

**SPARK STREAMING ASOSIDA REAL VAQTDA MA'LUMOTLARNI
QAYTA ISHLASH**

H.I. Eshankulov

*Texnika fanalari bo'yicha falsafa doktori (PhD), Buxoro davlat universiteti
professori, O'zbekiston Respublikasi*

E-mail: vivente_2006@mail.ru

R.B. Murodova

Buxoro davlat universiteti doktoranti, O'zbekiston Respublikasi

E-mail: x.i.eshonkulov@buxdu.uz

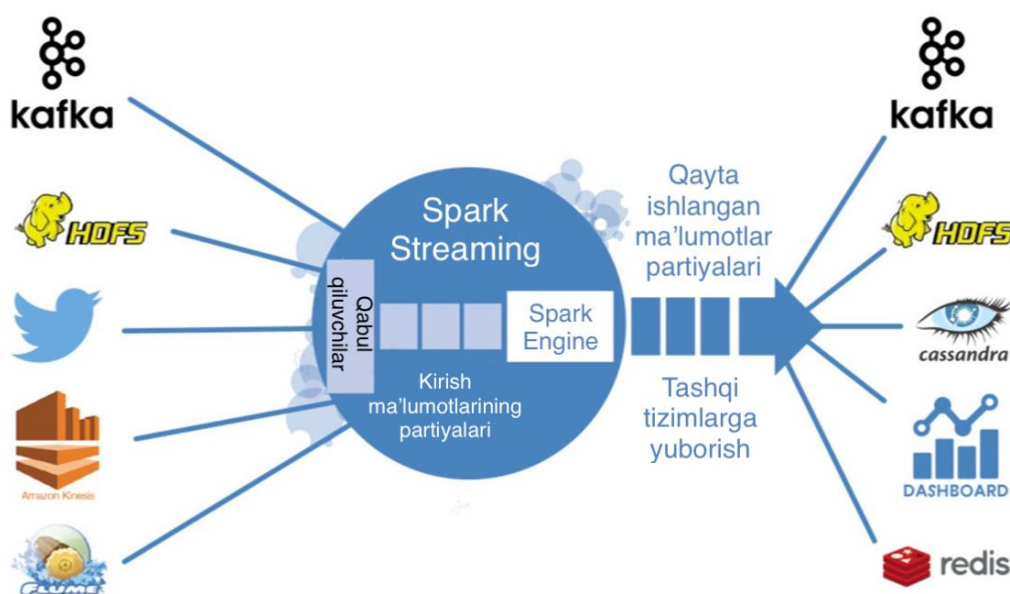
Annotatsiya. Ushbu maqolada real vaqt ma'lumotlarini qayta ishlash texnologiyalari, xususan, Spark Streamingning roli va afzalliklari ko'rib chiqiladi. Real vaqt tizimlarida ma'lumotlarni darhol qayta ishlash, tahlil qilish va qaror qabul qilishning ahamiyati o'rganilgan. Maqolada Spark Streamingning "batch" (partiya)lar shaklida ma'lumotlarni qayta ishlash usuli, turli ma'lumot manbalari bilan integratsiya imkoniyatlari va real vaqtda ma'lumotlarni olish va qayta ishlash jarayonlari taqdim etiladi. Shuningdek, Spark Streamingning moliya, tibbiyot, transport, energiya va sport sohalaridagi qo'llanilishi misollar bilan tushuntirilgan.

Kalit so'zlar: real vaqt ma'lumotlarini qayta ishlash, spark streaming, ma'lumot manbalari integratsiyasi, real vaqt tizimlari

