

MU5IN852
Bases de Données Large Echelle

data streaming et requêtes

Octobre 2020

Objectifs

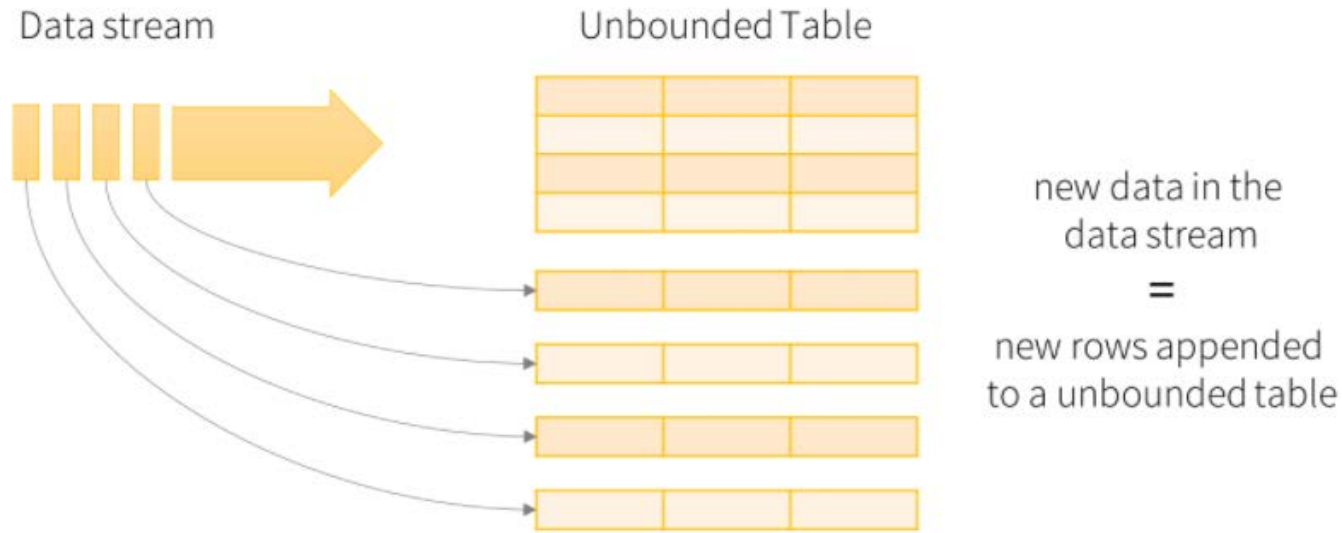
- Aperçu du data streaming
- Notion de fenêtre sur des flux
- Notion de requêtes continues
 - Jointures sur des fenêtres
- Perspectives

Biblio

- SIGMOD 2018
 - Titre: Structured Streaming: A Declarative API for Real-Time Applications in Apache Spark
 - Auteurs: Armbrust et al : Databricks, Stanford Univ
 - https://databricks.com/wp-content/uploads/2018/12/sigmod_structured_streaming.pdf

Contexte : Flux

- Données produites en continu
- Estampille
 - attribut **date d'événement** et/ou date d'arrivée



Data stream as an unbounded table

Motivations et Défis

- Requêtes incrémentales complexes à exprimer
 - Besoin de langage déclaratif
- Chaîne de traitement intégrée
- Défis opérationnels
 - Pannes et retards dus aux stragglers
 - tâches « à la traine »
 - Mise à jour des applis traitant un stream
 - Redimensionnement des ressources allouées
- Métrique de performance
 - Débit : nombre de tuples traités par minute
versus
 - Latence : temps de réponse d'une requête
 - date du résultat de la requête – date d'arrivée de la donnée

