

# Procesamiento streaming y en tiempo real

## Sistemas de Big Data

Ricardo García Ródenas  
Ricardo.Garcia@uclm.es



Sistemas de  
Big\_Data



Curso Especialización  
Inteligencia\_Artificial y  
Big\_Data

# Procesamiento streaming

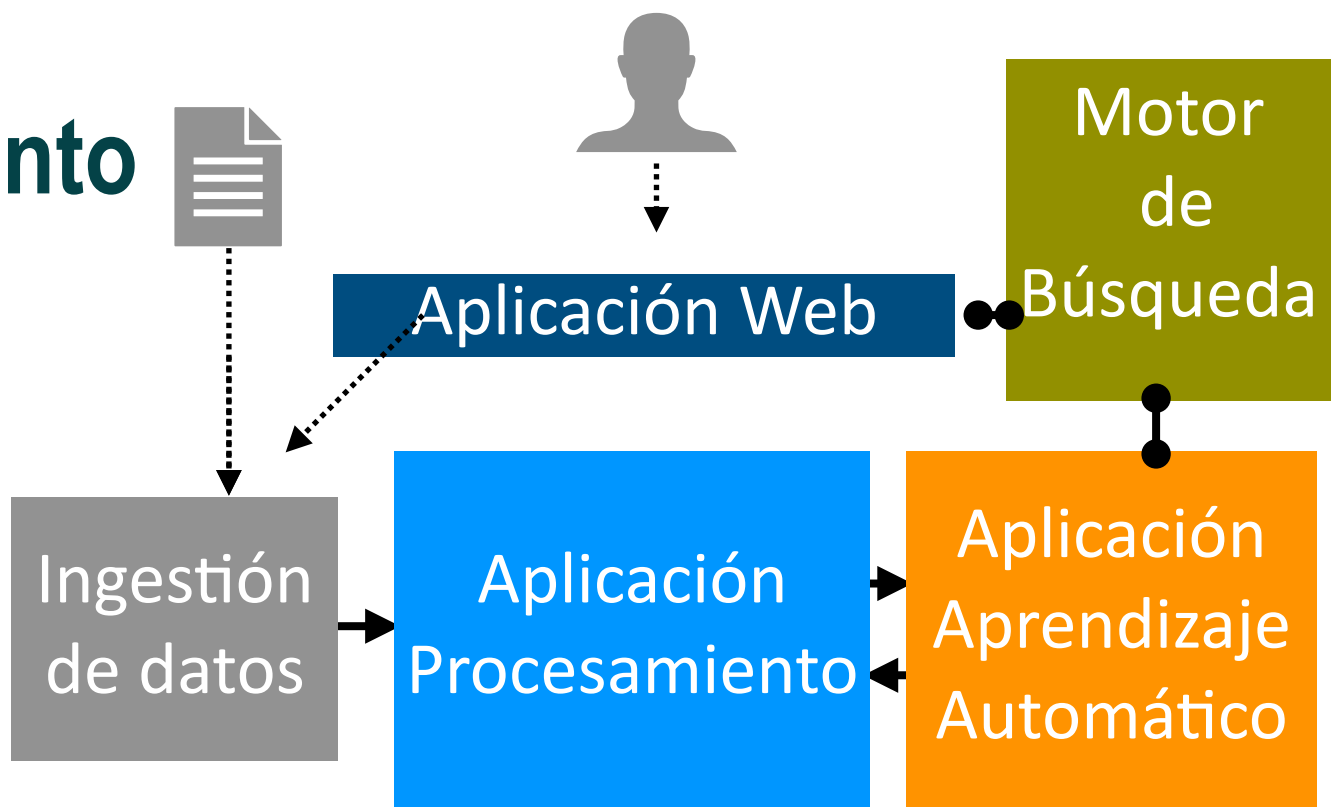
## Características

- Flujos de datos
- No hay limitación en **tiempo de procesamiento**
- No hay limitación en **tiempo generación output**
- Suficiente memoria
- Tasa de procesamiento  $>$  Tasa de entrada