Bugungi kunda barcha sohada ma'lumotlar oqimi juda oshib bormoqda. Bunday holda real vaqtda ma'lumotlarning intellektual tahlil qilish masalasi juda dozlarb sanaladi. Real vaqtda ma'lumotlarni qayta ishlash– ma'lumotlarning tezkor yoki kechikishsiz qayta ishlanishi jarayoni bo'lib, ushbu texnologiya orqali tahlil va qaror qabul qilish ham vaqtda amalga oshiriladi. Bu jarayon ko'pincha sensorlar, onlayn tizimlar yoki joriy voqealar oqimini kuzatish va ulardan ma'lumotlarni olish bilan bog'liq. Real vaqtda ma'lumotlarni qayta ishlashning bir nechta sohalardagi qo'llanilishi mavjud: moliya tizimlarida savdolar yoki tranzaksiyalar real vaqt rejimida

kuzatiladi va tahlil qilinadi; tibbiyotda esa sensorlar va qurilmalar orqali bemorlar holatini real vaqt rejimida kuzatish va ularga javob berish amalga oshiriladi; transport tizimlarida yo'llardagi tirbandlik real vaqt rejimida kuzatilib, avtomatik ravishda marshrut optimallashtiriladi; energiya sektorida esa energiya ishlab chiqarish va taqsimlashda real vaqtda ma'lumotlarni qayta ishlash energiya sarfini optimallashtirish va taqsimotning barqarorligini ta'minlash uchun ishlatiladi. Shuningdek, sport va o'yin sanoatida real vaqtda ma'lumotlarni qayta ishlash o'yin natijalarini tahlil qilish, statistika va yutuqlarni kuzatishda qo'llaniladi. Spark Streamingning real vaqt tizimlaridagi roli juda muhim, chunki bugungi rivojlanib borayotgan axborot olamida uning ahamiyati ortib bormoqda.

Spark Streaming – bu real vaqtda ma'lumotlarni qayta ishlash texnologiyasidir, unda kiruvchi ma'lumotlar "batch" (partiya)lar shaklida qayta ishlanadi. Bu esa ma'lumotlar oqimini kichik bo'laklarga bo'lib, ularni tezkor tahlil qilish va qayta ishlashni ta'minlaydi. Ma'lumotlar partiyalari juda tezkorlik bilan qayta ishlanadi. Spark Streaming, shuningdek, turli ma'lumot manbalari bilan oson integratsiya qilish imkoniyatiga ega. U Kafka, Flume, Kinesis, HDFS, JDBC va boshqa tizimlardan real vaqtda ma'lumotlarni olish va qayta ishlashni amalga oshiradi (1-rasm).



1-rasm. Ma'lumotlar oqimini qabul qilish

Ishlov berish imkoniyatlari: Spark Streaming, real vaqtda ma'lumotlar ustida turli statistik tahlil, agregatsiya, o'zgarishlarni aniqlash va boshqa murakkab

operatsiyalarni bajarish imkoniyatlarini taqdim etadi. Masalan, foydalanuvchilarning real vaqtdagi faoliyatlarini kuzatish yoki ma'lumotlarga asoslangan prognozlar qilish mumkin. Yuqori darajadagi API: Spark Streaming yuqori darajadagi API taqdim etadi, bu esa dasturchilarga qulay tarzda real vaqtda ma'lumotlarni qayta ishlash imkoniyatini beradi. API orqali operatsiyalarni oson bajarish, masalan, ma'lumotlarni filtrlash, agregatsiya qilish, bog'lanish va boshqalar. Qo'shimcha vositalar: Spark Streaming bilan birga Structured Streaming ham mavjud bo'lib, u ma'lumot oqimlarini (streaming) Sarka ekotizimidagi odatiy ma'lumotlar to'plami (batch) kabi ishlov berishga imkon beradi. Bu, statik va dinamik ma'lumotlarni bir xil tarzda ishlash imkoniyatini beradi. Apache Spark Streaming real vaqtda ishlov berishning yuqori samarali va kengaytiriladigan tizimi sifatida turli sohalarda, masalan, finans, e-commerce, monitoring tizimlari va boshqa ko'plab sohalarda qo'llaniladi.

ADABIYOTLAR RO'YXATI

1. Zhang, X., & Li, X. (2019). "A survey of real-time data processing systems and applications". IEEE Access, 7, 17482-17495. doi: 10.1109/ACCESS. 2019. 2898285.
2. Karim, M. R., & Ahmed, F. (2018). "Real-time big data stream processing with Apache Spark: A survey". International Journal of Computer Applications, 180(1), 18-25. doi: 10.5120/ijca2018916982.