

# Realizacija SOX *exchange*-a

Nenad Krajnović, Zoran Perović

**Sadržaj** — Jedan od osnovnih elemenata arhitekture Interneta danas predstavljaju i tačke razmene Internet saobraćaja. Pored razmene Internet saobraćaja, ovakva telekomunikaciona čvorišta su sve češće tačke razmene i ostalih vrsta telekomunikacionih servisa. U okviru rada predstavljeni su osnovni tehnički detalji realizacije *Serbian Open Exchange* – SOX čvorišta za razmenu Internet saobraćaja i ostalih telekomunikacionih servisa.

**Ključne reči** – Internet exchange, SOX, BGP, ethernet.

## I UVOD

Tokom svog razvoja, arhitektura Interneta se značajno menjala. Jedan od ključnih elemenata te arhitekture je i pojava Internet *exchange*-a (IX), tačaka za razmenu Internet saobraćaja između Internet servis provajdera. U početku su na IX-u saobraćaj razmenjivali samo veliki ISP-evi da bi se tokom vremena na ovim tačkama pojavili i *content* provajderi nudeći brzi pristup različitim sadržajima. SOX predstavlja pokušaj da se jedan takav nezavisan IX uspostavi i na teritoriji Srbije. U Srbiji je formirana kritična masa ISP-eva tako da se oseća nedostatak jednog nezavisnog komunikacionog čvorišta (ili mreže čvorišta) posredstvom kojih bi se razmenjivao domaći Internet saobraćaj. Druga bitna osobina modernih IX-eva je da su prerasli svoju osnovnu namenu – razmenu Internet saobraćaja. IX-evi danas predstavljaju komunikaciona čvorišta čijim posredovanjem se razmenjuju razne vrste IP baziranog ili kanalski baziranog saobraćaja. Za provajdere je to bio mali korak da pored Interneta, na postojećoj infrastrukturu počnu da razmenjuju i neki drugi saobraćaj. U okviru ovog rada predstavljene su tehničke osnove SOX-a, *Serbian Open Exchange*, tačke za razmenu Internet saobraćaja i ostalih telekomunikacionih servisa.

## II SOX ARHITEKTURA

U prošlosti su Internet čvorišta podrazumevala postojanje *telehouse* objekata u okviru kojih bi operatori i ISP-evi postavljali svoju opremu u cilju razmene saobraćaja. Takvo rešenje je zahtevalo veoma ozbiljnu infrastrukturu za smeštaj, napajanje i klimatizaciju opreme. U novije vreme su se pojavile realizacije IX-eva bazirane na drugačijim osnovama. Padom cena komunikacionih resursa, prvenstveno optičkih kablova i sistema, umesto centralizovane realizacije IX-a prešlo se na distribuiranu arhitekturu. Na lokacijama IX-a se postavlja samo komunikaciona oprema IX-a kao što je *ethernet switch*. Oprema korisnika se nalazi na njihovim

lokacijama a do lokacije IX-a se uspostavlja optički link. Takvim rešenjem je eliminisan zahtev za postojanjem infrastrukture za smeštanje velike količine komunikacione opreme korisnika a zadržana je funkcija razmene saobraćaja. Isti princip je korišćen i kod realizacije SOX-a.

Na početku svog razvoja SOX se sastoji od dva *ethernet switch*-a Extreme Summit x650 koji se nalaze na Elektrotehničkom fakultetu u Beogradu i u Beogradanci. *Switch*-evi su povezani u optički prsten za čiju realizaciju se koristi 10Gb/s *ethernet* tehnologija. Imajući u vidu da će se postojeći prsten, koji se trenutno sastoji samo od dva čvora, širiti, za realizaciju zaštite implementiran je EAPS (*Ethernet Automatic Protection Switching* [1]) koji obezbeđuje zaštitu saobraćaja sa istim kašnjenjem koje imamo kod SDH mreža. Korišćenjem EAPS-a obezbeđeno je prerutiranje saobraćaja za manje od 50 ms u slučaju da dođe do fizičkog prekida linka između SOX čvorišta. Osnovu SOX arhitekture čini *ethernet* tehnologija na L2 i L3 nivou. Povezivanje korisnika se obavlja primarno korišćenjem gigabit *ethernet* tehnologije a moguće je povezivanje i posredstvom *fast ethernet* i 10G *ethernet* tehnologije. Korisnici koji imaju opremu na lokaciji na kojoj se nalazi i oprema SOX-a mogu se povezati posredstvom bakarnih STP kablova ili optičkim kablom. Ako se korisnikova oprema ne nalazi na nekoj od lokacija gde SOX ima svoje komunikacione uređaje, tada se od nekog od postojećih operatora zakupljuju optička vlakna i uspostavlja se *Fast* ili Gigabit *Ethernet* link do lokacije SOX-a. Stvaran protok koji će biti na raspolaganju korisniku reguliše se softverski na portu Extreme x650 *switch*-a.

