

# Apache Kafka

Készítette: Szabó Bence

# Mi az Apache Kafka

- Elosztott üzenetküldő rendszer
- Probléma: A LinkedIn nem tudta a felhasználói interakciókkal kapcsolatos adatokat hatékonyan feldolgozni
- 2010: LinkedIn fejlesztette ki
- 2014: A fejlesztői kiváltak, megalapították a Confluent nevű céget
- Funkciói:
  - Nagy áteresztőképesség
  - Veszteségmentes
  - Skálázható
  - Adatfolyam feldolgozóképeség (Storm, Hadoop)
  - Támogatja, hogy egy üzenetre több feliratkozó legyen



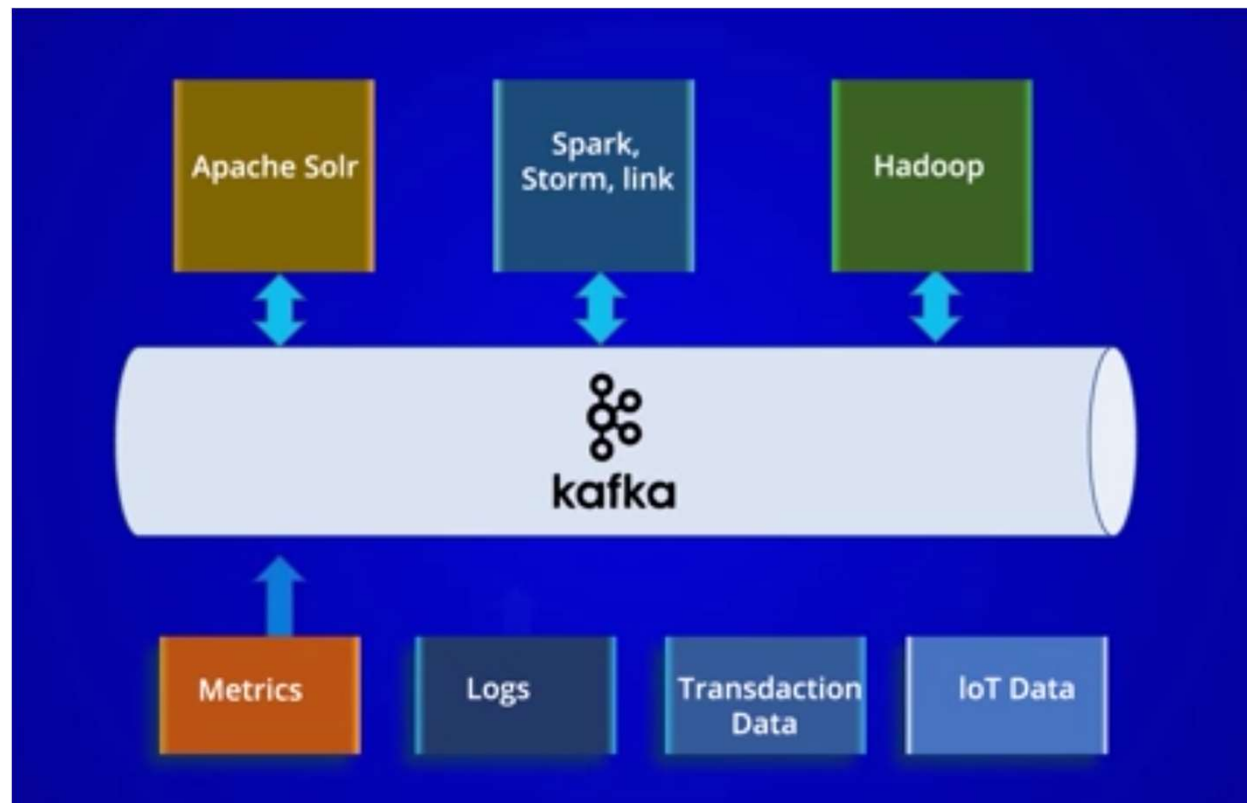
# Fogalmak

- Üzenet: adategység (Kulcs- érték párok)
- Batch: üzenetek gyűjteménye
- Fejezet (topic): Egy név, amely által vannak az üzenetek elküldve
- Partíció: Erre van egy fejezet felbontva
- **Producer: Üzenetek küldői**
- **Consumer: Üzenetek fogadói**
- Broker: Kafka szerverek
- Zookeeper: Elosztott rendszereket koordináló szolgáltatás
- Cluster: Egy zookeeper által kezelt szerverek összesége

