Turris:Sentinel

Nový systém pro sběr dat

Robin Obůrka • robin.oburka@nic.cz • 16.11.2018

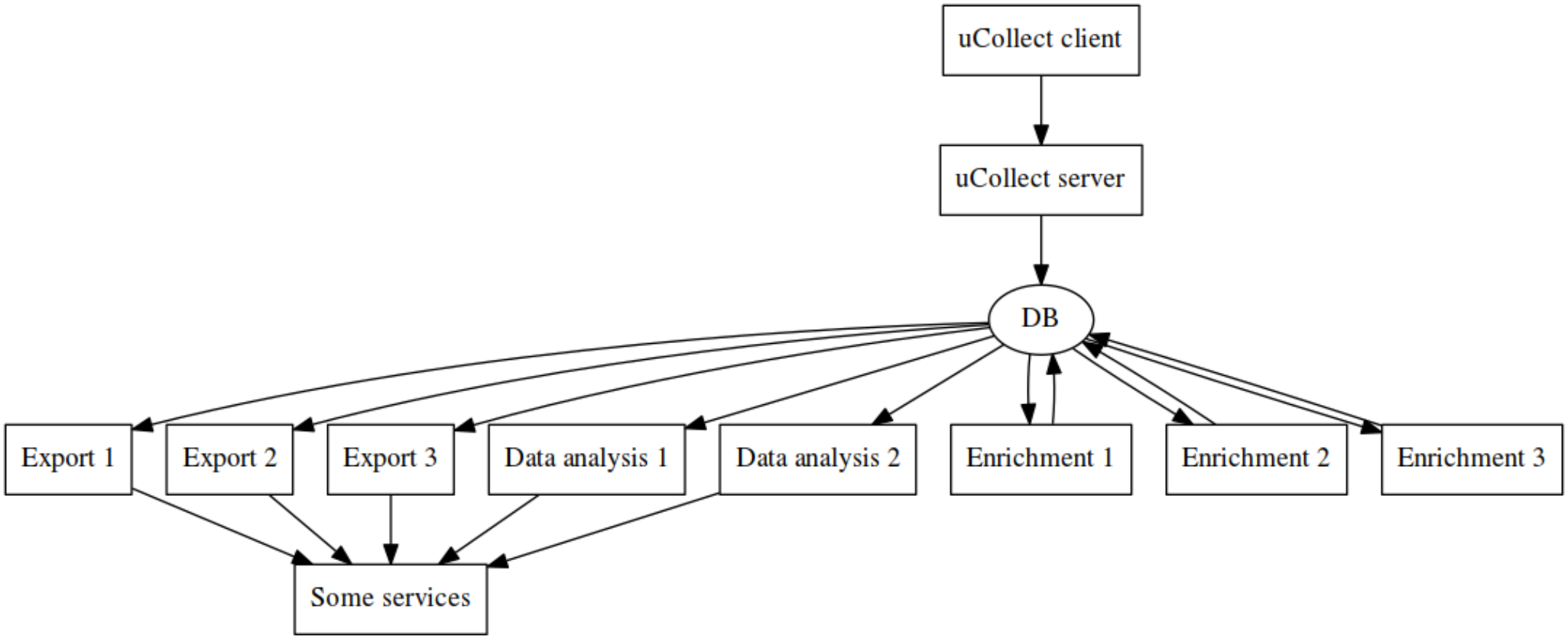
Projekt Turris a sběr dat

- Je s námi již od začátku
  - Turris je původně výzkumný projekt
  - Modré routery (Turris 1.0, Turris 1.1) rozdávané proti nájemní smlouvě
  - Dostupný i v Turris Omnia / MOX
- Výstupy
  - Greylist
  - dynamický firewall
  - Statistiky
  - „úlovky“
- Nebylo o nás moc slyšet
  - Aktuální systém „mele z posledního“
  - Náhrada starého systému za nový – uCollect → Sentinel

uCollect

- Ze všeho nejdříve
  - Původně výzkumný projekt pro malý počet zařízení
  - Řešení psané
    - Na daný počet zařízení
    - S danými výhledy a plány
  - Nevyužívané funkce omezují škálovatelnost
  - Špatná rozhodnutí
- Klient-server řešení
- Data ve velmi stresované databázi

uCollect


uCollect client
uCollect server
DB
Export 1
Export 2
Export 3
Data analysis 1
Data analysis 2
Enrichment 1
Enrichment 2
Enrichment 3
Some services

uCollect

- 2 úzká hrdla
  - Klientský protokol
    - Zbytečně složitý (nevyužívané možnosti)
    - Vázaný na jedno-instanční server
    - Nereálný přepis – po přepisu protokolu nic nezbude
  - Centrální DB
    - Za hranicí životnosti
    - Už nelze škálovat pomocí HW
    - Nelze přidávat další analýzy a klienty

Vize nového systému

- Základní motto: Připravit se na cokoliv
- Vše musí být škálovatelné dalším HW a dalšími instancemi
- Vše maximálně zjednodušit
  - Zahodit nevyužívané funkce v návrhu
  - Využít maximum existujících technologií
- Použít best practices, učit se, bavit se s lidmi co řeší podobné problémy

