

# Модели распределенной обработки графовых и потоковых данных на кластерах

Константин Абакумов

ИСП РАН

## Эволюция задач интеллектуального труда

1. Механизация ручного труда, переход к интеллектуальному труду
2. Появление компьютеров, упрощение интеллектуального труда за счёт автоматизации трудоёмких алгоритмов
3. Появление больших данных (информационный взрыв)
  - внедрение ИТ во все отрасли науки и промышленности
  - переход к цифровому представлению знаний и событий
  - развитие социальных Интернет-сервисов

Теперь нужно выполнять трудоёмкие алгоритмы на таких объёмах данных, которые человек не в состоянии обработать:

- **New York Stock Exchange** 1.5 ТБ в день, архив 8 ПБ
- **Large Hadron Collider** 41 ТБ в день, архив 15 ПБ
- **Large Synoptic Survey Telescope** 13 ТБ в день
- **Google** Архив 12 ПБ

## Требования

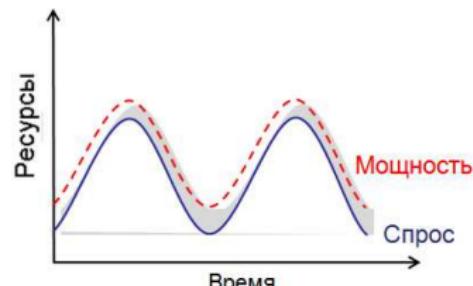
- Обработка больших объёмов данных
  - Перемещение кода к данным
  - Большая совокупная пропускная способность
- Устойчивая работа на ненадёжном оборудовании
  - Репликация данных
  - Восстановление после сбоев
- Хорошая масштабируемость
- Эффективный произвольный доступ к данным

## Преимущества кластерных SPMD-архитектур

- Дешёвое оборудование
- Простота развёртывания
  - Развитый стек свободного ПО
- Простота разработки алгоритмов и приложений
  - Абстракции высокого уровня
  - Интуитивно понятные парадигмы программирования
- Оптимальное использование ресурсов



