

The background of the slide is a light gray gradient. It is decorated with numerous realistic water droplets of various sizes. Some droplets are large and prominent, while others are small and subtle. They are scattered across the slide, with a higher concentration in the top-left and bottom-right corners. The droplets have a glossy, reflective surface with highlights and shadows, giving them a three-dimensional appearance.

# РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ: ОТ ТЕОРИИ К ПРАКТИКЕ

IGOR KHAPOV

CTO IBM RUSSIA/CIS

[KHAPOV.IGOR@RU.IBM.COM](mailto:KHAPOV.IGOR@RU.IBM.COM)

# О ЧЕМ ЭТОТ КУРС ?

- WHAT IBM IS ?
- FIRST DISTRIBUTED COMPUTING SYSTEM
- CLOUD COMPUTING VS DISTRIBUTED COMPUTING
- IAAS, PAAS, SAAS, CAAS
- MAPREDUCE - HADOOP, APACHE SPARK
- SERVERLESS









IBM занимается разработкой технологий и инновациями. Штаб-квартира корпорации находится в городе Армонк (шт. Нью-Йорк)

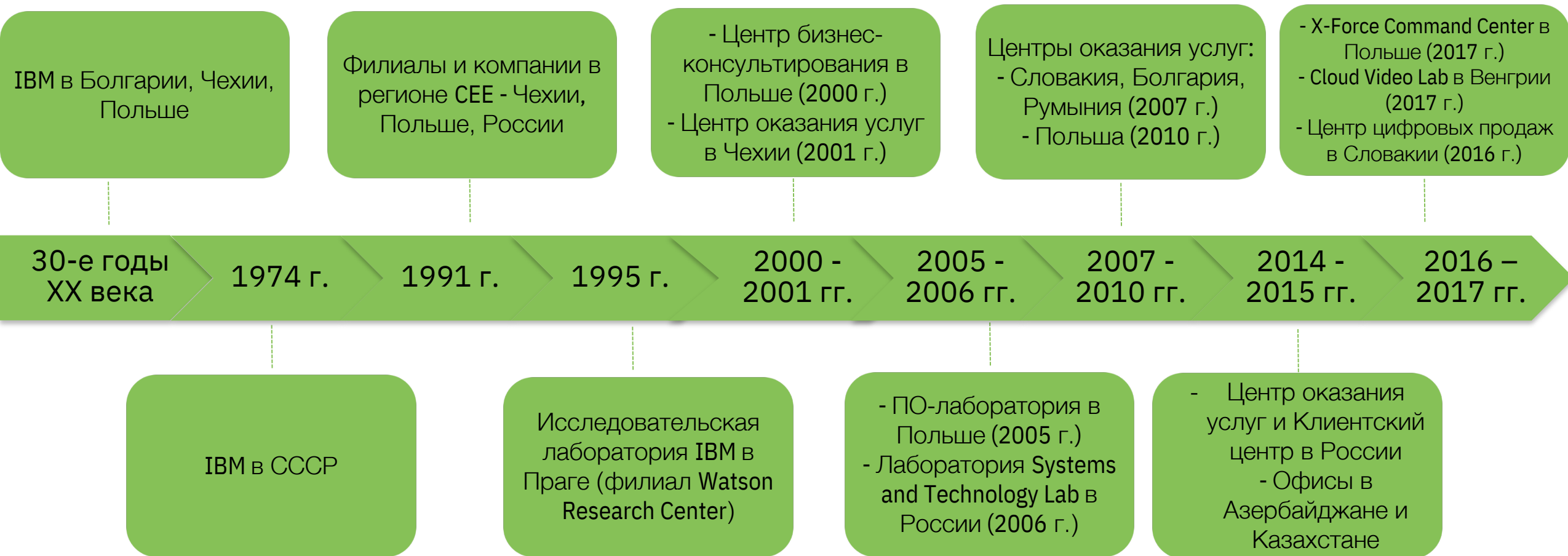
IBM – крупнейший в мире технологический работодатель. В IBM работает свыше 350 тыс. человек, которые обслуживают заказчиков из 170 стран мира

Доход компании в 2018 году составил \$79,6 млрд

IBM ежегодно инвестирует около 7% дохода в НИОКР и на протяжении 26 лет является лидером по числу патентов

Технологии и возможности IBM обладают потенциалом трансформации организаций, сообществ и качества жизни.

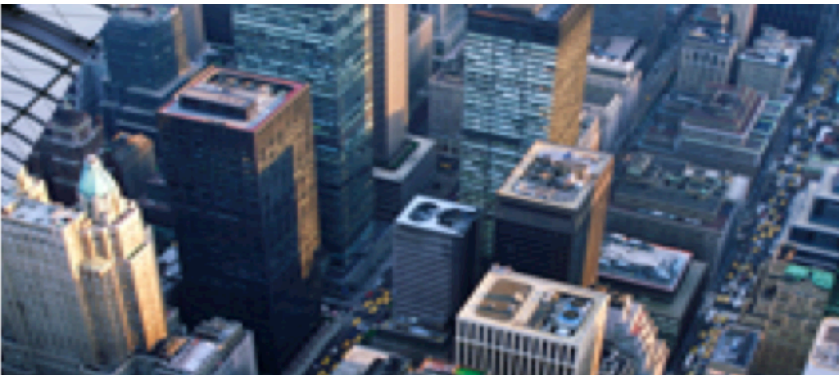
# Развитие IBM в регионе СЕЕ, в том числе в России и СНГ





# КОМПАНИЯ IBM ПОДГОТОВИЛАСЬ К НАСТУПЛЕНИЮ НОВОЙ ЭПОХИ, НАКОПИВ ОГРОМНЫЙ ОПЫТ В СФЕРЕ ТРАНСФОРМАЦИИ КОМПАНИЙ, ЦЕЛЫХ ОТРАСЛЕЙ И ОТДЕЛЬНЫХ ПРОФЕССИЙ

- > 97% крупнейших банков в мире используют продукты IBM в операционной деятельности
- > 80% бронирований поездок в мире обрабатывается системами IBM
- > 90% транзакций по кредитным картам в мире обрабатывается мейнфреймами IBM
- > 80% международных компаний в сфере розничной торговли используют решения IBM для розничной торговли
- > 83% крупнейших провайдеров услуг связи в мире являются клиентами IBM



Быстрое создание и монетизация масштабных когнитивных приложений для оказания финансовых услуг



Трансформация, реорганизация и изменение здравоохранения. Помощь в решении наиболее сложных проблем, таких как лечение рака



Использование данных для получения ценных сведений, выбора оптимального образа действий, повышения эффективности бизнеса и качества обслуживания клиентов в розничной торговле

# IBM Research и исследования глобального масштаба





# Направления и области исследований

Искусственный интеллект



Блокчейн



Квантовые  
вычисления



Интернет вещей



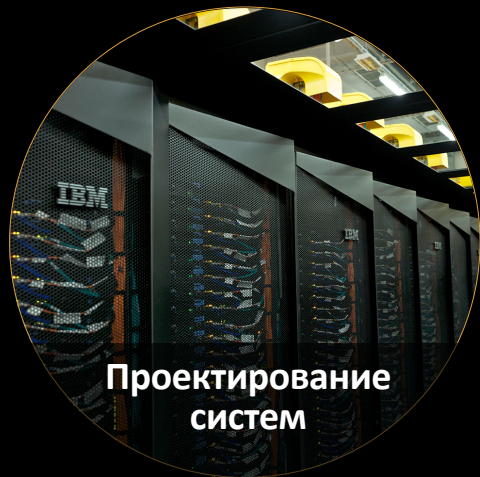
Информационная  
безопасность



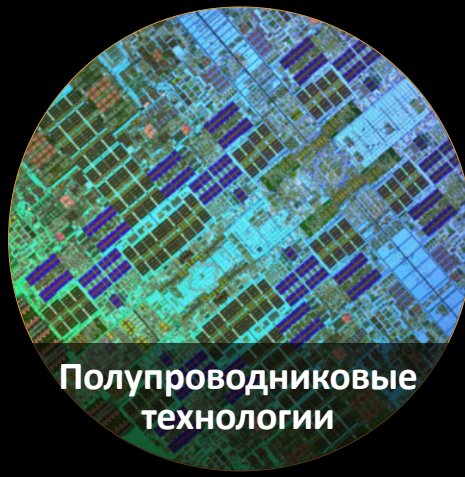
Облачные  
вычисления



Проектирование  
систем



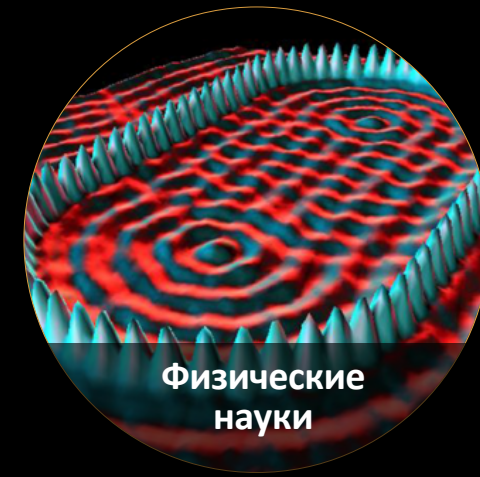
Полупроводниковые  
технологии



Биология



Физические  
науки





# ВЕДУЩИЕ КОМПАНИИ РОССИИ ИСПОЛЬЗУЮТ ПЕРЕДОВЫЕ РЕШЕНИЯ IBM ДЛЯ ЗАКРЕПЛЕНИЯ СВОЕГО ЛИДЕРСТВА



НОВЫЕ ПЛАСТЫ УГЛЕВОДОРОДОВ ВПЕРВЫЕ  
ВЫЯВЛЕНЫ САМООБУЧАЮЩЕЙСЯ  
ПРОГРАММОЙ, РАЗРАБОТАННОЙ НАУЧНО-  
ТЕХНИЧЕСКИМ ЦЕНТРОМ  
ПАО «ГАЗПРОМ НЕФТЬ» СОВМЕСТНО С IBM  
SERVICES



РОЗНИЧНАЯ СЕТЬ «МАГНИТ» ВЫБРАЛА  
ПЛАТФОРМУ IBM QRADAR ДЛЯ ПЕРЕХОДА К  
КОНЦЕПЦИИ ЦИФРОВОГО ДОВЕРИЯ



БЛОКЧЕЙН-ЛАБОРАТОРИЯ  
ПАО СБЕРБАНК В СОТРУДНИЧЕСТВЕ С  
IBM РАБОТАЕТ НАД СОЗДАНИЕМ  
БЛОКЧЕЙН-ПРИЛОЖЕНИЙ И  
ПРОДВИЖЕНИЕМ ЭТОЙ НОВОЙ  
ТЕХНОЛОГИИ В РОССИИ

Что мы знаем о прогнозировании технологий будущего?



