

Partners





Marco Pozzan

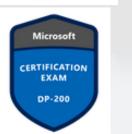
- Dal 2002 faccio consulenza sulla progettazione di data warehouse relazional e sulla progettazione di cubi Tabular/Multidimensional con strumenti Microsoft.
- Dal 2007 Community Lead di 1nn0va (https://www.facebook.com/1nn0va)
- dal 2014 MVP per SQL Server
- Dal 2017 mi occupo di modern data warehouse con prodotti Azure: Synapse, Azure Data Factory, Stream Analytics, Data Lake e affini. Sono Docente all'Università di Pordenone nel corso Architetture Big Data e DWH: Tecniche di modellazione del dato
- Dal 2020 Founder e CTO presso Cargo BI: start-up di insurance data analytics
- MCP, MCSA, MCSE, MCT SQL Server
- Relatore in diverse conferenze sul tema.
 - info@marcopozzan.it
 - @marcopozzan.it
 - www.marcopozzan.it



















SQL Server 2012/2014



Marco Parenzan

- Senior Solution Architect @ beanTech
- 1nn0va Community Lead (Pordenone)
- Microsoft Azure MVP



- Linkedin: https://www.linkedin.com/in/marcoparenzan/
- Slideshare: https://www.slideshare.net/marco.parenzan
- GitHub: https://github.com/marcoparenzan



















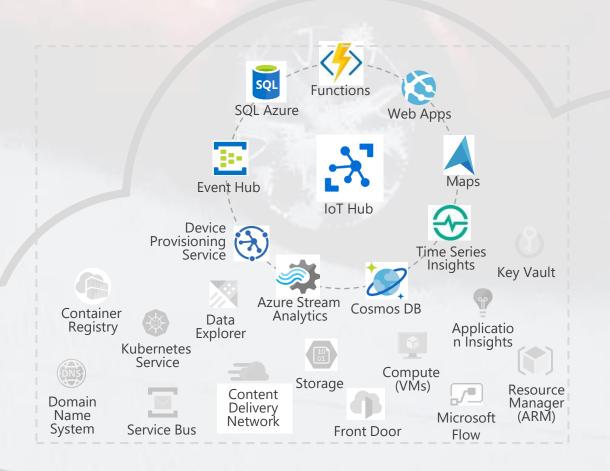
Agenda

- 1. Overview
- 2. Presentazione di Azure IoTCentral
- 3. Setup data export demo
- 4. Power BI Streaming Data Flow
- 5. Setup Data Flow
- 6. Anomaly detection
- 7. Invocare Anomaly detection da power query demo

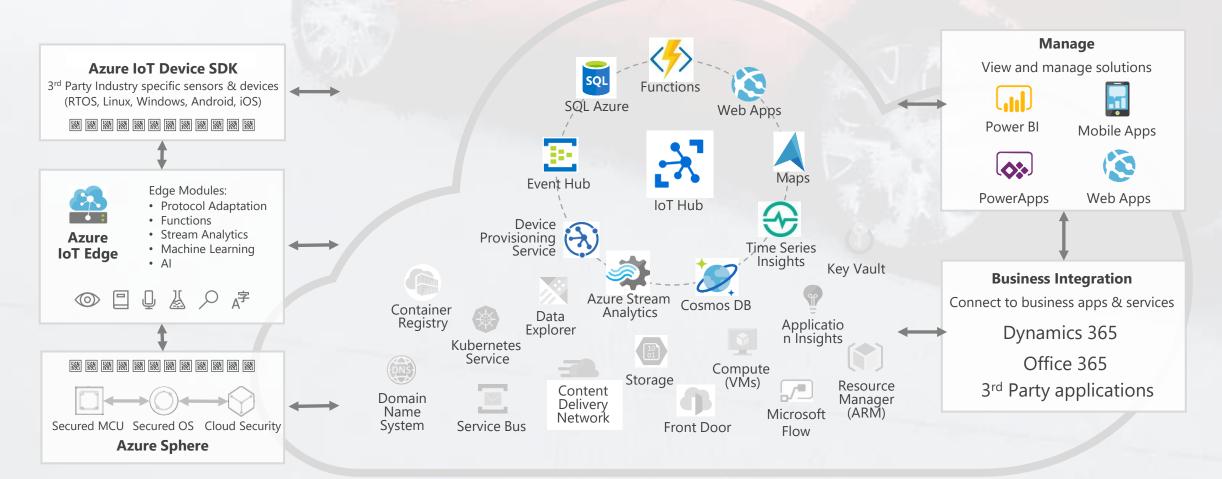
Overview

Facts

- Everyone «needs» IoT
- Everyone receive requests for IoT
- IoT can become complex
- It is difficult to organize a team for that
- Cost issues, manageability issues



