

Реализация обобщенных операторов реляционной алгебры в AsterixDB

Курсовая работа

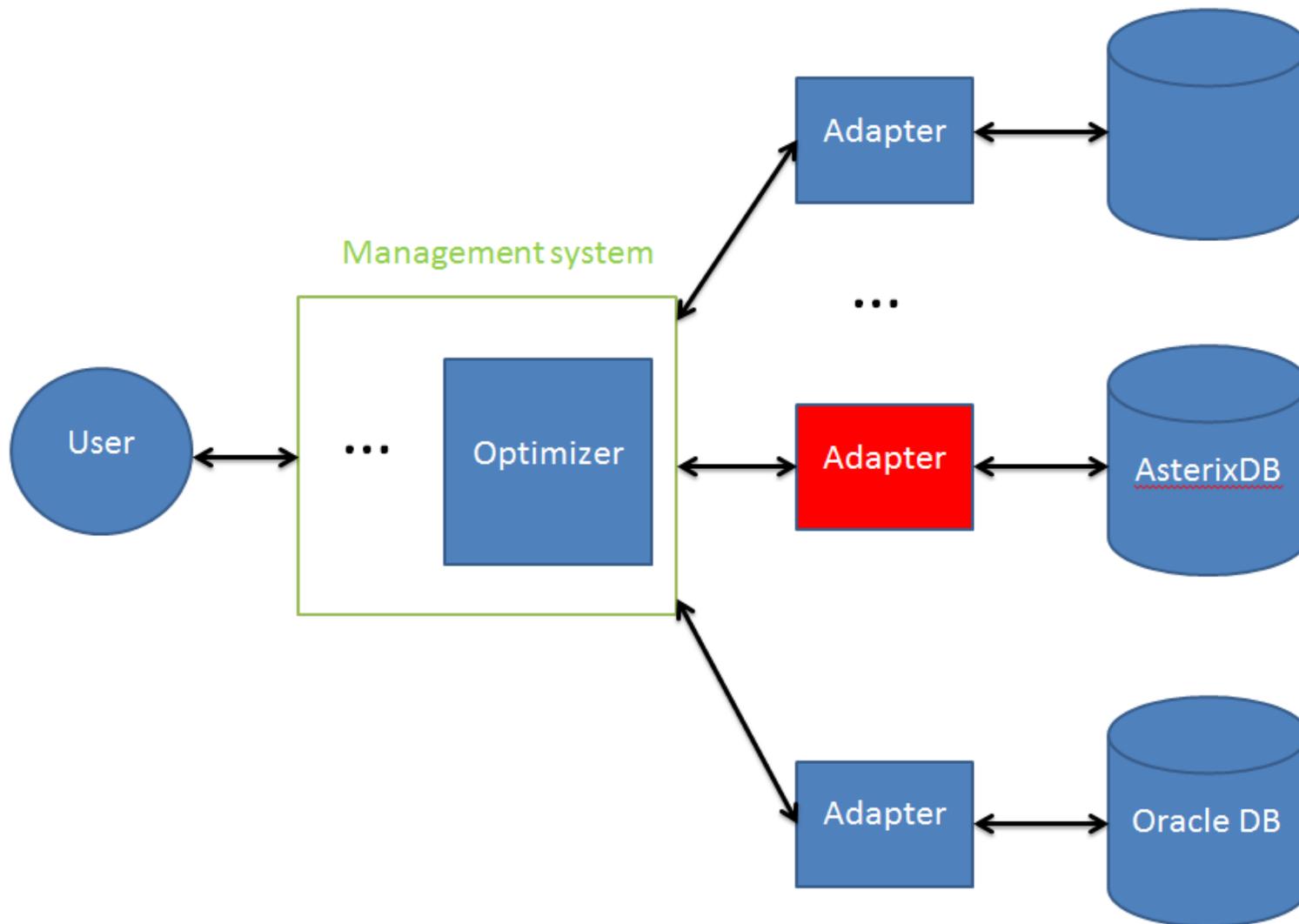
Н. Савельев(444 гр.)