Korišćenjem *ethernet* tehnologije korisnicima je na raspolaganju više različitih servisa. Osnovni servis zbog koga je i pokrenuta realizacija SOX-a je Internet *Exchange*. Ovaj server je realizovan korišćenjem L2 *ethernet* tehnologije. Svi korisnici SOX IX-a se nalaze u istom VLAN-u. SOX obezbeđuje dva *Route* servera čiji zadatak je razmena oglašavanja ruta putem BGP-4 protokola. Kod realizacije IX-a na L2 nivou veliki problem može da predstavlja postojanje velikog broja BGP *peering* konekcija što bi smo imali kada bi svaki korisnika IX-a morao da uspostavlja BGP *peering* sa svim ostalim korisnicima IX-a. Pored vremenski zahtevne operacija kontaktiranja velikog broja ISP-eva, javio bi se i problem grešaka u konfiguracijama tolikog broja rutera. Umesto toga, često rešenje kod L2 realizacije IX-a je postojanje *Route* servera. Njihov jedini zadatak je da razmenjuju rute koje oglašava svaki ISP. Na taj način svaki ISP koji dođe na SOX uspostavlja samo dve (zbog redundanse) BGP sesije sa dva *Route* servera SOX-a. Kroz te dve sesije ISP oglašava svoje rute ka SOX-u a od SOX-a dobija sve ostale rute koje postoje u tabeli rutiranja. Pošto se radi o BGP *peering*-u posredstvom *ethernet* mreže, *next-hop* atribut ruta se prilikom prolaska kroz *Route* servere SOX-a ne menja već u izvornoj formi (IP

Mr Nenad Krajnović, Elektrotehnički fakultet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (telefon: +381-11-3221419, e mail: krajko@etf.rs )

Zoran Perović, *Serbian Open Exchange*, Todora Dukina 78, 11000 Beograd, Srbija (telefon: +381-64-6118246, e-mail: iconsult@perke.net)

adrese ISP-eva koji oglašavaju svoje rute) se prosleđuje svim korisnicima SOX-a. Posledica toga je da korisnički saobraćaj ne prolazi kroz *Route* servere već se direktno putem L2 *ethernet* tehnologije razmenjuje između korisnika. Takvim rešenjem se koriste visoke performanse implementiranih *ethernet switch*-eva da bi se dobila isporuka saobraćaja sa kašnjenjem koje je manje od 1ms. S druge strane, obrada BGP oglašavanja, filtriranje ruta i slične zahtevne operacije su prepuštene *Route* serverima čija je procesorska snaga značajno jača u odnosu na *ethernet switch*-eve. Detalji implementacija *Route* servera su objašnjeni u sledećem poglavlju. Pošto se za realizaciju SOX IX-a koristi L2 *ethernet* tehnologija, postoji opasnost prolaska raznih kontrolnih paketa kao što su BPDU paketi između različitih korisnika. Takvi BPDU paketi mogu da dovedu do blokade porta na korisničkoj strani usled aktiviranja mehanizama *Spanning Tree* protokola. Da bi se to sprečilo, vrši se filtriranje saobraćaja ka i od korisnika tako da se propuštaju samo paketi koji imaju MAC adresu korisničkog rutera i *broadcast* MAC adrese.

SOX *Exchange* nije samo *Internet Exchange*, već on predstavlja tačku razmene proizvoljnog telekomunikacionog saobraćaja koji može da se upakuje u *ethernet* paket. Implementirano tehničko rešenje omogućava uspostavljanje L2 VPN konekcija sa garantovanjem potrebnog kvaliteta servisa. Korišćenjem VLAN-ova saobraćaj se može, formiranjem virtuelnih kanala, proslediti između bilo koje dve pristupne tačke u mreži. Time se omogućava razmena i ostalih vrsta saobraćaja koji nisu *Internet saobraćaj* a koriste IP tehnologiju za prenos. L2 VPN servis je inicijalno zamišljen za povezivanje dve tačke ali na isti način omogućava povezivanje i tri i više tačaka u jednu virtualnu L2 mrežu koja omogućava komunikaciju svakog sa svakim. Na ovaj način je moguće vršiti razmenu VoIP saobraćaj, distribuciju video kanala između kablovskih operatora, i slično.

