



POVEZANI ZNANJEM **(...25 GODINA)**

19. CARNetova korisnička konferencija CUC 2017
8. – 10. studenoga 2017. / Hotel Valamar Lacroma

Sadržaj ovog materijala isključiva je odgovornost
Hrvatske akademske i istraživačke mreže - CARNet.

Nova struktura podataka u nadzornom sustavu CARNeta



Marijo Tolić
mario.tolic@carnet.hr

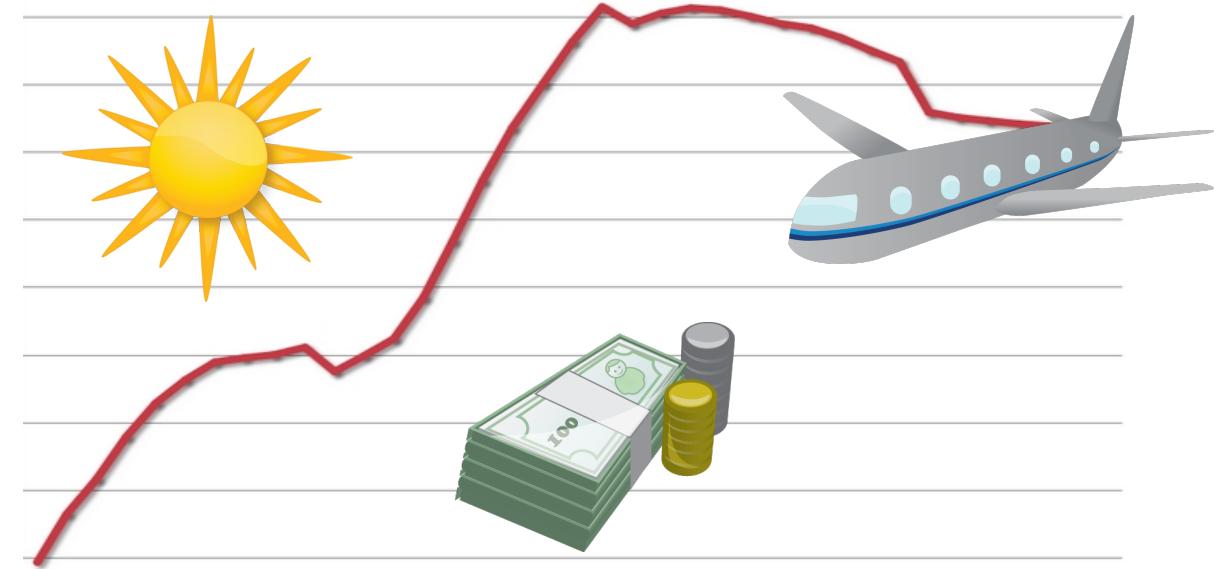
Sadržaj

- Time Series Data
- Time Series Databases
- Primjena u CARNetu
- RRD (Round Robin Database)
- Zašto InfluxDB?
- Implementacija
- Struktura InfluxDB podataka
- Primjena podataka



Time Series Data

- niz podataka indeksiranih u vremenskom slijedu
 - podatak sadrži vremensku oznaku
 - najčešće se podaci uzimaju u jednakim vremenskim razmacima
- prisutni svuda oko nas:
 - meteorologija
 - financije
 - crna kutija
 - IoT (The Internet of things)

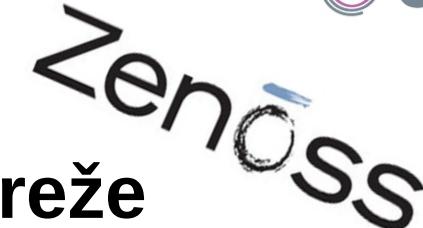


Time Series Databases

- baza za pohranu vremenski označenih podataka
- dijelimo ih na dvije osnovne skupine primjene:
 - zanima nas samo nedavna prošlost prikupljenih podataka
 - zanima nas veća količina podataka koju ćemo duže čuvati
- open source baze podataka:
 - **RRD**
 - Whisper/Graphit
 - openTSDB
 - KairosDB
 - Blueflood
 - Prometheus
 - **InfluxDB**
 - Elasticsearch
 - Btrdb
 - Riak TS



