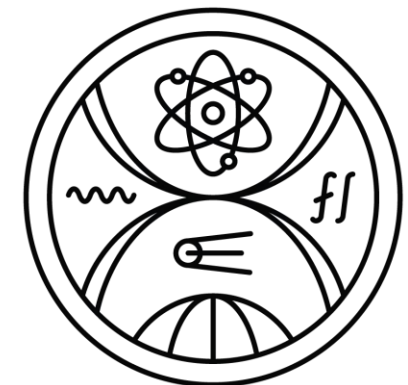


# Spracovanie dlhodobých meraní vybraných charakteristík Internetu

PETER TRENČANSKÝ

ŠKOLITEĽ: ING. DUŠAN BERNÁT, PHD.

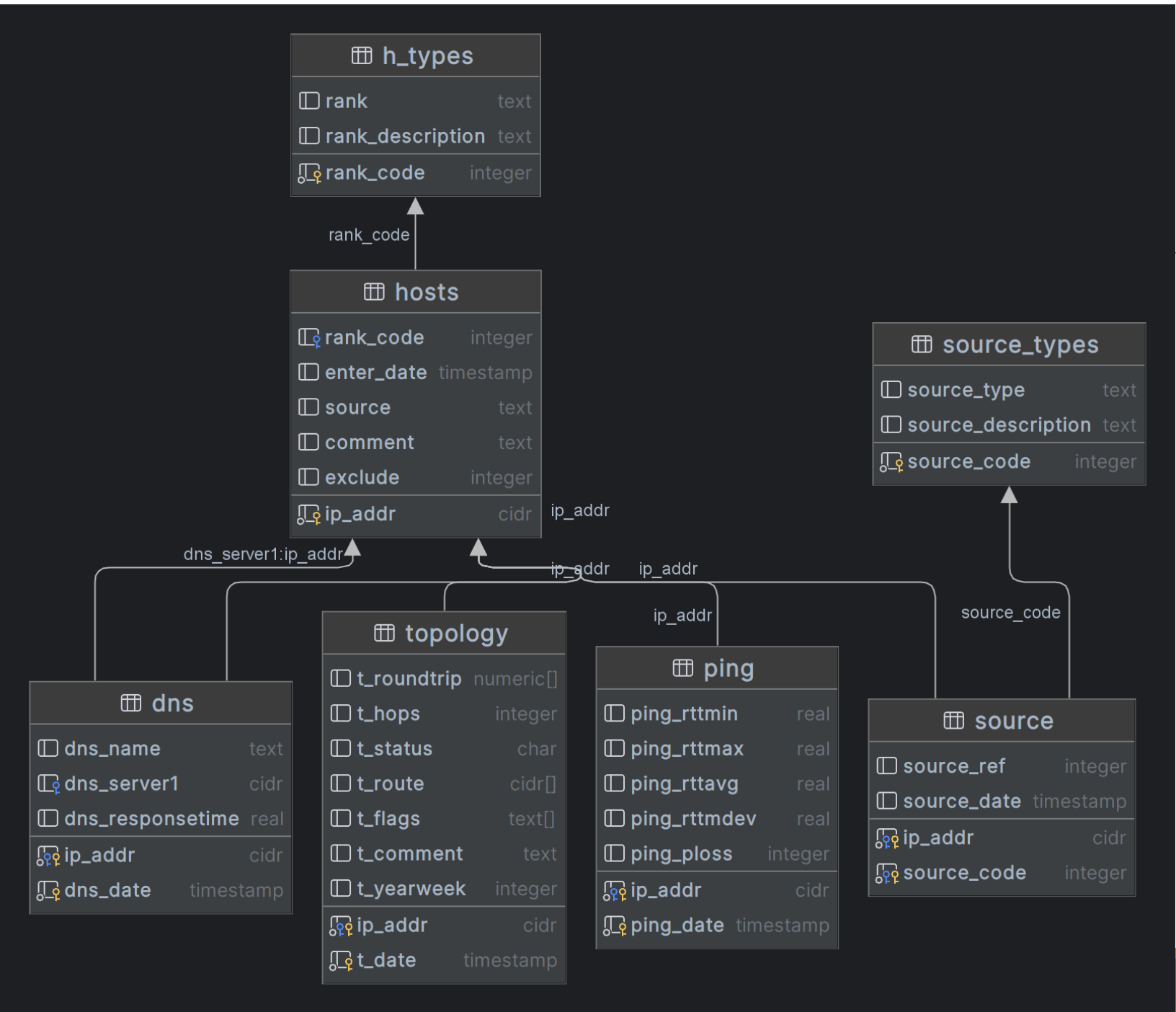


# Úvod

---

Máme databázu obsahujúcu dlhodobé merania internetu

Merania môžu slúžiť na pozorovanie vývoja internetu, ale aj na sledovanie zmien v čase prírodných katastrof alebo aj na sledovanie implementácií zmien v štandardoch



