

ISSN 2010-7242

O'ZBEKISTON RESPUBLIKASI FANLAR AKADEMIYASI

O'ZBEKISTON RESPUBLIKASI AXBOROT TEXNOLOGIYALARI  
VA KOMMUNIKATSIYALARINI RIVOJLANTIRISH VAZIRLIGI

**INFORMATIKA VA ENERGETIKA  
MUAMMOLARI**  
O'zbekiston jurnali

Узбекский журнал  
**ПРОБЛЕМЫ  
ИНФОРМАТИКИ И ЭНЕРГЕТИКИ**

Uzbek Journal

**OF THE PROBLEMS OF  
INFORMATICS AND ENERGETICS**

**4**

**2021**

FAN VA TEXNOLOGIYA



ЎЗБЕКИСТОН RESPUBLIKASI FANLAR АКАДЕМИЯСИ  
АКАДЕМИЯ НАУК РЕСПУБЛИКИ УЗБЕКИСТАН  
ЎЗБЕКИСТОН RESPUBLIKASI АХВОРОТ ТЕХНОЛОГИЯЛАРИ  
VA КОММУНИКАТСИЯЛАРИНИ РИВОЖЛАНТИРИШ  
VAZIRLIGI  
МИНИСТЕРСТВО ПО РАЗВИТИЮ ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ И КОММУНИКАЦИЙ РЕСПУБЛИКИ  
УЗБЕКИСТАН

ИНФОРМАТИКА  
VA ENERGETIKA  
MUAMMOLARI

4·2021

---

ПРОБЛЕМЫ  
ИНФОРМАТИКИ  
И ЭНЕРГЕТИКИ

*Журнал под таким названием издается с января 1992 г.  
по 6 номеров в год*

Fan va texnologiyalar nashriyot-matbaa uyi

ТАШКЕНТ– 2021

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Ш.Х. ФАЗЫЛОВ (*главный редактор*)  
Х.М. МУРАТОВ (*зам. главного редактора*)  
Б.М. АЗИМОВ (*ответственный секретарь*)  
Т.С. КАМАЛОВ (*ответственный секретарь*)  
С.Ф.АМИРОВ  
Р.А. ЗАХИДОВ  
Х.З. ИГАМБЕРДИЕВ  
М.А. ИСМАИЛОВ  
О.Х.ИШНАЗАРОВ  
М.М. КАМИЛОВ  
М.М.КАРИМОВ  
Х.А.МУХИТДИНОВ  
М.М.МУСАЕВ  
О.М. НАБИЕВ  
Т.Х. НАСЫРОВ  
Н. РАВШАНОВ  
А.А.РАДЖАБОВ  
Ш.Х.РАХИМОВ  
М.А.РАХМАТУЛЛАЕВ  
О.З.ТОИРОВ  
Р.Х.ХАМДАМОВ

*Адрес редакции:*

100173, г. Ташкент, ул. Фозилтепа, 226  
Телефоны 71-231-92-51, 71-231-92-46  
e-mail: [informatika-energetika@mail.ru](mailto:informatika-energetika@mail.ru)

© «Fan va texnologiyalar nashriyot-matbaa uyi»

ТАШКЕНТ – 2021

**B.B. ELOV, H.M.PRIMOVA**

## **BIG DATA ARXITEKTURASI**

Big Data – 2000-yil oxirida paydo bo‘lgan va an’anaviy ma’lumotlar bazalarini boshqarish tizimlari va Business Intelligence sinfining yechimlariga muqobil bo‘lgan, katta hajmdagi va sezilarli xilma-xillikdagi strukturlangan va strukturlanmagan ma’lumotlarni qayta ishlashga mo‘ljallangan dasturlash texnologiyasi hisonlanadi. Hozirda Big Data texnologiyalari hayotning turli sohalariga tobora chuqurroq kirib bormoqda, savdo, reklama, dam olish va hatto shaxsiy hayot tuzilishi haqidagi odatiy g‘oyalarni o‘zgartirmoqda. Mamalakarimizda elektron hukumat doirasida turli axbrot tizimlarining joriy etilishi natijasida ma’lumotlar bazasida katta hajmdagi ma’lumotlarning yuzaga kelishi kuzatilmoqda. Ushbu ma’lumotlarni qayta ishlash uchun Big Data texnologiyalaridan foydalanishga to‘g‘ri keladi. Katta hajmdagi ma’lumotlarni tahlil qilishning texnologik imkoniyatlarining paydo bo‘lishi bilan bog‘liq bo‘lgan ijtimoiy-iqtisodiy hodisa sifatida tilga olinadi. Ushbu maqolada Big Data arxitekturasini tuzilmasining tavsifi berilgan va katta hajmdagi ma’lumotlardan foydalanish mumkin bo‘lgan holatlarini xaritalash muammosi yechimi berilgan.

**Kalit so‘zlar:** katta hajmli ma’lumotlar, arxitektura, texnologiya, ma’lumotlar bazasi, monitoring, Big Data, Hadoop, MapReduce, data lakes, SQL.

**Б.Б.Элов, Х.М.Примова**

## **Архитектура Big Data**

Показано, что Big Data – технология программирования, появившаяся в конце 2000-х годов и являющаяся альтернативой традиционным системам управления базами данных и решениям класса Business Intelligence для обработки больших объемов и широкого спектра структурированных и неструктурированных данных. Установлено, что в настоящее время технологии Big Data все глубже проникают в различные сферы жизни, меняя привычные представления о структуре торговли, рекламы, досуга и даже личной жизни. Обосновано, что в результате внедрения различных информационных систем в рамках электронного правительства в нашей стране в базе данных накопилось большое количество данных. Отмечено, что для обработки этой информации необходимо использовать технологию Big Data, представляющую собой социально-экономическое явление, связанное с появлением технологических возможностей для анализа больших объемов данных. Рассмотрена структура архитектуры Big Data и предложено решение проблемы отображения экземпляров при наличии больших объемов данных.

**Ключевые слова:** большие данные, архитектура, технология, база данных, мониторинг, Big Data, Hadoop, MapReduce, озера данных, SQL.