Primjena u CARNetu



• Zenoss (RRD)

- besplatan i open source alat za **nadzor i analizu mreže**
- za pohranu i vizualizaciju podataka koristi **RRD**
- praćenje dostupnosti, servisa i performansi uređaja

• Grafana (InfluxDB)

- besplatan i open source alat za **analizu i vizualizaciju podataka**
- za izvor podataka koristi InfluxDB
- ostale izvore koje podržava:

- Prometheus
- MySQL
- Postgres
- Graphite
- Elasticsearch
- CloudWatch
- OpenTSDB



RRD (Round Robin Database)

- kružna baza podataka
- **prednosti:**
 - fiksna veličina arhive podataka
 - moćan alat za crtanje grafova
 - lako održavanje

	Koraka	Sekundi	Minuta	Sati	Dana	
5 min prosjek	1	1800000	3000	50	2.08	> 2 dana
30 min prosjek	6	1260000	21000	360	14.58	> 2 tjedna
120 min prosjek	24	5580000	93000	1550	64.58	> 2 mjeseca
24 h prosjek	288	1147680	1147680	19128	797	> 2 godine

- **nedostaci:**
 - vrijeme čuvanja podataka
 - rezolucije podataka
 - brzina prikaza i dohvatanja pohranjenih podataka



Zašto InfluxDB?

Tražili smo dugoročnu pohranu podataka bez gubitka rezolucije:

- može se nositi s velikim i učestalim zapisima i dohvatom podataka
- **rezolucija podataka** ne ovisi o učestalosti zapisa podataka
- korisnik definira željeni period čuvanja podataka
- jezik za dohvat podataka vrlo sličan **SQL** upitima
- moguće zapisivanje pojedinačnih i skupnih podatka
- **HTTP API** za zapisivanje i dohvaćanje podataka



Implementacija u CARNetu



- iskoristili postojeću strukturu podataka koju koristimo u nadzoru mreže
- izbjegli nepotrebno dodatno opterećivanje uređaja SNMP upitima
- podatke čitamo direktno iz Zenoss RRD datoteka

Zašto ovakvo rješenje?

- trenutno nam je zadovoljavajući 5 minutni interval prikupljanja podataka
 - **Tijek izvođenje jednog intervala pohrane podataka:**
 - **izvoz** identifikatora podataka koje želimo prikupljati
 - **dohvat** podataka iz RRD-a prema izvezenim podatcima
 - **priprema** za zapis dohvaćenih podataka
 - **zapisivanje** podataka u InfluxDB



Izvoz:

- naziv uređaja
- tip uređaja
- uređaj - port
- opis porta
- **svaka 3 h**

Priprema:

- prikupljeni podaci se pripremaju u grupirane txt datoteke
- preporuka je da pojedina datoteka ne sadrži **vise od 5000** podataka
- **svakih 5 min**

