

НТУУ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО»

«ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»

КАФЕДРА СИСТЕМНОГО ПРОЕКТУВАННЯ

NewSQL та Multi-model бази даних

Виконав: Леонов В. В. ДА-61

Науковий керівник: Булах Б. В.



Об'єкт дослідження

Методи зберігання та обробки даних у інформаційних системах

Предмет дослідження

NewSQL та NoSQL (в тому числі Multi-model) бази даних

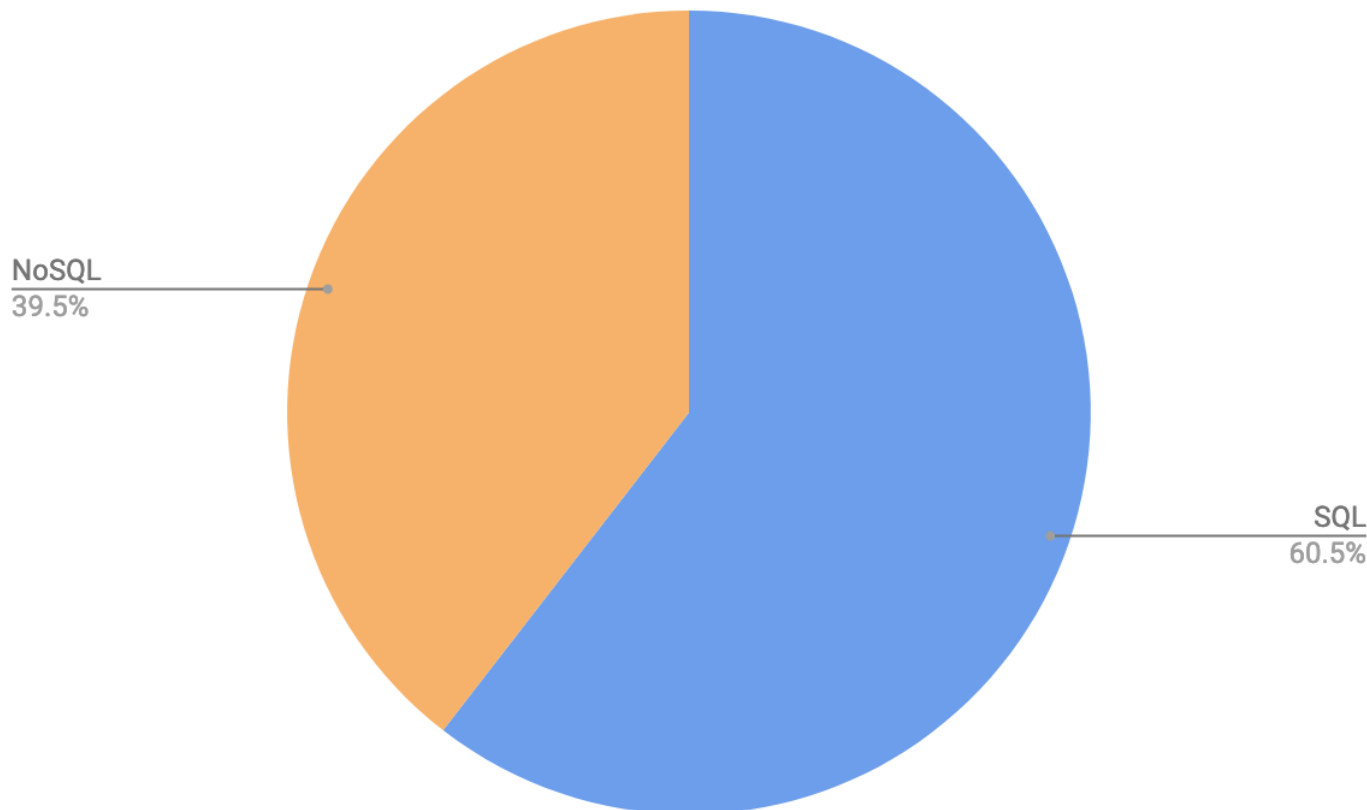
Мета

- провести дослідження характерних особливостей, сильних та слабких сторін для різних типів БД (NewSQL, ММ),
- реалізувати тестовий додаток, що взаємодіє з NewSQL або ММ-БД,
- з його допомогою проаналізувати особливості практичної роботи з такою БД.