**B.B. Elov, H.M.Primova**

## **Big Data architecture**

Big Data - is a programming technology that appeared in the late 2000s and is an alternative to traditional database management systems and Business Intelligence class solutions for processing large amounts and a wide variety of structured and unstructured data. Nowadays, Big Data technologies are penetrating more and more deeply into



various areas of life, changing the usual ideas about the structure of trade, advertising, leisure and even personal life. As a result of the introduction of various information systems in the framework of e-government in our country, there is a large amount of data in the database. To process this info, it is needed to use Big Data technology. It is referred to as a socio-economic phenomenon associated with the emergence of technological capabilities for the analysis of large amounts of data. This article describes the structure of the Big Data architecture and provides a solution to the problem of mapping instances where large amounts of data are available.

**Keywords:** Large data, architecture, technology, database, monitoring, Big Data, Hadoop, MapReduce, data lakes, SQL.

**Kirish.** Bugungi kunda respublikamizda elektron hukumat doirasida joriy qilinayotgan axborot tizimlarida juda katta hajmdagi ma'lumotlar (Big Data) shakllanyapti. Ma'lumotlar bazasidagi katta hajmdagi murakkab ma'lumotlarni tahlil qilish va qayta ishlash juda muhim masala hisoblanadi. Turli tashkilotlarda foydalnuvchilarning imkoniyatlari va vositalaridan foydalangan holda katta hajmdagi ma'lumotlar bazasini tashkil etishadi. Tashkilotlar uchun bu yuzlab gigabayt ma'lumotlar yoki yuzlab terabayt ma'lumotlar bo'lishi mumkin. Big Data ma'lumotlar bazasida yillar davomida jamlanib, o'zgarib boradi. Bundan tashqari, ma'lumotlar bilan ishlash uchun yangi imkoniyatlar yaratiladi <sup>1</sup>. Ba'zi ma'lumotlar katta bloklarda sekinlik bilan ma'lumotlar bazasiga qayd etiladi. Bunday ma'lumotlar o'nlab yillar davomida jurnal ma'lumotlari hisoblanadi. Ushbu maqolaa Big Data arxitekturasi tarkibiy qismlarini ko'rib chiqamiz.

Big Datani tahlil qilish uchun quyidagi amallar bajarilishi lozim:

- Aktiv bo'lmagan Big Data manbalarini paketli qayta ishlash;
- Real vaqt rejimida Big Datani qayta ishlash;
- Big Datani interaktiv o'rganish;
- Big Datani tahlil qilish va mashinali o'rganish.

Big Datani tahlil qilish maqsadida arxitekturada quyidagi biznes-jarayonlar o'z aksini topishi lozim:

- An'anaviy ma'lumotlar bazasidagi juda katta hajmdagi ma'lumotlarni qayta ishlash va saqlash;
- Big Datadagi strukturlanmagan ma'lumotlarni tahlil qilish va turli hisobotlarni shakllantirish maqsadida o'zgartirish;
- Real vaqt rejimida yoki ma'lumotlarni qabul qilishda biroz uzilishlarga ega bo'lgan holda o'zaro bog'liq bo'lmagan ma'lumotlar oqimlarini qayta ishlash.

#### **Big Data arxitekturasi komponentlari**

Big Datani arxitekturasi komponentlar quyidagi 1-rasmda keltirilgan bo'lib quyidagi komponentlardan iborat: *ma'lumotlar manbalari, ma'lumotlarni saqlash, real vaqtda ma'lumotlarni qabul qilish, qismni qayta ishlash, mashinali o'rganish, oqilarni qayta ishlash, ma'lumotlar bazasi tahlili, tahlil va hisobot.*

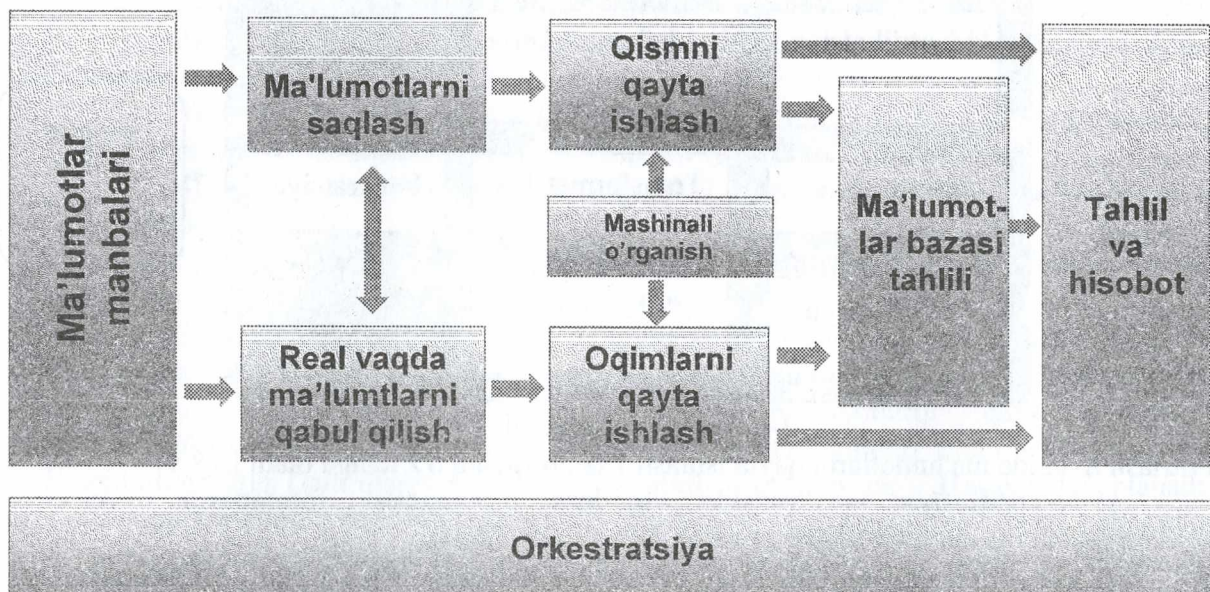
**Ma'lumotlar manbalari.** Barcha Big Data yechimlari bir yoki bir nechta ma'lumotlar manbalarini o'z ichiga oladi:

- Dasturiy ta'minot ma'lumotlarini saqlash ombori (masalan, relyatsion ma'lumotlar bazasi);
- Dasturiy ta'minotlar tomonidan shakllantiriladigan statik fayllar (masalan, web serverlardagi jurnal fayllari);
- Real vaqt rejimida uzatiladigan ma'lumotlar manbalari (masalan, online konferensiyalar).

