TARTU ÜLIKOOL

Füüsika-keemiateaduskond

Eksperimentaalfüüsika ja tehnoloogia instituut

Risto Rahu

CISCO NETFLOW ANALÜSAATOR

Magistritöö

Juhendaja: dr EERO VAINIKKO

Tartu 2006

SISUKORD.

Joonised:

lk.

Tabelid:

lk.

Sissejuhatus

Arvutivõrgu liikluse statistika kogumine on olnud aktuaalne alates interneti suurema leviku algusest 1990ndatel aastatel – üheks peapõhjuseks kindlasti sidefirmade ja interneti teenusepakkujate soov omada ülevaadet kasutatavate sidekanalite koormusest, mis aitab planeerida, kas , mida ja kuskohas on vaja võrgus paremaks teha.
Kõige lihtsam statistika on võrguliikluse andmemaht, mõõdetuna bittides või baitides. Ja kui siia juurde tuua aeg, saab ülevaate ka kiirusest ehk kasutatavast ribalaiusest. Marsruuterist (või muust võrguseadmest) läbimineva infohulga kokkuarvutamine on võrdlemisi lihtne tegevus ega nõua palju protsessoriaega, samuti pole probleeme niimoodi saadud informatsiooni esitamisega, näiteks piltidena veebilehel. Infovahetuseks kasutatakse enamasti SNMP protokolli ning võrguliikluse graafiliseks kujutamiseks on üks lihtsamalt häälestatavaid lahendusi Tobi Oetiker poolt valmistatud MRTG [] nimeline tarkvara (info aadressilt http://oss.oetiker.ch/mrtg/).

Kuid sellisel lihtsa baitide kokkuliitmise meetodil on puudused:

- Pole teada võrguressursi kasutamise jaotus arvutite/võrkude järgi
- Pole teada võrguressursi kasutamise teenuste järgi
- Puudub võimalus hiljem salvestatud info hulgast otsida midagi konkreetse arvuti või arvutivõrgu kohta.

Kui konkreetse arvuti või arvutivõrgu liikluse kohta statistika kogumine on põhiliselt turvalisuse alla kuuluv teema, siis võrguteenuste jaotuse teadmine annab olulist infot ka võrgu planeerimise kohta – näiteks kuhu võrgus paigutada vaheservereid (cahce server), tulemüürid , koormusejagajad jne.

Ülaltoodud probleeme aitavad lahendada põhiliselt kaks asja:

- Tulemüürid (i.k. *firewall*), mis peavad niigi arvet lubatud või keelatud arvutite, võrkude ja võrguteenuste kohta,
- Netflow , Sflow ja teised sarnase protokolli toetust omavad ning kesksele analüüsisüsteemile infot edastavad marsruuterid

Tulemüürid asuvad enamasti arvutivõrgu “välispiiril” ning läbi nende ei käi võrgusisene liiklus, seetõttu jääb ainult tulemüüridest saadava info korral teadmata sisemise võrgu liikluse statistika. Netflow [] (Cisco Systems poolt välja töötatud protokoll võrguliikluse info edastamiseks, seda toetavad veel ka Juniper jt võrguseadmete tootjad) ning Sflow [] (Foundry Networks poolt välja töötatud protokoll võrguliikluse info edastamiseks) võimaldavad saada statistikat ka sisevõrgu marsruuterit või võrgulülitit (i.k. *switch*) läbiva info kohta.
Lisaks marsruuteritele on Netflow (ja ka Sflow) infokogujaid saadaval ka eraldiseisvate seadmetena , mis võivad toimida kas pealtkuulamise rezhiimis või siis läbi seadme mineva liikluse infot analüüsides.
Kuna täpsema info kogumine võrguliikluse kohta , mis eeldab kõikide uuritava andmepaketi liikumise detailide läbivaatamist, on võrreldes lihtsa liitmisega (mida tehakse summaarse läbimineva infovoo teadasaamiseks) väga palju protsessoriaega nõudvam, on suuremal kiirusel (tänapäeva mõistes mitmed Gigabit/s) töötavad infokogujad (sensorid) reeglina keerulised ja ka väga kallid [].

Töö eesmärk

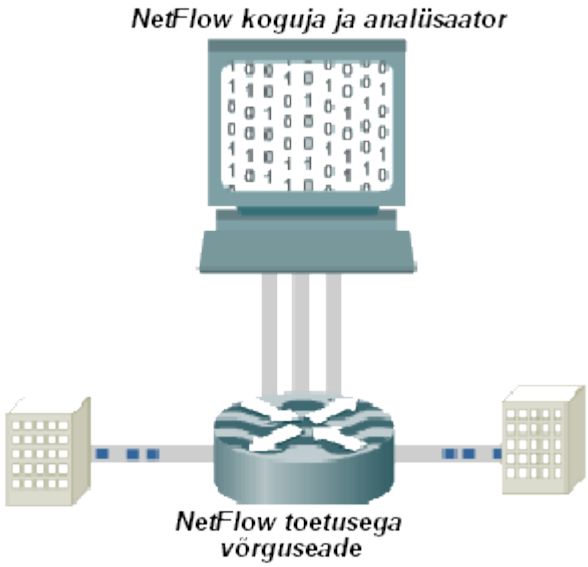
Konsulteerides Tartu Ülikooli infotehnoloogia osakonna arvutivõrgu peaspetsialisti hr Erkki Kukk-ega sai minu magistritöö eesmärgiks seatud Cisco marsruuteritel kasutatava NetFlow protokolli kasutamist võimaldava lahenduse väljatöötamine Tartu Ülikooli arvutivõrgu liikluse statistika ja analüüsi jaoks.
Töö tulemusena peab valmima lahendus, mis võimaldab marsruuterist tuleva info talletamist mingisugusesse andmebaasi, kus Netflow info peab olema talletatud sessioonidena , on näha kes-kellega suhtleb ja mis mahus (kellaeg, kestvus, lähte- ja sihtaadressid, pordinumbrid ning pakettide arv ja maht bittides).
Kui info on sellisel kujul baasis olemas, siis kasutajaliidese (soovitavalt veebipõhise) abil peab saama sellele päringuid esitada, mõned olulisemad oleksid:

1.) Kes kasutavad võrku kõige rohkem (top10 ip aadressidest, kelle maht on kõige suurem) või kes kasutavad kõige intensiivsemalt (top10 ip aadressidest, mille pakettide arv on kõige suurem)
2.) Milliseid teenuseid kasutatakse kõige rohkem
3.) Kui palju ribalaiust kasutatakse igas alamvõrgus
4.) Kas toimub DoS (teenusetõkestus - i.k. *denial of service*) rünnakuid ja kes on ründaja?
5.) Kõikides tabelites hiireklõps IP aadressile avab selle aadressi kõik sessioonid (mingis ajavahemikus, mida saab ka muuta).
6.) Suvalise hosti liiklus mingis ajahetkes

Lisaks päringutele peaks kasutajaliideses võimaldama visuaalselt kujutada liiklust piltidena, näiteks MRTG või RRD tarkvara abil.
Netflow versioonidest võiks olla toetatud versioonid 5 , 7 ja 9. Viimane neist on eriti huvipakkuv IPv6 liikluse analüüsimise jaoks.

Ülevaade kasutatavatest tehnoloogiatest

Kõige olulisemaks kasutatavaks abivahendiks töö koostamisel on firma Cisco Systems poolt välja töötatud NetFlow protokoll [].
Cisco IOS Netflow võimaldab detailselt kirjeldada võrguseadmest (marsruuterist või võrgulülitist) läbiminevat võrguliiklust, selle kasutusala on lai, ulatudes võrguliikluse statistikast kuni turvalisuse tagamise lahendusteni.



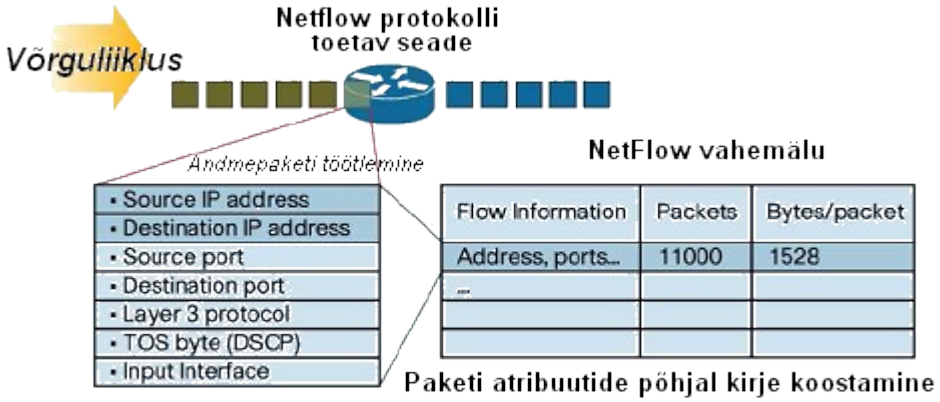
Joonis 1 Cisco NetFlow andmevahetuse ülevaade

Oluline on sealjuures, et võrguseadmest läbiminevate andmepakettide (vt Joonis 1) kohta saadetakse info ainult selle päise (mis sisaldab sihtaadresse, pordinumbreid jt.) ja suuruse kohta, andmepaketi sisu ei uurita. Edastatava infomahu ja protsessoriaja vähendamiseks kasutatakse sessioonipõhist koondamist (vt. lk) ning pistelist läbivaatust (i.k. *sampled netflow*), mille korral vaadeldakse sessiooni pakette ainult üle teatava intervalli. Selline lähenemine kaotab küll informatsiooni täielikkuse kuid tagab siiski piisava infohulga üldise statistika jaoks.
Põhiline Netflow väljund on vookirje (i.k. *flow record*). Vookirjeid tekitav seade (nimetatakse ka NetFlow sensor) saadab loodud vood ühele või mitmele kogumis ja/või analüüsiseadmele (nimetatakse ka NetFlow kollektor ja analüsaator).
Netflow võimaldab infot anda võrguliikluse kohta nii etherneti- , IP- kui transpordi tasemete kohta (laienduste abil veel ka rohkem).

