



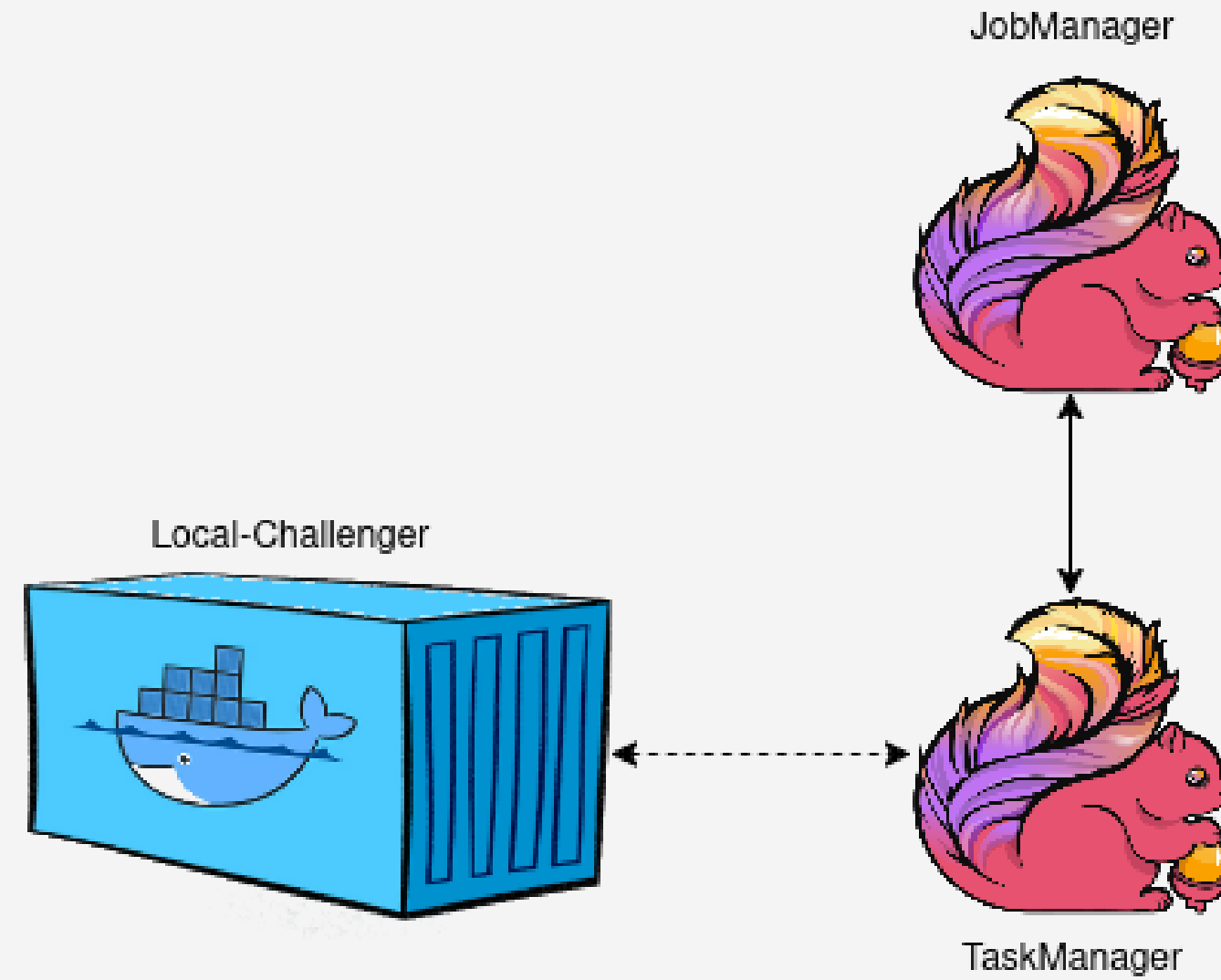
Analisi in tempo reale di difetti nella produzione L-PBF con Apache Flink

