



GraphQL og "Cost"



Klimadatastyrelsen



Konklusion fremhævet på mødet

- Cost modellen indgår som et element i den styring KDS udfører i forhold til Datafordelerens driftsleverandør
- Cost modellen giver information til anvendere om den forventede svartid på de tre kategorier af forespørgsler
 1. Simpel (100ms)
 2. Almindelig (1.000ms)
 3. Kompleks (2.500ms)
- **Cost modellen begrænser ikke anvenderes design af forespørgsler**
- Begrænsninger i design af forespørgsler udgøres af antal joints per opslag:
 - › Max 7 joins i **dybden**
 - › Max 17 joins **totalt**

HVAD ER "COST"?

"Cost" er en model, der er etableret med henblik på systematisk at kunne vægte og udtrykke kompleksiteten i en GraphQL-forespørgsel

"Cost"-modellen kan anvendes

- til at sammenligne inkompatible GraphQL-forespørgsler, idet modellen lægger et raster af faste operationelle kategorier ned over den enkelte forespørgsel
- til at værdisætte en GraphQL-forespørgsels indvirkning på driftskapaciteten
- til relativt at fastsætte den enkelte GraphQL-forespørgsels andel i driftsomkostningerne
- til at kunne prognosticere svartiden på en GraphQL-forespørgsel
- til at fastlægge SLA-mål for forskelligartede GraphQL-forespørgsler med ensartet cost

HVORDAN BEREGNES "COST" AF EN FORESPØRGSEL?

IBM har udarbejdet en GraphQL "cost"-vejledning. Vejledningen kan findes her:

<https://ibm.github.io/graphql-specs/cost-spec.html>

Vejledningen skal forstås som en guide, der definerer operationer, der kan give nyttig indsigt i forhold til at kunne udtrykke "kosten" ved at udføre GraphQL-forespørgsler

Vejledningen er anvendt som grundlag for "cost"-dannelsen i forbindelse med den Moderniserede Datafordeler.

COST-BEREGNING FOR DATAFORDELERENS GRAPHQL-TJENESTER MED FLEKSIBEL OPSLAGSLOGIK/FORMELSAMLING

Eksempel på formler til beregning af "cost" (uddrag af kontraktbilag vedr. servicemål)

For GraphQL-Tjenester med fleksibel opslagslogik er der overordnede to typer Forespørgsler: Forespørgsler der henter data gennem edges og Forespørgsler der henter data gennem nodes. Cost for Forespørgsler der henter data gennem edges beregnes som følger:

$$\text{Total Entitet cost} + \text{Total Filter Cost} + \text{paging cost} + \text{antal forespurgte resultater} * (\text{paging cost} + \text{Total Felt Cost})$$

Cost for Forespørgsler der henter data gennem nodes beregnes som følger:

$$\text{Total Entitet cost} + \text{Total Filter Cost} + \text{paging cost} + \text{antal forespurgte resultater} * \text{Total Felt Cost}$$

Beregning af Total Entitet cost, Total Filter cost og Total Felt cost er beskrevet nedenfor. Paging cost og antal forespurgte resultater er beskrevet i Tabel 9

Total Entitet cost er den samlede cost for den Entitet der efterspørges. Denne er enten lig med standardentitet cost eller stor Entitet cost. Standardentitet cost og stor Entitet cost er beskrevet i Tabel 9

Total Filter cost er den samlede cost for de filtre der er anvendes og beregnes som følger:

$$\begin{aligned} &\text{Antal almindelige filter} * \text{standardfilter cost} + \\ &\text{Antal geometriske filtre} * \text{geometrifilter cost} \end{aligned}$$

Standardfilter cost og geometrifilter cost er beskrevet i Tabel 9

COST-BEREGNING FOR DATAFORDELERENS GRAPHQL-TJENESTER MED FLEKSIBEL OPSLAGSLOGIK/FORMELSAMLING

Eksempel på formler til beregning af "cost" (uddrag af kontraktbilag vedr. servicemål)

For GraphQL-Tjenester med fleksibel opslagslogik er der overordnede to typer Forespørgsler: Forespørgsler der henter data gennem edges og Forespørgsler der henter data gennem nodes. Cost for Forespørgsler der henter data gennem edges beregnes som følger:

$$\text{Total Entitet cost} + \text{Total Filter Cost} + \text{paging cost} + \text{antal forespurgte resultater} * (\text{paging cost} + \text{Total Felt Cost})$$

