

Intelligens Data Science a gyakorlatban: Esettanulmány az elérhető Big Data-hoz



EXPERTISE WITH NO LIMITS

Big Data a mindennapokban

- Önvezető autók
 - Mesterséges intelligencia
 - Arcfelismerés
 - Nyelvfelismerés
-
- Ami közös a fentiekben:
Hosszú, drága projektek eredményei



EXPERTISE WITH NO LIMITS

Big Data a mindennapokban

- Monitoring
 - Business Intelligence
 - DevOps
 - Log-processing
-
- Ami közös a fentiekben:
Minden alkalmazásnak szüksége van rá valamilyen szinten



EXPERTISE WITH NO LIMITS

A környezet

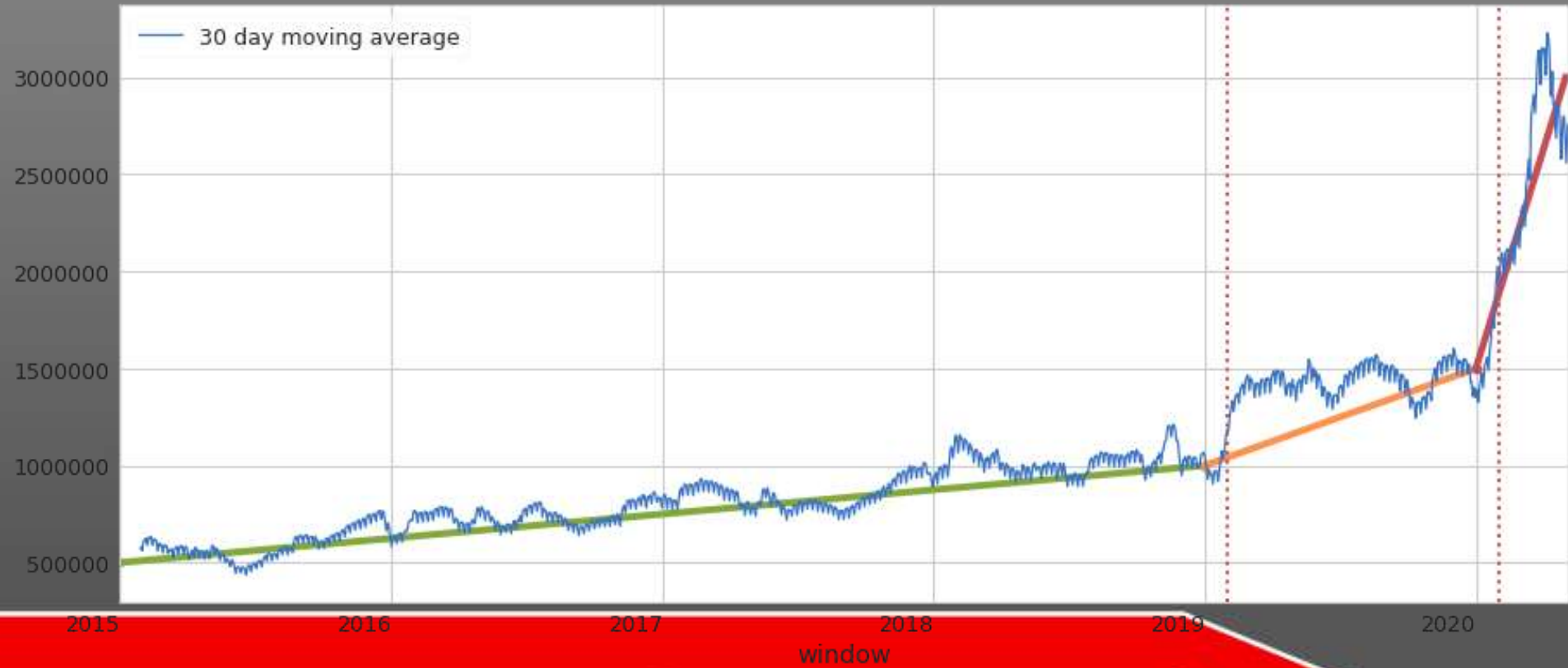
- Banki, bankközi, tőzsdei tranzakciók
- ~ 5-7 millió call naponta
- Microservices

- Kihívások:
Bottleneck azonosítása, teljesítmény becslése,
log elemzés, monitoring támogatása