Peamised Netflow poolt kasutatavad IP paketi atribuudid on järgmised:

- IP lähteaadress (i.k. *source address*)
- IP sihtaadress (i.k. *destination address*)
- Lähteport (i.k. *source port*)
- Sihtport (i.k. *destination port*)
- IP protokolli tüüp (i.k. *protocol type*)
- Teenuseklass (i.k. *class of service*)
- Marsruuteri või võrgulüliti liides (i.k. *interface*)

NetFlow vookirjed koondatakse sessiooni põhiselt - kõik sama lähte- ja sihtaadressiga , sama lähte- ja sihtpordi numbriga ning sama protokolli, võrguliidese ja teenuseklassiga paketid grupeeritakse üheks andmevooks ning arvutatakse selle voo kohta summaarne baitide ja pakettide hulk. Selline grupeerimine (nimetatakse ka sessioonipõhine identifitseerimine) on kergesti skaleeritav ka suure võrguinfo mahu korral , võimaldades info salvestamist sisemisse andmebaasi. Cisco marsruuterites nimetatakse seda NetFlow cache.
Suurte võrguinfo mahtude korral on väga oluline faktor ka iga konkreetse marsruuteri mälumaht ja protsessori jõudlus, kuna iga kirje Netflow vahemälus võtab ruumi (vähemasti mitukümmend baiti) ning nende voogude töötlemine nõuab protsessori aega.



Joonis 2 Cisco NetFlow sessioonipõhine filtreerimine

Netflow info töötlemine Cisco marsruuterites käib järgnevalt:

- Konfigureeritakse NetFlow info salvestamine NetFlow vahemällu (*cache*).
- Konfigureeritakse Netflow info väljastamine kogumisseadmele (*Netflow collector*).
- Netflow vahemälust otsitakse lõpetatud vood (sessioonid) ning edastatakse nende info kogumisserverile. Lõpetatuks loetakse voog, kui möödub teatud aeg viimase paketi saabumisest või saabub vastav lõppu teatav pakett.
- Korraga saadetakse kogujale UDP pakettidena tavaliselt 30...50 voo informatsioon.
- Netflow kogumistarkvara koostab saadud info põhjal raportid (kas hetkeseisu või eelnenud aja kohta).

NetFlow protokolli toetab suur hulk Cisco võrguriistvara (vt Tabel 1) ning ka teiste tootjate võrguseadmetel- või tarkvaral (näiteks Tabel 7ja Tabel 8).

Tabel 1 Netflow toetus Cisco riistvaral:

Cisco võrguseade	**Netflow toetus**
Cisco 800,1700 ja 2600	Jah
Cisco 1800,2800 ja 3800	Jah
Cisco 4500	Jah
Cisco 6500	Jah
Cisco 7200, 7300 ja 7500	Jah
Cisco 7600	Jah
Cisco 10000, 12000 ja CRS-1	Jah
Cisco 2900, 3500, 3660 ja 3750	Ei

Netflow protokollist on kasutusel palju erinevaid versioone, viimane antud hetkel on versioon 9.
Netflow kogumisseadmed (või tarkvara) võimaldab reeglina korraga töödelda mitme eri versiooni ning erinevatest allikatest tulevat NetFlow informatsiooni.

Cisco NetFlow versioon 9

Mõni sõna versiooni 9 omadustest, täpse ülevaate saab Cisco Systems võrguleheküljelt http://www.cisco.com/en/US/products/ps6601/products_white_paper09186a00800a3db9.shtml []
Põhiliseks erinevuseks eelmiste versioonidega on shabloonide (i.k. *template*) kasutamine – see võimaldab kasutada praegust versiooni voo kirjeldamisel ka võimalike tulevaste laienduste jaoks. Näiteks kui tuleb välja mingi uuendus Netflow infos, tuleb shabloone kasutavas tarkvaras täiendada ainult väike osa koodi, mis vastab uuele shablooninfo tüübile ja tegeleb konkreetset tüüpi info töötlemisega.
Tabel 2 ja Tabel 3 esitavad lihtsustatult Netflow andmepaketi struktuuri:

Tabel 2 NetFlow versioon 9 andmepakett:

Paketi päis	Shabloon nr1	Andmed 1	Andmed 2	...	Andmed N	Shabloon nr 2	Andmed 1	...

Paketi päises antakse edasi kasutatava NetFlow versioon, antud juhul 0x0009, paketiga edastatavate shabloon- ja andme kirjete koguhulk , infot saatva seadme töötamise aeg (i.k.*system uptime*) , väljastatava paketi järjekorranumber ning NetFlow väljastaja identifikaator (mis aitab üheselt eristada sama seadme poolt saadetavaid voogusid).

Tabel 3 NetFlow shabloonikirje lihtsustatud struktuur:

16 bitine väärtus:
FlowSet ID=0
Kirje pikkus
Shablooni number (1)
Andmeväljade arv
Andmevälja 1 tüüp
Andmevälja 2 tüüp
Andmevälja 2 pikkus
.....
Andmevälja N pikkus
Shablooni number (2)
Andmeväljade arv
Andmevälja 1 tüüp
.....
Andmevälja N pikkus
.....
Shablooni number (N)
.....

Iga shablooni kirje määrab ära vastava andmekirje struktuuri, vt Tabel 4, väike väljavõte võimalike andmekirjete tüüpidest koos pikkuse ja selgitusega, on toodud Tabel 5 (see on ühtlasi ka programmi NfDump poolt toetatud NetFlow andmetüüpide nimekiri).

Tabel 4 NetFlow andmekirje lihtsustatud struktuur:

16 bitine väärtus:
FlowSet ID=vastava shablooni number
Kirje pikkus
Kirje 1 , andmevälja 1 väärtus
Kirje 1 , andmevälja 2 väärtus
.....
Kirje 1 , andmevälja N väärtus
Kirje 2 , andmevälja 1 väärtus
Andmevälja 1 tüüp
.....
Kirje 2 , andmevälja N väärtus
Kirje M, andmevälja väärtus
.....
Kirje M, andmevälja N väärtus
Täitebitid

Tabel 5 Mõningad NetFlow andmetüübid (NfDump poolt toetatud andmetüübid):

Andmevälja tüüp	**Väärtus**	**Andmevälja pikkus**	**Selgitus**
IN_BYTES	1	N (vaikimisi 4)	Konkreetse IP *flow* kirje sisenevate baitide loendur, pikkusega N x 8 bitti.
IN_PKTS	2	N (vaikimisi 4)	Konkreetse IP *flow* kirje sisenevate pakettide loendur, pikkusega N x 8 bitti.
FLOWS	3	N	IP *flow* koondatud kirjete arv; vaikimisi N 4
PROTOCOL	4	1	IP protokoll
SRC_TOS	5	1	Teenuse tüübi bait, loetuna sisendliidese pealt
TCP_FLAGS	6	1	IP *flow* jaoks kumulatiivne TCP lippude seis
L4_SRC_PORT	7	2	TCP/UDP lähtepordi number
IPV4_SRC_ADDR	8	4	IPv4 lähteaadress
INPUT_SNMP	10	N	Sisendliidese number; vaikimisi N on 2 baiti
L4_DST_PORT	11	2	TCP/UDP sihtpordi number
IPV4_DST_ADDR	12	4	IPv4 sihtaadress
OUTPUT_SNMP	14	N	Väljundliidese number; vaikimisi N on 2 baiti
SRC_AS	16	N (vaikimisi 2)	Siseneva paketi BGP autonoomse süsteemi number, N on kas 2 või 4
DST_AS	17	N (vaikimisi 2)	Väljuva paketi BGP autonoomse süsteemi number, N on kas 2 või 4
LAST_SWITCHED	21	4	Ajahetk alates arv süsteemi käivitamisest, millal voo paketti viimati töödeldi
FIRST_SWITCHED	22	4	Ajahetk alates arv süsteemi käivitamisest, millal voo esimest paketti töödeldi
IPV6_SRC_ADDR	27	16	IPv6 lähteaadress
IPV6_DST_ADDR	28	16	IPv6 sihtaadress

Kõikides Tabel 5 toodud andmevälja pikkuste korral on tegemist N*8 bitiste väärtustega, kui pole öeldud teisiti.

Kasutatava välja pikkus sõltub infot edastava seadme rollist ja koormusest võrgus – näiteks kui on tegemist väikese arvutivõrgu keskmarsruuteriga, on täiesti piisav 32 bitiste (N=4) andmeväljade kasutamine, samas suure arvutivõrgu korral tuleb kindlasti kasutada juba 64 bitiseid väljapikkuseid, et tagada korrektne summeerimine.
Üheks väga oluliseks detailiks NetFlow versioon 9 juures on shablooni valikute kirje (i.k *options template format*). Kirje formaat on toodud .

