

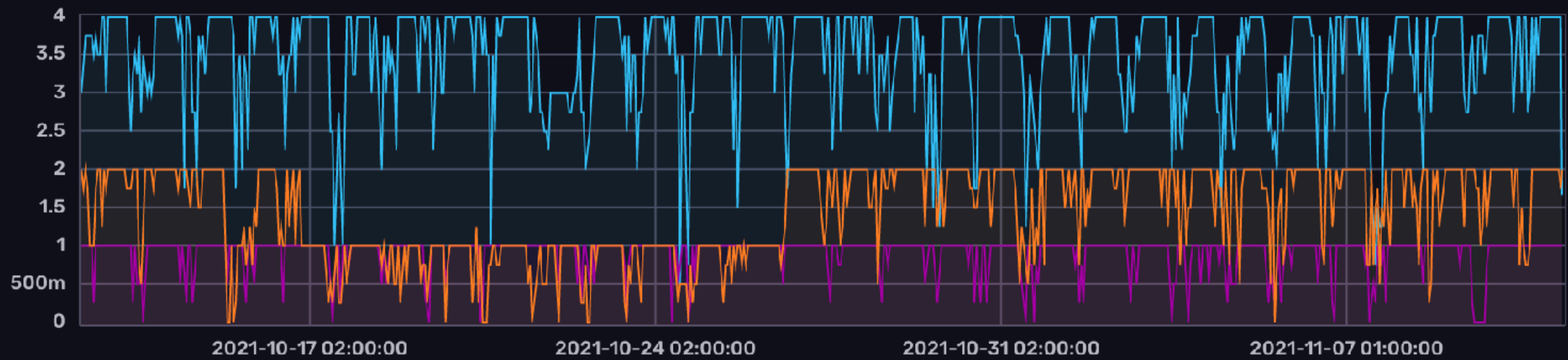
MONITORING IN THE CLOUD

ERIK LUPANDER

CADEC 2022.02.02 | CALLISTAENTERPRISE.SE

CALLISTA

\$1 +



MIDDLE AGED MEN...

CROSSFIT



CC: [HTTPS://WWW.FLICKR.COM/PHOTOS/RUNARE/13472386673](https://www.flickr.com/photos/runare/13472386673)

CYCLING



PADEL



ELECTRIC CARS



CC: WIKIMEDIA

ELECTRIC CARS

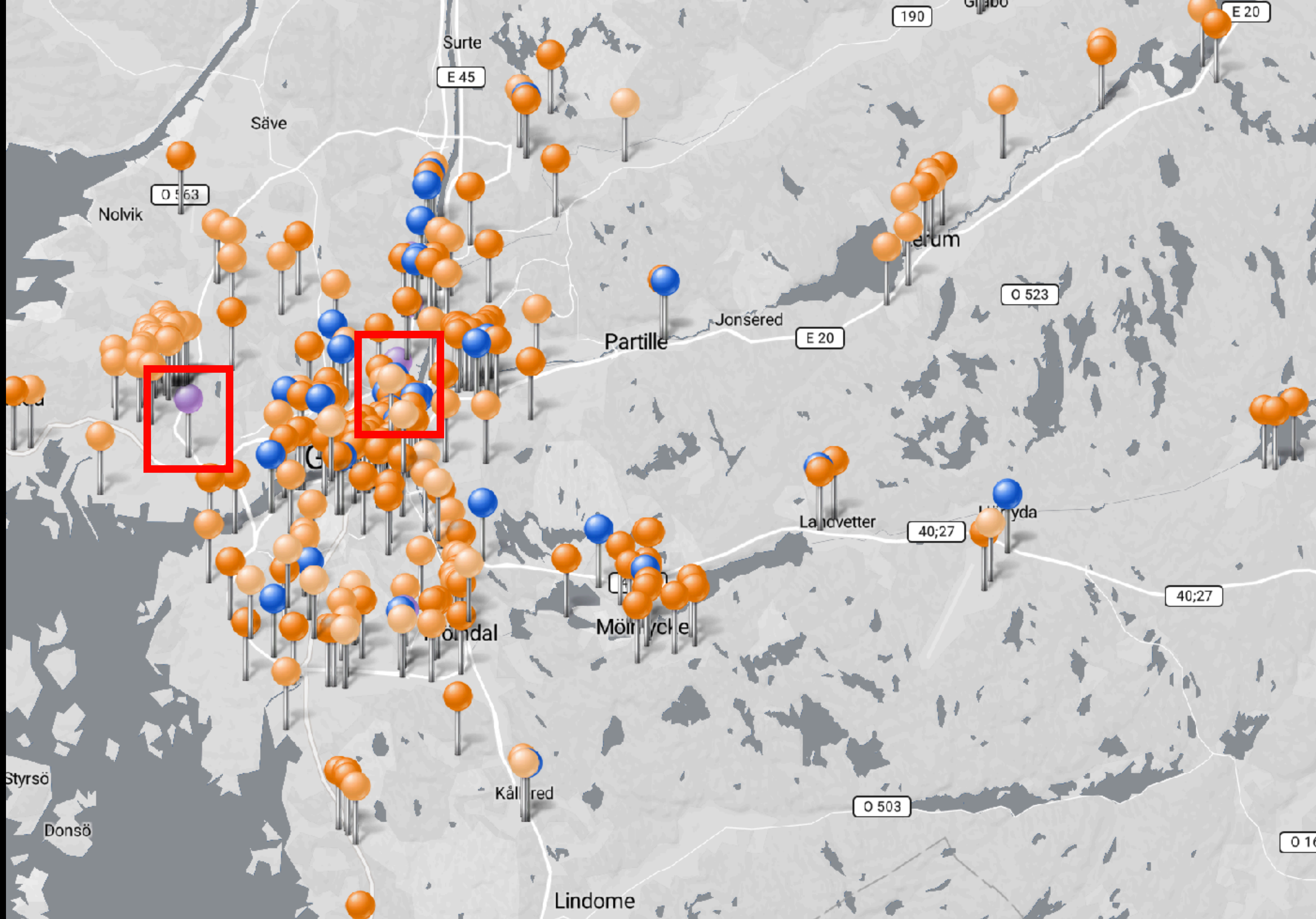


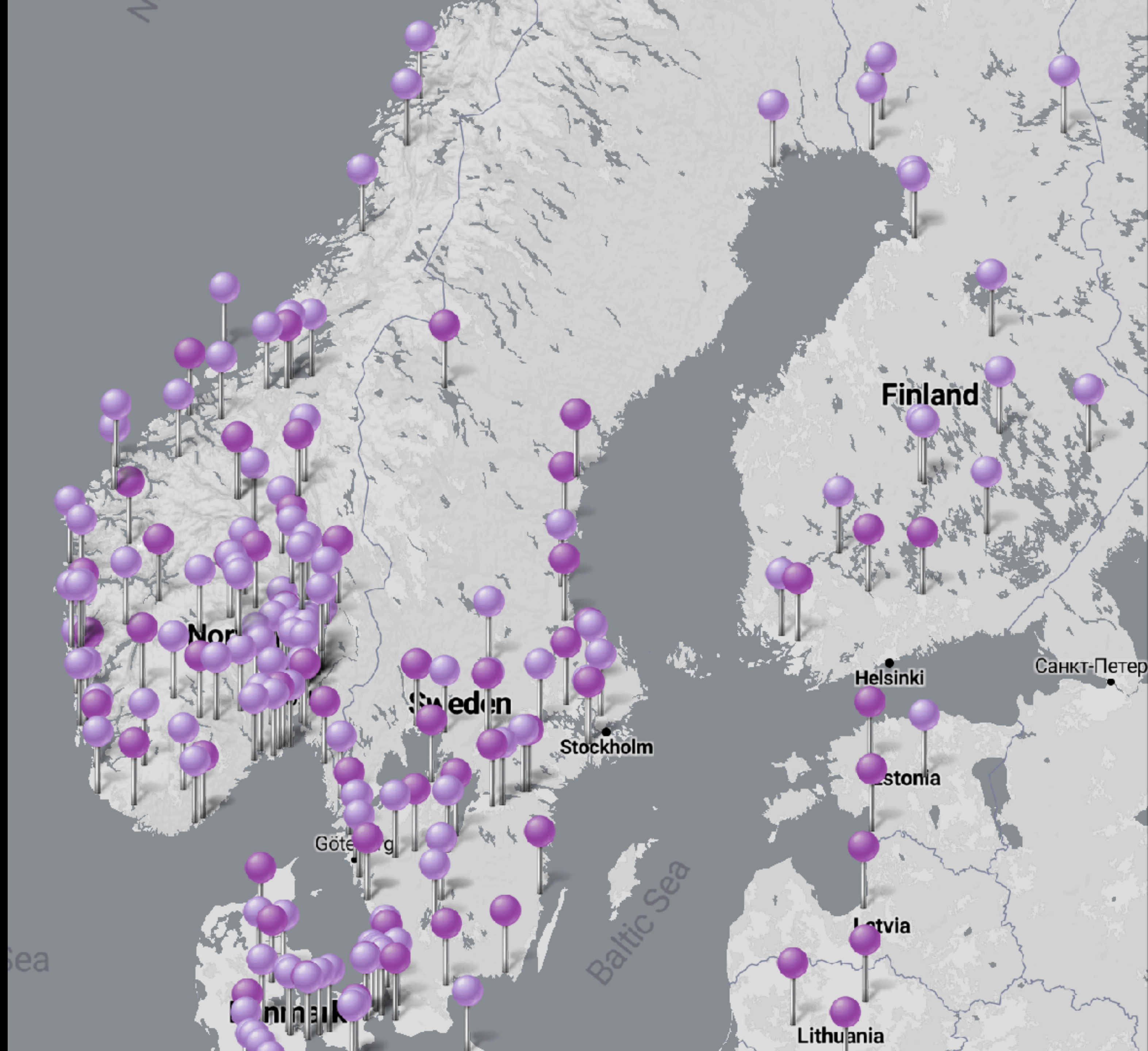
CC: MATTI BLUME

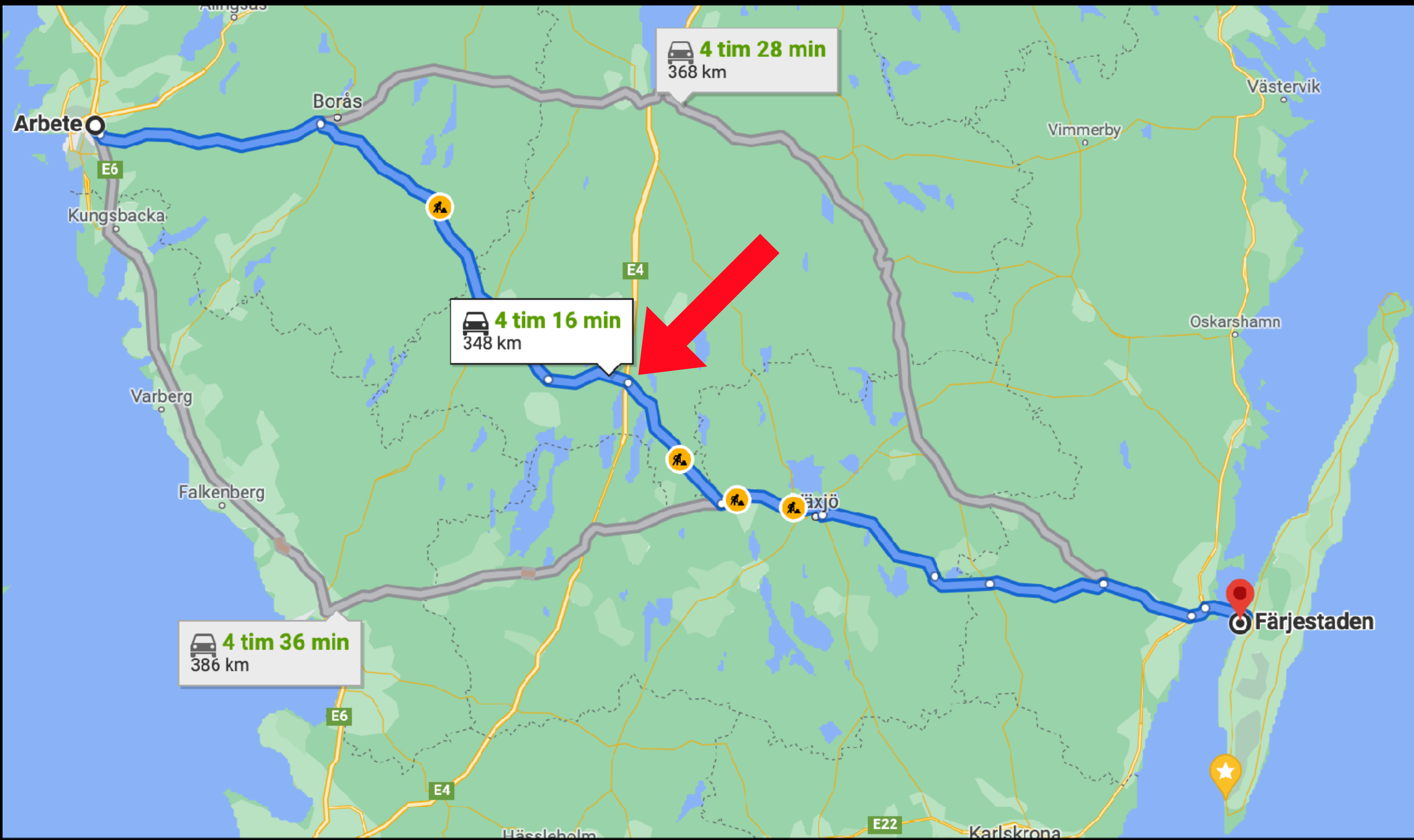
TODAY'S TOPIC:

**THE ELECTRIC CAR
CHARGING PROBLEM**











CCO: MAX PIXEL

HOW TO AVOID CHARGING QUEUES?