EXPERTISE WITH NO LIMITS

A környezet



EXPERTISE WITH NO LIMITS

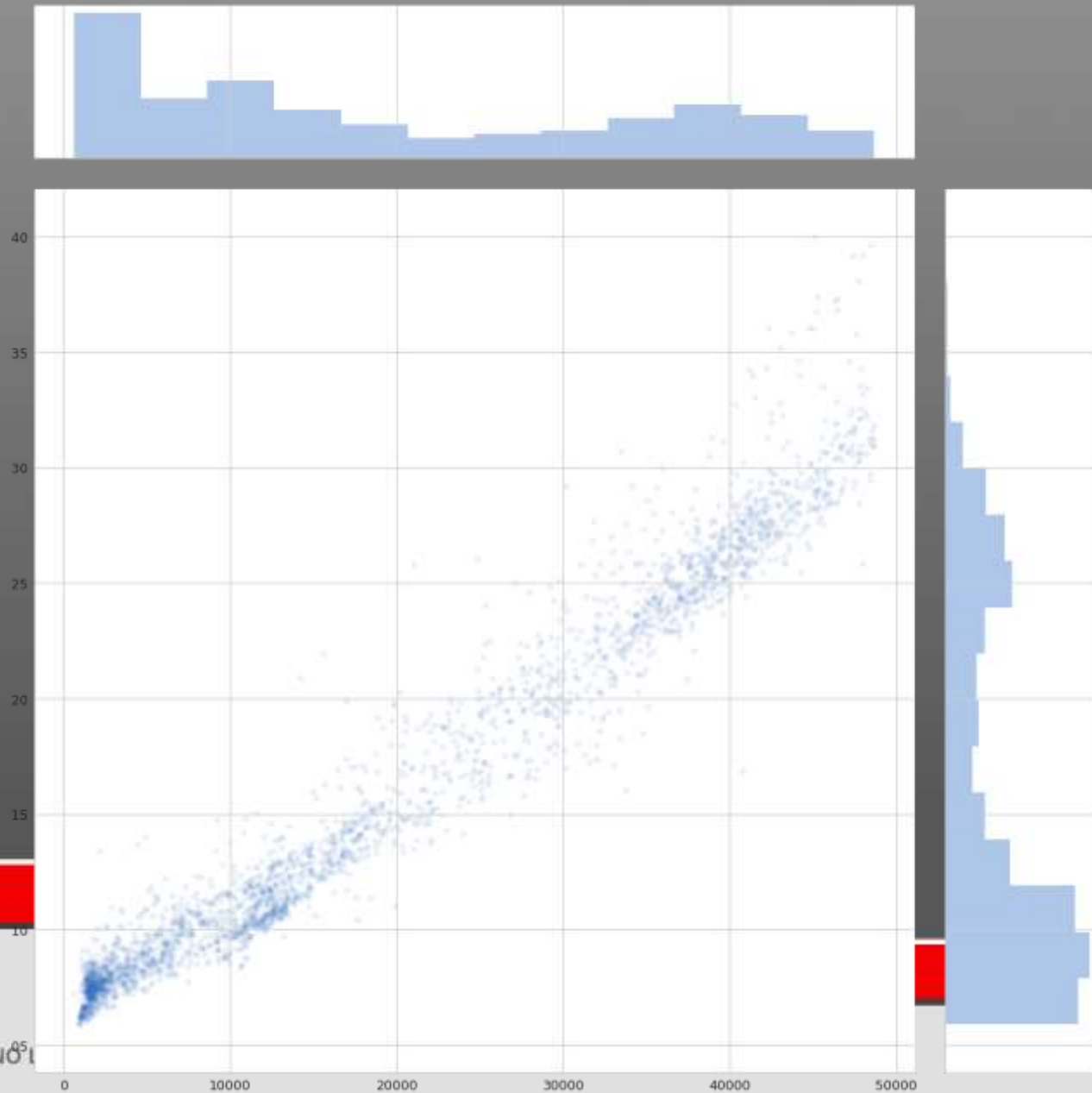
CPU adatok hasznosítása

- A legtöbb cégnek, amely üzemeltet valamilyen alkalmazást, vannak adatai a szerverei állapotáról, valamint az alkalmazások terheléséről.
- Ezek az adatok sok esetben csak pillanatnyi monitoringra használtak.