Что мы знаем о прогнозировании технологий будущего?



1994 г.

# Что мы знаем о прогнозировании технологий будущего?



1994 г.



2007 г.

Что мы знаем о прогнозировании технологий будущего?



1999 г.



# Что мы знаем о прогнозировании технологий будущего?



1999 г.



2013 г.

Что мы знаем о прогнозировании технологий будущего?



2000 г.



# Что мы знаем о прогнозировании технологий будущего?



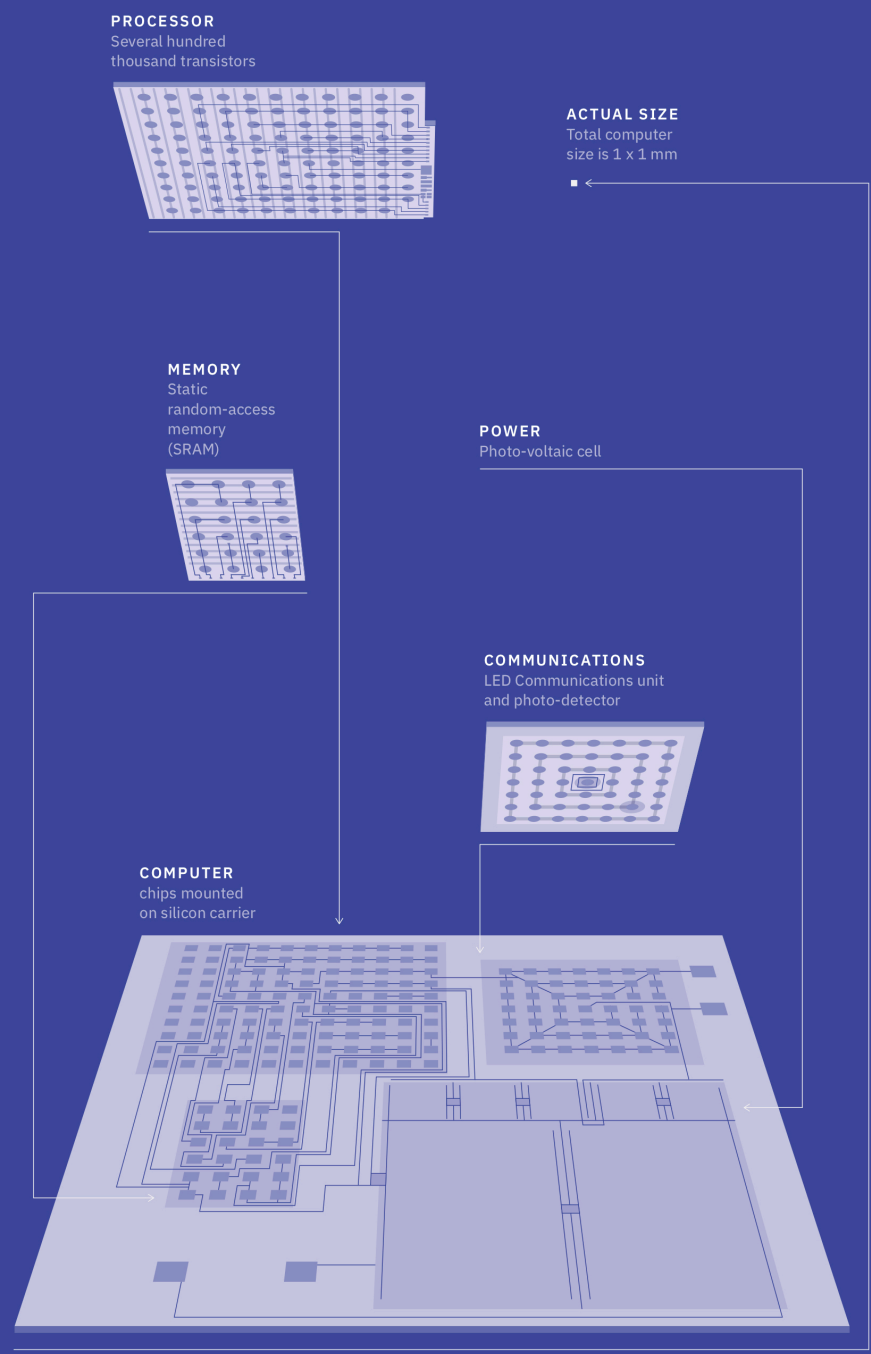
2000 г.



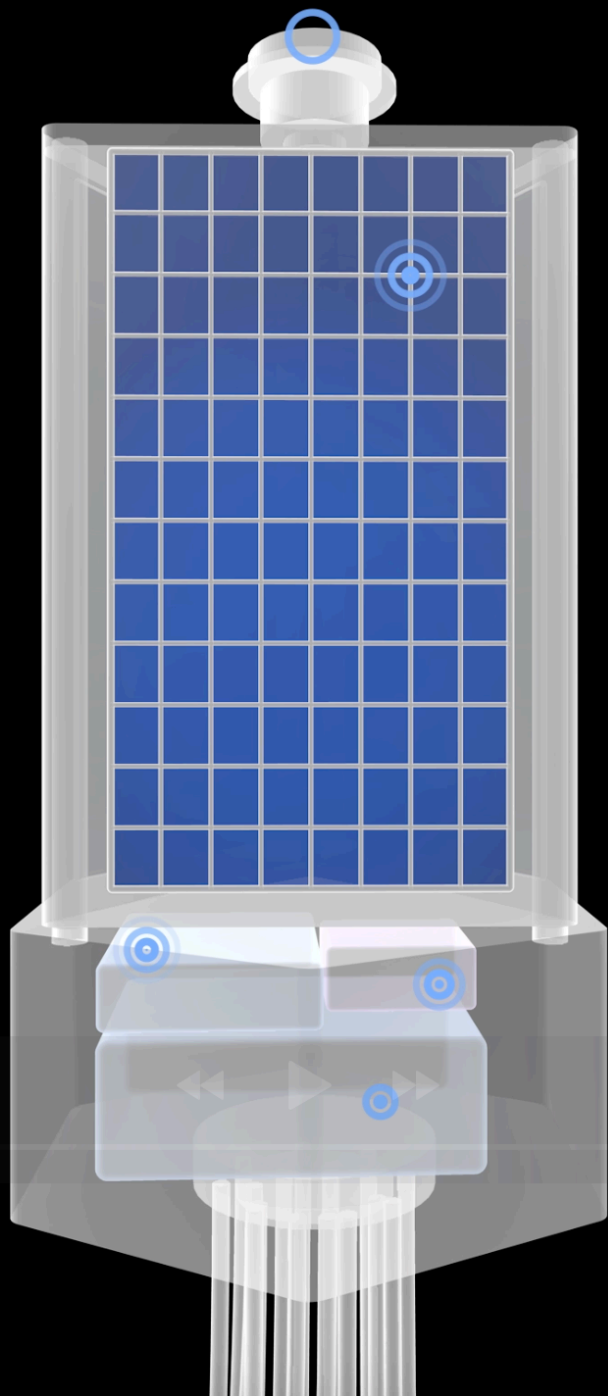
2015 г.