Marco Lorenzini - 0353515

Gianluca Ronzello - 0354333



Architettura

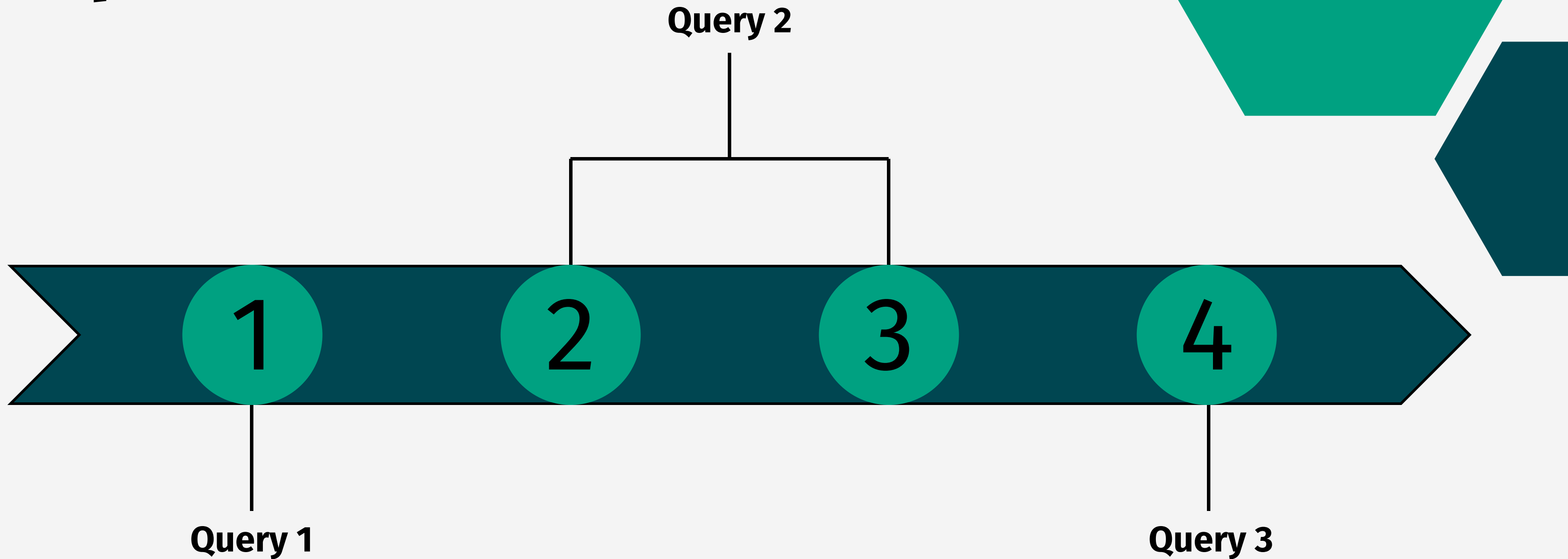


Sorgente

- **Client**
- **Source**
- **Reader**
- **Enumerator**
- **BatchDeserializer**



Pipeline



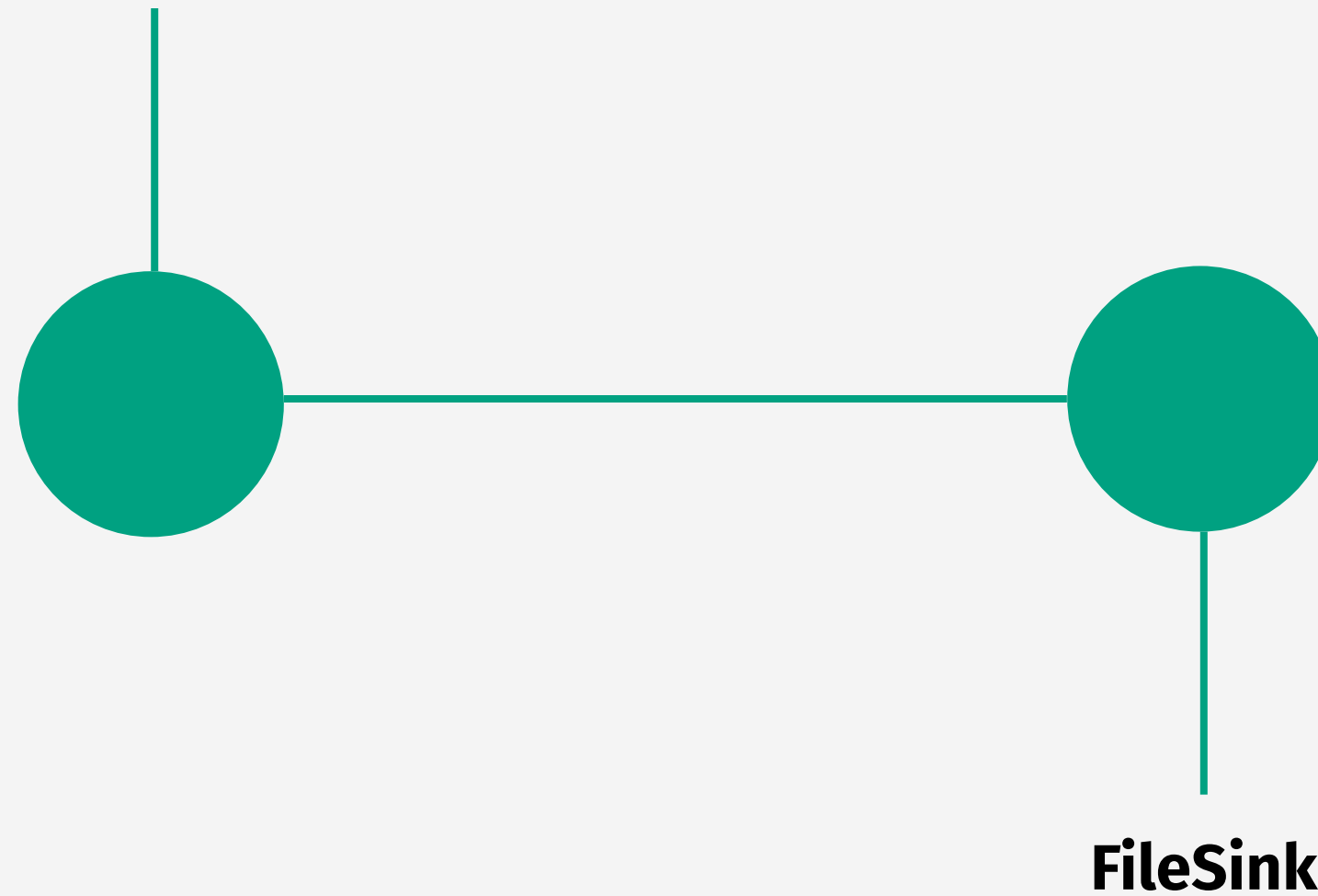
Query 1

- Identificare i punti che si trovano al di fuori del intervallo di validità [5000, 65000], questi verranno esclusi nei successivi stadi di elaborazione
- Contare i punti saturi



Query 1 - Implementazione

Calcolo validMask e punti saturi



Query 2

Per ogni **tile**, mantenere una finestra scorrevole degli ultimi 3 layer.