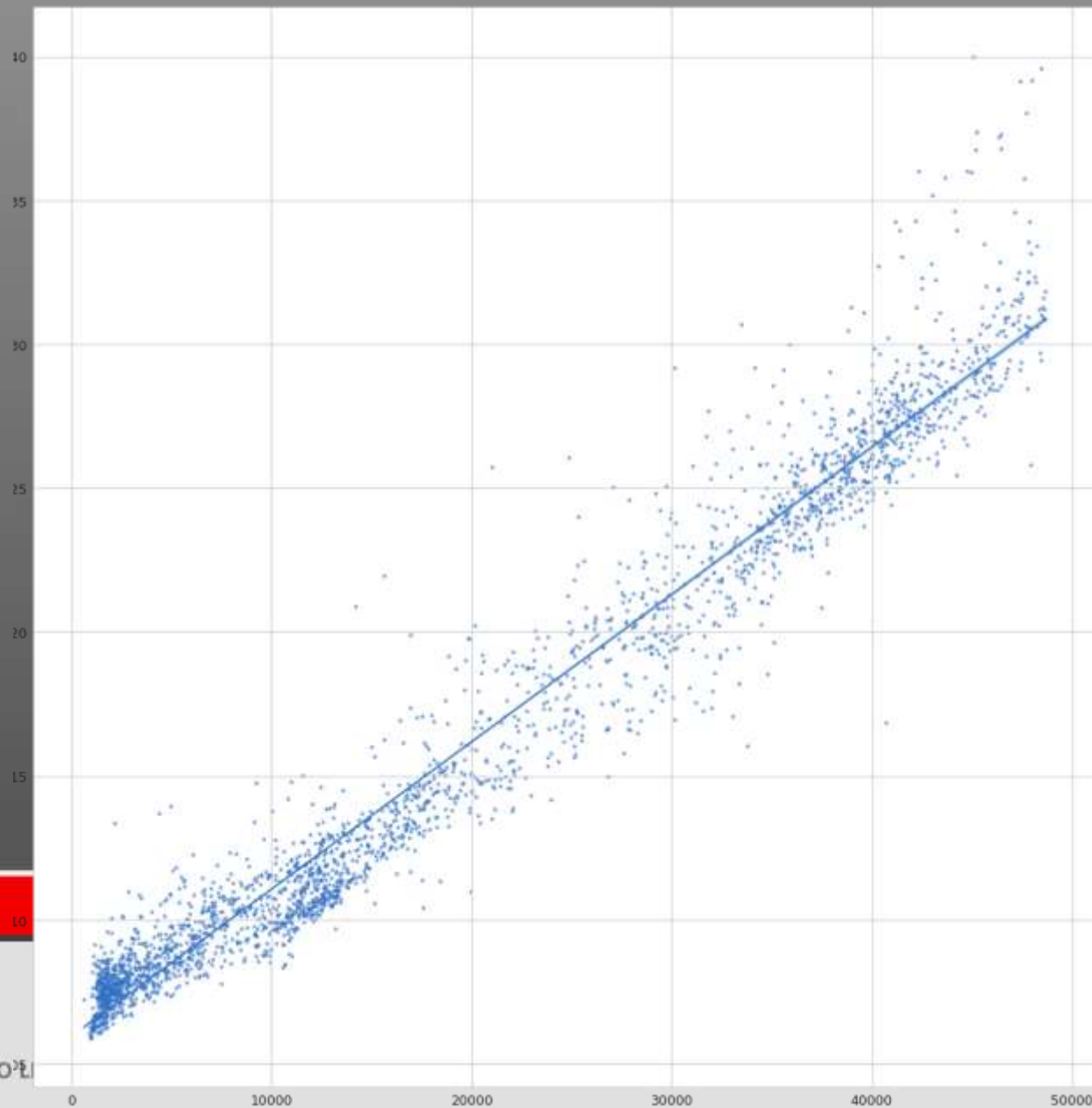


EXPERTISE WITH NO LIMITS

- A heatmap-en jól láthatóak a különböző kilengések és a normál viselkedés.
- Ezen adatok alapján egy lineáris kapcsolat sejthető.