# Самый маленький компьютер в мире







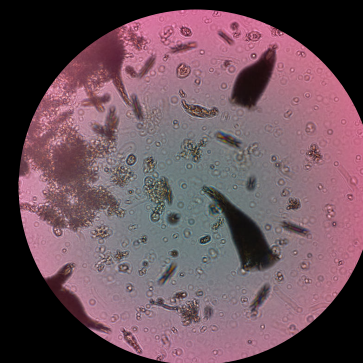
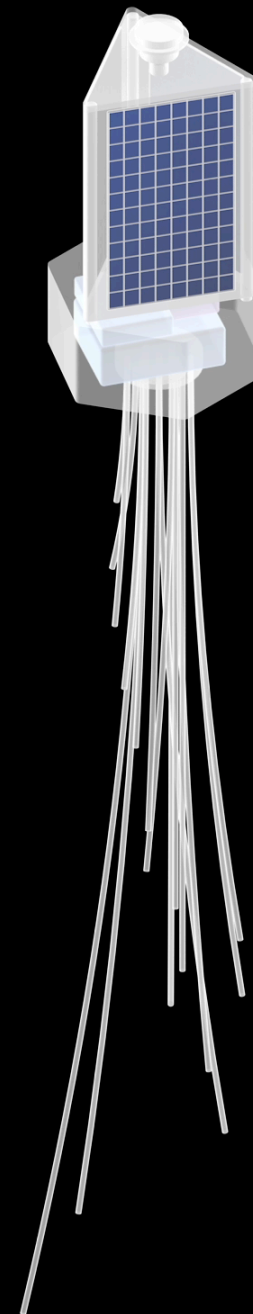
Антенна

Солнечная  
батарея

Батарейки  
и электроника

3D-микроскоп

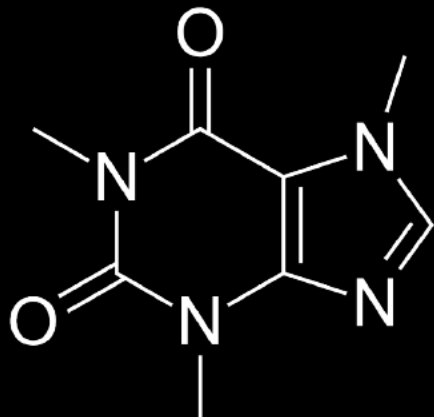
Насос



Планктон под  
микроскопом IBM

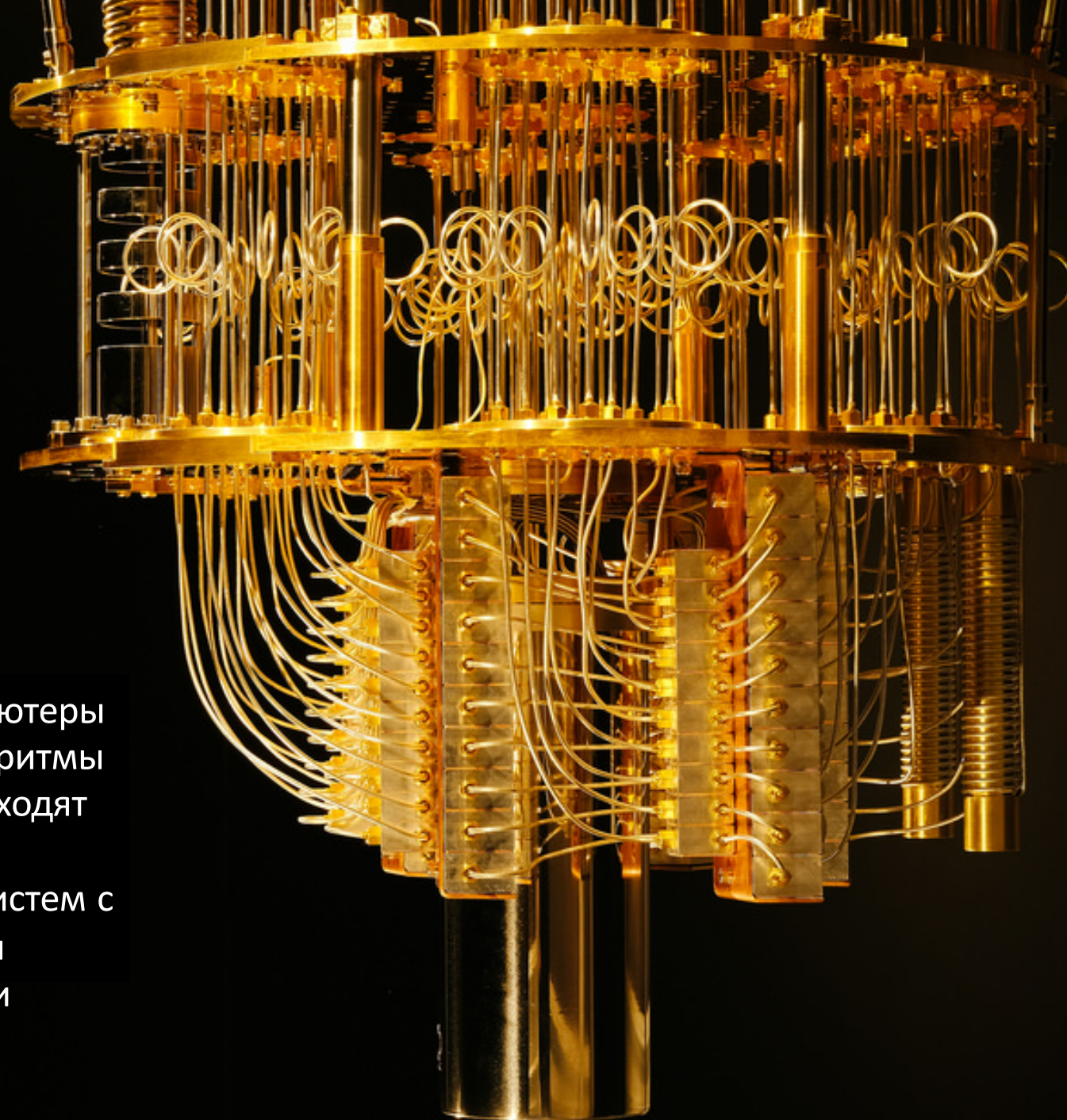


# Квантовые вычисления скоро станут мейнстримом



Классические компьютеры требуют экспоненциального масштабирования для симуляции и хранения квантовых состояний химических молекул

Квантовые компьютеры и квантовые алгоритмы оптимально подходят для задач моделирования систем с квантовыми состояниями





IBM Q Experience:  
свободный облачный  
доступ уже сегодня

~80 000  
пользователей

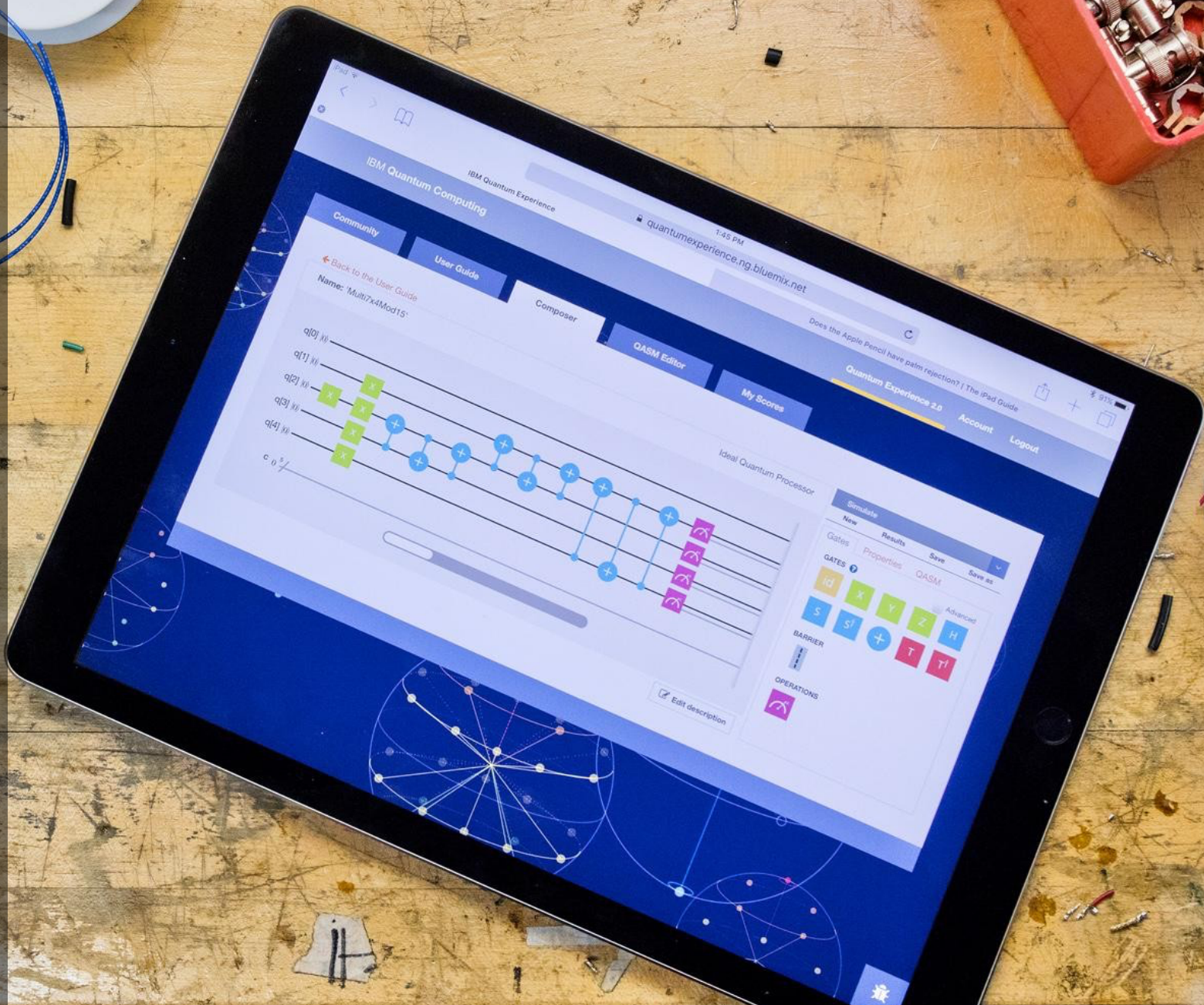
143 страны

3 млн запусков

>60 научных статей

1 500 университетов

300 высших школ





# ONE OF THE FIRST DISTRIBUTED SYSTEM - IBM 3790

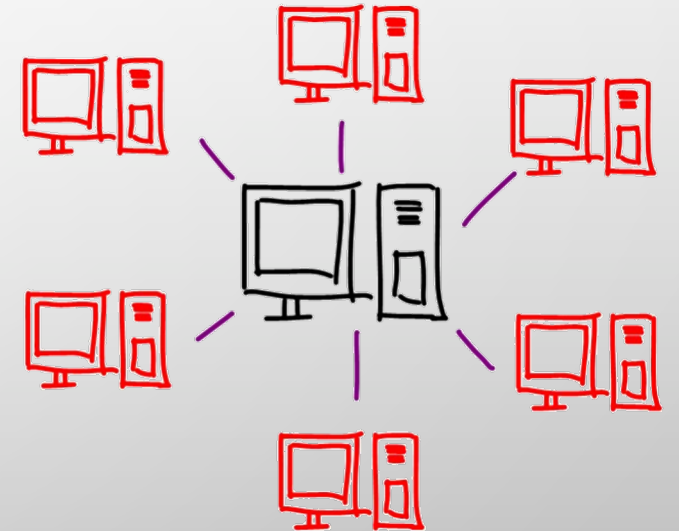
THE **IBM 3790** COMMUNICATIONS SYSTEM, DEVELOPED BY [IBM](#)'S DATA PROCESSING DIVISION (DPD), WAS ANNOUNCED IN **1974** AND SUPPORTED:

- UP TO 16 IBM [3277](#) DISPLAY STATIONS
- AN INTEGRATED [FLOPPY DISK](#) UNIT
- UP TO FOUR 3793 KEYBOARD-PRINTERS
- UP TO THREE 3292 AUXILIARY CONTROL UNITS
- AN INTEGRATED 120 LINES PER MINUTE (LPM) [LINE PRINTER](#)
- A [SYNCHRONOUS DATA LINK CONTROL](#) (SDLC) COMMUNICATIONS INTERFACE
- A 1200 BAUD INTERNAL OR EXTERNAL MODEM



# DISTRIBUTED COMPUTING VS CLOUD COMPUTING

- **DISTRIBUTED COMPUTING** IS WHEN MULTIPLE AUTONOMOUS MACHINES COMMUNICATE THROUGH A CENTRAL NETWORK TO PERFORM A COMMON GOAL.
- **CLOUD COMPUTING** IS MORE ABOUT COMPUTING AS A SERVICE, THAT IS GIVEN TO A COMPUTER OVER A NETWORK.



# CLOUD COMPUTING

IS THE DELIVERY OF ON-DEMAND COMPUTING  
RESOURCES — EVERYTHING **FROM APPLICATIONS TO**  
**DATA CENTERS** — OVER THE INTERNET ON A PAY-  
FOR-USE BASIS.







## CLOUD ...

- ... ENABLED
- ... READY
- ... BASED
- ... FIRST
- ... NATIVE



## *DISTRIBUTED ...*

- ... SYSTEM
- ... PROGRAMMING
- ... ALGORITHM



# CLOUD NATIVE

- AN **APPLICATION ARCHITECTURE** DESIGNED TO USE THE ***STRENGTHS*** AND ACCOMMODATE THE ***CHALLENGES*** OF A STANDARDIZED **CLOUD** ENVIRONMENT , INCLUDING THESE CONCEPTS: ELASTIC SCALING, IMMUTABLE DEPLOYMENT, DISPOSABLE INSTANCES, AND LESS PREDICTABLE INFRASTRUCTURE.



# WHAT IT MEANS TO BE *CLOUD NATIVE*

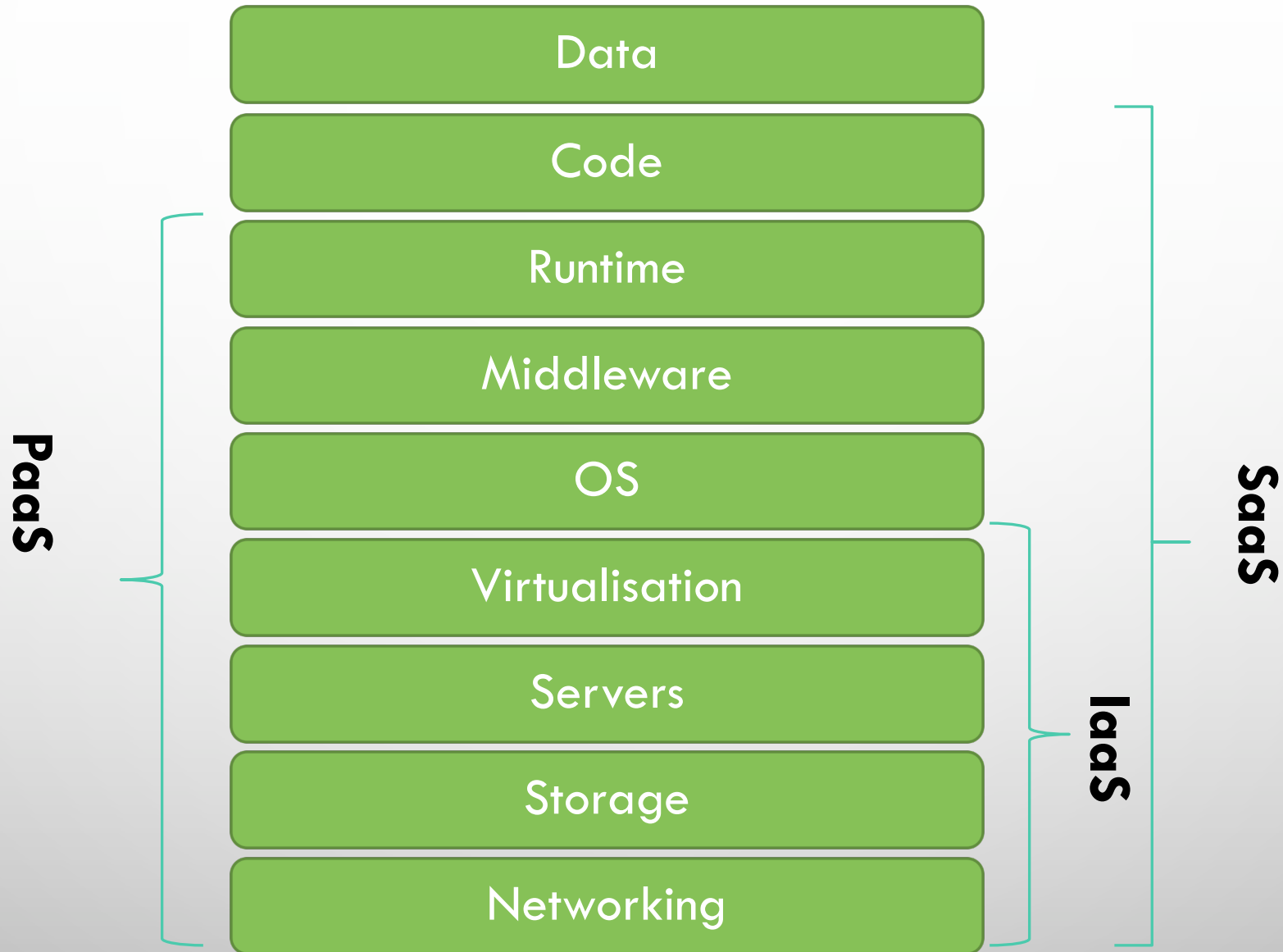
- CLEAN CONTRACT WITH UNDERLYING OS – **MAXIMUM PORTABILITY**
- SCALE **ELASTICALLY** WITHOUT SIGNIFICANT CHANGES TO TOOLING AND DEV PRACTICES
- **RESILIENT TO INEVITABLE FAILURES** IN THE INFRASTRUCTURE AND APPLICATION
- USE **CLOUD SERVICES** – STORAGE, QUEUING, CACHING,...
- RAPID AND REPEATABLE **DEPLOYMENTS** – MAXIMIZE AGILITY
- **AUTOMATED SETUP** – MINIMIZE TIME AND COST FOR NEW DEVELOPERS