CALLISTA



CHARGER AVAILABILITY METRICS!



AND SO OUR JOURNEY BEGINS...

WHERE IS THE DATA??

IONITY Mariestad

Ulriksdal, Mariestad, Sweden

Laddare

350 kW CCS

SE*ION*E303501

8,70 kr/kWh

SE*ION*E303502

8,70 kr/kWh

SE*ION*E303503

8,70 kr/kWh


SE*ION*E303504

8,70 kr/kWh


IONITY 

 **4 / 4**

 Tillgänglig

 Tillgänglig

 Tillgänglig

 Tillgänglig



SCREEN SCRAPING?!?!?



PUBLIC DOMAIN: FREESVG.ORG/PUKE-MAN

× Headers Preview Response Initiator Timing Cookies

▼ General

Request URL: https://adm.chargefinder.com/status/3qg2q

Request Method: GET

Status Code: ● 200

Remote Address: 52.16.172.234:443

Referrer Policy: strict-origin-when-cross-origin

× Headers Preview Response Initiator Timing Cookies

```
▼ [{id: "SE*ION*E303501", status: 2, price: "8,70 kr/kWh", free: null},...]  
  ▼ 0: {id: "SE*ION*E303501", status: 2, price: "8,70 kr/kWh", free: null}  
      free: null  
      id: "SE*ION*E303501"  
      price: "8,70 kr/kWh"  
      status: 2  
  ▶ 1: {id: "SE*ION*E303502", status: 3, price: "8,70 kr/kWh", free: null}  
  ▶ 2: {id: "SE*ION*E303503", status: 2, price: "8,70 kr/kWh", free: null}  
  ▶ 3: {id: "SE*ION*E303504", status: 2, price: "8,70 kr/kWh", free: null}
```

THE SOLUTION

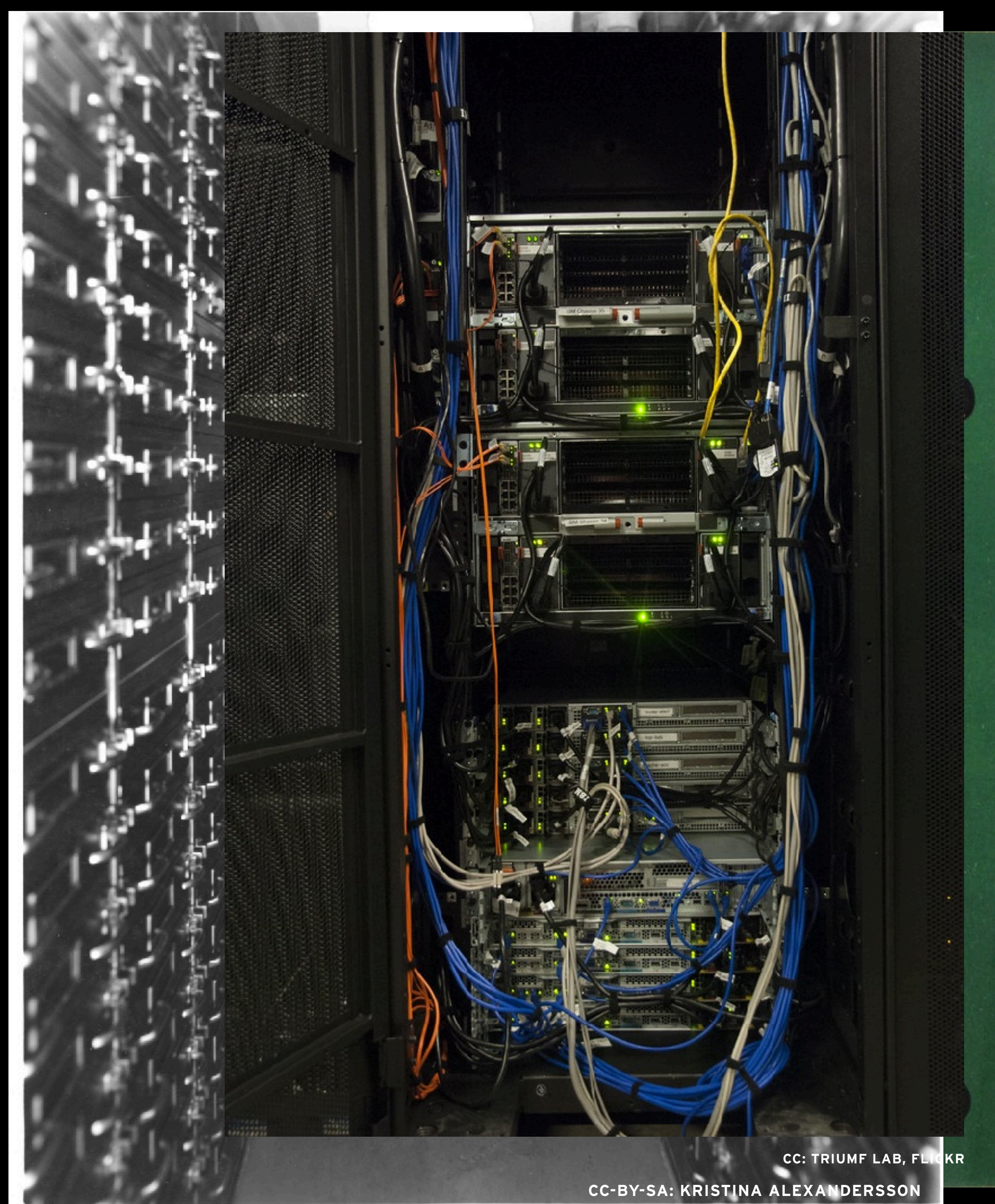
- Call Chargefinder's API every 15 minutes for the ~20 charging sites I'm interested in for the upcoming winter season.
- Store the data for later querying

 ChargeFinder



MY REQUIREMENTS

- Not too expensive...
- Not in my closet
- Long data retention
- Powerful querying



CC: TRIUMF LAB, FLIKR

CC-BY-SA: KRISTINA ALEXANDERSSON

USE MANAGED INFRASTRUCTURE IN THE CLOUD



AWS LAMBDA

THE OTHER PIECE...

CHOICE OF DATABASE

**WHAT KIND OF DATA ARE WE GOING
TO STORE?**

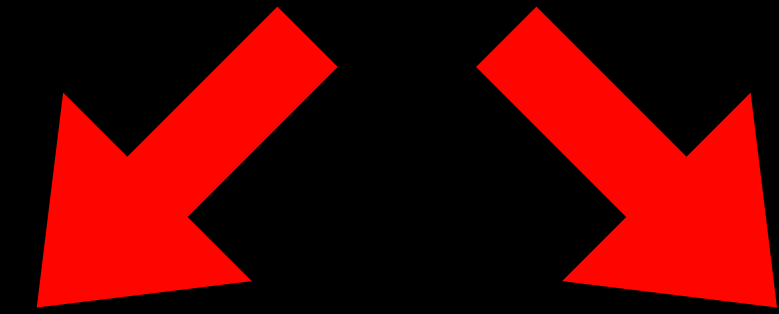
JSON FROM CHARGEFINDER...

...WHICH WE TRANSFORM...

...INTO DATA "ROWS"

Site	Time	Available
Ionity Mariestad	2021-11-12T12:00:00	3
Ionity Mariestad	2021-11-12T12:15:00	2
Ionity Mariestad	2021-11-12T12:30:00	4

INDEX AND QUERY FRIENDLY!



Site	Time	Available	Day of week	Hour of day
Ionity Mariestad	2021-11-12T23:45:00	3	Friday	23
Ionity Mariestad	2021-11-13T00:00:00	2	Saturday	0
Ionity Mariestad	2021-11-13T00:15:00	4	Saturday	0

**THIS LOOKS LIKE A
TIME SERIES**

TO ME!

TIME SERIES DATABASES



AWS TIMESTREAM



AZURE TIME SERIES
INSIGHTS

CHOICE OF TIME SERIES DATABASE

- Relational- and document databases often provide time-series storage as well
 - I decided to focus on dedicated Time-Series databases
- AWS TimeStream, AWS Managed Service for Prometheus and InfluxDB Cloud
 - Fully managed
 - Zero up-front cost
- InfluxDB Cloud offers very powerful querying through its Flux query and scripting language as well as many visualization types