Système de streaming

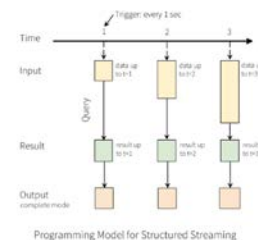
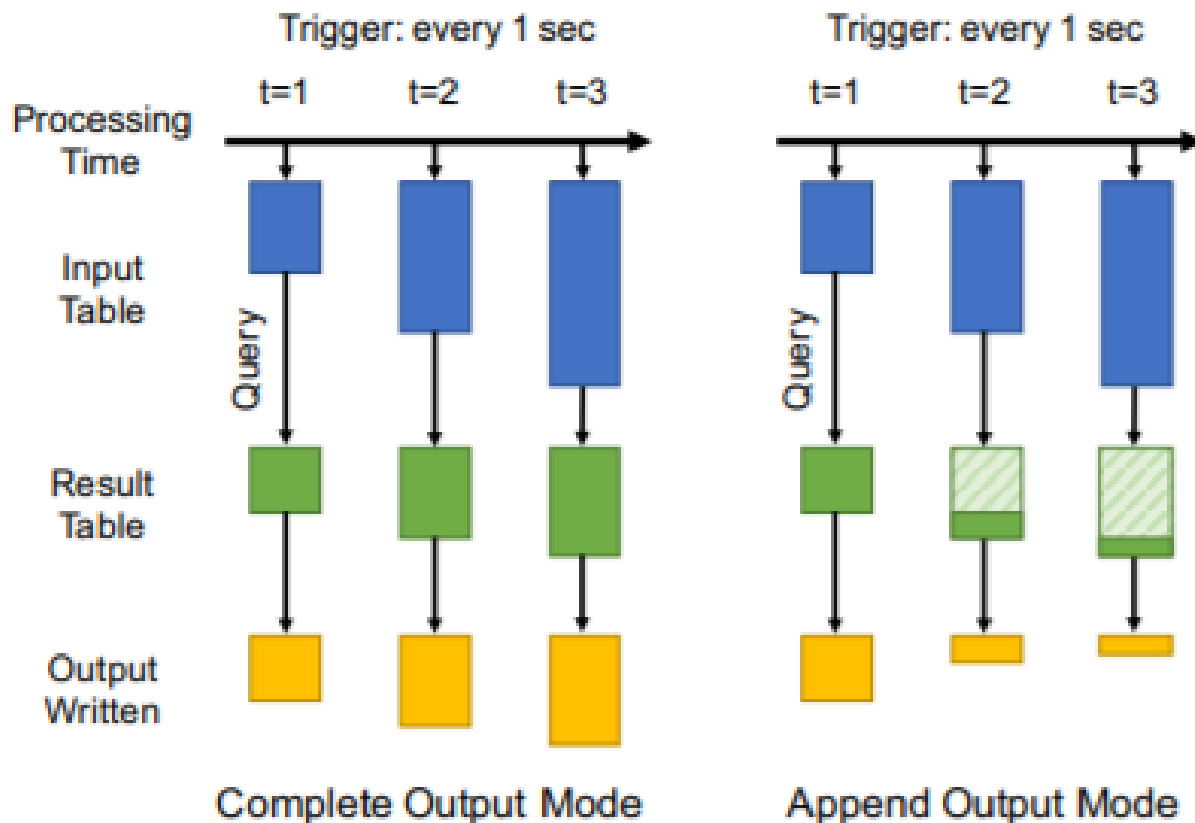
- Système de gestion des flux et des requêtes
 - Scalable : architecture distribuée
- Gestion des flux entrants
 - Tolérance aux pannes
 - Stockage temporaire des flux
 - Possibilité de répéter l'arrivée d'un flux
 - Propriété sémantique
 - Chaque tuple arrive une et une seule fois
 - Appelée : exactly once
 - Exemple : Kafka
- Gestion des flux sortants
 - Tolérance aux pannes
 - Ecriture idempotente : 1 ou plusieurs invocations d'une écriture produit le même résultat

