távközlési szolgáltatók újabban komoly kihívásokkal néznek szembe. A felhasználók egyre nagyobb sávszélességű technológiákat szeretnének használni (ez igaz a vezetékes és a mobil hálózatokra is), emellett terjednek a felhőalapú szolgáltatások. Ezek a szolgáltatások átrajzolják az eddig megszokott forgalmi irányokat, mert a hagyományos „észak-dél" (letöltés) viszonylatok helyett a „kelet-nyugat" (adatközpontok közötti és az adatközpontokon belüli) irányok is nagyon fontosak lesznek. A fenti követelményeknek megfelelő hálózatot pedig változatlan vagy egyenesen csökkenő költségekkel kell megvalósítani.

Cikkünkben áttekintjük az ilyen hálózatok építőelemeit és architektúráját.

Alapvetően három kérdést fogunk vizsgálni:

- A hardverelemek fejlődését, illetve azt, hogy milyen lehetőség van egyszerűbb, olcsóbb hálózati elemeket létrehozni.
- A szoftvertechnológia, illetve a virtualizáció és az automatizálás fejlődését. Sok előrelépés történt az eszközök menedzsment-interfészeinek hatékonyabbá tétele terén (orchestration, Netconf-protocol, Yang-adatmodellek). Sok funkciót meg lehet teljesen szoftveres alapon valósítani és virtualizált környezetben használni.
- Az architektúra fejlődését, illetve egyszerűsödését, és ezt egy létező, feltörekvő technológián keresztül (Segment Routing) mutatjuk be.

2. A router/switch hardver fejlődése

A nagy teljesítményű routerek, amelyek szolgáltatói környezetben is alkalmazhatóak, hosszú ideig azonos séma szerint készültek: a gyártó néhány év alatt kifejlesztett egy ASIC-generációt, amely képes volt a több millió IP-prefix, a lapkánként több Tbit/s sávszélesség kezelésére, és e köré az ASIC köré épült fel a hardver. Az operációs rendszer általában valamilyen Linux disztribúcióból származott, és a HAL (hardware abstraction layer) segítségével több platformra is adaptálható volt. Egy ilyen folyamat költséges és hosszú, az eredménye azonban olyan (drága...) eszköz lehet, ami bizonyos tulajdonságaival felülmúlja a konkurens termékeket.

A piaci igényt felismerve bizonyos gyártók elkezdtek általános használatra alkalmas, de nagy skálázhatóságú chipeket fejleszteni. Ezek először az adatközponti eszközökben terjedtek el, ma már szolgáltatói routerekben is megtaláljuk őket. Ilyeneket fejleszt és gyárt pl. a Broadcomm, a Marvell vagy az Innovium. Az eredmény egy polcról levehető architektúra („merchant silicon"). A gyártó általában egy API-t is ad a chipekhez, így az alkalmazók könnyen adaptálhatják a már meglevő szoftvereiket. Az eredmény: rövidebb fejlesztési idő, olcsóbb eszközök, alacsonyabb energiafelhasználás. A dolognak természetesen árnyoldala is van: a „merchant silicon" routerek hasonló paraméterekkel fognak rendelkezni, a gyártók csak a szoftver segítségével tudják megkülönböztetni a termékeiket.

A fent leírt építőelemek és „merchant silicon" routerarchitektúrák fejlődése töretlen, ma már minden távközlési szolgáltatói igényt ki tudnak elégíteni (aggregáció, gerinchálózat, ISP peering).

Az *1. ábra* összefoglalja az egyik népszerű „merchant silicon" ASIC különböző változatainak a paramétereit.

Meg kell még jegyezni, hogy a 100 és 400 GE interfészek elterjedésével az optikai interfészek, modulok ára egyre magasabb hányadot képvisel a router teljes árában. A koherens WDM-technológiák nagyon sokat fejlődtek, ma már 400 Gbit/s sebességű optikai modulok is elérhetők, koherens transceiverrel.

Vannak olyan gyártók, amelyek kizárólag a „merchant silicon"-csipekre alapozzák a termékeiket. Vannak olyanok is, amelyek fejlesztik a saját ASIC generációikat, és emellett sokkal olcsóbban kínálnak „merchant

silicon”-alapú termékeket is. A folyamat kiélezte a gyártók közötti versenyt, a vásárlók számára pedig előnyösebb pozíciókat eredményezett.