+

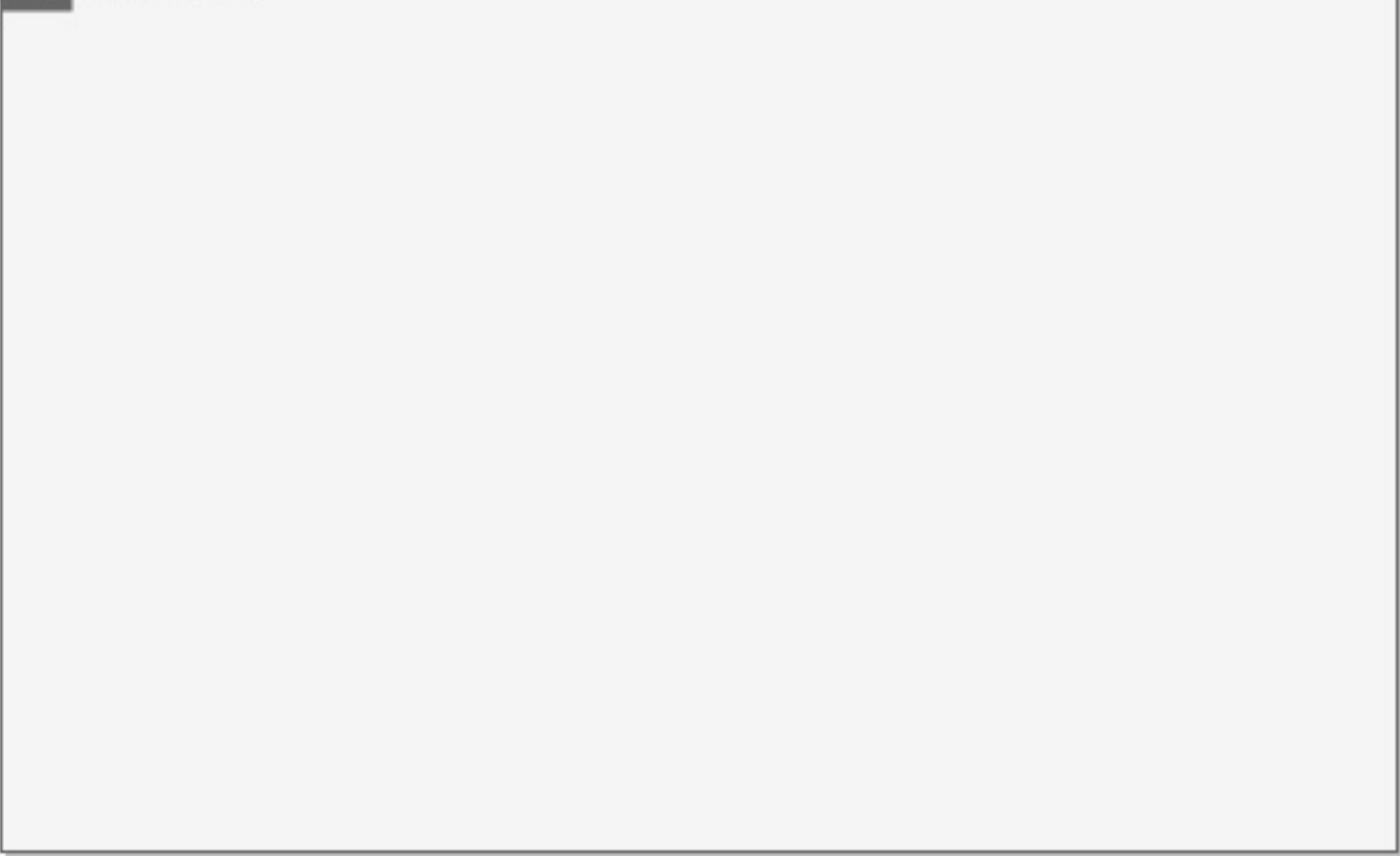


influxdb

ARCHITECTURE



AWS Cloud



CDK stack



AWS Cloud

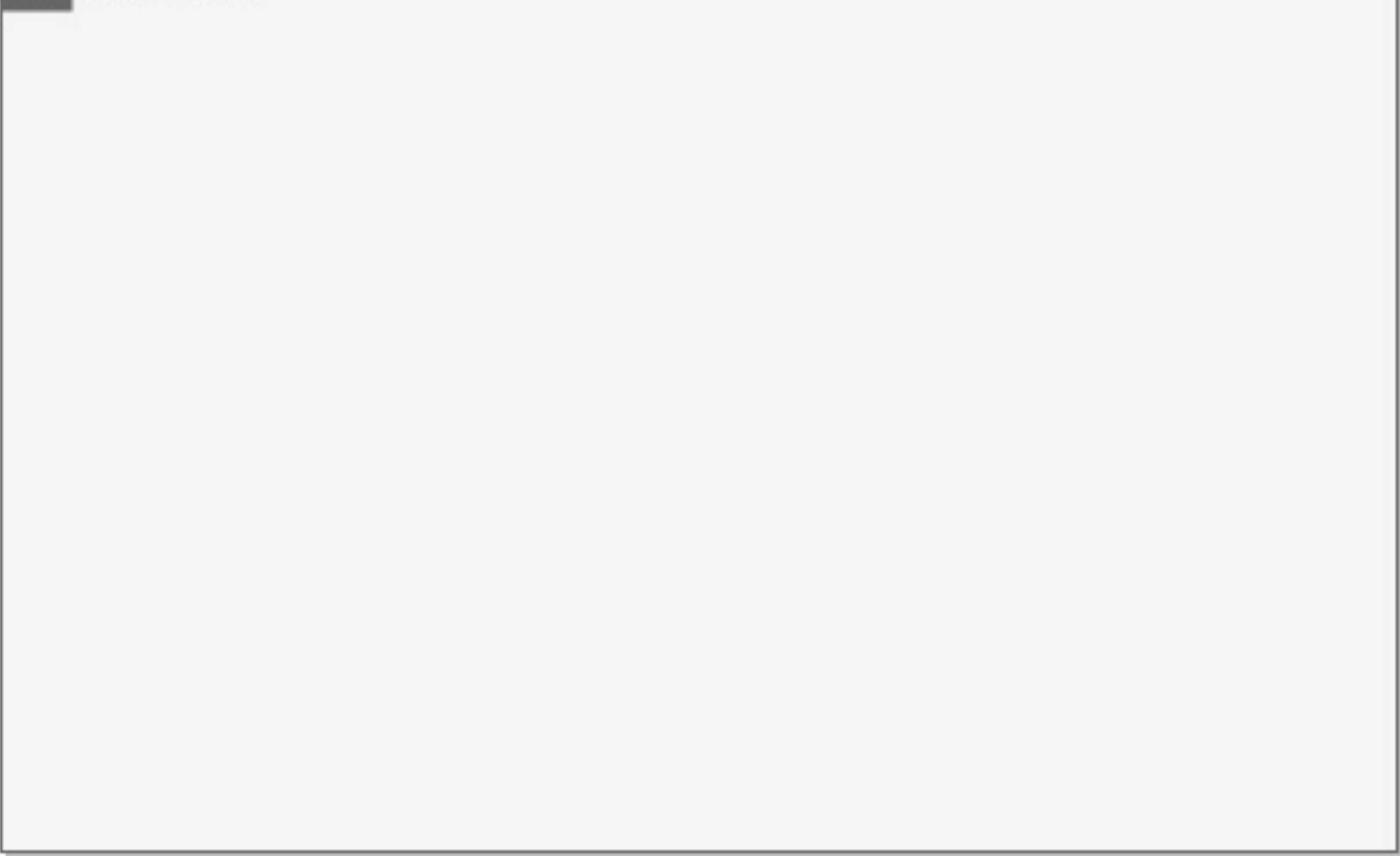
```
export class ChargerStatusStack extends cdk.Stack {  
  constructor(scope: cdk.Construct, id: string, props?: cdk.StackProps) {  
    super(scope, id, props);  
  }  
}
```