# Tabuľky

---

Pre nás dôležité tabuľky – hosts – zoznam všetkých IP adries v meraní  
- ping výsledky meraní doby odozvy

# Úvod

---

Merania prebiehali z jedného uzla nachádzajúceho sa na Slovensku

Databáza v súčasnosti obsahuje dáta z období apríl 2008 až november 2022

Dáta do nej môžu pribúdať

Cieľom je vytvoriť aplikáciu na predspracovanie a vizualizáciu výsledkov meraní

# Potreba predspracovania

---

Databáza má v súčasnosti asi 80 GB a bude sa zväčšovať, preto by bolo časovo nemožné robiť zložitejšie dopyty v reálnom čase. (Např. `SELECT count(*) FROM ping;` - trvá 1m35s)

Dáta ktoré nás budú zaujímať vopred spracujeme a uložíme do databázy, z ktorej sa budú ľahko vyťahovať.

Agregované dáta – zmenšený objem

Na to budú použité servisy bežiacie na pozadí servera, ktoré budú automatizovane pravidelne spúšťané aby prispracovávali pribúdajúce dáta

Spracovávanie vytvára nové dáta, nemaže staré dáta

# Typy meraní

## Ping

```
C:\Users\Peter> ping uniba.sk

Pinging uniba.sk [158.195.6.138] with 32 bytes of data:
Reply from 158.195.6.138: bytes=32 time=20ms TTL=55
Reply from 158.195.6.138: bytes=32 time=31ms TTL=55
Reply from 158.195.6.138: bytes=32 time=38ms TTL=55
Reply from 158.195.6.138: bytes=32 time=17ms TTL=55

Ping statistics for 158.195.6.138:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 17ms, Maximum = 38ms, Average = 26ms
```

## Traceroute

```
C:\Users\Peter> tracert uniba.sk

Tracing route to uniba.sk [158.195.6.138]
over a maximum of 30 hops:

  0  2 ms    1 ms    1 ms  internet.o2 [10.0.1.138]
  1  *        *        *    Request timed out.
  2  29 ms   37 ms   39 ms  10.109.20.1
  3  26 ms   17 ms   20 ms  90.176.30.9
  4  21 ms   18 ms   19 ms  O2BS-gw2.six.sk [192.108.148.43]
  5  33 ms   37 ms   39 ms  Sanet-gw1.six.sk [192.108.148.10]
  6  31 ms   18 ms   18 ms  sav-bratislava.sanet2.sk [194.160.8.16]
  7  27 ms   17 ms   20 ms  158.195.1.18
  8  35 ms   20 ms   18 ms  158.195.1.85
  9  19 ms   18 ms   19 ms  www.uniba.sk [158.195.6.138]

Trace complete.
```

# Spracovanie

---

Aplikácia bude mať podobu webovej aplikácie

Frontend – Angular – vizualizácie pomocou máp

Backend – ASP.NET Core – webová služba pre sprostredkovanie dát pre frontend

- služby oriebežne spracovávajúce údaje z databázy
- aplikácia na prvotné uloženie geografických dát