3. Szoftver, virtualizáció, automatizálás

A routerek szoftverarchitektúrája szintén drámaian fejlődött az elmúlt években. A szabványosításra való törekvés mellett ehhez jelentősen hozzájárult az Open Source technológiák fejlődése és adaptálása.

Az eszközök menedzselhetőségében igazi áttörés következett be: az elavult, nehézkes, vállalati használatra kifejlesztett SNMP-protokoll helyett jött a Netconf, amely tranzakcióalapú, és képes absztrakt adatmodellekkel dolgozni (YANG). Az eredmény az, hogy az egyes eszközök, sőt komplett szolgáltatások is leírhatók YANG-modellekkel. Ez lehetővé teszi különböző gyártók termékeinek ugyanabban a hálózatban való alkalmazását. Ha a megfelelő YANG-modellek rendelkezésre állnak, akkor a szerviz-logikát tartalmazó „orchestrator” az összes eszköz számára el tudja küldeni a szükséges Netconf-parancsokat és ellenőrizheti is azok sikeres végrehajtását. A Netconf-protokoll mellett más, egyszerűbb API-k is elterjedtek a routerek világában, pl. RESTCONF vagy a JSON-alapúak is.

A hálózat működési paramétereinek vizsgálata és azok elemzése is új szintre került. A hagyományos SNMP-alapú lekérdezések helyett ma már elterjedt a „streaming telemetry”, amely gyakorlatilag azt jelenti, hogy a felügyeleti rendszer meghatározhatja, hogy milyen paraméterekre kíváncsi és ezt az eszközök folyamatosan küldik. Ezt egyes esetekben a hardware is támogatja. Így akár IP-flow-szintű adatok is kinyerhetők. A legelterjedtebb protokollok a már említett Netconf és a gRPC (Google RPC).

A szolgáltatásmenedzsment csúcsán egy olyan szoftver van, amelyet orkesztrátornak hívunk. Ez a szoftver képes megérteni a szolgáltatási vagy alkalmazási logikát, és képes ezeket az egyes eszközöket – mindegy, hogy hardver vagy virtualizált elemekről beszélünk – úgy konfigurálni, hogy a kívánt szolgáltatás létrejöjjön.

Sok olyan hálózati funkció van, amit szoftveralapon is meg lehet valósítani. Az ezt leíró szabványokat, architektúrákat egységesen Network Function Virtualizationnak (NfV) hívja a szakirodalom.

Fontos kérdés, hogy mit érdemes virtualizálni? Természetesen minden kontrollsík-funkciót (routing protokollok, traffic engineering stb.) lehetséges, általában olyan feladatokat érdemes, ahol sok és komplex számítási igény van. Az SDN-technológia hőskorában (úgy 10 éve) úgy gondolták, hogy a teljes kontrollsík virtualizálva lesz és a routert és a kontrollert össze kell kapcsolni egy protokollal, ami csak a csomagtovábbításra vonatkozó információt továbbítja. Ez volt az OpenFlow. Korlátozottan terjedt el, mert kiderült, hogy jobb, ha megmarad a router kontrollsíkja, de azt szabványos interfészekkel ki kell nyitni, és így előtte elképzelhetetlen szolgáltatásokat lehet bevezetni (erről a későbbiekben lesz még szó). A szolgáltatói router, különösen a gerinchálózati eszközök az infrastruktúra kritikus elemei. Célszerű, ha meghagyjuk őket valamennyire autonóm üzemmódban. Ezt a koncepciót egyesek „hybrid SDN”-nek hívják.

