

Apache Hadoop 2.0 (YARN).

**Последние
тенденции в
обработке больших
данных
(BigData)**

Архитектура Hadoop 1.0

Apache Oozie (Workflow)

Hive
DW System

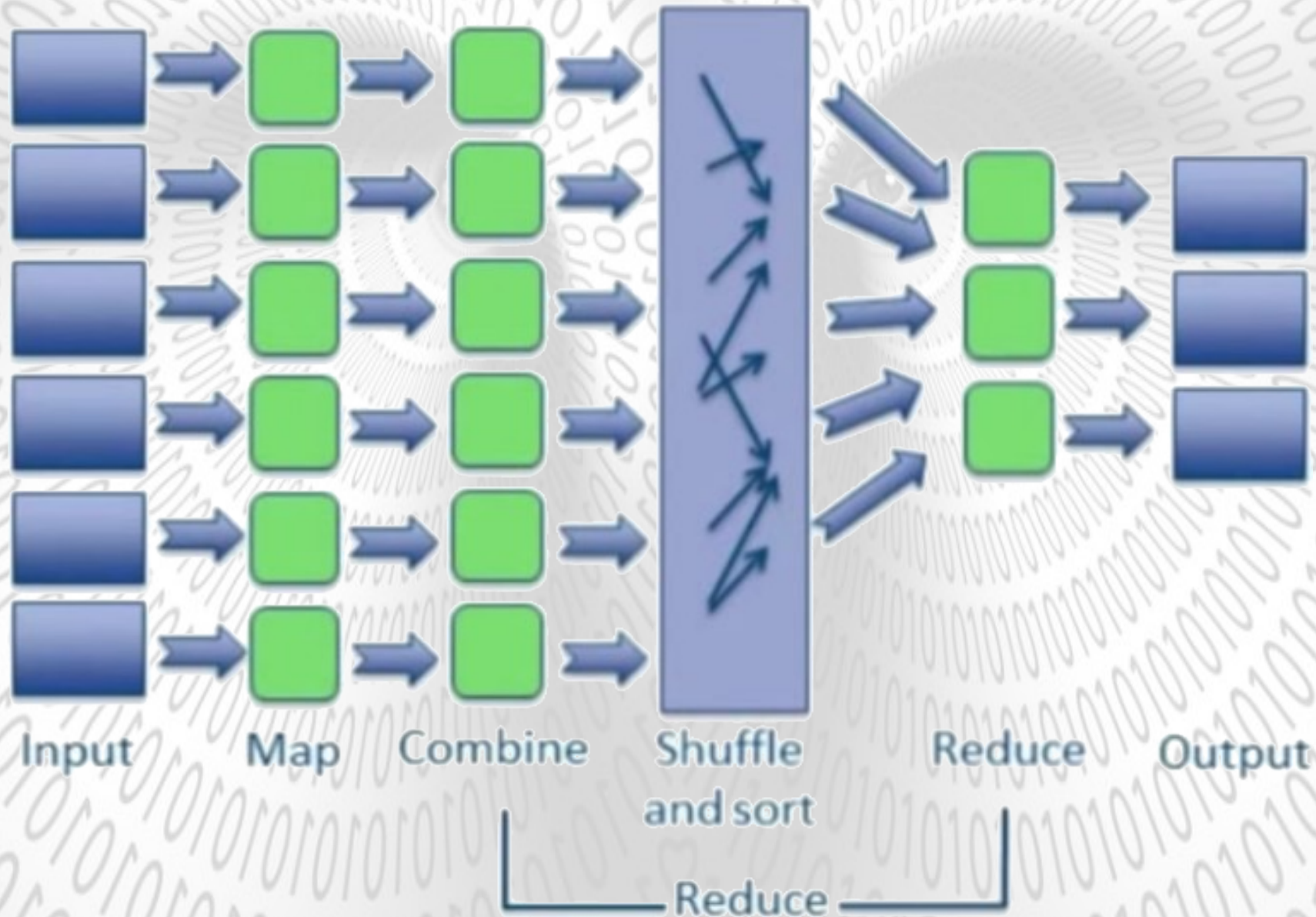
Pig Latin
Data Analysis

HBase

MapReduce Framework

HDFS
(Hadoop Distributed File System)

MapReduce DataFlow



Стек технологий C2R (MRv1)