# Vizualizácie

---

Mapy

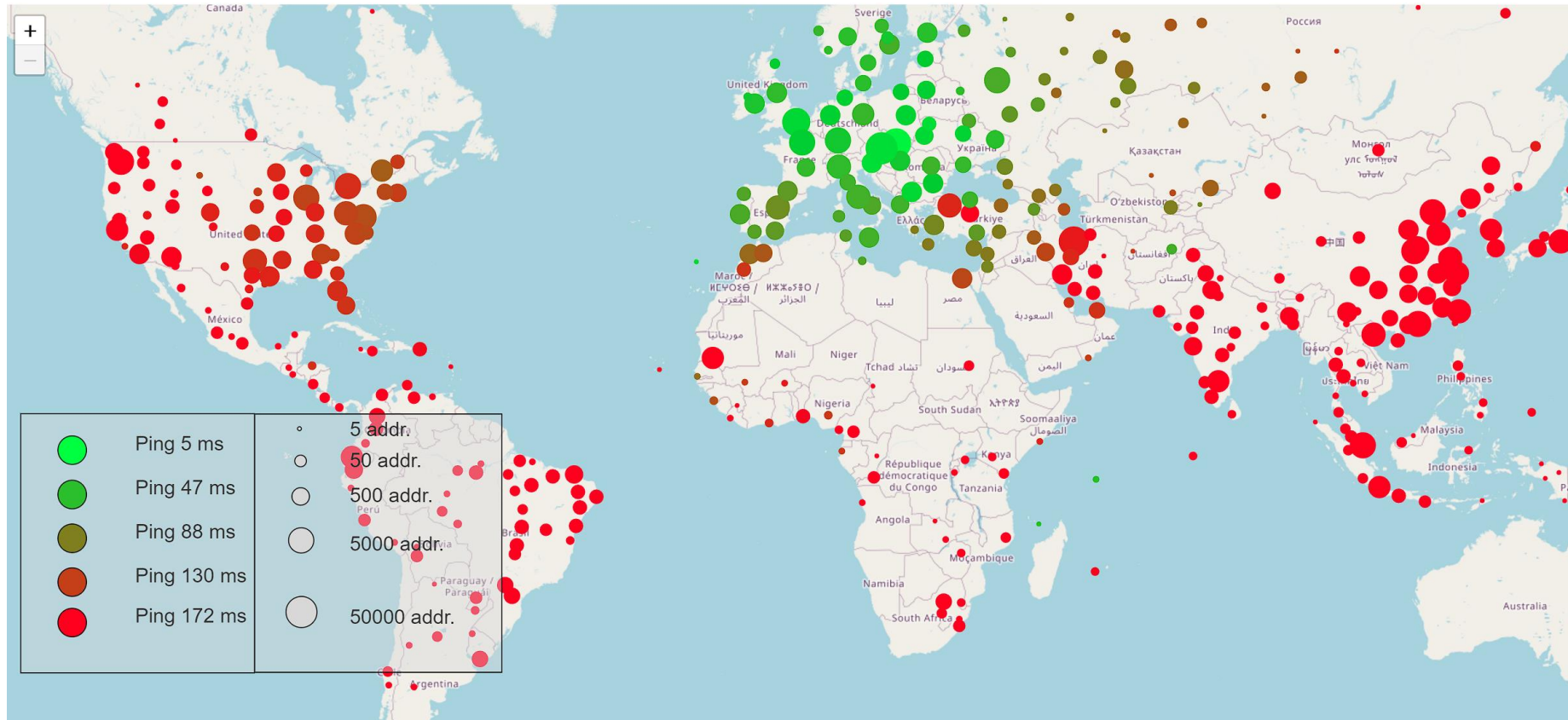
Dáta agregované podľa polohy na mape

Informácie o polohe boli získané z verejnej databázy o polohe IP adries

Dáta pochádzajú z údajov poskytovateľov internetu

Aplikácia nie je závislá od konkrétnych dát, geografické dáta jej stačí poskytnúť pri prvom spustení