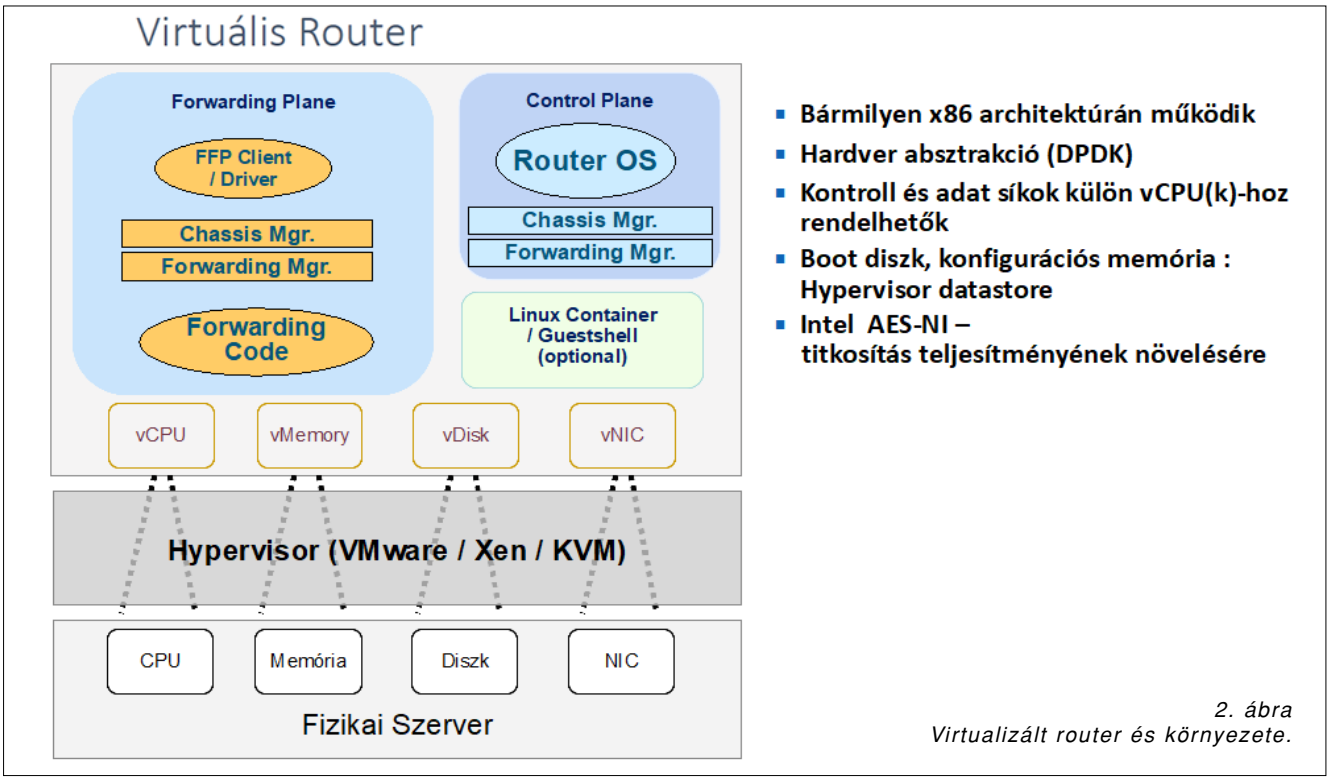
Az adatsík nehezebben virtualizálható, bár ma már ez is lehetséges. Az Intel DPDK (Data Plane Development Kit) megjelenése áttörést jelentett: az általános CPU-k alkalmassá váltak IP-csomagok hatékony továbbítására is. A fejlődés valóban robbanásszerű: ma már minden komolyabb gyártónak van szoftveralapú routere, sőt ma már a második generációról beszélhetünk. Az első generációs eszközök általában egy meglevő routercsaládból indultak ki, és az abban levő hardvert emulálták egy szoftverréteg segítségével, ami egy szabványos CPU-n futott. A második generáció már a szabványos szerverarchitektúrák (Intel, AMD) sajátosságait figyelembe véve sokkal nagyobb teljesítményre képes. Ma nem ritka a CPU-socketenkénti 10 Gbit/s teljesítmény sem.

A *2. ábrán* egy szoftveralapú, teljesen virtualizált router architektúráját láthatjuk.

1. ábra

Egy népszerű „merchant silicon” ASIC különböző típusainak jellemző paraméterei (Broadcom BCM88 sorozat).