Научный руководитель: Б. А. Новиков

Задачи

- Адаптер
- Трансляция операторов
- Модель стоимостей

Архитектура



AsterixDB BDMS

- Данные в формате JSON
- Мощный язык запросов
- Параллельное исполнение запросов
- Поддержка запросов с нечеткой логикой
- Возможность обработки внешних данных
- Горизонтальное масштабирование

Трансляция

- Конфигурируемые операторы реляционной алгебры:
 - Фильтрация/выборка
 - Соединение
 - Агрегирование
 - ...
- Пример

Пример трансляции

XML query configuration

```
- <query>
- <operation>
  <name>join</name>
- <paramaters>
- <expression>
  <name>equal</name>
- <arg>
  <name>field</name>
  <scope>1</scope>
  <arg>department_id</arg>
</arg>
- <arg>
  <name>field</name>
  <scope>2</scope>
  <arg>id</arg>
</arg>
</expression>
</paramaters>
</operation>
- <arguments>
- <query>
- <operation>
  <name>data</name>
- <parameters>
  <source>Employees</source>
</parameters>
</operation>
</query>
- <query>
- <operation>
  <name>data</name>
- <parameters>
  <source>Departments</source>
</parameters>
</operation>
</query>
</arguments>
</query>
```



AsterixDB Query Language(~XQuery)

```
for $var1 in dataset Employees
  for $var2 in dataset Departments
    where $var1.department_id = $var2.id
return {
  $var1,
  $var2
};
```

Модель стоимостей

- По плану запроса определяет примерное время его исполнения на выбранной платформе
- Используется оптимизатором для определения оптимального плана
- Учитывает особенности исполнения запроса на исполняющей платформе

Модель стоимостей: фильтрация

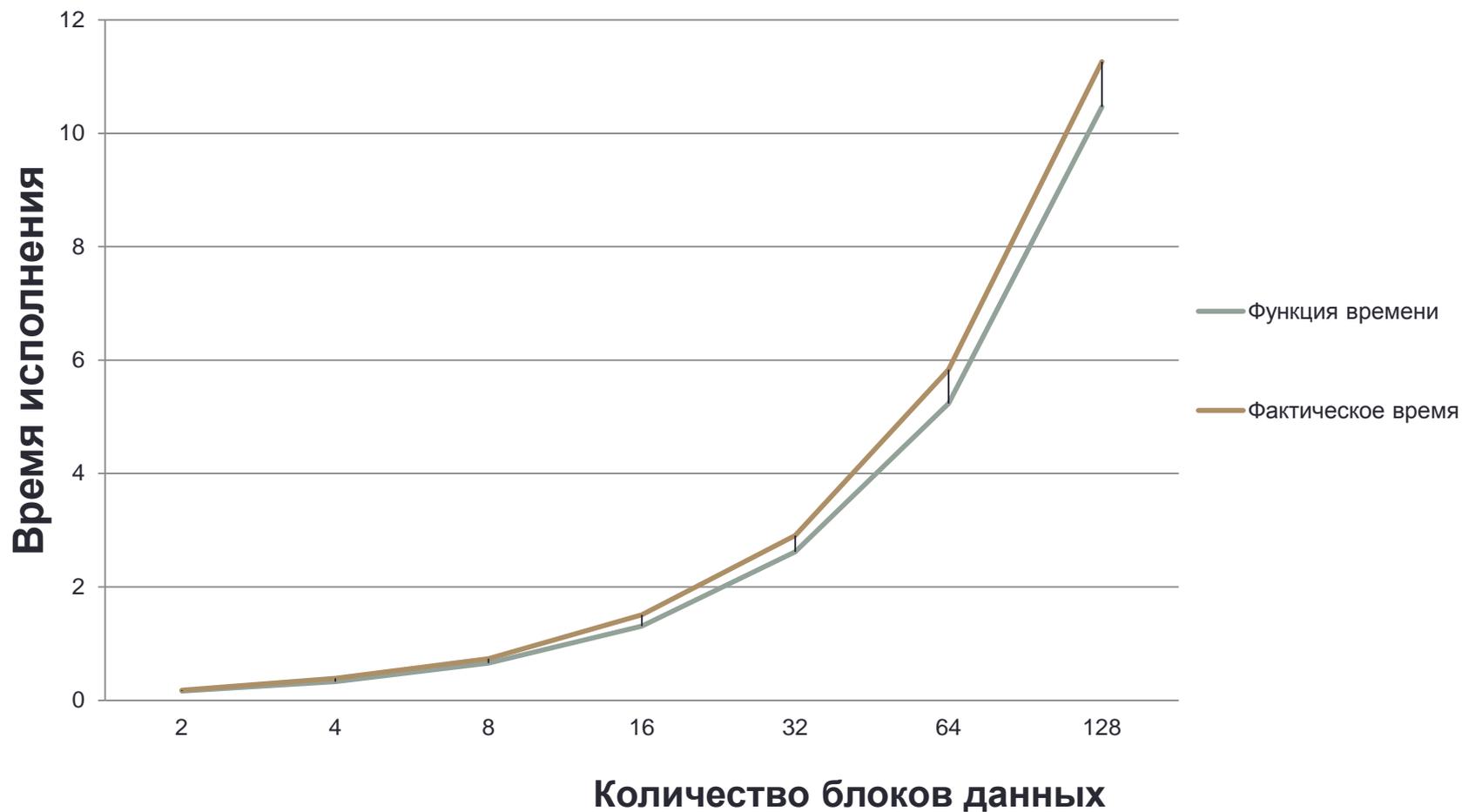
- Вычисление какой-либо функции для каждого элемента отношения. Нет зависимости от других.
- Линейное время исполнения.
- Функция времени исполнения:

$$t_f = \frac{C_b N_b + C_e |N|}{P}$$

- N_b - количество блоков памяти, занимаемых отношением N
- C_b - время считывания одного блока
- C_e - временная константа оператора
- P – число рабочих узлов системы AsterixDB

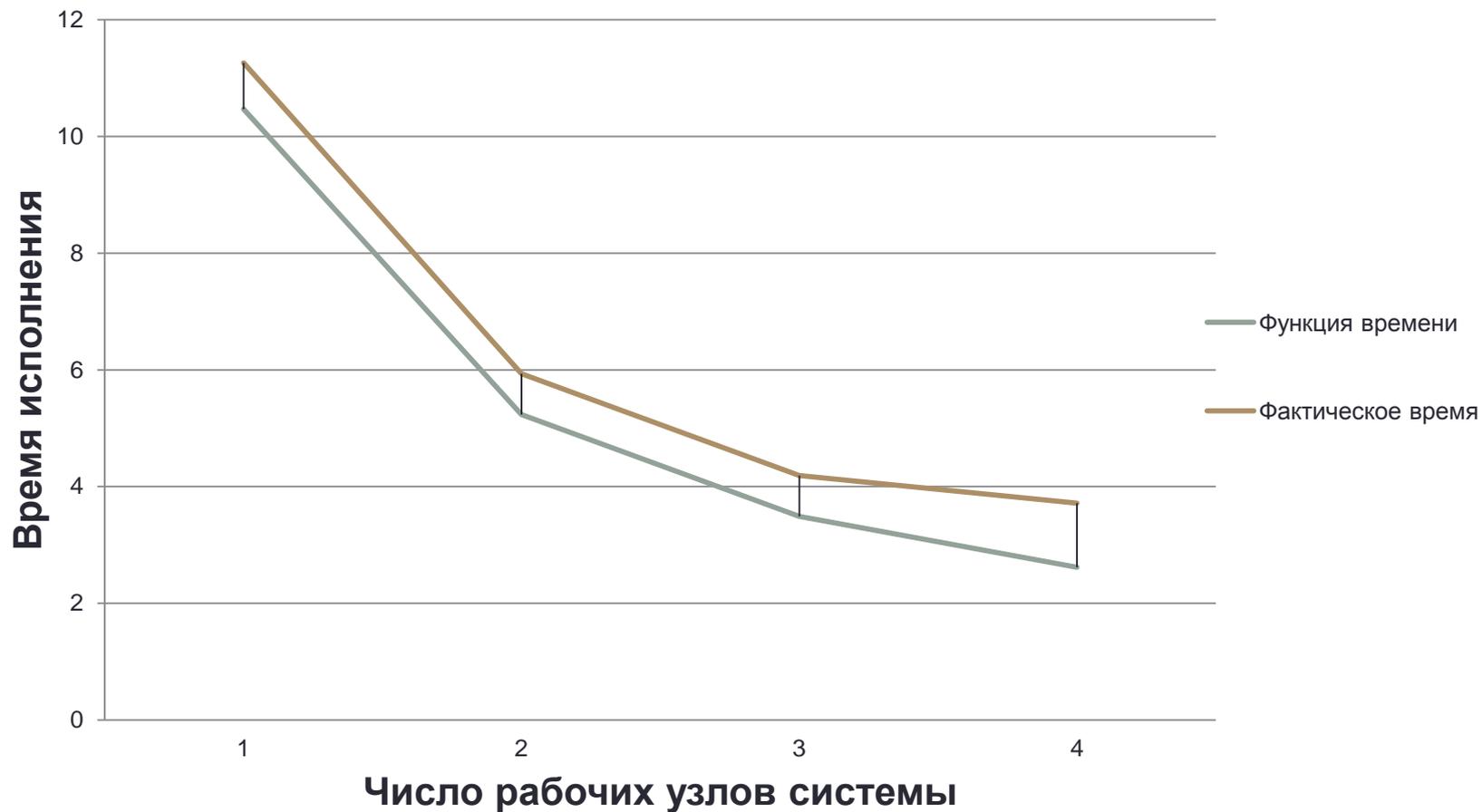
Фильтрация

Сравнение для одного узла



Фильтрация

Сравнение для нескольких узлов



Модель стоимостей: соединение

- Алгоритм параллельного соединения:
 - Разные части отношения хранятся на разных узлах системы
 - Каждому рабочему узлу предоставим все части меньшего отношения
 - На каждом узле выполним соединение меньшего отношения с частью большего соединения(локальное соединение)
 - Объединим результаты со всех узлов
- Алгоритмы локального соединения:
 1. Hash join
 2. Sort-merge join
 3. Nested loop join