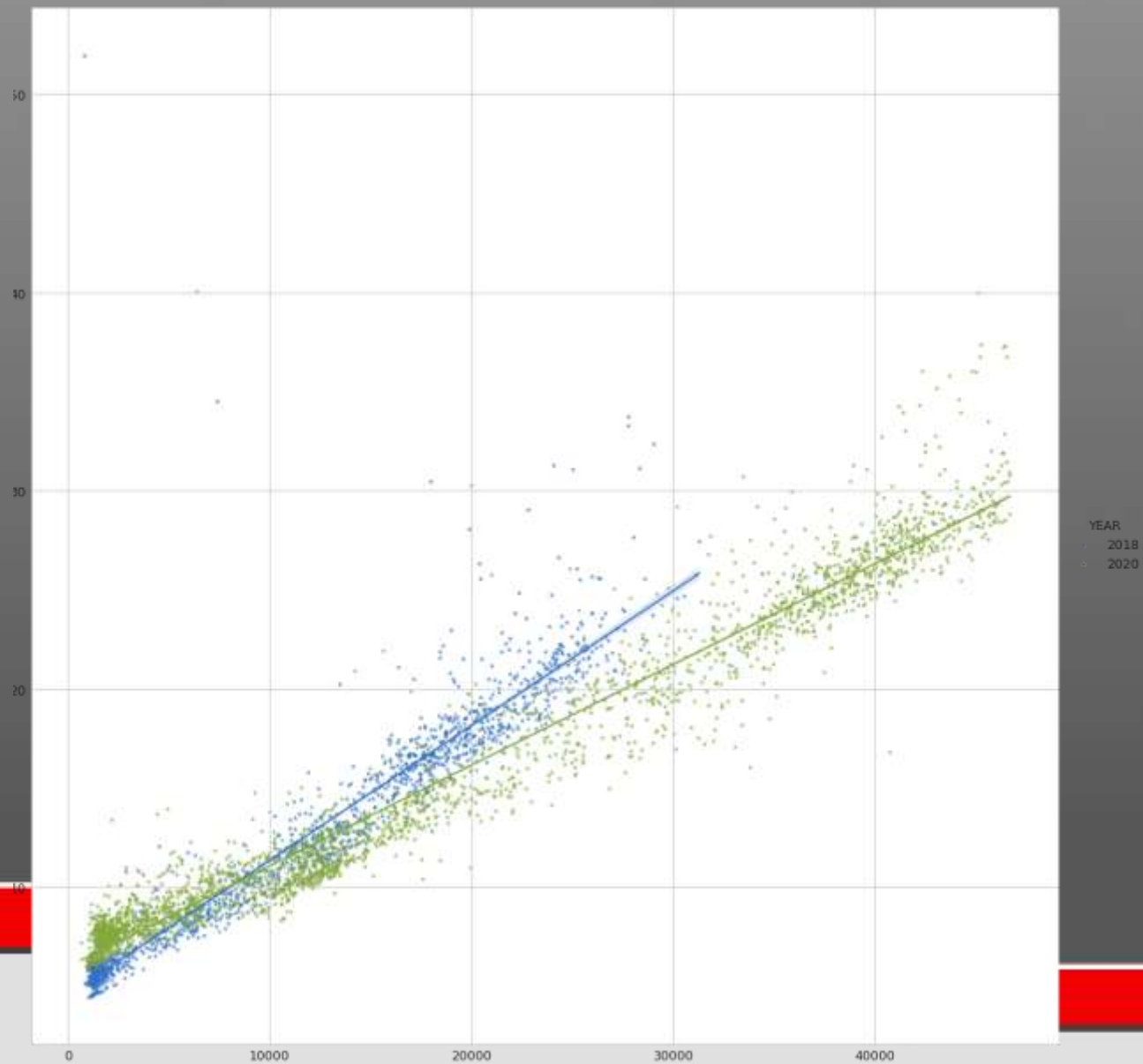


- A CPU adatokkal elkészíthető az adott alkalmazás CPU profile-ja.
- Ezzel becslések készíthetők a teljes terhelési értékekre.



EXPERTISE WITH NO²E

- A CPU adatokkal elkészíthető az adott alkalmazás CPU profile-ja.
- Ezzel becslések készíthetők a teljes terhelési értékekre.



CPU adatok hasznosítása

- Az eredmények pontosságát jól mutatja, hogy 2020 nyarán a rendszer egy burst során eljutott a közel 100%-os terheltségig, becslésünk pár százalékkal lőtte alá a tényleges értéket.