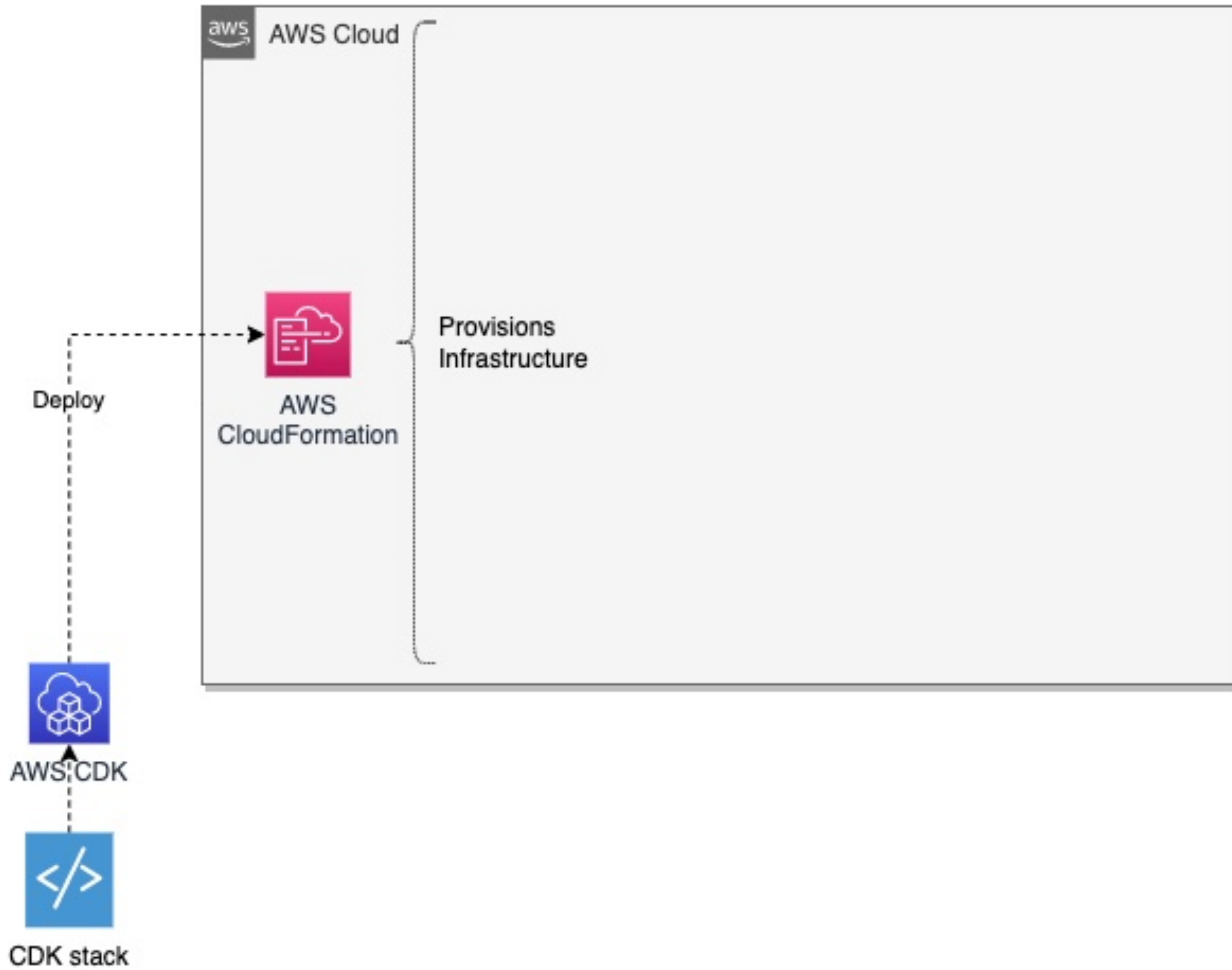
CDK stack

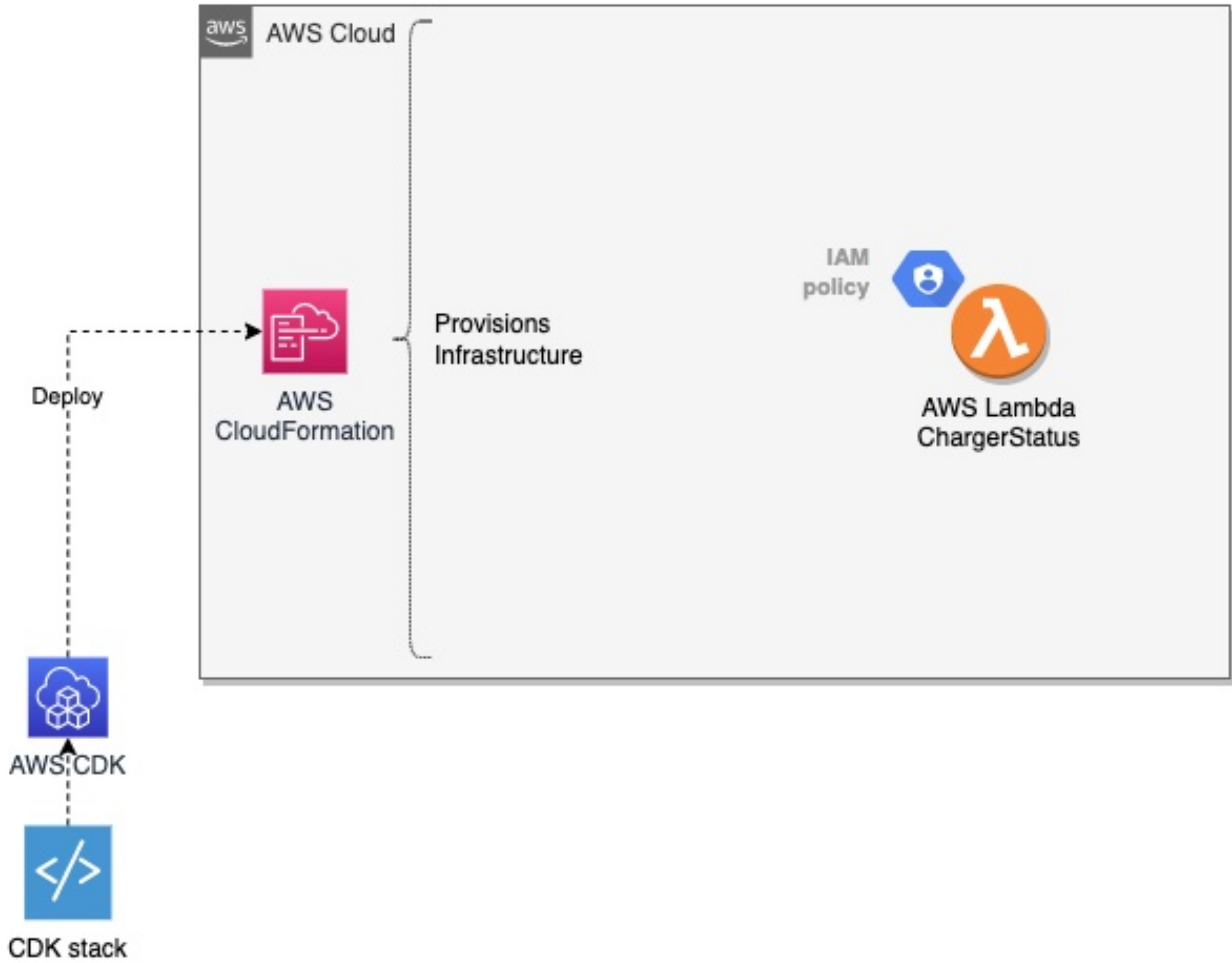


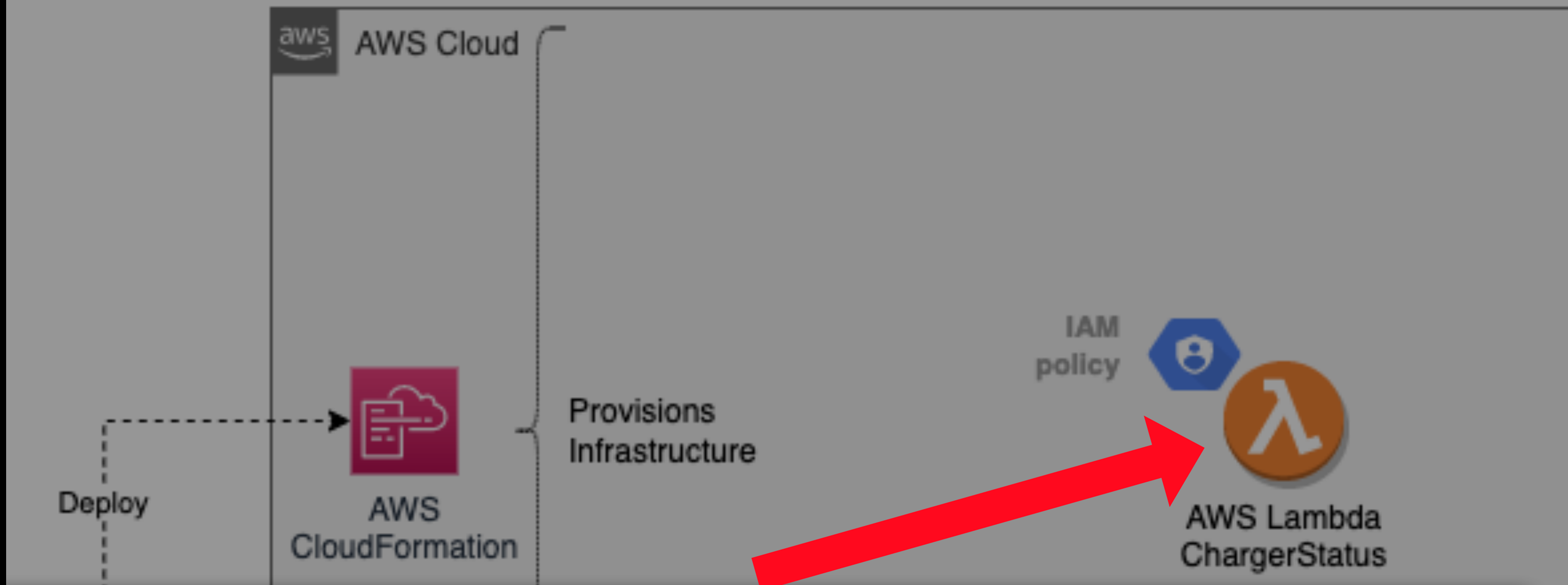
AWS Cloud



CDK stack



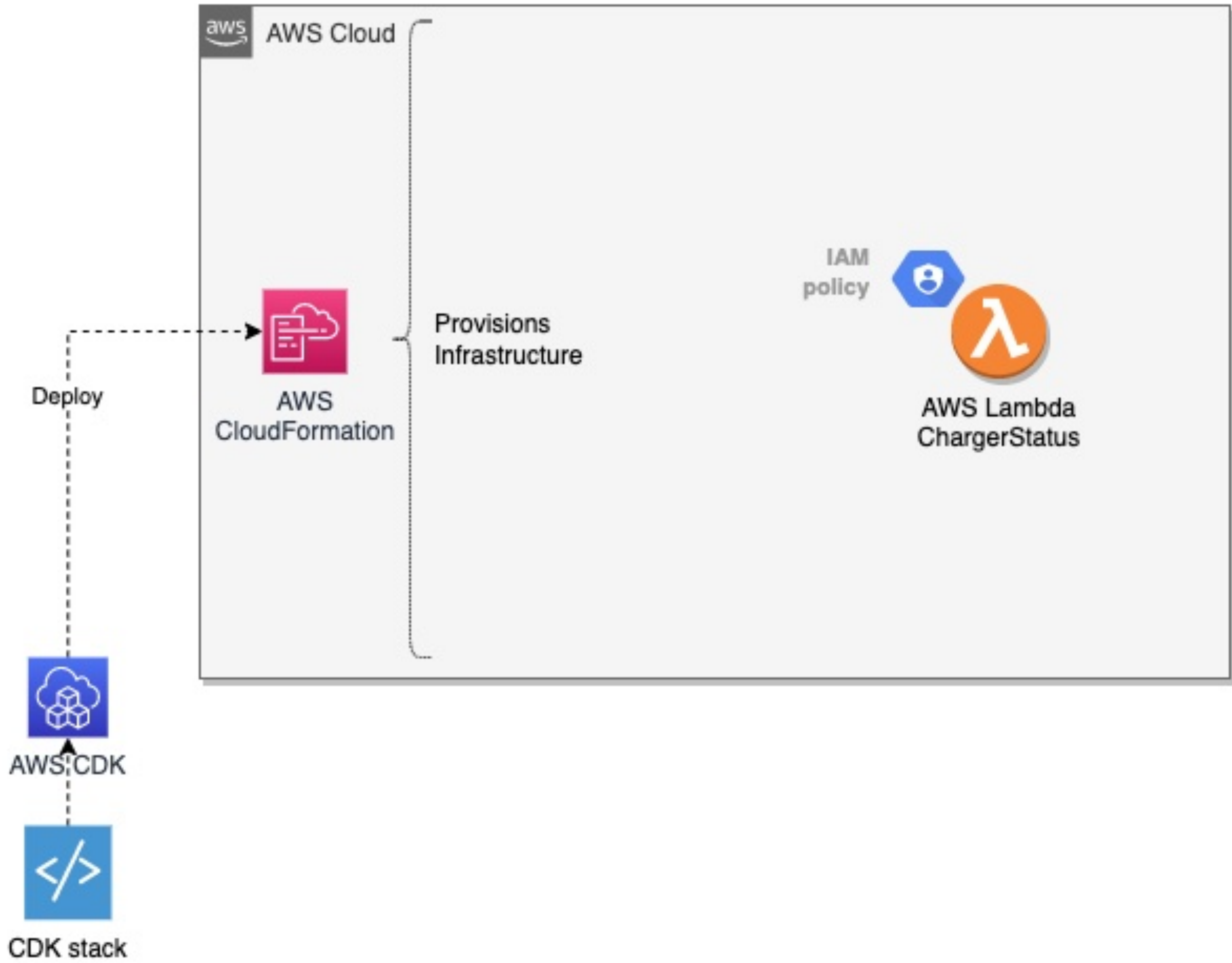


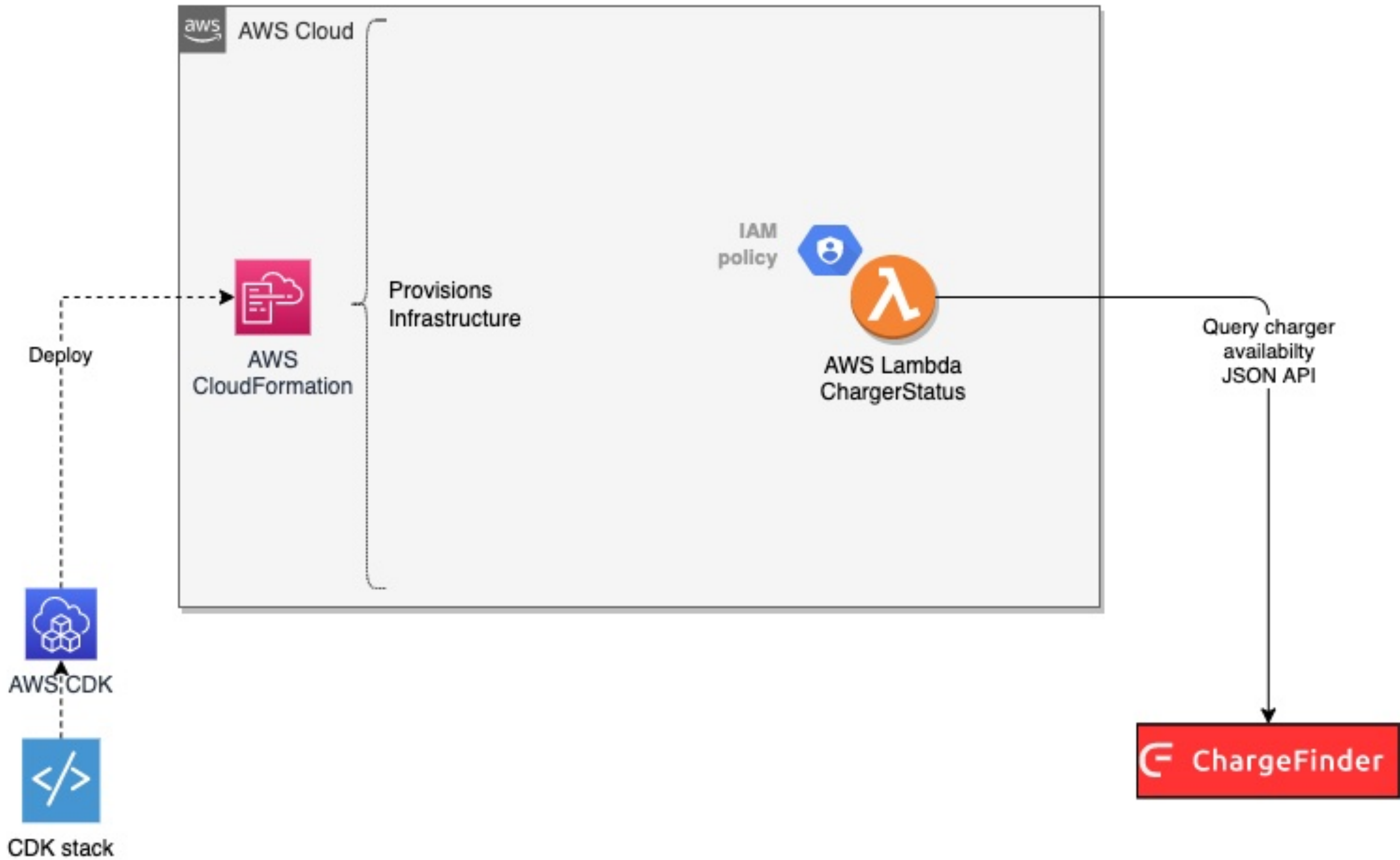


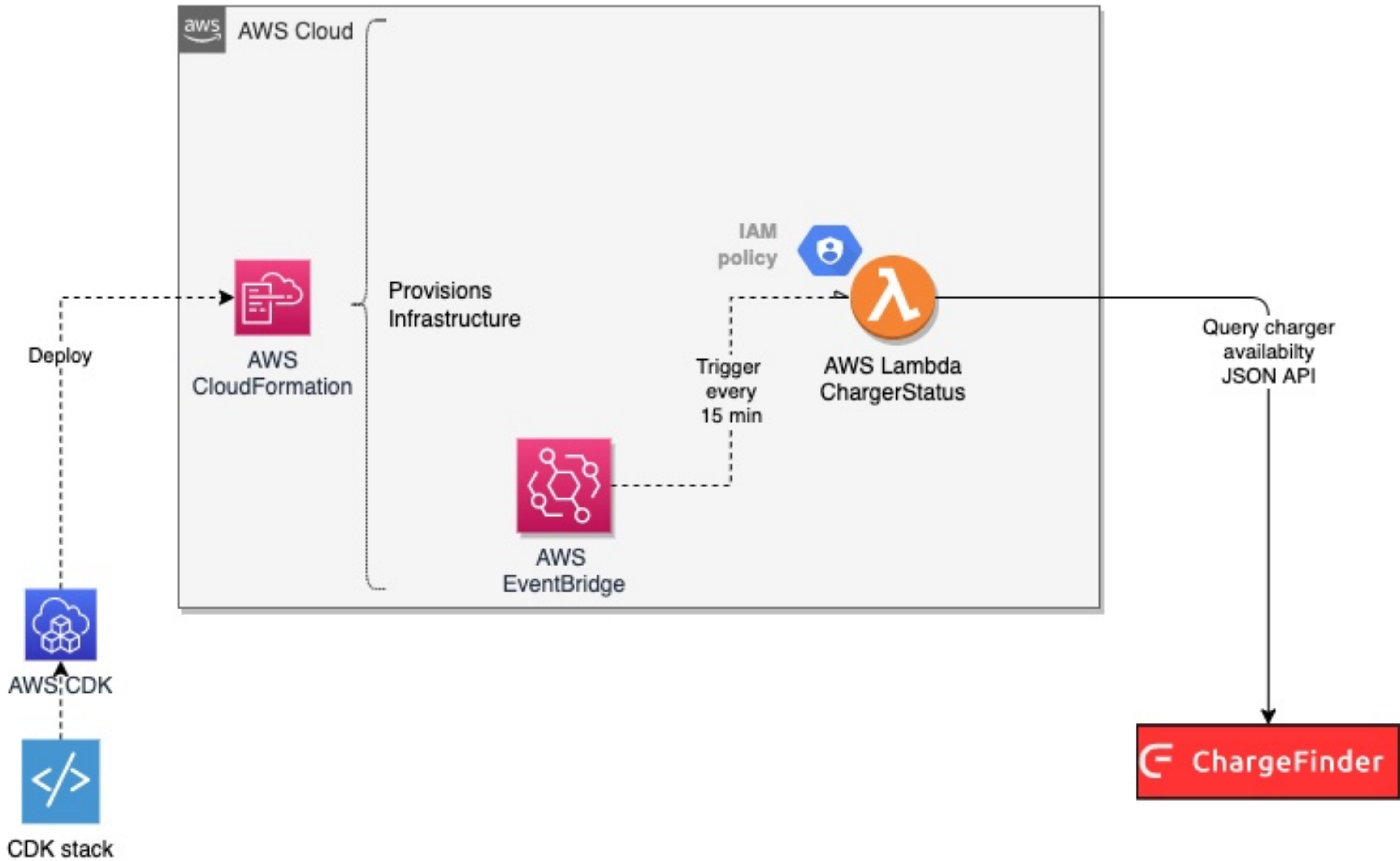
```
// main is called when a new lambda starts, so don't
// expect to have something done for every query here.
func main() {
    logrus.Info("init charger status recorder")

    // load secrets etc. Will panic on errors.
    configure()

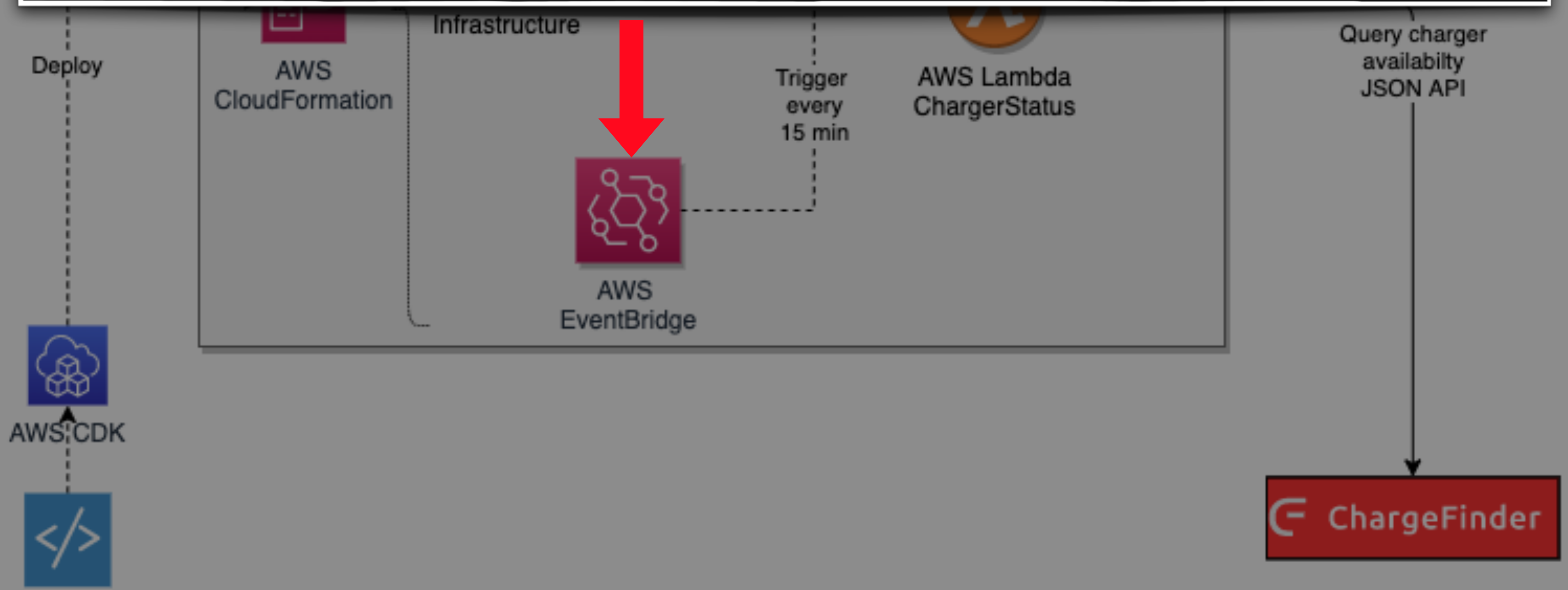
    lambda.StartWithContext(context.Background(), handler)
}
```

