

Anycast

Lubor Jurena
CEO
jurena@skhosting.eu



Michal Kolárik
System Administrator
kolarik@skhosting.eu

O nás

- Registrátor
- Webhosting
- Serverové riešenia
- Správa infraštruktúry
- Všetko sa dá :-)

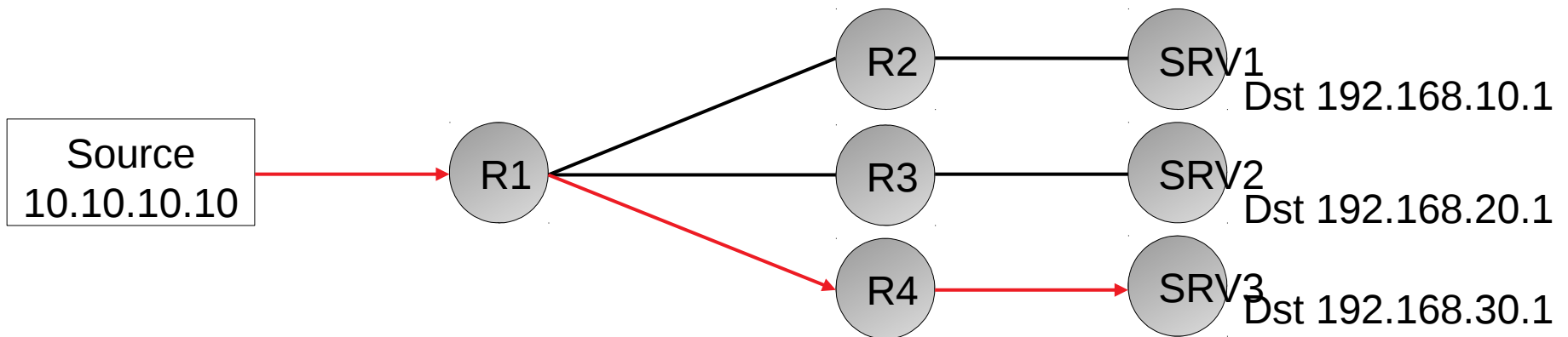
Index

- Čo je to anycast?
- Ako prevádzkovať vlastnú anycast sieť?
- Problémy – technické, právne
- Monitoring

Čo je to anycast?

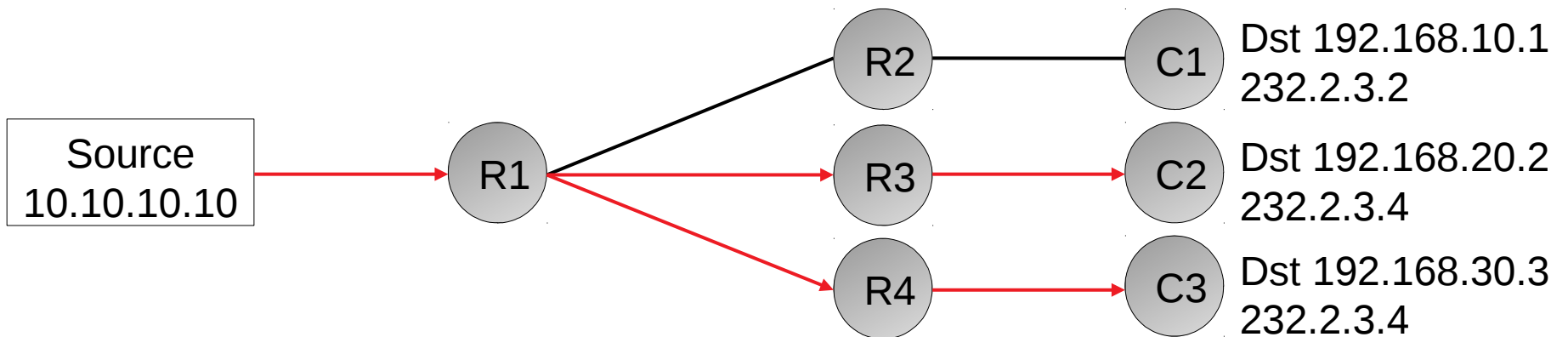
Anycast nie je ...