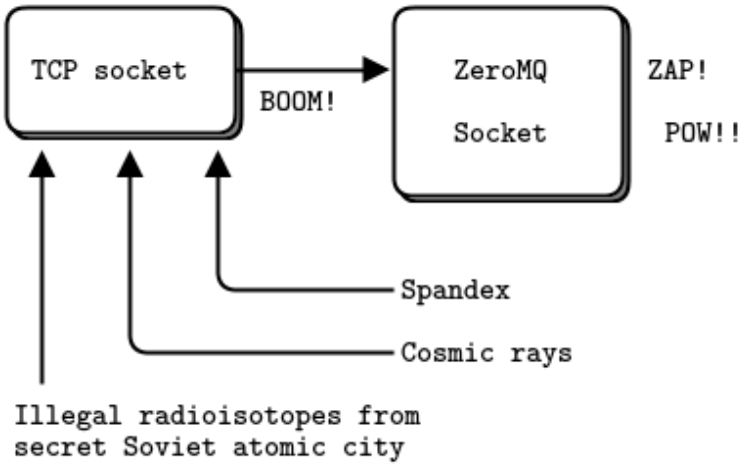
Sentinel

- „Celým“ jménem Turris:Sentinel
- „Microservice architektura“
  - Použití message queue technologií
  - Metoda skládačky
  - Nano-, piko-, …

Sentinel – použité technologie

- MQTT pro ingress dat
  - Standardní řešení daného problému
  - Broker (aktuálně Mosquitto)
  - Publish/Subscribe
  - Přihlášení pomocí klientského certifikátu, šifrování
  - Automatizovaná CA pro ověření klienta
  - Prozatím jeden uzel
  - Existuje spousta SW s podporou clusteringu

Sentinel – použité technologie



- ZMQ „uvnitř“ sítě
  - Jiný přístup
  - Brokerless
  - „TCP na steroidech“
  - Komunikační patterny - PUB/SUB, PUSH/PULL, REQ/REP,…
  - Funkcí zadarmo v rámci vhodně zvoleného patternu
  - V některých věcech náročnější
    - Nevhodné použití pro uzly, které nemám pod správou
    - Zabezpečení není moc elegantní
    - Aplikace postrádá spoustu informací
    - Občas musí něco hlídat aplikační protokol

Sentinel – náš standard

- SN „framework“
  - Trochu atypická knihovna
  - Požírá trochu více z prostředí než je zvykem
  - … tj. I věci, které obvykle náleží aplikaci
  - Samotný box poskytuje pouze logiku
  - Fungující krabičky i v rozmezí 10, 20 řádků
- Dominuje PUSH/PULL pattern
  - Load-balancing zadarmo
  - Alespoň nějaké garance doručení
- Proudové zpracování
  - Z DB ideálně vůbec nečíst
  - Vše co potřebujeme dělat „za běhu“

Sentinel – náš standard

- Jednoduchost
  - Máme i velmi naivní boxy
  - Některé funkce vznikají poskládáním vhodných boxů za sebe
- „Routing“ zpráv podle „msg type“
  - Hierarchický identifikátor zprávy
- Ansible
  - Centrální správa nastavení
  - Protipól SN
- Systemd
  - Myslíme si o něm své, ale
  - … celkem dobrý pomocník

Sentinel – architektura

- Intenzivní vývoj (dokončujeme stage 1)
- 3 základní části
- Certifikátor
  - Automatizovaná CA
  - Klient žádá o certifikát
  - Ověříme klienta
    - Náš router → ATSHA204A
    - Nebo jinak (komu chceme dát přístup)
    - Cesta k otevření i jiným architekturám
  - Vystavíme certifikát

Sentinel – architektura

- Pipelines
  - Samotné proudové zpracování dat
  - Pro každý typ dat jedna pipeline
  - Činnosti
    - Příjem
    - Validace
    - Obohacení
    - Distribuce analýzám
    - Archivace

Sentinel – architektura

- DynFW
  - První a zatím jediná „analýza“
    - Sběr informací z mnoha zdrojů
    - Přiřazování skóre zlým IP adresám
      - Každý typ příchozích dat má jinou „váhu“
    - Po překročení limitu jde adresa do blacklistu
  - Fáze „ladění koleček“
  - Vše realtime – včetně publikace dat
- Viz ukázka schématu a komentář k němu

DynFW – vyzkoušejte si!

- https://gitlab.labs.nic.cz/turris/sentinel/dynfw-example-client/
- Bacha na
  - Sice realtime, ale aktuálně pouze 2 zdroje dat:
    - Minipoty – malý počet klientů (testujte!)
    - HaaS – čteme jednou za půl hodiny
  - DynFW neopakuje eventy
    - 1 uživatel může 1 adresu nahlásit v rámci 1 analýzy pouze jednou
    - Svoji aktivitu uvidíte pouze jednou jedinkrát za časový úsek
    - … ano, je to opravdu feature, nemáme to rozbité
- Prosím, čtěte README, kde je popsán protokol, jak se má klient správně chovat

Co můžete testovat

- Sentinel a DynFW na routeru
  - https://forum.turris.cz/t/trying-new-data-collecting-system-sentinel/8637
- DynFW v RPM pro vaši distribuci
  - https://software.opensuse.org/download.html?project=home:-miska-&package=sentinel-dynfw-client
- Jak zabalit klienta pro vaši distribuci
  - https://build.opensuse.org/package/show/home:-miska-/sentinel-dynfw-client

Děkuji za pozornost

Robin Obůrka • robin.oburka@nic.cz