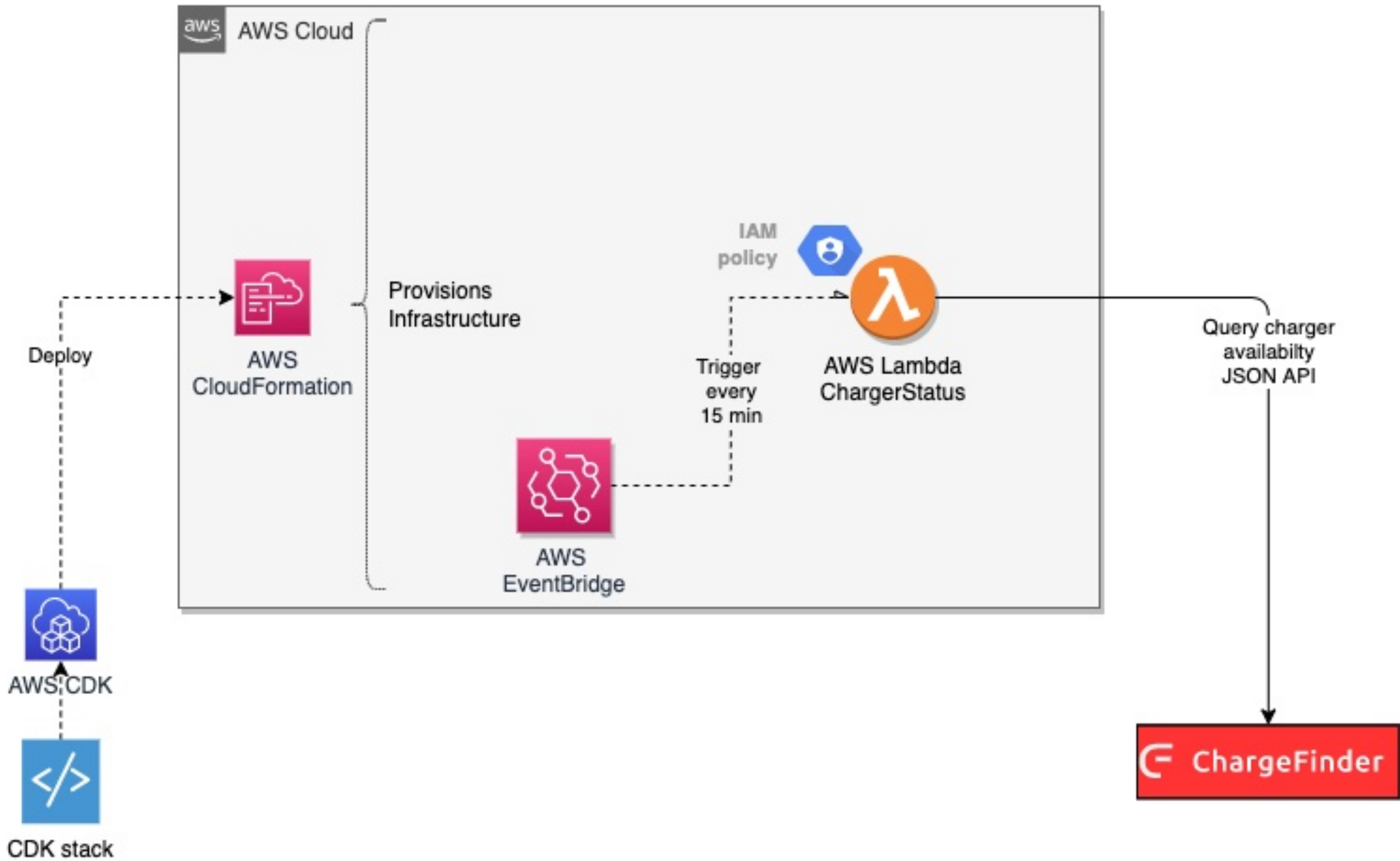


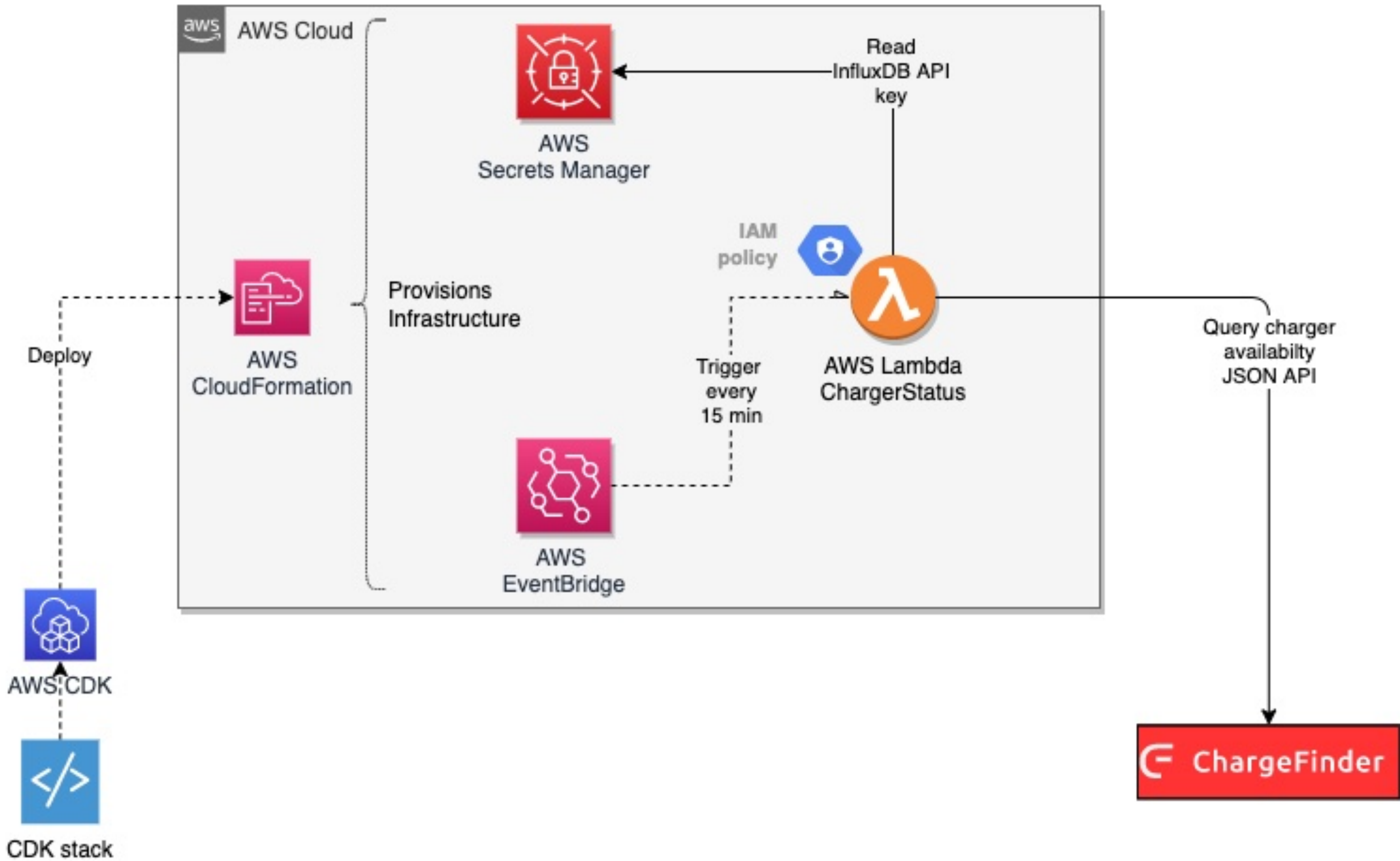


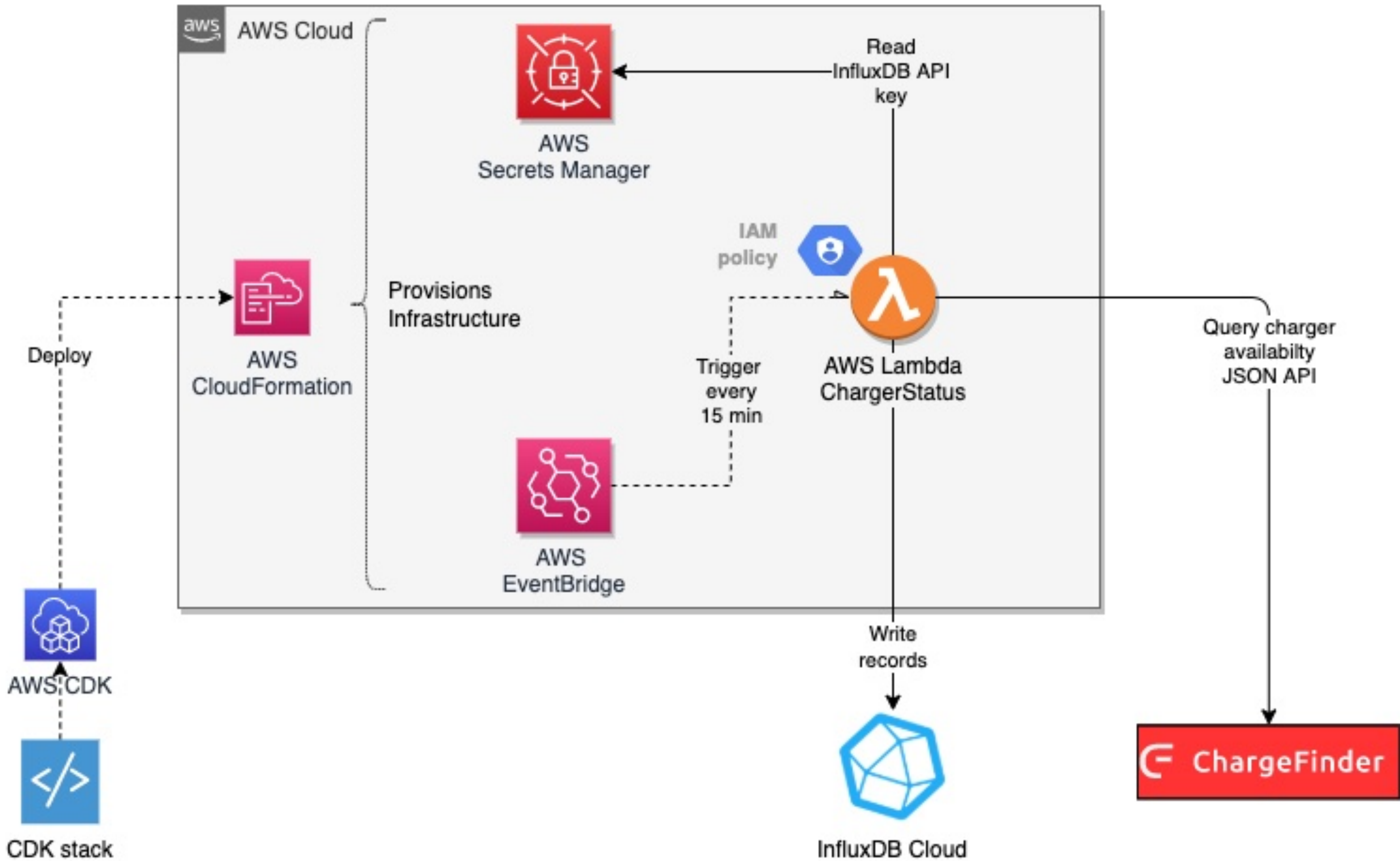


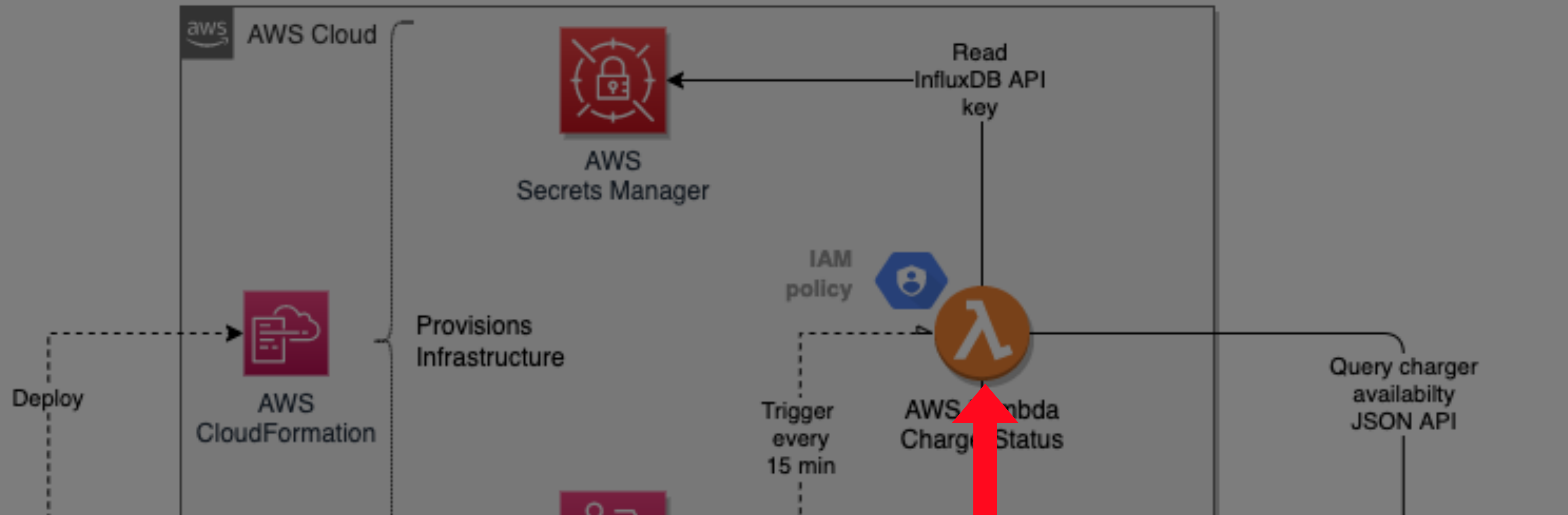
```
const rule = new ruleCdk.Rule(this, "collect_rule", {  
  description: "Invoke every 15 minutes",  
  schedule: Schedule.expression("cron(0/15 * * * ? *)")  
});  
rule.addTarget(new targets.LambdaFunction(chargerStatusLambdaFn))
```