Актуальність

Бази даних є однією з ключових технологій у сучасних інформаційних системах. Для різних даних та вимог до їх обробки та зберігання використовуються різні типи та моделі баз даних. Вибір найкращого типу та моделі (або моделей) баз даних може складати труднощі для розробників. Дослідження має виявити переваги та недоліки в різних новітніх типах БД з метою спрощення вибору кращої альтернативи під конкретні вимоги до даних та їх обробки.

Розподіл використання різних типів баз даних



highscalability.com, 06.04.2019

Для досягнення мети

- Проведено огляд історії розвитку баз даних
- Оглянуто різні типи баз даних
- Проведено аналіз різних моделей баз даних
- Проведено порівняння різних СКБД
- Обрано СКБД для тестування
- Проведено тестування СКБД

Типи баз даних та моделі

Типи:

- SQL
- NoSQL
- NewSQL

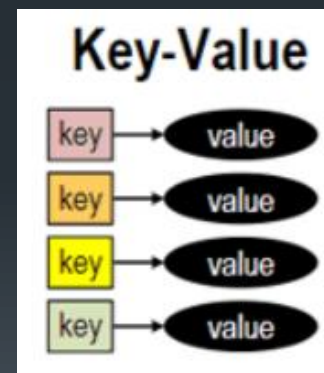
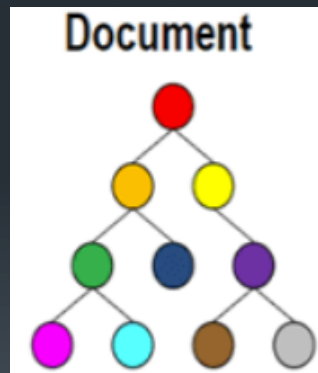
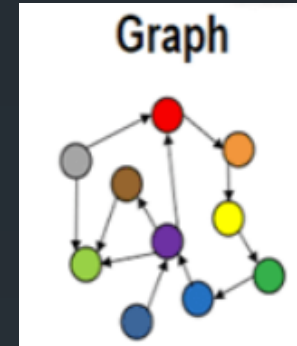
Моделі:

- Реляційна
- Документна
- Ключ-значення
- Графова

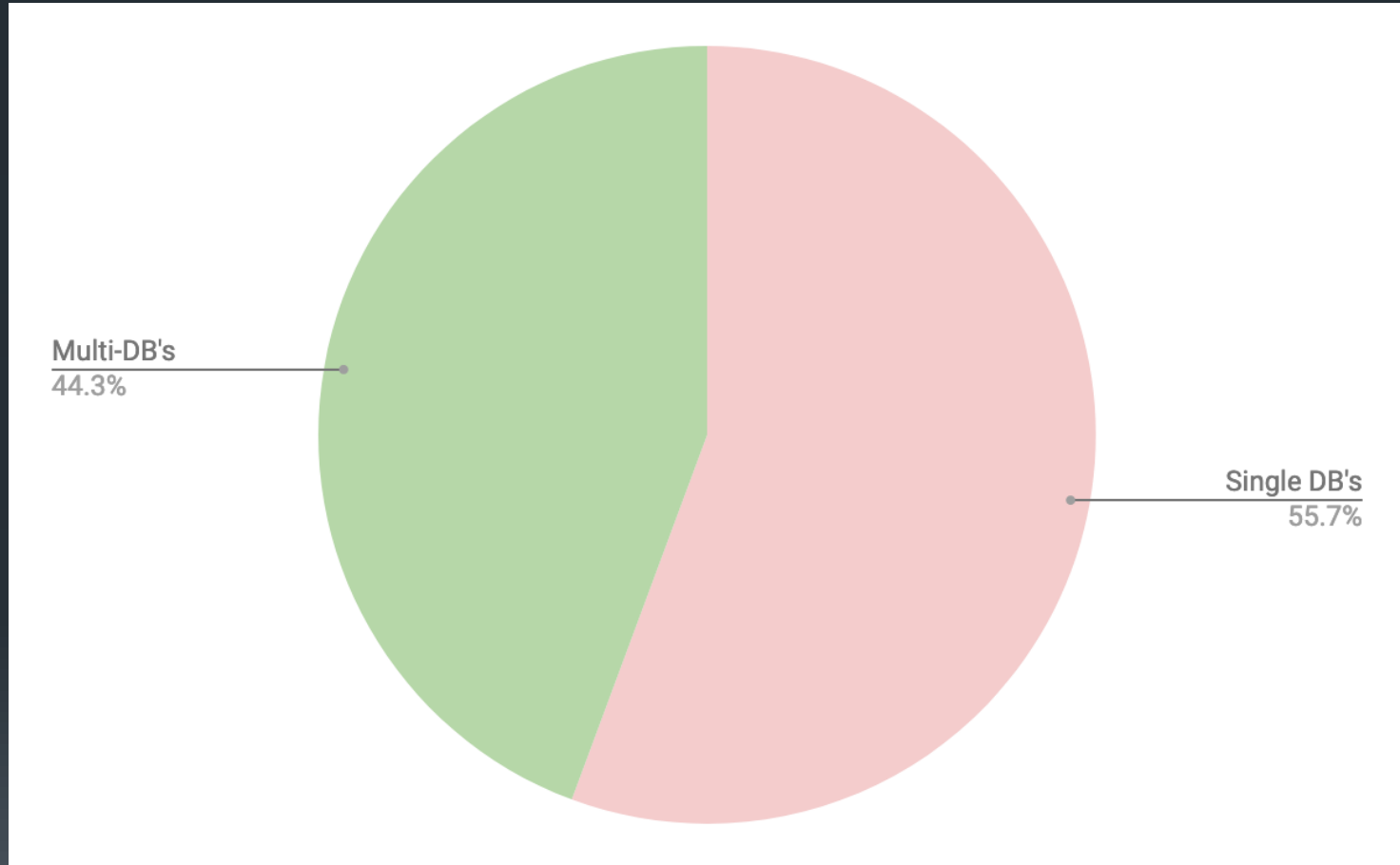
Порівняння різних типів БД

	SQL	NoSQL	NewSQL
Модель	Реляційна	Ключ-значення, документна, стовпцева, графова і т.д.	Реляційна
Підтримка SQL	Підтримує SQL (у різних СУБД може відрізнятись імплементація)	Немає підтримки SQL	Підтримує розширений функціонал SQL
Забезпечення ACID	Так	Ні	Так
Підтримка OLTP	Неефективна для OLTP	Має підтримку OLTP, але не найкращий варіант	Забезпечує повний функціонал OLTP баз даних з високою ефективністю
Масштабованість	Вертикальна	Вертикальна	Горизонтальна
Можливість розподілення	Ні	Так	Так
Обробка запитів	Легко обробляє прості запити, провалюється при складних запитах	Краще ніж SQL при складних запитах	Високоєфективно обробляє великі та малі запити
Наявність схеми	Використовує жорстку та чітко визначену схему даних.	Не покладається на схему при записі даних	Має чітку схему зберігання даних

Моделі баз даних



Розподіл використання ММ БД



highscalability.com, 06.04.2019

Наявні моделі зберігання даних Multi-model СКБД

	Реляційна	Документна	Графова	Ключ-значення	Інші
Oracle Database	+	+	+	+	Сховище об'єктів
Microsoft SQL Server	+	+	+	-	_____
MarkLogic Server	+	+	+	-	Сховище бінарних файлів
ArangoDB	-	+	+	+	_____
OrientDB	+	+	+	+	_____

Порівняння NewSQL СКБД

	NuoDB	Clustrix DB	VoltDB	CockroachDB	MemSQL	Altibase	Pivotal Gemfire XD
Зручність використання	7.2	7.4	7.1	7.4	6.8	6.3	6.7
Функціонал	6.3	8.5	4.8	7.0	6.9	6.2	7.1
Розширені функції	9.0	8.6	4.5	7.2	6.8	9.0	6.5
Інтеграція	7.6	8.4	9.2	9.4	6.8	8.8	6.4
Швидкодія	7.9	7.7	8.0	9.0	6.7	8.9	7.8
Підтримка користувачів	8.0	10	9.0	7.6	7.6	7.8	7.6
Загалом	7.7	8.4	7.1	7.9	6.9	7.8	7.0