# DISTRUBUTED COMPUTING

- COMPUTING
- DATABASES
- INTEGRATION



# COMPUTING



# COMPUTING

- DISTRIBUTED MONOLITH
- SOA
- MICROSERVICE
- EDGE COMPUTING
- BLOCKCHAIN
- SERVERLESS
- MAPS/REDUCE

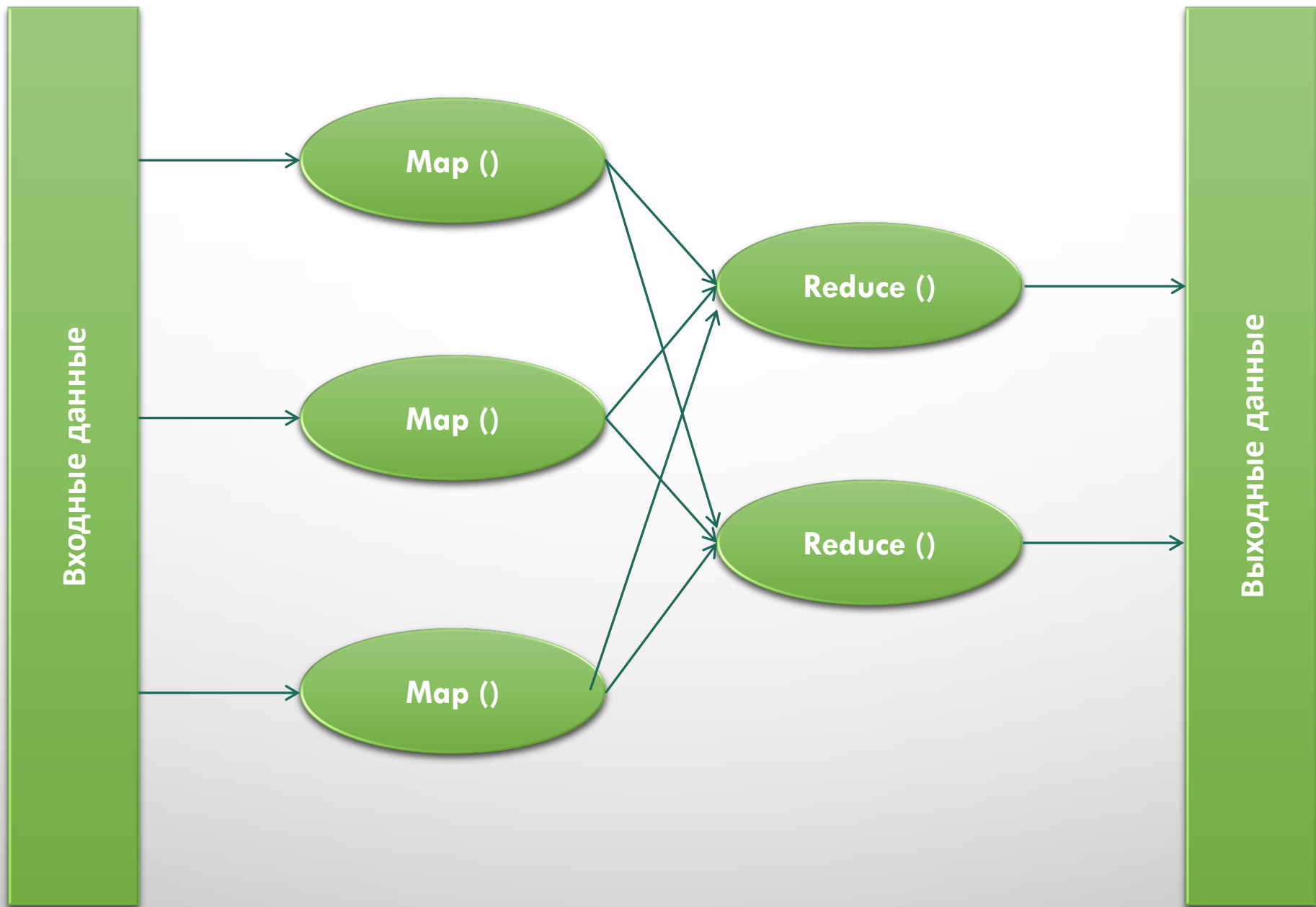
# MapReduce

- ПРОГРАММНАЯ МОДЕЛЬ ПРЕДНАЗНАЧЕННАЯ ДЛЯ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ ЗА СЧЕТ РАЗДЕЛЕНИЯ ЗАДАЧИ НА НЕЗАВИСИМЫЕ ЧАСТИ.

JEFFREY DEAN, SANJAY GHEMAWAT.

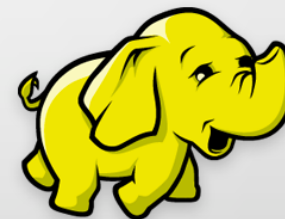
*“SIMPLIFIED DATA PROCESSING ON LARGE CLUSTERS”* 2004





# HADOOP ЭТО

СВОБОДНО РАСПРОСТРАНЯЕМЫЙ НАБОР БИБЛИОТЕК И УТИЛИТ ДЛЯ  
РАЗРАБОТКИ И ВЫПОЛНЕНИЯ РАСПРЕДЕЛЕННЫХ ПРОГРАММ,  
РАБОТАЮЩИХ НА КЛАСТЕРАХ ИЗ СОТЕН И ТЫСЯЧ ВЫЧИСЛИТЕЛЬНЫХ  
УЗЛОВ.