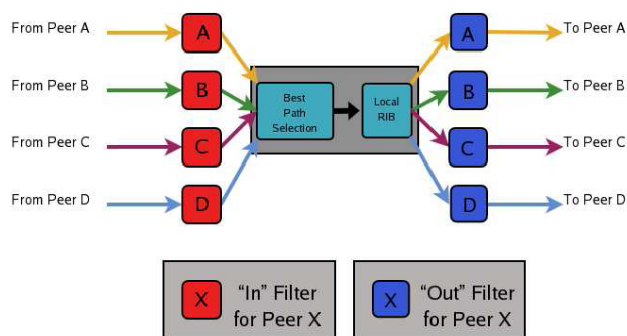
Da bi koristio različite servise SOX mreže, korisnik može da se na SOX poveže samo putem jednog linka. Korišćenjem VLAN-ova za multipleksiranje saobraćaja iz različitih virtuelnih putanja, mogu se koristiti različiti servisi. Pošto implementirani mrežni uređaji omogućavaju kontrolu protoka i QoS-a po VLAN-u, za svaki servis je moguće dogovoriti poseban kvalitet.

Iz prethodno izloženog se vidi da SOX možemo da posmatramo i kao svojevrsnu berzu telekomunikacionih servisa i usluga. Na primer, *content* provajder umesto da uspostavlja poseban komunikacioni link do svakog od potencijalnih korisnika svojih usluga, može da uspostavi vezu sa SOX-om i da posredstvom SOX servisa svoje usluge ponudi svim ostalim korisnicima mreže. Na ovaj način se smanjuje i cena servisa koje provajderi nude jer je i cena realizacije značajno niža.

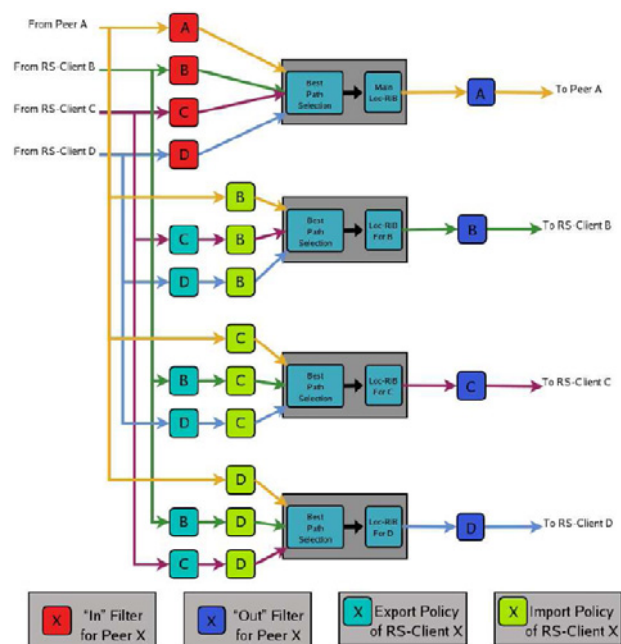
### III ROUTE SERVERI NA BAZI QUAGGA SOFTVERA

Postojanje *Route* servera predstavlja suštinsku razliku između SOX mreže i bilo koje druge *ethernet* mreže koja se koristi za razmenu *Internet saobraćaja*. U slučaju velikog broja učesnika IX-a njihovo direktno povezivanje u smislu BGP *peering*-a predstavlja složen administrativno-tehnički zadatak. Prvi problem je povezati sve administratore u cilju uspostavljanja *peering*-a a drugo

je da svako od njih ispravno konfigurise svoju opremu. Umesto toga se uvodi *Route* server posredstvom kog se vrši razmena ruta. Specifičnost *Route* servera u odnosu na neki ruter koji bi vršio ulogu centralne baze ruta se ogleda u tome što *Route* server može da vrši selektivno formiranje tabele rutiranja, za svakog učesnika ponaosob. Na Sl. 1 prikazan je način formiranja i distribucije ruta u tipičnom BGP ruteru. Sve dotle dok nema preklapanja oglašavanja koje stiže od korisnika, tj. dok se do nekog odredišta može stići samo na jedan način, rešenje sa klasičnim BGP ruterom će dobro funkcionisati. Problem se pojavljuje kada se pojavljuju višestruke rute za isto odredište koje imaju različitu metriku. Prilikom formiranja lokalnog RIB-a (*Routing Information Base* – Sl. 1) ruter će sprovesti proces izbora najbolje rute u skladu sa definicijom BGP-a i samo jednu rutu smestiti u RIB. S druge strane, ako bi korisnici želeli da koriste neku od alternativnih ruta i to svaki korisnik bi imao drugačije želje, problem bi bio nerešiv. Upravo ovde se koriste dodatne funkcije koje omogućava Quagga softverski paket [2]. Na Sl. 2 prikazano je na koji način se u Quagga softveru formiraju tabele rutiranja i kako se one oglašavaju korisnicima sa kojima je uspostavljen BGP *peering*.