C2R

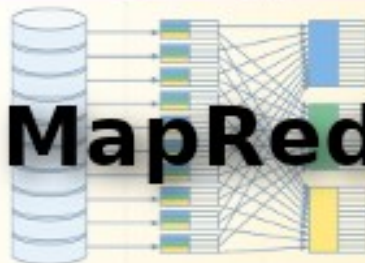


lily

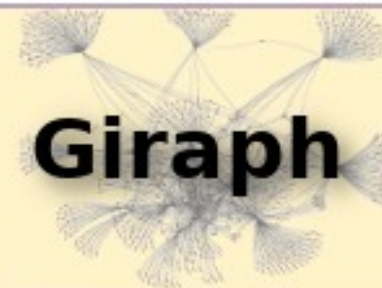
Apache
Solr



HBase



MapReduce



Giraph

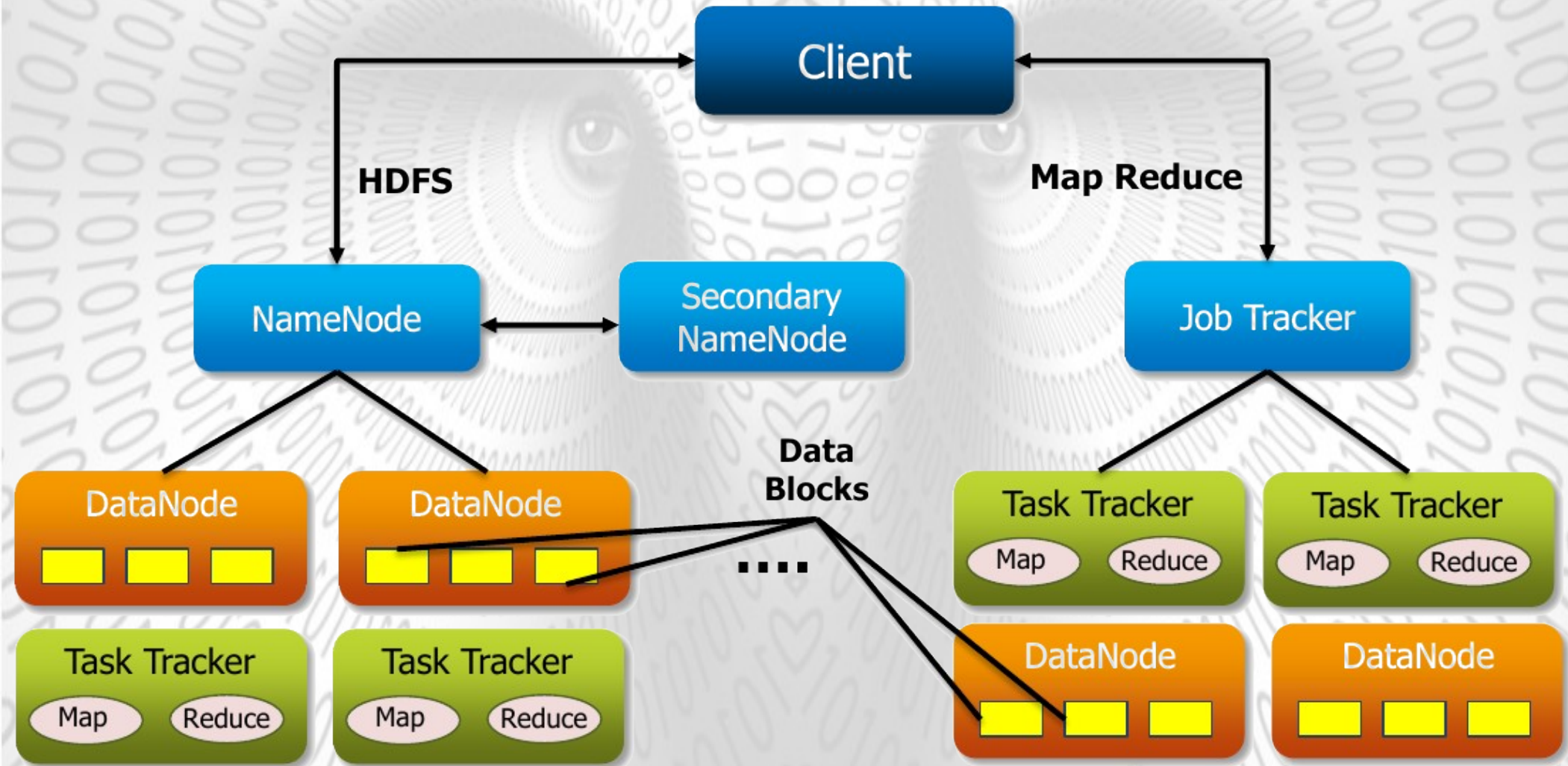


HDFS



C2R

Hadoop 1.0



Ограничения Hadoop 1.0

- ✓ Отсутствие в HDFS горизонтального масштабирования для NameNode
- ✓ Отсутствие режима высоконадежного функционирования (High Availability - HA) для NameNode
- ✓ Повышенная нагрузка на JobTracker
- ✓ Слабая поддержка не-MapReduce приложений
- ✓ Не поддерживается мультиаренда

Проблемы Hadoop 1.0

Проблема

Описание

NameNode — невозможность горизонтального масштабирования

Единственный сервер *NameNode* и файловая система, ограниченная размером оперативной памяти сервера поддерживающего *NameNode*.

NameNode — отсутствует поддержка высоконадежного функционирования (High Availability - HA)

NameNode — точка отказа. В случае его падения требуется ручное восстановления через вторичный (*Secondary*) *NameNode*