## Управление данными и системой

ZooKeeper

Chukwa

## Доступ к данным

Pig

Hive

Avro

## Обработка данных

MapReduce

## Хранение данных

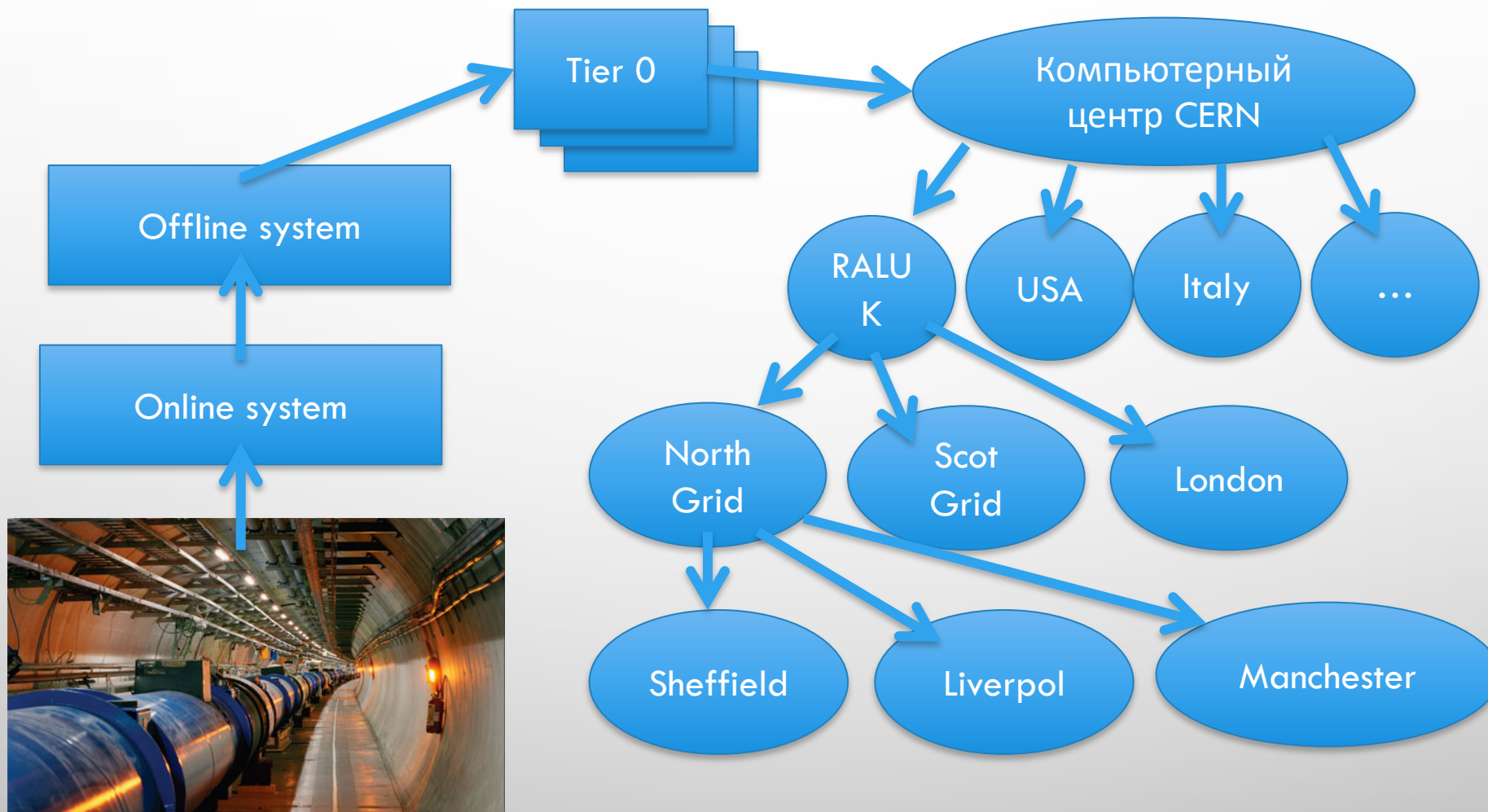
HDFS

HBASE



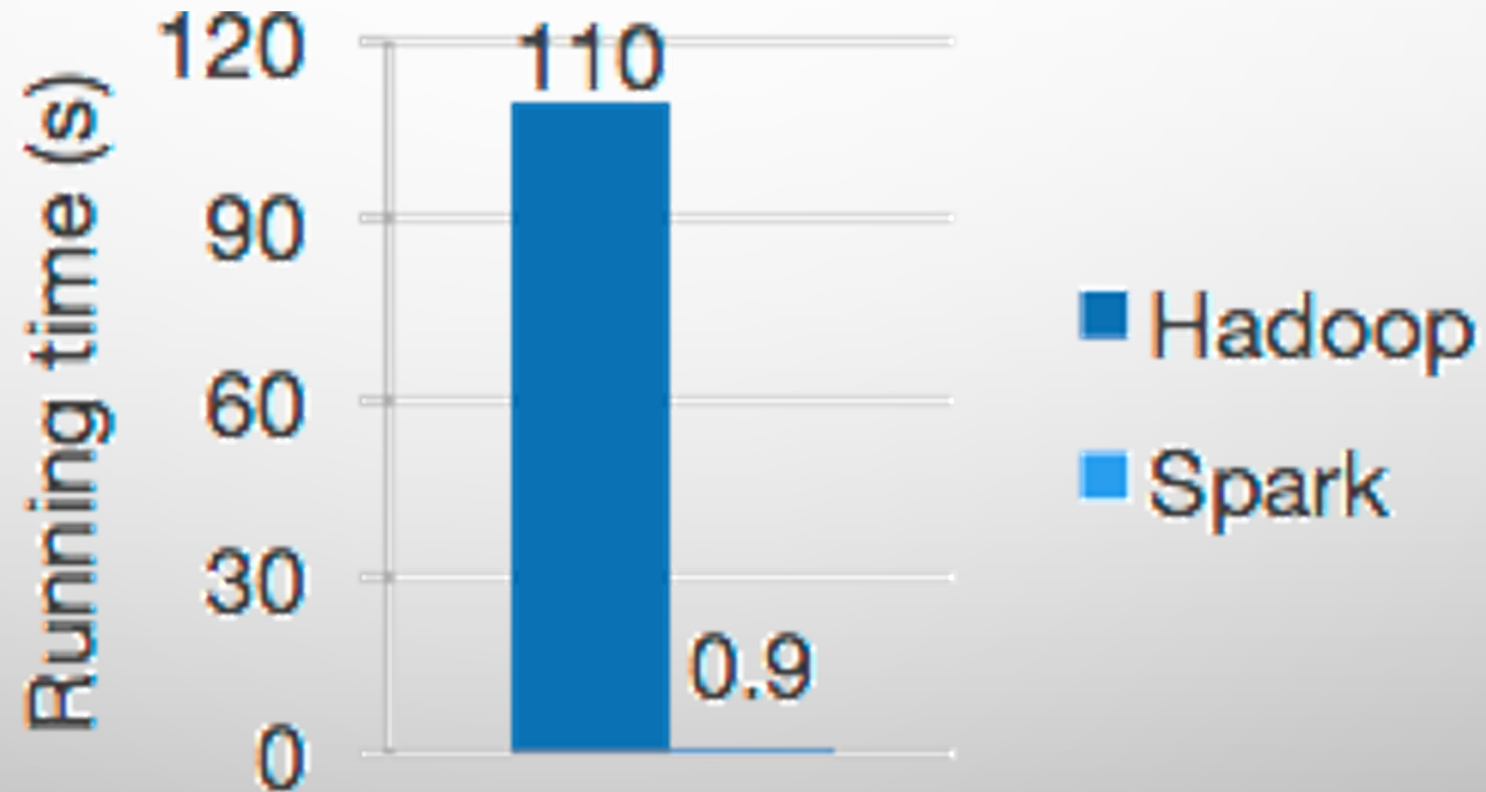
**Кому это нужно?**

БОЛЬШОЙ АДРОННЫЙ КОЛЛАЙДЕР ИСПОЛЬЗУЕТ HADOOP ДЛЯ РАССЧЕТОВ РЕЗУЛЬТАТОВ. ДАННЫЕ ТОЛЬКО В HDFS



# SPARK

ПРОГРАММНЫЙ КАРКАС ДЛЯ РЕАЛИЗАЦИИ РАСПРЕДЕЛЕННОЙ ОБРАБОТКИ. В ОТЛИЧИЕ ОТ HADOOP ХРАНИТ ДАННЫЕ В ОПЕРАТИВНОЙ ПАМЯТИ И КАК СЛЕДСТВИЕ УСКОРЯЕТ ВЫПОЛНЕНИЕ НЕКОТОРЫХ ЗАДАЧ.





# ReactJS and Flux



Workload Transformation Tool

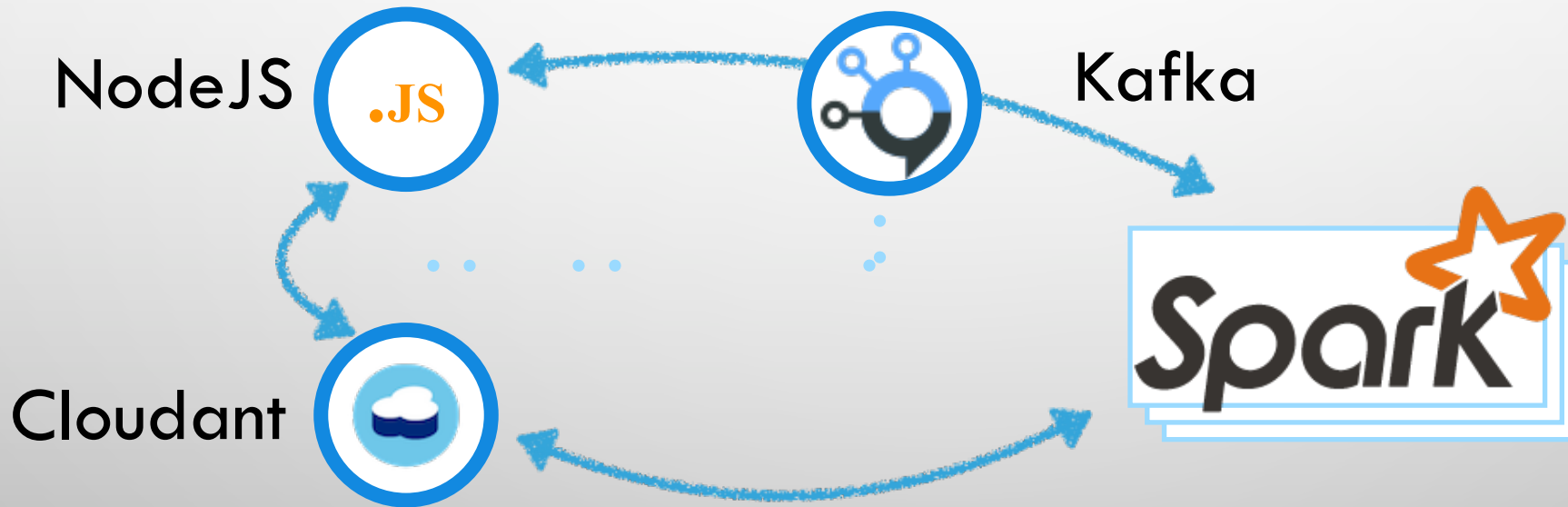
Dashboard Upload Reports Account

Latest activity [ Migration to the SoftLayer ]

Run Plugin1

id	image name	os name	os version	migratable
1	server_7	Windows	06.01.7601	No
2	server_2	Windows	5.0.2195	No
3	server_11	Windows	06.01.7601	No
4	server_8	unknown	unknown	No
5	server_3	Windows	5.0.2195	No
6	server_4	Windows	06.01.7601	No
7	server_12	Windows	06.01.7601	No
8	server_6	Windows	06.01.7601	No
9	server_10	Windows	06.01.7601	No
10	server_1	Linux	unknown	No
11	server_5	Windows	06.01.7601	No
12	server_9	Windows	06.01.7601	No