Sl. 1. – Prijem i oglašavanje ruta u tipičnom BGP ruteru ([2])



Sl. 2. – Prijem i oglašavanje ruta u Quagga *Route* serveru ([2])

U slučaju da se koristi *Route* server sa Quagga softverom, imamo da se na *Route* serveru za svakog korisnika, sa kojim se uspostavlja BGP *peering*, formira posebna tabela rutiranja - RIB. Pored standardnog filtriranja ruta u odlaznom i dolaznom smeru, Quagga omogućava pojedinačno filtriranje oglašavanja za svakog korisnika ponaosob. Korišćenjem *import* i *export* polisa, Quagga omogućava sledeće: *import* polisama se kontroliše koje rute prilikom prijema će biti ubačene u koju tabelu rutiranja (RIB); *export* polisama se kontroliše za svakog korisnika šta će da oglasi ka kom drugom korisniku. Uvođenjem *import* i *export* polisa proces filtriranja se od korisnika seli na *Route* server. U kombinaciji sa *community* atributima ovaj proces se može dodatno automatizovati tako da korisnik ne mora da ima direktan pristup *Route* serveru da bi izvršio promenu oglašavanja i filtriranja saobraćaja.

Druga bitna osobina Quagga softvera je da ima mogućnost brisanja sopstvenog AS broja iz *AS path* atributa prilikom oglašavanja. Ova osobina je veoma bitna zbog činjenice da se izbor najbolje rute u BGP procesu odlučivanja najčešće obavlja na osnovu dužine *AS path* atributa. Uvođenje *Route* servera sa svojim AS brojem definitivno produžava *AS path* atribut za jedan što u krajnjem može da ima da putanja preko IX-a bude nepovoljnija što nije bio cilj. Da bi se ovo izbeglo koristi se funkcija brisanja sopstvenog AS broja iz *AS path* atributa. Krajnji rezultat je da *AS path* atribut ima vrednost kao da su ISP-evi uspostavili direktan BGP *peering* između sebe. Mali problem kod korišćenja ove funkcije je činjenica da specifikacija BGP protokola ovo ne smatra za regularno ponašanje. Cisco Systems je od verzije 12.0 IOS-a uveo ograničenja za ovakav rad. U slučaju da se pokuša uspostaviti BGP *peering*-a sa Cisco ruterom, čim se dobije prvo ovakvo oglašavanje, Cisco ruter će raskinuti BGP *peering*. Zbog toga moramo dozvoliti korišćenje ovakve "neregularnosti" time što ćemo eksplicitno dozvoliti uspostavu BGP sesije sa susedom koji oglašava rute a da u njima nema njegovog AS broja. To se postiže komandom "no bgp enforce-first-as" u delu BGP konfiguracije. Kod Juniper rutera se ne pojavljuje ovaj problem, jer ruter bez problema prihvata i ovakve rute.

### III POVEZIVANJE SA INTERNETOM

Jedan od servisa koji je u planu da bude ponuđen korisnicima SOX-a je i povezivanje sa velikim svetskim Internet *Exchange* čvorištima. Za svakog ISP-a u Srbiji je relativno skupo da sam realizuje sopstveni link do DECIX-a, LINX-a, AMSIX-a i sličnih velikih komunikacionih čvorišta. Razvojni plan SOX-a predviđa da SOX realizuje takve linkove i da ovde u Srbiji ponudi povezivanje na ova velika čvorišta. Ovakvo povezivanje može da bude realizovano barem na dva načina. Jedan način predviđa da se između korisnika SOX-a i svetskog IX-a uspostavi L2 tunel kroz koji bi korisnik SOX-a imao direktan pristup tom IX-u, od njega dobio parametre za rad i sa njim dogovario politiku rutiranja saobraćaja. Drugi način realizacije ovog servisa podrazumeva formiranje ekstenzije svetskog IX-a na SOX-u na isti način na koji se realizuje i domaći IX. Formirao bi se poseban VLAN-u kome bi bili svi korisnici SOX-a koji žele pristup