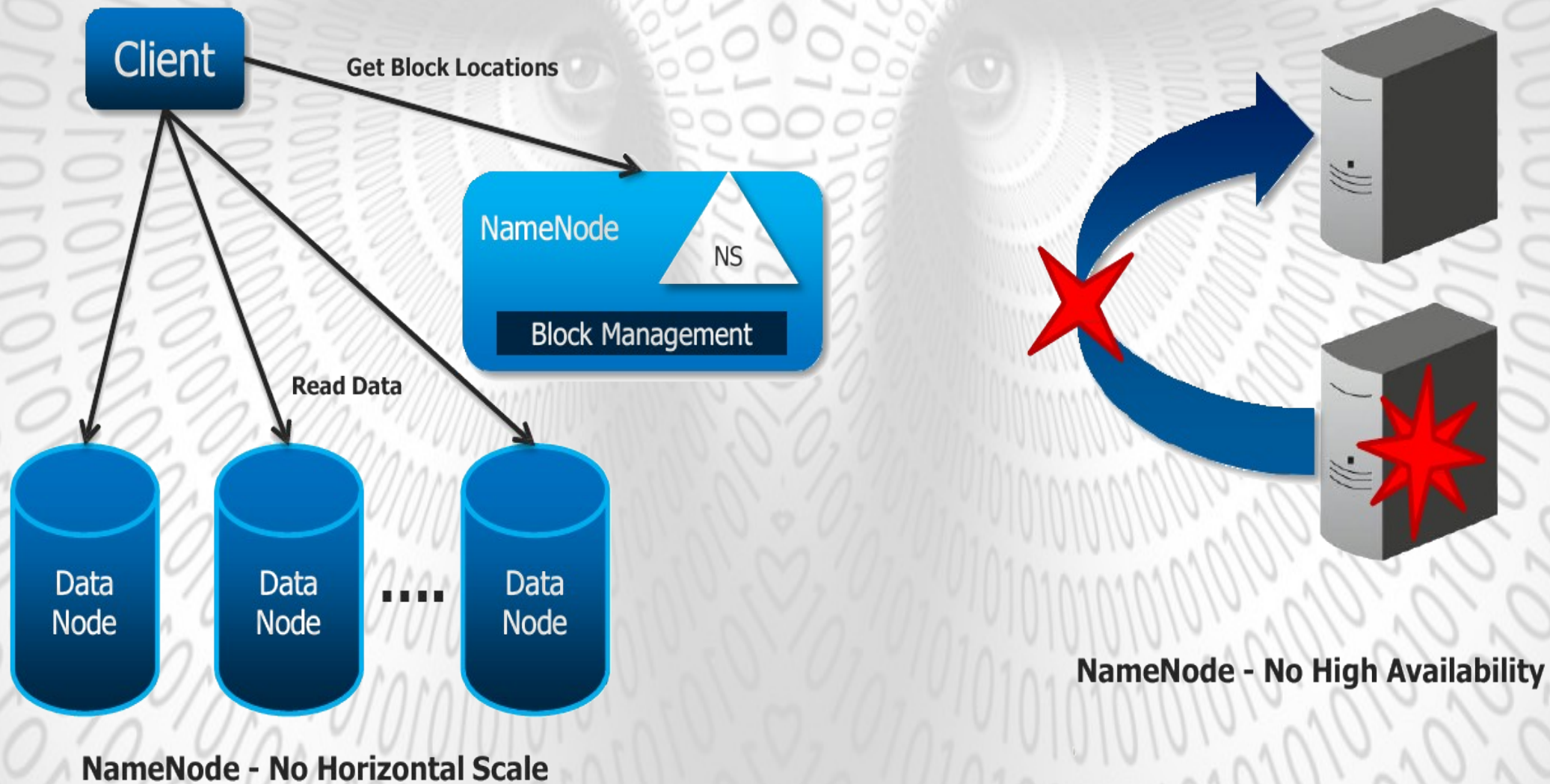
JobTracker — перегрузка функциями.

Большую часть времени *JobTracker* тратит на поддержку жизненного цикла MapReduce-задач.

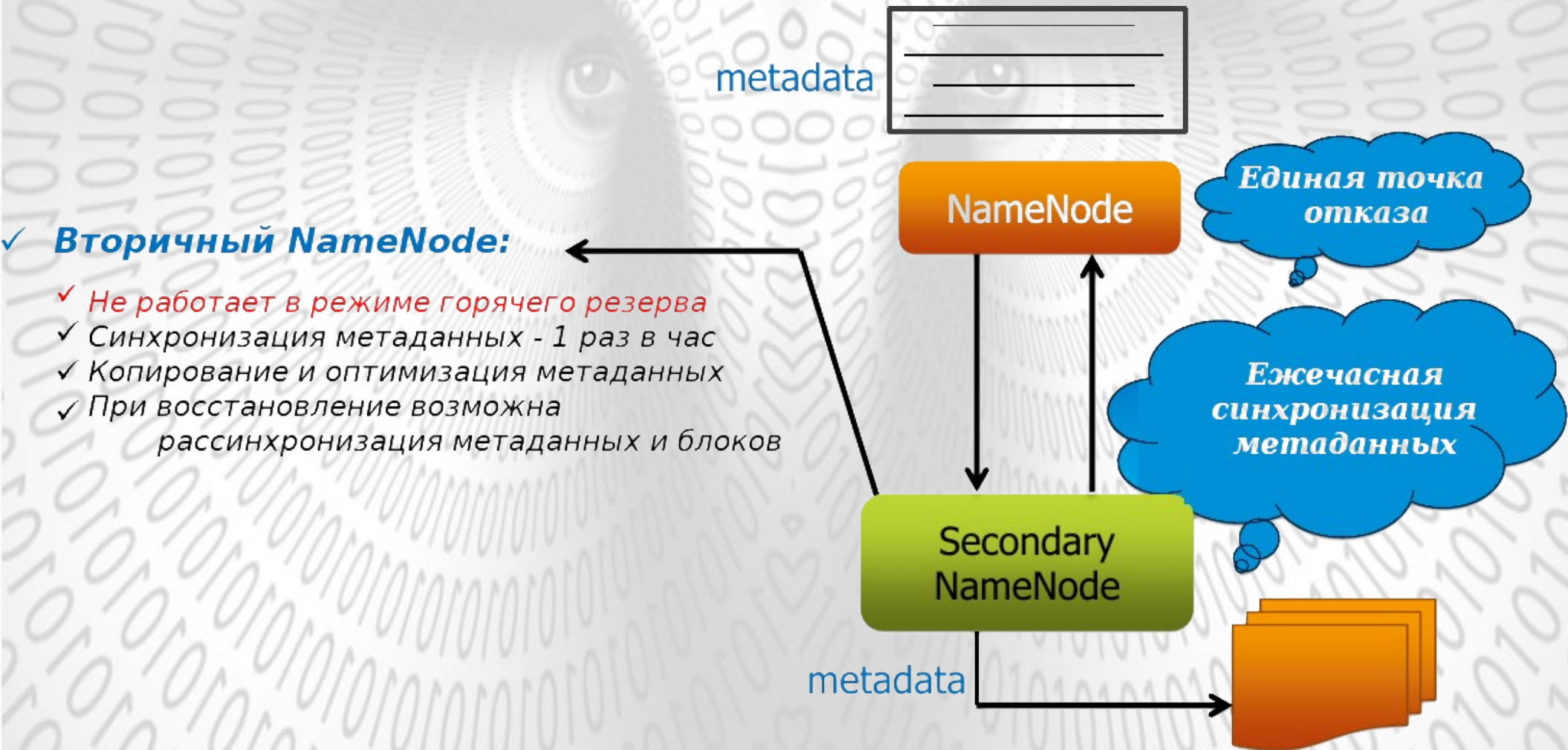
MapReduce version 1.0 (MRv1) — Поддерживает только функционирование *MapReduce java*-задач.