Modèle d'exécution des flux

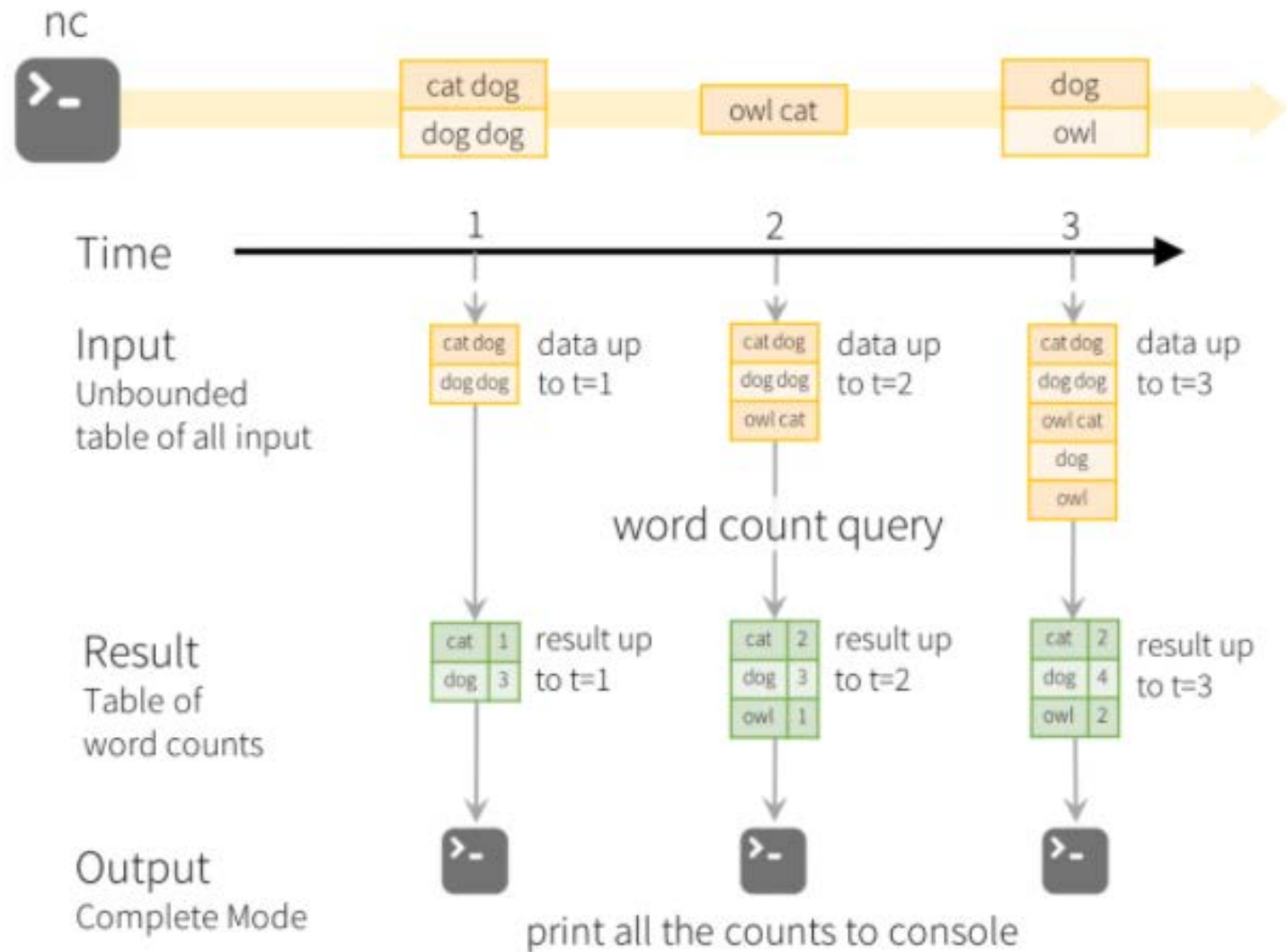
- Source
 - Ensemble ordonné de tuples
 - Taille infinie
 - Ordre partiel si une source est distribuée
- Requête
 - Exécutée à l'instant t
 - Posée sur toutes les données d telles que $d.date \leq t$
 - Exécution **périodique** ou continue
- Sortie : 3 modes possibles
 - Résultat complet à chaque instant t
 - Ajout de nouveaux tuples dans le résultat
 - Mise à jour du résultat