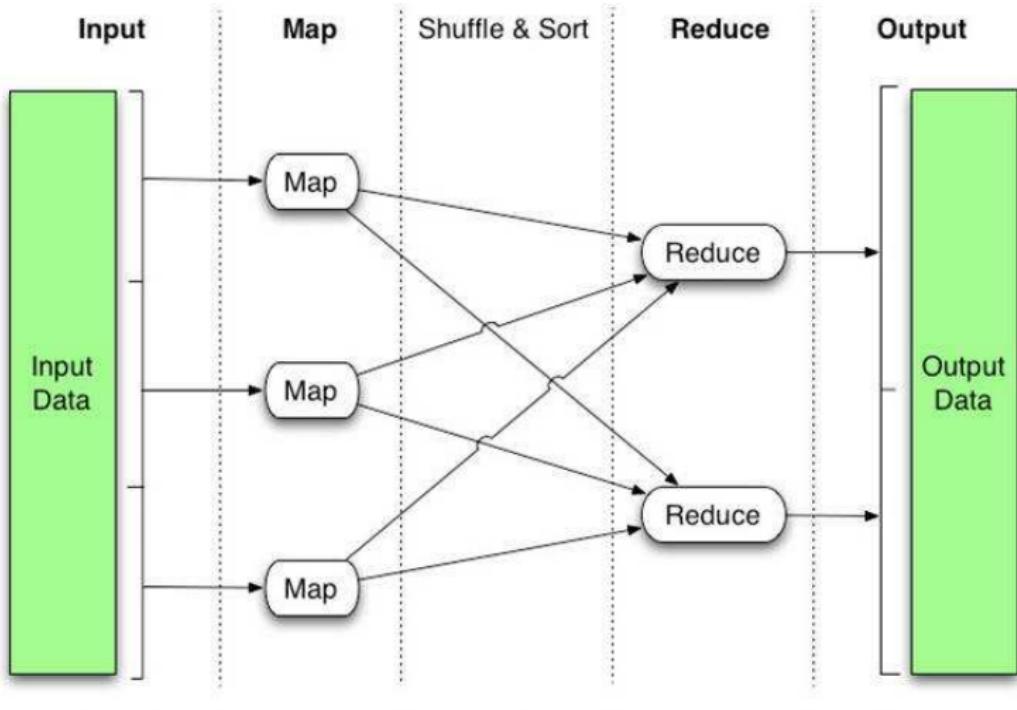
Обычный вычислительный центр



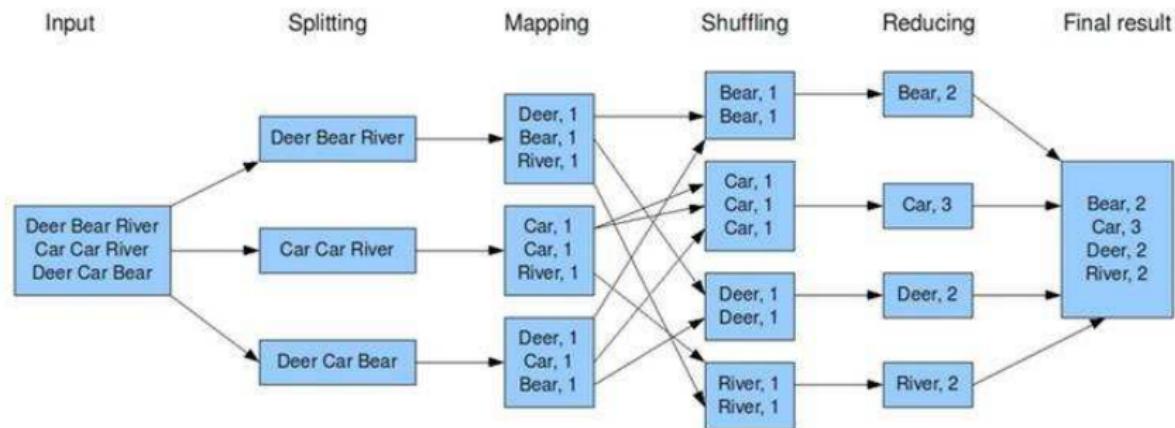
Облачный ЦОД

- Хранение HDFS, HBase, Cassandra, FlockDB
- Обработка MapReduce, Giraph, Spark, Storm, Pig
- Аналитика Mahout, GraphLab
- Инфраструктура ZooKeeper, Mesos, Avro, Chukwa
- Интеграция Sqoop