---

<sup>1</sup> <https://docs.microsoft.com/ru-ru/azure/architecture/data-guide/big-data/>





1-rasm. Big Data arxitekturasi komponentlari

Big Data ning ixtiyoriy yechimi muayyan masalani hal qilishga qaratilgan va o'ziga xos tarzda unikal hisoblanadi. Ma'lumotlar turli xil ichki va tashqi manbalar orqali **Big Data**ning saqlash tizimiga qabul qilinishi mumkin. Manbalar *turi, hajmi va tezligiga* ko'ra farqlanadi. Big Data texnologik steki yordamida kiruvchi ma'lumotlar “*ma'lumotlar ko'llari*” (**data lakes**<sup>2</sup>) deb ataladigan katta “*ma'lumotlar to'plami*”ga (**big data sets**<sup>3</sup>) birlashtiriladi. Ma'lumot manbalarining ikkita asosiy toifasi mavjud [1,2]:

1) *An'anaviy* — standartlar va protokollardan iborat ma'lumotlar. Ushbu turdagi barcha manbalarda foydali ma'lumotlar juda yuqori zichligi bilan ajralib turadi. Masalan:

- HTTP veb-servislar
- Relyatsion ma'lumotlar bazalari
- Fayllar
- XML

2) *Zamonaviy* — so'nggi 10 yil ichida faol foydalanishda bo'lgan qayta ishlanmagan “xom-ashyo” tipidagi ma'lumotlar turlari bo'lib, “an'anaviy” turdan farqli o'laroq, ma'lumotlarning juda *xilma-xilligi, hajmi va tezligi* bilan ajralib turadi. Shuningdek, har qanday qarorlar qabul qilinadigan ma'lumotlarning ulushi “*zamonaviy*” turdagi ma'lumotlarda ancha past va odatda 10% dan oshmaydi. Masalan:

- Turli xil qurilmalardan **uzluksiz tarzda** qabul qilinadigan ma'lumotlar
- **Call-center yozuvlari**, shu jumladan, ovozni aniqlash orqali olingan yozuvlar
- **GPS qurilmalari** orqali qabul qilinadigan ma'lumotlar
- **Sotsial tarmoqlardagi** xabarlar

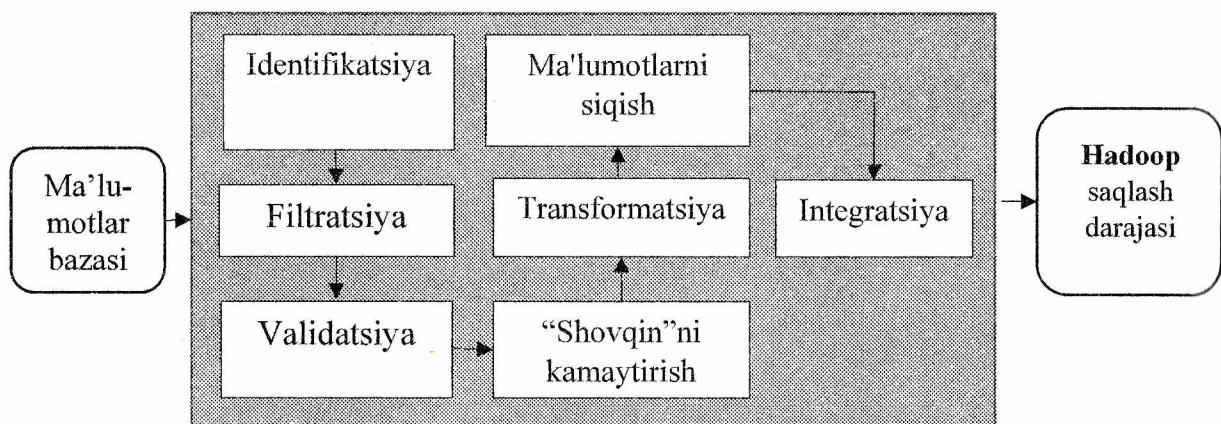
#### Ma'lumotlarni filtrlash tizimi

Nitin Sawant, Himanshu Shanlarning “Big Data Application Architecture Q&A” nomli maqolasida *Ingestion Layer* deb nomlangan filtrlash tizimi turli xil manbalardan *Big Data* tizimiga kiradigan juda katta miqdordagi ma'lumotlar uchun eshik bo'lib xizmat qiladi. Filtrlash tizimining asosiy maqsadi qabul qilinadigan ma'lumotlarni keraksiz narsalardan tozalash va shu bilan *Big Data* platformasida ma'lumotlarni qayta ishlashga tayyorlashdir [3–7].

<sup>2</sup> **Data lake** (*ma'lumotlar “ko'li”*) — strukturlangan va strukturlanmagan katta hajmdagi ma'lumotlar bazasi

<sup>3</sup> **Big Data sets**- Ma'lumotlar bazasining bir yoki bir nechta katta hajmdagi ma'lumotlarni o'zida saqlovchi jadvalar to'plami