EXPERTISE WITH NO LIMITS

Monitoring

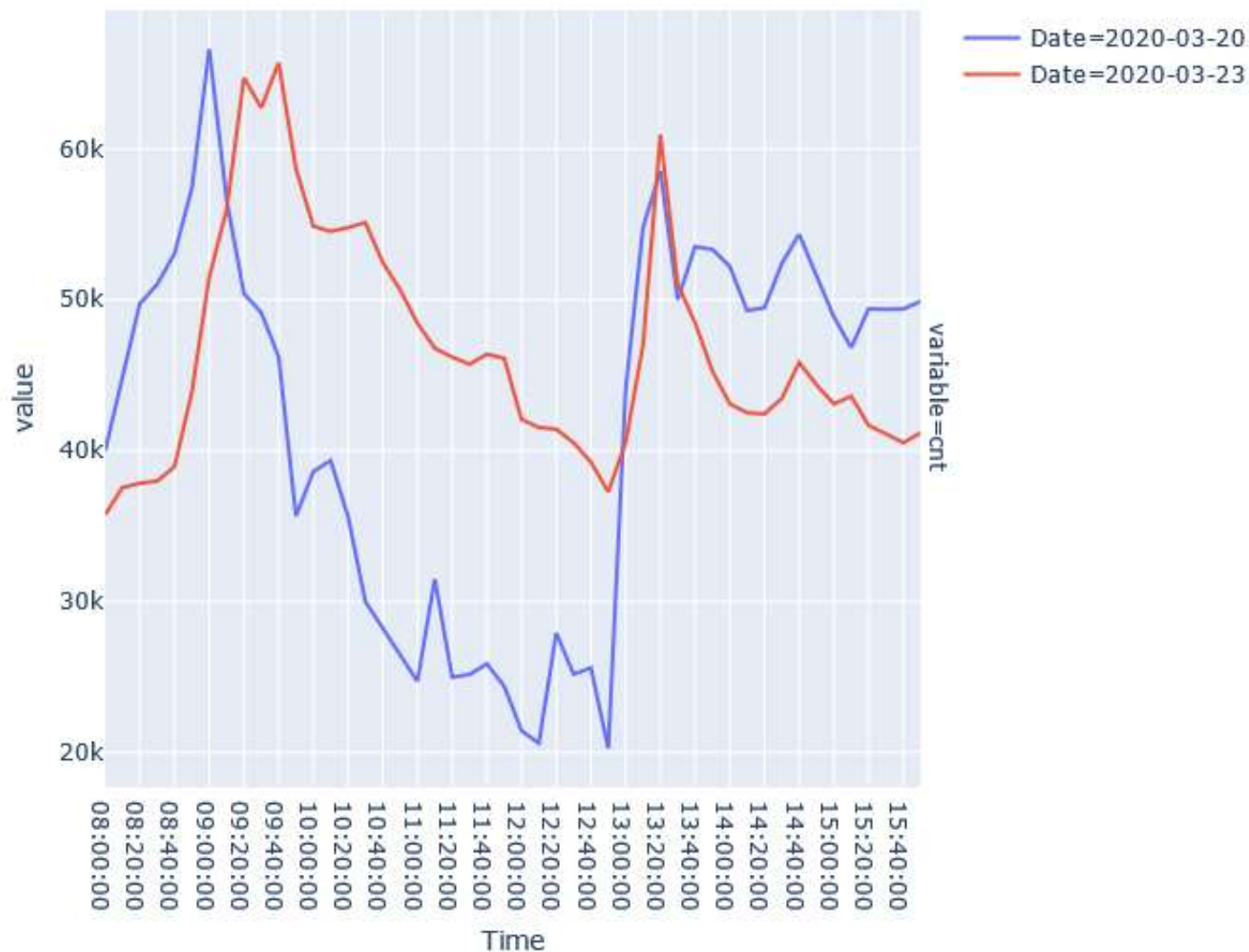
- A monitoringhoz rendkívül sok eszköz rendelkezésünkre áll (Prometheus, ElasticSearch/Kibana, stb.)
- Ezen eszközök nagy része támogat különböző AI megoldásokat.
- A leghasznosabb oldala ezen eszközöknek a custom metrikák definiálása.