	Jericho	Jericho +	Jericho2	Jericho2C	Jericho2C+
Sávszélesség Gbps	720	900	4,800	2,400	7,200
Teljesítmény PPS	720M	835M	2B	1B	2.83B
Interfészek	24x 25G+36x 12.5G	48x25G+24x12.5G	96x 50G	32x50G+96x25G	144x 50G
Fabric Interfész	36x 25G	48x 50G	112x 50G	48x 50G	192x 50G
Chip technológia	28nm	28nm	16nm	16nm	7nm
OCB	16MB	16MB	32MB	32MB	32MB
Buffer	4GB (GDDR)	4GB (GDDR)	8GB (HBM)	4GB (HBM)	8GB (HBM)
Virtual output queue	96K	96K	64K per core	128K per core	256K per core
Számlálók	256K	256K	192K	384K	384K
Fogyasztás	120W	150W	300W-350W	150-200W	450W
PTP támogatás (on Chip)	Class B	Class B	Class B	Class C	Class C
Macsec	No	No	No	No	Yes



2. ábra
Virtualizált router és környezete.

Az NfV általában valamilyen jól bevált virtualizációs környezetet használ (VMware, Openstack és hasonlók). A virtualizációs rendszer és az erőforrásmenedzsment gondoskodik arról, hogy mindig megfelelő számú virtuális router, load balancer, firewall stb. álljon rendelkezésre az adott terhelésnek megfelelően. Az NfV előnye a hatalmas rugalmasság: ha bővíteni kell a rendszert, ugyanolyan szabványos szervereket kell venni és üzembe helyezni. A virtualizáció legújabb iránya a konténeralapú virtualizáció (az egyik legelterjedtebb konténertechnológia a Docker). A konténeres technológia előnye, hogy a Hypervisor által okozott többlet-CPU- és memóriaigény nagyrészt kiküszöbölhető. A konténeres rendszerek erőforrás-menedzsmentje, policy-menedzsmentje és hálózati integrációja ma már megoldott.

Az NfV egyik sikertörténete a mobil internet egyik alapvető építőeleme, a Mobile Packet Core. Erre alkalmazhatók a fentebb leírtak: CPU-intenzív, sok előfizető forgalmát kell egyszerre feldolgozni, és a transzporthálózat bevezeti a forgalmat az adatközpontba, ahol azt fel lehet dolgozni. A MPC-megoldások eleinte még célhardvereket használtak, aztán a fent leírt fejlesztések már lehetővé tették a tisztán virtualizált („cloud native”) megvalósítást is.

Ma már az 5G-mobilhálózatok úgynevezett „cloud native” architektúrát alkalmaznak. Ez azt jelenti, hogy a rádiós hardvert kivéve minden funkció virtualizálva van. Ez hatalmas előnyt jelent a fejlesztőnek és a vásárlónak/üzemeltetőnek is. Azonos hardverelemekből kell építkezni, igen jól skálázható a megoldás és nagyon magas üzembiztonság érhető el.

A programmatikus interfészek, a streaming-telemetria és az SDN-kontroller alkalmazása lehetővé teszi bizonyos funkciók automatizálását, csökkentve ezzel a hálózat üzemeltetésének költségeit. Nagyon sok alkalmazási példa van erre, ezek közül egyet emelnénk ki. *Az IP és az optikai hálózat kontrollsíkjának integrálása* eddig nem látott funkciók megvalósítását teszi lehetővé. Ehhez a következő technológiák szükségesek:

- Koherens DWDM-átvitel, hangolható transceiverekkel.
- IP- és DWDM-integráció a routereken.
- Programozható hullámhosszkapcsolók a DWDM-hálózatban.
- Programmatikus interfészek az IP- és DWDM-doménben.
- SDN-kontroller.

Ha a fentiek rendelkezésre állnak, az SDN-kontroller képes észlelni a routerből érkező telemetria-jelekből, hogy egy adott DWDM-linken rosszabbodik az átvitel minősége (romló SNR, romló hibaarány). Ha ezek az értékek a beállított tartományokból kiesnek, a rendszer képes arra, hogy új optikai adatutat keressen (hisz az optikai szálak topológiája is adott), és ezt az optikai doménben automatikusan kialakítsa. Ezek után az IP-forgalom más irányban, más topológián keresztül továbbítódik. A mai korszerű rendszerekben az átkapcsolás n x 10 sec nagyságrendű lehet.