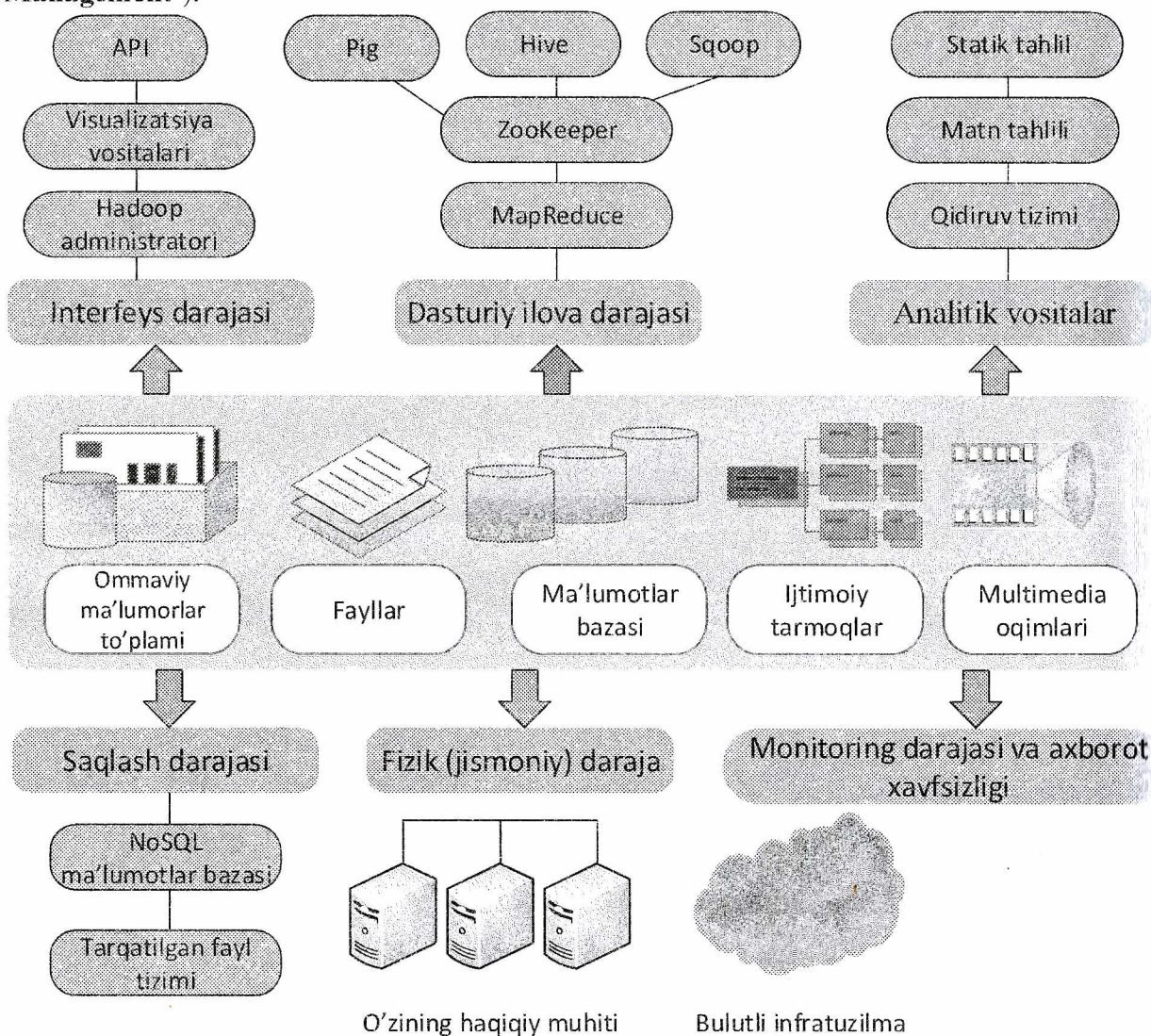


2-rasm. Filtrlash tizimi

Filtrlash tizimida ma'lumotlarni qayta ishlash 7 ta bosqichni o'z ichiga oladi [6,8,9]:

1) *Identifikatsiya* — avvaldan ma'lum ma'lumotlar formatlarini yoki strukturlangan ma'lumotlar uchun standart formatlarni aniqlash.

2) *Filtratsiya* — tashkilotga tegishli bo'lgan kiruvchi ma'lumotlar. Muvofiqlik ma'lumotlarni boshqarish bo'yicha asosiy tizim asosida baholash (**Master Data Management**<sup>4</sup>).



3-rasm. Big data arxitekturasi

<sup>4</sup> **Master Data Management** — kompaniyaning asosiy ma'lumotlarini doimiy ravishda aniqlaydi va ma'lumotlarni boshqaradi



1) *Validatsiya* — ma'lumotlarni tekshirish va tahlil qilish.

2) "*Shovqin*"ni kamaytirish — ma'lumotlarni tozalash va turli xil shovqinlarni minimallashtirish.

3) *Transformatsiya* — ma'lumotlar to'plamlarini tarkibiy qismlarga ajratish, ma'lumotlarni tekislash, denormalizatsiya qilish va yig'ishni ta'minlash.

4) *Ma'lumotlarni siqish* — avvalgi tahlil natijalarini yo'qotmasdan ma'lumotlar hajmini kamaytirish.

5) *Integratsiya* — filtrlangan ma'lumotni **Hadoop**<sup>5</sup> saqlash qatlamiga tushirishni ta'minlash.

**Saqlash darajasi.** Big Data texnologiyalari ma'lumotlarni saqlash darajasida taqsimlangan saqlash tizimidan foydalaniladi. Taqsimlangan saqlash tizimi nafaqat xatolarga chidamlilikni ta'minlaydi balki ma'lumotlarni qayta ishlashni *parallellashtirishga* imkon beradi, bu esa Big Data hajmlari bilan ishlashda juda muhimdir. Big Data dunyosidagi asosiy saqlash texnologiyasi — **Hadoop** hisoblanadi. Hadoop yordamida Big Data dasturlari to'g'ridan-to'g'ri jismoniy serverlarga kirmasdan saqlash tugunlari bilan o'zaro ta'sir qiladi [3]. Hadoop ikkita asosiy komponentlarga ega:

1) Taqsimlangan va yuqori pog'onadagi mashtablashgan **HDFS (Hadoop Distributed File System)** fayl tizimi.

2) So'rovlarni paketli rejimda qayta ishlashni ta'minlaydigan **MapReduce**<sup>6</sup> qism tizimi [10–12].

Hadoop — kutubxonalar va fremvorklariga ega bo'lgan, yuzlab yoki minglab tugunli klasterlarda tarqatiladigan dasturlarni ishlab chiqishi va bajariladigan bepul tarqatadigan xizmatlar to'plamidir. Hadoop — *Apache Software Foundation*ning eng yuqori darajadagi loyihasi bo'lib, u *Big Data*ing asosiy saqlash texnologiyasi hisoblandi. Hadoop 4 ta asosiy moduldan iborat:

- *Hadoop Common*;
- *HDFS (Hadoop Distributed File System)*;
- *YARN*;
- *Hadoop MapReduce*.