Tabel 6 NetFlow shablooni valikute kirje struktuur:

16 bitine väärtus:
FlowSet ID=1
Kirje pikkus
Shablooni number
Valikute skoobi kogupikkus
Valikväärtuste kogupikkus
Skoobi 1 tüüp
Skoobi 1 väljapikkus
.....
Skoobi N tüüp
Skoobi N väljapikkus
Valikväärtuse 1 tüüp
Valikväärtuse 1 väljapikkus
.....
Valikväärtuse N tüüp
Valikväärtuse N väljapikkus
Täitebitid

- FlowSet ID=1 aitab eristada shablooni kirjet andmekirjetest, millede puhul FlowSet ID on alati suurem,kui 255.
- Kirje pikkus näitab kirje kogupikkust, sealhulgas sisaldades ka pikkuse väärtuse baite ning FlowSet ID väärtuse baiti.
- Shablooni number eristab üheselt antud võrguseadme kohta konkreetse shablooni. Shablooni number on alati suurem, kui 255.
- Valikute skoobi kogupikkus näitab kõikide sisalduvate skoobi väljade pikkust baitides.
- Valikute kogupikkus näitab kõigi kirjeldatud valikute osa pikkust baitides.
- Skoobi <N> tüüp näitab ära, millele vastav valikväärtuse kirje viitab.

  Defineeritud on järgmised väärtused
  - 0x0001 - süsteem
  - 0x0002 - võrguliides

- 0x0003 - liidese kaart
- 0x0004 - netflow vahemälu
- 0x0005 - shabloon

Valikväärtuse <N> tüüp on numbriväärtus, mis kirjeldab antud skoobi kohta väljatüübi. Tüübid on samad, mis eelnevalt andmetüüpide Tabel 5 poolt kirjeldatud.

NetFlow väljastamise konfigureerimine

Kõige lihtsam viis Cisco marsruuteritel Netflow töötlemise ja edastamise aktiveerimiseks ühe võrguliidese peal näeb välja alljärgnev []:

Router#configure terminal
Router(config)# interface fastethernet 0/0
Router(config-if)# ip route-cache flow

Et öelda marsruuterile kuhu Netflow info saata,tuleb anda järgmised käsklused:
Router(config)# ip flow-exportip flow-export version 5
Router(config)# ip flow-cache timeout active 5
Viimase korraldusega tükeldame pikaajalised Netflow kirjed 5-minuti pikkusteks osadeks (valida võib intervalli 1 kuni 60 minutit vahel)
Riistvaralise marsruutimise toetusega seadmetel, nagu seda on Cisco 6500/7600, tuleb kindlasti lubada "NDE" kasutamine lisaks tavalisele Netflow väljastamisele. NDE (i.k. *Netflow Data Export*) on riistvaraline variant Netflow väljastamisest Cisco Catalyst 6500/7600-OSR seadmetel. Konfiguratsiooni näide:
Router(config)# mls flow ip interface-full
Router(config)# mls flow ipv6 interface-full
Router(config)#mls nde sender version 5

NB! Praeguseks hetkeks ei ole IPv6 NDE veel realiseeritud, kuid on võimalik vaadata "*online*" kirjeid "show mls netflow ipv6" korralduse abil käsurealt.

Lisaks Netflow voogude väljastamisele on Cisco marsruuteritel olemas võimalus Netflow info küsimiseks ja töötlemiseks ka käsurealt (CLI) või ka SNMP protokolli vahendusel.
Kõige lihtsam konfiguratsioon on alljärgnev

Router#configure terminal
Router(configure)#ip flow-top-talkers
Router(configure-flow-top-talkers)# top 50
Router(configure-flow-top-talkers)# sort by bytes
Router(configure-flow-top-talkers)# end

Siia juurde saab lisada käskusid filtreerimise kohta (IP aadress või võrk, port, paketi suurused jne).
Tulemuste vaatamiseks käsklus
Router>show ip flow top-talkers
Tulemus võiks olla näiteks selline:

SrcIf	SrcIPaddress	DstIf	DstIPaddress	Pr SrcP DstP	Pkts
Gi0	202.158.192.34	Local	202.6.112.239	2F 0000 0000	45K
Gi1	202.6.112.5	Null	233.29.147.222	11 800E DC91	8733
Gi1	202.6.112.12	Gi2	203.59.48.154	2F 0000 0000	2624
Gi2	203.19.110.254	Gi1	202.6.112.12	11 D606 2707	2339
Gi2	203.19.110.254	Gi1	202.6.112.5	11 D606 2707	2339

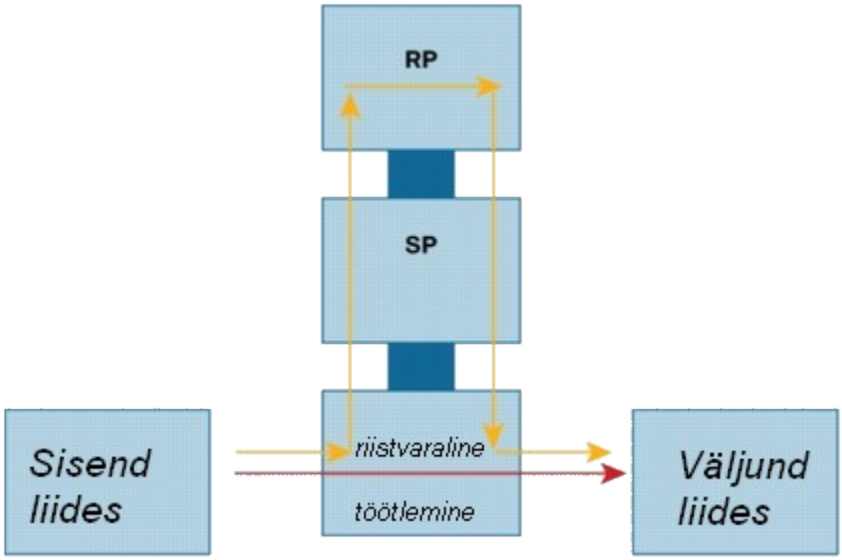
Ülevaade kasutatavatest seadmetest ja arvutivõrgu topoloogia ülesehitus

Tartu ülikooli arvutivõrk koosneb tänapäeval peamiselt gigabit-etherneti tehnoloogial ühendatud võrgulülititest (switch), palju kasutatakse VLAN arhitektuuri. Hoonetevahelised ühendused on teostatud põhiliselt valguskaablite abil.
Kogu arvutivõrgu IP marsruutimine toimub põhiliselt ühes (suures) marsruuteris, milleks on firma Cisco Systems seade Cisco Catalyst 6506 switch []. See tagab ka minu töö teostamisel väärtusliku suuremahulise Netflow infovoo, millega saab testida loodava lahenduse tööd ka suurte koormuste all.

Mõned Cisco Catalyst 6506 põhiparameetrid:

- toetab *Supervisor engine* riistvara, kasutusel Supervisor Engine-2, 2GE, MSFC-2 / PFC-2.
- suur valik etherneti mooduleid (100Mbit/s, 1Gbit/s ja 10Gbit/s)
- Kõneside moodulid
- Flex Wan ja ATM moodulid laivõrgu ühendamiseks
- *Multi Gigabit services* moodulid erinevate võrguteenuste (sh sisupõhised suunamised, tulemüür(*firewall*) ,ründetuvastus (*IDS*), IPSec VPN ühendused, võrguanalüüs, SSL teenuste kiirendamine jt.) teostamiseks seadet läbival liiklusel.
- *Suur jõudlus - 256Gbit/s (switching), 30Mpps*

Netflow info väljund seadmel kahte teed pidi (vt Joonis 3): riistvaraliselt (toetatud Netflow v5 ja v7) ja tarkvaraliselt (toetatud Netflow v5,v7 ja v9) ruuteri moodulist. Enamus pakette läbib ainult riistvaralise töötlemise, paketid, mille sihtkoht või töötlemise kulg pole veel teada läbivad ka teenuse protsessori (SP) ja marsruutingu protsessori (RP) moodulid. Edaspidi kirjeldatakse järgmiste sarnaste pakettide töötlus juba riistvara poolele ära ning ülejäänud paketid läbivad ainult riistvaralist osa. Selline töötlemise viis tagab suure jõudluse koos vajaliku paindlikkusega.



Joonis 3 Pakettide liikumine Cisco Catalyst 6500 seadmes

Tartu Ülikooli arvutivõrgus on kasutusel peamiselt IPv4 aadressruum, mis on enamasti avalikud (välised) IP aadressid, kuid kasutusel on ka sisemised (privaatsed) IP aadressid.
Kokku on IPv4 võrgusegmente kasutusel umbes 40.
Paralleelselt on käigus ja arendamisel ka IPv6 aadressruumi kasutamine.

Ülevaade valdkonna olemasolevatest lahendustest

Cisco Netflow info kogumiseks ja töötlemiseks on olemas palju tarkvarapakette, igal suuremal võrguseadmete tootjal on reeglina olemas oma süsteem, kuhu kuulub ka Netflow analüsaator.
Alljärgnevalt väike ülevaade peamistest Cisco Netflow töötlemiseks mõeldud valmislahendustest või tarkvarapakettidest. Tabelis 2 toodud nimekirjas on valik Cisco Systems poolt [,] välja pakutavast kommertstoodetest mis sobib NetFlow info töötlemiseks, tabelis on ära mainitud tarkvara sihtgrupp, platvorm ning hinnaklass (madal on alla $7000, keskmine $7000-$25000, kõrge >$25000).