A fenti művelet az SDN-technológia alkalmazása nélkül kézi beavatkozást és akár napokat is igényelhet.

4. Korszerű szolgáltatói hálózati architektúra

A távközlési szolgáltatók ma alapvetően MPLS-technológiát használnak a hálózataikban. Az MPLS egy absztrakciót használ, az egyes irányokat (IP-prefix), szolgálta-

tásokat vagy tunneleket egy-egy címkével (label) azonosítja. A hozzárendelést az LDP-, az RSVP-, vagy a BGP-protokoll végezheti. Az eredmény egy stabil, skálázható, de komplikált hálózat. Az idők folyamán igény mutatkozott az egyes forgalomtípusok megkülönböztetésére és egy adott útvonalon történő továbbítására. Ezt a problémát képes megoldani a Traffic Engineering, de ez egy új elem bevezetésével járt a protocol stack-be: ez az RSVP (Resource Reservation Protocol). Az MPLS Traffic Engineering nem tudott széles körben elterjedni, mert az újabb elem a protocol stack-ban és a routereken jelentkező plusz CPU- és memóriaigény nehézkessé tette a használatát (az RSVP-protokoll „stateful”: az összes részt vevő routeren nyilván kell tartani a tunneleket).

A Segment Routing (továbbiakban SR) [1] az MPLS adatsíkjának változatlanul hagyásával (ezáltal natív IPv6-hálózaton is működik, ennélfogva jövőálló!), de a kontrollsík jelentős egyszerűsítésével jött létre. Ma már gyakorlatilag minden vezető gyártó támogatja, ezért kvázi szabványként is tekinthetünk rá. A Segment Routing legfontosabb tulajdonságai a következők:

- MPLS- vagy IPv6-adatsík (ez utóbbi esetén nem label stack, hanem IPv6-címek rendezett listája van az IPv6 routing-header-ben).
- Source routing, a label impozíciót végző router által az IP-csomagra csatolt „label stack” alapján történik a forgalom irányítása.
- Nincs LDP-protokoll, a címkék (Segment ID) allokálása és terítése az IGP-protokoll feladata. IS-IS és OSPF kiterjesztések segítségével történik.
- Minden szolgáltatás, ami MPLS-hálózaton működik, SR-hálózaton is fog (L2VPN, L3VPN) működni, minden változtatás nélkül.
- Nem szükséges újabb protokoll a Traffic Engineering megvalósításához. SDN-kontrolleren keresztül, szabványos API-n (Path Computation Element Protocol) bármilyen „tunnel” beprogramozható.

A *3. ábra* mutatja az MPLS és a SR közötti különbségeket. Ahhoz, hogy az SR-technológia működését pontosan megértsük, talán érdemes röviden áttekinteni a két mechanizmus működését.