Your options for building IoT solutions

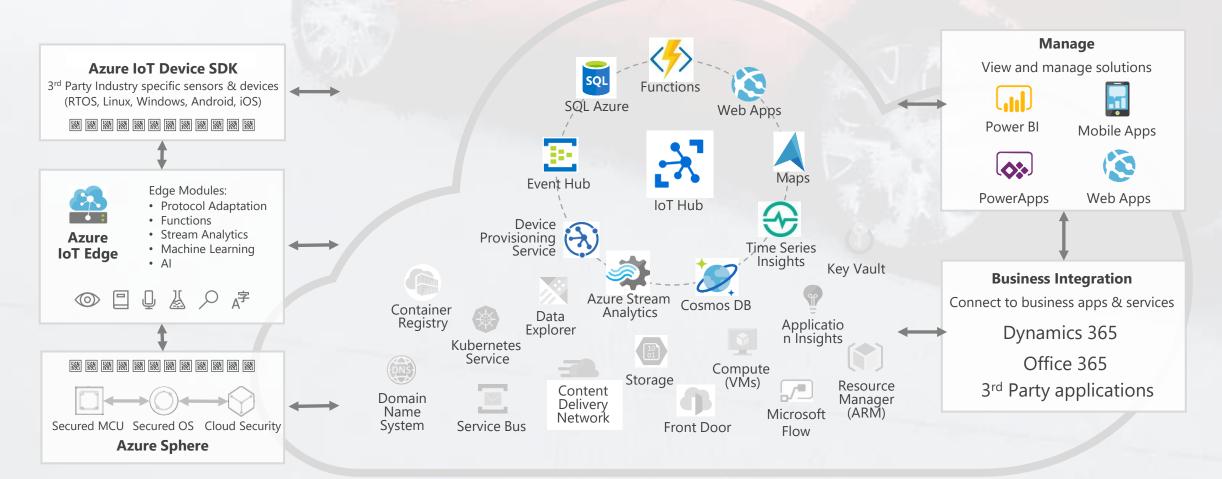


Needs

- Software as a Service approach
- Power BI set the line
- Any option for IoT in this space?

Azure IoTCentral

Your options for building IoT solutions



Your options for building IoT solutions



Azure IoT Central Values

- SaaS approach
- Scenario-based approach
- Pricing per Device (with message tier)
- IoT PnP, Device Provisioning Service
- IoT Hub full compatibility
- Dashboards
- Internal Data Consumption→Single Service
- REST API
- Routing to external paths

Data Logger and Azure IoT Edge



WISE-4250

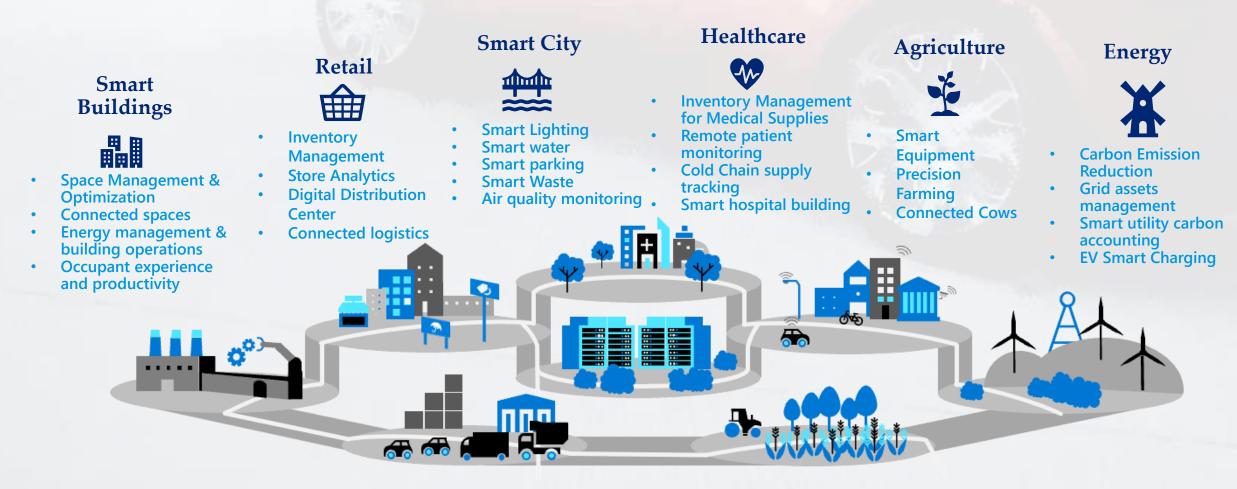
https://devicemodels.azure.com/dtmi/advant ech/wise_4250as_s231-1.expanded.json



ECU-150

https://github.com/Azure/azure-iot-device-ecosystem/blob/master/get_started/embedded-linux-ecu-1251-c.md

IoT Central can help you quickly innovate and build IoT applications for your industry