EXPERTISE WITH NO LIMITS

- A monitoring megoldások fő célja, hogy megtaláljuk az átlagos működéstől való eltéréseket és ezek okait

Request count (10 min aggregates)



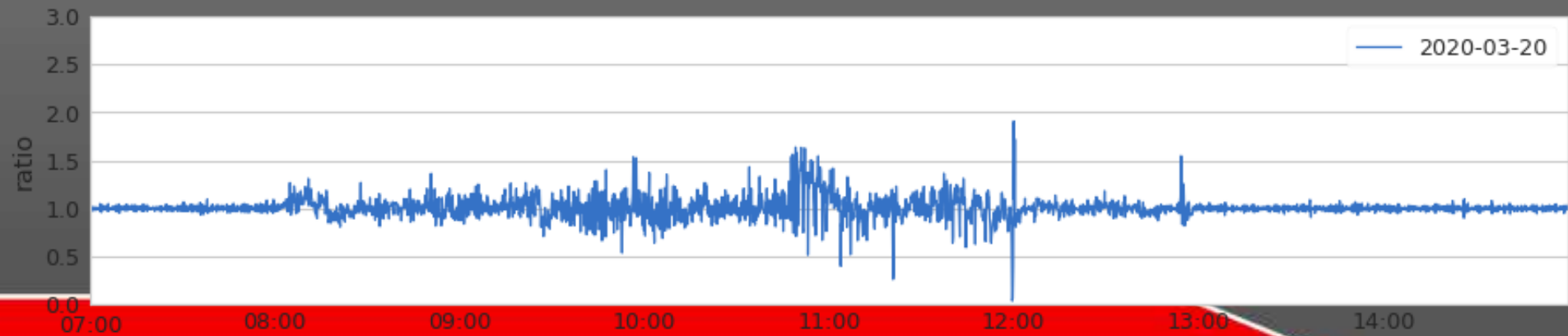
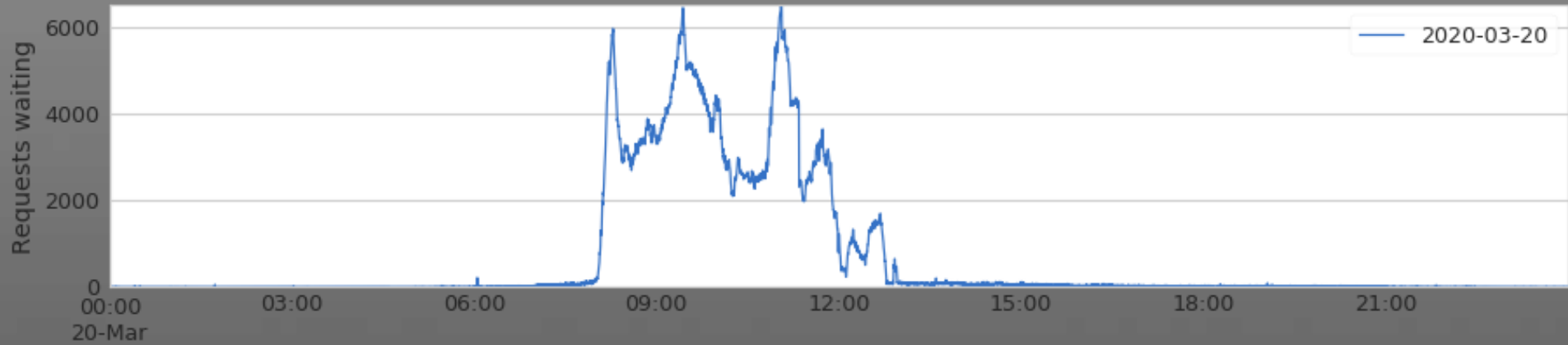
Monitoring

- Ezen cél sokkal összetettebb mint elsőre hangzik
- Az átlagok folyamatosan mozoghatnak, egy-egy kilengő érték nagyban befolyásolhatja
- Folyamatosan frissíteni kell a benchmarkokat a minél pontosabb eredmények érdekében
- A benchmarkingnál az alap adatok minősége és mennyisége kritikus kérdés



