



▶ DATA MANAGEMENT AND STORAGE IN IoT

Riferimenti: *Data Management and Storage IoT* 

Alessandro Schiavo

► TABLE OF CONTENTS

01

ARCHITETTURE : CLOUD, EDGE COMPUTING

02

DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION

03

DATA ANALYTICS

04

DATA RETENTION POLICY

05

SECURITY AND PRIVACY

06

SCALABILITY

Alessandro Schiavo

ARCHITETTURE : CLOUD, EDGE COMPUTING

Cloud Computing

- + Astrazione di manutenzione, allocazione e inizializzazione delle risorse
- + Diversificazione dei servizi offerti
- + Risorse computazionali (concettualmente) illimitate
- + Astrazione requisiti tecnici come sicurezza, privacy e retention policies
- + Pay-per-use
- Non fornisce priorità alle richieste
- Latenza non trascurabile e non diversificabile
- Vincolato alle condizioni offerte dal provider

Edge Computing

- + Nodi fisicamente vicini alle sorgenti dati
- + Latenza diversificabile fornendo connessioni ad-hoc alle sorgenti
- + Capacità di applicare trasformazioni sui dati e memorizzazione sui nodi
- + Scelta del modello gerarchico (cluster, albero)
- Manutenzione non trascurabile
- Sincronizzazione delle memorie
- Gestione dei requisiti tecnici infrastrutturali (sicurezza, privacy, scalabilità)

ARCHITETTURE : CLOUD, EDGE COMPUTING

RICERCHE SELEZIONATE

- Barcelo et al. in [21] propone un framework che tratta le reti di nodi come un problema di flusso riducendo con un fattore di al massimo 80% le risorse
- Zhang et al. [25] comprende il framework EVAPS utilizzando sia i concetti di **Cloud** che **Edge** per eliminare eventuali dati non utili al processo di prevenzione ed incremento della sicurezza delle strade urbane.



▶ **DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION**

La Data Management è l'insieme delle tecniche per la gestione dei flussi di dati e di tecniche atte a contenere i costi di gestione dei dati stessi.



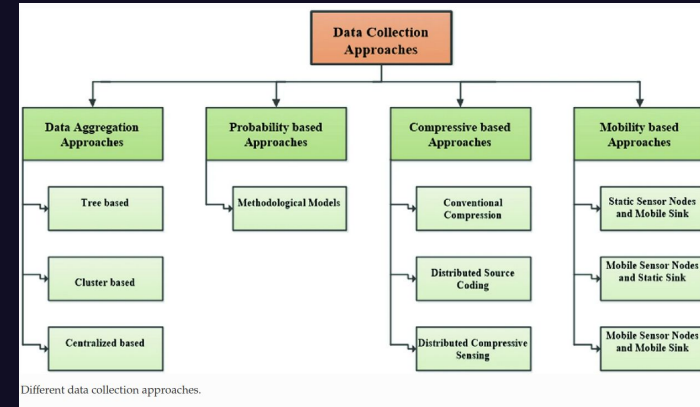
Alessandro Schiavo

DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION

DATA COLLECTION

I metodi della Data Collection delineano le caratteristiche con i quali i dati sono caricati:

- Priorità di memorizzazione e caricamento
- Compressione dei dati grezzi
- Standardizzazione dei valori rilevati
- Tecniche di ridondanza e accessibilità dei dati grezzi
- Applicazione a livello firmware e/o software



DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION

DATA AGGREGATION

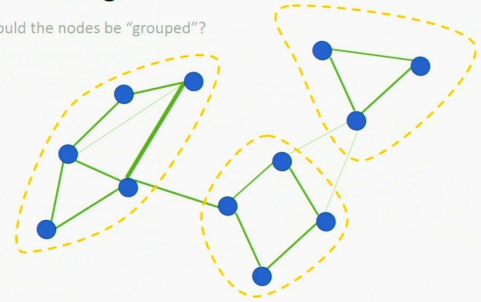
Costituisce una serie di metodi utili per aggregare i dati tenendo conto dell'infrastruttura e dei dati disponibili. Le tecniche si dividono per: aggregazione centralizzata, per cluster, per alberi.