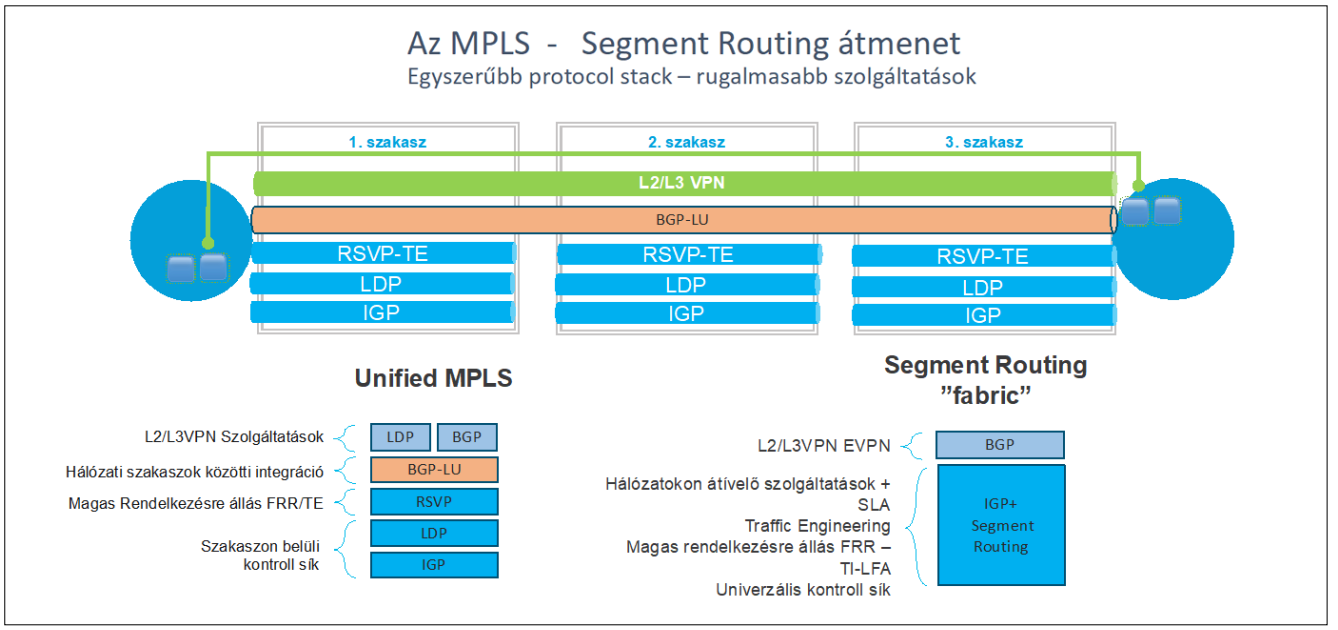
Az MPLS-hálózatokban valamilyen IGP-routingprotokoll (ISIS vagy OSPF) kiszámítja a routingtáblát. Az LDP-protokoll ezekhez címkéket rendel és azokat elküldi a szomszédos routereknek. Ha van Traffic Engineering, az RSVP-protokoll kiszámítja a beállított feltételeknek megfelelő utakat és kialakítja a tunneleket a hálózatban. Ezek tetején ott a BGP, ami a szolgáltatásokért felelős (L2VPN, L3VPN, IPv4, IPv6). Ha nagyon sok hálózati szakaszt kell összekötni, akkor még ott van a BGP-LU (labelled unicast), a skálázhatóság növelésére. Ez a protokollstack látható az ábra bal alsó sarkán.

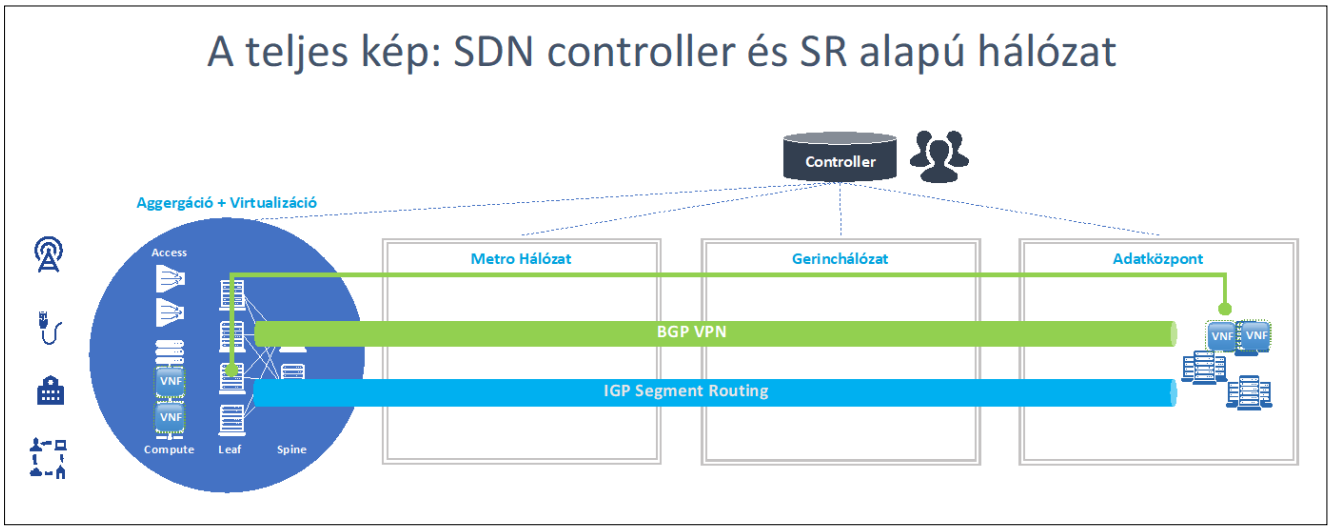
A SR (ábra, jobb alsó sarok) sokkal egyszerűbb! Az IGP-protokoll-kiterjesztések (OSPF és ISIS) magukban hordozzák a SID- (Segment ID) információkat, erről az összes routernek lesz információja a konvergencia befejeződése után. (Többféle SID létezik: prefix SID, adjacency SID, ez lokális, két router közötti link azonosítására szolgál). Ezek után minden MPLS-alapú szolgáltatás működőképes. Ha nagyon gyors konvergencia a követelmény, rendelkezésre áll a TI-LFA (Topology Independent Loop Free Alternate). Ez lényegében azt jelenti, hogy az IGP-protokoll mintegy előre lemodellezi az egyes linkek megszakadását és előre megszerkeszti az ilyenkor használatos label stack-et.