Огромное число хранимых в файловой системе HDFS данных не могут быть использованы другими приложениями.

Отсутствие масштабирования и наличие точки отказа



NameNode — единая точка отказа



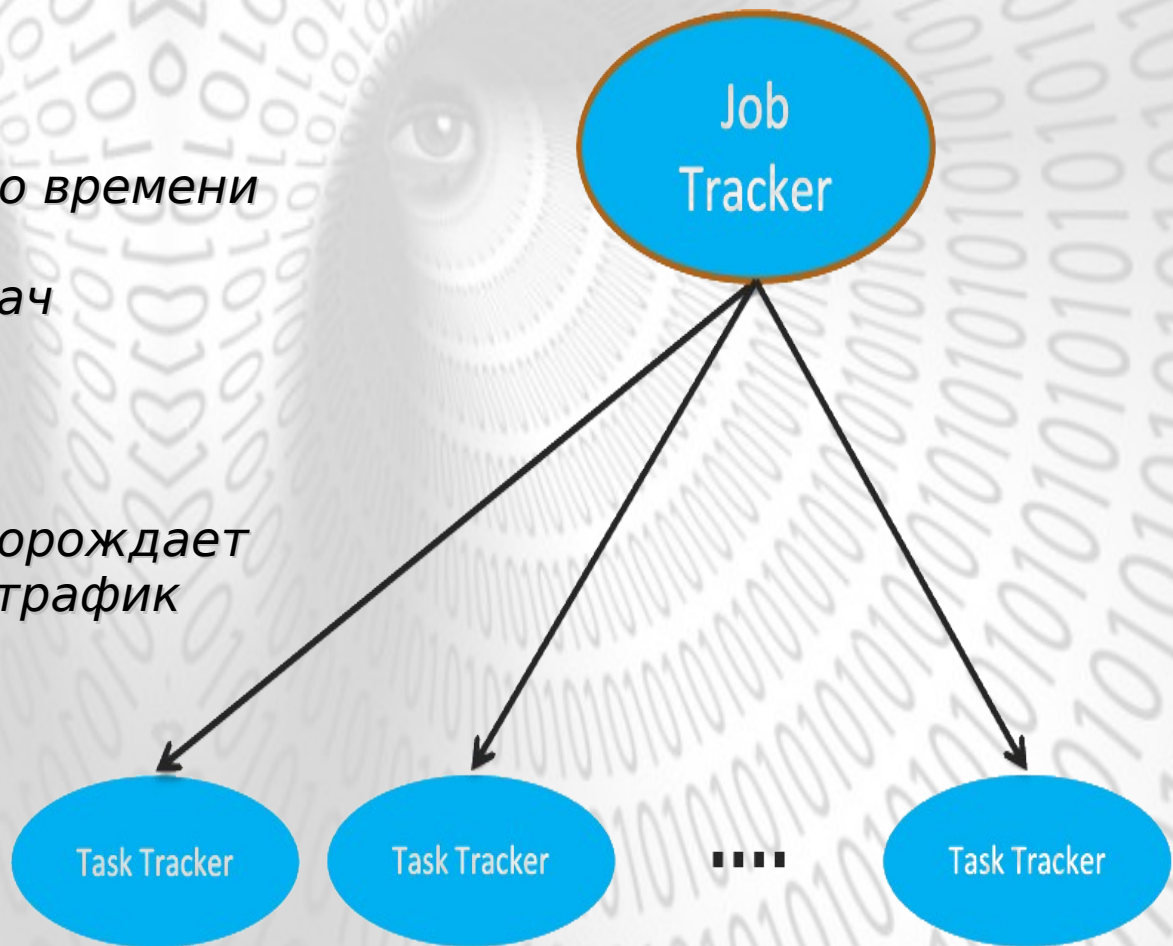
Перегрузка JobTracker

CPU:

- ✓ Большая часть процессорного времени тратится на поддержку жизненного цикла тысяч задач MapReduce

Сеть:

- ✓ Единая точка координации порождает большой централизованный трафик



Непредсказуемость больших кластеров

При достижении кластера размера 4000 серверов:

- ✓ Каскадные сбои:

Выход из строя сервера данных (DataNode) приводит к резкому снижению производительности кластера из-за повышения нагрузки на сеть при репликации недоступных блоков.

- ✓ Мультиаренда:

При увеличении размеров кластера возникает необходимость в поддержке различных независимых файловых систем и различных моделей обработки данных.