Modèle d'exécution périodique

Exécution appelée micro-batch



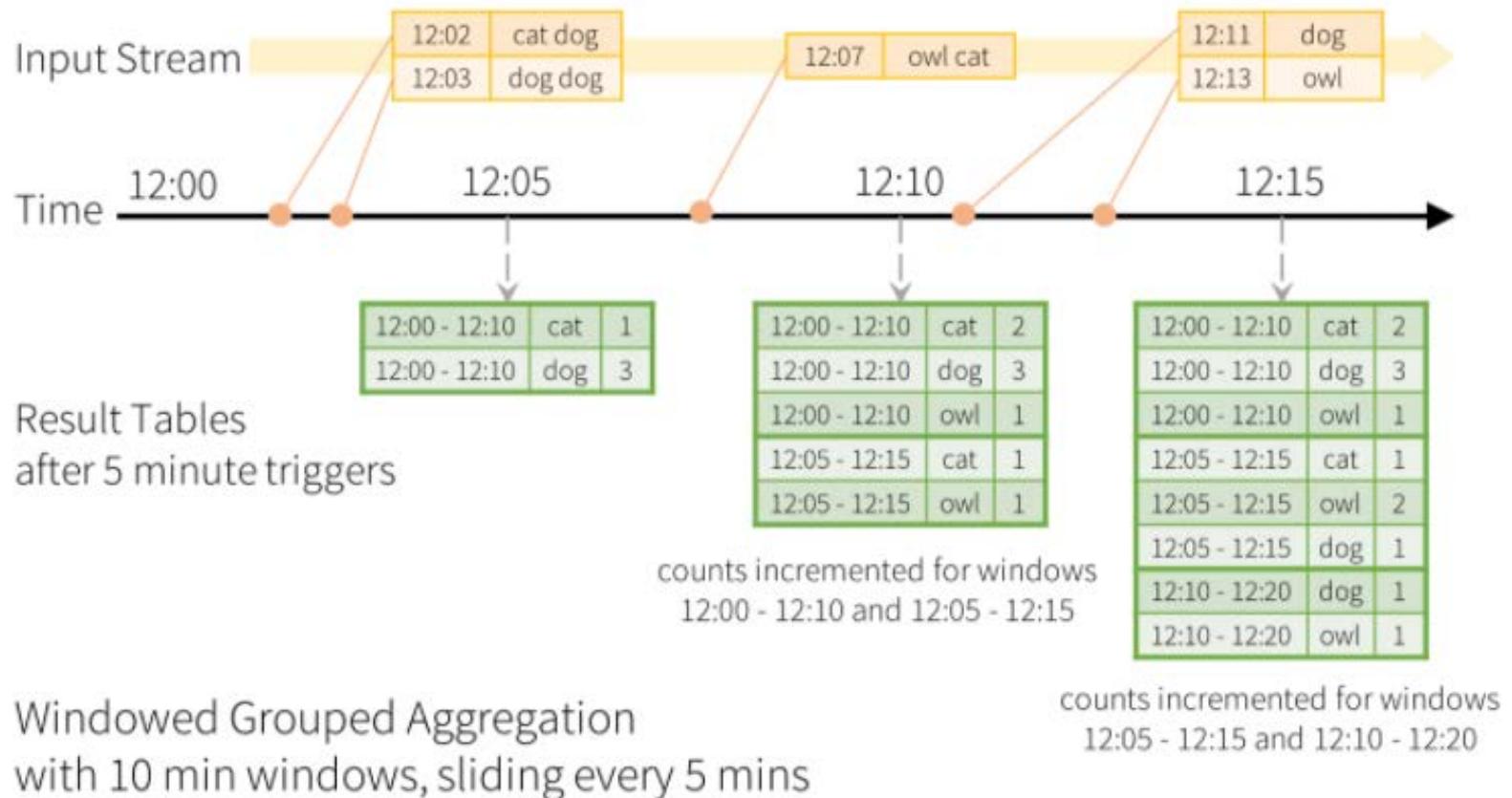
Example : word count sur un flux



Fenêtrage temporel

- Basé sur la date de l'événement
 - Attribut du flux entrant
- GROUP BY WINDOW attribut taille décalage
 - Attribut de type date
 - Taille de la fenêtre
 - Décalage entre deux dates consécutives de début de fenêtre
- Superposition des fenêtres si décalage < taille