Tabel 7 Valik Cisco NetFlow toetavat tasulisi valmislahendusi või tarkvara

Toote nimetus	Peamine kasutusala	Sihtgrupp	Tarkvara	Hinnaklass
Cisco NetFlow Collector	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Linux, Solaris	Keskmine
Cisco CS-Mars	Turva , monitooring	Suured- ja keskmise suurusega ettevõtted	Linux	Keskmine
AdventNet	Võrguliikluse analüüs	Suured- ja keskmise suurusega ettevõtted	Windows	Madal
Apoapsis	Võrguliikluse analüüs	Suurettevõtted	Linux	Keskmine
Arbor Networks	Turva / Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	BSD	Kõrge
Caligare	Võrguliikluse / Turva analüüs	Suurettevõtted, Teenusepakkuja	Linux	Keskmine
Crannog Software	Võrguliikluse analüüs	Suured- ja keskmise suurusega ettevõtted	Windows	Madal
CA Software	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Windows	Kõrge
Evident Software	Võrguliikluse analüüs, maksustamine	Suurettevõtted	Linux	Kõrge
HP	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Linux, Solaris	Kõrge
IBM Aurora	Võrguliikluse analüüs /Turva	Suurettevõtted, Teenusepakkuja	Linux	Keskmine
InfoVista (Crannog)	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Windows	Kõrge
IsarNet	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Linux	Keskmine
*Micromuse	Võrguliikluse analüüs	Suurettevõtted, Teenusepakkuja	Solaris	Kõrge
NetQoS	Võrguliikluse /Turva Analüüs	Suurettevõtted	Windows	Kõrge
Valencia Systems	Võrguliikluse analüüs	Suurettevõtted	Windows	Kõrge
Wired City	Võrguliikluse analüüs	Suurettevõtted	Windows	Kõrge

Ühe kommertstarkvaralahenduse, Cisco CS-Mars [] lühiülevaate leiab ka käesoleva töö lisast, lk 38,”Lisa 2 – Ülevaade CS-MARS lahendusest”.
Tegemist ei ole ainult NetFlow töötlemisele orienteeritud lahendusega vaid hoopis universaalse võrguhalduslahendusega, millega saab hallata suurt andmesidevõrku paljude erinevate (tootjate) võrguseadmetega. CS-Mars võimaldab saada hea ülevaate nii sündmustest võrgu suures plaanis kui laskuda ka üksiku turvaintsidendi detailidesse.
Tartu Ülikooli IT osakonna plaanides on ka Cisco CS-Mars lahenduse testimine, mis tõenäoliselt ei toimu siiski mitte enne käesoleva aasta sügist.

Kuna kommertstarkvara (tasulise tarkvara) kasutamine minu töö juures ei olnud võimalik ning seega keskendusin ainult vabavaraliste lahenduste otsimisele.

Tabel 8 Valik Cisco NetFlow toetavat tasuta tarkvara

Toote nimetus	Peamine kasutusala	Kommentaar	Tarkvara
Cflowd	Võrguliikluse analüüs	Arendamine peatatud	Unix
Flow-tools	Netflow kollektor	Laiendusvõimalused	Unix
Flowd	Netflow kollektor	Toetab netflow v9	BSD , Linux
FlowScan	Raportid Flow-tools paketile		Unix
NetFlow Guide	Raportid		BSD , Linux
NetFlow Monitor	Võrguliikluse analüüs	Toetab netflow v9	Linux
NTOP	Netflow kollektor	Toetab netflow v9	Unix
Panoptis	Turvamonitooring		Unix
Stager	Raportid Flow-tools paketile		Unix

Kõik minu poolt edaspidi käsitletud programmid on avatud lähtekoodiga ning edasiarendatavad , publitseeritud GNU GPL [] või BSD litsentsi [] alusel.

Täpsemalt uurides Tabel 8 toodud tarkvarapakettide detailsemat informatsiooni , tutvudes Sveitsi haridus- ja teadusvõrgu poolt kokku kogutud voopõhiste võrguhaldustarkvara nimekirjaga [] ning pidades silmas püstitatud eesmärgi saavutamiseks vajalikke kriteeriume jäi edasisele uurimisele kolm tarkvarapaketti:

- NTOP Netflow Probe (http://www.ntop.org/overview.html) []
- NfSen / NfDump (http://nfsen.sourceforge.net/) [,]
- PMACCT (http://www.ba.cnr.it/~paolo/pmacct/) []

Esimene neist oli ilusa veebipõhise liidesega, kuid ei salvestanud vooandmeid andmebaasidesse, teine tarkvara tundus keerulisem ja kirjutas andmeid andmebaasi asemel kettale failidesse , kolmas tundus kõige sobivam, kirjutades infot SQL andmebaasi.
Kõik väljavalitud paketid töötasid minu kasutada olnud Linux platvormil ning toetasid Cisco NetFlow versioon 9 vooinfot.

Minu lahendus töö eesmärgi saavutamiseks

Koostöös TÜ Infotehnoloogia osakonnaga (hr Erkki Kukk jt.) oli mul töö eesmärgi saavutamiseks kasutada üks Debian Linux operatsioonisüsteemiga serverarvuti (Intel Pentium III, 800MHz, 512MB RAM ja 20GB RAID0 HDD) mis oli täielikult minu kasutuses, võimalusega tarkvara juurde lisada / kompileerida.

Töö käigus paigaldasin serverile ja kasutasin alljärgnevat tarkvara:

- Apache veebiserver (*versioon 1.3.33*)
- MySQL server (*versioon 4.0.24*)
- PhpMyAdmin veebiliides andmebaasi haldamiseks (*versioon 2.6.2*)
- Pdnsd nimeserveri vahemälu [] (*versioon 1.1.11-par*)
- Perl (*versioon 5.8.4 , koos lisamoodulitega*)

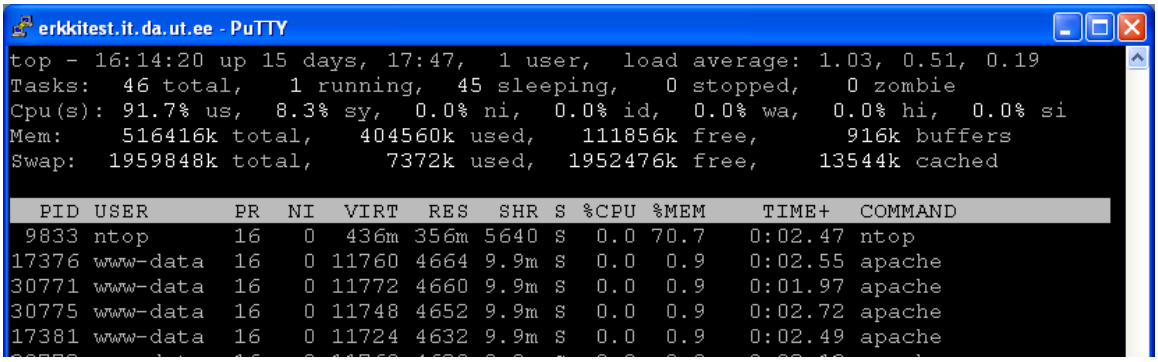
Ülikooli keskmarsruuteri (Cisco Catalyst 6506) poolt genereeritavatele Netflow voogudele ligipääsemiseks oli serveril konfigureeritud eraldi võrguliides.

Esimene etapp: Alustarkvara sobivuse hindamine

NTOP tarkvara.

NTOP on kombineeritud tarkvaralahendus võrguliikluse statistika esitamiseks ja analüüsimiseks. Kirjutatud perl/php/python vahenditega.
Kuna NTOP tarkvara häälestamine tundus kõige lihtsam, sai esmalt proovitud selle tarkvara tööd. Peale tarkvara seadistamist ja ülikooli keskmarsruuterist testserveri suunas saadetavate Netflow versioon 9 ja 7 voogude käivitamist tundus esmapilgul kõik kenasti töötavat. Tarkvara sorteeris Netflow infost välja

Joonis 4 Ekraanipilt serverarvuti koormuse kohta NTOP programmi korral



enimkasutatavad võrguteenused, serverid ja kliendid, kuid tarkvara jõudlus ning mälukasutus osutus antud lahenduse jaoks ebapiisavaks – juba paari minutiga oli programm kasutanud ~70% vaba mälu mahust ning ~90% protsessoriajast (vt Joonis 4) ning tõenäoliselt just sellepärast lõpetas tarkvara mingil ajahetkel lihtsalt töö, või seiskus veebiliides ja ei suutnud enam infot väljastada. NTOP tarkvara näitas aga väga head statistikat sissetuleva NetFlow liikluse kohta – sealt sai välja lugeda kuipalju pakette vastu võetud on , kirjete jaotus versioonide kaupa ning ülevaade vigadest. Kokkuvõttes tuli siiski tunnistada, et NTOP tarkvarapakett ei sobi püstitatud eesmärgi täitmiseks.