- Unicast
- Jeden cieľový server



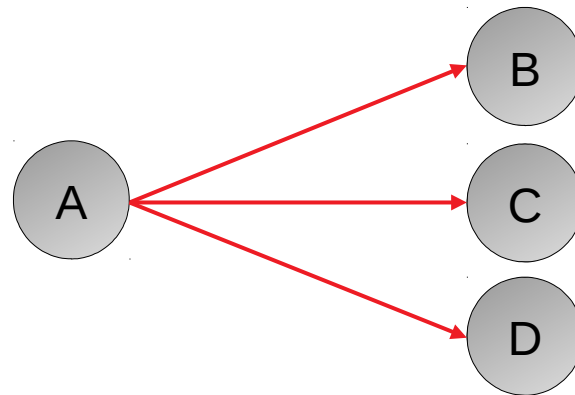
Anycast nie je ...

- Multicast
- Jeden zdroj a skupina príjemcov
- IPv4 224.0.0.0/4



Anycast nie je ...

- Broadcast
- Napr. ARP



Anycast nie je ...

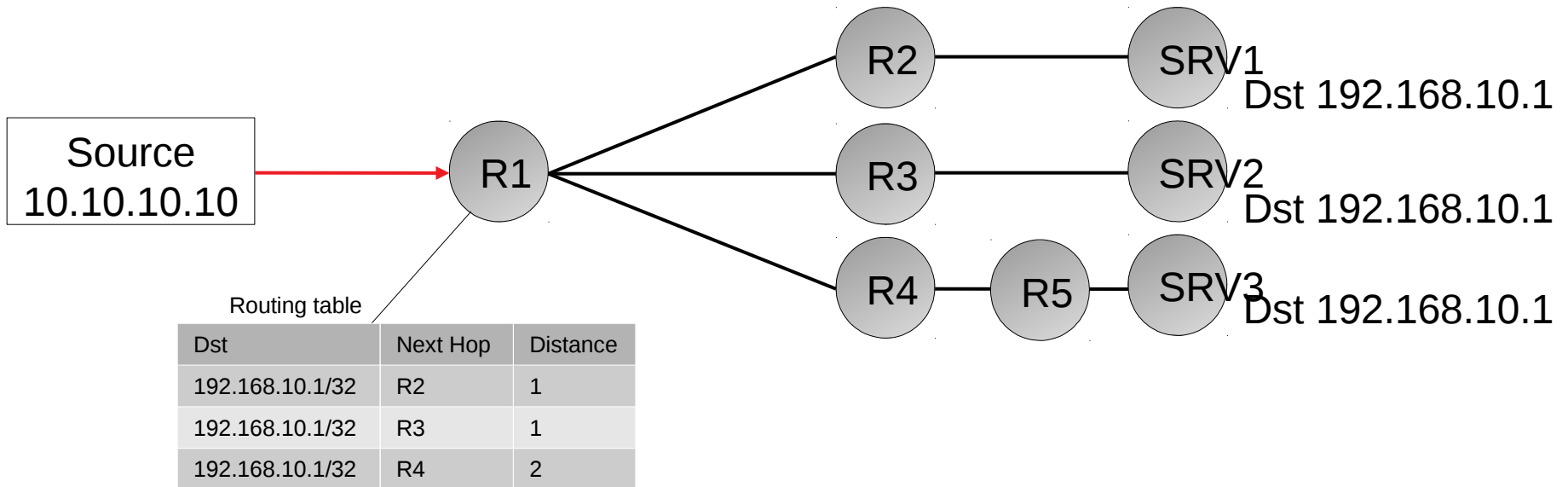
- Technológia
- Protokol
- Nevyžaduje špeciálne technológie / zariadenia
- ...

Anycast JE ...