# Zobrazenie údajov zoskupených podľa zemepisných súradníc



# Zobrazenie údajov zoskupených podľa zemepisných súradníc

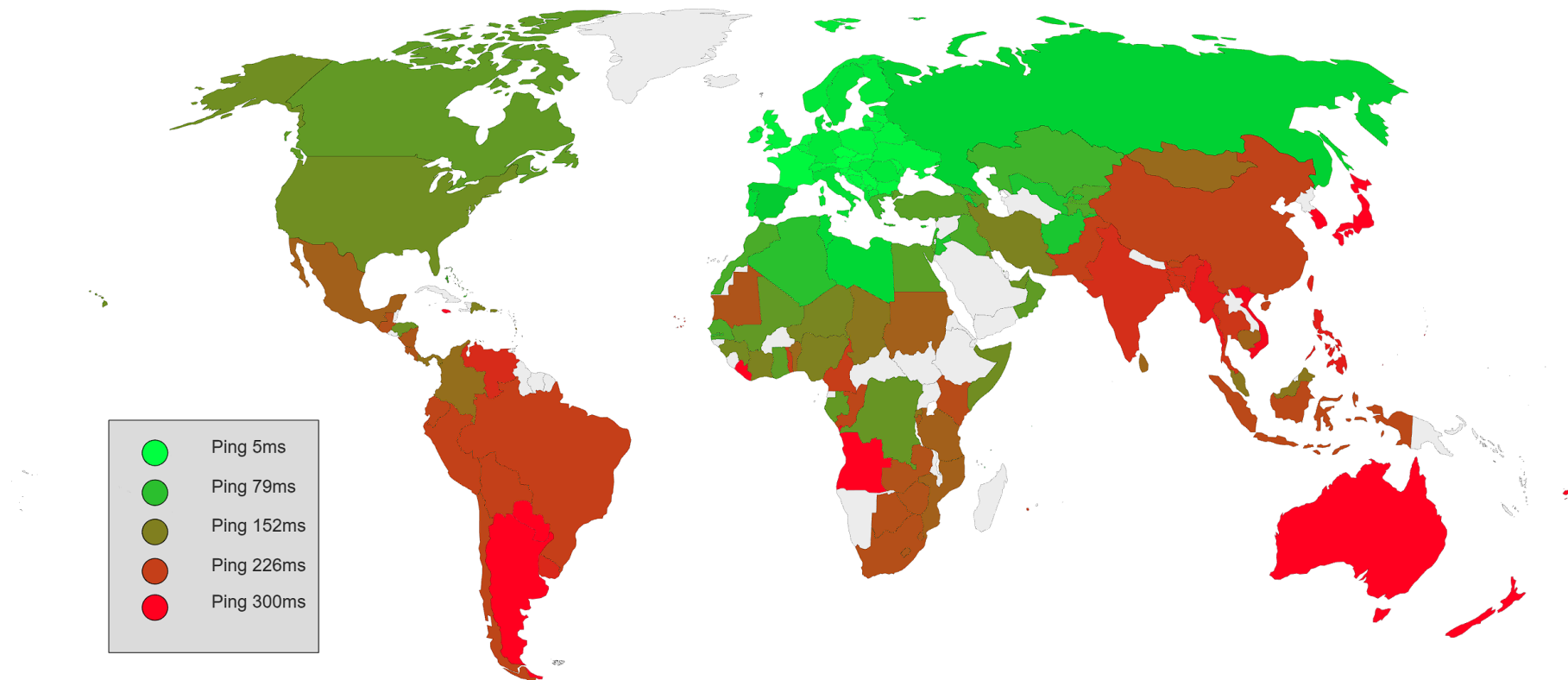
---

Možnosť voľby škály farieb

Možnosť voľby zobrazenia minimálnej, maximálnej alebo priemernej doby odozvy

# Zobrazenie údajov zoskupených podľa krajín

---



# Otázky do diskusie

---

Pre prácu s časovými radmi existuje open-source softvér TimeScaleDB, implementovaný ako nadstavba PostgreSQL. Jeho použitie by značne zjednodušilo architektúru riešenia z práce, pokiaľ by časy vykonávania dotazov spĺňali požiadavku na interaktivitu aplikácie. Bola zvažovaná možnosť použitia TimeScaleDB?

- použitie sme nezvažovali
- TimeScaleDB by bolo vhodné použiť pre databázu s meraniami, k tej však máme len read-only prístup

# Otázky do diskusie

---

Aká je veľkosť databázy s meraniami, v počte záznamov? Aké sú časy vykonávanie relevantných dotazov na tejto databáze v PostgreSQL? Aké indexy sú použité v tejto databáze?

Tabuľka	Počet záznamov	Indexy
hosts	281 330	Primárny kľúč na ip_addr
topology	71 187 905	Primárny kľúč (t_addr, ping_date), Index na t_date,
ping	563 605 275	Primárny kľúč na (ip_addr, ping_date), Index na ping_date, Index na ping_date pre null hodnoty

# Bežný select

---

```
SELECT avg(ping_rttavg), min(ping_rttmin), max(ping_rttmax), count(h.ip_addr)
FROM ping p
JOIN hosts h on h.ip_addr = p.ip_addr
WHERE p.ping_date between ('2019-05-11 00:00:00') AND ('2019-05-18 00:00:00')
      AND exclude = 0 and h.ip_addr NOT BETWEEN '10.0.0.0' AND '10.255.255.255'
      AND h.ip_addr NOT BETWEEN '172.16.0.0' AND '172.31.255.255'
      AND h.ip_addr NOT BETWEEN '192.168.0.0' AND '192.168.255.255'
      AND (EXISTS (SELECT * FROM topology t WHERE t.ip_addr=h.ip_addr AND t_status<>'E' LIMIT
1)
          OR EXISTS (SELECT * FROM ping p WHERE p.ip_addr = h.ip_addr AND ping_ploss
BETWEEN 0 AND 99))
GROUP BY DATE(h.enter_date)
```

12m 0s

# Zaujímavé pozorovania

---

Doba odozvy rastie s rastúcou vzdialenosťou

Nie však priama úmera

Turecko – oblasť Istanbulu a Ankary – neprimerane vysoká odozva, zvyšok krajiny je v norme

Austrália – medzi rokmi 2011 a 2022 sa priemerná doba odozvy skrátila takmer na polovicu

Japonsko – zemetrasenie v roku 2011 výrazne neovplyvnilo namerané hodnoty

USA – počas hurikánu Sandy stúpila najvyššia nameraná hodnota o 12 sekúnd