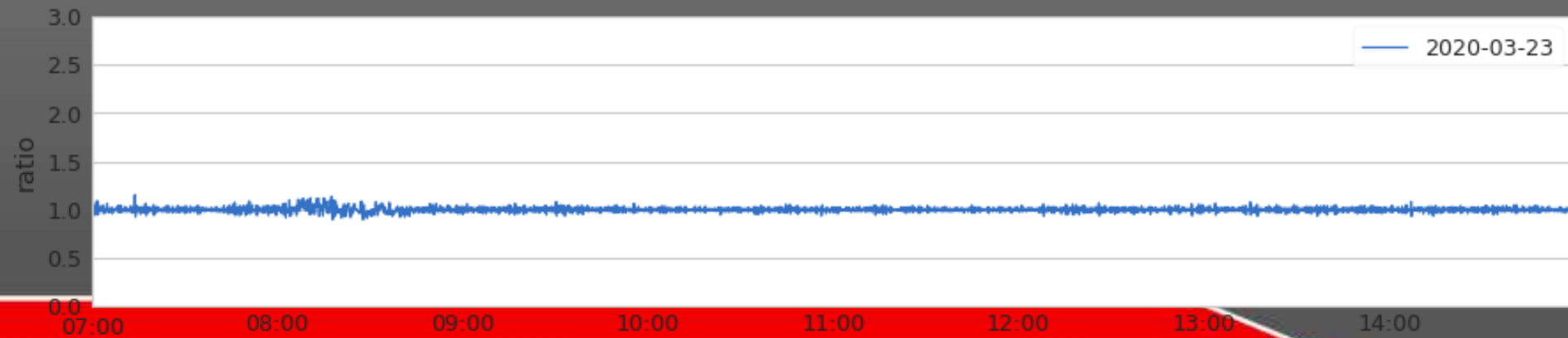
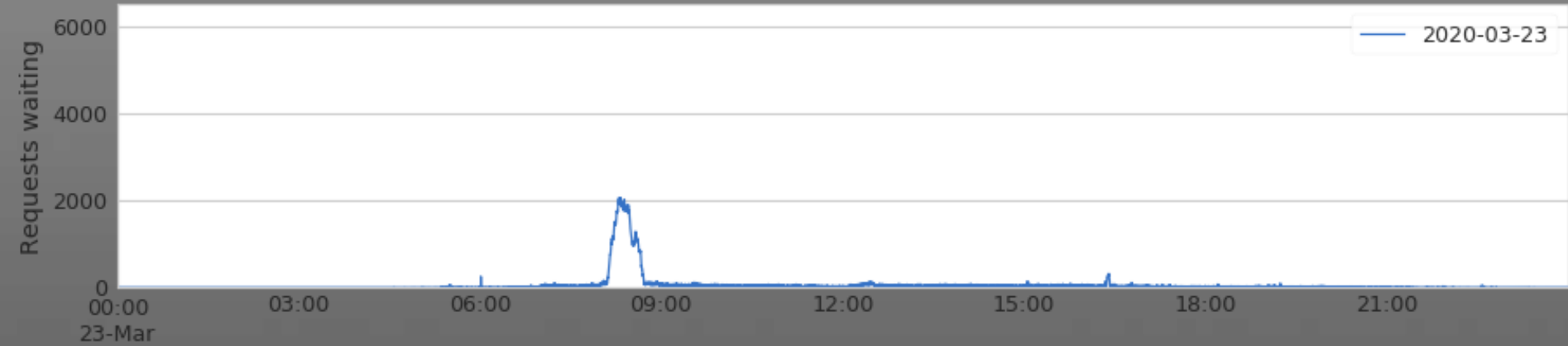
EXPERTISE WITH NO LIMITS