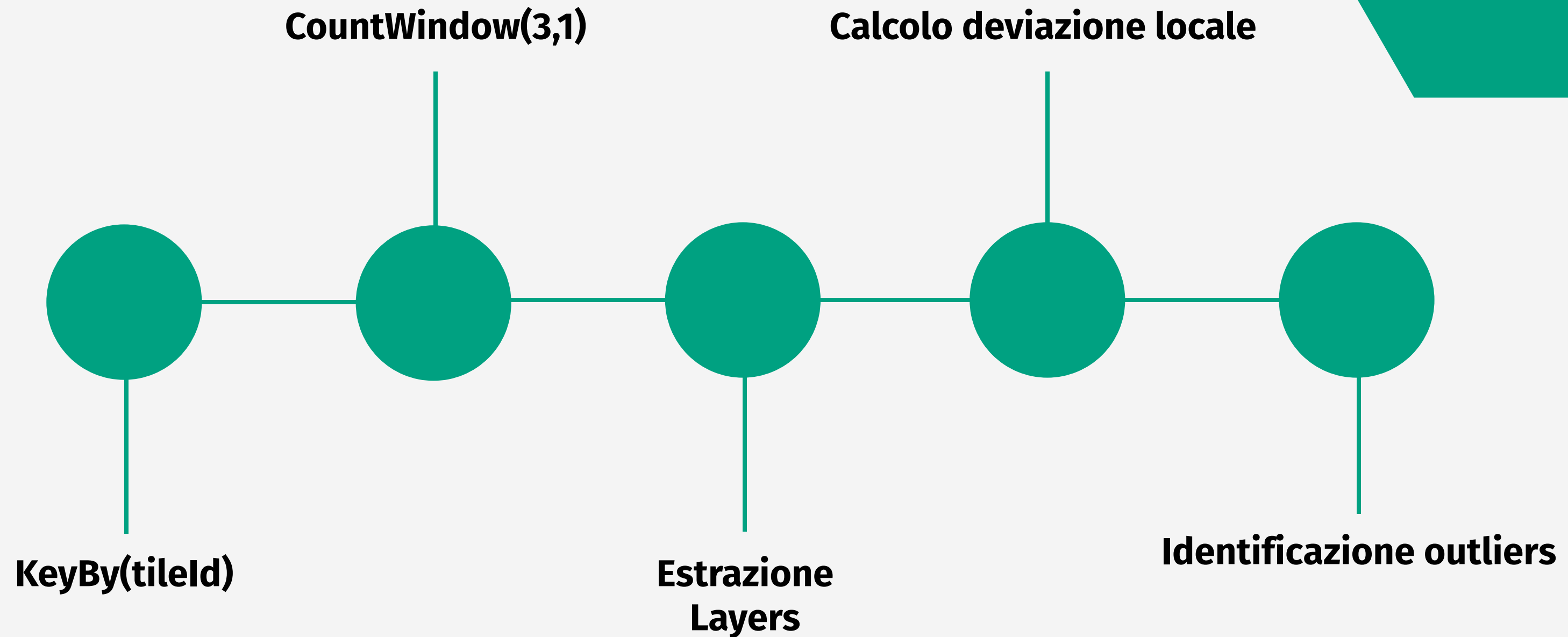
Questa finestra temporale consente l'analisi dell'evoluzione della temperatura tra layer consecutivi. Per ogni punto **p** nel layer più recente di una finestra sulla tile, calcolare la **deviazione di temperatura locale**, che viene definita come la differenza assoluta tra:

- La temperatura media dei vicini prossimi di un punto, cioè di tutti i punti con distanza di Manhattan $0 \leq d \leq 2$ da p, considerando i tre layer.
- La temperatura media dei vicini esterni, cioè di tutti i punti con distanza di Manhattan $2 < d \leq 4$ da p, considerando i tre layer.

Classificare il punto come outlier se la sua deviazione di temperatura locale supera la soglia di 6 000.



Query 2 - Implementazione



Query 3

I punti outlier identificati nella Q2 devono essere clusterizzati usando l'algoritmo **DBSCAN4**, usando la **distanza euclidea** come metrica.