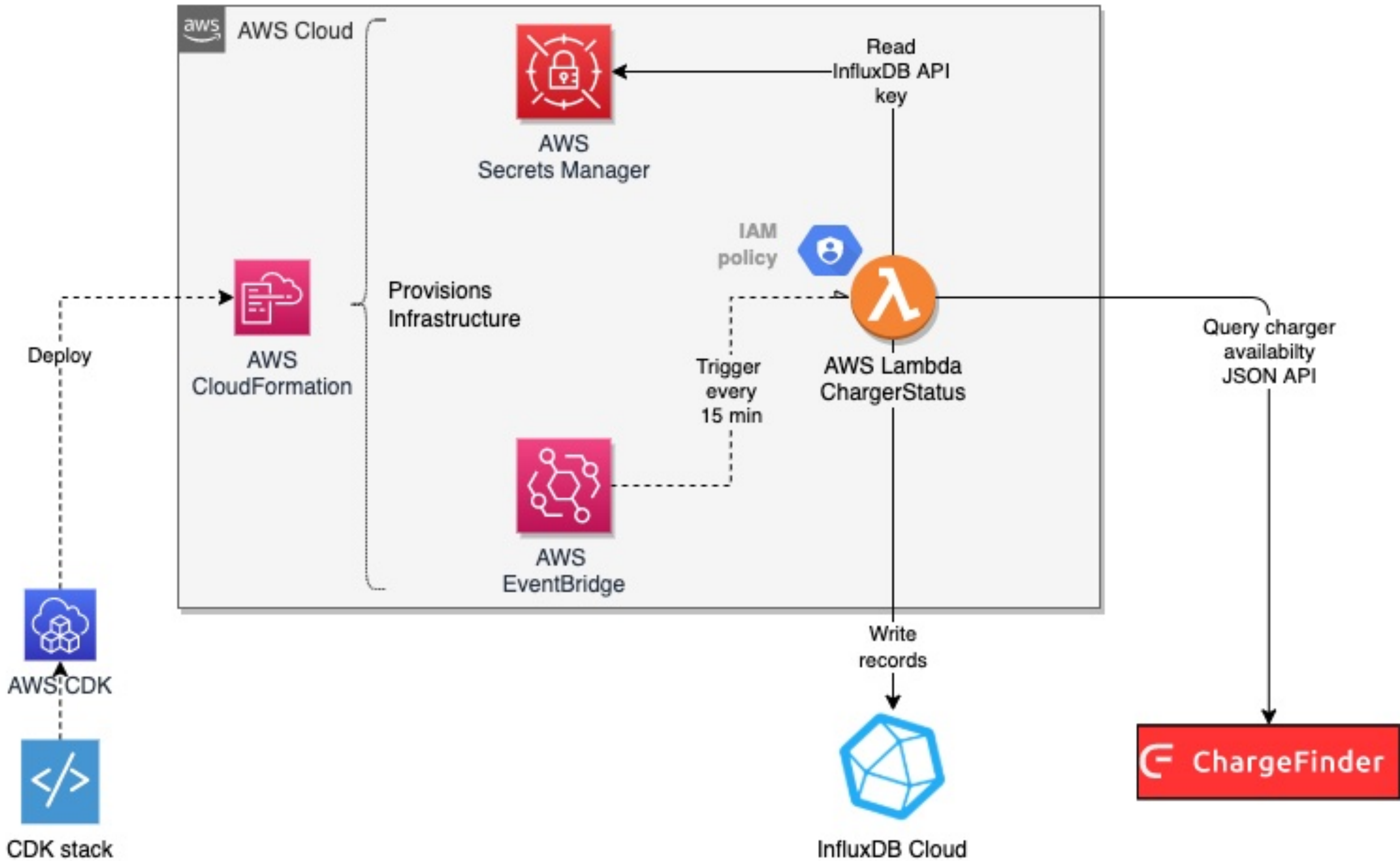








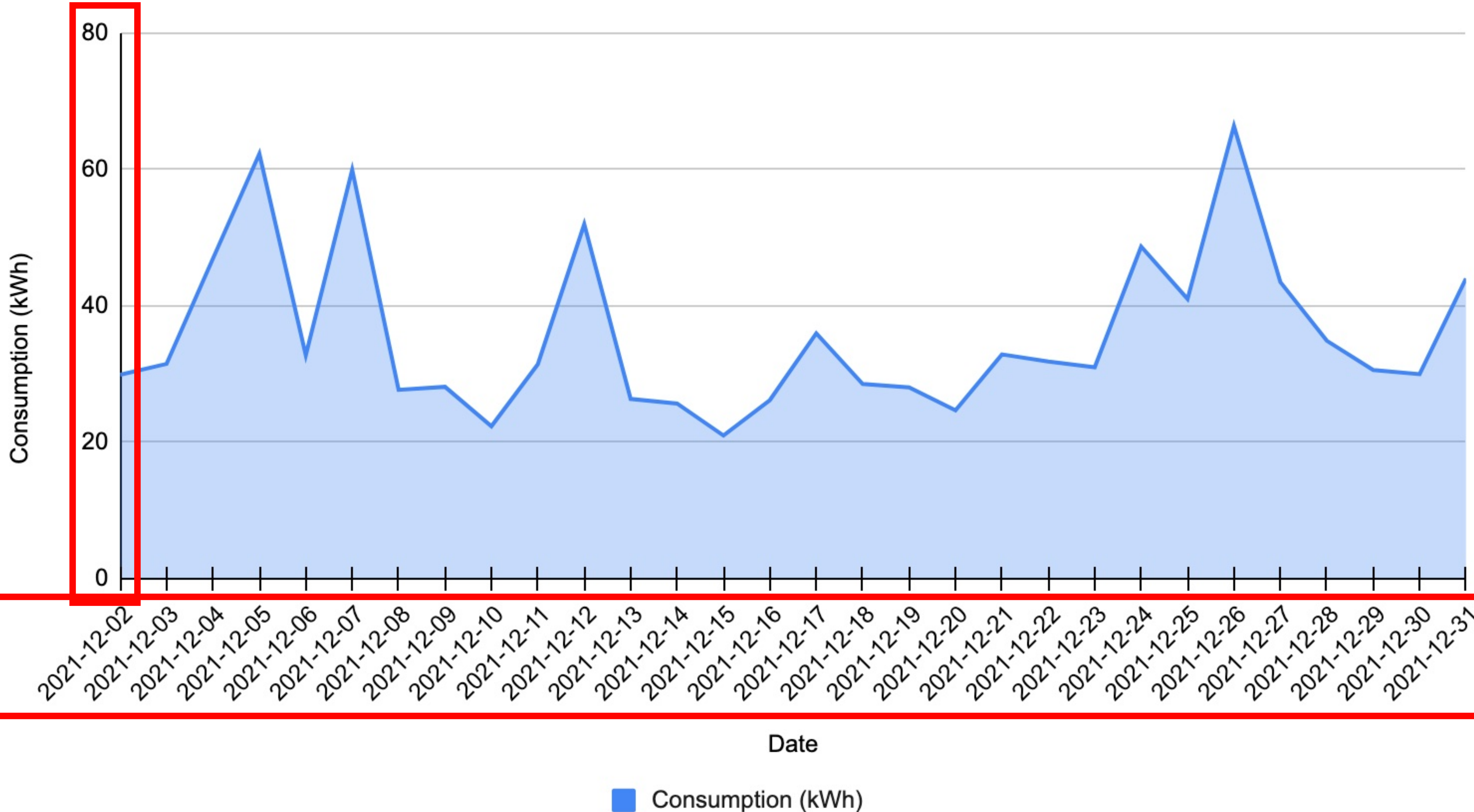
```
p := influxdb2.NewPointWithMeasurement("charger availability").  
AddTag("site", r.SiteName).  
AddTag("weekday", time.Now().Weekday().String()).  
AddTag("hour of day", strconv.Itoa(time.Now().Hour())).  
AddField("available", r.Available).  
SetTime(time.Now())  
  
iw.writeApi.WritePoint(p)
```



WHAT IS A TIME-SERIES DATABASE?

**TO ANSWER THIS,
WE NEED TO UNDERSTAND:**

WHAT IS A TIME SERIES?



WHAT IS A TIME SERIES?

- A **series name** describing the what we are **observing**:
 - Temperature, Stock quote, Number of available chargers, CPU load...
- However...
 - Which thermometer?
 - Which stock?
 - Which charging site?
 - Which **CPU** of which **Server** in which **data center**?
- A time series also needs **metadata** describing the context of the data points:
 - commonly known as a **tag** or **label**

CHARGER AVAILABILITY TIME SERIES

- We are observing **availability**
- Our data has three **tags**:
 - Charging site
 - Weekday (Monday to Sunday)
 - Hour of day (0-24)

Site	Time	Available	Day of week	Hour of day
Ionity Mariestad	2021-11-12T23:45:00	3	Friday	23
Ionity Mariestad	2021-11-13T00:00:00	2	Saturday	0
Ionity Mariestad	2021-11-13T00:15:00	4	Saturday	0

**LET'S COMPUTE A TIME SERIES
KEY!**

TIME SERIES KEY

TIME SERIES KEY

MEASUREMENT + SITE + DAY_OF_WEEK + HOUR

series label

tags

label value

tag values

"Availability" + "lonity Mariestad" + "Saturday" + 23

EXAMPLE SERIES KEY

"availability : Ionity Mariestad : Saturday : 23"

"availability : Ionity Mariestad : Thursday : 23"

"availability : Ionity Mariestad : Thursday : 21"

HOW MANY TIMES SERIES IN THE TABLE?

SERIES 1: [Ionly Mariestad] + [Friday] + [23]

SERIES 2: [Ionly Mariestad] + [Saturday] + [0]

Site	Time	Available	Day of week	Hour of day
Ionly Mariestad	2021-11-12T23:45:00	3	Friday	23
Ionly Mariestad	2021-11-13T00:00:00	2	Saturday	0
Ionly Mariestad	2021-11-13T00:15:00	4	Saturday	0

TIME SERIES KEYS

- Uniqueness of tags means that for "Ionomy Mariestad" we will have
 - 1 site
 - 7 weekdays
 - 24 hours per day
 - => 168 time series per charging site

Q site=Max Alingsås

ity hour_of_day = 13 site = Max Alingsås weekday = Wednesday
ity hour_of_day = 21 site = Max Alingsås weekday = Sunday
ity hour_of_day = 9 site = Max Alingsås weekday = Sunday
ity hour_of_day = 1 site = Max Alingsås weekday = Tuesday
ity hour_of_day = 9 site = Max Alingsås weekday = Saturday
ity hour_of_day = 7 site = Max Alingsås weekday = Tuesday
ity hour_of_day = 12 site = Max Alingsås weekday = Thursday
ity hour_of_day = 21 site = Max Alingsås weekday = Friday
ity hour_of_day = 10 site = Max Alingsås weekday = Wednesday
ity hour_of_day = 2 site = Max Alingsås weekday = Thursday
ity hour_of_day = 17 site = Max Alingsås weekday = Tuesday
ity hour_of_day = 14 site = Max Alingsås weekday = Thursday
ity hour_of_day = 0 site = Max Alingsås weekday = Sunday
ity hour_of_day = 0 site = Max Alingsås weekday = Thursday
ity hour_of_day = 19 site = Max Alingsås weekday = Tuesday
ity hour_of_day = 0 site = Max Alingsås weekday = Friday

_time	site	hour_of_day	weekday	available
2021-11-24 14:00:00 ...	Max Alingsås	13	Wednesday	1,25
2021-12-01 14:00:00 ...	Max Alingsås	13	Wednesday	1,25
2021-12-08 14:00:00 ...	Max Alingsås	13	Wednesday	1,75
2021-12-15 14:00:00 ...	Max Alingsås	13	Wednesday	1,50
2021-12-22 14:00:00 ...	Max Alingsås	13	Wednesday	1,50
2021-12-29 14:00:00 ...	Max Alingsås	13	Wednesday	2
2022-01-05 14:00:00...	Max Alingsås	13	Wednesday	1
2022-01-12 14:00:00 ...	Max Alingsås	13	Wednesday	2

Data points for a single series aggregated per hour (UTC)

