



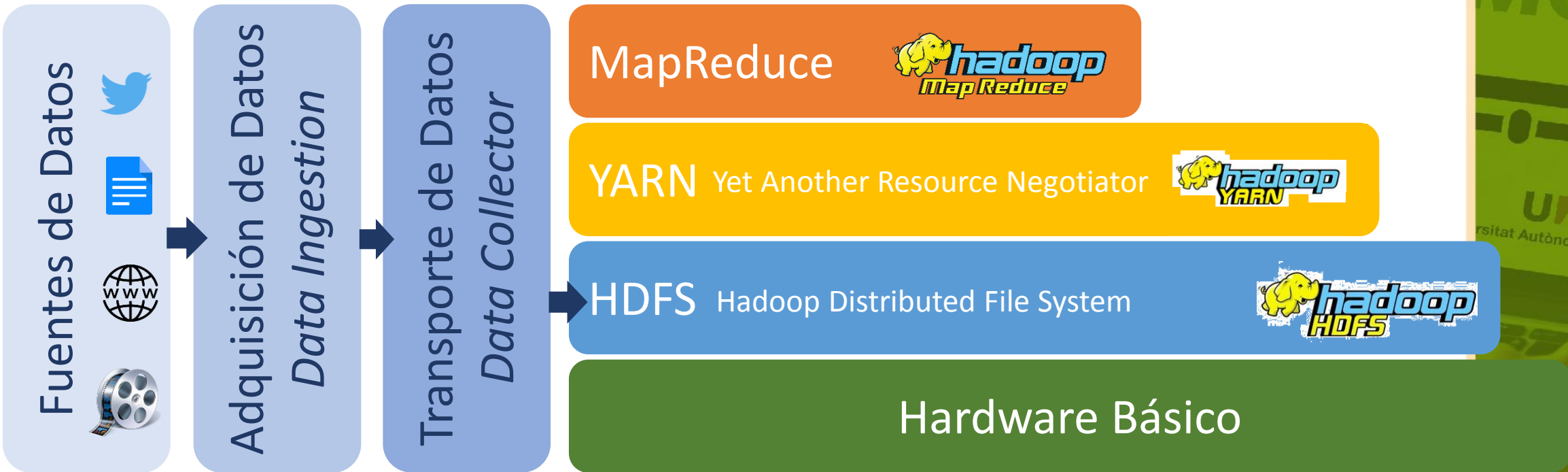
UAB

Universitat Autònoma
de Barcelona

Tomás Margalef

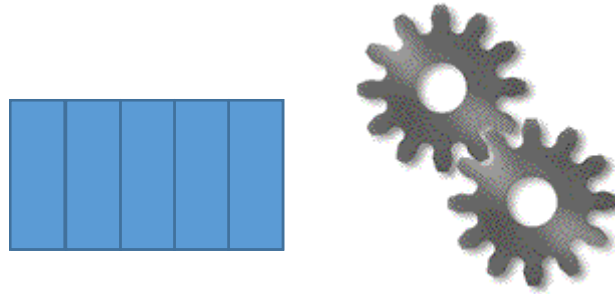
Adquisición de datos

Adquisición de datos

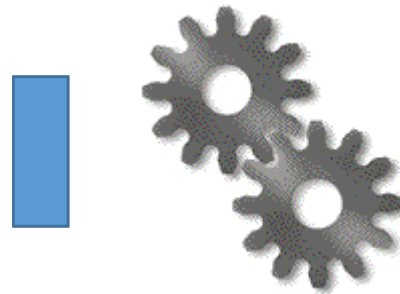


Adquisición de datos

- Procesamiento por lotes (*Batches*)

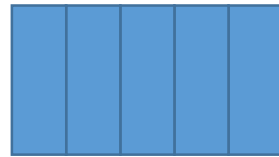


- Procesamiento en tiempo real (*Streaming*)



Adquisición de datos

- Procesamiento por lotes (*Batches*)



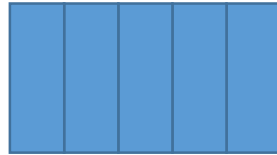
El modelo MapReduce se ajusta al procesamiento por lotes

- Procesamiento en tiempo real (*Streaming*)



Adquisición de datos

- Procesamiento por lotes (*Batches*)