Azure IoT Central



Get Connected

Connect IoT devices to the cloud faster than with any other platform



Device connectivity and management



Edge support



Plug-and-Play



End to End Security



Stay Connected

Reconfigure and update devices with centralized device management



Telemetry ingestion and command & control



Dashboards, visualization & insights



Monitoring rules & triggered actions



User roles and permissions



Transform

Bridge the gap with connectors and extensibility APIs



Industry-focused
Application Templates



White Labeling

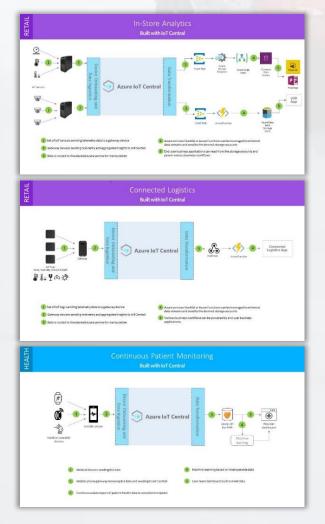


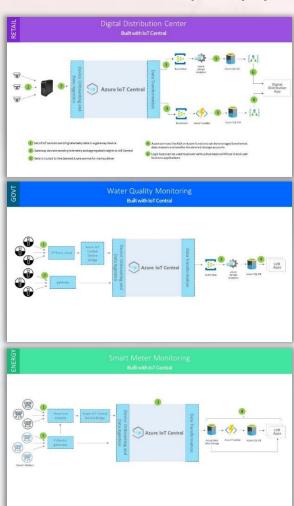
Multitenancy

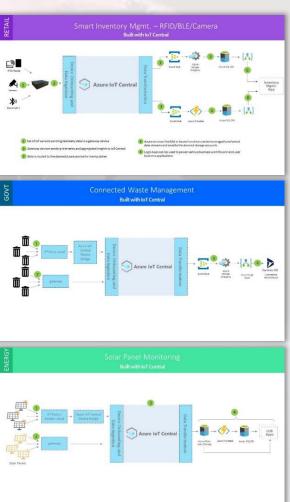


Extensibility & Customization IoT Central public APIs

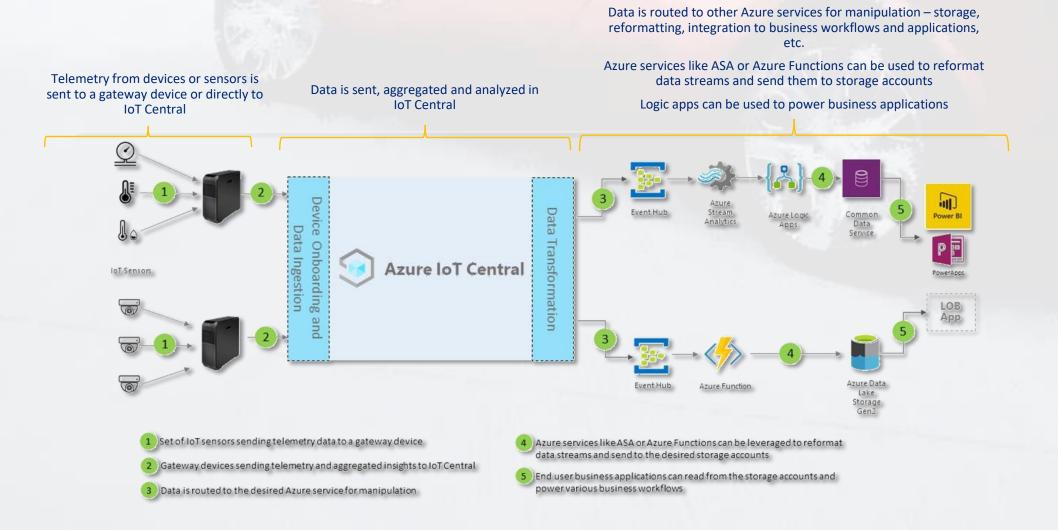
"Build with" patterns (from Industry Application Templates)





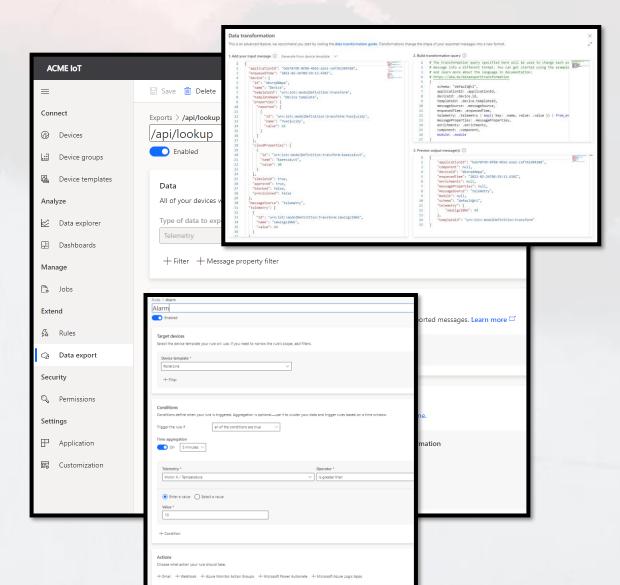


Common Architectures



The integration model of IoT Central

- Built upon the IoT Hub infrastructure
- Advantage from internal data consumption by IoT Central
 - Already include some services you always add
- New from IoT Hub
 - Device Provisioning Service
 - WebHooks
 - Rules (Stream Analytics)
 - REST API to query Central data
 - REST API to manage IoT Central
 - Data Transformation besides enrichment in data routing



Pricing Model

	Standard Tier 0	Standard Tier 1	Standard Tier 2
Price per device	€0,07 al mese	€0,34 al mese	€0,59 al mese
Monthly device message allocation*	400 messages	5,000 messages	30,000 messages
Included free quantity per IoT Central application	2 free devices (800 included messages)	2 free devices (10,000 included messages)	2 free devices (60,000 included messages)
Use case	For devices sending a few messages per hour	For devices sending a few messages per hour	For devices sending a message every few minutes
Overage pricing per 1K messages1	€0.060 per 1K messages	€0.013 per 1K messages	€0.013 per 1K messages
*Total message allocation is shared across all devices in an IoT Central application			

Setup IoT Central Export

4. Power BI Streaming Data Flow

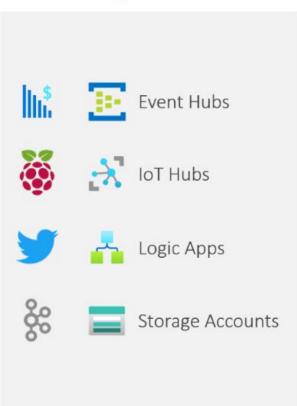
Che cosa sono gli streaming dataflow

reformatting, integration to business workflows and applications, etc. Azure services like ASA or Azure Functions can be used to reformat Telemetry from devices or sensors is data streams and send them to storage accounts Data is sent, aggregated and analyzed in sent to a gateway device or directly to IoT Central Logic apps can be used to power business applications IoT Central Data Transformation Stream Analytics Data Data Ingestion Service Onboarding Azure IoT Central IoT Sensors Event Hub Azure Function Set of IoT sensors sending telemetry data to a gateway device Azure services like ASA or Azure Functions can be leveraged to reformat data streams and send to the desired storage accounts Gateway devices sending telemetry and aggregated insights to IoT Central End user business applications can read from the storage accounts and power various business workflows Data is routed to the desired Azure service for manipulation

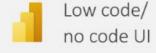
Data is routed to other Azure services for manipulation - storage,