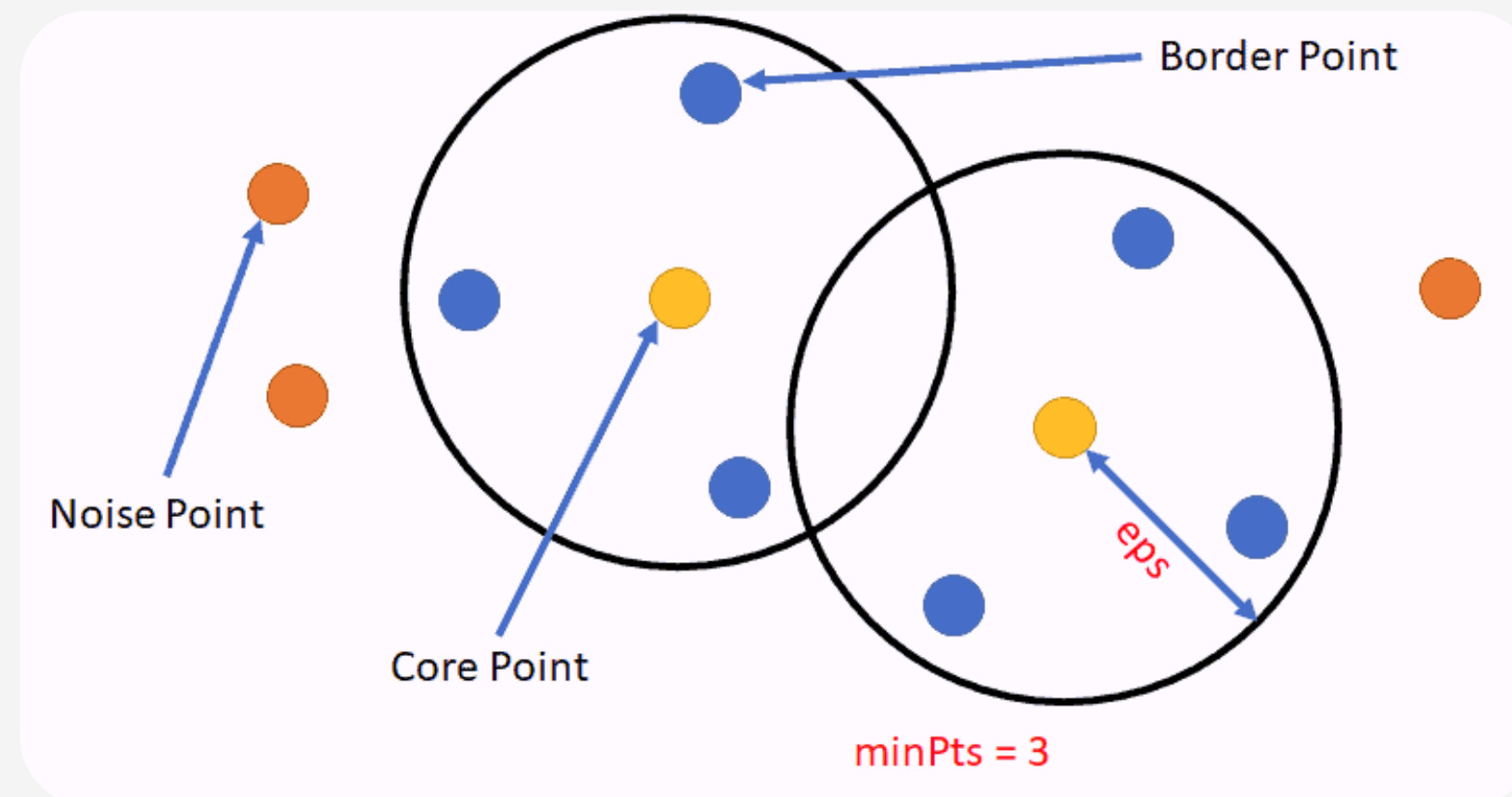


Query 3 - DBSCAN

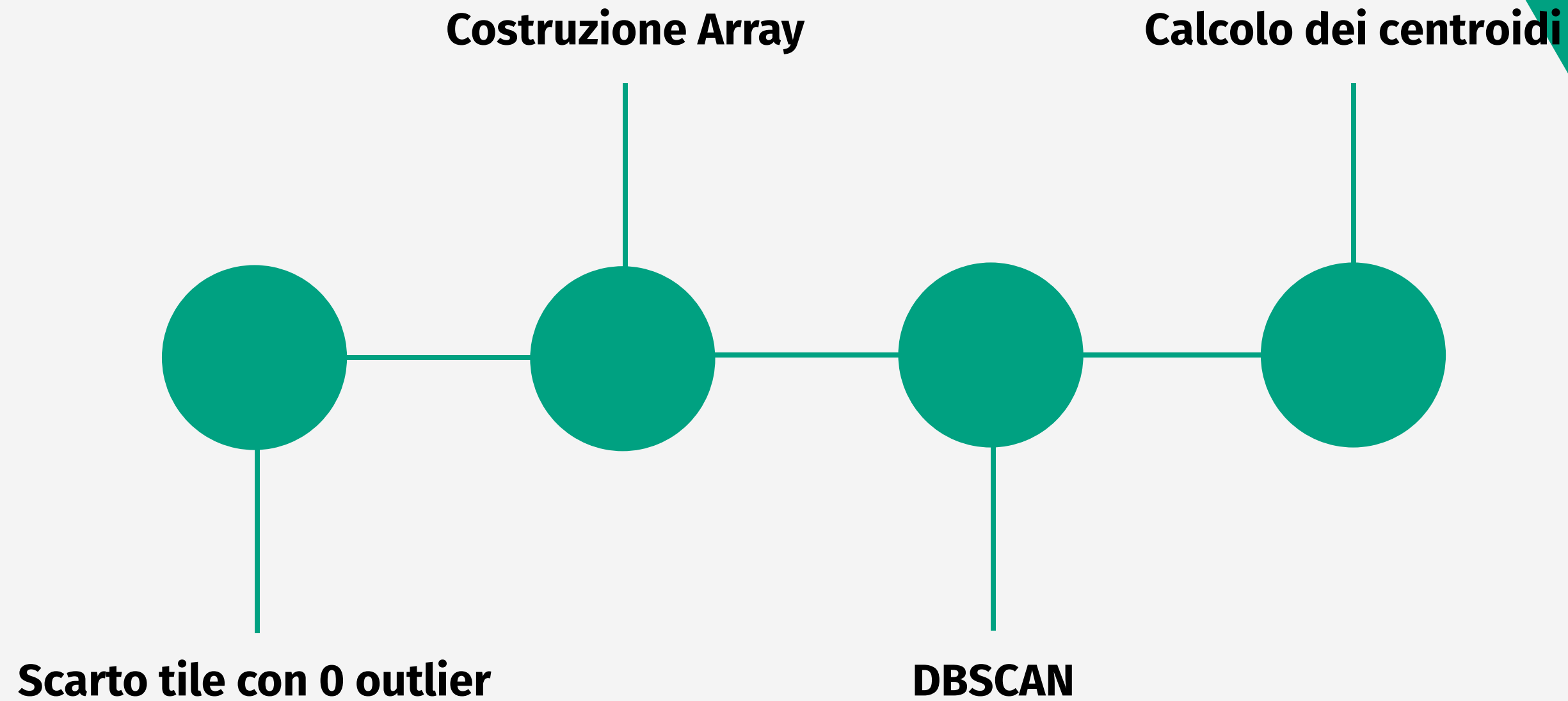
DBSCAN è un algoritmo di clustering basato sulla densità che individua regioni ad alta densità di punti separandole da regioni a bassa densità (considerate rumore).

I suoi due parametri principali sono:

- **ϵ (eps)**: raggio di vicinato attorno a ciascun punto.
- **MinPts**: numero minimo di punti richiesti per considerare una regione “densa”.