Соединение: Hash join

Алгоритм:

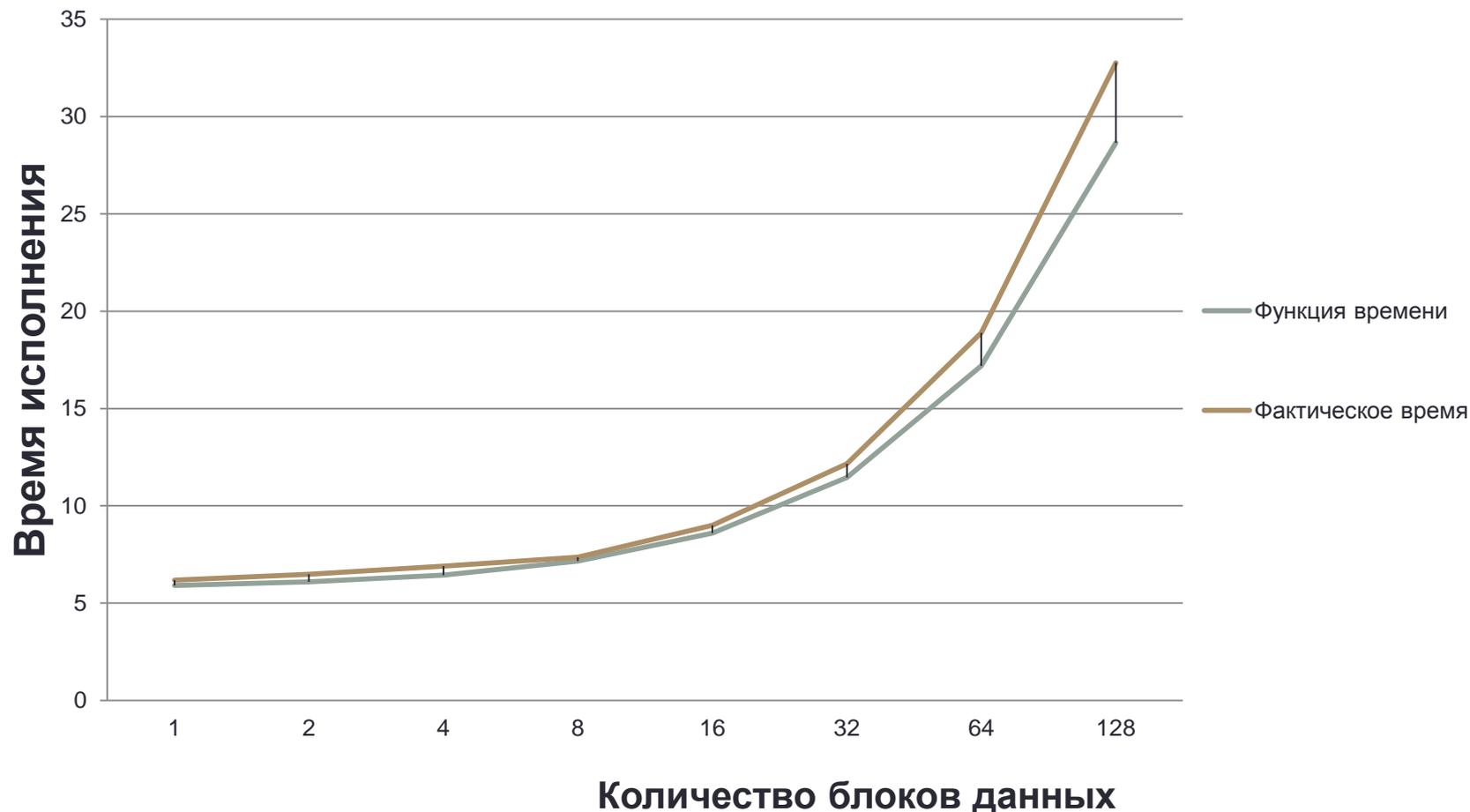
1. Построение хэш-таблицы по отношению с меньшим объемом
2. Проход по второму отношению и с обращением к хэш-таблице

$$t_{hj} = \frac{C_b}{P} ((P - 1)N_b + M_b) + C_e \left(|N| + \frac{|M|}{P} \right)$$