Неиспользуемые данные



- ✓ Терабайты и петабайты данных используются только для операций MapReduce

Hadoop 2.0

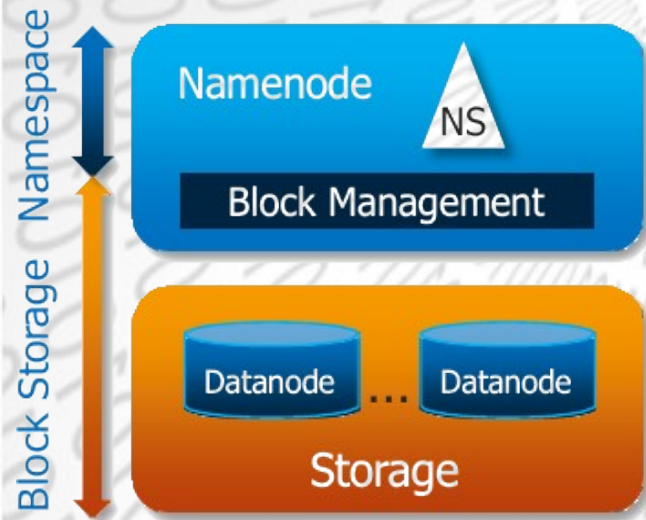
Новый функционал	Hadoop 1.x	Hadoop 2.0
HDFS Federation	Один сервис NameNode и файловая система	Произвольное множество NameNode и файловых систем
NameNode High Availability (HA)	Отсутствует	Поддерживает высоконадежное функционирование
YARN — контроль задач и мультиаренда	JobTracker, TaskTracker	ResourceManager, ApplicationManager's, Nodemanager's, Container's

Другой существенный функционал Hadoop 2.0:

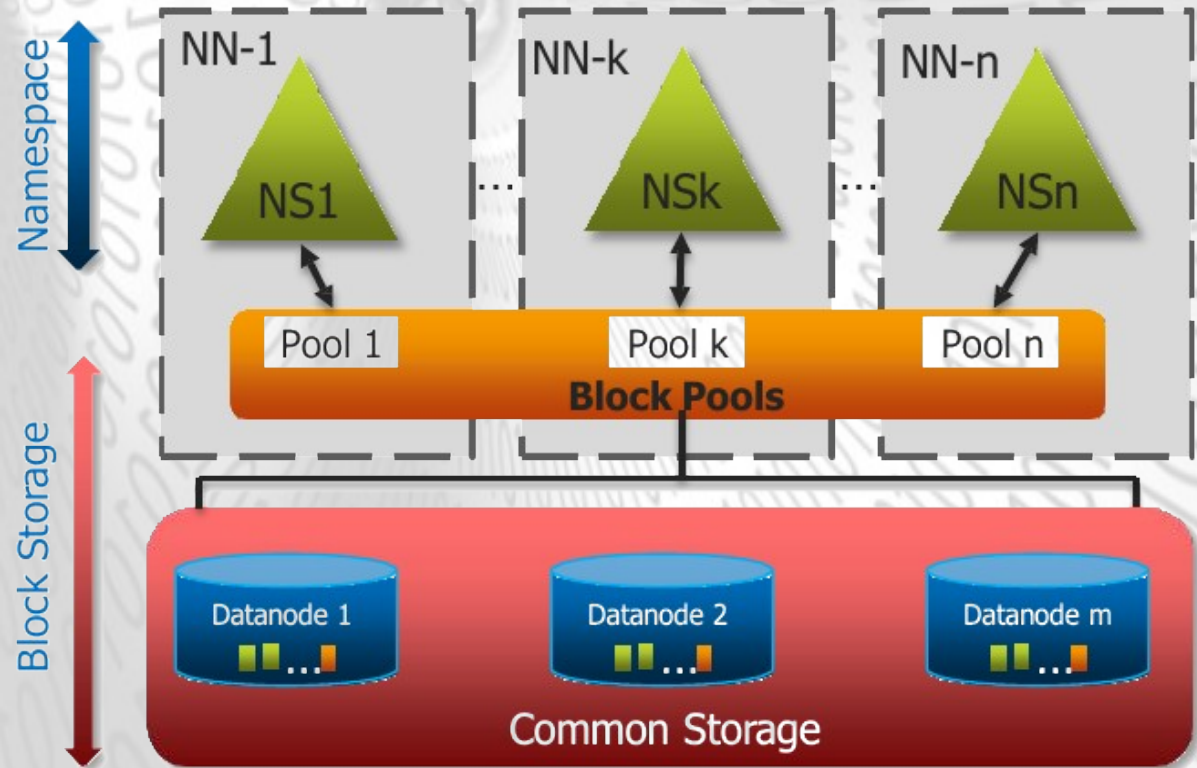
- ✓ HDFS Snapshot's;
- ✓ Бинарная совместимость с задачами MapReduce v1;
- ✓ Глубокая интеграция с проектами Apache Hadoop;
- ✓ Доступ к данным HDFS через NFSv3;
- ✓ Возможность развертывания в среде MS Windows.