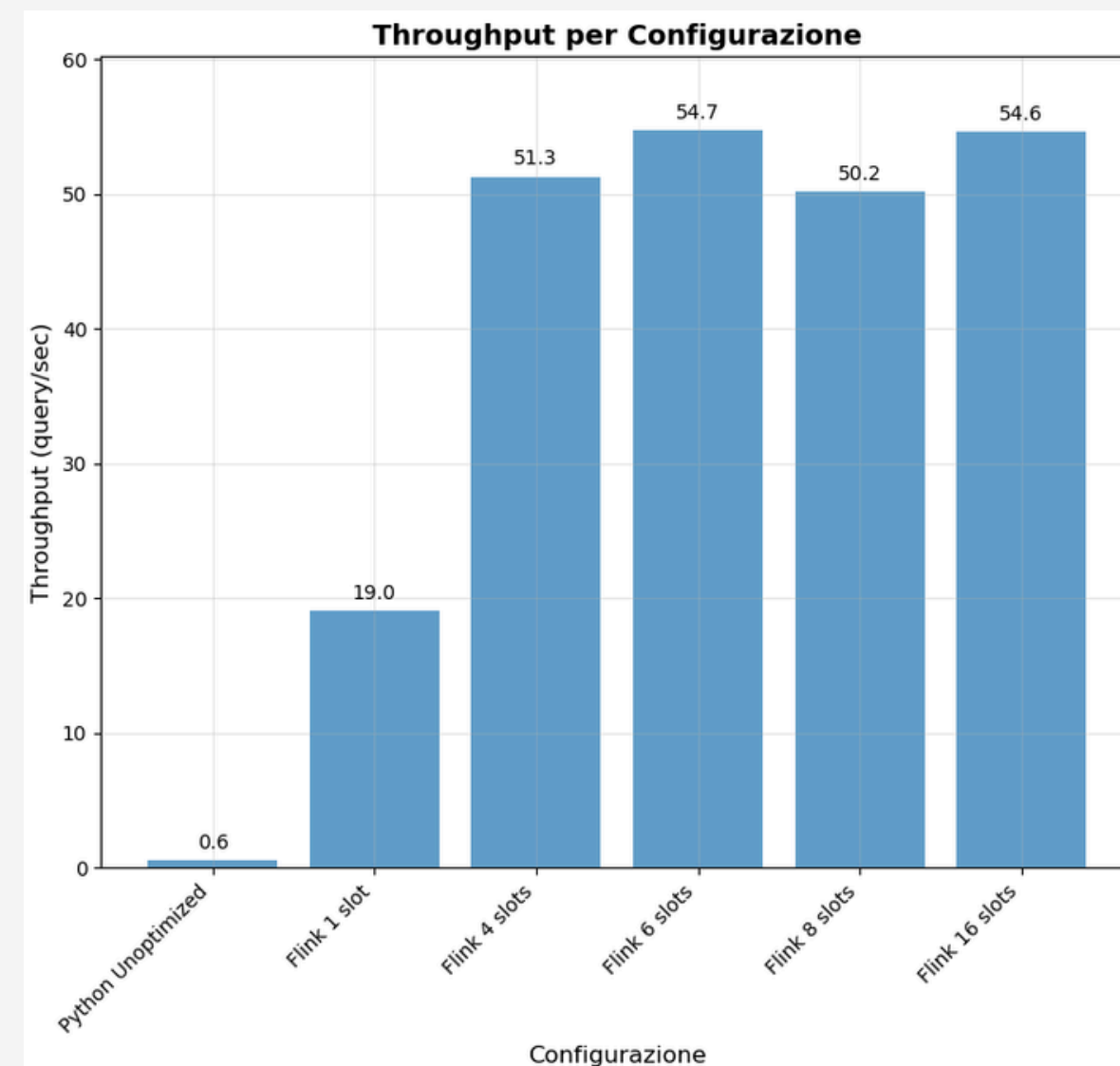
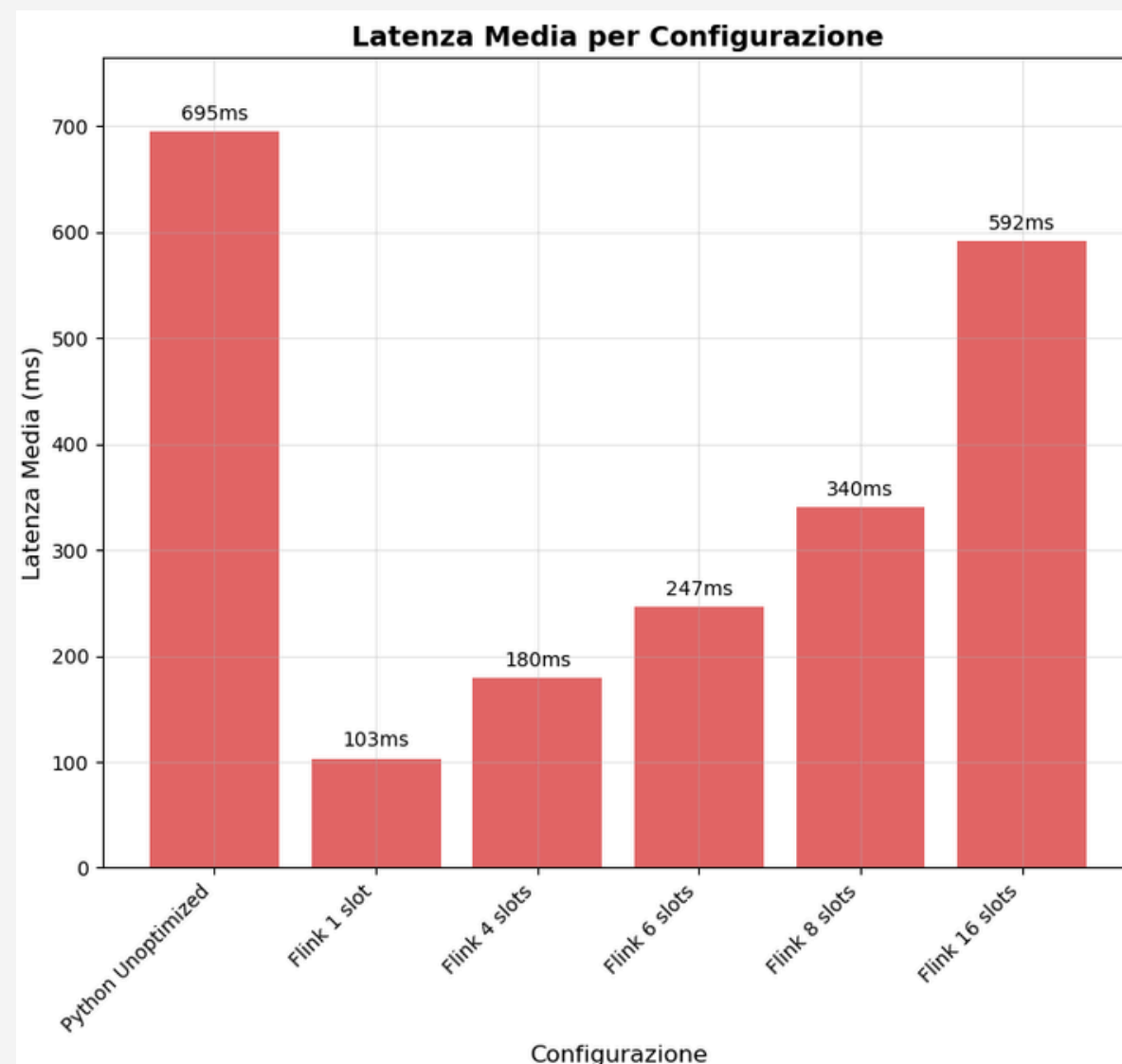
Query 3 - Implementazione



Risultati

Vediamo ora i risultati ottenuti facendo variare il numero di slot tra {1,4,6,8,16} in termini di:

- **Latenza:** Tempo che intercorre tra il momento in cui un evento arriva al sistema e quello in cui la sua elaborazione produce un risultato.
- **Throughput:** Volume di dati che il sistema riesce a elaborare in un'unità di tempo.



**Thank
You!**