# Mire használható

- Üzenetküldés
- Aktivitás követés
- Log és metrikák gyűjtése
- Log aggregálás
- Adatfolyam továbbítás
- Logolás

# Kafka helye a big data-ban



# Szerkezet

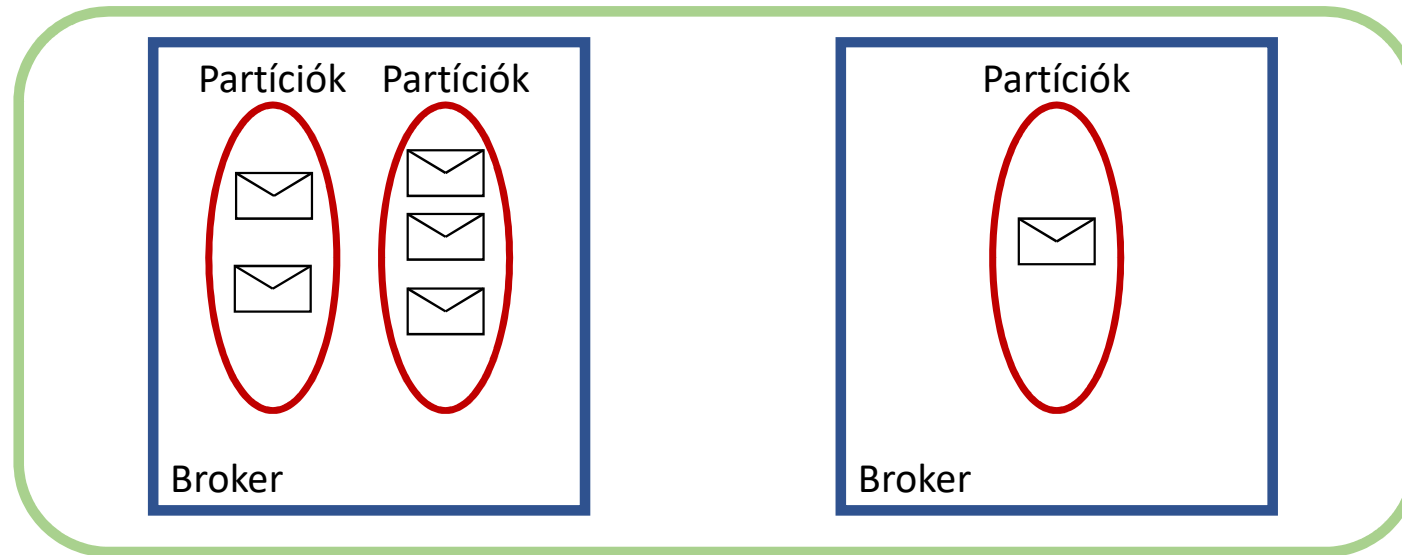


Cluster

Producer

Producer

Producer



Consumer

Consumer

Consumer

Consumer

# Üzenetküldés

- Típusai:
  - **Fire-and forget:** itt nem várja meg a sikeres kézbesítését az üzenetnek
  - **Szinkron küldés:** itt megvárja a visszaigazolást a sikeres kézbesítésről
  - **Asszinkron küldés:** itt nem várja meg a visszaigazolást, de egy callback-en keresztül megkapja a visszaigazolást
- Producer kötelező paraméterei:
  - bootstrap.server: Kafka szerver címe
  - key.serializer
  - value.serializer

# Ipari felhasználások

- LinkedIn
- Twitter/ X ? (Stormmal együtt)
- Airbnb
- Coursera
- Netflix
- Oracle
- Spotify
- Wikimedia



# Integráció

- Apache Storm
- Spark
- Flume (demo)

# Köszönöm a figyelmet

- Források:
  - LearnkartS Apache Kafka specializáció két tanfolyama (1.; 4.):  
<https://www.coursera.org/specializations/complete-apache-kafka-course>
  - Apache Kafka hivatalos oldala: <https://kafka.apache.org/powered-by>
  - Icon-icons: <https://icon-icons.com/icon/apache-zookeeper-logo/168599>
  - Confluent:  
<https://yt3.googleusercontent.com/UL39JvqfMj2FQShUibj3l0dkOXqiHBxXyLSey4VFue5Vyem0bOTaWsGMX2S2b3Q6fNLNnMxV9w=s900-c-k-c0x00ffffff-no-rj>
- Apache Kafka github repository: <https://github.com/apache/kafka>