Тестування ArangoDB



```
const QueryBuilder = () => {
  const priv = {
    // private fields and methods
  };

  const pub = {
    getQueryTree() {
      return priv.queryTree;
    },

    fromSchema(schema) {
      priv.mainCollectionName = schema.title;
      priv.queryTree.loop = `FOR ${schema.title}Item in ${schema.title}`
      priv.queryTree.sorting = `SORT ${schema.title}Item.id`;
      // some more code

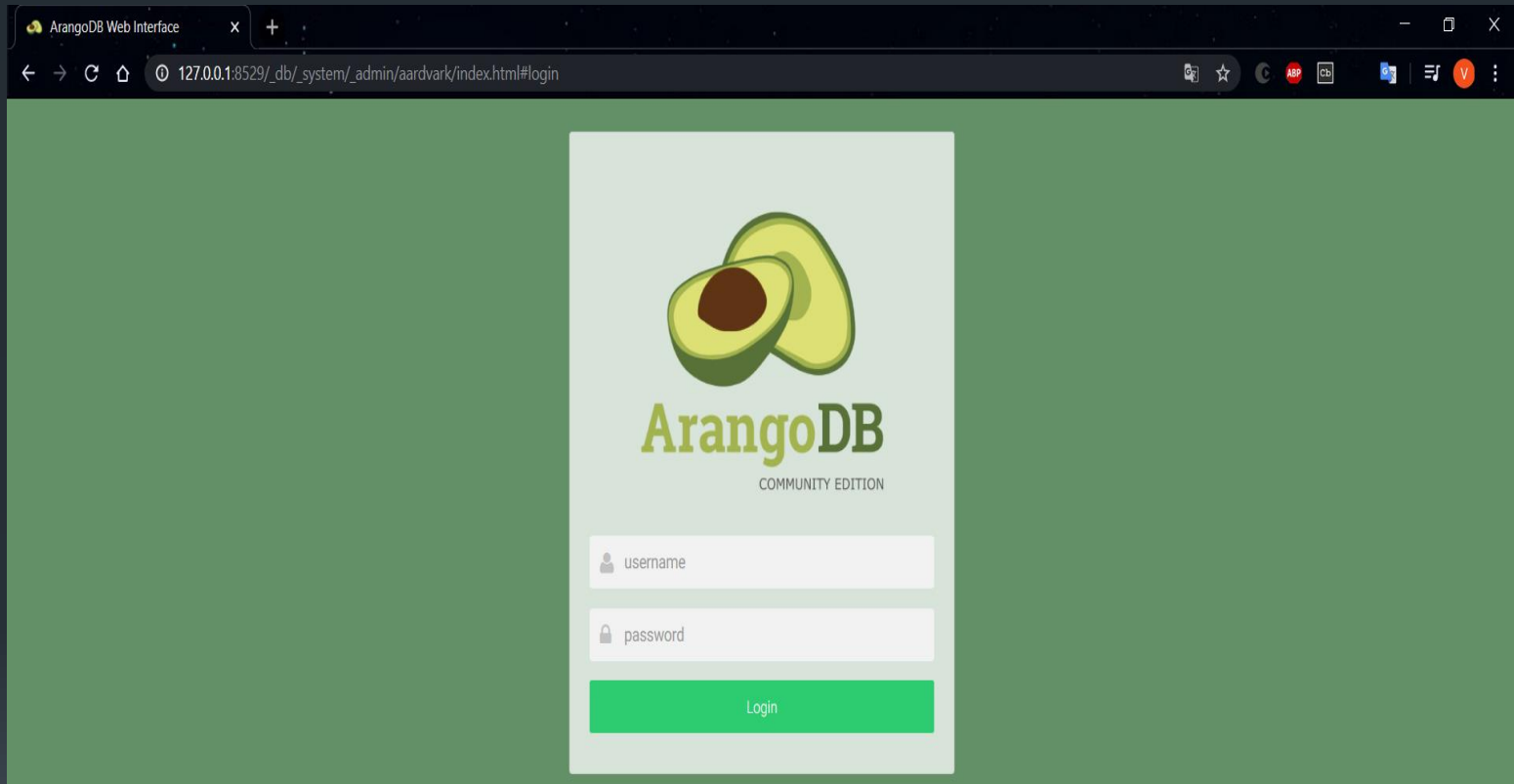
      return pub;
    },

    withLimit(offset, count) {
      // some code
    },

    byId(id) {
      // some code
    },

    byIdentifiers(identifiers) {
      // some code
    }
  };
};
```

Тестування ArangoDB



AQL

Filtering rows / documents from a table / collection, with projection

SQL:

```
SELECT CONCAT(firstName, " ", lastName)
  AS name, gender FROM users
WHERE active = 1;
```

AQL:

```
FOR user IN users
  FILTER user.active == 1
  RETURN {
    name: CONCAT(user.firstName, " ",
                  user.lastName),
    gender: user.gender
  }
```

Тестування ArangoDB

The image shows a screenshot of the ArangoDB Web Interface. The browser address bar displays the URL `127.0.0.1:8529/_db/_system/_admin/aardvark/index.html#collections`. The interface includes a sidebar with navigation options: DASHBOARD, COLLECTIONS, VIEWS, QUERIES, GRAPHS, SERVICES, USERS, DATABASES, REPLICATION, LOGS, and SUPPORT. The main content area shows the 'New Collection' dialog box with the following fields:

- Name*:**
- Type:**
- Advanced:**

At the bottom of the dialog box, there are two buttons: **Cancel** and **Save**. The background interface shows the 'COLLECTIONS' section is active, and a 'loaded' status is visible in the bottom right corner.