© WTA Project 2015

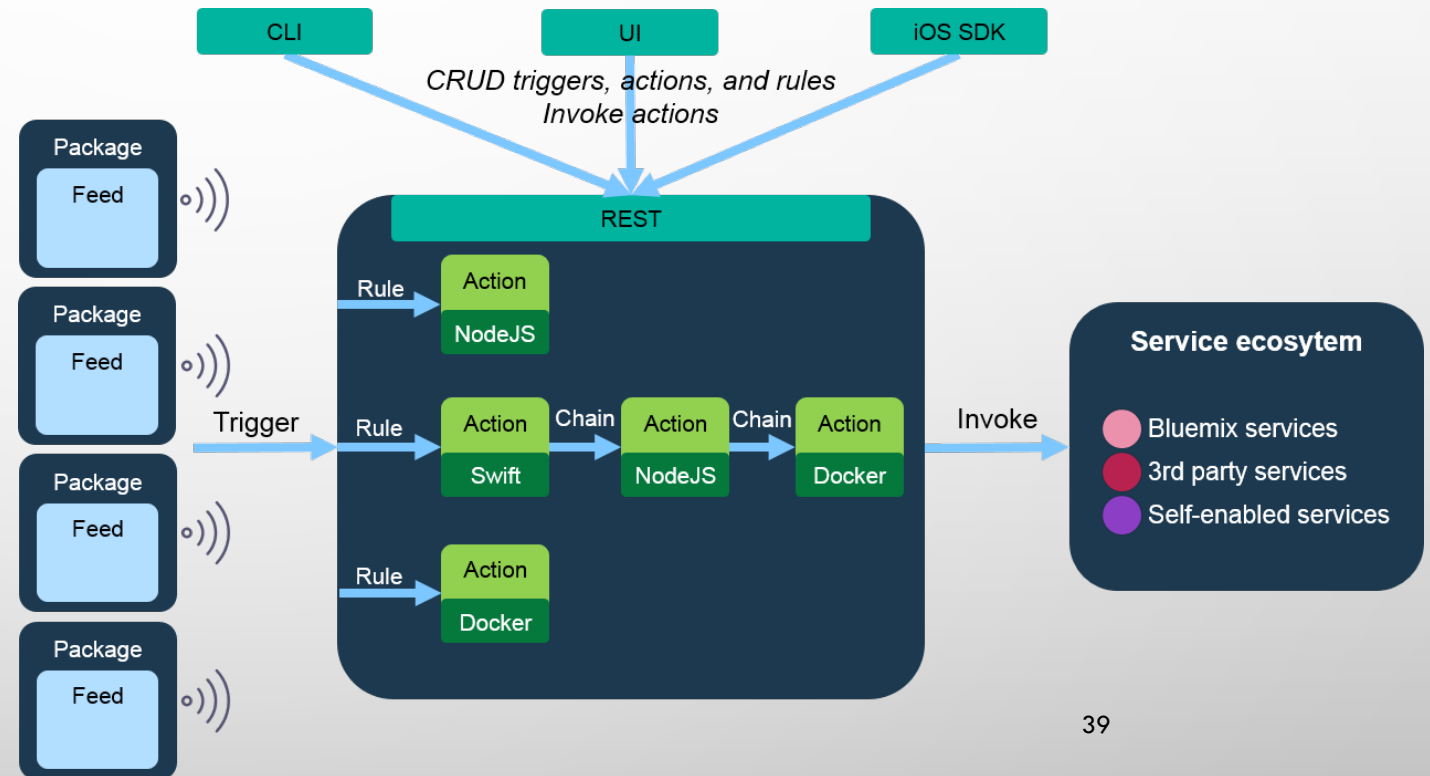


# SERVERLESS

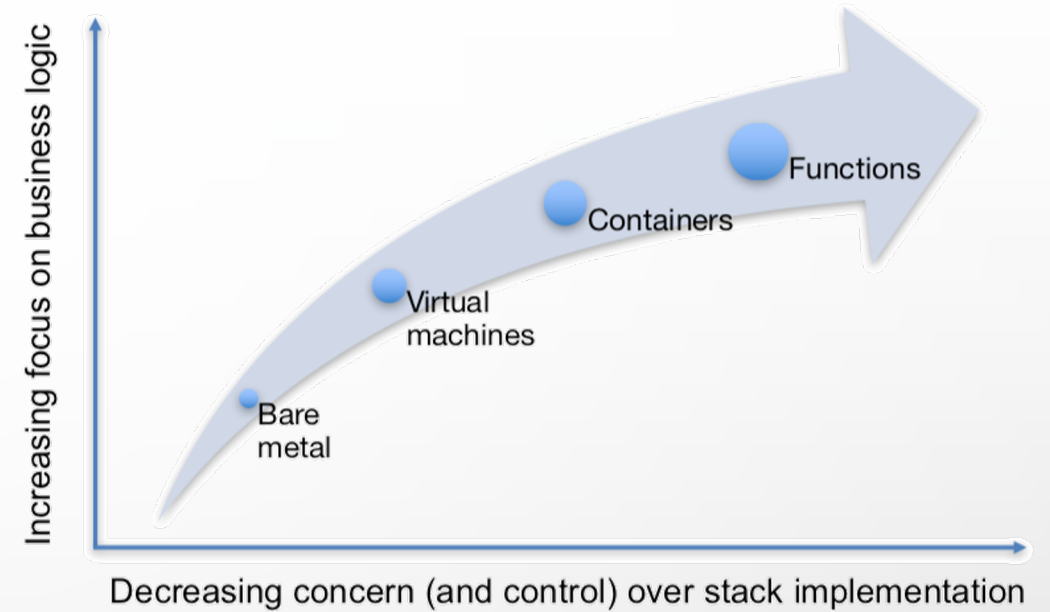
**Serverless** is cloud computing execution model which provides new agnostic level to start business logic. It simplify development and DevOps processes and reduce complexity of maintaining application layers.

## Serverless platform

**Serverless platform** supports a programming model in which developers write functional logic (called **Actions**), in any supported programming language, that can be dynamically scheduled and run in response to associated events (via **Triggers**) from external sources ( **Feeds**) or from HTTP requests.



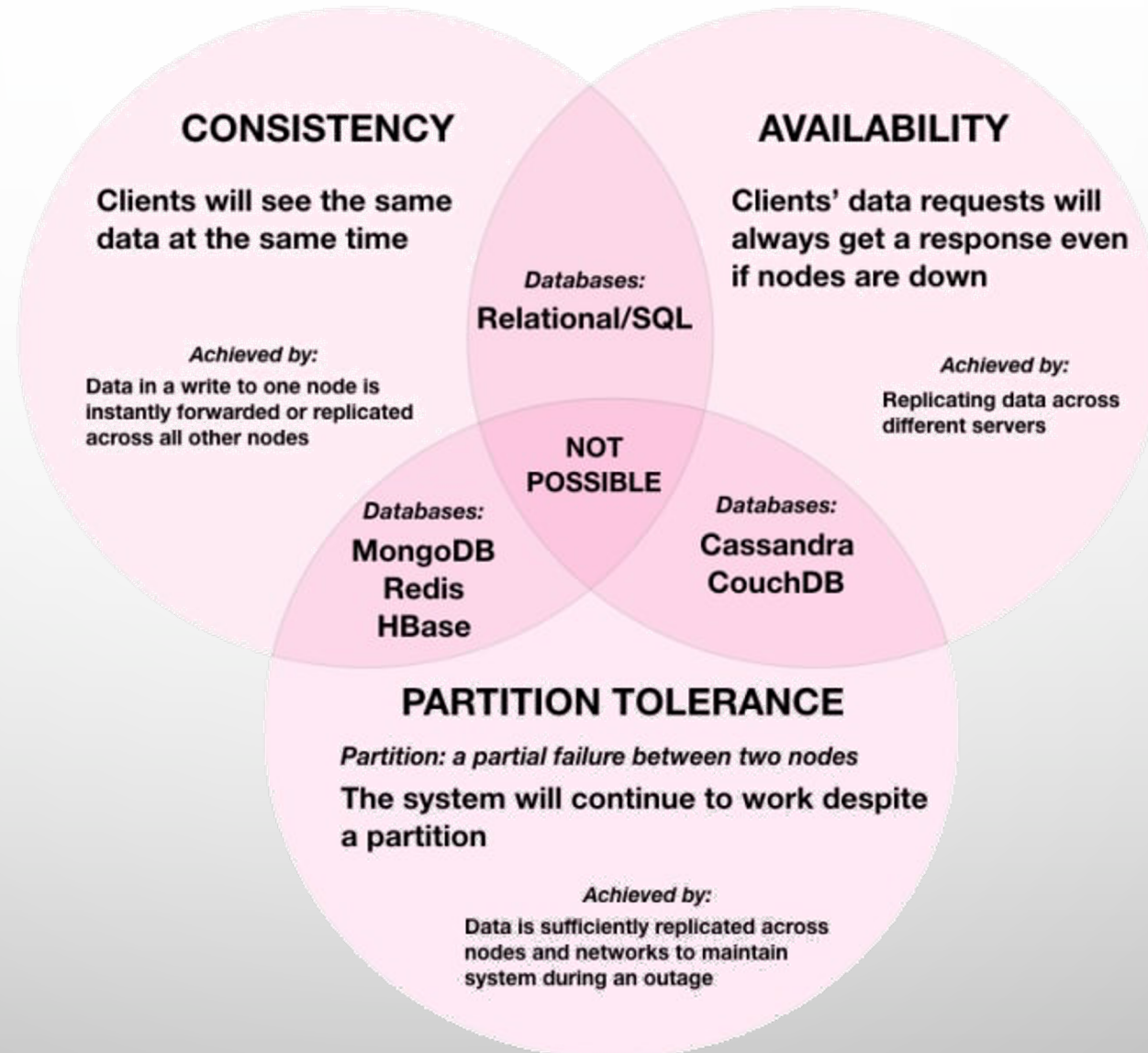
# WHY SERVERLESS





# DISTRIBUTED DATABASES

- CAP THEOREM



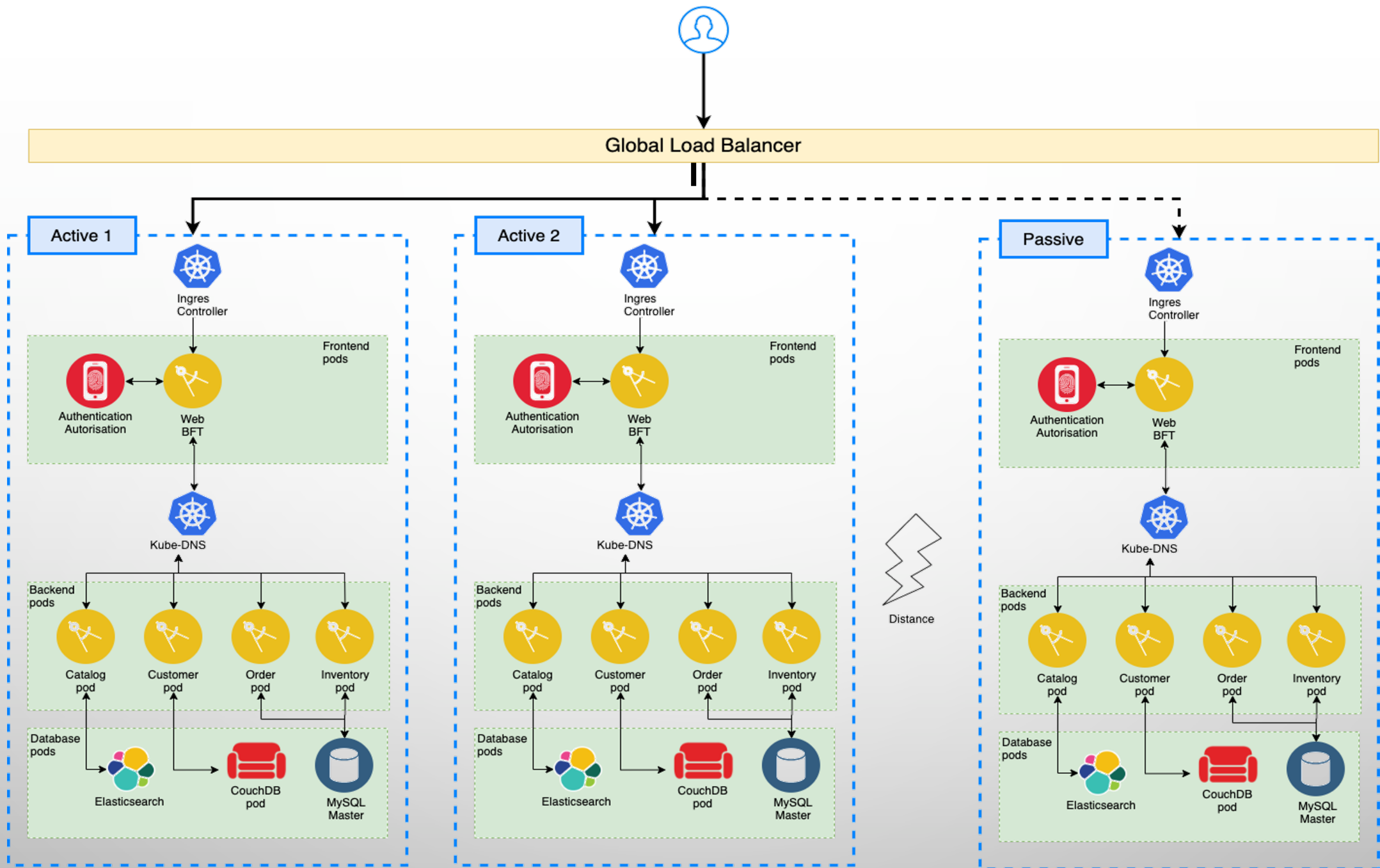
WHAT TO DO ?

