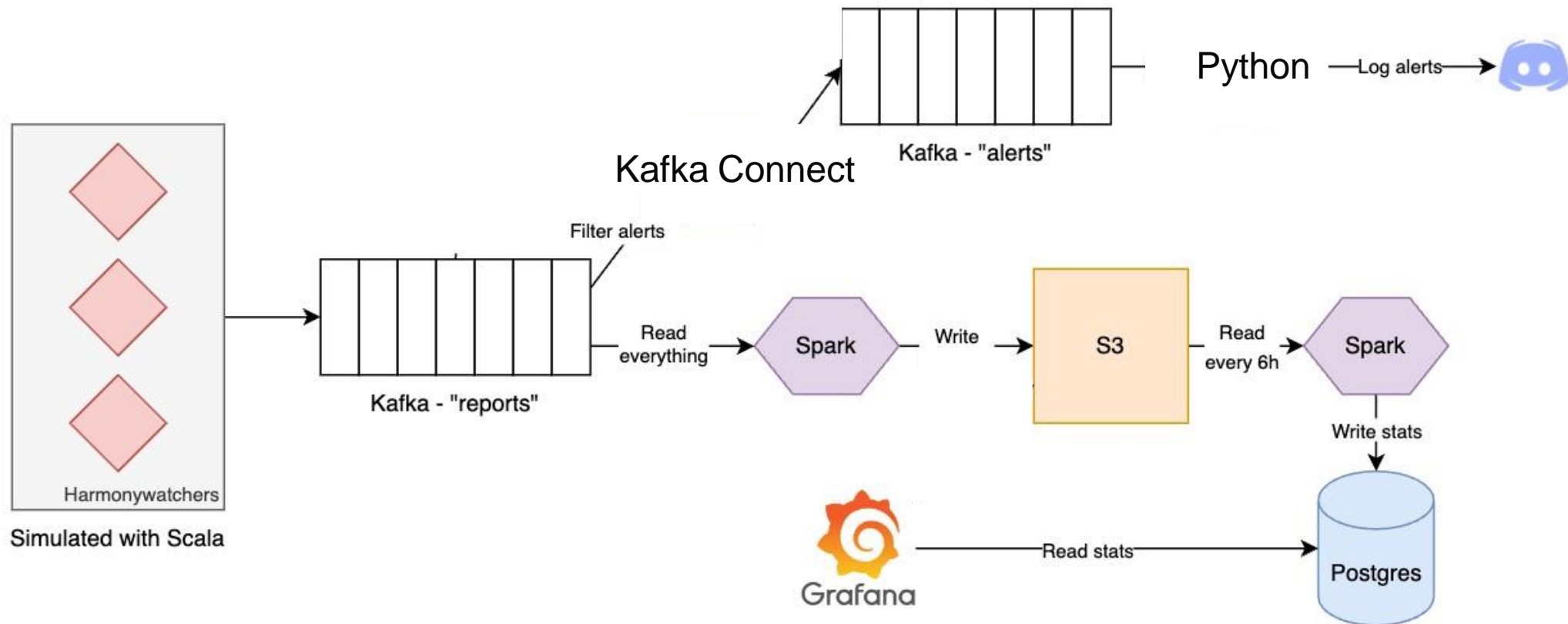


# SOUTENANCE INDE2

Emon Barberis, Paul Hartmann, Nicolas Schmitt

# SUMMARY

1. Architecture
2. Components details
  - a. Messages
  - b. Alerts
  - c. Storage
  - d. Statistics
3. Demo



# ARCHITECTURE

# MESSAGES

Generate reports :

- Random alcohol level between 0 and 3 g/L
- Random location in Paris
- One data every 1000ms

```
You, 2 seconds ago | 1 author (You)
case class Report(latitude: Double,
  longitude: Double,
  date: LocalDateTime,
  alcoholLevel: Float){

  override def toString: String = {
    s"latitude: $latitude, longitude:" +
    s" $longitude," +
    s" dayOfWeek: $date," +
    s" alcoholLevel: $alcoholLevel"
  }
}
```

# ALERTS

Score < 1.5 == ALERT

Kafka Connect Job : read from the reports to topics and put it in alerts topics instantly

Discord Job: Read from alert topics and log to discord



# STORAGE

- Datalake: AWS S3
- Read in Kafka “report” topic every **6 hours** → 4 files / day
- Write using json format
- Partioned by day:
- For testing/demo purposes, read Kafka every minute

Amazon S3 > Buckets > harmonylandrff > Reports/ > year=2023/ > month=6/ > day=20/

# STATISTICS

Spark Job:

- Run every day
- Write stats in Postgres database
- Grafana plugged on the database

DEMO