Che cosa sono gli streaming dataflow

Streaming data sources

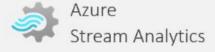


Streaming dataflows





Streaming engine



Power BI

Real-time reports



Hot (real-time) data



Cold (historical) data

Real-time reports



Automatic page refresh and change detection

All Power BI visualizations available

Che cosa sono gli streaming dataflow

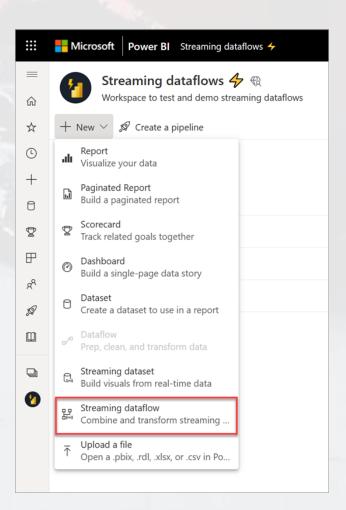
• Il servizio consente esperienze drag-and-drop senza codice.



- streaming dataflow in Power BI consente alle organizzazioni di:
 - Prendere decisioni quasi in tempo reale. Le organizzazioni possono essere più agili e intraprendere azioni significative sulla base delle informazioni più aggiornate.
 - Democratizza i dati in streaming. Le organizzazioni possono rendere i dati più accessibili e più facili da interpretare con una soluzione senza codice riducendo le risorse IT.
 - Accelera il time to insight utilizzando una soluzione di analisi dello streaming end-to-end integrata con l'archiviazione dei dati di Bl.

Creazione di uno streaming dataflow

- Puoi aggiungere e modificare tabelle nel streaming dataflow direttamente dall'area di lavoro in cui è stato creato il flusso di dati.
- A causa della natura dei dati in streaming, c'è un flusso continuo in entrata. L'aggiornamento è costante o infinito a meno che non lo si interrompa
- Puoi avere un solo **un tipo di flusso di dati per area di lavoro.** Se disponi già di un flusso di dati regolare nell'area di lavoro Premium, non potrai creare un flusso di dati in streaming (e viceversa).

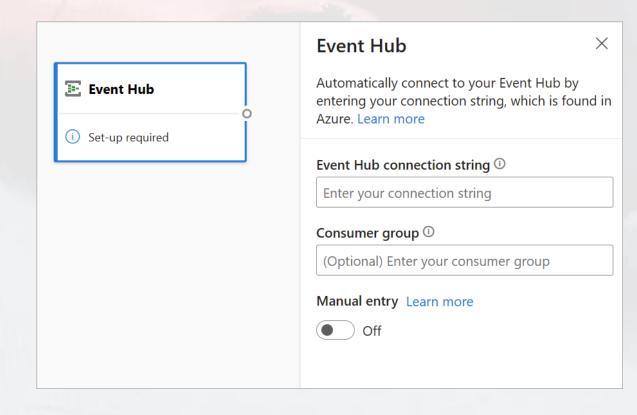


Componenti dello streaming dataflow

- Input
- Trasformazioni
- Output

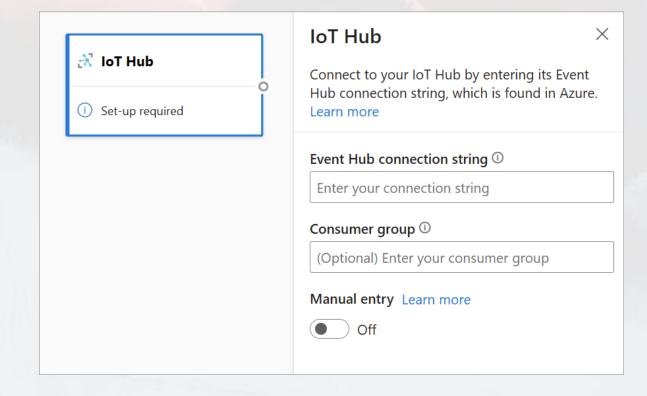
Input: Event Hub

- Event Hub è una piattaforma di streaming per big data e un servizio di importazione di eventi.
- Può ricevere ed elaborare milioni di eventi al secondo.
- I dati inviati a un hub eventi possono essere trasformati e archiviati utilizzando qualsiasi provider di analisi in tempo reale o adattatori di batch/archiviazione.
- Viene visualizzata una scheda nella vista del diagramma, incluso un riquadro laterale per la sua configurazione.



Input: lot Hub

- L'hub IoT è un servizio gestito ospitato nel cloud.
- Funge da hub centrale di messaggi per le comunicazioni in entrambe le direzioni tra un'applicazione loT e i suoi dispositivi collegati.
- Puoi connettere milioni di dispositivi e le relative soluzioni back-end in modo affidabile e sicuro. Quasi tutti i dispositivi possono essere collegati a un hub IoT.



Considerazioni su Event Hub e IoT Hub

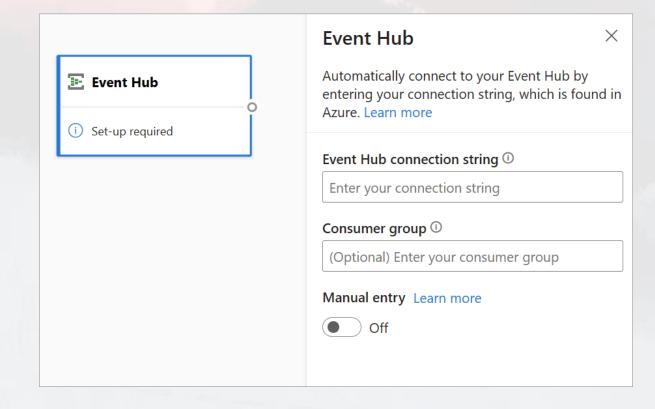
Quando usi i dati di flusso da Hub eventi o Hub IoT, hai accesso ai seguenti campi temporali dei metadati nel flusso di dati di streaming:

- EventProcessedUtcTime : la data e l'ora in cui l'evento è stato elaborato.
- EventEnqueuedUtcTime : la data e l'ora di ricezione dell'evento.

