

Verbale Esterno

21-12-2023

Contatti: swateng.team@gmail.com

Versione: 0.1



Registro delle Modifiche

Versione	Data	Descrizione	Autore	Ruolo
0.1	09-01-2024	Prima stesura	Nancy Kalaj	Responsabile

Indice

1 Partecipanti	4
1.1 Partecipanti di <i>SWAT Engineering</i>	4
1.2 Partecipanti di <i>Sync Lab</i>	4
2 Sintesi Elaborazione Incontro	4
2.1 Migliorie apportate al PoC	4
2.2 Analisi dei Requisiti	5
2.3 Consigli per la presentazione RTB	5

1 Partecipanti

Inizio incontro 15:30

Fine incontro 16:00

Luogo incontro Chiamata Google Meet

1.1 Partecipanti di *SWAT Engineering*

Nome	Durata presenza
Caregnato Simone	0.5h
Costantin Riccardo Alberto	0.5h
D'Ovidio Giacomo	0.5h
Kalaj Nancy	0.5h
Rango Matteo	0.5h
Toniolo Riccardo	0.5h

1.2 Partecipanti di *Sync Lab*

- Daniele Zorzi
- Fabio Pallaro
- Andrea Dorigo

2 Sintesi Elaborazione Incontro

2.1 Migliorie apportate al PoC

L'incontro si è aperto con la valutazione delle modifiche apportate al PoC; in particolare, il team ha dimostrato l'utilizzo della funzione aggregata "movingAverage" per ottimizzare le query (questo per evitare di far fare operazioni di trasformazioni sui dati a Grafana, ossia al visualization layer), l'aggiunta dei sensori di pioggia e l'aggiunta di descrizioni ai pannelli della dashboard per renderla più facilmente comprensibile. La Proponente ha valutato positivamente le modifiche dimostrate e ha consigliato anche l'implementazione di alcune funzionalità aggiuntive:

- Il filtraggio dei dati: sfrutta le "Grafana variables";
- Tabelle contenenti i dati grezzi dei sensori: contrariamente ai dati aggregati che vengono sfruttati per fare analisi di tipo statistico di business, i dati grezzi vengono utilizzati nel caso in cui si abbia bisogno di visualizzare un singolo evento di interesse; ad esempio, nel caso in cui si visualizzi un allarme all'interno della dashboard, l'anomalia corrispondente sarebbe facilmente rintracciabile all'interno di una delle tabelle contenente i dati grezzi;

- Aggiunta di altre tipologie di sensori per incrementare l'impatto visivo della dashboard.

2.2 Analisi dei Requisiti

Il team ha voluto esprimere un dubbio riguardo all'inclusione di una stima dell'ordine di grandezza dei dati in grado di essere gestiti dal sistema all'interno dell'*Analisi dei Requisiti v1.0*; in risposta, la Proponente ha suggerito di considerarlo come un requisito di performance e di impostarne il valore in seguito ad alcuni test volti a determinare la portata massima dei dati che il sistema è in grado di gestire senza subire rallentamenti o impedimenti visibili (i test dovrebbero durare almeno qualche ora). La Proponente prevede che il sistema sarà in grado di gestire intorno ai 70-80 dati generati al secondo o anche di più.

2.3 Consigli per la presentazione RTB

La Proponente ha infine fornito alcuni consigli per affrontare al meglio la prima revisione RTB:

- Rendere la dashboard più dinamica restringendo l'intervallo temporale di visualizzazione dei dati: in particolare, è stato suggerito di impostare un intervallo di tempo ridotto all'inizio della presentazione per mostrare l'andamento dei dati in ingresso e di allargarlo in seguito;
- Incorporare all'interno della dimostrazione del PoC tutte le componenti dell'architettura: i generatori di dati che pubblicano gli eventi, il flusso di dati che attraversa Kafka, i dati in ingresso in ClickHouse in formato pseudo-tabellare e infine la loro visualizzazione finale in Grafana;
- Motivare la scelta di utilizzare Grafana per un progetto che prevede event-driven visualization: Grafana si basa su un refresh molto rapido dei singoli grafici della dashboard (anche ogni 0.5 secondi), mentre altri strumenti come PowerBI o SuperSet sono impostati su un refresh meno frequente (ogni 10-30 secondi) poiché si concentrano maggiormente sulle aggregazioni e sull'analisi dei dati rispetto alla loro visualizzazione in real time;
- Motivare la scelta di utilizzare ClickHouse per l'elaborazione di dati di tipo time series: nonostante ClickHouse sia un motore OLAP, esso viene spesso utilizzato anche per gestire dati in formato time series in quanto si prestano bene ad operazioni su grandi moli di dati in continuo arrivo da fonti come Kafka e simili.

Luogo: Padova (PD)

Data: 23/01/2024

Firma: *Falzi Daniele*