*HDFS* bir martalik yozish va ko'p martalik o'qish (*write-once-read-many*) tamoyili asosida qurilgan va blokli tuzilishga ega bo'lib, uning har bir blokida faylni yoki faylning qismini saqlash mumkin. Shuni ta'kidlash kerakki, *HDFS* fayllariga biz to'g'ridan-to'g'ri an'anaviy fayl tizimlarida bo'lgani kabi murojaat qilish mumkin emas. *HDFS*da samarali ishlash uchun juda murakkab maxsus dasturlarni ishlab chiqish lozim.

*Big Data* tizimlarida muvayyan bir soha bo'yicha strukturlangan ma'lumotlarni saqlash uchun (masalan, hujjatlar uchun) **NoSQL**<sup>7</sup> (*Not only SQL*) MBBTdan foydalaniladi. NoSQL MBBTning to'rtta asosiy toifasi mavjud [13–16]:

*MBBTning oddiy turi* — Ma'lumotlar bazasida ma'lumotlarga kalit yordamida (*key-value*) murojaat qilish

1) *Ustun (maydon)larga asoslangan ma'lumotlar bazalari* — tuzilmalanmagan ma'lumotlarni indeksatsiya qilish vositasini beradi va asosan web-dasturlarda xizmat qiladi. Bunday MBBTga **H base**<sup>8</sup>ni misol qilib keltirish mumkin.

2) *Hujjatlarga yo'naltirilgan ma'lumotlar bazasi* — ierarxik tuzilmaga ega ma'lumotlarni saqlashga imkon beradi. Masalan, **Mongo DB**<sup>9</sup>da JSON formatidagi ob'yektlarni saqlash mumkin.

<sup>5</sup>**Hadoop** — arzon kompyuterlarda katta hajmdagi ma'lumotlarni (petabayt miqyosida) tarqatishni ta'minlaydigan ochiq kodli dasturiy ta'minot bazasi

<sup>6</sup>**MapReduce** — katta hajmdagi ma'lumotlarni ommaviy qayta ishlash texnologiyasi

<sup>7</sup>**NoSQL** — moslashuvchan ma'lumotlar modeli bilan ma'lumotlarning miqyosli saqlashini (ma'lumotlar bazasini) amalga oshiradi.

<sup>8</sup>**H Base** — NoSQL turiga mansub ochiq kodli MBBT.

<sup>9</sup>**Mongo DB** — hujjarlarga asoslangan asoslangan MBBT.



3) *Graf shaklidagi ob'ektlarni saqlashga xizmat qiladigan ma'lumotlar bazalarida* ma'lumotlar orasidagi aloqalarni saqlash mumkin. Ushbu NoSQL MBBT klassi *tarmoqli modellashtirishda*, shuningdek tavsiya servislarini yaratishda foydalanilishi mumkin.

**Fizik (jismoniy) daraja.** Jismoniy daraja – saqlash darajasi uchun infratuzilma bo'lib, *Big Data* infratuzilmasining gorizontallash imkonini taqdim etadi. Bu bir tomondan yuqori pog'onadagi ishonchlilikni ta'minlaydi, boshqa tomondan esa an'anaviy vertikal integral yechimlarga nisbatan pastroq xarajatni talab etadi. Gorizontallashning yuqori darajasi tugunlar resurslarini taqsimlamaydigan va ular bilan ziddiyatga olib kelmaydigan, hech narsaga yaramaydigan arxitekturadan foydalanishni ta'minlaydi. *Big Data* tizimlarida ushbu natijaga ma'lumotlar va ularni boshqarish funksiyalari bir tugunda joylashganligi sababli erishiladi.

**Dasturiy (ilova) daraja.** Dasturiy daraja, shuningdek **platforma** darajasi sifatida ham talqin qilinib, ma'lumotlardan foydalanishni tartibga soladi va so'rovlarni boshqaradi. Ushbu darajadagi asosiy texnologiya *MapReduce* hisoblanadi. Muayyan dasturdan qat'iy nazar, odatdagi *MapReduce* yechimi quyidagicha ishlaydi:

- ma'lumotlarni qayta ishlash maqsadida mijoz dasturidan amalga oshirilgan so'rovni bajarish uchun *Map* funksiyasi ma'lumotlarni qismlarga ajratadi va muvozanatlash va xatolarga bardoshlilik talablarini hisobga olgan holda ularni tizim tugunlariga tarqatadi.

- Tugun nuqtalaridagi masalalar yakuniga etganda *Reduce* funksiyasi har bir bo'lakdan ishlov berish natijasini birlashtiradi va mijoz dasturiga qaytaradi.

Dasturiy darajada quyidagi vositalarni qo'llash mumkin:

- *Hive* – bu Hadoop doirasida amalga oshirilgan ma'lumotlar ombori hisoblanib, *sql*ga o'xshash interfeysga ega va katta hajmdagi ma'lumotlarni birlashtirish imkoniyati taqdim etadi.

- *Pig* – MapReduce ishlarini yanada samarali loyihalashtirishda yordam beradigan va katta hajmdagi to'plamlarini qayta ishlashga mo'ljallangan skript dasturlash tili.

- *HBase* – ustunlar asosida amalga oshiriladigan NoSQL ma'lumotlar bazasi. Analoglardan farqli o'laroq, *HBase Hadoop*da ishlashga moslashgan.

- *Sqoop* – buyruqlar satriga asoslangan vosita bo'lib, skrukturlanmagan katta hajmdagi ma'lumotlarni skrukturlangan (masalan, ma'lumotlar bazasini relyatsion boshqarish tizimiga) o'tkazish uchun mo'ljallangan.

- *ZooKeeper* – turli xildagi *Hadoop* nusxalari uchun koordinator hisoblanadi. *ZooKeeper* tizim tugun nuqtalarini sinxronizatsiyalash va nosozliklardan himoya qilish imkoniyatini taqdim etadi.

**Monitoring darajasi va axborot xavfsizligi.** *Big Data* komponentlarining doimiy ishlashini ta'minlash maqsadida har bir komponentning avtonomligi tamoyili asosida shakllantirilgan monitoring tizimidan foydalanish lozim. Ushbu vazifani bugungi kunda ommabop hisoblangan *Ganglia* va *Nagios* ochiq manbali monitoring tizimlari amalga oshiradi [17–20].