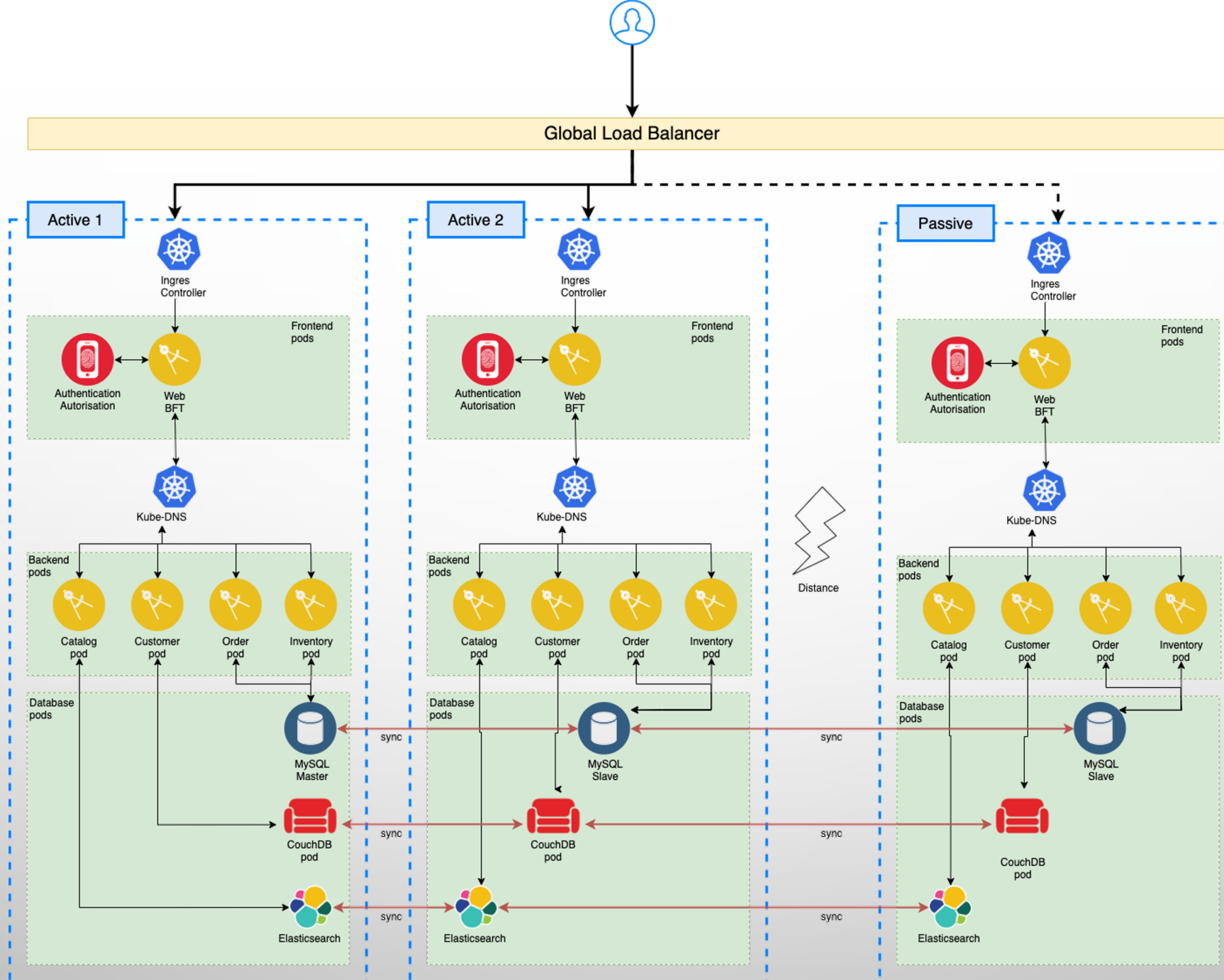
# WORK AS ARCHITECT

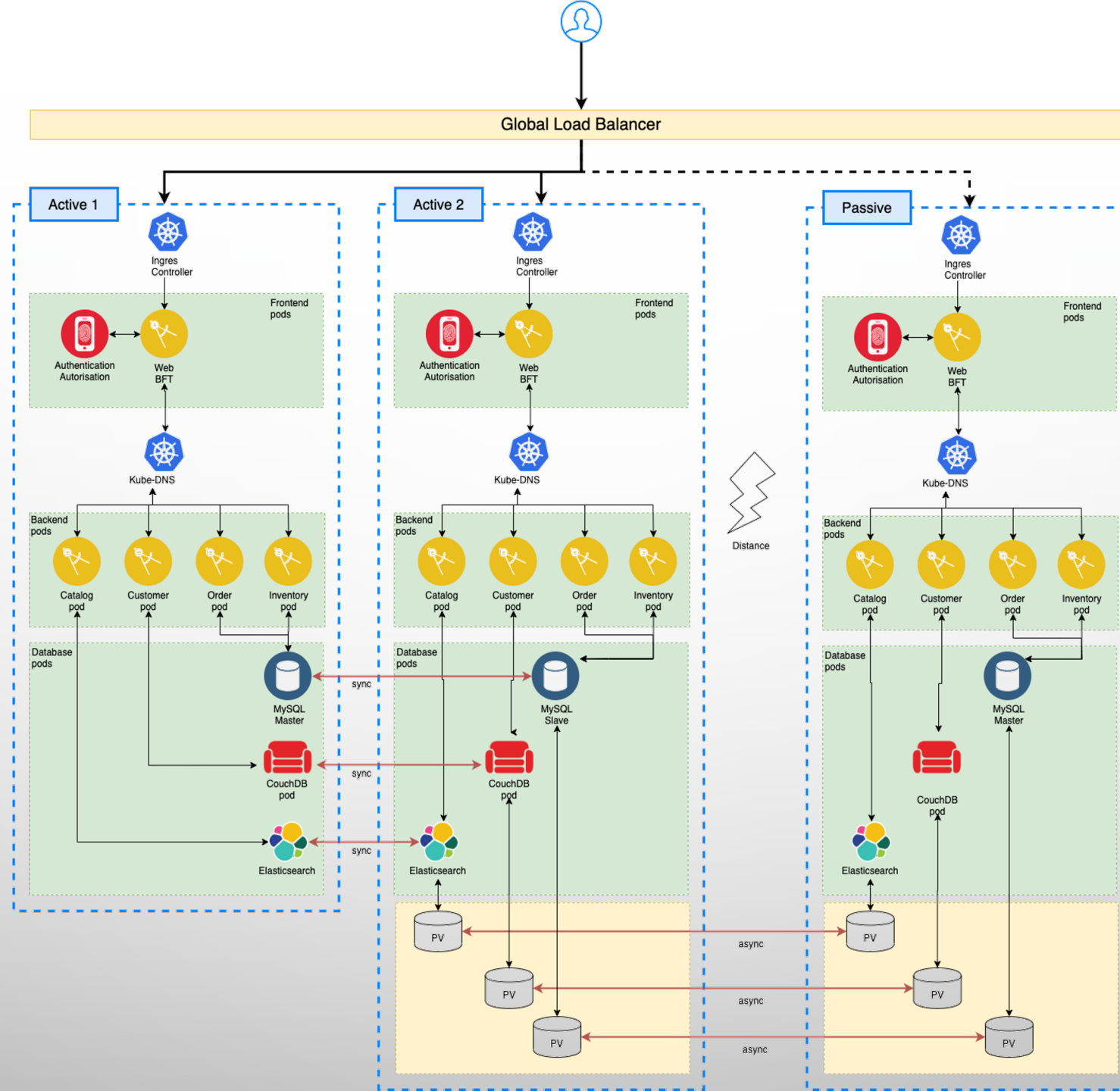
- COLLECT REQUIREMENTS
- KNOW THEORY
- KNOW TECHNOLOGIES













HOW TO START ?

# IBM ARCHITECTURE CENTER

[LINK](#)

# WHAT IS NEXT ?

- HIGH AVAILABLE AND HIGH LOAD SYSTEMS. CLOUD COMPUTING
- MICROSERVICE AND CONTAINERS
- CLUSTERS, ORCHESTRATION AND MODERN DEVSECOPS
- BLOCKCHAIN AS DISTRIBUTED SYSTEM
- CLOUD AND FOG COMPUTING





# THANK YOU

- KHAPOV IGOR CTO IBM RUSSIA/CIS
- KHAPOV.IGOR@RU.IBM.COM