Cost for Forespørgsler der henter data gennem nodes beregnes som følger:

$$\text{Total Entitet cost} + \text{Total Filter Cost} + \text{paging cost} + \text{antal forespurgte resultater} * \text{Total Felt Cost}$$

Beregning af Total Entitet cost, Total Filter cost og Total Felt cost er beskrevet nedenfor. Paging cost og antal forespurgte resultater er beskrevet i Tabel 9

Total Entitet cost er den samlede cost for den Entitet der efterspørges. Denne er enten lig med standardentitet cost eller stor Entitet cost. Standardentitet cost og stor Entitet cost er beskrevet i Tabel 9

Total Filter cost er den samlede cost for de filtre der er anvendes og beregnes som følger:

$$\begin{aligned} &\text{Antal almindelige filter} * \text{standardfilter cost} + \\ &\text{Antal geometriske filtre} * \text{geometrifilter cost} \end{aligned}$$

Standardfilter cost og geometrifilter cost er beskrevet i Tabel 9

COST-BEREGNING FOR DATAFORDELERENS GRAPHQL-TJENESTER MED FLEKSIBEL OPSLAGSLOGIK/VÆRDIER

Eksempel på værdier til beregning af "cost" (uddrag af kontraktbilag vedr. servicemål)

Kategorier	Cost type	Beskrivelse	Cost
Entitet	Standardentitet cost	Cost for en Entitet.	1.000
	Stor Entitet cost	Cost for at hente information fra en tabel med >= 10.000.000 rækker.	5.000
Felt	Standard felt cost	Standard cost for et felt.	1
	Geometri felt cost	Cost for geometriske felter.	10
	Sammensat felt basis cost	Basis cost for sammensatte felter, dvs. felter der har en type som består af et eller flere felter.	10
	Cost for felt med én-til-én relation	Cost for at hente et felt der repræsenterer en en-til-en relation (hver Entitet relaterer til præcis én anden Entitet).	5.000
	Cost for felt med én-til-mange relation	Cost for at hente et felt der repræsenterer en en-til-mange relation (en enkelt Entitet relaterer til flere andre Entiteter).	20.000
	Cost for felt med én-til-én relation til store tabeller	Cost for at hente et felt der repræsenterer en relation til en stor tabel med >= 10.000.000 rækker.	30.000
	Cost for felt med relation til Entiteter med kompleks bitemporalitet	Cost for at hente et felt der repræsenterer en relation til en Entitet med en Komplex Bitemporalitet.	40.000
Filter	Standardfilter cost	Cost anvendt for filtre.	2.000
	Geometrifilter cost	Cost for filtre anvendt på geometrifelter.	5.000
Operator	Cost for lette operatorer	Cost for operatorerne eq, in.	1.000
	Cost for tunge operatorer	Cost for operatorerne gt, ngt, gte, ngte, lt, nlt, lte, nlte	4.000
	Paging cost	Basis-cost når "edges", "nodes" eller "node" er del af Forespørgslen.	10

COST-BEREGNING FOR DATAFORDELERENS GRAPHQL-TJENESTER MED FLEKSIBEL OPSLAGSLOGIK/VÆRDIER

Eksempel på værdier til beregning af "cost" (uddrag af kontraktbilag vedr. servicemål)

Kategorier	Cost type	Beskrivelse	Cost
Entitet	Standardentitet cost	Cost for en Entitet.	1.000
	Stor Entitet cost	Cost for at hente information fra en tabel med >= 10.000.000 rækker.	5.000
Felt	Standard felt cost	Standard cost for et felt.	1
	Geometri felt cost	Cost for geometriske felter.	10
	Sammensat felt basis cost	Basis cost for sammensatte felter, dvs. felter der har en type som består af et eller flere felter.	10
	Cost for felt med én-til-én relation	Cost for at hente et felt der repræsenterer en en-til-en relation (hver Entitet relaterer til præcis én anden Entitet).	5.000
	Cost for felt med én-til-mange relation	Cost for at hente et felt der repræsenterer en en-til-mange relation (en enkelt Entitet relaterer til flere andre Entiteter).	20.000
	Cost for felt med én-til-én relation til store tabeller	Cost for at hente et felt der repræsenterer en relation til en stor tabel med >= 10.000.000 rækker.	30.000
	Cost for felt med relation til Entiteter med kompleks bitemporalitet	Cost for at hente et felt der repræsenterer en relation til en Entitet med en Komplex Bitemporalitet.	40.000
Filter	Standardfilter cost	Cost anvendt for filtre.	2.000
	Geometrifilter cost	Cost for filtre anvendt på geometrifelter.	5.000
Operator	Cost for lette operatører	Cost for operatørerne eq, in.	1.000
	Cost for tunge operatører	Cost for operatørerne gt, ng, gte, ngte, lt, nlt, lte, nlte	4.000
	Paging cost	Basis-cost når "edges", "nodes" eller "node" er del af Forespørgslen.	10



