Engineer Data

30 dagers strømdata på 180 ms

Utviklet varianter for timesaggregering av strømforbruk i Python, SQL og Bash + Jq; Python-varianten behandlet ett JSONL-datasett for 30 dager på 180 ms i ytelsestesten.

2025

180 ms

for å aggregere en måned med måleravlesninger

Situasjon. Selskapet behandler store mengder heterogene strømdata fra flere datakilder, hver med sin egen rapporteringsfrekvens — fra timesintervaller til 15-minuttersintervaller, og i noen tilfeller uregelmessige tidsstempler. Denne variasjonen gjør det utfordrende å lage et sammenhengende, sammenlignbart datasett. For å støtte presise energianalyser må disse dataene normaliseres til konsistente forbruksverdier per time, aggregert per sone.

Oppgave. Målet var å ta rå hendelsesdata fra et historisk datasett (i JSONL-format) og konvertere dem til strømforbrukstall justert til hele timer, strukturert som én rad per time og per sone. Konkret innebar oppgaven:

- Justere tidsstemplede data til strenge timesintervaller,
- Aggregere total strømproduksjon innenfor hvert intervall ved å summere mix-verdiene,
- Ta med utveksling over landegrensene ved å legge til import og trekke fra eksport,
- Levere de endelige verdiene i et strukturert, skalerbart format egnet for videre analyse.
- Løsningen måtte også være effektiv nok til å skalere til store tidsvinduer (30+ dager) og på tvers av flere land/soner.

Handling. For å løse oppgaven ble tre ulike varianter av løsningen implementert med Python, JQ (for JSON-behandling på kommandolinjen) og SQL, hver optimalisert for ulike bruksområder:

Python ble valgt for sin fleksibilitet, enkle datamanipulering og evne til å håndtere transformasjoner i minnet effektivt. Med rene Python-funksjoner ble det bygget en pipeline som:

- Parset JSONL-filene til strukturerte datarammer
- Samplet tidsseriedataene om til timesintervaller
- Aggregerte produksjonen og beregnet netto strømforbruk (produksjon + import – eksport)
- Eksporterte resultatene som CSV eller lastet dem inn i en lettvekts SQLite-database for inspeksjon.

JQ ble brukt til å lage en rask løsning med minimale avhengigheter for kommandolinjemiljøer, bygget som en kjede av JQ-filtre.

I PostgreSQL ble JSONL-dataene importert, normaliserte tabeller opprettet og en serie SQL-spørringer skrevet.

For å evaluere ytelsen ble hver løsning benchmarket med historiske datasett på både 1 dag og 30 dager.

Resultat. Python-implementasjonen viste seg å være den raskeste og mest skalerbare løsningen, og fullførte:

- Behandling av data for 1 dag på bare 34 ms
- Datasettet for 30 dager på 180 ms

Dette bekreftet at den egnet seg til å håndtere lengre tidsrom og samtidig holde behandlingstiden under ett sekund. Den ga også tydelig, vedlikeholdbar kode som lett kunne utvides til behandling av flere soner eller integreres i en ETL-pipeline.

Samlet sett viste tilnærmingen med flere verktøy fleksibilitet i verktøybruken, sterk ytelsesoptimalisering og robust håndtering av utfordringer med tidsjustering og aggregering i reelle datapipelines for strømdata.

Bruk erfaringen i prosjektet deres

Bruk «30 dagers strømdata på 180 ms» som utgangspunkt for en lignende utfordring i organisasjonen deres. Tjenestene bak arbeidet er tilgjengelige for prosjektet deres. Finne spørringene og de manglende indeksene bak et system som har blitt tyngre for hvert år, og rette dem som betyr noe.

Se tjenestene våre Ta kontakt →

Se bak resultatet

Les om situasjonen, beslutningene og resultatet i eksemplet. Vi kan drøfte hva som passer deres rammer, og hva som krever en annen tilnærming. 30 dagers strømdata på 180 ms — 180 ms for å aggregere en måned med måleravlesninger

Se tjenestene våre Ta kontakt →

Trenger du raske og effektive datapipelines?

Snakk om prosjektet med ingeniøren som har ansvaret for leveransen. Vi avtaler omfang, prisgrunnlag og overlevering skriftlig før arbeidet begynner.

Rolle i oppdraget: Plattform- og dataingeniør. Det bygger på Databaser, PostgreSQL, SQL, Data engineering, Datapipelines (ETL/ELT), Dataanalyse, Automatisering og CI/CD, Ytelsesoptimalisering og Python. Og nå som dere er her: vi er ganske gode på Optimalisering av databaseytelse, Utvikling av datapipelines (ETL/ELT) og Dataanalyse og BI-dashbord.

Ta kontakt Se tjenestene våre

Ett stykke av **en lengre oversikt** — og den ligger i sin helhet her.

Hver oppføring er skrevet på samme måte: situasjonen, oppgaven, hva vi gjorde, og hva som endret seg.

Se hele oversikten

Definer forbedringen dere trenger

Ta med et representativt datasett, spørsmålene det skal besvare og dagens flaskehals. Vi kan avtale målinger av kvalitet, gjennomstrømning og gjenoppretting.

Se hva det koster Ta kontakt →

Trenger du raske og effektive datapipelines?

Spør om beslutningene bak «30 dagers strømdata på 180 ms». Dere snakker med folk som kan forklare arbeidet og vurdere deres eget prosjekt.

Se tjenestene våre

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

Ta med deg denne siden

Visittkort

Laget med kjærlighet av D.

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

Opphavsrett © 2025 – nå · Engineer ApS · Alle rettigheter forbeholdt.

Skrevet ut fra https://data.engineer.company/nb/portfolio/a-month-of-meter-readings-in-180-milliseconds/