*Big Data* taqsimlangan tizimida axborot xavfsizligini ta'minlash an'anaviy tizimlarga nisbatan ancha murakkab vazifa hisoblanadi. Asosiy qiyinchilik tizim xavfsizligini yaxlit tarzda ta'minlashda hisoblanadi. Chunki ko'plab tugunlarda jurnal ( $\log^{10}$ )lar hosil bo'lganda va yakuniy natijani shakllantirish foydalanuvchi tomonidan tekshirilmasdan amalga oshiriladi. *Big Data* tizimida axborot xavfsizligini minimal ta'minlash choralarini sifatida quyidagilar tavsiya etiladi:

- 1) Tizim tugun nuttalariga kirishda autentifikatsiya
- 2) Tugunlar o'rtasida xavfsiz aloqani o'rnatish
- 3) Ma'lumot saqlagichlarida fayllarni shifrlash

---

<sup>10</sup>**Log (jurnallar)** — foydalanuvchi va dasturning ba'zi harakatlarini qayd qilish jurnali

4) Ishonchli sertifikatlar xizmatidan foydalanish

5) Arxitekturaning barcha darajalarida taqdimlangan *log*lar tizimidan foydalanish

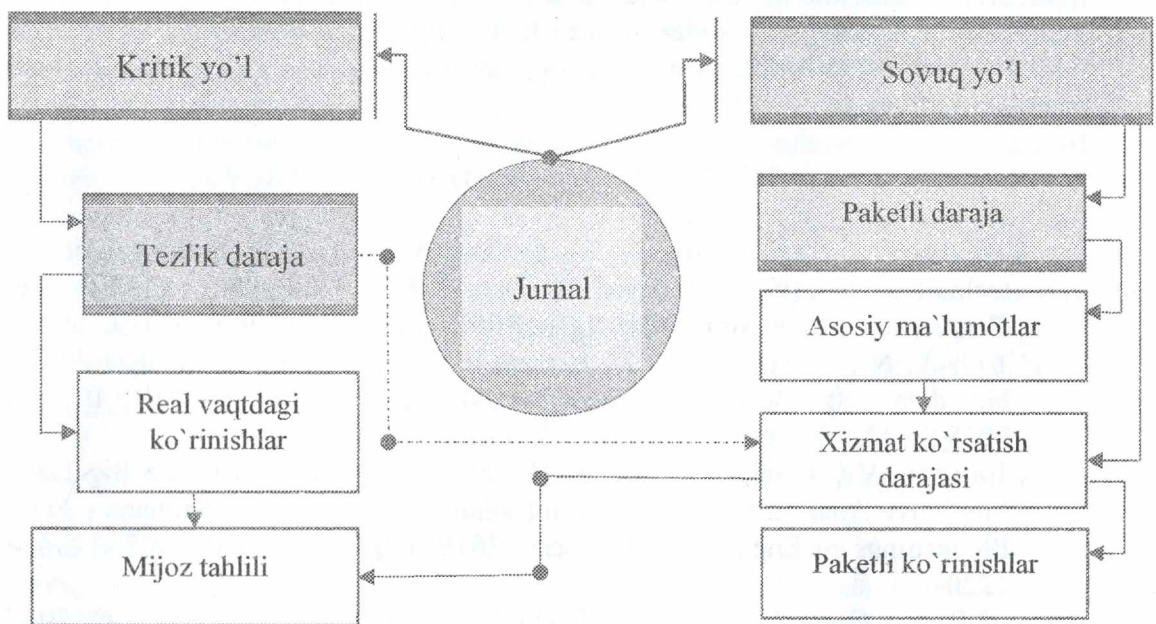
**Lambda arxitektura.** Big Datada mijozlarning so'rovlarini bajarishda juda ko'p vaqt kerak bo'ladi. Ushbu so'rovlarni real vaqt rejimida bajarib bo'lmaydi. So'rovlarni qayta ishlashda ma'lumotlar manbalarini parallel qayta ishlashga mo'ljallangan MapReduce algoritmlaridan foydalaniladi. Hosil bo'lgan natijalar alohida tarzda saqlanadi va so'rovlarni bajarish uchun ishlatiladi.

Ushu yondashuvning kamchiligi shundaki, ma'lumotlarni qayta ishlash jarayonida kechikish kuzatiladi. Agar ma'lumotlarni qayta ishlash bir necha soatni tashkil etsa, hosil qilingan natija ma'naviy eskirgan (zaruriyat yo'qolgan) holga kelishi mumkin. Shu sababli ko'p hollarda real vaqt rejimidagi ba'zi natijalarni tahlil natijalar paketi bilan birlashtirish va foydalanish lozim.

Nathan Marz tomonidan Big Dataning Lambda arxitekturasi birinchi bo'lib taklif qilingan bo'lib, ikki turdagi ma'lumotlar oqimini yaratish orqali ushbu muammoni hal qilinadi [21]. Axborot tizimiga qabul qilinadigan barcha ma'lumotlar quyidagi ikki yo'ldan o'tadi [22]:

1. *Paketli daraja (sovuq yo'l).* Barcha qabul qilinayotgan ma'lumotlar qayta ishlanmagan tarzda saqlanadi va ular paket ko'rinishida qayta ishlanadi. Ushbu tarzda qayta ishlash natijalari paket ko'rinishida saqlanadi.

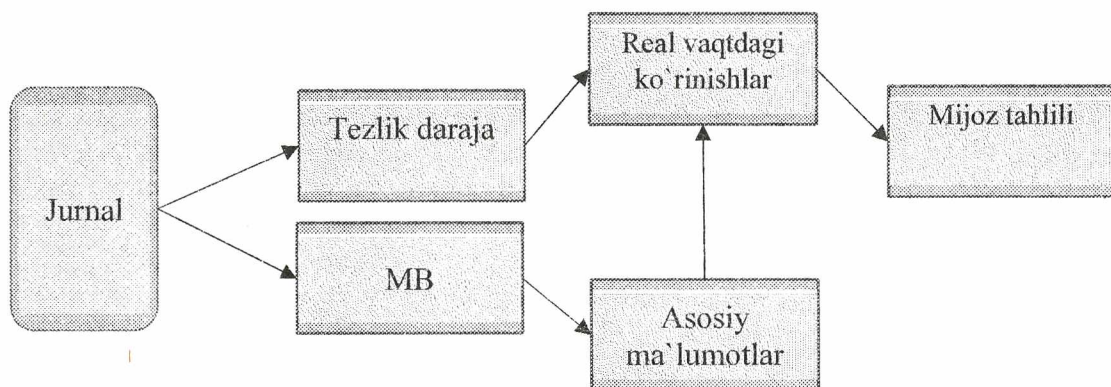
2. *Tezlashtirish darajasi (kritik yo'l).* Barcha qabul qilinayotgan real vaqt rejimida tahlil qilinadi. Ushbu darajada natijalarni juda tez qaytarilishini ta'minlaydi.