PMACCT

Edasi otsustasin proovida SQL andmebaasi kirjutavat tarkvara.
Pmacct on tarkvarapakett võrguliikluse analüüsimiseks, võimaldades toimida nii Netflow või Sflow kogujana kui ka passiivse statistikakogujana otse etherneti võrguliidesel. Programmi kohta leiab kena ülevaate autori ettekandest käesoleva aasta 26.aprillil RIPE aastakonverentsil Istanbulis [].
PMACCT programmi juures huvitas mind eelkõige Netflow kollektori funktsionaalsus, muid osasid programmist ma ei uurinud.
Kõik paketi programmid on kirjutatud C keeles, mis peaks tagama suure jõudluse.
Peale tarkvara kompileerimist, MySQL andmebaasi häälestamist ning vajalike veebiliideste seadistamist sain asuda reaalsete katsetuste kallale.
Esimeses etapis tekitas segadust vajalike reeglite seadmine tarkvarale, vaikimisi oli programm valmis infot sorteerima (ja salvestama) ainult ühe parameetri, näiteks IP aadress järgi. Leides vajalikud võtmed konfiguratsiooni juhendist saavutasin vajaliku info (kellaaeg, siht- ja lähteaadressid, siht- ja lähtepordid , infohulk) salvestamise andmebaasi. Programmi töö testimiseks kirjutasin lihtsa PHP skripti, mis otsis andmebaasis oleva info seest välja 10 suurema andmemahuga IP ühendust ja 10 kõige enam bitte liigutavat IP aadressi.
Alustades katseid ainult Netflow versioon 9 voogudega, tundus see tarkvara töötavat probleemideta – andmebaasi tekkisid kirjed, php skript kuvas seal vajaliku info.

Joonis 5 Ekraanipilt MySQL andmebaasi administreerimisliidesest, pmacct programmi katsetamise tulemusena

Kuid see rõõm ei kestnud kaua – peale Netflow versioon 7 voogude (mida on umbes 95% ülikooli keskmarsruuteri toodetavast netflow liiklusest) aktiveerimist oli tarkvara töö umbes 1.5h pärast peatunud – põhjuseks MySQL andmebaasi liidese seiskumine.

Peale täpsemaid vaatlusi selgus, et MySQL andmebaasis umbes 5,6 miljonit kirjet ja andmebaasi failid olid kokku umbes 1.3GB mahuga (siit saame ühe kirje suuruseks ~232 baiti), vt Joonis 5. Iga päringu tegemine andmebaasi võttis aega minuteid..
Oli jälle selge, et selle tarkvaraga (ja andmebaasimootoreid kasutades) püstitatud ülesannet lahendada pole mõistlik, sest isegi kui antud programm kirjutada ringi analoogiliselt NFDUMP poolt kasutatava 5 minuti pikkuste andmebaaside kasutamisega, tuleks umbes 4 kordne vahe andmebaasi suuruses nfdump kasuks.

NFSEN ja NFDUMP

Tarkvarapakett koosneb kahest osast – NfDump pakett on netflow voogude kollektor (*nfcapd* programm) koos töötlusutiliitidega (*nfdump* töötlemiseks, *nfprofile* vaadete genereerimiseks jt), NfSen on kasutajaliides NfDump tarkvarale, võimaldades selle poolt kirjutatud andmete põhjal näidata kasutajale graafiliselt ülevaadet olukorrast ning teostamaks otsingufunktsioone. Täpsem ülevaade paketist on toodud käesoleva dokumendi eraldi alajaotises “lk 30,Tarkvarapakettide Nfdump ja NfSen lühiülevaade” ning “lk 28, Netflow analüsaatori veebiliidese kasutusjuhend”.

Programmid on kirjutatud C (NfDump) ja PHP ning Perl (NfSen) keeles.
Ülevaade tarkvarapaketi omadustest on leitav autori ettekandest eelmisel aastal 2-6 mail Stokholmis toimunud RIPE (Réseaux IP Européens) 50ndal kokkusaamisel [].
Tarkvara esialgne häälestamine nõudis mõningate komponentide (Perl, RRDtool) kompileerimist kuid sujus kokkuvõttes kenasti.
Nfdump paketi häälestamine ei olnud keeruline ning nüüd sai juba näha detailsemalt mis mahus infovooga on Netflow puhul antud ülesande juures tegemist.
Päevasel ajal, kl 11.40 tuli netflow 5 minuti kirjeid sisaldava faili suuruseks ~27 MB.

```
erkkitest:/data/nfsen/profiles/live/ciso6500# ls -la nfcapd.200603211410
-rw-r--r--  1 www-data www-data 27929052 2006-03-21 14:15
nfcapd.200603211410
```

edasi lugesin read üle:

```
erkkitest:/data/nfsen/profiles/live/ciso6500# nfdump -r nfcapd.200603211410 | wc -l
537088
```

Siit on lihtsasti leitav voogude arve ühes sekundis , milleks tuli umbes 1790 kirjet, ühe kirje suurus ~52 baiti.
Siit veel üks järeldus – kui on soov saada ülevaadet pikema perioodi netflow info kohta, tuleb arvestada väga suure infohulga (mitmed gigabaidid informatsiooni!) läbitöötamisega, mis on väga aeganõudev ülesanne. Näiteks suudab testserver (sõltuvalt kasutatavast filtrist, muidugi) ühes sekundis töödelda umbes 350000 kirjet, seega võib 2GB andmehulga läbitöötlemiseks kuluda ~1.58 tundi.
Ka tuli välja üks oluline probleem – nimelt on testserveris kasutada 20GB kõvaketas, millest kasulikku vaba ruumi oli ~15GB. Seega jätkub kettaruumi maksimaalselt 46 tunni informatsiooni salvestamiseks.
Siinkohal tuli appi tarkvarapaketis juba sisaldunud võimalus seada maksimaalset andmefailide suurust kettal, mille ületamisel hakatakse vanu (kuupäeva/kellaaja järgi) kustutama, kuni saavutatakse minimaalne vajalik vaba kettamaht (mida saab ise seada, vaikimisi on see 90% maksimaalsest mahust).
Ka muudelt omadustelt oli tarkvarapakett igati sobiv minu töö ülesande täitmiseks – programmi kood ja veebiliides oli kirjutatud arusaadavalt ning tundus täiendatav. Kõik see võimaldas mul edasise töö jaoks keskenduda just NfSen / NfDump tarkvarapaketi täiendamisele.

Teine etapp: Tarkvara täiendamine

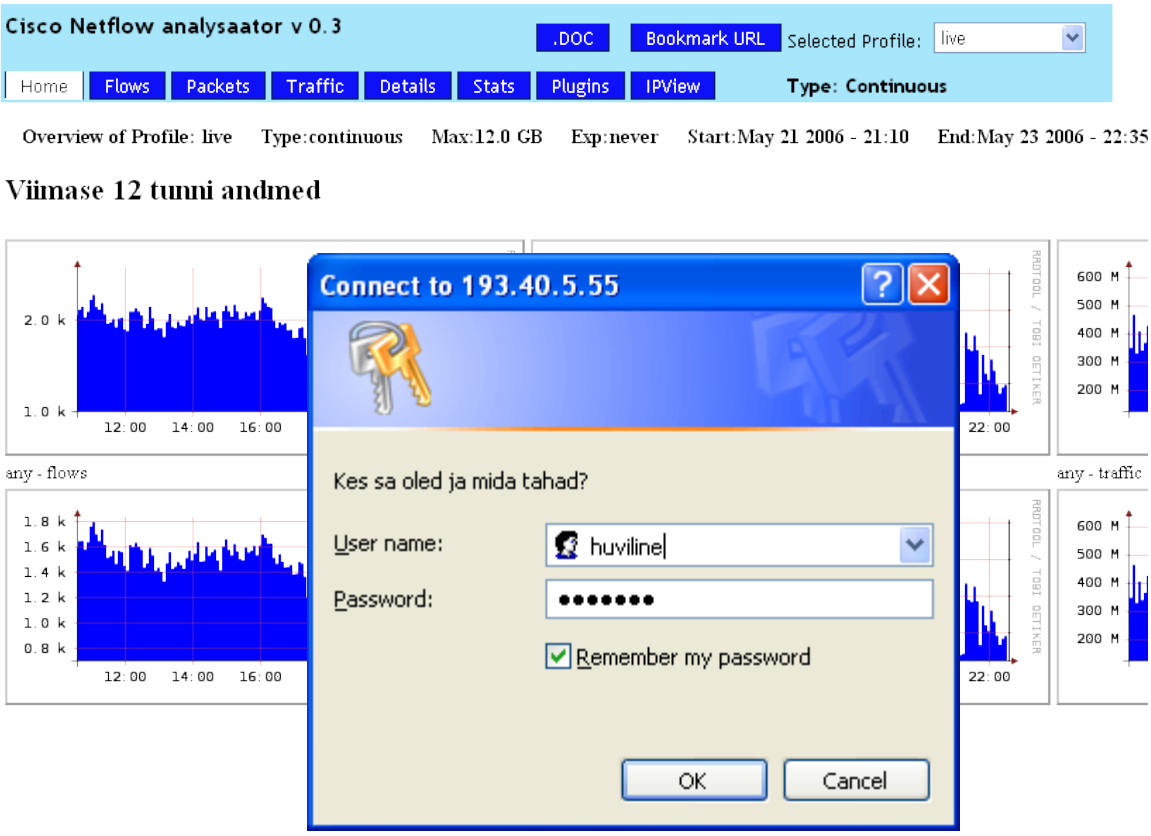
Kuigi NfSen ja NfDump tarkvarapaketid sisaldasid ja võimaldasid pea-aegu kõike olulisemat minu töö lähteülesandes kirja pandut, sai töö edasiseks peaülesandeks paketi veebiliidese ja alamprogrammiosa täiendamine.