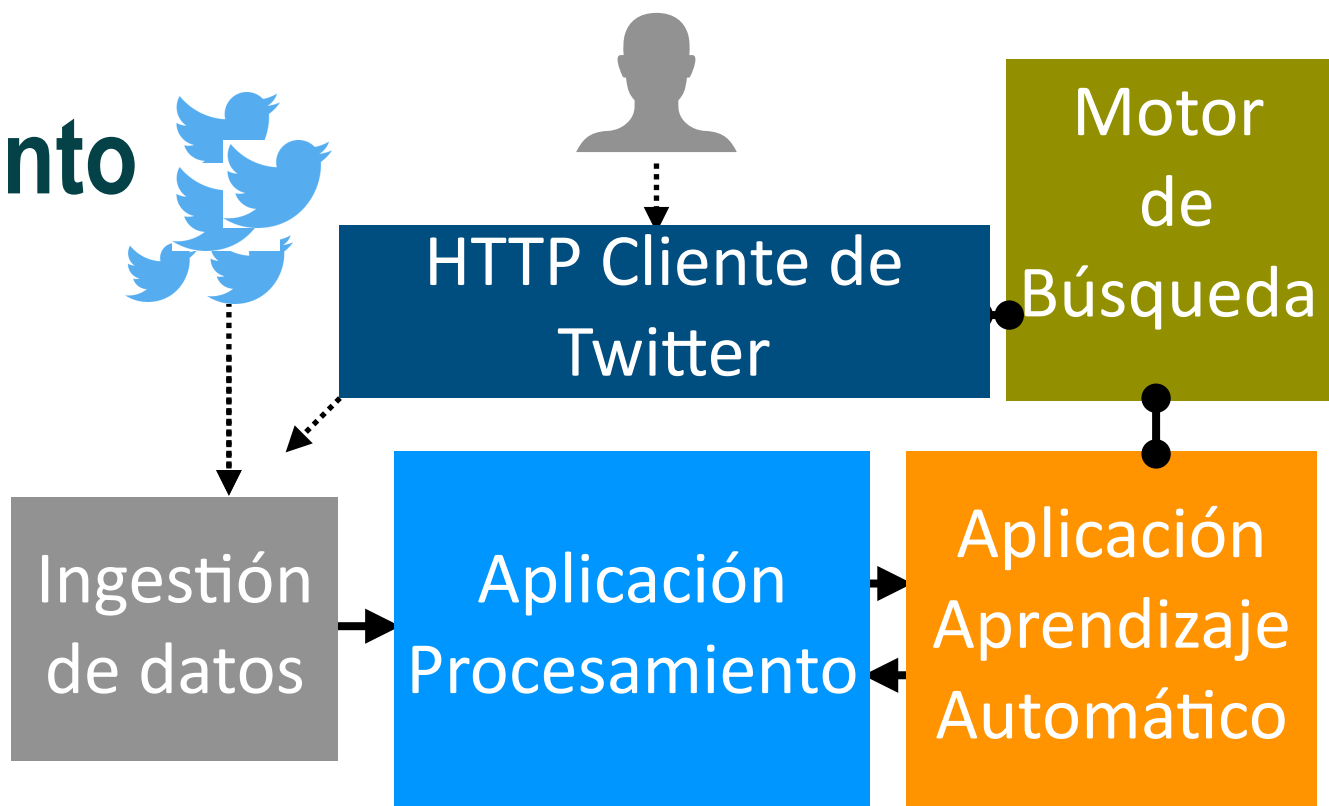
# Procesamiento streaming

## Ejemplo



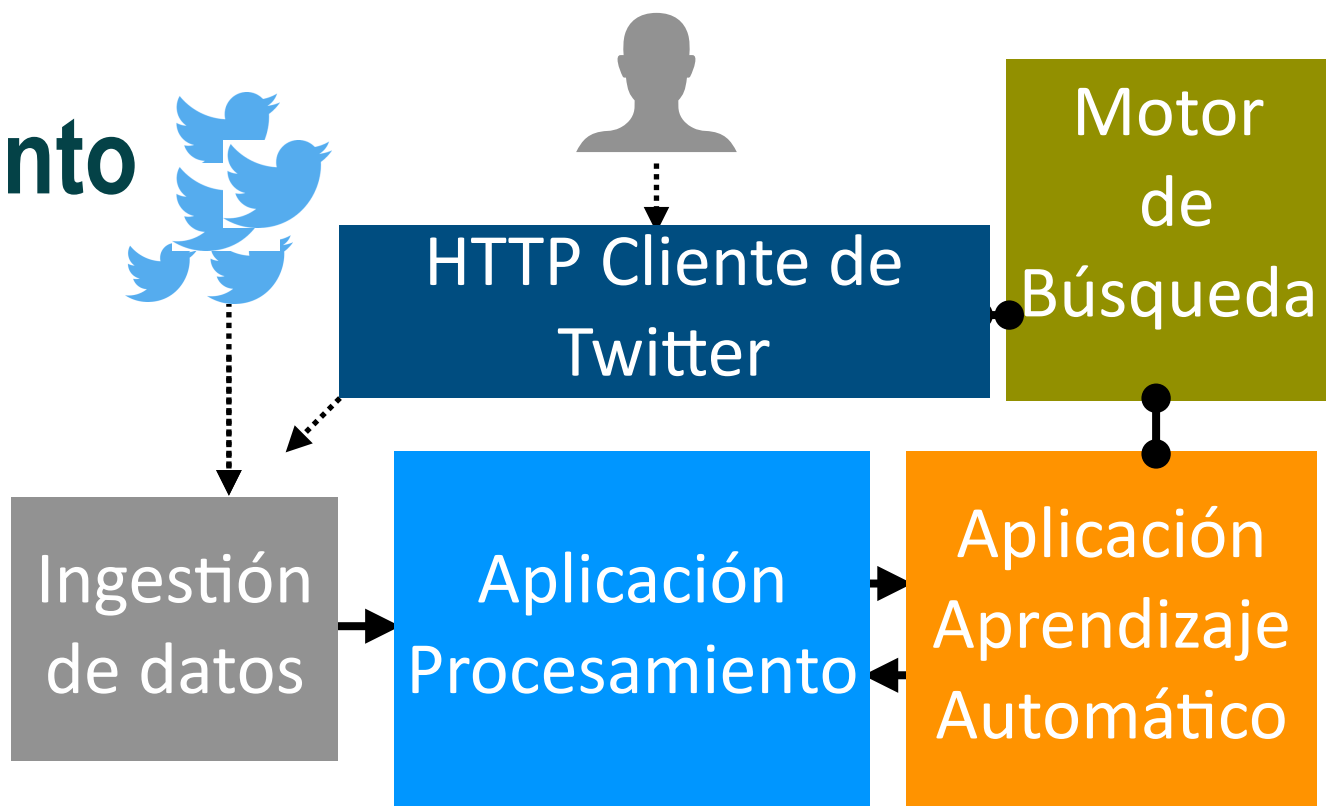
# Procesamiento streaming

## Ejemplo



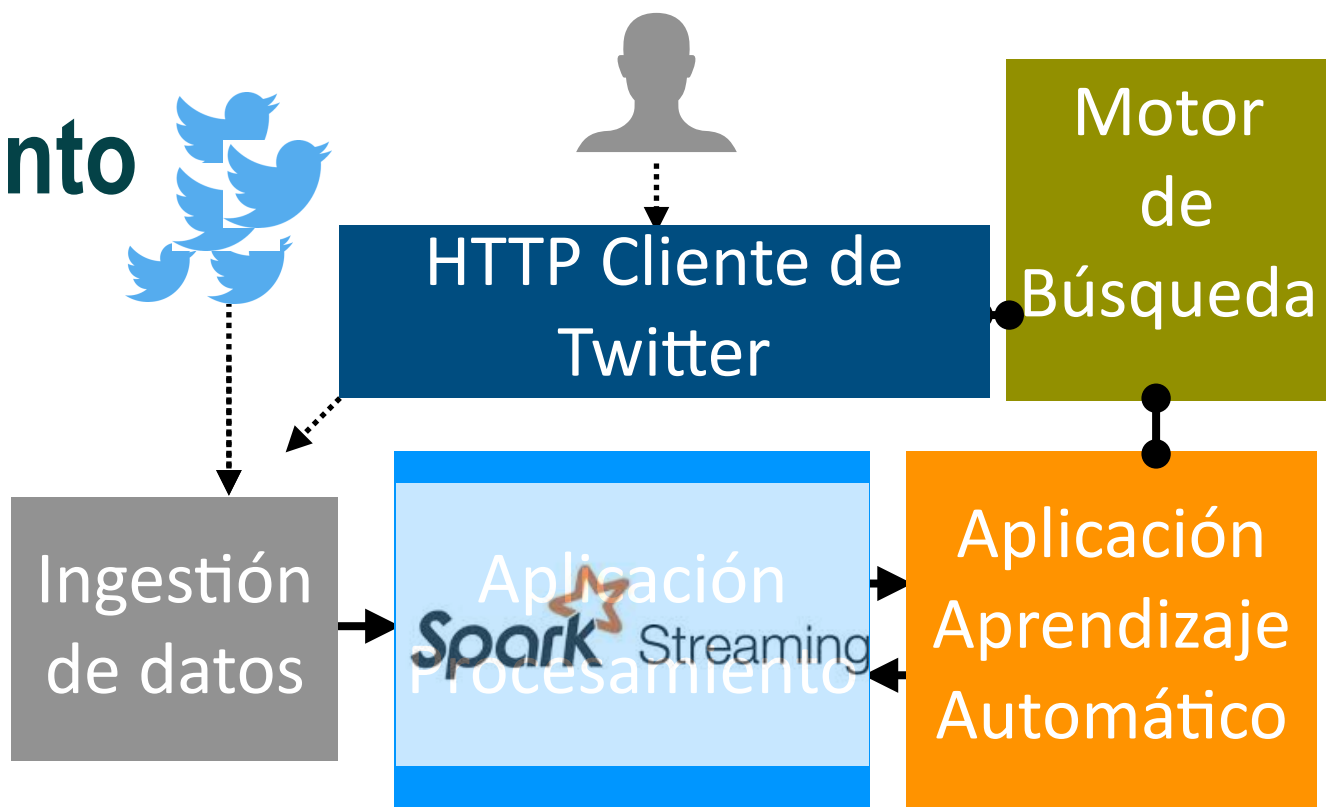
# Procesamiento streaming

## Ejemplo



# Procesamiento streaming

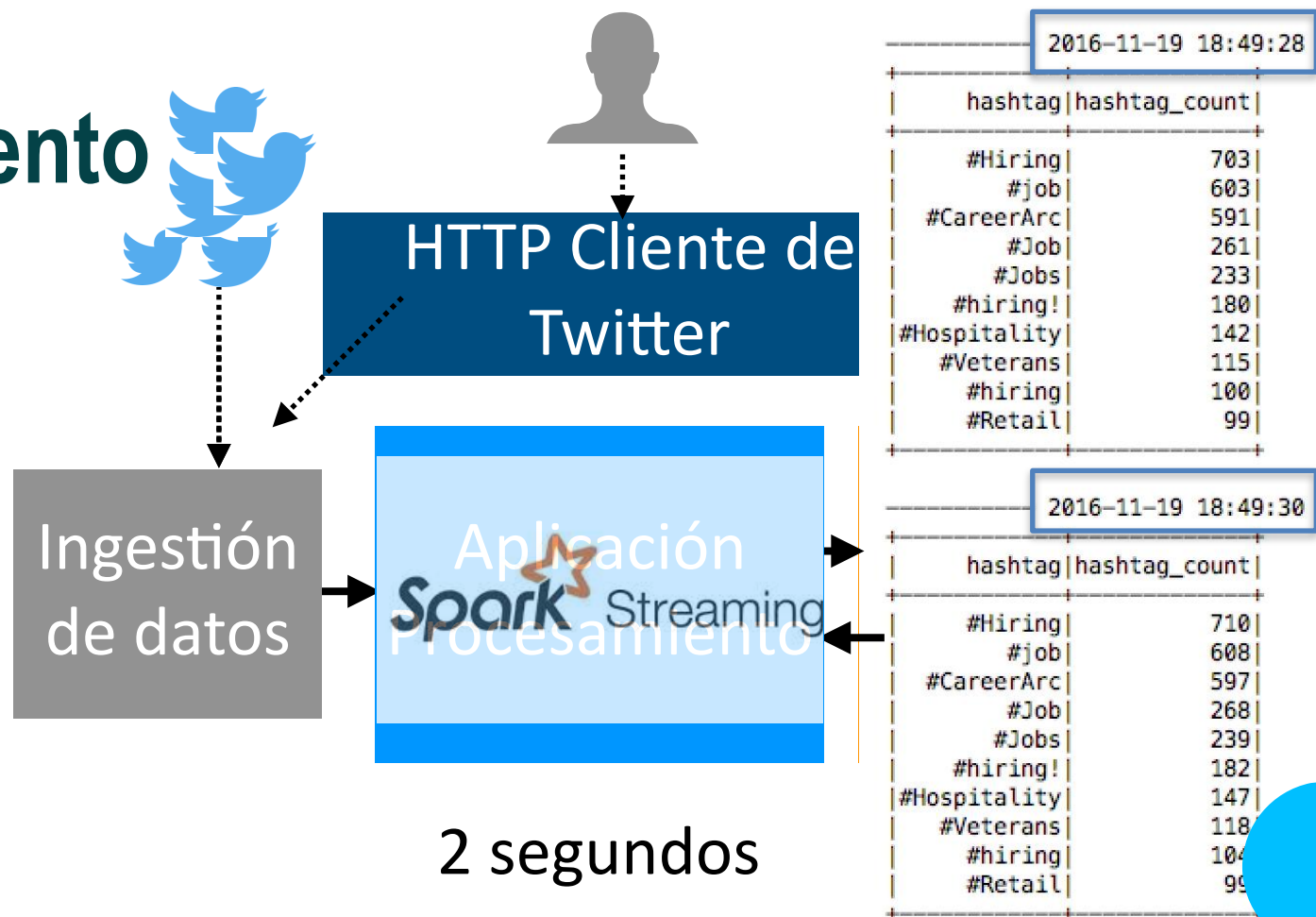
## Ejemplo



2 segundos

# Procesamiento streaming

## Ejemplo



# Procesamiento en tiempo real

## Características

- Flujos de datos