Nessuno di questi campi apparirà nell'anteprima di input. Devi aggiungerli manualmente.

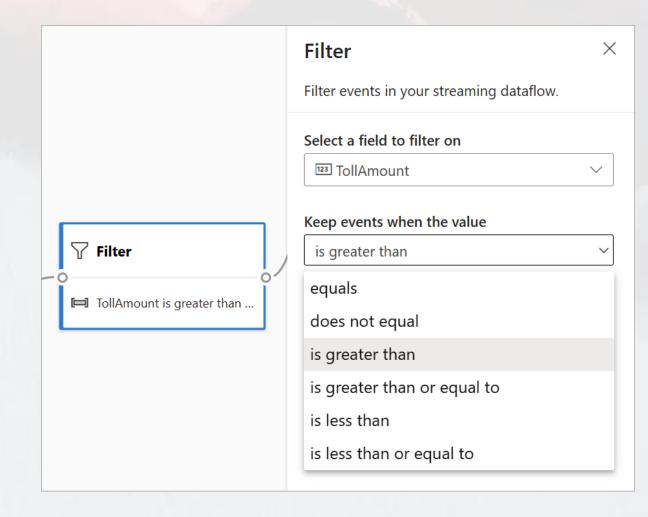
Input: Blob Storage

- L'archiviazione BLOB è ottimizzata per l'archiviazione di enormi quantità di dati non strutturati.
- Possiamo usare i BLOB di Azure come input di streaming/riferimento.
- I BLOB in streaming vengono generalmente controllati ogni secondo per streaming. Mente i BLOB di reference (<= 50 MB) viene caricato solo all'inizio dell'aggiornamento



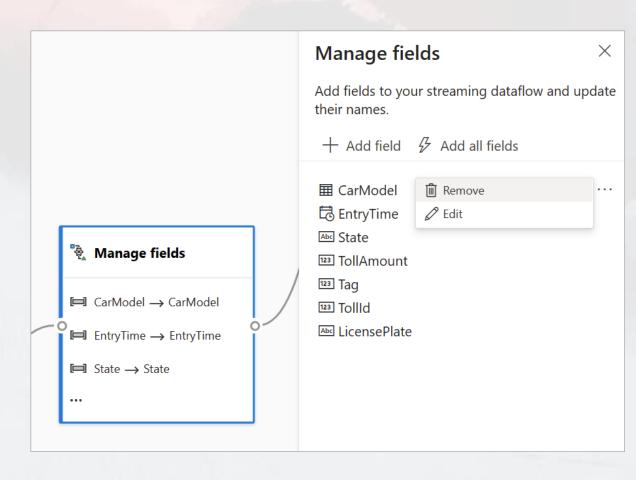
Trasformazione: Filtro

- Utilizzare la trasformazione **Filtro** per filtrare gli eventi in base al valore di un campo nell'input.
- A seconda del tipo di dati (numero o testo), la trasformazione manterrà i valori che corrispondono alla condizione selezionata.



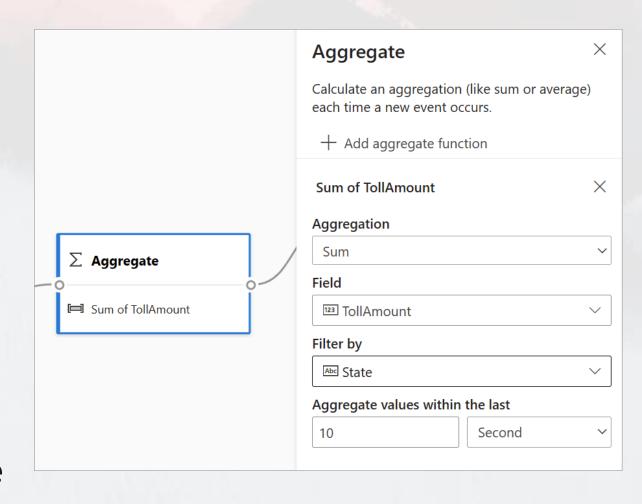
Trasformazione: Gestisci i campi

- La trasformazione **Gestisci campi** consente di aggiungere, rimuovere o rinominare i campi provenienti da un input o da un'altra trasformazione.
- Le impostazioni nel riquadro laterale ti danno la possibilità di aggiungerne uno nuovo selezionando **Aggiungi** campo o aggiungendo tutti i campi contemporaneamente.



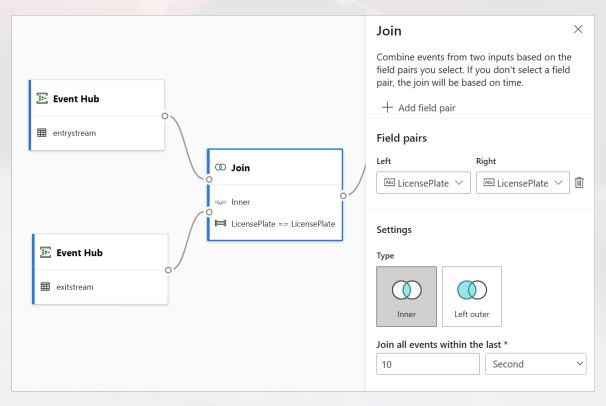
Trasformazione: Aggregato

- È possibile utilizzare la trasformazione **Aggrega** per calcolare un'aggregazione (**Sum**, **Minimum**, **Maximum** o **Average**) ogni volta che si verifica un nuovo evento in un periodo di tempo.
- Puoi avere una o più aggregazioni nella stessa trasformazione.
- Possiamo aggiungere il periodo di tempo durante il quale verrà calcolata l'aggregazione e filtrare per altra dimensione