- RFC 1546: „Host Anycasting Service“
- AΩ koreňovej DNS infraštruktúry
- Všetky servery vystupujú pod jednou IP adresou
- O smerovaní rozhoduje smerovací protokol

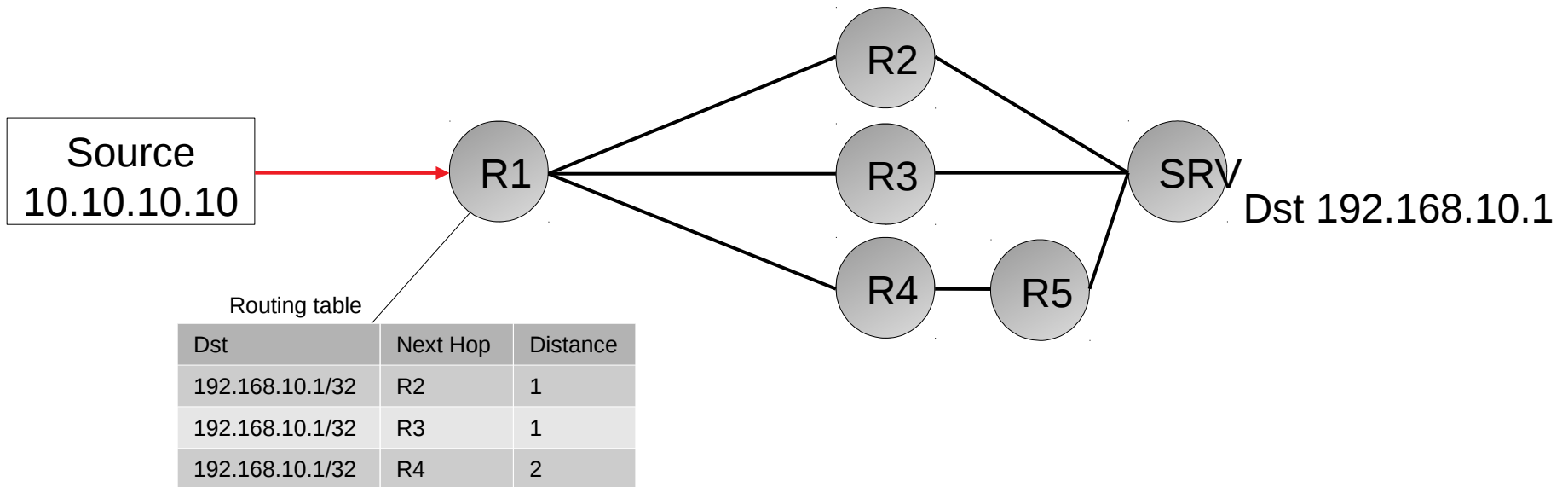
Anycast JE ...

- Layer 3 routing



Anycast JE ...

Ako to vidí smerovač?



AMS -> F Root

- lubor@ams:~\$ traceroute f.root-servers.net
- traceroute to f.root-servers.net (**192.5.5.241**), 30 hops max, 60 byte packets
- 1 * * *
- 2 108.61.99.1 (108.61.99.1) 0.322 ms 0.545 ms 0.402 ms
- 3 ams-ix.as13335.net (80.249.211.140) 1.044 ms 1.037 ms 1.056 ms
- 4 f.root-servers.net (192.5.5.241) **0.904 ms 0.921 ms 0.908 ms**
- lubor@ams:~\$ dig hostname.bind @f.root-servers.net chaos txt
- ...
- hostname.bind. 60 CH TXT "AMS.cf.f.root-servers.org"
- ...

SAO -> F Root

- lubor@sao:~\$ traceroute f.root-servers.net
- traceroute to f.root-servers.net (**192.5.5.241**), 30 hops max, 60 byte packets
- 1 191.96.70.1 (191.96.70.1) 0.266 ms 0.162 ms 0.220 ms
- 2 177-234-154-33.static.hostdime.com (177.234.154.33) 0.173 ms 0.169 ms 0.241 ms
- 3 as30122.saopaulo.sp.ix.br (187.16.217.105) 0.680 ms 177-234-156-254.static.hostdime.com (177.234.156.254) 0.204 ms as30122.saopaulo.sp.ix.br (187.16.217.105) 0.673 ms
- 4 f.root-servers.net (192.5.5.241) **0.664 ms 0.662 ms 0.654 ms**
- lubor@sao:~\$ dig hostname.bind @f.root-servers.net chaos txt
- ...
- hostname.bind. 60 CH TXT "gru1a.f.root-servers.org"
- ...