Obiettivi:

- Riduzione delle risorse impiegate
- Maggior compressione e accuratezza dei dati fruibili
- Riduzione delle ripetizioni tra intervalli temporali diversi
- Evitare collisioni tra dati diversi ma di simile interpretazione

Graph clustering

How should the nodes be "grouped"?

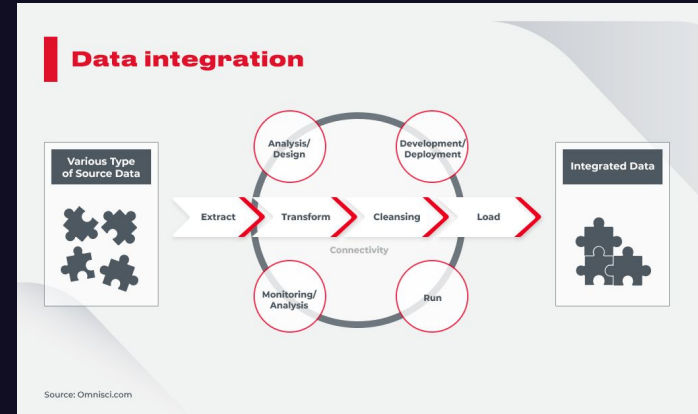


DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION

DATA INTEGRATION

Si definisce come il processo di unione di dati provenienti da più sorgenti. Sono applicate tecniche di :

- Estrazione delle feature di rilievo
- trasformazione/mapping uniformando i dati con quelli esistenti
- Caricamento nei sistemi di memoria per l'analisi dei dati



DATA MANAGEMENT: COLLECTION, AGGREGATION, INTEGRATION

RICERCHE SELEZIONATE

- You et al. in [23] propone il calcolo dell'etichetta in base alle risorse energia e computazione del dispositivo locale e dal guadagno ottenuto dalla richiesta per l'intera rete
- Rezaeibagha et al. in [16] propone un meccanismo per collezionare e accumulare dati provenienti da dispositivi IoT indossabili utilizzano la crittografia. La peculiarità del meccanismo è che consente di effettuare operazioni di analisi dei dati direttamente dai dati crittografati



► DATA ANALYTICS

La Data Analytics comprende una serie di strumenti per compiere inferenza sui dati con modelli predittivi e per visualizzare i dati raccolti



Alessandro Schiavo

DATA ANALYTICS

Le operazioni di analisi, estrazione dei modelli e visualizzazione sono applicate su dati caricati a diversi livelli: memoria, BI, analisi massiva.

I metodi di analisi sono applicati in funzione della struttura dei dati e dai pattern da rilevare, alcuni algoritmi sono: KNN, clustering, analisi predittiva.

La visualizzazione dei dati risulta onerosa data la mole di dati e la variabilità delle strutture dati da sintetizzare. In commercio esistono numerosi strumenti che la semplificano, come: Tableau, Kibana.



RICERCHE SELEZIONATE

- Wang et. al in [26] propone il metodo DGE per separare le caratteristiche multidimensionali e facilitare la rappresentazione del dato
- Zhong et. al in [27] propone un metodo per visualizzare in real-time i Big Data memorizzati nel cloud

A series of thin, light blue and white lines forming a complex geometric pattern in the top-left corner of the slide.

► DATA RETENTION POLICY

Definisce in che modo, dove e per quanto tempo sono conservati i dati, il loro formato, il mezzo su cui vengono conservati, chi vi ha accesso e che cosa succede nel caso in cui si verifichi un accesso non autorizzato a tali dati

A series of thin, light blue and white lines forming a complex geometric pattern in the bottom-right corner of the slide.