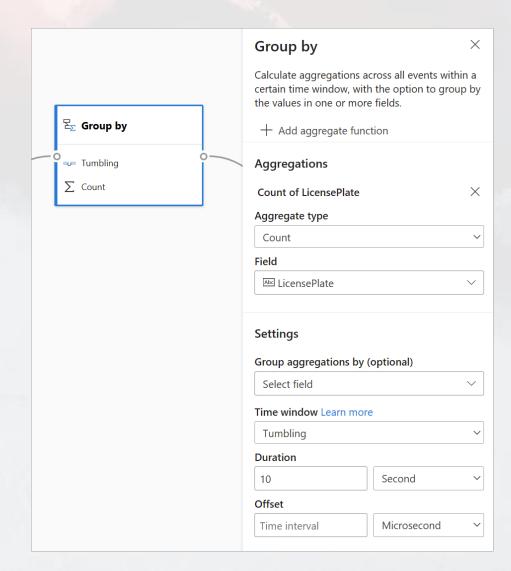
Trasformazione: Aggiungi (Join)

- Usare il Join per combinare eventi da due input in base alle coppie di campi selezionate. Come per i join regolari, hai diverse opzioni per la logica di join:
 - Inner join
 - Left join
- Infine, seleziona in quale periodo di tempo desideri che venga calcolato il join. In questo esempio, il join esamina gli ultimi 10 secondi.
- Per impostazione predefinita, sono inclusi tutti i campi di entrambe le tabelle.

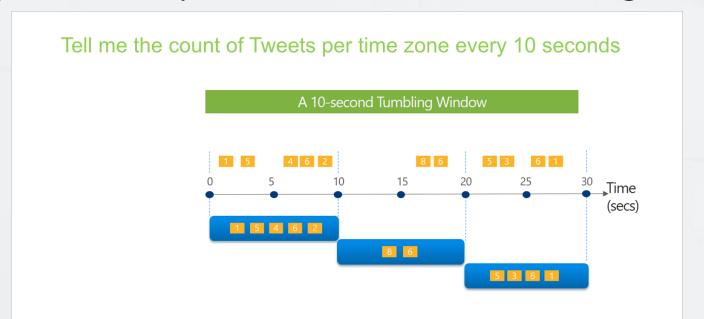


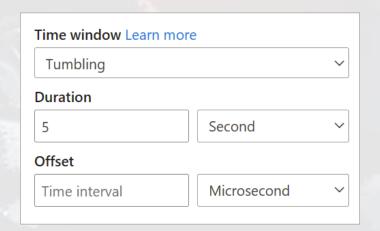
Trasformazione: Raggruppa per

- Utilizzare la trasformazione **Raggruppa per** per calcolare le aggregazioni di tutti gli eventi arrivati entro un determinato intervallo di tempo.
- Include anche opzioni di finestre temporali più complesse. Puoi aggiungere più di un'aggregazione per trasformazione.
- Le aggregazioni disponibili in questa trasformazione sono: Media, Conteggio, Massimo, Minimo, Percentile (continuo e discreto), Deviazione standard, Somma e Varianza.



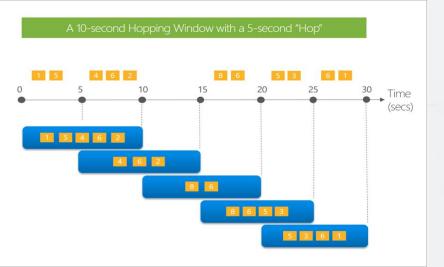
• Il tipo di finesta **tumbling (cascata)** è il tipo più comune di finestra temporale. Le caratteristiche chiave delle finestre a tumbling sono che si ripetono, hanno la stessa durata (**duration**) e non si sovrappongono. Un evento non può appartenere a più di una finestra di tumbling.

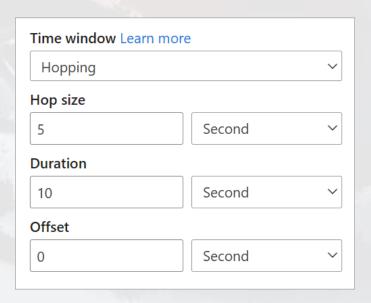




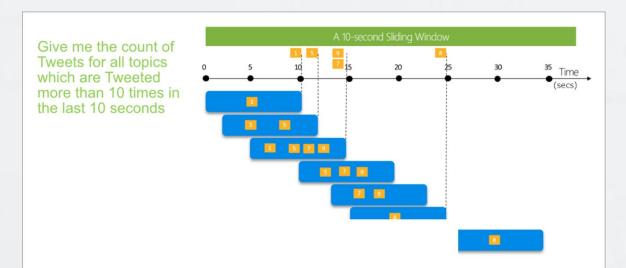
• Le finestre **hopping** (**salto**) servono per saltare in avanti nel tempo di un periodo fisso. Puoi pensarli come finestre **tumbling** che possono sovrapporsi ed essere emesse più spesso delle dimensioni della finestra. Gli eventi possono appartenere a più di un set di risultati per una finestra **hopping**.

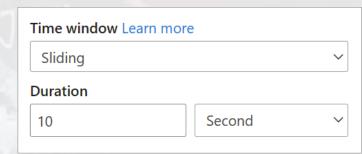
Every 5 seconds give me the count of Tweets over the last 10 seconds



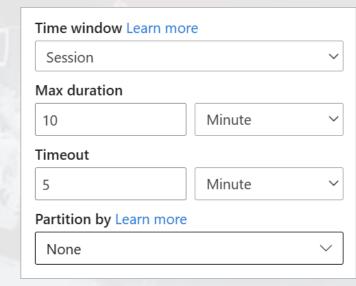


• Le finestre **sliding** (**scorrevole**), a differenza delle finestre a **tumbling** o **hopping**, calcolano l'aggregazione solo per i momenti in cui il contenuto della finestra cambia effettivamente. Quando un evento entra o esce dalla finestra, viene calcolata l'aggregazione. Quindi, ogni finestra ha almeno un evento.