Töö tulemust saab vaadata serveri (vt lk 19) veebiserveri vahendusel, aadressilt

http://193.40.5.55/nfsen/nfsen.php

kasutajanimi: huviline

parool: huviline



Joonis 6 Cisco NetFlow analüsaatori veebiliides

Loodud programmi struktuur leiab pikemalt käsitlemist käesoleva töö lisas 1, lk 37 , “Lisa 1 – Cisco NetFlow analüsaatori programmi struktuur”.

Üheks esimestest eesmärkidest oli luua ülevaatlik lehekülg, millelt on võimalik näha hetke võrgustatistikat ning liides, millega teostada otsinguid. Graafikutest, mida programm vaikimisi koostab jääb samuti väheseks, kuna ei peeta arvestust üksikute TCP / UDP teenuste kaupa ning puudub võimalus graafiku koostamiseks konkreetse hosti kohta.

Põhilised muudatused ja uued funktsioonid olid:

- Programmile tuli lisada TOP10 statistika võrgukasutuse, teenuste ja alamvõrkude kohta
- Kõik tabelid tuli teha “klikitavaks” astmelise filtreerimise jaoks
- Kõikide kuvatavate IP aadresside ja võrguteenuste lahendamine nimedeks
- Lisaks tabelitele veel täiendavad graafikute (sektorgraafikud) tekitamine
- Veebiliideses IP aadressi, pordi ja protokolli põhine filtreerimine
- Lisada statistika erinevate (saab muuta, milliste) IP portide , IP aadresside ja võrgusegmentide kohta

Töö käigus tuli täiendada ja juurde kirjutada koodi PHP keeles kuid osa tarkvarast (graafikute genereerimine ja veebiliidese lisamoodulid) tuli täiendada ka Perl keeles. Kuna programmi NfDump väljundiks on kas tekstifail või binaarfail, otsustasin andmevahetuse tabelite ja graafikute kuvamiseks teha läbi tekstifailide. Siinkohal aitas palju kaasa nii PERL kui ka PHP keelte puhul tugeval tasemel olevad tekstinfo töötlemise vahendid.
Programmeerimiseks vajaliku info saamisel oli peamiseks abivahenditeks interneti otsingumootorid (*google.com* , *altavista.com* jt) ja programmide kodulehed, lühitutvustused ja kasutusjuhendid (PHP programmeerimise kasutusjuhend [] , Perl programmeerimise kasutusjuhend []. jt).
Graafikute genereerimisel kasutasin NfSen paketis juba rakendust leidnud , Tobi Oetiker arendatud tarkvara RRDTool [,] vahendeid ning lisana Nico Puhlmann poolt programmeeritud PHP moodulit “PieGraph” [] , mida tuli programmi jaoks siiski veidi kohendada.
IP aadresside nimelahenduse realiseerisin PHP keele vahendite ning Linux käsurea programmi “host” abil, nimepäringute kiirendamiseks olin serverisse installeerinud DNS vahemälu serveri PDNSD [].
Et tagada kõigis kuvatavates tabelites ka TCP / UDP pordinumbrite kuvamine teenusenimedena, lisasin programmi koodi vastava tõlkimise alamfunktsiooni.

Funktsioon tagastab kindla formaadiga tekstifailist etteantud pordinumbri kohta kirjelduse, kui pordinumbri kohta info puudub, tagastab pordinumbri.
Otsingu nimekiri on muudetavas tekstifailis, mille koostamise aluseks võtsin firma Neohapsis poolt koostatud nimekirja [] , mida saab kasutada GPL litsentsi alusel vabalt.
Portide kohta sain usaldusväärseid andmeid ka IANA (*The Internet Corporation for Assigned Names and Numbers*) kodulehel [], kus oli võrreldes eelmise allikaga neid aga hulga vähem.

Edasised täiendused ja arengusuunad

Kuigi lahendus võimaldab kõiki ülesande püstitusel välja toodud olulisemaid eesmärke, on edasise täiendamise ja arendamise jaoks tööd veel palju.
Mõned arendamist vajavad teemad on:

- IPv6 liikluse jälgimise osa kontroll
- Reaalajas info kogumine marsruuterist
- Töö mitme sisendi korral
- Kogutud info väljastamine andmebaasi või faili
- Andmevahetus teiste võrguhaldusprogrammidega

Täiendavat uurimist vajaksid veel:

- Süsteemi skaleerimine. Mis saab, kui tegemist on suure ja keeruka arvutivõrguga, kus on palju (>10) Netflow infoallikaid?
- Info mitmekordse loendamise küsimused mitmetasemelises arvutivõrgus – kui sama pakett läbib mitut võrguseadet.
- Võimalus hajutada salvestusressurssi (serverite kettamassiivid üle võrgu)
- Suuremahuliste otsingute hajutamine / jagamine arvutivõrgu vahendusel mitmeks alamosaks.

Netflow analüsaatori veebiliidese kasutusjuhend

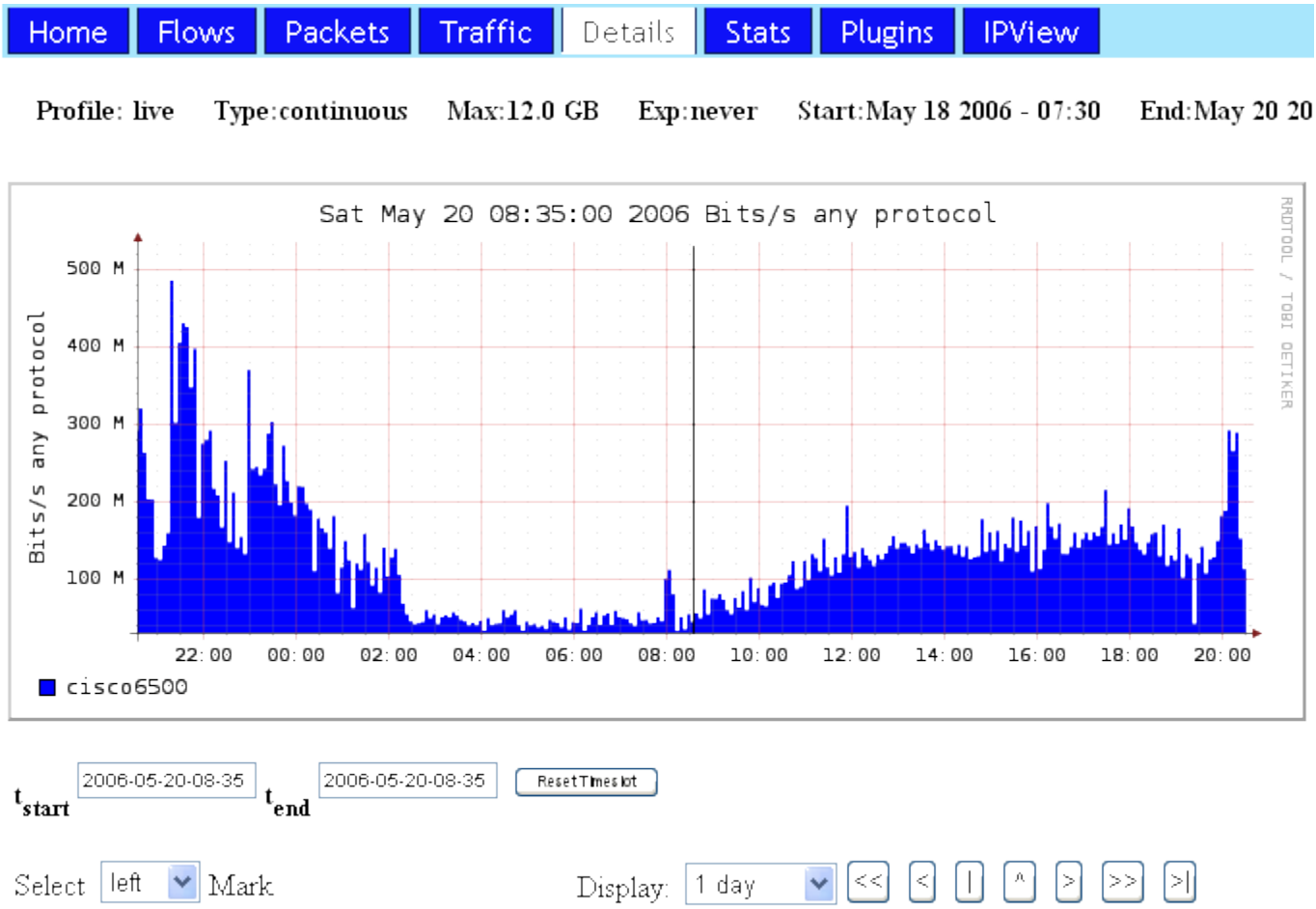
Netflow analüsaatori veebiliides põhineb NfSen tarkvarapaketi täiendustel ja lisadel, seetõttu on kasutatavad suurem osa paketi NfSen funktsionaalsust.
Veebiliidese töö on jagatud kaheksaks alamjaotiseks (vt Joonis 7)

Joonis 7 Cisco NetFlow analüsaatori valikuriba

Hetkel aktiivne alamjaotis on valikuribal kujutatud heledana.
Alamjaotiste detailne kirjeldus:

- “Home” – Annab ülevaate viimase 5 minuti statistikast
- “Flows” – Näitab graafikuid Netflow kirjete kohta
- “Packets” – Näitab graafikuid pakettide kohta
- “Traffic” – Näitab graafikuid võrguliikluse mahu kohta baitides
- “Details” – Võimaldab valida ajavahemikku, kuvab valitud ajavahemiku kohta statistika ning võimaldab kasutajal sisestada käsklusi NfDump programmile täitmiseks
- “Stats” – Kuvab andmed aktiivse profiili kohta
- “Plugins” – Lisaprogrammide leht, minu lahenduses on siin võimalik “pordidjms” nimelise lisamooduli abil vaadata võrguressursside kasutust üksikute TCP/UDP portide, hostide ja alamvõrkude kaupa.
- “IPView” – IP aadresside, IP protokolli ja TCP/UDP pordinumbri järgi info filtreerimine, võimalik muuta ajavahemikku (tekstiliselt) ja kuvatava info mahtu.