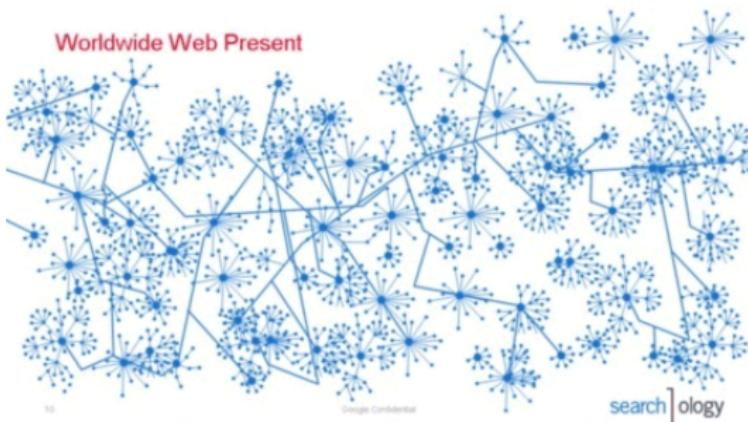
# Парадигма MapReduce



# MapReduce: подсчёт слов в документах

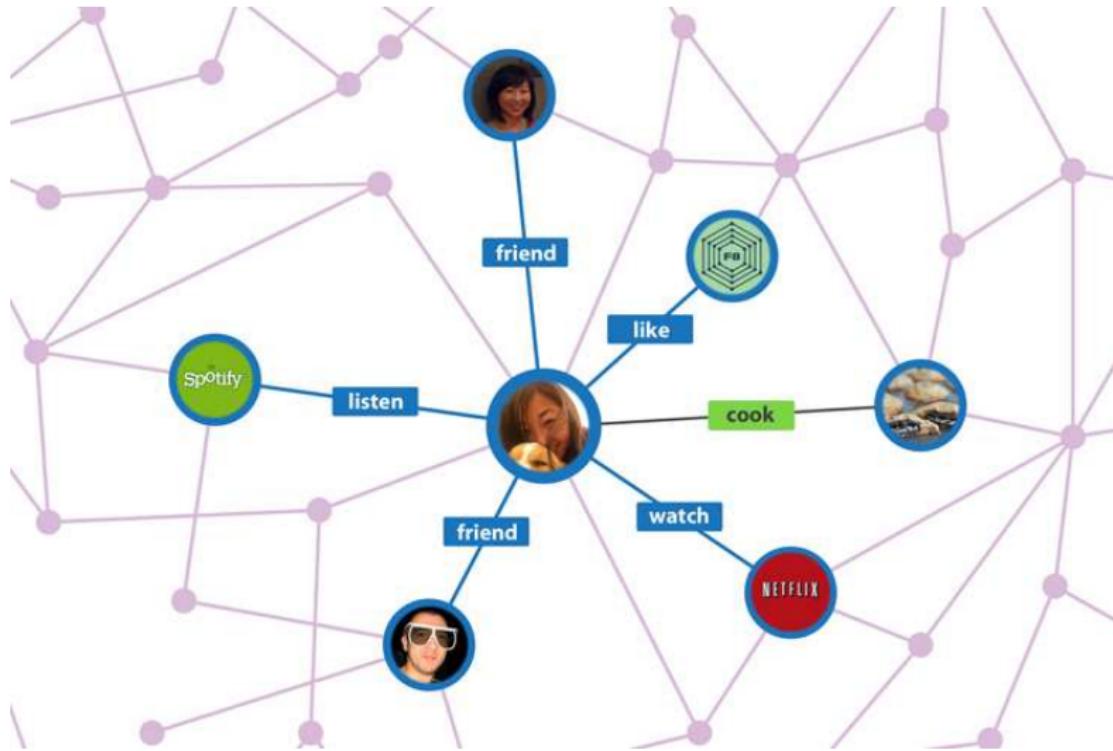


## Примеры масштабных графов



- Веб-графы
- Социальные графы: Facebook, Twitter, вКонтакте
- Графы потока управления
- Графы взаимодействия белков

# Facebook Open Graph



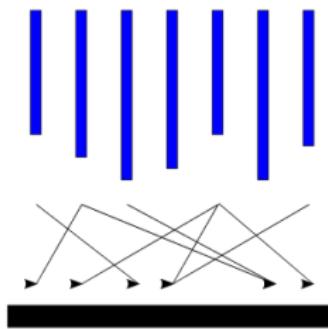
# Google Pregel: обработка графовых данных на кластере

## Модель Bulk Synchronous Parallel (BSP)

**Вход** Список вершин с атрибутами + список рёбер с атрибутами

**Алгоритм** Последовательность супершагов (итераций)

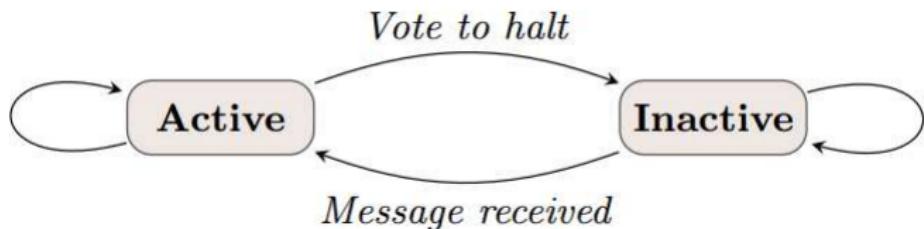
**Выход** Изменённый граф или рассчитанные статистики



