# Technologies for Advanced Programming (AA 2023-2024)

## Progetto conclusivo

### Descrizione

Design e implementazione di un’architettura software in grado di processare e analizzare in tempo reale un flusso dati provenienti da una o più sorgenti.

La soluzione deve fare uso degli strumenti tecnologici di stream processing analizzati durante il corso:

* **Ingestion** (Log Stash, Fluentd, Nifi…)
* **Streaming** Apache Kafka
* **Processing** Spark Structured Streaming
* **Machine Learning** tramite Spark Mlib (plus) o servizi esterni integrati
* **Indexing** dei dati su Elastic Search
* **Visualization** tramite Kibana
* *Opzionali:*
  + *Documentazione tramite Notebook Jupyter e Jupyter Books*
  + *Uso di docker compose / k8s*
  + *Tecnologie alternative (Opensearch, Grafana, …)*
  + *Esecuzione della soluzione in Cloud*
  + *Piattaforma di Data Catalog*
  + *Utilizzo di Kafka Connect/Stream*

### Modalità di esame

Il pacchetto con il codice sorgente deve essere disponibile su Github.

È consigliabile un colloquio preliminare con il docente per la definizione del progetto, in particolare in caso di progetto di gruppo (max due studenti) è necessario per verificare la congruità.

In sede di esame lo studente ha a disposizione 30 minuti per

* Descrivere il design della soluzione tramite presentazione
* Effettuare una demo live
* Rispondere a domande sulle componenti utilizzate

**Griglia di valutazione**

|  |  |  |
| --- | --- | --- |
| Task | % | Max |
| Esecuzione funzionante della soluzione | 75% | 22.5 |
| Presentazione | 20% | 6 |
| Documentazione/Code Style | 15% | 4.5 |
| Plus | 10% | 3 |

*Buon lavoro!*