BTS -> F Root

- `lubor@bts:~$ traceroute f.root-servers.net`
traceroute to f.root-servers.net (**192.5.5.241**), 30 hops max, 60 byte packets
- 1 ba-gw1-vl300.skhosting.eu (185.66.200.254) 0.660 ms 0.774 ms 0.872 ms
- 2 46.229.239.89.host.vnet.sk (46.229.239.89) 0.708 ms 0.775 ms 0.889 ms
- 3 109.74.147.188.host.vnet.sk (109.74.147.188) 47.392 ms 47.377 ms 47.386 ms
- 4 nixsk1.cloudflare.com (194.30.187.171) 2.093 ms 2.078 ms 2.064 ms
- 5 f.root-servers.net (192.5.5.241) **1.952 ms 1.902 ms 1.892 ms**
- `lubor@bts:~$ dig hostname.bind @f.root-servers.net chaos txt`
- ...
- `hostname.bind. 60 CH TXT "VIE.cf.f.root-servers.org"`
- ...

Dôvody pre anycast

- Fail-over
- Zlepšenie odozvy
- Zvýšenie robustnosti
 - Load balancing
- Ochrana pred útokmi

Ako nasadiť anycast?

- IP adresy
- Konfigurácia anycast servera
- Routing
 - Quagga, BIRD, ...

IP Adresy

- Minimálne 2 IP adresy
 - 1 IP anycastová
 - 1 IP unicastová
- Anycastovú IP spravidla nasadíme na loopback
- `ifconfig lo0:1 192.168.10.1/24`

Konfigurácia anycast servera

- Upravenie firewall-u
- Pripravenie služieb na anycast
- Synchronizácia všetkých anycast serverov
 - Centrálna správa konfigurácie
 - Rovnaký obsah vo všetkých lokalitách

Routing

- ~~Statický smerovací protokol~~
- Dynamický smerovací protokol
 - BGP

Routing

▪ bird.conf

```
filter ipout {
    if net = 192.168.10.0/24 then accept;
}
filter ipin {
    if net = 0.0.0.0/0 then accept;
}
router id 192.168.10.1;
protocol direct {
    interface "loopback0";
}
protocol kernel {
    persist;
    scan time 20;
    import all;
    export all;
}
protocol device {
    scan time 10;
}
protocol bgp anycast {
    source address 192.168.1.1;
    export filter ipout;
    import filter ipin;
    graceful restart on;
    local as 65001;
```

Problémy pri prevádzke

Technické

- Dodávateľia
 - „So why not using our IPs ? will this affects your service ???“
 - ISP pripojení na druhom konci sveta
- BGP algoritmus
 - Najkratšie neznamená najrýchlejšie
- Ktorý server má problém?

Právne

- Štátne obmedzenia
- Čo je zakázané v USA nemusí byť v EU
- Middle East

Monitoring

Prečo monitorovať?

- Potreba byť včas informovaný o výpadku servera a služby
- Zníženie reakčného času v prípade problémov
- História dát o serveri

Ako monitorovať anycast?

- Monitorujeme každý server zvlášť (služby, prostriedky)
- To nestačí, potrebujeme vedieť či sú jednotlivé serveri vo svete dostupné a ako sú dostupné.
- Potrebujeme monitorovať každý náš server z rôznych lokalít.
- Monitorovanie odozvy z rôznych lokalít na našu anycast sieť.

Možnosti monitorovania anycastu

- Platené služby monitorovania CDN sietí
- RIPE Atlas sondy umiestnené v lokalitách po svete, umožňujúce spúšťať merania za kredit.

Prečo RIPE Atlas sondy?

- Cena
- Možnosť výberu konkrétnej sondy s približnou polohou
- Viacero možností pre meranie
- Dobrá dostupnosť (nie všade rovnaká)

Ako vybrať sondy?