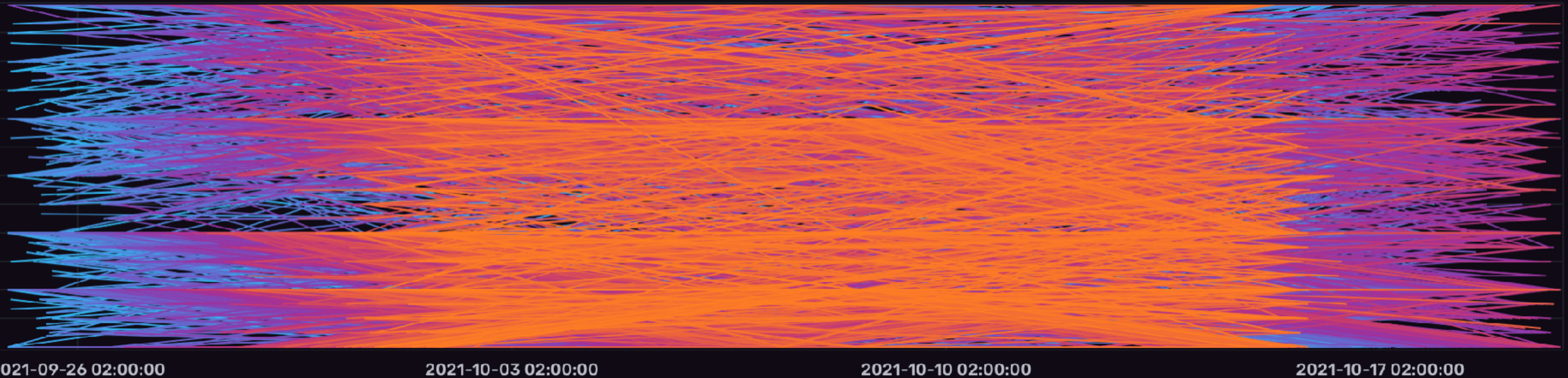
All series for site=Max Alingsås

_time	site	hour_of_day	weekday	available
2021-11-24 13:15:00 ...	Max Alingsås	13	Wednesday	1
2021-11-24 13:30:00 ...	Max Alingsås	13	Wednesday	1
2021-11-24 13:45:00 ...	Max Alingsås	13	Wednesday	1
2021-11-24 14:00:00 ...	Max Alingsås	13	Wednesday	2
2021-12-01 13:15:00 ...	Max Alingsås	13	Wednesday	1
2021-12-01 13:30:00 ...	Max Alingsås	13	Wednesday	1
2021-12-01 13:45:00 ...	Max Alingsås	13	Wednesday	1
2021-12-01 14:00:00 ...	Max Alingsås	13	Wednesday	2
2021-12-08 13:15:00 ...	Max Alingsås	13	Wednesday	2

Data points without per-hour aggregation

TIMES SERIES CARDINALITY

- If we're tracking ~20 charging sites each having 168 distinct time series, we'll have ~3660 time series.
- Without aggregation, filtering, grouping and proper visualization this data is rather useless



THE FLUX QUERY LANGUAGE

FLUX

- A general functional data scripting and query language (primarily for InfluxDB)
- Operates on Data Sources, not just InfluxDB data
 - InfluxDB, CSV, SQL
- Similar to ETL, composable streams etc.
- Supports
 - custom functions
 - pivot, join
 - map, reduce
 - histograms
 - much more...



INFLUXDATA


```
from(bucket: "chargerstatus")
```

```
|> range(start: -90d)
```

```
|> filter(fn: (r) => r["site"] == "Ionity Mariestad"  
                or r["site"] == "McDonalds Kristinehamn"  
                or r["site"] == "Bilmetro Noret")
```

```
|> group(columns: ["site"])
```

```
|> aggregateWindow(every: 4h, fn: mean)
```



**SO - WHAT IS A TIME SERIES
DATABASE!?!**

TIME SERIES DATABASES

- The secret sauce is the **columnar store model**
- Data is structured in a read-friendly manner suitable for querying huge data sets
- Writing records needs many writes
- Worth looking into!!

ID	NAME	SIZE	AGE
1	Erik	M	43
2	Lance	L	32
3	Angela	S	71

ID
1
2
3

NAME
Erik
Lance
Angela

SIZE
M
L
S

AGE
43
32
71

TIME SERIES DATABASES

- Data is very suitable for **compression**
 - Low variance
- Facebook whitepaper:
 - <https://www.vldb.org/pvldb/vol8/p1816-teller.pdf>
 - needs 16 bytes per metric data point
 - Compressed, they need 1.37 bytes on average! (about 11 bits)
- Time series data is 3 tuples:
 - series key (label + tags)
 - » timestamp
 - » value



PUBLIC DOMAIN: OPENCLIPART

delta-of-delta	delta
1634677228	1634677228
899	899
	901
0	899
2	901
0	899
2	901
1	
1	900
1	900

2 bytes per int16

1 bytes per int8

_time	site	_value
2021-10-19 21:00:28 UTC	Bilmetro Noret	4
2021-10-19 21:15:27 UTC	Bilmetro Noret	4
2021-10-19 21:30:28 UTC	Bilmetro Noret	4
2021-10-19 21:45:27 UTC	Bilmetro Noret	4
2021-10-19 22:00:28 UTC	Bilmetro Noret	4
2021-10-19 22:15:27 UTC	Bilmetro Noret	4
2021-10-19 22:30:28 UTC	Bilmetro Noret	4
2021-10-19 22:45:28 UTC	Bilmetro Noret	4
2021-10-19 23:00:28 UTC	Bilmetro Noret	4
2021-10-19 23:15:28 UTC	Bilmetro Noret	4
2021-10-19 23:30:27 UTC	Bilmetro Noret	4
2021-10-19 23:45:28 UTC	Bilmetro Noret	4
2021-10-20 00:00:28 UTC	Bilmetro Noret	4
2021-10-20 00:15:28 UTC	Bilmetro Noret	4

8 bytes per int64

- delta or delta-of-deltas can also be represented as value X repeated N number of times
- Can lead to less than 1 bit used per value :)

_time	site	_value
2021-10-19 21:00:28 UTC	Bilmetro Noret	4
2021-10-19 21:15:27 UTC	Bilmetro Noret	4
2021-10-19 21:30:28 UTC	Bilmetro Noret	4
2021-10-19 21:45:27 UTC	Bilmetro Noret	4
2021-10-19 22:00:28 UTC	Bilmetro Noret	4
2021-10-19 22:15:27 UTC	Bilmetro Noret	4
2021-10-19 22:30:28 UTC	Bilmetro Noret	4
2021-10-19 22:45:28 UTC	Bilmetro Noret	4
2021-10-19 23:00:28 UTC	Bilmetro Noret	4
2021-10-19 23:15:28 UTC	Bilmetro Noret	4
2021-10-19 23:30:27 UTC	Bilmetro Noret	4
2021-10-19 23:45:28 UTC	Bilmetro Noret	4
2021-10-20 00:00:28 UTC	Bilmetro Noret	4
2021-10-20 00:15:28 UTC	Bilmetro Noret	4

TIMES SERIES DATABASES - MORE THAN METRICS?

**PERHAPS THE
QUESTION
TO ASK IS:**

WHAT IS A METRIC?

■ MORE THAN METRICS?

- Traditionally, we've collected technical metrics from servers such as CPU, memory usage and request/response durations
- Business metrics have belonged to BI solutions
- Perhaps we'll see time series databases in the BI domain?



CC BY-SA 4.0: WIKIMEDIA COMMONS

*”Virtually every data mart
is a time series”*

- RALPH KIMBALL, 1997

DELL

NO. 12

10¢

Walt Disney's
UNCLE \$CROOGE

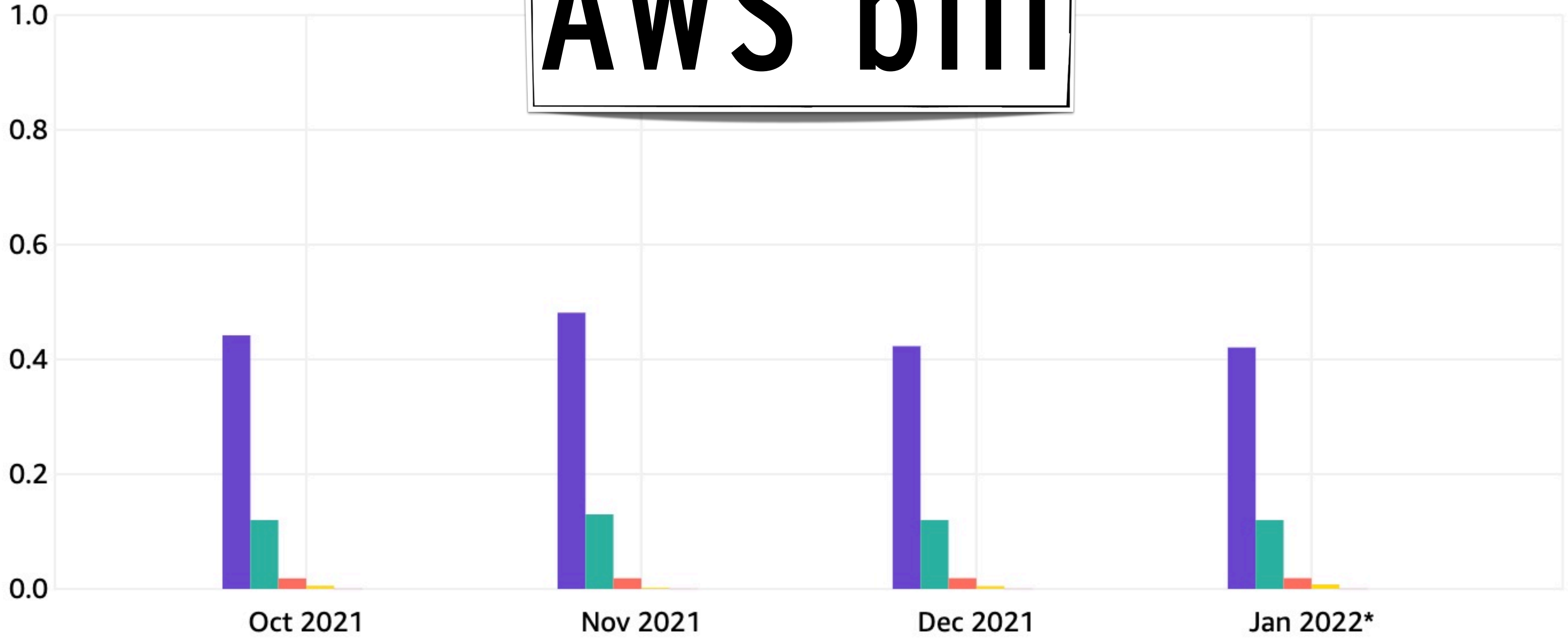


COPYRIGHT, WALT DISNEY PRODUCTIONS

CC: TOM SIMPSON, FLICKR

AWS bill

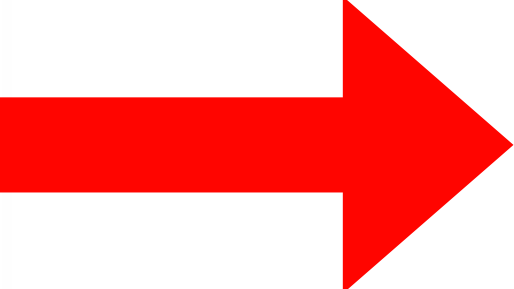
Costs (\$)





Gross Amount:	\$0.00
Tax:	\$0.00

Total:
282 696 entries

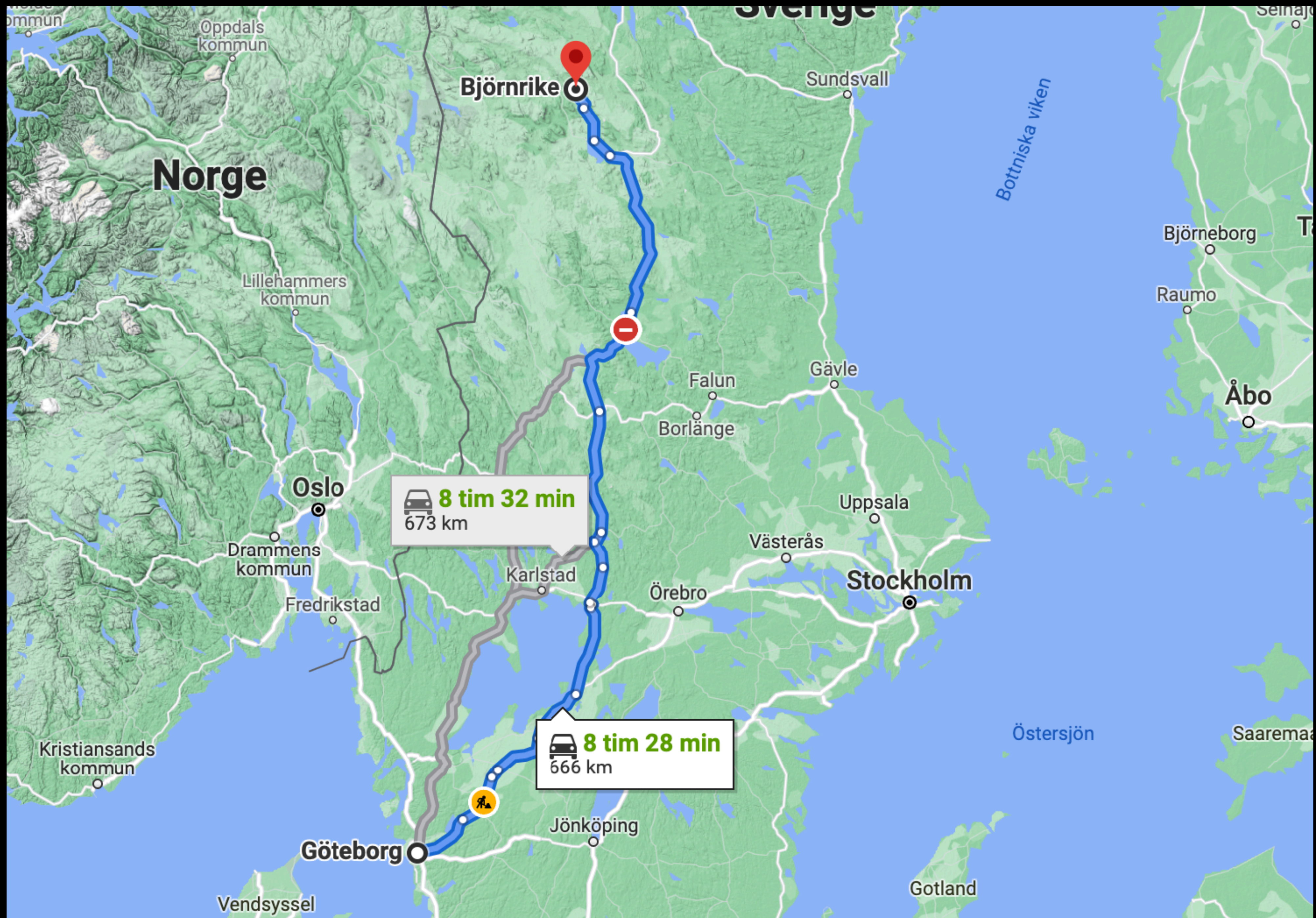
Invoice Due (USD):	 \$0.00
---------------------------	---