Cost-beregning og performance

Forespørgsler placeres indenfor en af følgende tre kategorier

1. Simpel (100ms)
 2. Almindelig (1.000ms)
 3. Kompleks (2.500ms)
- Følgende anvendelse er de mest indflydelsesrige faktorer ift. beregning af cost:
 - › Antal resultater, der forespørges om
 - › Filtrering på geometrier
 - › Generel filtrering
 - › Join-dybde

"COST" SOM GRUNDLAG FOR SERVICEMÅL

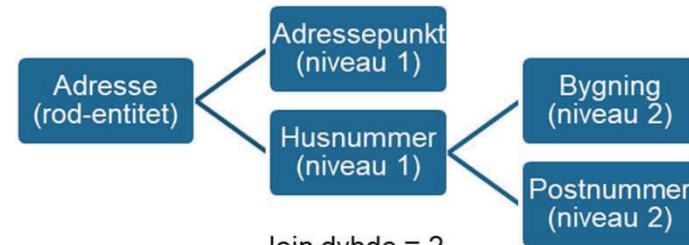
Forespørgsels-kategori	Beskrivelse/evt. forudsætninger	Opfyldelsesgrad i %	Maks. svar-tid i sek.
Simpel (S)	Forespørgsel der, ud fra inputparameter entydigt identificerer en specifik Entitet, fremsøger og returnerer denne, som ikke medtager præ-autentificerede credentials.		
	S1: Beregnet cost er i intervallet 0 op til 150.000 og der dannes ikke en relation til en Entitet med Komplex Bitemporalitet og Forespørgslen oprinder ikke i en Entitet med Komplex Bitemporalitet.	90	100 ms
Almindelig (A)	Forespørgsel der, ud fra inputparametre, fremsøger data fra flere Entiteter og returnerer disse, samt opslag der går på tværs af Registre, hvor opslag sker med nøgleværdier.		
	A1: Beregnet cost er i intervallet 150.000 op til 100.000.000.000 og der dannes ikke relation til en Entitet med Komplex Bitemporalitet og Forespørgslen oprinder ikke i en Entitet med Komplex Bitemporalitet.	90	1,0
	A2: Beregnet cost er i intervallet 0 op til 190.000 og der dannes relation til en Entitet med Komplex Bitemporalitet og Forespørgslen oprinder ikke i en Entitet med Komplex Bitemporalitet.	90	1,0
Kompleks (K)	Forespørgsel der fremsøger afledte data ud fra angivne inputparametre, og hvor resultatet er en sammensat mængde af data, eller hvor data fremsøges fra flere forskellige Registre.		
	K1: Beregnet cost er over 100.000.000.000 og der dannes ikke en relation til en Entitet med Komplex Bitemporalitet og Forespørgslen oprinder ikke i en Entitet med Komplex Bitemporalitet.	90	2,5
	K2: Beregnet cost er over 190.000 og der dannes relation til en Entitet med Komplex Bitemporalitet og Forespørgslen oprinder ikke i en Entitet med Komplex Bitemporalitet.	90	2,5

Begrænsning af kompleksitet

- Der er etableret en begrænsning på antallet af mulige joins pr. opslag
 - › Max 7 joins i **dybden**
 - › Max 17 joins **totalt**
- Join dybde: antal niveauer af relationelle forbindelser mellem entiteter i en enkelt forespørgsel



Join dybde = 2
Joins totalt = 2



Join dybde = 2
Joins totalt = 4

- Begrænsning af mulige joins medfører en implicit grænse for mulig cost pr. Opslag
 - › Performance testens opslag med højeste cost: 24.917.441.470.472.410.000.000.000