- RIPE Atlas používa 2 typy:
 - Anchors
 - Probes

Probes and Anchors

Probes

There are thousands of probes in the RIPE Atlas network. Probes are small, USB-powered hardware devices that hosts connect to an Ethernet port on their router. They perform active measurements, including ping, traceroute, DNS, SSL/TLS and NTP. The data they collect is aggregated and made publicly available.



[Learn More](#)

Anchors



[Learn More](#)

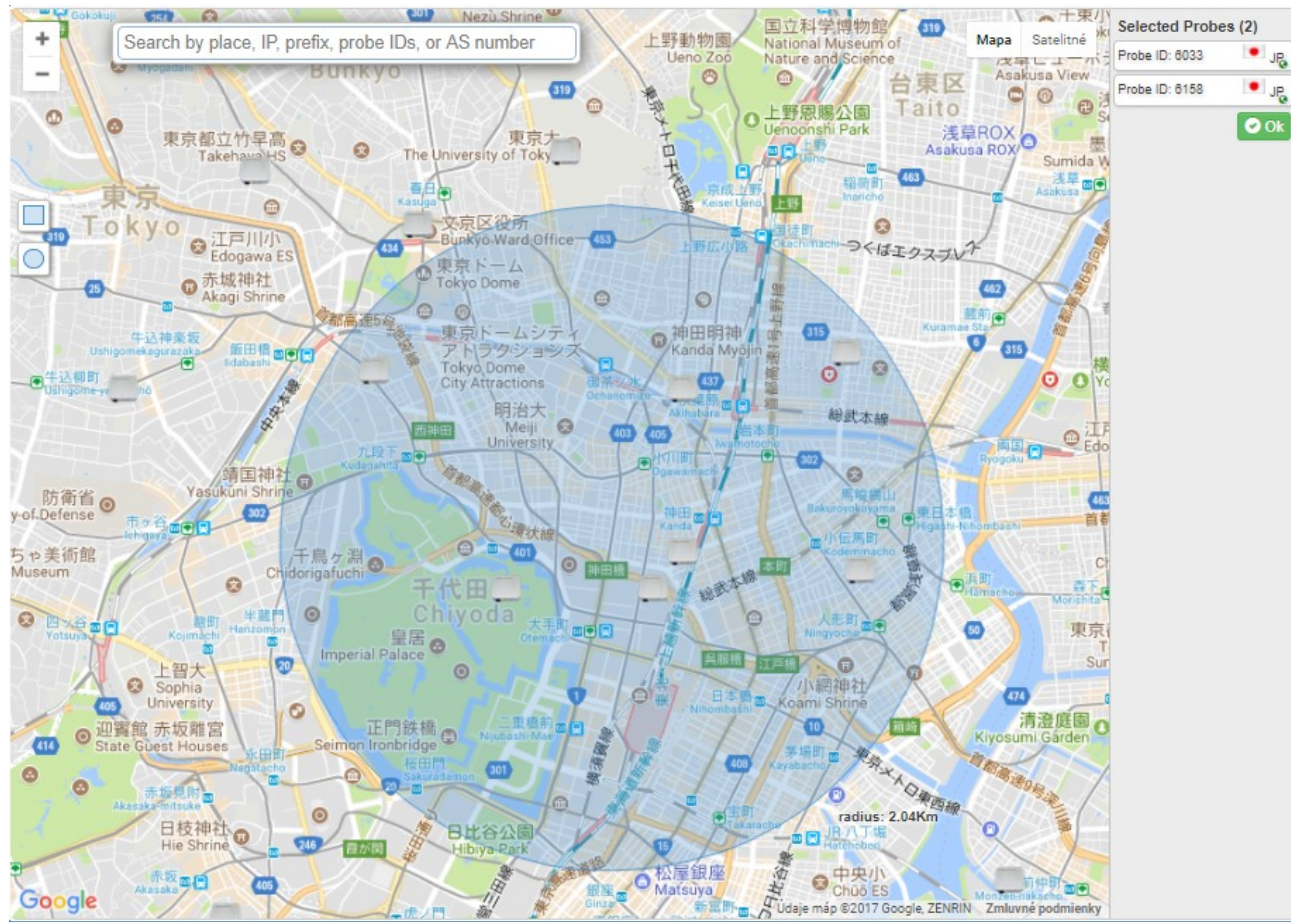
RIPE Atlas anchors are both enhanced RIPE Atlas probes with more measurement capacity, as well as regional measurement targets within the greater RIPE Atlas network. As such, RIPE Atlas anchors provide valuable information about the local and regional connectivity and reachability of the Internet.