Alamjaotis “Details” võimaldab graafiliselt navigeerida olemasolevas Netflow infos, kuvada statistikat valitud ajahetke või ajavahemiku kohta ning esitada täitmiseks käske NfDump programmile võimalusterohke liidese abil.



Joonis 8 Detailse vaate aken, navigatsiooni instrumendid.

Joonis 8 on toodud vaate “Details” põhilised navigatsioonivõimalused, kõige olulisem nendest on (vasaku hiirenupu abil) klikitav graafik, millelt saab valida uuritava ajavahetke. Graafik saab kujutada aktiivse ajavahemiku kohta edastatavate baitide, pakettide või Netflow kirjete hulka.

Kui on aga soov uurida ajavahemikku, saab “*Select right Mark*” valiku abil järgmise hiireklikiga graafikule valida lõpp-ajahetke, alghetkeks jääb enne seda aktiivne olnud ajahetk.

Sisendkastide “$\mathbf{t_{start}}$” ja “$\mathbf{t_{end}}$” abil saab samuti uuritavat ajahetke või ajavahemikku muuta, peale uue ajavahemiku (formaat **aaaa-kk-pp-hh-mm** , kus **aaaa** on neljakohaline aasta number, **kk** on kahekohaline kuu number, **pp** on kahekohaline kuupäev , **hh** kahekohaline tunni number ning **mm** minutid) sisestamist Enter klahviga muutub uus ajahetk- või vahemik aktiivseks ning kuvatakse sellele vastav graafik. Nupp “**reset timeslot**” algväärtustab aktiivse ajahetke (ehk graafiku keskkohaväärtuse) praegusest kellaajast kuus tundi minevikku.

Tarkvarapakettide Nfdump ja NfSen lühiülevaade

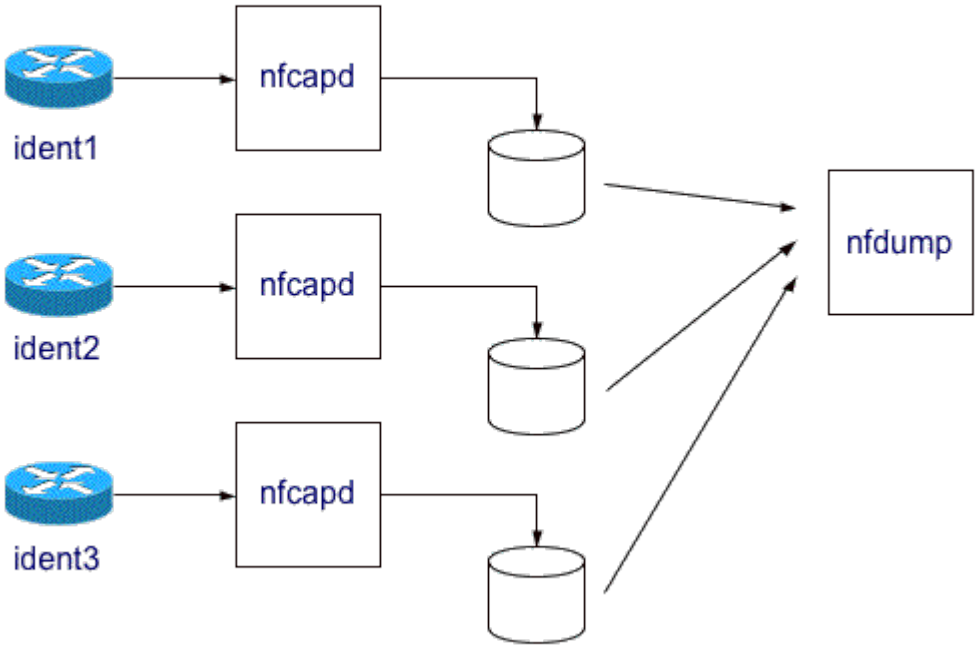
Programmid NfDump ja NfSen on välja töötatud Peeter Haag poolt , kes on ametis Sveitsi teaduse-ja hariduse andmesidevõrgu (*SWITCH*) turvaintsidentide rühma insenerina.

Programmi arendamise võttis Peeter Haag ette põhiliselt kolmel põhjusel []:

- Olemasolevad tarkvarad olid liiga aeglased - Oli vaja väga kiirelt töötavat programmi Netflow info töötlemiseks
- Netflow infot tuli väga-väga palju – tuleb organiseerida väga hea Netflow info hoidmise süsteem teatud perioodi jooksul
- Senised programmid olid liiga keerulised – tuli luua lihtne ja paindlik programm, mis võimaldaks filtreerida ja andmeid töödelda

Arvestades ülaltoodud nõudmisi valmiski programmipakett NfSen , NfDump, Nfcapd, kogu paketi esimene avalik versioon tuli välja 20.09.2004a.

Programm töötleb Cisco NetFlow andmeid kolmekihiliselt (vt Joonis 9)



Joonis 9 Andmete kogumine NfDump paketi abil

Marsruuteritest tuleva info võtab vastu nfcapd protsess, mis salvestab Netflow kirjed kettale failidesse (failinimi koostatakse kuupäeva, kellaaja ja protsessi identifikaatori kombinatsioonina kujul , *näiteks nfcapd.200605190300* ainult ühe protsessi korral).

NfDump programm võimaldab eelnevalt salvestatud informatsiooni töödelda , kasutades filtreerimist, summeerimist , sorteerimist ja ka anonümiseerimist. Väljundinfo võidakse salvestada tekstifailidesse või binaarfailina järgmise NfDump sessiooni sisendiks.
Nfcapd ja Nfdump tarkvara on kirjutatud C keeles, mis tagab andmete töötlemise kiiruse.
Paketti kuulub ka valik Perl keeles kirjutatud skripte andmeprofiilide töötlemiseks ja andmete haldamiseks (sh. kustutamine aegumistähtaja või lubatud maksimaalse kettamahu ületamise korral).

NFSen on graafiline veebiliides Nfdump netflow tarkvarale.

Joonis 10 NfSen lahenduse struktuuri ülevaade

Koos NfDump tarkvaraga moodustub terviklahendus Netflow informatsiooni haldusest, milles on olemas vahendid ka tegevuste automatiseerimiseks ja teadete edastamiseks.

Nfsen võimaldab:

- Kuvada NetFlow informatsiooni, kirjete, pakettide ja baitide kaupa sorteeritult, kasutades RRD (*Round Robin Database*) tarkvara.
- Hõlpsasti navigeerida andmete hulgas
- Töödelda Netflow infot (kasutaja saab ise ette anda parameetrid) valitud ajavahemikus
- Luua Netflow voogudest erinevaid vaateid, profiile, seda kõike ka eelnenud perioodide (ajaloo) kohta.
- Kirjutada ise lisamooduleid (i.k *plugins*), mida käivitatakse regulaarse ajavahemiku järel.
- Edastada informatsiooni e-posti teel
- Võimaldab anonümiseerida IP aadresse statistika publitseerimise tarvis.

Nfsen võimaldab läheneda Netflow andmetele läbi mitme erineva liidese – võimaldab kasutada nii käsurea eeliseid, kui annab ka graafilise ülevaate andmete kohta. NfDump ja NfSen tarkvarapaketid on saadaval *sourceforge.net* veebikeskkonna vahendusel [,] ning jagatakse BSD tarkvaralitsentsi alusel.

Kokkuvõte

Üheks tähtsamaks järelduseks minu töö tegemisel on tõdemus, et Cisco NetFlow protokoll on väga võimas vahend andmesidevõrgu liikluse jälgmisel. Netflow abil on võimalik teha nii lihtsaid süsteeme, mis peavad arvestust marsruuterist (või muust Netflow toetusega võrguseadmest) läbimineva infohulga üle , kui ka keerulisi haldussüsteeme, mis on võimelised suure täpsusega toimima ka väga suurte arvutivõrkude- ja sidemahtude korral.
Kõige olulisemaks Netflow omaduseks on tema infomahukus – kirjed sisaldavad pea-aegu kogu informatsiooni võrguseadet läbiva paketi kohta. See annab võimaluse Netflow kasutamiseks ka mitmesuguste võrgurünnakute tuvastamisel, näiteks teenusetõkestamine (*DoS*) ja usside ,troojade või viiruste rünnete puhul.
Töö alguses püstitatud eesmärgi (luua veebiliides Netflow info kogujale ja analüsaatorile) täitmine kulges üldjoontes edukalt.
Töö käigus tutvusin põhiliselt interneti vahendusel kuid mõne lahenduse puhul ka ise katsetades olemasolevate NetFlow võrguhaldus- ja jälgimislahendustega ning seadistasin Tartu Ülikooli IT osakonna poolt antud serverarvutile tarkvarapaketi “NfSen ja NfDump” põhjal arendatud Netflow analüsaatori tarkvara. Tarkvara täiendamine on võimalik tänud GNU GPL [] ja BSD [] tarkvaralitsentsile ning oma töö raames muutsin veebiliidese ülesehitust (ja sisu) ning programmeerisin ülesande püstitusele vastavad lisafunktsioonid ning täiendavad graafikute esitused.
Lahenduse väljatöötamise käigus selgusid ka võimalikud edasiarendamist vajavad teemad (töö suurte andmemahtudega, skaleermine, IPv6 liikluse osa arendamine).