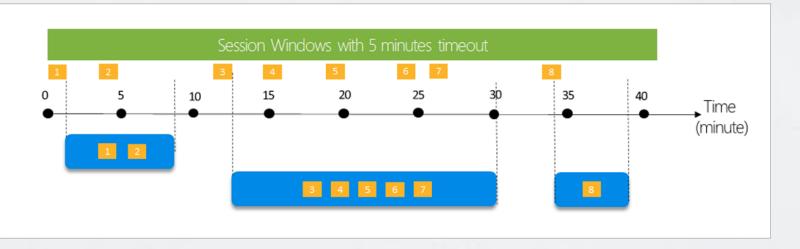




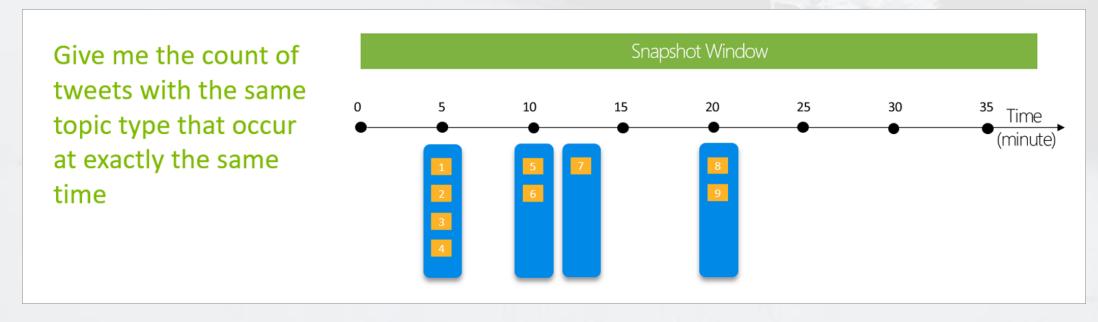
- Le finestre di **session (sessione)** sono il tipo più complesso. Raggruppano eventi che arrivano in orari simili, filtrando i periodi di tempo in cui non ci sono dati. Per questo è necessario fornire:
 - Un timeout: quanto aspettare se non ci sono nuovi dati.
 - Una durata massima: il tempo più lungo in cui verrà calcolata l'aggregazione se i dati continuano ad arrivare.
- Puoi anche definire una partizione, se lo desideri.



Tell me the count of Tweets that occur within 5 minutes to each other

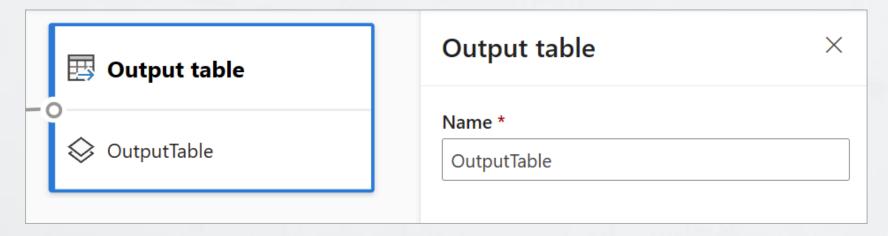


• Le finestre di **snapshot (istantanea)** raggruppano gli eventi che hanno lo stesso timestamp. A differenza di altre finestre, uno snapshot non richiede alcun parametro perché utilizza il tempo del sistema.



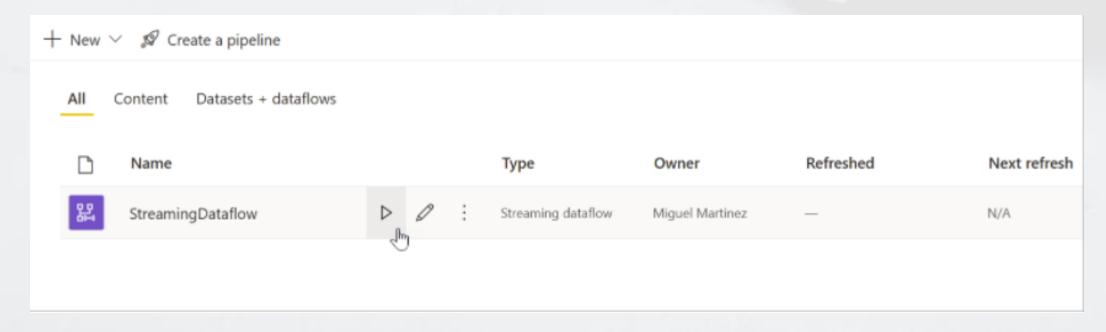
Output: Table

- Dopo che sei pronto con input e trasformazioni, è il momento di definire uno o più output. A partire da luglio 2021, i flussi di dati in streaming supportano un solo tipo di output: una tabella Power Bl.
- Questo output sarà una tabella del flusso di dati (ovvero un'entità) che puoi usare per creare report in Power BI Desktop. Devi unire i nodi del passaggio precedente con l'output che stai creando per farlo funzionare. Dopodiché, tutto ciò che devi fare è nominare la tabella.



Eseguire un flusso

• Per avviare il flusso di dati in streaming, prima salva il flusso di dati e vai all'area di lavoro in cui lo hai creato. Passa il mouse sopra il streaming dataflow e seleziona il pulsante di riproduzione che appare. Un messaggio pop-up informa che il flusso di dati in streaming è in fase di avvio.