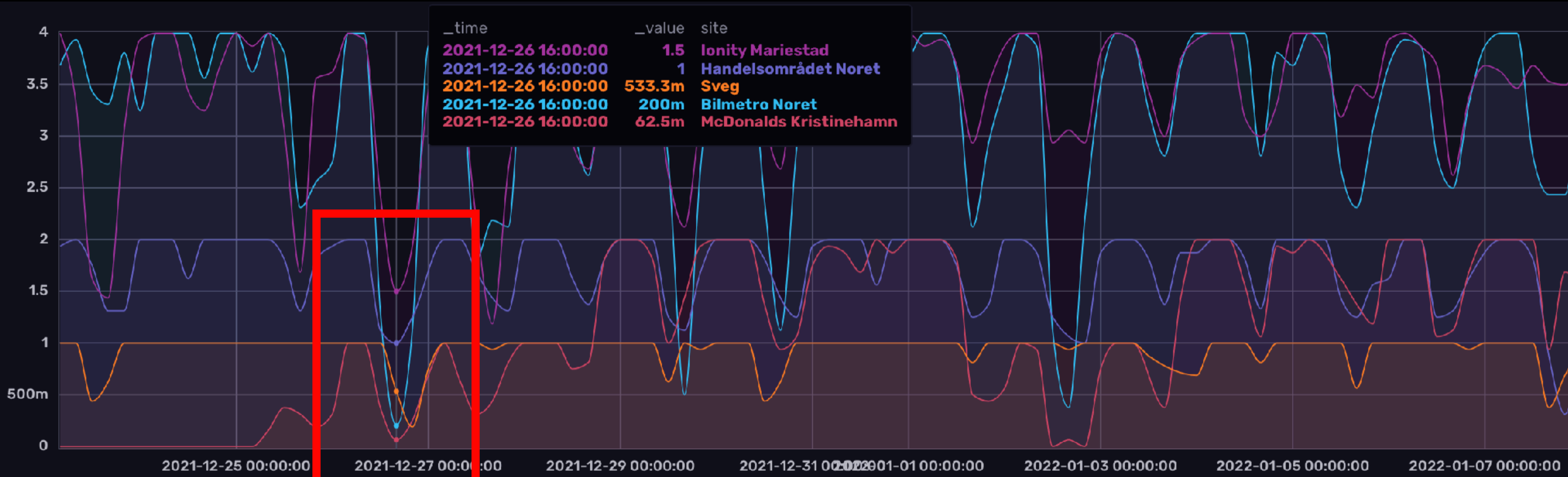
FINAL WORDS

- A technology stack built around *AWS* services and the Cloud Development Kit provides a really cost-effective way to build and deploy services in the cloud.
- Time Series databases are picking up traction - both in the traditional metrics and IoT domains, as well as emerging as an alternative for business-oriented metrics such as event streams and some scenarios otherwise typically provided by Business Intelligence software.

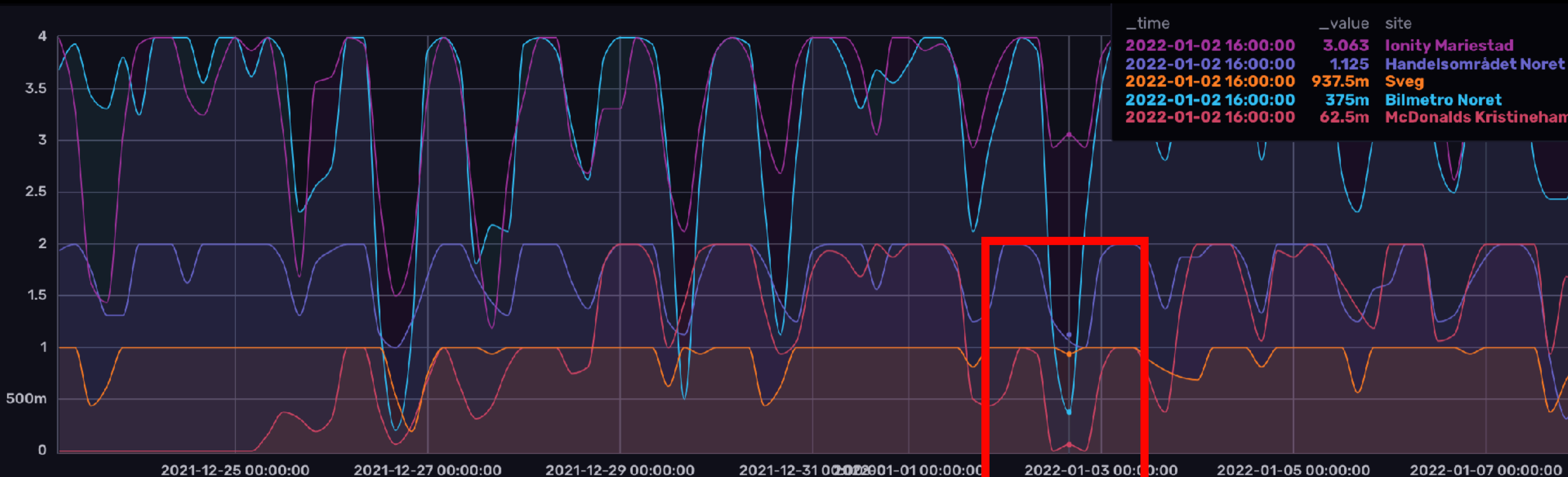
WAIT A MINUTE?!?!

**WHAT ABOUT THE
CHARGER STATS?**

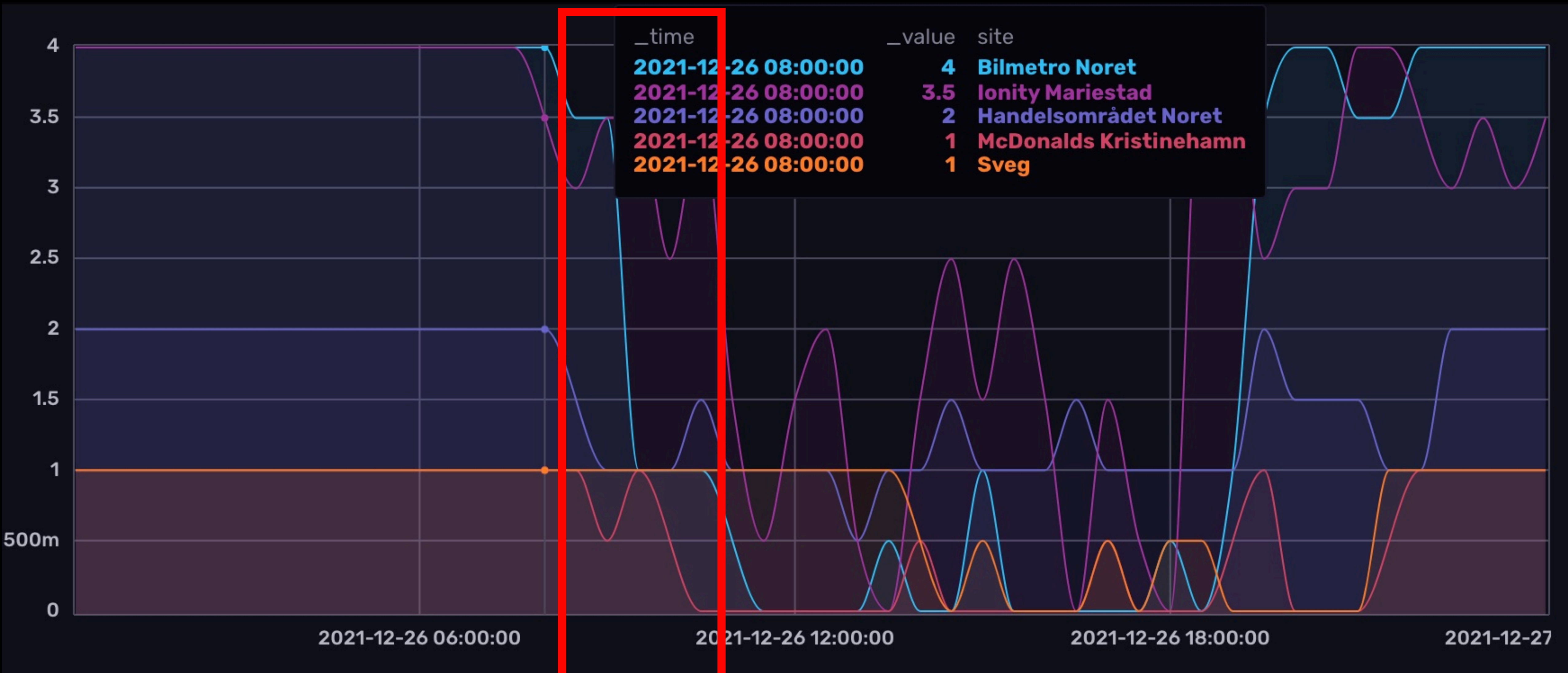




sunday 26th



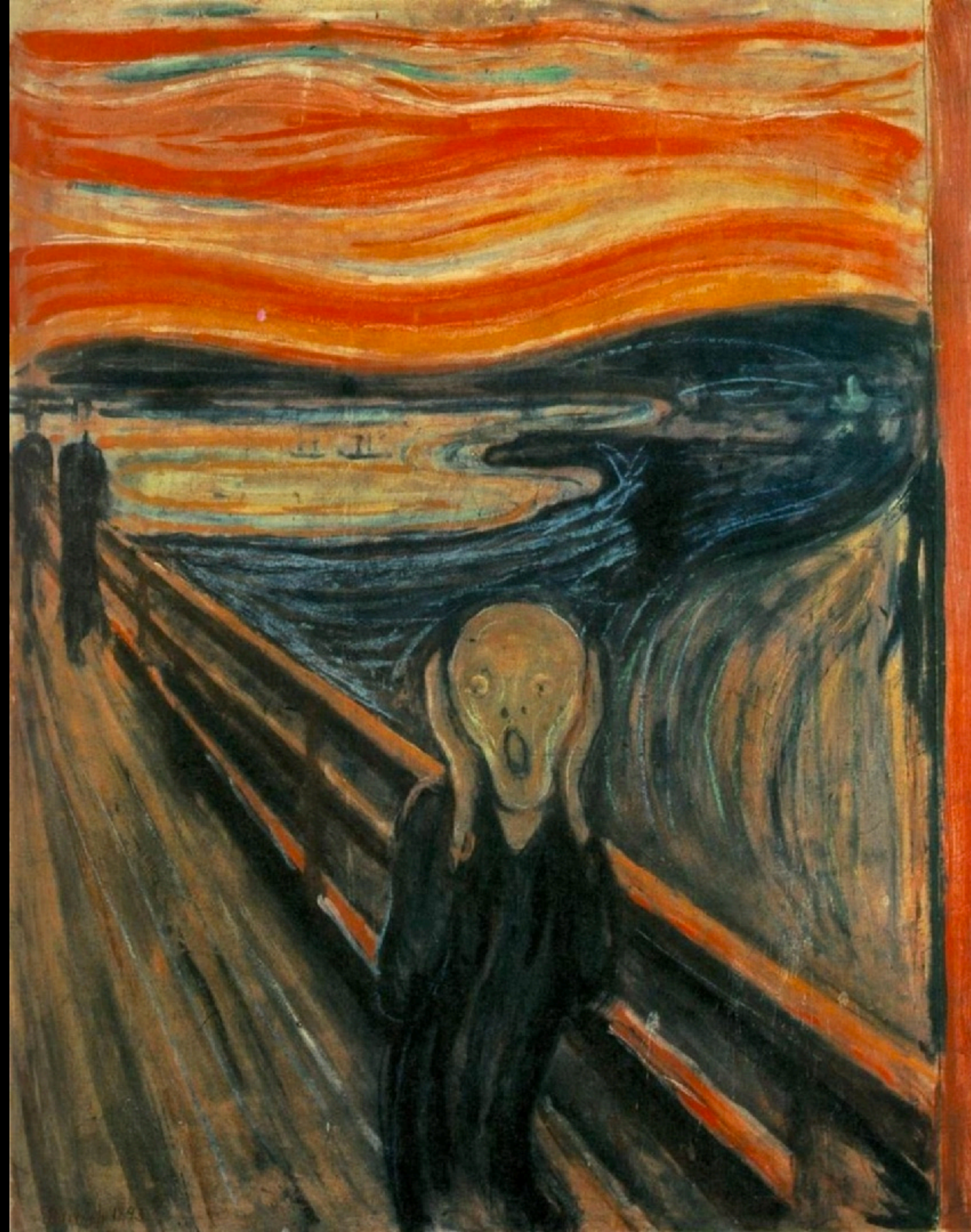
sunday 2nd



09:00

FINAL WORDS - CHARGER AVAILABILITY

- Planning is key for a smooth journey
- And while I really like and enjoy my electric car...
- ... I'll borrow a diesel car for the upcoming ski trip. :(



THANK YOU!

RESOURCES

- Facebook's whitepaper
 - Link: <https://www.vldb.org/pvldb/vol8/p1816-teller.pdf>
 - Summary: <https://jessicagreben.medium.com/four-minute-paper-facebooks-time-series-database-gorilla-800697717d72>
- ChargeFinder:
 - <https://chargefinder.com>

QUESTIONS?