Mapper

Mapper

Mapper

Mapper

Mapper

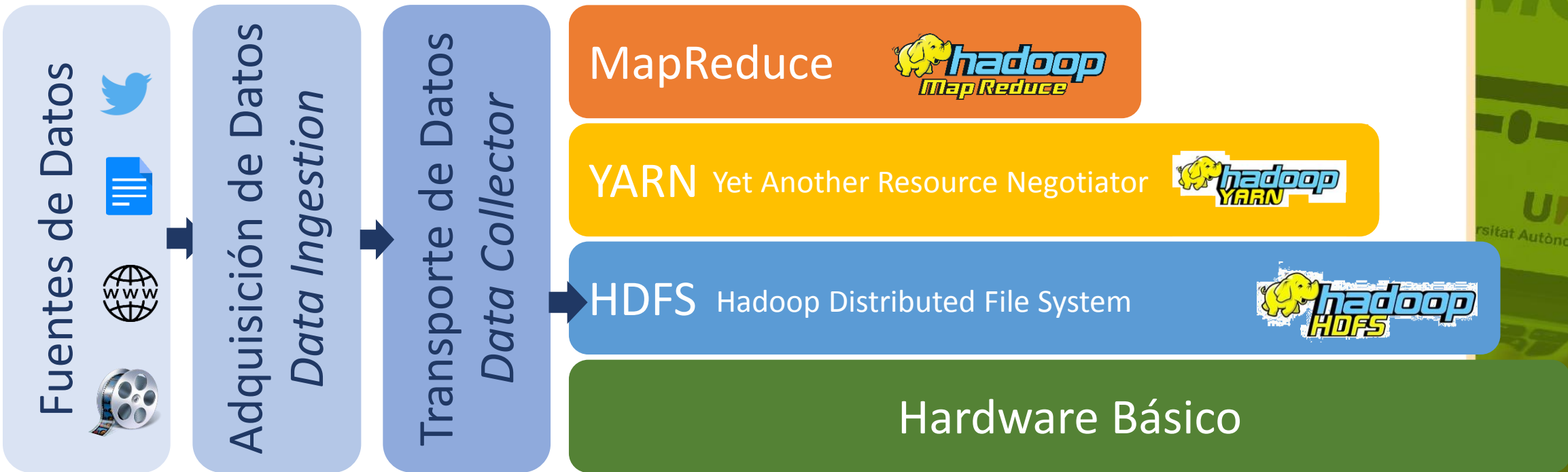


Adquisición de datos

- *Macro Batch*: Más de 15 minutos.
- *Micro Batch*: Entre 2 y 15 minutos.
- *Near Real-Time Decision Support*: Entre 2 segundos y 2 minutos.
- *Near Real Time Event Processing*: Entre 100 milisegundos y 2 segundos.
- *Real-Time*: Menos de 100 milisegundos.



Adquisición de datos



Adquisición de datos

- Recolección de información procedente de múltiples fuentes.
 - Extracción de los “datos”.
 - Validación de los datos.
- Integración de los datos en el sistema de procesamiento.
 - Colocar los datos en los destinos apropiados.



Desafíos de la Adquisición de datos

- Enorme cantidad de fuentes de datos en crecimiento.
 - Diferentes tipos de datos y formatos.
 - Generación de datos a gran velocidad.
- Fuentes de datos en constante evolución.
 - Cambios en los datos generados.
 - Cambios en la semántica de los datos.
 - Detección de los cambios en los datos.



Requisitos del sistema de Adquisición

- Gestionar nuevos tipos de datos, tecnologías y aplicaciones.
- Asegurar la corrección, consistencia y valor de los datos.
- Permitir la rápida integración de datos.
- Escalar de acuerdo con el volumen de los datos de entrada.
- Agregar datos de entrada para reducir el volumen.



Consideraciones del sistema de Adquisición

- Ancho de banda de la red (*Network Bandwidth*).
- Posibilidad de fallos en la red.
- Tecnologías y sistemas heterogéneos.
- Conversión de datos de distinto formato.





MOOC

Adquisición de datos



UAB
Universitat Autònoma de Barcelona

MOOC
Escola de
Postgrau

UAB

Universitat Autònoma
de Barcelona