svetskom IX-u. BGP *peering* bi se uspostavljao sa *Route* serverima SOX-a koji bi za te potrebe formirali posebnu tabelu rutiranja koja bi sadržavala rute dobijene od svetskog IX-a i rute koje SOX treba u ime svojih korisnika da oglasi ka svetskom IX-u. U realizaciji ovog servisa SOX bi se pojavio samo kao posrednik između domaćih korisnika i svetskih IX-eva.

### IV ETHERNET EXCHANGE

Savremene svetske tendencije prepoznaju Internet *Exchange Point* (IXP-Evropski naziv), tj. *Network Access Point* (NAP-Američki naziv) kao elemente savremenih telekomunikacija koji značajno unapređuju poslovanje svih učesnika na telekomunikacionom tržištu. Gotovo redovno svoje postojanje IXP, tj. NAP ostvaruju u tzv. *carrier-neutral colocation facility*, koji se u Srbiji rado nazivaju *Telehouse*. Veliki Američki i Evropski operateri su prepoznali koristi od „virtuelnih“ povezivanja, pa su čak i pojedine kompanije koje se bave pružanjem „virtuelnih“ usluga međupovezivanja postigla strateška partnerstva, tj. proširili svoje usluge na šire geografsko područje. Svetska kriza donekle je usporila tempo izgradnje savremenih optičkih pristupnih mreža, pa je saradnja između operatera i nepreklapanje izgrađene infrastrukture došlo u prvi plan, zajedno sa pitanjima interoperabilnosti, kontrole kvaliteta veza, nadzora i korektivnog održavanja itd... Savremene računarske mreže dominantno počivaju na Ethernet protokolima, pa su se razvojem *Carrier Ethernet* seta standarda, pre svega kroz MEF (*Metro Ethernet Forum*), ali i kroz druga standardizaciona tela stekli uslovi za pojavu novog koncepta: *Ethernet Exchange*, koji koristeći sve predosti visoko skalabilne, pouzdane i protokolima opisane tehnologije, uz izuzetno nisku cenu po Mbps donosi novu vrednost na telekomunikaciono tržište međupovezivanja. Primera radi 20 operatera u gradu za svoje *full mesh* povezivanje treba da upotrebe 191 optički put i 382 GEth portova, od kojih će mnogi naravno raditi daleko ispod svog kapaciteta. Upotrebom *Ethernet Exchange*-a potrebno je 20 optičkih prenosnih puteva i 40 GEth portova, odnosno, za redundantno rešenje (primenjeno u SOX topologiji) 40 optičkih puteva i 80 GEth portova za potpuno povezane mreže u kojima je moguće izvršiti separaciju servisnih platformi i postići veliku pouzdanost, velike bitske protoke i možda najvažnije, brza uspostava međupovezivanja. U tom smislu se SOX može nazvati i *Ethernet Exchange* te ponuditi i prethodno navedene servise.

### V ZAKLJUČAK

Iz prethodnog izlaganja se može videti na koji način je realizovan SOX *exchange*. Korišćenje *ethernet* tehnologije omogućava veliku fleksibilnost i dobre performanse u radu. Umesto da se saobraćaj rutira na trećem nivou, on se komutira (*switching*) na drugom sloju. Korišćenjem standardnih mehanizama za garantovanje QoS-a u *ethernet switch*-evima, obezbeđuje se realizacija servisa sa potrebnim komunikacionim parametrima. U smislu daljeg razvoja SOX može da postane berza za razmenu telekomunikacionih usluga jer bi se na jednom mestu

mogli naći svi veliki kreatori i korisnici Internet servisa kao što su Internet servis provajderi i *content* provajderi.

#### LITERATURA

- [1.] S.Shah, M.Yip, "Extreme Networks' Ethernet Automatic Protection Switching (EAPS) Version 1", RFC 3619
- [2.] Kunihiro Ishiguro et al. "Quagga documentation", <http://www.quagga.net/>
- [3.] "The IX Point Best Current Practice", <http://www.euro-ix.net/>

*Abstract* – In this paper authors presents their own experience in realization SOX Exchange. It's described all main technical features that are used in SOX with details important for SOX users.

#### REALIZATION OF SOX EXCHANGE

Nenad Krajnović, Zoran Perović