**Супершаг** - предопределённая функция, исполняемая всеми вершинами:

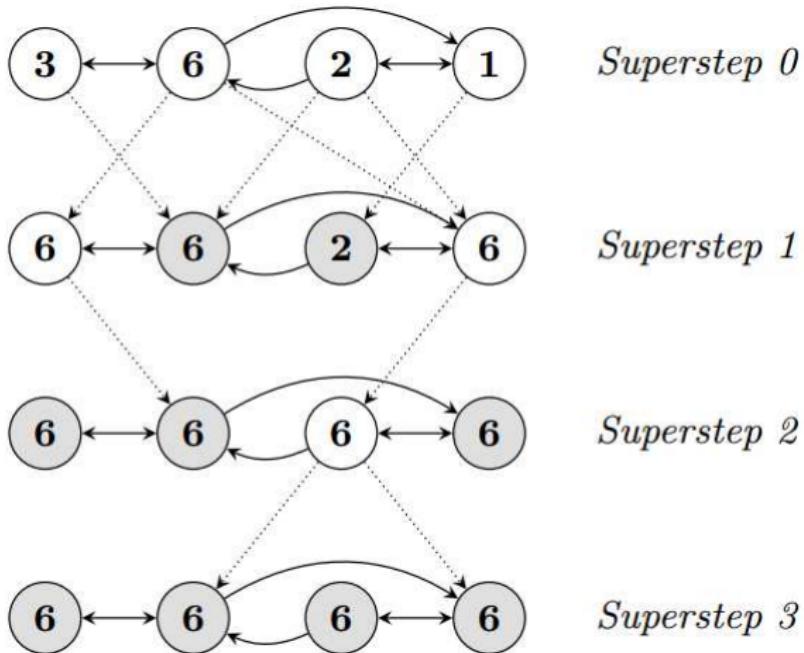
- приём сообщений с предыдущего супершага
- отправка сообщений (обычно соседним вершинам)
- изменение состояния вершины и/или соседей
- изменение структуры графа (добавление/удаление/изменение вершин/ребёр)

## Модель вычислений



- На супершаге 0 все вершины активны
- Только активные вершины участвуют в очередном супершаге
  - Вершина становится неактивной путём голосования за завершение процесса вычислений
  - Реактивация возможна при получении сообщения
- Процесс останавливается единогласным решением всех вершин

# Google Pregel: обработка графовых данных на кластере



## Google Pregel: открытые аналоги

- Apache Giraph
- Bagel
- GoldenOrb
- Apache HAMA
- Phoebus
- ...

# Потоки данных и события



A screenshot of a Twitter search results page. The search term is "phlipgreesen". The results show numerous tweets from a single account, each containing a short message and a small profile picture. The tweets are timestamped and appear to be a continuous stream of updates.

twitter

We may have to recall the new iPhone. This, I did not expect.

— Steve Jobs (@stevejobs) October 20, 2010

 **ceoSteveJobs**  
Steve Jobs

© 2010 Twitter. About Us | Contact | Help | Status | Cookies | API | Business | Help | Jobs | Terms | Privacy



## New Truck Route Opens This Week

A new all-weather trucking route for the CN rail system, one north of the border, has opened, making it easier to move grain and other crops between the U.S. and Canada. The route was a distant dream until a massive snow storm last winter closed the existing route through the Rocky Mountains. Charles O. Pugh of Wisconsin-based Cargill Inc. says the new route will reduce shipping costs by 10 percent. The new route will also cut travel time by two days. The route goes from the U.S. port of Duluth, Minn., to the Canadian port of Vancouver, B.C. It will take about 10 days to travel the 3,000 miles of water route of the Great Lakes and the St. Lawrence Seaway. The route will be used primarily for grain, says C. G. Farwell and D. M. Kennedy, who represent the Canadian port of Vancouver. Farwell says the new route will open up new markets for Canadian grain producers. Farwell and Kennedy say the new route will be more efficient than the existing routes through the Rocky Mountains.

## Области применения

- **Фондовые рынки**

Прогнозирование наиболее выгодных стратегий для трейдеров

- **Телекоммуникационные системы**

Выявление и блокировка подозрительной активности

- **Социальные сервисы**

Аналитика по размещенному контенту. Количество просмотров, статистика по различным критериям