Ha Traffic Engineering szükséges, akkor (legalább) két lehetőség van:

- Az IGP-protokoll különböző paraméterekkel is ki tud számolni utakat (késleltetés, linkek vagy routerek elkerülése stb.). Ezek a label stack-ek előre kiszámíthatók. Ami nagyon fontos, csak az úgynevezett Head End (a kezdő) routernek kell rendelkeznie ezzel az információval, nincs protokoll-interakció a többi routerrel.

3. ábra Az MPLS és Segment Routing technológia összehasonlítása.





4. ábra SDN-alapú, Segment Routing-ot alkalmazó szolgáltatói hálózat.

- A másik megoldás, hogy egy SDN-kontroller az összes hálózati szakaszra kiszámolja a kívánt utat. Ezek után az egyes label stack-eket a megfelelő routerekbe egyszerűen letölti. Ehhez két fontos építőelem szükséges:
  – A kontrollernek el kell küldeni az aktuális hálózati topológiát és a linkek attribútumait. Erre alkalmas – a már az SR-technológia előtt kifejlesztett – BGP LS-(BGP Link State) protokoll.
  – A másik építőelem az SR-szegmensek sorozatának, azaz a label stack-nek a beprogramozása a routerekbe: erre való a PCE- (Path Computation Element) protocol.

Fontos megjegyezni, hogy utóbbi két mechanizmus csak a menedzsment- és kontrollsíkban jelenik meg. A hálózat alapvetően képes üzemelni a kontroller leszakadása esetén is (csökkentett szolgáltatásokkal...).

A fenti leírásnak megfelelő hálózat vázlata látható a *4. ábrán*.

5. Összefoglalás

Cikkünkben megvizsgáltuk a távközlési szolgáltatók hálózati infrastruktúrájának legújabb trendjeit. A növekvő adatátviteli igények és a rugalmasabb szolgáltatásmenedzsment kisebb OPEx- és CAPEx-szintek mellett valósíthatók meg, ha alkalmazzuk ezeket az eszközöket.

- Meg kell vizsgálni az olcsóbb, „merchant silicon” alternatívákat a nagy teljesítményű routerek kiválasztása esetén.
- Meg kell vizsgálni, hogy melyik szolgáltatás virtualizálható, és hogy ezt központi, vagy elosztott architektúrában célszerű-e megtenni?
- Olyan eszközöket célszerű választani, amelyek korszerű API-okkal rendelkeznek és beilleszthetők egy közös orkesztrációs rendszerbe.
- A Segment Routing az MPLS valós alternatívája. A jelenlegi hálózatok átmigrálhatók. A hálózat egyszerűbb lesz és több szolgáltatást képes nyújtani.

Hivatkozások

[1] RFC 8402, Segment Routing Architecture, Clarence Filsfils, Stefano Previdi (eds.), *https://datatracker.ietf.org/doc/html/rfc8402*

A szerzőről

KÁKONYI ISTVÁN 16 éve a Cisco munkatársaként rendszermérnöki, majd architekt pozícióban, túlnyomórészt távközlési szolgáltatókkal foglalkozott. Szakterülete az IP routing/switching, MPLS, távközlési szolgáltatók hálózati architektúrája és az SDN. CCIE és DevNET associate specializációkkal rendelkezik.