- Suficiente memoria
- Tasa de procesamiento  $>$  Tasa de entrada





# Procesamiento en tiempo real

Características  
diferente  
streaming

- Hay limitación en **tiempo** de procesamiento o de obtención **output**
- No se pueden perder datos



# Procesamiento Batch/ streaming



## Motivación

- ¿Cuanto **tiempo** necesita un cluster MapReduce con 20 nodos para leer 1 TB?
- Lectura disco duro 100-200 MB/s
- 1 TB =  $10^6$  MB
- MB de disco por nodo =  $10^6 / 20$

$$\text{Tiempo} = \frac{10^6}{20 \cdot 200} = 250 \text{ s}$$

# Procesamiento en tiempo real



## Motivación

- ¿Cuanto **tiempo** necesita un esquema MapReduce con 20 nodos para leer 1 TB **en memoria RAM**?
- Lectura memoria 10 GB/s =  $10^4$  MB/s
- 1 TB =  $10^6$  MB
- MB de disco por nodo =  $10^6 / 20$

$$\text{Tiempo} = \frac{10^6}{20 \cdot 10^4} = 5 \text{ s}$$

# Procesamiento en tiempo real

Computación  
en memoria

- MapReduce
- Consultas en tiempo real

Precio memoria RAM 1TB= 20000 \$



# Procesamiento streaming y en tiempo real

- Características
- Rendimiento

## Sistemas de Big Data

Ricardo García Ródenas  
Ricardo.Garcia@uclm.es



Sistemas de  
Big\_Data



Curso Especialización  
Inteligencia\_Artificial y  
Big\_Data