Тестування ArangoDB

The screenshot displays the ArangoDB Web Interface in a browser window. The address bar shows the URL `127.0.0.1:8529/_db/_system/_admin/aardvark/index.html#graphs`. The interface includes a sidebar with navigation options: DASHBOARD, COLLECTIONS, VIEWS, QUERIES, GRAPHS (highlighted), SERVICES, USERS, DATABASES, REPLICATION, LOGS, and SUPPORT. The main content area shows the 'Create Graph' dialog box, which is currently open. The dialog has two tabs: 'Examples' and 'Graph'. The 'Graph' tab is active, showing the following fields:

- Name*:** (with an information icon and a plus icon to its right)
- Edge definitions*:** (with an information icon to its right)
- fromCollections*:** (with an information icon to its right)
- toCollections*:** (with an information icon to its right)
- Vertex collections:** (with an information icon to its right)

At the bottom of the dialog, there are two buttons: 'Cancel' and 'Create'.

Тестування ArangoDB

```
ArangoContext arango = new ArangoContext( "Server=http://localhost:8529;Realm=myproject;User=root;Password=ПАРОЛЬ;");
//Створення колекції:
await arango.CreateCollectionAsync("Назва бази", "назва колекції", ArangoCollectionType.Document);

Створення графу:
await arango.CreateGraphAsync("Назва бази", new ArangoGraph
{
    Name = " Назва графу",
    EdgeDefinitions = new List<ArangoEdgeDefinition>
    {
        new ArangoEdgeDefinition
        {
            Collection = "колекція ребер графу",
            From = new List<string> {"collection"},
            To = new List<string> {"collection"}
        }
    }
});
//Виконання та ітерація результатів запитів:
var result = arango.QueryAsync<JObject>("database",
"FOR user IN test FILTER user.gender == 'Male' SORT user.first_name RETURN user");
result.Wait();
foreach (var item in result.Result)
{
    Console.WriteLine(item.ToString());
}
```

Тестування ArangoDB

```
C:\Users\Admin\source\repos\temp\temp\bin\Debug\temp.exe
{"id": 540,
 "first_name": "Yorgos",
 "last_name": "Djorvic",
 "email": "ydjorvicez@mit.edu",
 "gender": "Male",
 "ip_address": "85.114.120.188"
}
{
  "_key": "1042",
  "_id": "test/1042",
  "_rev": "_aoKsKcW-Ae",
  "id": 631,
  "first_name": "Zachery",
  "last_name": "Mesnard",
  "email": "zmesnardhi@miibeian.gov.cn",
  "gender": "Male",
  "ip_address": "82.31.43.11"
}
{
  "_key": "1207",
  "_id": "test/1207",
  "_rev": "_aoKsKCe-_e",
  "id": 796,
  "first_name": "Zedekiah",
  "last_name": "Coady",
  "email": "zcoadym3@noaa.gov",
  "gender": "Male",
  "ip_address": "225.19.119.96"
}
```

Результати тестування швидкодії

	Запит всіх рядків	Запит з фільтруванням по id	Запит з пошуком підрядків	Запит із сортуванням по рядку
Час виконання у мс (1000 записів)	0.988	1.379	1.683	3.529
Час виконання у мс (10 000 записів)	9.634	23.869	31.240	51.126

Висновки

- Було проведено огляд історії розвитку технологій баз даних
- Було проведено порівняння різних типів баз даних
- Було проведено аналіз різних моделей баз даних та випадків їх використання
- Було проведено порівняння СКБД у кожному класі
- Було проведено огляд та тестування багатомодельної СКБД ArangoDB та API для доступу до даних

Висновки (подальший розвиток)



- Провести тестування інших СКБД
- Збільшити вибірку даних для запитів
- Розширити набір запитів для тестування
- Провести тестування на розподіленій системі



Дякую за увагу