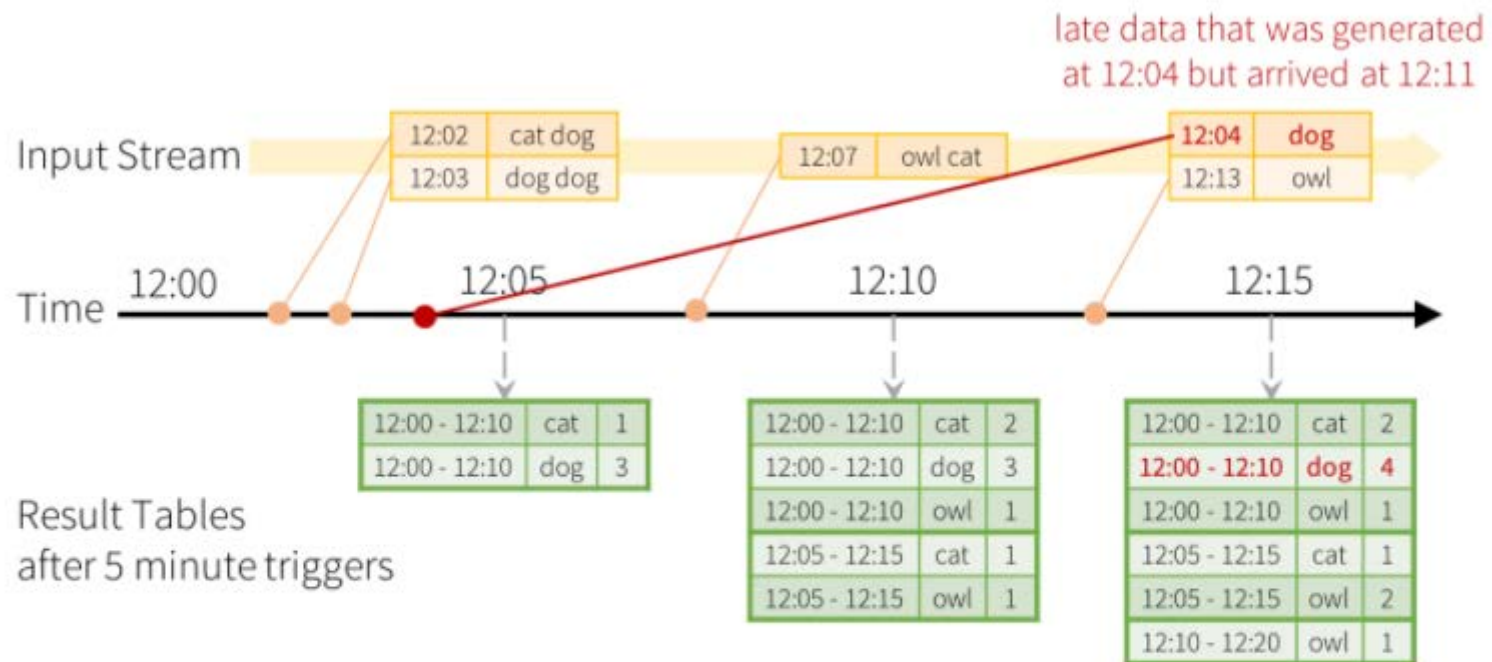
Exemple de fenêtrage



Flux borné

- Problème : Garantir la complétude des résultats sur des flux potentiellement infinis ?
- Solution : spécifier une contrainte sur la date d'événement par rapport à la date courante
 - Taille d'une plage de validité du flux
 - Date événement > date d'événement la plus récente – taille plage valide
- Notion appelée Watermarking
- Syntaxe : WITH WATERMARK attribut, taille
 - Attribut de type date
 - Taille de la plage « valide »