Ako vybrať sondy? ...

- Systémové tagy:
 - Anchor
 - IPv4 works
 - IPv4 capable
 - IPv4 60d



Ako vybrat' sondy ? ...



Aký typ merania použiť?

- Traceroute – pre výber vhodných sond (3 sondy)

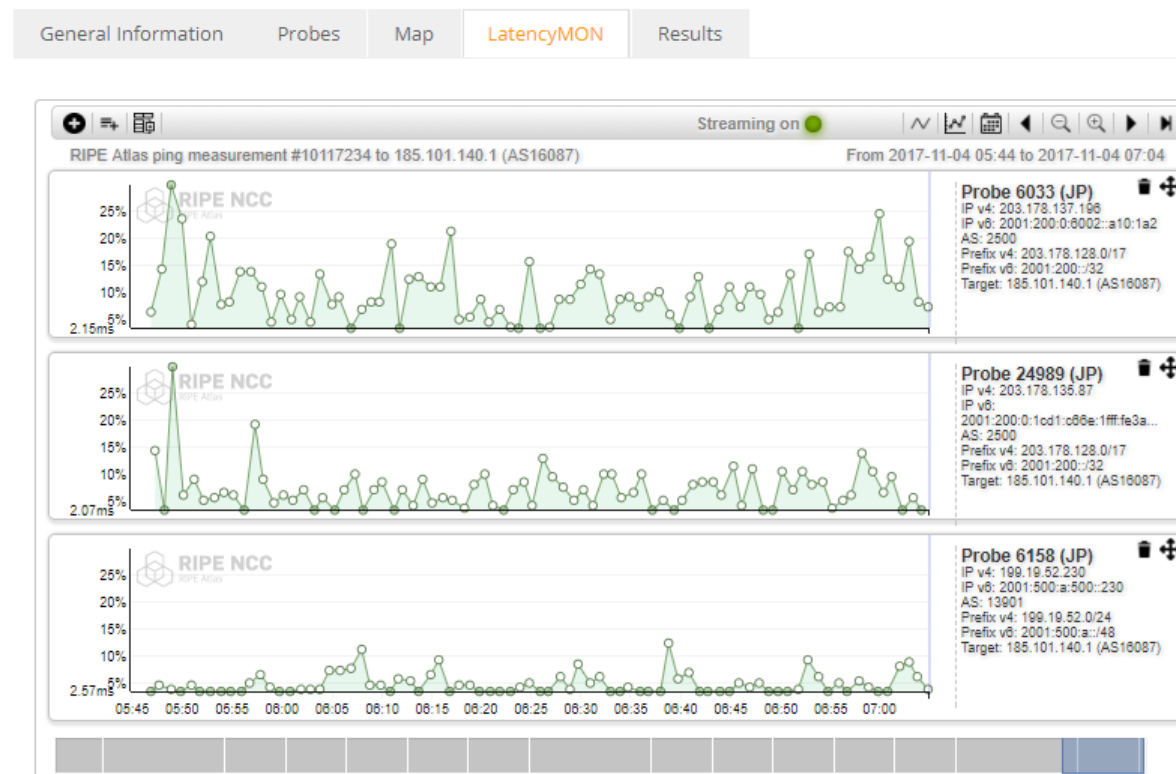
⚡ ANYCAST test Japan 185.101.140.1 (AS16087)