Hadoop 2.0: Federation

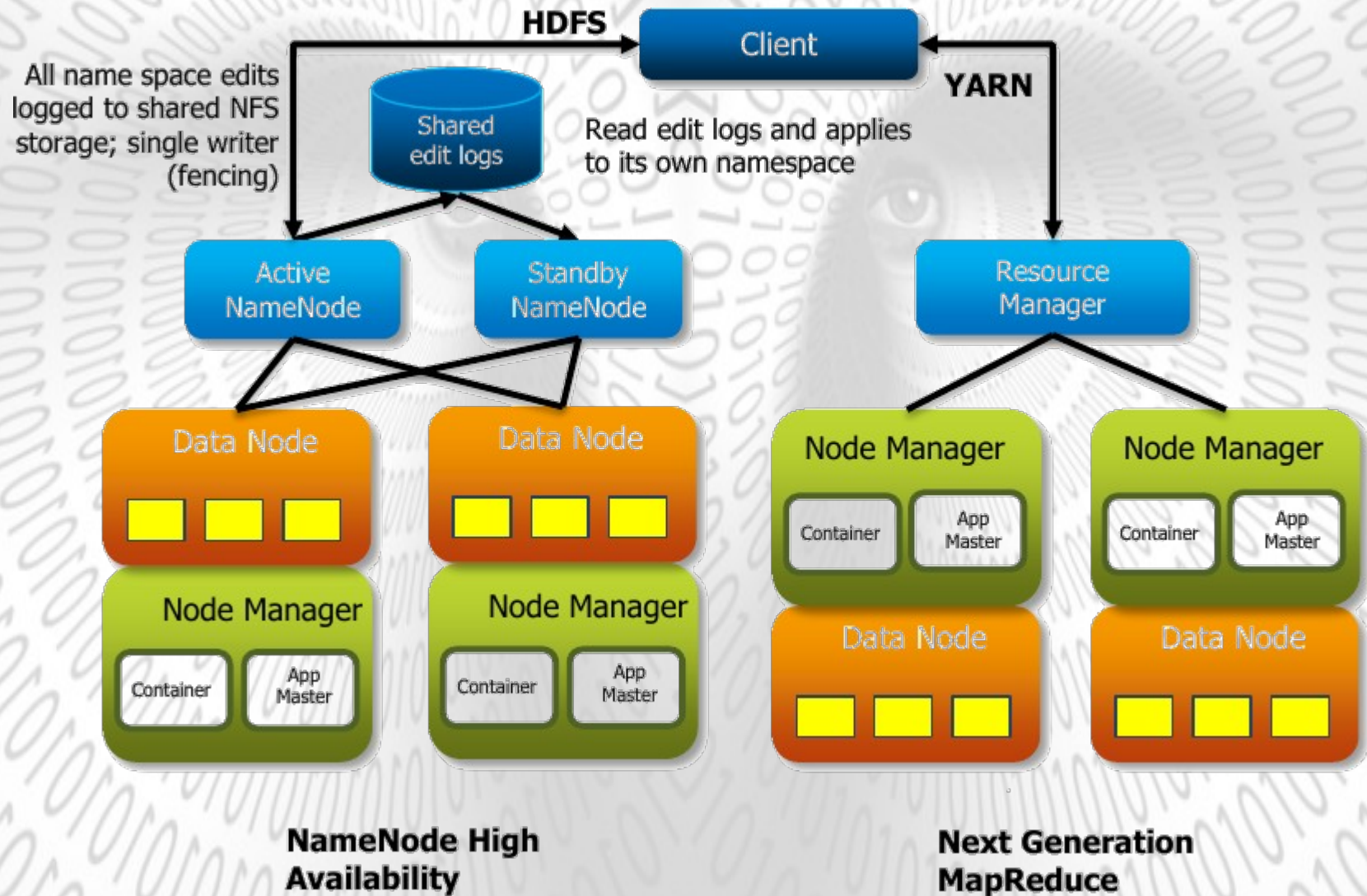
Hadoop 1.0



Hadoop 2.0

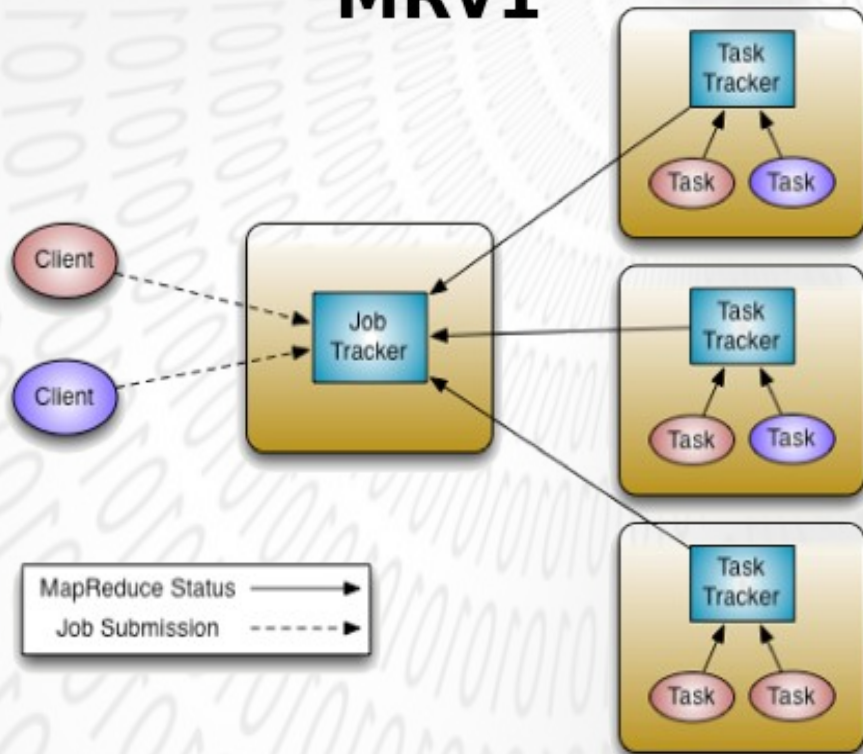


Hadoop 2.0 High Availability

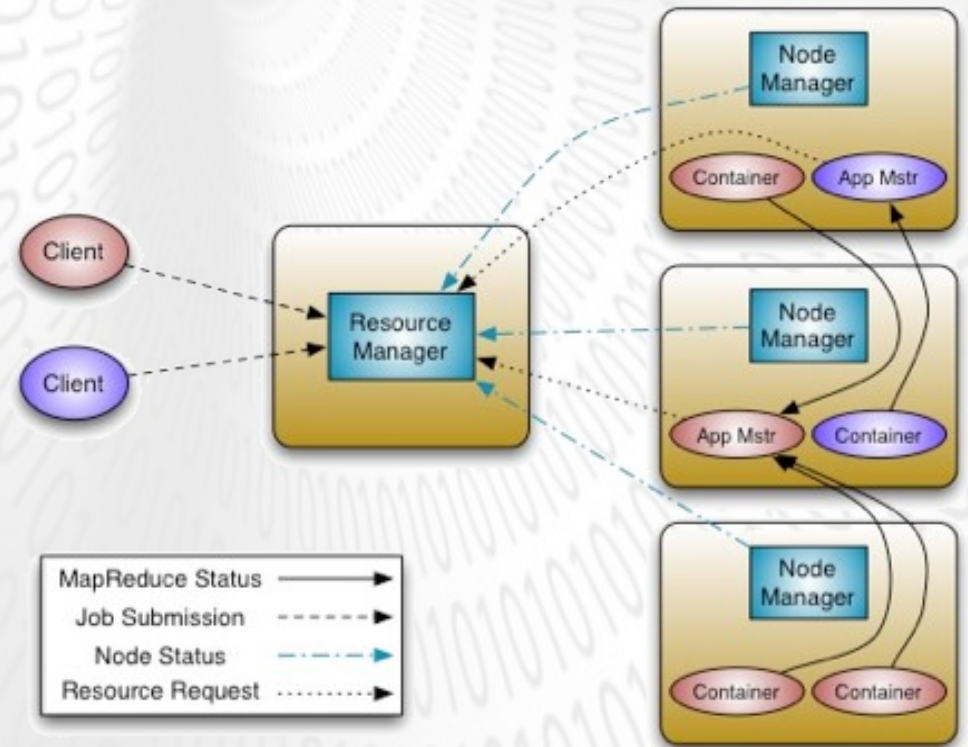


Нadoop 2.0 YARN — КОНТРОЛЬ ЗАДАЧ

Нadoop1, Нadoop2 MRv1



MRv2/YARN