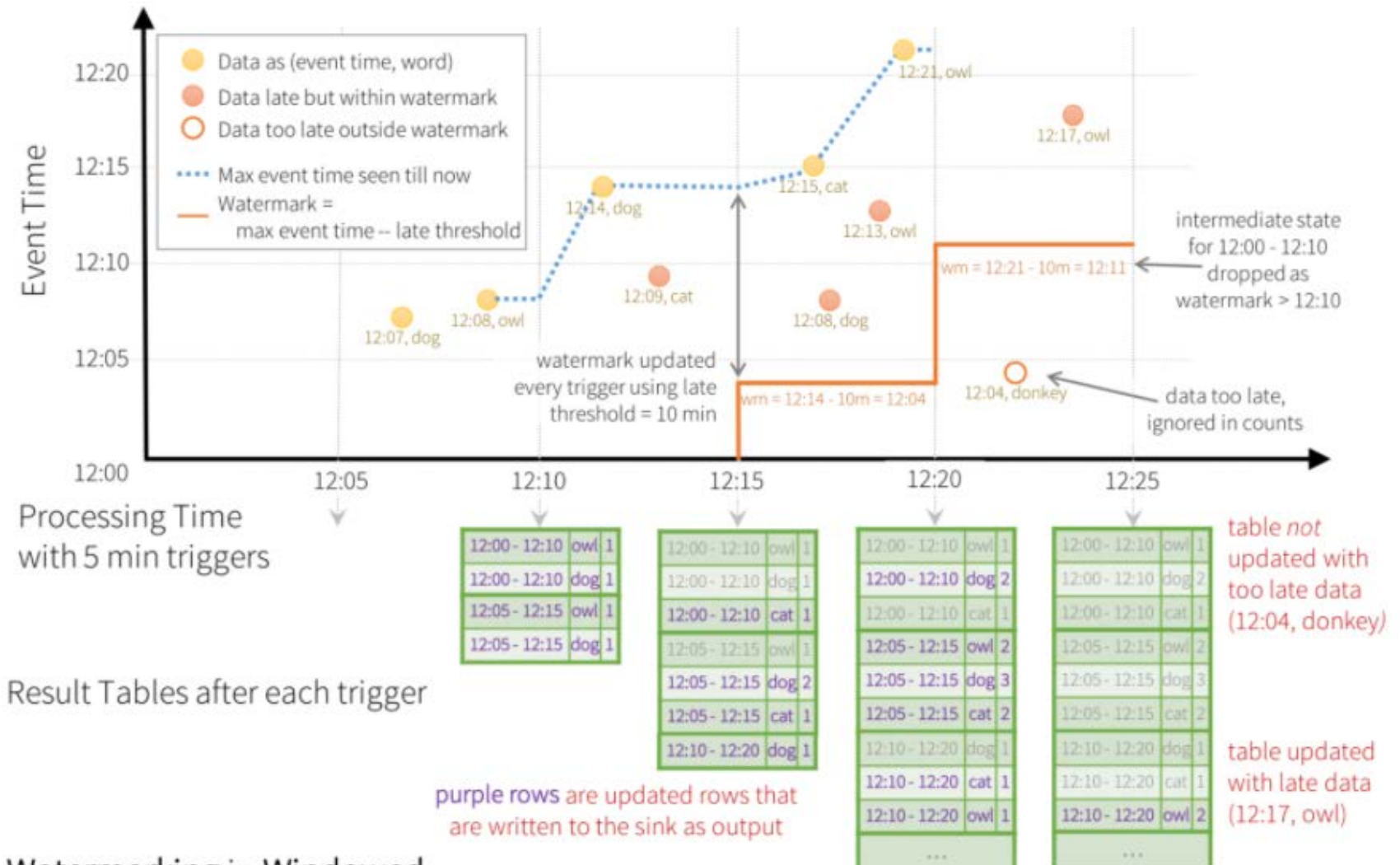
Flux borné : Exemple 1/3



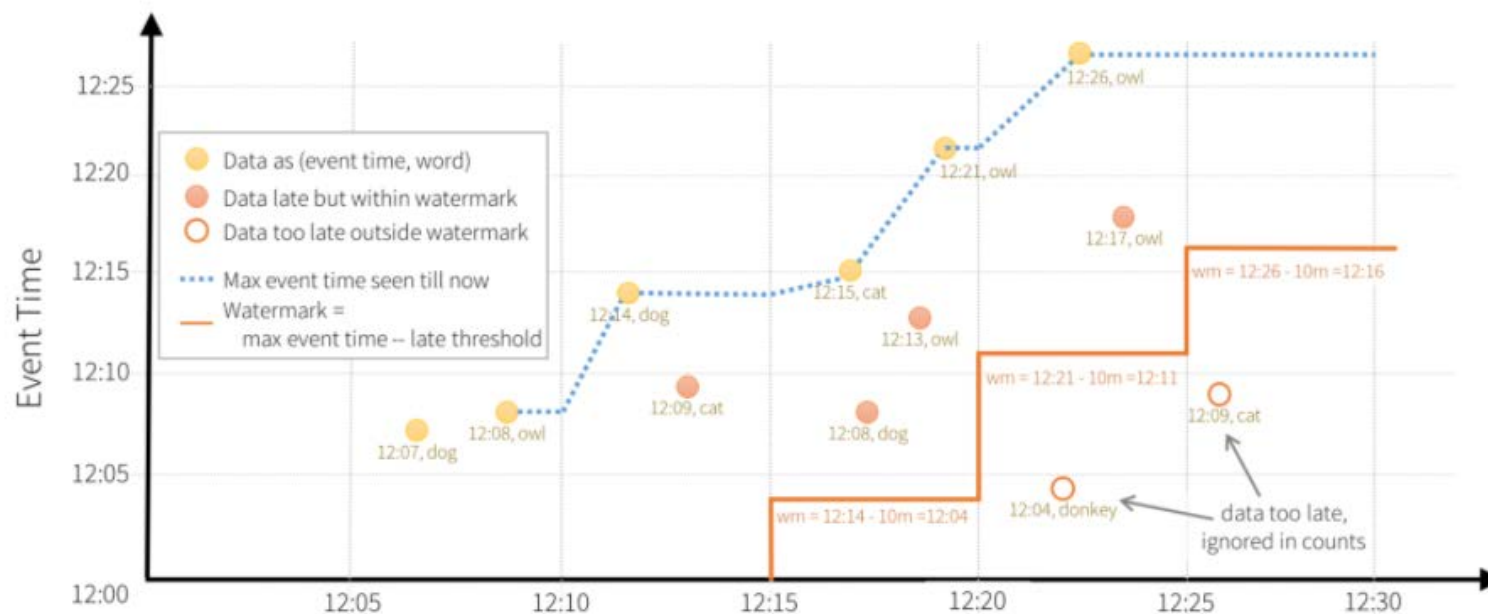
counts incremented only for window 12:00 - 12:10

Late data handling in
Windowed Grouped Aggregation

Example 2/3



Example 3/3



partial counts for window 12:00 - 12:10 maintained as internal state while waiting for late data, so not yet added to result table

final counts for 12:00 - 12:10 added to table when watermark > 12:10, late data counted, and intermediate state for window dropped

12:00 - 12:10	owl	1
12:00 - 12:10	cat	1
12:00 - 12:10	dog	2

12:00 - 12:10	owl	1
12:00 - 12:10	cat	1
12:00 - 12:10	dog	2
12:05 - 12:15	owl	2
12:05 - 12:15	cat	2
12:05 - 12:15	dog	3

Result Tables after each trigger

Watermarking in Windowed Grouped Aggregation with Append Mode

Jointure sur des flux

- Jointure entre 1 flux et une table
- Jointure entre 2 flux
 - Jointure naturelle
 - Jointure externe ?

Extensibilité

- Fonction définie par l'utilisateur
- Evaluation incrémentale sur des fenêtre avec recouvrement
 - **Etat à maintenir** entre deux évaluations de requêtes

Biblio et perspectives

- SIGMOD
 - Structured Streaming
- Spark Structured Streaming
 - Programming guide
 - <https://spark.apache.org/docs/latest/structured-streaming-programming-guide.html>
- Académique :
 - VLDB 2015 : Google DataFlow Model
 - BigData 2018 : BigSR: real-time expressive RDF stream reasoning on modern Big Data platforms
 - <https://ieeexplore.ieee.org/document/8621947>