- N_b - количество блоков памяти, занимаемых отношением N
- M_b - количество блоков памяти, занимаемых отношением M
- C_b - время считывания одного блока
- C_e - временная константа оператора
- P - число рабочих узлов системы AsterixDB

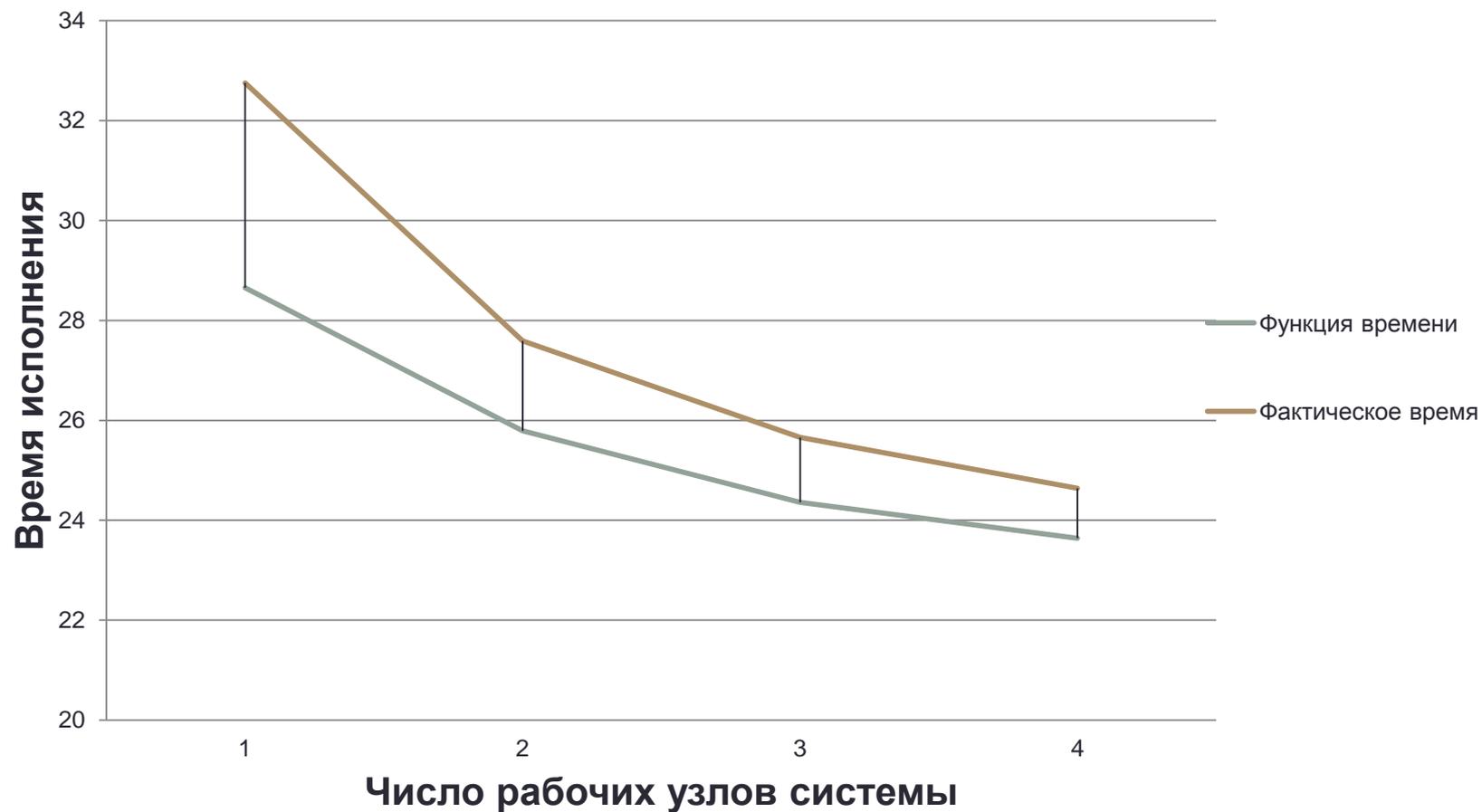
Модель стоимостей: соединение

Сравнение для одного узла



Модель стоимостей: соединение

Сравнение для нескольких узлов



Итоги

- Реализован адаптер для соединения AsterixDB с оптимизатором:
- Реализованы трансляция и модель стоимостей для таких операторов реляционной алгебры как фильтрация, соединение