Alessandro Schiavo

DATA RETENTION POLICY

Gli strumenti di memorizzazione sono affiancati ad una serie di regolamentazioni alle quali i dati sono sottoposti. Le normative possono essere aziendali, settoriali o territoriali come la GDPR per gli stati membri dell'UE.

La complessità dell'attuazione delle normative è correlata alle strutture di memorizzazione utilizzate e dalla presenza di eventuali supporti offline



Alessandro Schiavo

DATA RETENTION POLICY

RICERCHE SELEZIONATE

- Diène et al. [28] propone di crittografare ogni file con una chiave univoca e di utilizzarla per controllare la durata del file in memoria; gestendo in maniera centralizzata le chiavi, il sistema può sapere in ogni momento lo stato dei file sia su dispositivi online che offline



► **SECURITY AND PRIVACY**

I dispositivi IoT sono generalmente dotati di limitate capacità hardware che impediscono o limitano l'attuazione di strumenti di sicurezza.

Anche la varietà dei dispositivi IoT e le numerose soluzioni architetturali contribuiscono ad innalzare la difficoltà di attuazione dei sistemi di sicurezza



Alessandro Schiavo

SECURITY AND PRIVACY

La sicurezza dei sistemi IoT coinvolge l'applicazione di strumenti di sicurezza ad ogni livello:

- Hardware: accessi, energia, porte aperte
- reti di trasmissione: crittografia, autenticazione
- supporti di memoria: privacy, accessi
- visualizzazione dei dati: anonimizzazione

La privacy estende il concetto di sicurezza integrando anche le preferenze della persona sui dati archiviati e il rispetto delle normative legali.

La sicurezza di una infrastruttura IoT comprende anche la gestione di attacchi con software e/o dati mirati



RICERCHE SELEZIONATE

- Ruckebush et al. [35] che propone aggiornamenti remoti parziali divisi in tre livelli di priorità
- Vidgren et al. [32] ha illustrato come un avversario potrebbe compromettere i dispositivi IoT abilitati per ZigBee leggendo le chiavi trasmesse non crittografate durante l'inizializzazione
- Hatzivasilis et al. [34] propone una confronto di 52 chip per la crittografia sui dispositivi IoT

A series of thin, light blue lines forming a circuit-like pattern in the top-left corner of the slide. The lines include vertical, horizontal, and diagonal segments, with some small circles at the junctions.

► **SCALABILITY**

Con scalabilità sono intese una serie di funzionalità che una infrastruttura IoT dovrebbe integrare, come la possibilità di aggiungere/rimuovere dispositivi eterogenei e seguire la domanda delle risorse

A series of thin, light blue lines forming a circuit-like pattern in the bottom-right corner of the slide. The lines include horizontal, vertical, and diagonal segments, with some small circles at the junctions.

Alessandro Schiavo

SCALABILITY

Per scalabilità s'intendono requisiti tecnici e funzionali del tipo:

- Scalabilità delle risorse:
 - Orizzontale: distribuzione del carico
 - Verticale: dinamicità delle risorse impiegate
- Allocazione delle risorse automatizzato
- Controllo delle pipeline dei dati
- Supporto a distribuzione/ parallelizzazione dei processi
- Tolleranza agli errori



RICERCHE SELEZIONATE

- Sarkar et al. in [38] propone un'architettura distribuita (DIAT) in grado di assecondare la scalabilità e interoperabilità del sistema
- Miorandi et al. in [36] propone uno studio sull'allocazione delle risorse automatizzato

Abstract geometric lines in the top-left corner, including a vertical line, a horizontal line, and a diagonal line, with a small circle at the end of the diagonal line.

Fine
grazie per l'attenzione

Abstract geometric lines in the bottom-right corner, including a horizontal line, a vertical line, and a diagonal line, with a small circle at the end of the diagonal line.

Alessandro Schiavo