# Поиск популярных хештегов в твитах

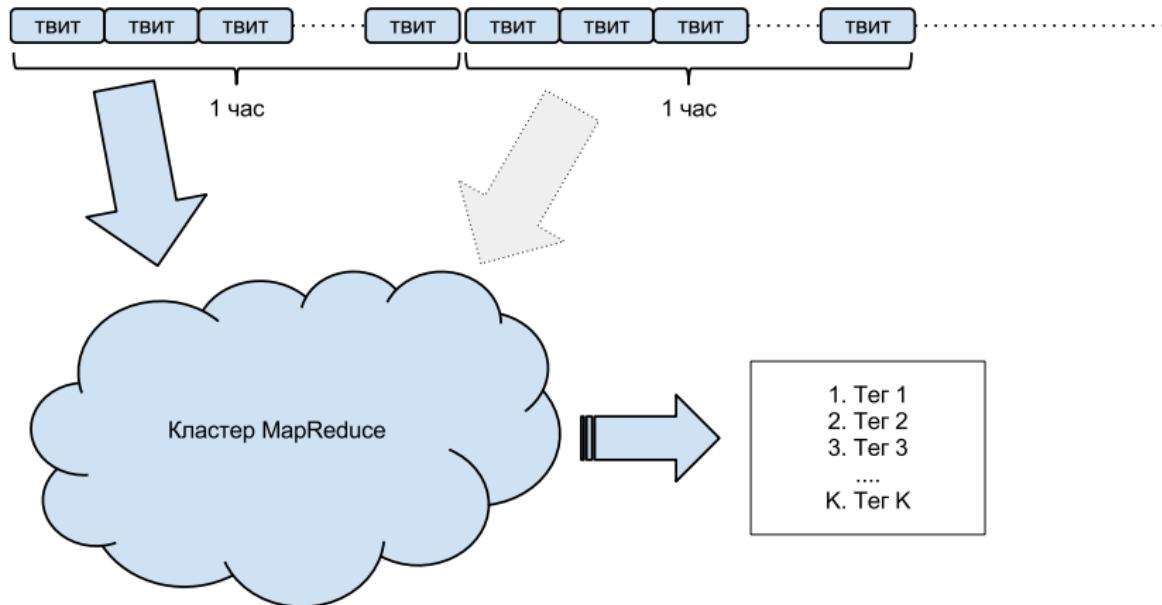
The screenshot shows a search interface for finding popular hashtags in tweets. On the left, there is a list of tweets from various users, each with a profile picture, name, and timestamp. The tweets are as follows:

- Google\_Forum (@GoogleForum)** 10h: Опенсорс платформа Google 2.0, чтобы уметь получать позитивный груз из интернета. Говорите и анализируйте! поддержка для дружеских.
- Я (@yandex\_russia)** 11h: Наша команда выпустила Яндекс.Браузер 2.0 — новый уровень поиска с обновленным интерфейсом магазина, быстрый и удобный поиск у вас на телефоне!
- Sergey\_KH\_09 (@Sergey\_KH\_09)** 10h: Google Glass - Google Glasses, angles and TNG. youtube.blogspot.ru/2012/11/longway... , see more news@news.commeencheque.org/
- Google\_Forum (@GoogleForum)** 10h: Google+ Community Павла Кончакова Ольга Ткачукова нашла своего первенца в рейтинге участников конкурса "Деды для Google" googleplus.google.com
- Sergey\_KH\_09 (@Sergey\_KH\_09)** 17h: Google Glass - Google Glass databases with Google Glass, Google Glass and Google+ Instagram.com/SD4uR9C6n, see more news@news.commeencheque.org/
- Ivan Shishkin (@ivanishik)** 18h: New Google Glass devices like google glass create new patterns, new ways to google.commeencheque.org/
- Google\_Forum (@GoogleForum)** 14h: Видео о грядущем таймкапе премьер новых игр (Интервью с Аланом Сильвером, Новая жизнь - Link Craft & Guardian of Light, игра в стиле StarCraft).
- Я (@yandex\_russia)** 10h: Яндекс Карты и Яндекс Навигатор научили автомобилистов в Москве, Новосибирске и Екатеринбурге ездить по новым линиям общественного транспорта. Яндекс.Карты и Яндекс.Навигатор!

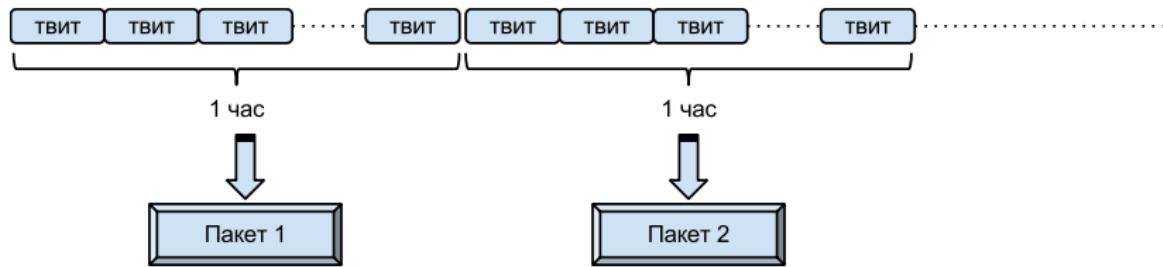
On the right, there is a sidebar titled "Trends - Change" which lists three trending hashtags:

- #LittleMixAlbum
- #QuandJetaisEnfant
- #LaCagasCuando

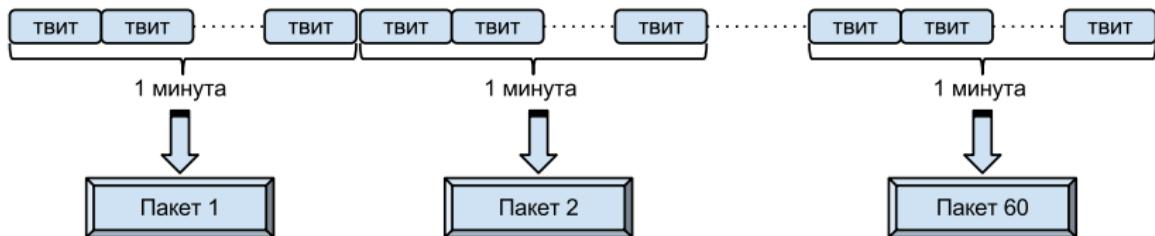
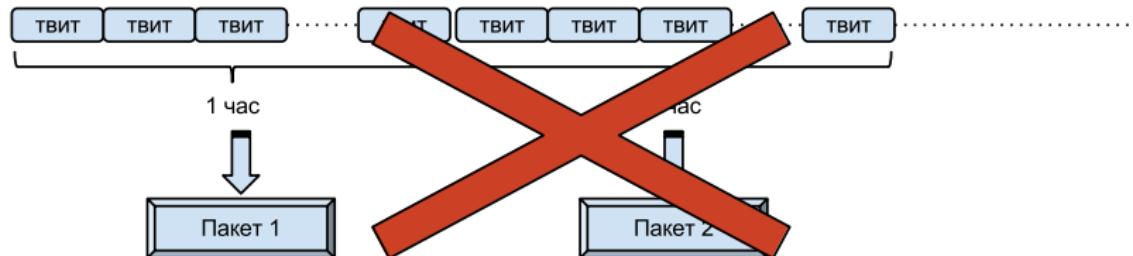
# Распределенная пакетная обработка



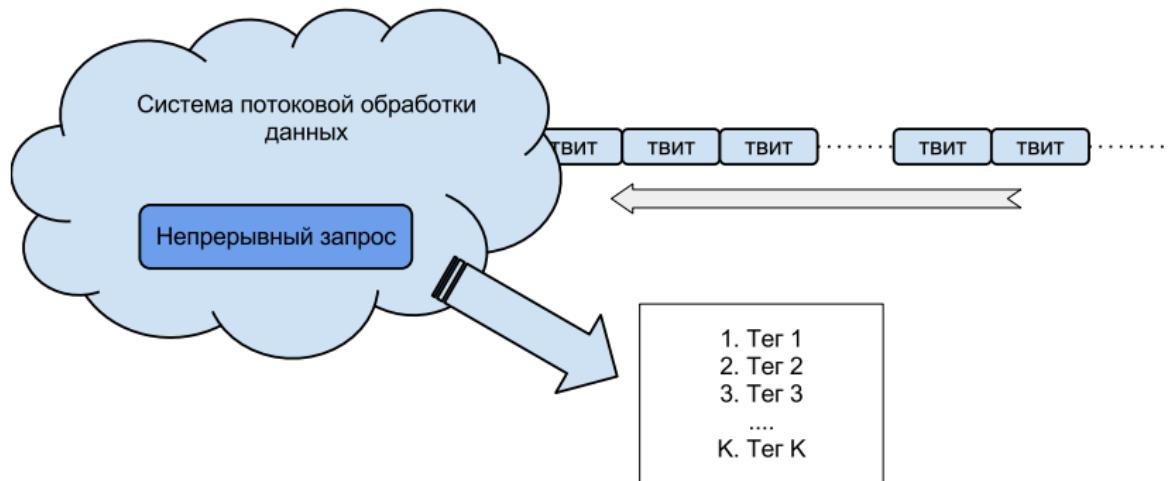
# Распределенная пакетная обработка



# Распределенная пакетная обработка



# Распределенная потоковая обработка



# Системы потоковой обработки

- S4
- Storm
- StreamBase
- PAF (PatternBuilders Analytic Framework)
- Microsoft StreamInsight
- ...