JobTracker vs ResourceManager

Hadoop 1.0 (MRv1)

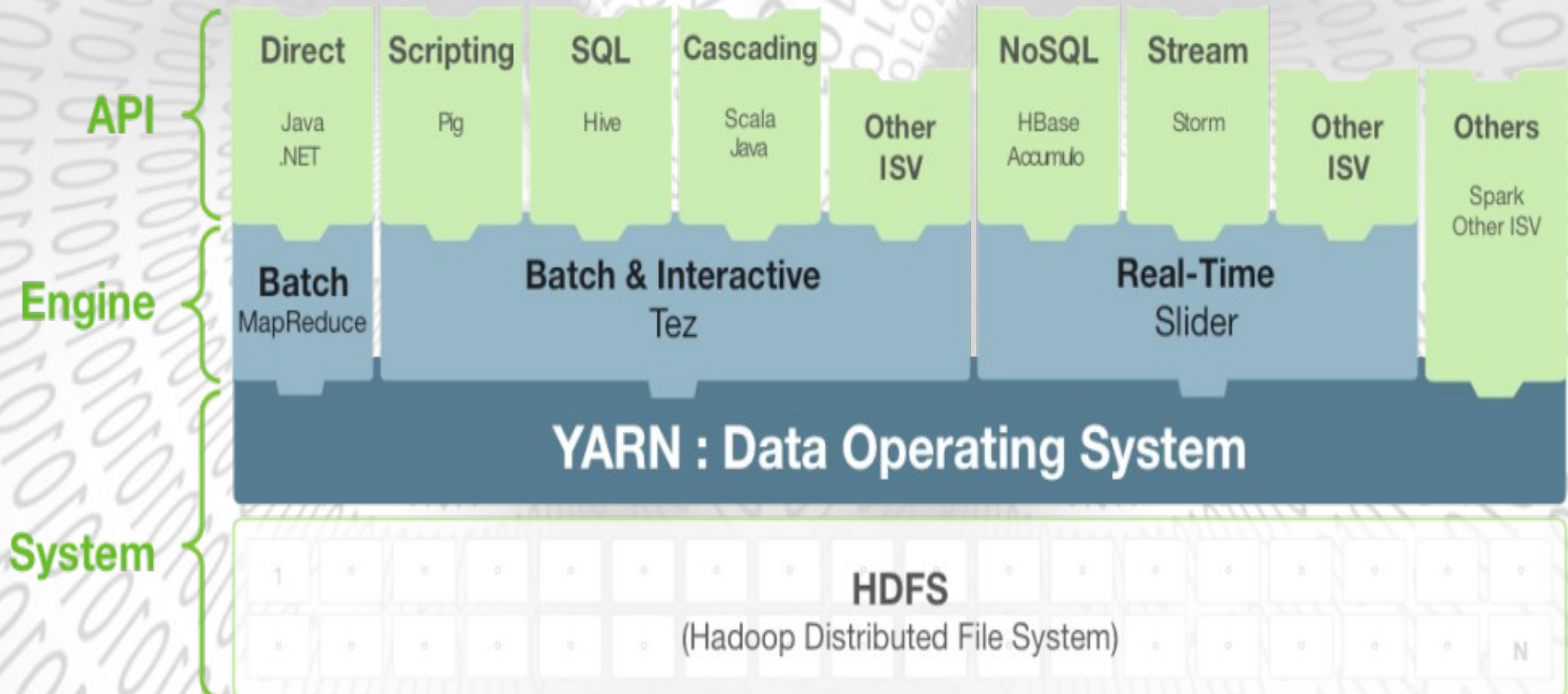
- ✓ Только MapReduce
- ✓ Организация линейных очередей задач
- ✓ Каждая очередь получает % ресурсов кластера
- ✓ FIFO-управление очередью
- ✓ Неуправляемое распределение данных (единый HDFS)

