MarkLogic 调研

@rh01 https://github.com/rh01

摘要

多模型数据库（Multi-Model Database）是新一代的数据库，与只支持单一数据模型的传统数据库有所不同，多模型数据库是一种在统一、综合的平台下同时支持多种不同的数据模型的数据库，这些数据模型可包括传统的关系模型和 NoSQL 数据模型，并且多模型数据库具有自己一种或多种的查询语言，并不仅仅依赖于传统的 SQL 查询语言，这使得数据组织、管理和操作变得十分的简单和便捷。在企业中，拥有结构良好的数据和基于 NoSQL 技术的综合数据平台对用户是有益的 [2] ，这种方法显著地降低了集成，迁移，开发，维护和运营等问题. 因此，在本论文中，我们将介绍多数据模型的数据库代表-MarkLogic, 将会介绍它的底层架构，所支持的数据模型，以及在底层和顶层如何访问、管理和操作数据，最后设计 Benchmark 来对 MarkLogic 进行评测，并与文档数据库 MongoDB 进行比较.

关键字：Multi-Model Database; MarkLogic; MongoDB

1 MarkLogic 简介

MarkLogic

- A brief motivation and introduction to the topic of the project.
- Discuss history of the technology, mention related technologies (if relevant).
- A brief account of the results that have been obtained in the project.
- A one paragraph overview at the end, explaining how the rest of the report has been organised.

This rest of this report is organised as follows: Section 2 gives an ....

2 MarkLogic 重要特性

About 4 pages that introduces in (sufficient) depth the key concepts and architecture of the technology. May use a running example to introduce the technology.

This part and other parts of the report probably needs to refer to figures. Figure 1 from [1] just illustrates how figure can be included in the report.

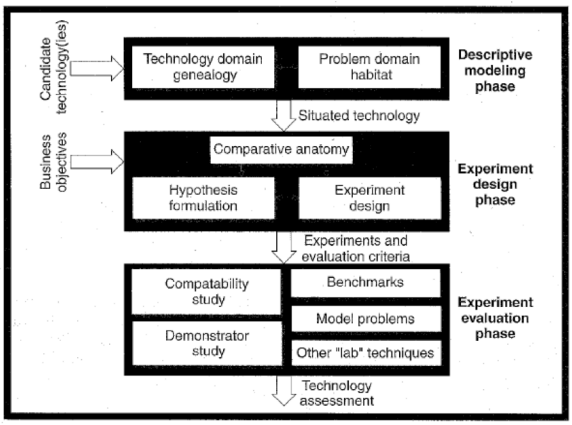

图 1: Software technology evaluation framework.

3 编程 API

About 5 pages that gives:

1. High-level view of the demonstrator and its purpose.
2. Details of how the demonstrator has been implemented.
3. May involve presentation of code snippets.

The example below shows how you may include code. There are similar styles for many other langages - in case you do not use Java in your project. You can wrap the listing into a figure in case you need to refer to it. How to create a figure was shown in Section 2.

```
1  public class BoksVolum {
2
3    public static void main(String[] args) {
4
5      int b, h, d;
6      String btext, htext, dtext;
7
8      [ ... ]
9
10     int volum = b * h * d;
11
12     String respons =
13          "Volum [" + htext + "," + btext + "," + dtext + "] = " + volum;
14
15   }
16 }
```

Config	Property	States	Edges	Peak	E-Time	C-Time	T-Time
22-2	A	7,944	22,419	6.6 %	7 ms	42.9%	485.7%
22-2	A	7,944	22,419	6.6 %	7 ms	42.9%	471.4%
30-2	B	14,672	41,611	4.9 %	14 ms	42.9%	464.3%
30-2	C	14,672	41,611	4.9 %	15 ms	40.0%	420.0%
10-3	D	24,052	98,671	19.8 %	35 ms	31.4%	285.7%
10-3	E	24,052	98,671	19.8 %	35 ms	34.3%	308.6%

表 1: Selected experimental results on the communication protocol example.

4 测试环境和测试结果

About 3 pages that:

Describes the software used to establish the test-bed and for implementing the demonstrator prototype.

Explains what experiments have been done and the results.

For some reports you may have to include a table with experimental results are other kinds of tables that for instance compares technologies. Table 1 gives an example of how to create a table.

5 结论

Concludes on the project, including the technology, its maturity, learning curve, and quality of the documentation.

The references used throughput the report should constitute a well chosen set of references, suitable for someone interesting in learning about the technology.

参考文献

[1] A.W. Brown and K. C. Wallnau. A framework for evaluating software technology. *IEEE Software*, 13(5):39–49, September 1996.

[2] Jiaheng Lu and Irena Holubová. Multi-model data management: What's new and what's next? In *EDBT*, pages 602–605, 2017.