4-rasm. Big Dataning Lambda arxitekturasi

**Kappa arxitektura.** Big Data Lambda arxitekturasi kamchiligi uning murakkabligidir. Ma'lumotlarni qayta ishlash mantig'i ikki xil usulda qo'llanadi: *sovuq* va *kritik*. Ikkala usul arxitektura boshqaruvini murakkablashtiradi. Jay Kreps tomonidan Kappa arxitekturasi lambda arxitekturasi sifatida taklif qilingan [23]. U lambda arxitekturasi bilan bir xil maqsadga ega, ammo muhim farqli jihatli bor: qayta ishlangan ma'lumotlar oqimlari bitta yo'ldan o'tadi.





5-rasm. Big Dataning Kappa arxitekturasi

**Xulosa.** Hozirda ma'lumotlarni o'z vaqtida operativ qayta ishlash kompaniyaning asosiy vazifasi bo'lib, keng ko'lamlı ishlov berish va sifatli ma'lumotlarni olish kompaniyaning o'z ma'lumotlari bilan o'sishiga imkon beradigan muhim omil hisoblanadi. Ma'lumotlar hajmi kattalashganda va qayta ishlash platformasini noldan tiklash qimmatga tushganda, ko'plab kompaniyalar ma'lumotlarni qayta ishlash tizimining muammolariga duch kelishadi. Ma'lumotlarning parallelligi va o'lchovliligi printsipli boshidanoq puxta o'ylangan va ishlab chiqilgan bo'lishi kerak. Katta hajmdagi ma'lumotlarni qayta ishlash uchun Big Data arxitekturasi to'g'ri tashkil etish muhim ahamiyat kasb etadi.

#### FOYDALANILGAN ADABIYOTLAR

1. Kovalev I., Nezhmetdinov R. & Kvashnin D. Big data analytics of the technological equipment based on Data Lake architecture. MATEC Web of Conferences. 2019. <https://doi.org/10.1051/matecconf/201929800079>.
2. Cobb A. N., Benjamin A. J., Huang E. S. & Kuo P. C. Big data: More than big data sets. Surgery (United States), 2018. <https://doi.org/10.1016/j.surg.2018.06.022>.
3. Chauhan V., Gaur R., Tiwari A. & Shukla A. Real-time Big Data and Predictive Analytical Architecture for healthcare application. Sadhana - Academy Proceedings in Engineering Sciences, 2019. <https://doi.org/10.1007/s12046-019-1220-z>.
4. Lai B. C., Chen C. Y., Hsin Y. Da & Lin B. Y. A Two-Directional Big Data Sorting Architecture on FPGAs. IEEE Computer Architecture Letters. 2020. <https://doi.org/10.1109/LCA.2020.2993040>.
5. Malhotra S., Doja M. N., Alam B. & Alam M. Big data analysis and comparison of big data analytic approaches. Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017. <https://doi.org/10.1109/CCAA.2017.8229821>.
6. Lydia E. L., Sekhar G. C., Chevuru M. B., Ramya D. & Vijaya Kumar K. Text mining with apache hadoop over different hadoop clusters architectures. International Journal of Recent Technology and Engineering, 2019. <https://doi.org/10.35940/ijrte.B1866.078219>.
7. Sunny B. K., Janardhanan P. S., Francis A. B. & Murali R. Implementation of a self-adaptive real time recommendation system using spark machine learning libraries. 2017 IEEE International Conference on Signal

- Processing, Informatics, Communication and Energy Systems, SPICES, 2017. <https://doi.org/10.1109/SPICES.2017.8091310>.
8. Alshammari H., Lee J. & Bajwa H. H2hadoop: Improving hadoop performance using the metadata of related jobs. *IEEE Transactions on Cloud Computing*, 2018. <https://doi.org/10.1109/TCC.2016.2535261>.
  9. Hussain T., Sanga A. & Mongia S. Big Data Hadoop Tools and Technologies: A Review. *SSRN Electronic Journal*, 2019. <https://doi.org/10.2139/ssrn.3462554>.
  10. Dean J. & Ghemawat S. MapReduce. *Communications of the ACM*, 2010. <https://doi.org/10.1145/1629175.1629198>.
  11. Dean J. & Ghemawat S. MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 2018. <https://doi.org/10.1145/1327452.1327492>.
  12. Kalia K. & Gupta N. Analysis of hadoop MapReduce scheduling in heterogeneous environment. In *Ain Shams Engineering Journal*, 2021. <https://doi.org/10.1016/j.asej.2020.06.009>.
  13. Gupta S. & Narsimha G. Efficient query analysis and performance evaluation of the Nosql data store for big data. *Advances in Intelligent Systems and Computing*, 2017. [https://doi.org/10.1007/978-981-10-2471-9\\_53](https://doi.org/10.1007/978-981-10-2471-9_53).
  14. Ethirajan D. & Purushotham. Big data performance evaluation for effective E-governance applications // *Journal of Advanced Research in Dynamical and Control Systems*, 2019. <https://doi.org/10.5373/JARDCS/V11I1SP10/20192995>.
  15. Mathew A. B. & Kumar S. D. M. Novel research framework on SN's NoSQL databases for efficient query processing // *International Journal of Reasoning-Based Intelligent Systems*, 2015. <https://doi.org/10.1504/IJRIS.2015.072959>.
  16. Potey M., Digrase M., Deshmukh G. & Nerkar M. Database Migration from Structured Database to non- Structured Database // *International Journal of Computer Applications*, 2015.
  17. Moustafa N. A Systemic IoT-Fog-Cloud Architecture for Big-Data Analytics and Cyber Security Systems: A Review of Fog Computing. In *arXiv*. 2019.
  18. Andrade R., Torres J. & Flores P. Management of information security indicators under a cognitive security model. 2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018. <https://doi.org/10.1109/CCWC.2018.8301745>.
  19. Yang K., Lee D., Kim K. & Yoon H. Analysis of Security Threat and Security Requirements of the Big data System // *Journal of Security Engineering*, 2016. <https://doi.org/10.14257/jse.2016.12.08>.
  20. Ahmad S., Yasin A. & Shafi Q. DDoS attacks analysis in big data (hadoop) environment. *Proceedings of 2018 15th International Bhurban Conference on Applied Sciences and Technology, IBCAST*, 2018. <https://doi.org/10.1109/IBCAST.2018.8312270>.
  21. <https://ajstorm.medium.com/the-lambda-architecture-simplified-a28e436fa55e>.
  22. Марц Натан, Уоррен Джеймс. Большие данные. Принципы и практика построения масштабируемых систем. ISBN978-5-8459-2075-1.368. С.2016.
  23. <https://jonboulineau.me/blog/architecture/kappa-architecture>.