General Information	Probes	Map	TraceMON (beta)	Openipmap (beta)	Results				
Probe	ASN (IPv4)	ASN (IPv6)		Time (UTC)	RTT	Hops	Success		
234	2497	55391	🔴 🌐	2017-11-02 10:45	75.140		13	✓	🔍
2277	2514		🔴 🌐	2017-11-02 10:45	165.288		23	✓	🔍
4754	17506		🔴 🌐	2017-11-02 10:45	108.340		15	✓	🔍
6033	2500	2500	🔴 🌐	2017-11-02 10:45	2.341		12	✓	🔍
6158	13901	13901	🔴 🌐	2017-11-02 10:45	2.713		8	✓	🔍
11267	2497	4704	🔴 🌐	2017-11-02 10:45	79.438		13	✓	🔍
12790	2514		🔴 🌐	2017-11-02 10:45	175.122		24	✓	🔍
14393	2510	2516	🔴 🌐	2017-11-02 10:45	9.094		14	✓	🔍
14810	4713		🔴 🌐	2017-11-02 10:45	4.677		17	✓	🔍
24310	2907		🔴 🌐	2017-11-02 10:45	3.514		14	✓	🔍
24989	2500	2500	🔴 🌐	2017-11-02 10:45	2.539		13	✓	🔍
26698	2527		🔴 🌐	2017-11-02 10:45	9.444		13	✓	🔍
28352	17506		🔴 🌐	2017-11-02 10:45	107.245		16	✓	🔍
32335	2527	2527	🔴 🌐	2017-11-02 10:45	3.287		12	✓	🔍

Aký typ merania použiť? ...

- Ping – meranie odozvy na server

*⁴ ANYCAST Japan 185.101.140.1 (AS16087)



Ako spracovať dáta z merania ?

- Spracovanie status-check dát
- Predspracované dáta z ping monitoringu
- Každú minútu - update

```
{  
  "total_alerts": 0,  
  "global_alert": false,  
  "probes": {  
    "6033": {  
      "source": "Probes: 6033,6158,24989",  
      "last_packet_loss": 0,  
      "last": 2.585,  
      "alert": false  
    },  
    "6158": {  
      "source": "Probes: 6033,6158,24989",  
      "last_packet_loss": 0,  
      "last": 2.915,  
      "alert": false  
    },  
    "24989": {  
      "source": "Probes: 6033,6158,24989",  
      "last_packet_loss": 0,  
      "last": 2.405,  
      "alert": false  
    }  
  }  
}
```