Hadoop 2.0 (YARN)

- ✓ Иерархия очередей для управления ресурсами в пределах каждой организации
- ✓ Эластичность — ресурсы для очередей предоставляются в предсказуемом порядке
- ✓ Гарантированное выделение ресурса: под каждую очередь выделяется доля ресурсов кластера
- ✓ Мультиаренда — установка лимитов для каждого приложения
- ✓ Ресурсно-ориентированное управление — при необходимости приложение может запросить дополнительные ресурсы
- ✓ Безопасность — защита данных от несанкционированного доступа

Hadoop 2.0 Slider

Applications



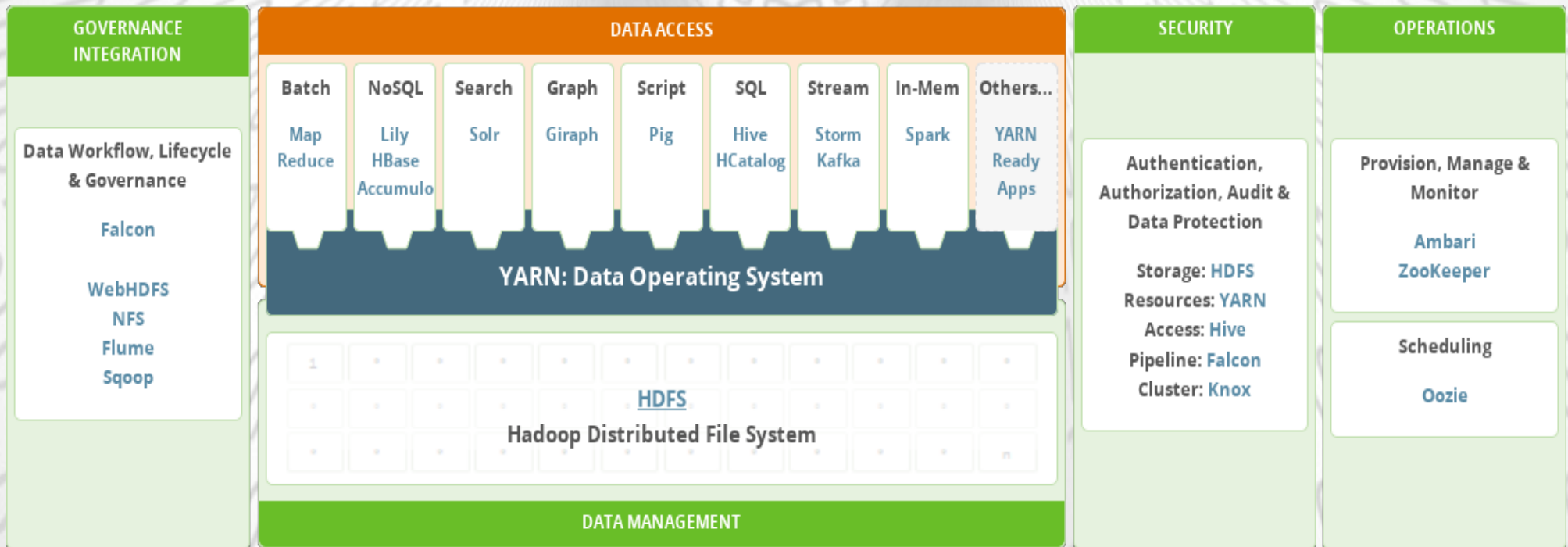
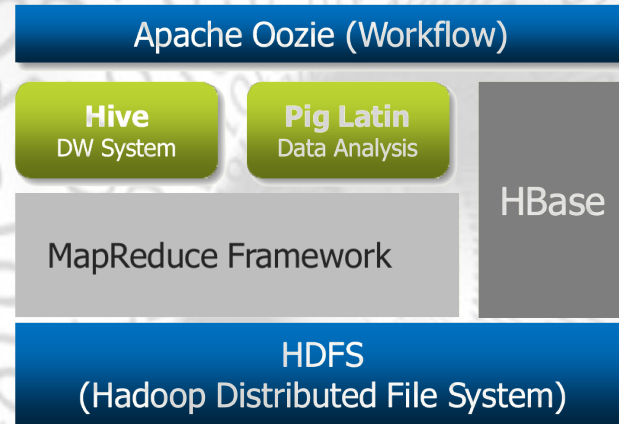
Обработка потоков (Stream)

STORM

- ✓ Надзор реального времени — обработка непрерывных потоков данных
- ✓ Распределенная, Масштабируемая, Гибкая, Отказоустойчивая и Надежная
- ✓ Независима от языков (JSON-based Java, Sparc, Ruby, Python, ...)
- ✓ USECASE: обработка непрерывных потоков данных, распределенный вызов процедур (RPC), непрерывные вычисления

Hadoop 2.0 (YARN) Product Platform

Hadoop 1.0 ->



Наши компетенции

DATA ACCESS

Batch

Map
Reduce

Script

Pig

SQL

Hive/Tez
HCatalog

NoSQL

HBase
Lily

Stream

Storm

Search

Solr

Others

In-Memory
Analytics
ISV Engines

YARN : Data Operating System

HDFS

(Hadoop Distributed File System)

DATA MANAGEMENT

SECURITY

Authentication
Authorization
Accounting
Data Protection

Storage: HDFS
Resources: YARN
Access: Hive, ...
Pipeline: Falcon
Cluster: Knox

OPERATIONS

Provision,
Manage &
Monitor

Ambari
Zookeeper

Scheduling

Oozie

ООО НЕВОД, ГК ИВС
Костарев Алексей
kaf@nevod.ru
kaf@ics.perm.ru



**СПАСИБО
ЗА
ВНИМАНИЕ**