## СОДЕРЖАНИЕ

### Информатика и управление

Д.К.Мухамедиева, А.Ю.Нурумова. Blowup решения для квазилинейных параболических уравнений на основе автомоделного анализа решений.....	3
Д.Т. Мухамедиева, А.Х.Мирзарахмедова. Разработка модели определения оптимального количества городского пассажирского транспорта...	8
И.Х.Сиддиков, Н.У.Утеулиев, Р.Ж.Алламуратов.Разработка кинетической модели процесса аммонизации рассола.....	13
Б.М.Азимов, С.А.Мухамматов. Моделирование движения горизонтально-шпиндельного хлопкоуборочного аппарата и оптимизация ее параметров.....	19
Э.Урунбаев.Математический метод построения оптимального корректора эвристических алгоритмов.....	28

### Энергетика

О.З. Тоиров, Т.М. Саъдуллаев. Современное состояние и вопросы энерго- и ресурсосбережения в буровой установке горнодобывающей промышленности.....	35
Ж.С. Ахатов, К.А. Самиев, У.Б. Шаропов, Х.С. Ахмадов, А.С. Халимов, Э.Т. Жураев, Х.Ф. Сайфиева, Т.И. Жураев. Получение водорода методом термохимических циклов на основе $\text{CH}_4$ и $\text{CeO}_2$ при концентрированном потоке солнечного излучения.....	45
М.С. Якубов, К.Х. Турдибеков, А.Х. Суллийев, И.А. Каримов, М.Ж. Мирасадов. Совершенствование информационно-измерительного комплекса диагностирования объектов тягового электроснабжения при высокоскоростном движении.....	54
Н.О. Усмонов, А.Р. Иванисова. Моделирование тепло- и массообмена двухступенчатых испарительных воздухоохладителей.....	63

### Информационные и телекоммуникационные технологии

Н.С.Маматов, Н.А.Ниёзматова, Ш.Ш.Абдуллаев, П.Б.Нуримов, А.Н.Самижонов, К.К. Ережепов. Создание системы автоматического распознавания узбекской речи.....	72
Д.Э.Эшмурадов, Т.Д.Элмурадов, А.А.Сулайманов.Программное и информационное обеспечение процессов управления воздушным движением .....	85
Б.Б.Элов Х.М.Примова. Архитектура Big Data.....	91

## МУНДАРИЖА

### Информатика ва бошқарув

Д.К.Мухамедиева, А.Ю.Нурумова. Ечимни автомодел таҳлили асосида параболик кўринишдаги квазичизиқли тенгламанинг Blow up ечими.....	3
Д.Т.Мухамедиева, А.Х.Мирзарахмедова. Шахар йўловчилар транспортининг оптимал сонини аниқлаш моделини ишлаб чиқиш.....	8
I.N.Siddikov, N.U.Uteuliev, R.J.Allamuratov. Eritmani ammoni-zatsiyalash jarayonining kinetik modelini ishlab chiqish.....	13
Б.М.Азимов, С.А.Мухамматов. Горизонтал шпинделли пахта териш аппарати ҳаракатини моделлаштириш ва параметрларини оптималлаштириш.....	19
Е. Urunbaev. Evristik algoritmlarning optimal tuzatuvchisini qurishning matematik usuli.....	28

### Энергетика

О.З. Тоиров, Т.М. Саъдуллаев. Тоғ -қон саноатидаги бурғулаш қурилмаларининг энергия ва ресурс тежамкорлик масалалари ва ҳозирги ҳолати.....	35
Ж.С. Ахатов, К.А. Самиев, У.Б. Шаропов, Х.С. Ахмадов, А.С.Ҳалимов, Э.Т. Жўраев, Х.Ф. Сайфиева, Т.И. Жўраев. Мужассамлашган қуёш нуллари оқимида $\text{CH}_4$ ва $\text{CeO}_2$ асосида термохимёвий цикл орқали водород олиш.....	45
М.С. Якубов, К.Х. Турдибеков, А.Х. Суллийев, И.А. Каримов, М.Ж. Мирасадов. Юқори тезликдаги ҳаракат таркиби тортиш электр таъминоти объектларини диагностикалаш учун ахборот-ўлчов мажмуасини такомиллаштириш.....	54
Н.О.Усмонов, А.Р.Иванисова. Икки босқичли буғланма ҳаво совитгичлар исиклик ва масса алмашинувини моделлаштириш.....	63

### Ахборотли ва телекоммуникацияли технологиялар

Н.С.Маматов, Н.А.Ниёзматова, Ш.Ш.Абдуллаев, П.Б.Нуримов, А.Н.Самижонов, К.К. Ережепов. Ўзбек тилидаги нутқни автоматик таниб олиш тизимини яратиш.....	72
Д.Э.Эшмурадов, Т.Д.Эльмурадов, А.А.Сулайманов. Ҳаводаги ҳаракатни бошқариш жараёнларининг дастурий ва информацион таъминоти.....	85
В.В. Elov X.M.Primova. Big data arxitekturasi.....	91



Индекс 879

ISSN 2010-7242 ПРОБЛЕМЫ ИНФОРМАТИКИ И ЭНЕРГЕТИКИ. №4 2021