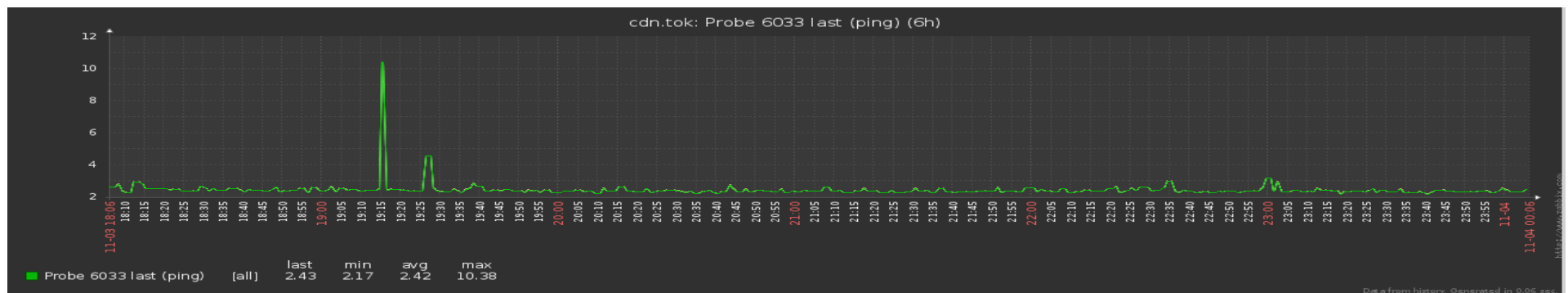
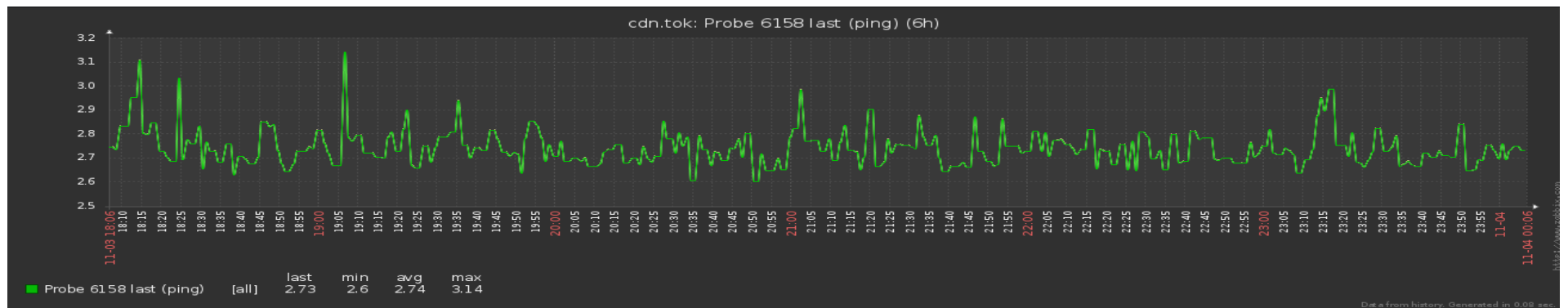
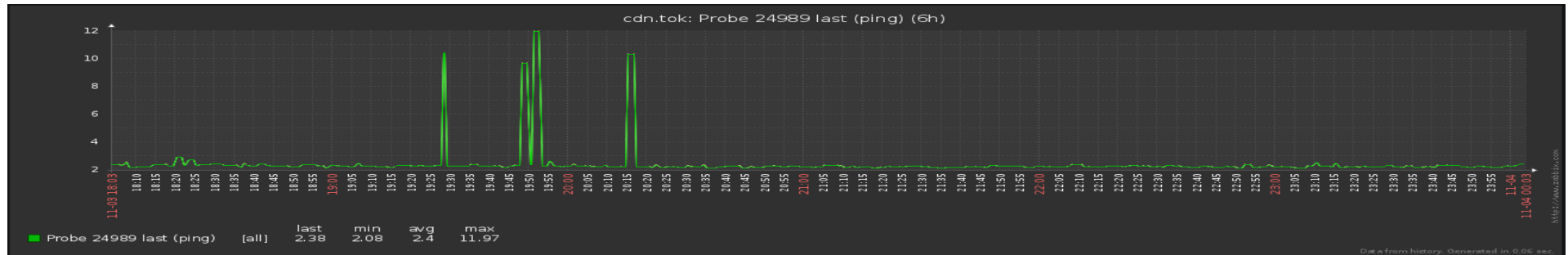
Spracovanie merani do Zabbix-u

Name	Last check	Last value	Change	
cdn-zabbix (9 Items)				
Probe 24989 last_packet_loss	2017-11-04 00:04:02	0		Graph
Probe 24989 last (ping)	2017-11-04 00:04:02	2.38		Graph
Probe 24989 alert	2017-11-04 00:04:02	False		History
Probe 6158 last_packet_loss	2017-11-04 00:04:02	0		Graph
Probe 6158 last (ping)	2017-11-04 00:04:02	2.75	+0.01	Graph
Probe 6158 alert	2017-11-04 00:04:02	False		History
Probe 6033 last_packet_loss	2017-11-04 00:04:02	0		Graph
Probe 6033 last (ping)	2017-11-04 00:04:02	2.32		Graph
Probe 6033 alert	2017-11-04 00:04:02	False		History

Kedy informovať o probléme ?

- Pri zvýšení odozvy nad 50ms v časovom rozhraní 2 minút
- Ďalšie možnosti:
 - možnosť informovania o “ALERTE” z RIPE (nemusí byť vhodné pre anycast)
 - Možnosť informovania v prípade problému súčasne na viacerých alebo všetkých sondách

Príklad nasadenia



Ďakujeme za pozornosť

Ľubor Jurena
CEO
jurena@skhosting.eu



Michal Kolárik
System Administrator
kolarik@skhosting.eu