Käesoleva töö kooseisus on vaadeldud põgusalt Cisco Netflow protokolli ülesehitust, olemasolevaid valdkonna lahendusi , esitatud kasutusjuhend loodud analüüsitarkvarale ning toodud tarkvarapaketi “NfSen” lühitutvustus.

Cisco NetFlow analyzer

Master thesis
Risto Rahu

Abstract

The present master’s thesis gives an basic overview of Cisco NetFlow [] protocol based network statisics collection and overview of a program named “Cisco Netflow analysaator” which is used in Taru University’s IT department with Cisco Catalyst 6506 switch.
Program “Cisco Netflow analysaator” is based on Peeter Haag’s program “NfSen” [] and uses many other tools (NfDump,RRDtool, libgd) to collect and process netflow data.
This work was initiated by University’s IT department, who needs to have system (software & hardware), which can collect and process large amount of Netflow traffic originating from campus network core router - Cisco Catalyst 6506.

Kasutatud viited, kirjandus

1. Tobi Oetiker's MRTG - The Multi Router Traffic Grapher
   http://oss.oetiker.ch/mrtg/

2. Tobi Oetiker's RRDtool -Logging & Graphing:
   http://oss.oetiker.ch/rrdtool/

3. RRD Tutorial:
   http://oss.oetiker.ch/rrdtool/tut/rrdtutorial.en.html

4. NfSen - Netflow Sensor:
   http://nfsen.sourceforge.net/

5. Peeter Haag - “NetFlowDUMP – NFDUMP”
   http://nfdump.sourceforge.net/

6. Paolo Lucente – “pmacct: a small set of IPv4/IPv6 accounting and aggregation tools”,
   http://www.ba.cnr.it/~paolo/pmacct/

7. Paolo Lucente , “pmacct, a new player in the network management arena “, 04.2006 , RIPE 52 meeting plenary session,
   http://www.pmacct.net/ripe52-plenary-pmacct.pdf

8. NTOP - a network traffic probe
   http://www.ntop.org/overview.html

9. Cisco IOS NetFlow:
   http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html

10. Cisco NetFlow Version 9 Flow-Record Format
    http://www.cisco.com/en/US/products/ps6601/products_white_paper09186a00800a3db9.shtml

11. Introduction to Cisco IOS NetFlow - A Technical Overview
    http://www.cisco.com/en/US/products/ps6601/products_white_paper0900aecd80406232.shtml

12. SFlow - Making the Network Visible
    http://sflow.org/

13. Cisco Catalyst 6506 Switch
    http://www.cisco.com/en/US/products/hw/switches/ps708/ps710/index.html

14. IANA port numbers
    http://www.iana.org/assignments/port-numbers

15. Neohapsis port list
    http://www.neohapsis.com/neolabs/neo-ports/neo-ports.html

16. Nico Puhlmann , “piegraph.class.php”
http://www.jaxoo.de/exec/hacks

17. Cisco IOS NetFlow White Papers
http://www.cisco.com/en/US/products/ps6601/prod_white_papers_list.html

18. SWITCH - Network Monitoring and Analysis software list
http://www.switch.ch/tf-tant/floma/software.html

19. Peeter Haag ettekanne RIPE (Réseaux IP Européens) 50ndal kokkusaamisel 03.05.2005a Stokholmis, “Watch your Flows with NfSen and NFDUMP”,
http://www.ripe.net/ripe/meetings/ripe-50/presentations/ripe50-plenary-tue-nfsen-nfdump.pdf

20. PHP online manual, http://www.php.net/manual/en/

21. PERL (*Practical Extraction and Report Language*) manual,
http://www.perl.com/doc/manual/html/pod/perl.html

22. PDNSD – “a proxy DNS server with permanent caching”,
http://www.phys.uu.nl/~rombouts/pdnsd/index.html

23. Cisco CS-MARS: Cisco Security Monitoring, Analysis and Response System
http://www.cisco.com/go/mars

24. GNU General Public License (versioon 2)
http://www.gnu.org/copyleft/gpl.html

25. The BSD License template
http://www.opensource.org/licenses/bsd-license.php

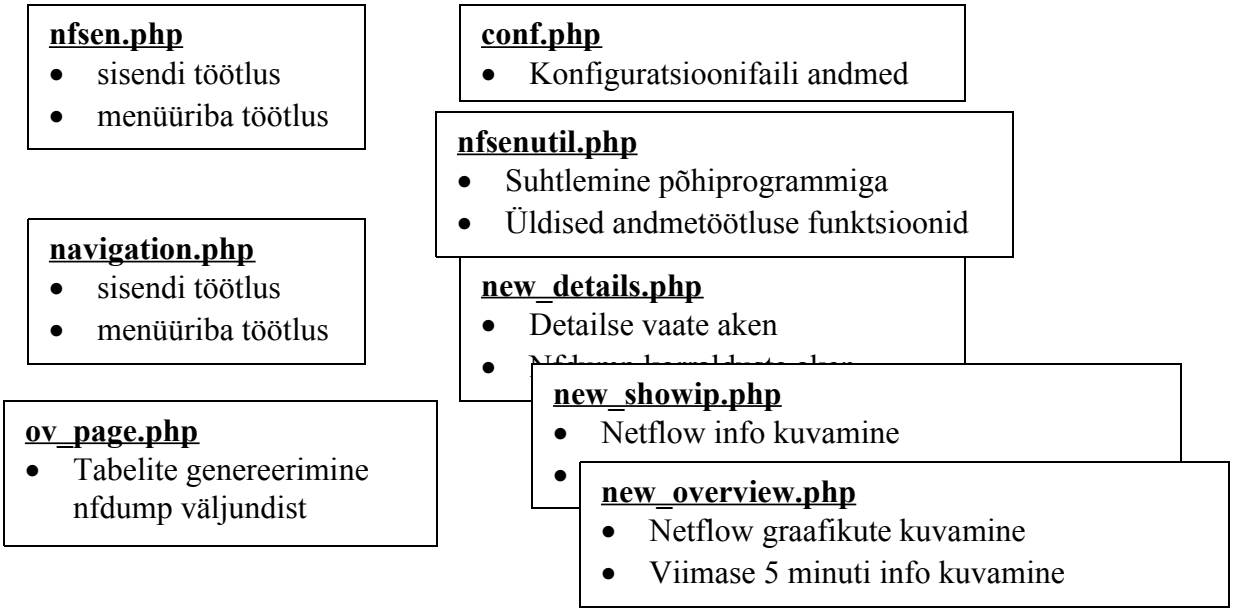
26. ™

Lisa 1 – Cisco NetFlow analüsaatori programmi struktuur

Programm koosneb kahest osast:

- veebiliides, mis suhtleb kasutajaga
- lisaprogrammid NfSen põhimootori juures, mis töötavad taustal ning töötlevad automaatselt sisendinfot.

Veebiliides on kirjutatud PHP keeles, lisaprogrammid PERL keeles.



Joonis 11 Cisco Netflow analüsaatori veebiliidese struktuur

Ncsnocdisndcinonsdiniosdniosd

Ndsnlkclknknlknlsdknlsdknlknlcsnklcs

Lisa 2 – Ülevaade CS-MARS lahendusest

Toode CS-MARS [] (*Cisco Security Monitoring, Analysis and Response System*) on Firma Cisco Systems Inc poolt välja töötatud valmissüsteem andmesidevõrgu turvakeskuse jaoks. CS-MARS tagab ülevaate ning võimalused turvaintsidentidega tegeleda ka väga suure andmesidevõrgu korral.

Joonis 12 CS-MARS GC välisvaade

Cisco CS-MARS lahendust on saadaval kuues eri versioonis:
CS-Mars 20, CS-Mars 50, CS-Mars 100e, CS-Mars 100, CS-Mars 200 ja CS-Mars GC (Global Controller).
Eelistena võib välja tuua:

- RAID 1+0 salvestusmassiiv
- Integreeritud andmebaasirakendus (Oracle)
- Ei vaja täiendavaid infokogumisseadmeid
- Netflow toetus
- Võrgutopoloogia tuvastamine, L2 ja L3 tasemel.
- Sündmuseid võimalik detialiseerida kuni MAC aadressi tasemeni

Erinevate mudelite omaduste võrdlus on toodud Tabel 9.

Tabel 9 CS-MARS erinevate mudelite võrdlus:

CS-MARS Mudel:	**20**	**50**	**100e**	**100**	**200**	**Global Controller**

Sündmusi/sek	500	1000	3000	5000	10000	-
Netflow kirjeid/sek	15000	25000	75000	150000	300000	-
RAID salvestusmaht	120GB	120GB	750GB	750GB	1TB	1TB
Kõrgus seadmekapis	1 RU	1 RU	3 RU	3 RU	4 RU	4 RU

Ncsnocdisndcinonsdiniosdniosd

Ndsnlkclknknlknlsdknlsdknlknlcsnklcs