



ByteOps.swe@gmail.com

Verbale Esterno · Data: 23/02/2024

Informazioni documento

Luogo	Google Meet
Orario	16:30 - 17:00
Redattore	N. Preto
Verificatore	E. Hysa
Amministratore	F. Pozza
Destinatari	T. Vardanega R. Cardin
Partecipanti interni	A. Barutta E. Hysa R. Smanio D. Diotto F. Pozza L. Skenderi N. Preto
Partecipanti esterni	A. Dorigo

Il responsabile: Lisien Skenderi

Registro delle modifiche

Versione	Data	Autore	Verificatore	Dettaglio
0.0.1	23/02/2024	N. Preto	R. Smanio	Redazione documento

Indice

ByteOps

Febbraio 23, 2024

Contents

1	Revisione del periodo precedente	4
2	Ordine del giorno	4
2.1	Ottimizzazione database e query	4
2.2	Pianificazione test da implementare	4
2.3	Verosimiglianza delle misurazioni	5
3	Attività da svolgere	5

1 Revisione del periodo precedente

In data 20/02/2024 si è tenuto l'incontro di revisione RTB_G con il Professor Vardanega, il quale ha espresso un giudizio positivo sul lavoro svolto fino a quel momento. Tuttavia, rispetto all'ultimo SAL_G , non sono state avviate nuove attività a causa dei precedenti impegni accademici, come l'esame di Ingegneria del Software, che diversi membri del gruppo hanno dovuto affrontare il 22 febbraio. I componenti del team che non erano coinvolti nell'esame hanno continuato a lavorare sul progetto, iniziando l'implementazione delle correzioni indicate dal Professor Vardanega e conducendo sessioni di brainstorming per pianificare le future attività relative alla PB_G .

2 Ordine del giorno

2.1 Ottimizzazione database e query

Si è discusso sulle modalità per ottimizzare il $database_G$ e le query per migliorare le prestazioni del sistema.

È emerso che le materialized view attualmente adottate, sebbene potenzialmente utili in un contesto ipotetico con grandi volumi di dati, potrebbero risultare complesse da mantenere ed eccessivamente sofisticate per il nostro contesto specifico.

Di conseguenza, la $Proponente_G$ ha suggerito di valutare l'opportunità di utilizzare il database in modo più tradizionale e di effettuare le aggregazioni di dati tramite query, evitando così di dover calcolare le aggregazioni incrementalmente. Questo approccio potrebbe consentire di ottenere gli stessi risultati senza aggiungere eccessiva complessità alla configurazione del database. È stato inoltre stabilito che successivamente ai test di performance, verrà valutata l'opportunità di utilizzare o meno funzionalità più avanzate.

2.2 Pianificazione test da implementare

Si sono analizzati possibili test da effettuare prestando particolare attenzione ai *Test di integrazione $_G$* sul database e ai test sull'integrità dei dati.

Uno dei test pianificati riguarda la verifica della coerenza tra i dati generati dai $sensori_G$ e quelli memorizzati sia su $Apache Kafka_G$ che nel database $Clickhouse_G$. Poiché non sono disponibili $librerie_G$ predefinite che consentano di eseguire compiti così specifici, l'azienda proponente ha suggerito lo sviluppo di una suite di test di confronto personalizzata. Questo strumento è essenziale per garantire l'integrità dei dati lungo l'intero percorso di trasmissione e memorizzazione.

Per quanto riguarda gli stress test, la decisione è stata quella di generare considerevoli volumi di dati al secondo, e successivamente verificare le prestazioni in seguito ad operazioni e query su tali dati. L'obiettivo è identificare il miglior approccio per la gestione ottimale del database in situazioni di elevato carico.

2.3 Verosimiglianza delle misurazioni

È stato trattato il quesito riguardante la necessità di rendere plausibili i dati generati, arrivando alla conclusione che l'attenzione dovrebbe essere rivolta alla correttezza dei dati piuttosto che alla loro verosimiglianza.

Poiché il focus del progetto non verte sulla simulazione di dati realistici, bensì sulla gestione di dati in un contesto di *Big Data*_G, la verosimiglianza dei dati non costituisce un requisito fondamentale; al contrario, sarebbe addirittura interessante poter gestire dati non realistici al fine di testare la capacità del sistema nell'affrontare dati errati.

3 Attività da svolgere

- Predisporre l'ambiente per essere nelle condizioni di effettuare le diverse tipologie di test;
- Eseguire stress test sul database per valutare la sua efficienza;
- Eseguire test di verifica sui dati per verificarne coerenza tra le varie componenti del sistema;
- Sviluppo dello strumento di verifica della coerenza tra i dati in uscita ed in entrata tra le componenti del sistema;
- Implementazione calcolo di punteggio della salute della città (opzionale).

Luogo e Data: Padova, 23/02/2024

Firma referente Sync Lab:

Folzi Daniele