Engineer Data

30 dagars eldata på 180 ms

Utvecklade varianter för timvis aggregering av elförbrukning i Python, SQL och Bash + Jq; Python-varianten behandlade en JSONL-datamängd för 30 dagar på 180 ms i prestandatestet.

2025

180 ms

för att aggregera en månads mätaravläsningar

Situation. Företaget bearbetar stora volymer heterogena eldata från flera datakällor, var och en med sin egen rapporteringsfrekvens — från timintervall till 15-minutersintervall och i vissa fall oregelbundna tidsstämplar. Variationen gör det svårt att skapa ett sammanhängande och jämförbart dataset. För att möjliggöra korrekt energianalys måste dessa data normaliseras till konsekventa förbrukningsvärden per timme, aggregerade per zon.

Uppgift. Målet var att ta rå händelsedata från ett historiskt dataset (i JSONL-format) och omvandla dem till elförbrukningssiffror justerade till hela timmar, strukturerade som en rad per timme och zon. Konkret innebar uppgiften att:

- justera tidsstämplade data till strikta timintervall,
- aggregera den totala elproduktionen inom varje intervall genom att summera värdena i produktionsmixen,
- ta med gränsöverskridande utbyte genom att lägga till import och dra av export,
- leverera slutvärdena i ett strukturerat, skalbart format som lämpar sig för vidare analys.
- Lösningen behövde dessutom vara tillräckligt effektiv för att skala till långa tidsfönster (över 30 dagar) och till flera länder/zoner.

Handling. För att lösa uppgiften implementerades tre olika varianter av lösningen, i Python, JQ (för JSON-bearbetning på kommandoraden) och SQL, var och en optimerad för olika sammanhang:

Python valdes för sin flexibilitet, sin smidiga datahantering och sin förmåga att effektivt hantera transformationer i minnet. Med rena Python-funktioner byggdes en pipeline som:

- tolkade JSONL-filerna till strukturerade dataframes
- omsamplade tidsseriedata till timintervall
- aggregerade produktionen och beräknade nettoelförbrukningen (produktion + import – export)
- exporterade resultaten som CSV eller läste in dem i en lättviktig SQLite-databas för inspektion.

JQ användes för att skapa en snabb lösning med minimala beroenden för kommandoradsmiljöer, byggd som en kedja av JQ-filter.

JSONL-data importerades till PostgreSQL, där normaliserade tabeller skapades och en serie SQL-frågor skrevs.

För att utvärdera prestandan benchmarkades varje lösning med historiska dataset över både 1 dag och 30 dagar.

Resultat. Python-implementationen visade sig vara den snabbaste och mest skalbara lösningen och klarade:

- bearbetning av 1 dags data på bara 34 ms
- ett dataset på 30 dagar på 180 ms

Det bekräftade att den lämpar sig för längre tidsperioder med en bearbetningstid under en sekund. Den gav också tydlig och underhållbar kod som enkelt kunde byggas ut till bearbetning av flera zoner eller integreras i en ETL-pipeline.

Sammantaget visade angreppssättet med flera verktyg på flexibilitet i verktygsvalet, stark prestandaoptimering och robust hantering av tidsjustering och aggregering i verkliga datapipelines för eldata.

Använd erfarenheten i ert projekt

Använd ”30 dagars eldata på 180 ms” som utgångspunkt för en liknande utmaning i er organisation. Tjänsterna bakom arbetet finns tillgängliga för ert projekt. Hittar databasfrågorna och de saknade indexen bakom ett system som har blivit tyngre för varje år, och åtgärdar dem som spelar roll.

Se våra tjänster Hör av er →

Se bakom resultatet

Läs om situationen, besluten och resultatet i exemplet. Vi kan diskutera vad som passar era förutsättningar och vad som kräver ett annat angreppssätt. 30 dagars eldata på 180 ms — 180 ms för att aggregera en månads mätaravläsningar

Se våra tjänster Hör av er →

Behöver du snabba och effektiva datapipelines?

Prata om ert projekt med ingenjören som ansvarar för leveransen. Vi avtalar omfattning, hur priset beräknas och överlämning skriftligt innan arbetet börjar.

Roll i uppdraget: Plattforms- och dataingenjör. Det bygger på Databaser, PostgreSQL, SQL, Data engineering, Datapipelines (ETL/ELT), Dataanalys, Automatisering och CI/CD, Prestandaoptimering och Python. Och nu när ni är här: vi är ganska bra på Optimering av databasprestanda, Utveckling av datapipelines (ETL/ELT) och Dataanalys och BI-dashboards.

Hör av er Se våra tjänster

En del av **en längre historik** — och den finns i sin helhet här.

Varje post är skriven på samma sätt: situationen, uppgiften, vad vi gjorde och vad som förändrades.

Se hela historiken

Definiera förbättringen ni behöver

Ta med ett representativt dataset, frågorna det ska besvara och dagens flaskhals. Vi kan avtala mått på kvalitet, genomströmning och återställning.

Se vad det kostar Hör av er →

Behöver du snabba och effektiva datapipelines?

Fråga om besluten bakom ”30 dagars eldata på 180 ms”. Ni talar med personer som kan förklara arbetet och bedöma ert eget projekt.

Se våra tjänster

welcome@engineer.company +45 50 10 61 84 Köpenhamn, Danmark

Ta med dig den här sidan

Visitkort

Omsorgsfullt skapad med kärlek av D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Upphovsrätt © 2025 – nu · Engineer ApS · Alla rättigheter förbehållna.

Utskrivet från https://data.engineer.company/sv/portfolio/a-month-of-meter-readings-in-180-milliseconds/