Monitoring



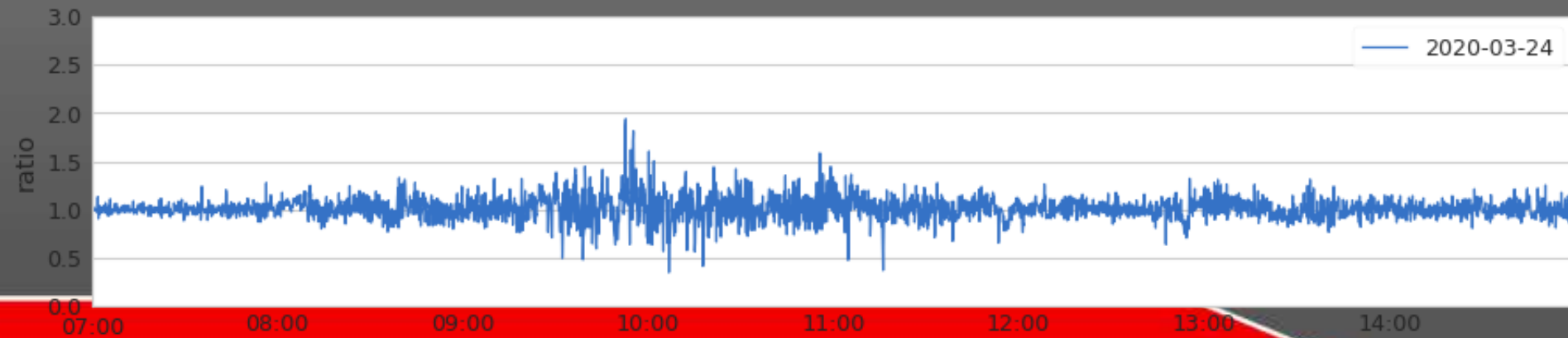
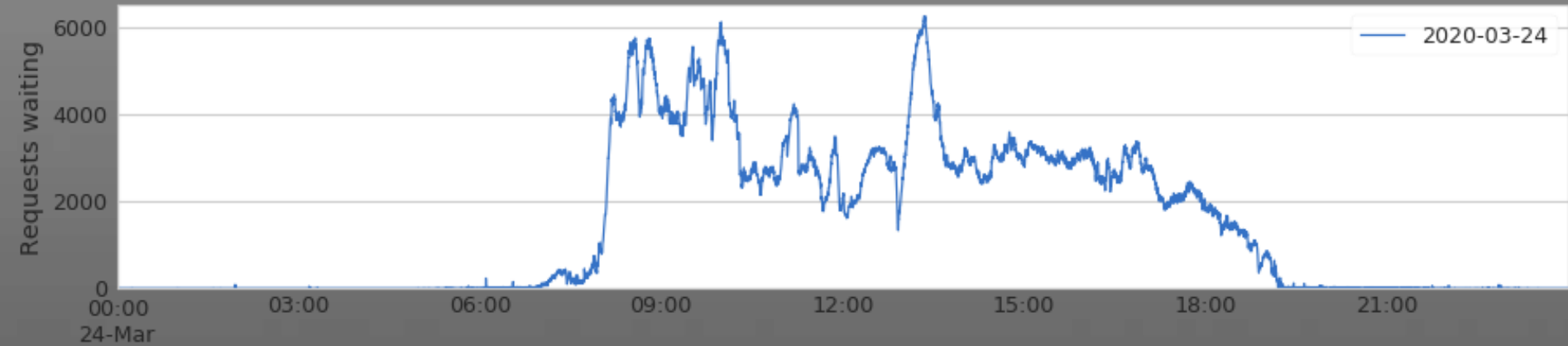
EXPERTISE WITH NO LIMITS

Monitoring



EXPERTISE WITH NO LIMITS

Monitoring



EXPERTISE WITH NO LIMITS

Monitoring

- A definiált metrikáknak köszönhetően sikerült megtalálni a hibás service-t, mely egy frissítésnek köszönhetően lassult le
- Ezen megoldások a rendszerek komplex monitoringjával könnyíthetik a rendszerek karbantartását a problémás részek azonosításával, gyorsabb reakcióidővel.



EXPERTISE WITH NO LIMITS

Adatok forrása

- Általában is minél több adat áll rendelkezésre, annál átfogóbb képet kaphatunk a rendszereinkről. A fenti elemzések az alábbi adatforrásokból dolgoztak:
- Audit adatbázis
- Oracle Diagnostics Pack
- Oracle Enterprise Manager Cloud Control (EMCC)
- Java rendszerek logjai

További lehetőség elemzésre:

Oracle Cloud Observability and Management Platform



EXPERTISE WITH NO LIMITS

Lehetséges buktatók

- Az előadásban bemutatott két eredmény rendkívüli mennyiségű és minőségű rendelkezésre álló adatot használt.
- Minden matematikai, statisztikai elemzés sarokkövét jelenti az adatok mennyisége és minősége.
- A fenti elemzések könnyedén replikálhatóak bármely rendszeren a megfelelő adatmennyiség rendelkezésre állása esetén.



EXPERTISE WITH NO LIMITS

Konklúzió

- A fenti példák pár hetes projektek eredményei, ad-hoc események elemzései, melyek eredményeit a gyakorlatban tudtuk tovább használni a napi működés során.
- Két fős team munkájának eredménye.
- A megfelelő minőségű és mennyiségű adatok esetén ezen elemzések más, erőforrásigényes folyamatokat is kiegészíthetnek, helyettesíthetnek (stressztesztek, átfogó rendszertesztek update-k folyamán).



EXPERTISE WITH NO LIMITS