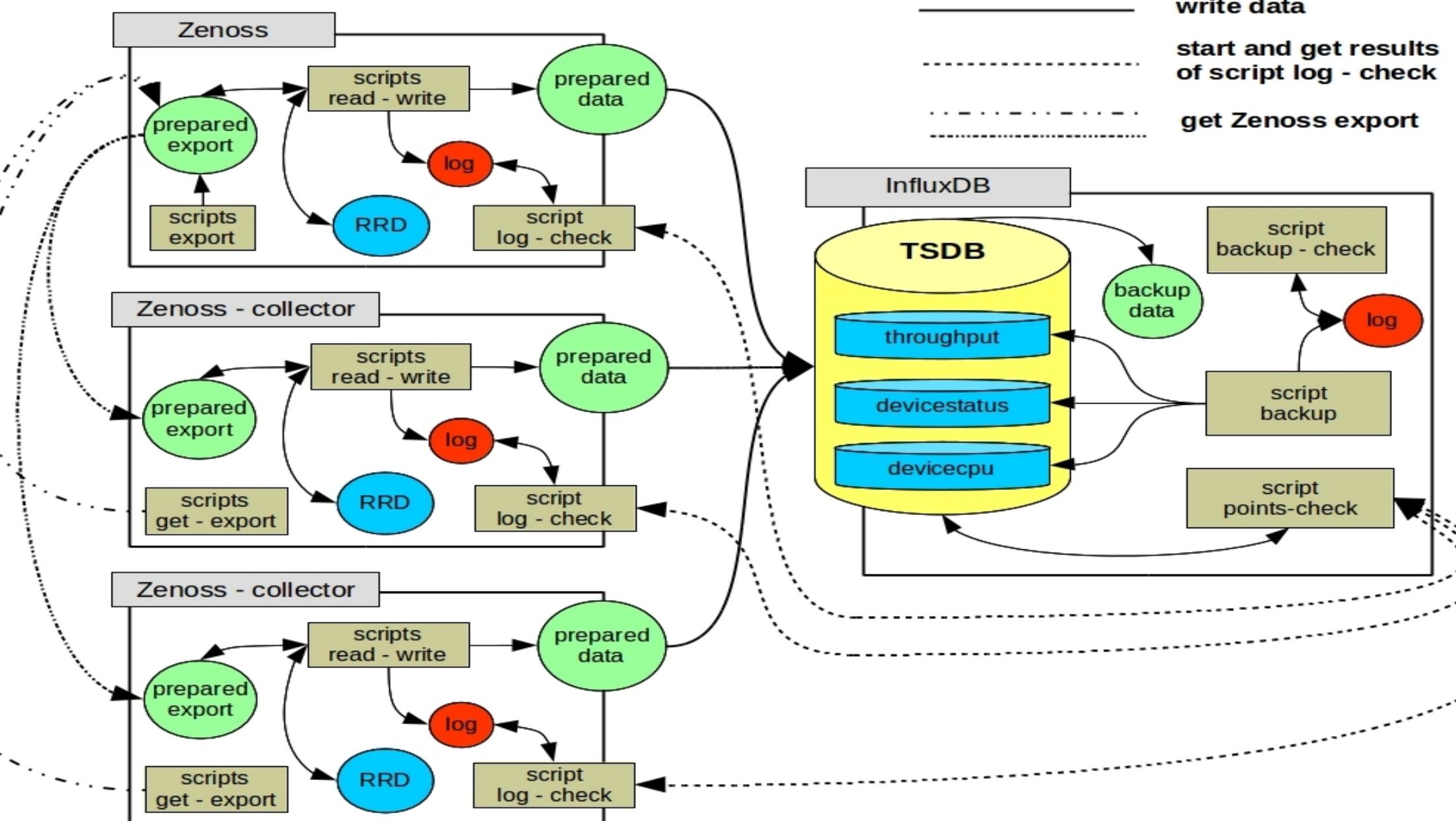
Dohvaćanje:

- na osnovu izvezenih podataka vadimo zadnju zapisanu vrijednost iz RRD datoteke
- **svakih 5 min**



Zapisivanje:

- koristeći HTTP API predajemo svaku pojedinu datoteku
- **svakih 5 min**



Struktura InfluxDB podataka

Točka mjerenja je struktura koju sastavljamo od željenih **tagova**, njihovih vrijednosti, **podatka** koji želimo zapisati i **vremenske** oznake

```
<measurement>[,<tag-key>=<tag-value>...]<field-key>=<field-value>[,<field2-key>=<field2-value>...] [unix-nano-timestamp]
```

- mogućnost dodavanja **novih tagova** u točku mjerenja
- inicijalno dodan **tip podatka** u točku mjerenja se ne može mijenjati
- **vremenska oznaka** je opcionalni atribut
- važno dobro planiranje tagova:
 - lakši dohvrat željenih podataka
 - učinkovitiji prikaz podataka
 - bolja iskoristivost podataka u kasnijim analizama



izostavljanjem se
upisuje trenutak
zapisivanja u bazu

Struktura podataka - primjer



```
InteOctInOut,device=CN-ACC-BORONGAJ-01,  
interface=TenGigabitEthernet0_1,direction=IN  
value=151547354533117 1478773512000000000
```

- **measurement:** InteOctInOut
- **tag-key:** device, tag-value: CN-ACC-BORONGAJ-01
- **tag-key:** interface, tag-value: TenGigabitEthernet0_1
- **tag-key:** direction, tag-value: IN
- **tag-key:** NaN, tag-value: NaN
- **field-key:** value, field-value: 151547354533117
- **unix-nano-timestamp:** 1478773512000000000



Primjena podataka u CARNetu



- prikupljamo podatke:
 - opterećenja **CPU** uređaja
 - **dostupnost** uređaja
 - **propusnost** portova

Zapisujemo **27897** točaka
mjerjenja svakih **300 sec**
(05.11.17.)