Connettersi a un streaming dataflow

• Con la versione di luglio 2021 di Power BI Desktop, è disponibile un nuovo connettore denominato **Dataflow** di dati da usare. Come parte di questo nuovo connettore, per lo streaming dei flussi di dati, vedrai due tabelle che corrispondono all'archiviazione dei dati descritta in

precedenza.

Get Data	
Search	Power Platform
All	Power BI datasets
File	Power BI dataflows
Database	Common Data Service (Legacy)
Power Platform	Dataverse
Azure	P Dataflows
Online Services	
Other	

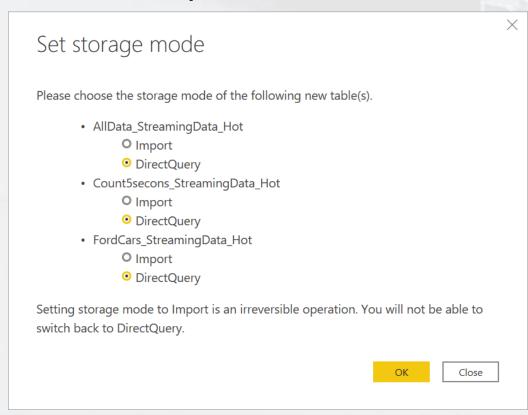
Connettersi a un streaming dataflow

I Streaming Dataflow salvano i dati nelle due posizioni seguenti. L'uso di queste fonti dipende dal tipo di analisi che stai cercando di fare:

- Hot (analisi in tempo reale): quando i dati arrivano in Power BI dai Streaming Dataflow, i dati vengono archiviati in una posizione attiva a cui puoi accedere con elementi visivi in tempo reale. La quantità di dati salvata in questo archivio dipende dal valore definito per Durata conservazione nelle impostazioni Streaming Dataflow. Il valore predefinito (e minimo) è 24 ore.
- Cold (analisi storica): qualsiasi periodo di tempo che non rientra nel periodo definito per Durata di conservazione viene salvato in conservazione a freddo (BLOB) in Power Bl per essere utilizzato se necessario. (solo con il connettore dei flussi di dati (Beta) di Power Platform)

Connettersi a un streaming dataflow

 Quando ti viene chiesto di scegliere una modalità di archiviazione, seleziona DirectQuery se il tuo obiettivo è creare oggetti visivi in tempo reale.



Limitazioni

- Per creare ed eseguire streaming dataflow è necessario un abbonamento a Power BI Premium (capacità o PPU).
- È consentito un solo tipo di flusso di dati per area di lavoro (streaming o regolare).
- Non è possibile collegare dataflow regolari e streaming dataflow.
- Capacità inferiori ad A3 non consentono l'uso di streaming dataflow.
- Se i dataflow o il motore di calcolo avanzato non sono abilitati in un tenant, non è possibile creare o eseguire streaming dataflow.
- Le aree di lavoro connesse a un account di archiviazione non sono supportate.
- Ciascun streaming dataflow può fornire fino a 1 megabyte al secondo di throughput.

Licenze

- Il numero di streaming dataflow consentiti per tenant dipende dalla licenza utilizzata:
- Per le capacità Premium, utilizzare la formula seguente per calcolare il numero massimo di streaming dataflow consentiti in una capacità:
 - Numero massimo di flussi di dati in streaming per capacità = vCore nella capacità x 5
 - Ad esempio, P1 ha 8 vCore: 8 * 5 = 40 flussi di dati in streaming.
- Per Premium per utente, è consentito un flusso di dati in streaming per utente. Se un altro utente desidera utilizzare un flusso di dati in streaming in un'area di lavoro PPU, avrà bisogno anche di una licenza PPU.

Setup Data Flow

Anomaly Detection

Anomaly Detection

- Anomaly detection is the process of identifying unexpected items or events in data sets, which differ from the norm.
- And anomaly detection is often applied on unlabeled data which is known as unsupervised anomaly detection.

Anomaly Detection in Time Series

- In time series data, an anomaly or outlier can be termed as a data point which is not following the common collective trend or seasonal or cyclic pattern of the entire data and is significantly distinct from rest of the data. By significant, most data scientists mean statistical significance, which in order words, signify that the statistical properties of the data point is not in alignment with the rest of the series.
- Anomaly detection has two basic assumptions:
- Anomalies only occur very rarely in the data.
- Their features differ from the normal instances significantly.

Many Opportunities in Azure

- Programmatic:
 - ML.NET (in .NET space)
 - All Python space
- Data Platform
 - Spark and DataBricks
 - Azure Data Explorer
- All these solutions are in the PaaS space
- Any option in SaaS space?

Azure Cognitive Services

- Cognitive Services brings AI within reach of every developer—without requiring machine-learning expertise. All it takes is an API call to embed the ability to see, hear, speak, search, understand, and accelerate decision-making into your apps. Enable developers of all skill levels to easily add AI capabilities to their apps.
- Five areas:
- Decision
- Language
- Speech
- Vision
- Web search

Anomaly Detector

Identify potential problems early on.

Content Moderator

Detect potentially offensive or unwanted content.

Metrics Advisor PREVIEW

Monitor metrics and diagnose issues.

Personalizer

Create rich, personalized experiences for every user.

Anomaly Detector

 Through an API, Anomaly Detector ingests time-series data of all types and selects the best-fitting detection model for your data to ensure high accuracy. Customize the service to detect any level of anomaly and deploy it where you need it most -- from the cloud to the intelligent edge with containers. Azure is the only major cloud provider that offers anomaly detection as an AI service.

It seems too much simple... ®

Anomaly Detector Features

- Online sample
 - https://algoevaluation.azurewebsites.net/
- Notebooks
- .NET
- REST API
- Univariate and Multivariate analysis
- Univariate analysis
 - LastPoint
 - Batch
 - Trend
- Multivariate analysis

Invocare Anomaly detection API

Invocare Anomaly detection da power query - demo

Conclusioni