## Storm - Определения

- **Источники**

Отвечают за получение данных из потока

- **Обрабатывающий элемент**

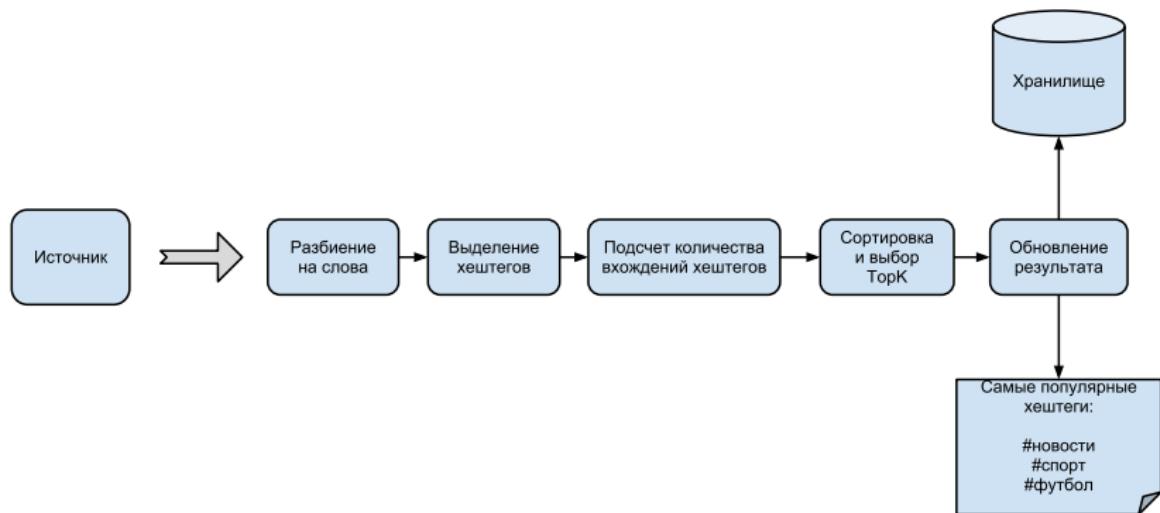
Процесс, выполняющий некоторую обработку событий

- **Топология**

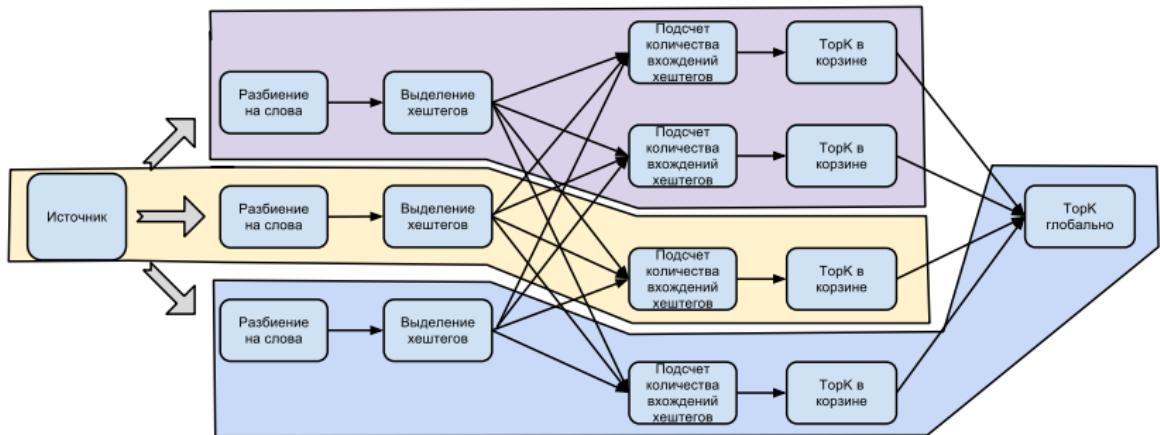
Граф связей между источниками и обрабатывающими элементами.

Полностью описывает решение задачи

## Пример топологии для поиска популярных хештегов



# Выполнение топологии в распределенной среде



Вопросы?