- prikaz podataka: **Grafana**
 - lako pretraživi **Dashboardi** zahvaljujući definiranim tagovima

- **periodični izvještaji**, skripte
 - tjedni, mjesecni, godišnji i na zahtjev
- **dugoročno praćenje trendova prometa**
 - bolje planiranje budućih “last mile” i “backbone” kapaciteta

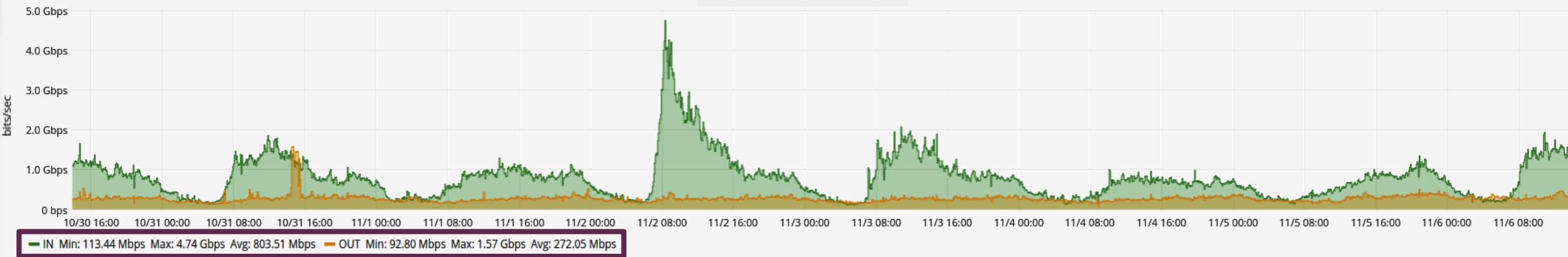


Host Port

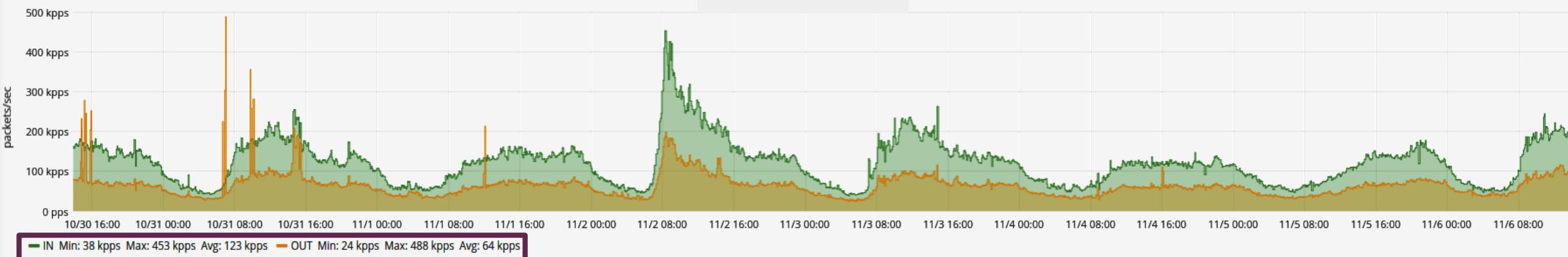
▼ Description:

opis porta

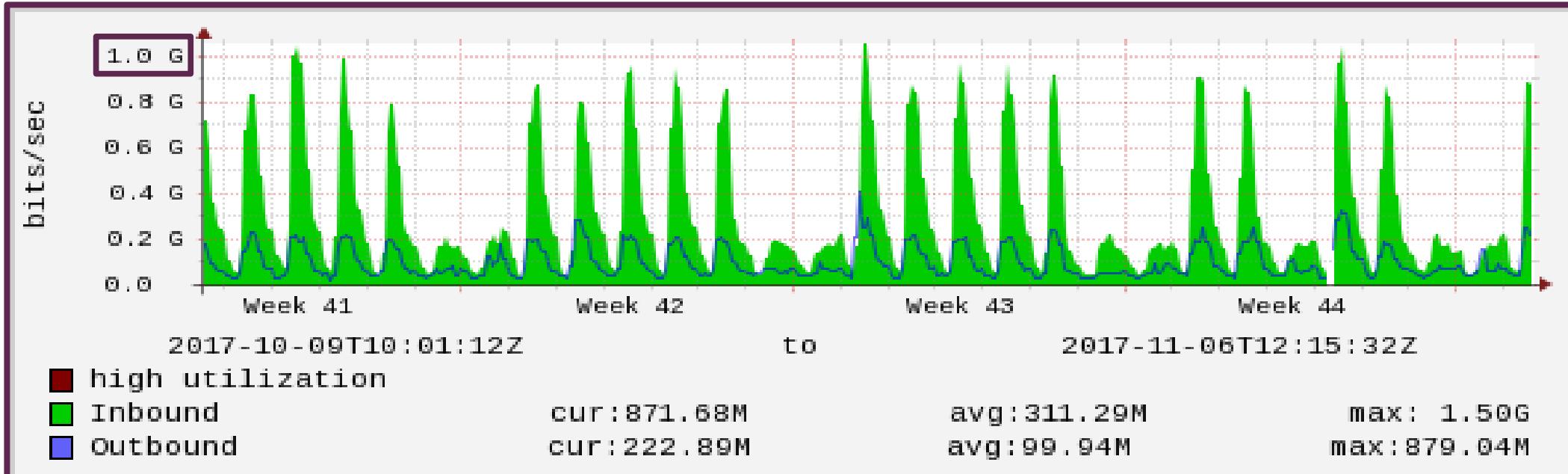
▼ OctInOut



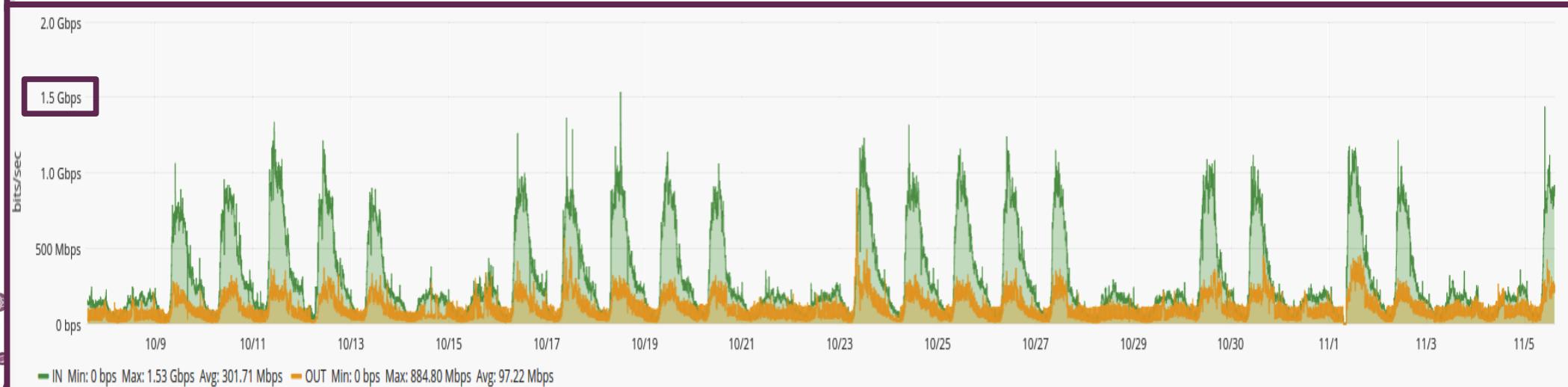
▼ PacInOut



Zenoss



Grafana





CUC2017

CARNetova korisnička konferencija

POVEZANI ZNANJEM (...25 GODINA)

19. CARNetova korisnička konferencija CUC 2017
8. – 10. studenoga 2017. / Hotel Valamar Lacroma



PITANJA?

HVALA NA PAŽNJI!