****

**Data Docment Transfer**

**介绍**

***上海锦木信息技术有限公司***

**商业信用**

|  |
| --- |
| **声明：**  **该文档由锦木信息技术有限公司提交。文中的所有信息均为本公司机密信息，仅供下文中被呈送方使用，务请妥善保管并且仅在与项目有关人员范围内使用，未经本公司明确做出的书面许可，不得为任何目的、以任何形式或手段（包括电子、机械、复印、录音或其他形式）对本文档的任何部分进行复制、存储、引入检索系统或者传播。**  **锦木信息保留所有权利。** |

**文档属性**

|  |  |
| --- | --- |
| **属性** | **内容** |
| 客户名称 |  |
| 项目名称 |  |
| 项目编号 |  |
| 文档主题 |  |
| 文档副标题 |  |
| 文档编号 |  |
| 文档版本 | 1.0 |
| 版本日期 |  |
| 文档状态 |  |
| 内部参照 |  |
| 作者 | 李和平 |

**文档变更**

|  |  |  |  |
| --- | --- | --- | --- |
| **版本** | **修订日期** | **修订人** | **描述** |
| 1.0 | 2021-12-15 | 李和平 | 撰写文档 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**文档送呈**

|  |  |
| --- | --- |
| **客户姓名或客户单位名称** | **目的** |
|  |  |
|  |  |

目录

**一.项目介绍 3**

**二.功能介绍 4**

2.1 高性能 4

2.2同步模式 5

2.3断点续传 6

2.4支持多版本 6

2.5 DDL操作 6

2.6 Oplog延迟 6

2.7同步范围 6

**三.测试与案例 7**

**四.公司介绍 11**

# 一 项目介绍

DDT是由上海锦木信息技术有限公司(以下简称“锦木信息”)结合众多客户需求和锦木信息多年 MongoDB 一线服务及研发经验所推出新一代MongoDB数据库迁移同步工具。

DDT是一个以JAVA语言开发的通用的传输软件，具有高健壮性、高传输性、高可用性的特点，可以快速稳定的迁移数据，帮助用户进行实现数据备份，实时迁移，备份容灾等业务操作。用户也可自定义配置参数来实现场景不同高效传输。

由于目前MongoDB副本集内置的主从同步对于这种业务场景有较大的局限性，为此，我们开发了DDT同步工具，可以应用在实例，机房、跨数据中心复制，满足灾备和多活需求。

传统MongoDB数据同步仅支持相同架构的数据同步，DDT支持单节点，复制集，分片三种类型架构的数据相互传输。帮助用户实现类似于复制集点到分片 分片到单节点的数据同步方案

DDT的核心实时同步：我们利用对OPLOG日志的解析和应用，高效安全的实现实时同步。

源端MongoDB可以是单实例、副本集或分片，而目标可以是 mongod 或 mongos。如果源端是副本集，我们建议从次要/隐藏节点中获取数据以减轻主节点压力。如果源端是分片，每个分片都应该连接到 DDT。

# 

# 二 功能介绍

DDT 具有简单、安全、适用性强、多功能 、高性能等特点。

|  |  |  |
| --- | --- | --- |
| DDT功能模块及描述 | | |
| 功能模块 | 功能项 | 描述 |
| 同步模式 | 全量同步 | 全量同步表,仅同步程序启动前的源表数据 |
| 实时同步 | 实时同步源端产生的oplog |
| 全量和增量同步 | 全量同步后,仅同步同步期间对源表进行的操作。实时同步的开始时间为全量同步的开始时间,实时同步的结束时间为全量同步的结束时间 |
| 全量和实时同步 | 全量同步后,开始进行实时同步,实时同步的开始时间为全量同步的开始时间 |
| 附加特性 | 延迟同步 | 在实时同步期间可以延迟设置读取oplog |
| 同步DDL | 在实时同步期间，用户可以自定义同步某些DDL操作 |
| 源表索引 | 在全量同步期间，用户可以设置是否在总数据同步60%以后，开始同步源表的索引列表 |
| 多表并行 | 在全量同步时，设置多源表同时同步还是单表依次同步 |
| 同步源表名单 | 使用正则表达式设置需要的同步表名单 |
| 时间区间实时同步 | 实时同步时，可以设置同步某时间区间内的oplog |
| MongoDB版本 | 支持版本 | DDT支持3.2，3.4，3.6，4.0，4.4，5.0，6.0版本的 |

**2.1高性能**

1. **高效数据校验**

- 确保数据总量一致

- 确保数据信息一致

- 确保数据索引一致

- 确保数据结构一致

1. **多种同步方案**

- 全量数据复制

- 实时数据同步

- 增量数据同步

- 自定义同步范围

- 复合数据同步方案

1. **高速同步机制**

- 100%传输带宽利用

- 可控CPU利用率

- 内存使用率可配置

- 支持多表并传

1. **小巧稳定高效**

- 体积小巧

- 断点续传

- 支持多版本MongoDB同步

**2.2同步模式**

同步模式：全量，实时，全量和增量，全量和实时。 其中增量是指有时间范围限制的Oplog实时同步。

全量同步：将源端MongoDB库表进行切分查询，查询到的数据多线程批写入到目标MongoDB库。在全量同步模式下，原则上来说提供资源数越多，QPS越高。

实时同步：将数据从源MongoDB复制到另一个MongoDB，以构建冗余复制。它从源 MongoDB数据库中获取 oplog，并在目标 mongo db数据库中重放日志。

**2.3 断点续传**

source端MongoDB意外宕机时重启，DDT依然能够正常 无丢失的同步数据。当DDT意外关闭时，下次启动DDT时，能自动追踪到断点处，且进行追加重传。

**2.4 支持多版本**

目前DDT支持3.2到6.0的MongoDB。新版本的时许表，桶表均 可靠支持传输同步。

**2.5 DDL操作**

在实时同步期间，用户可以自定义同步某些DDL操作。同时DDL操作也会被记录在日志中，方便审查操作。

**2.6 Oplog延迟**

延迟同步Oplog可以方便用户进行故障切换。

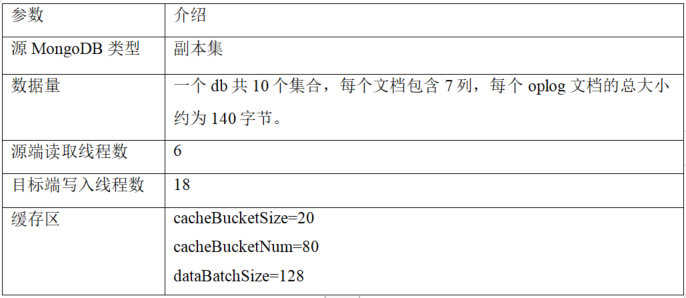
**2.7 同步范围**

实时同步时，可以设置同步Oplog的开始结束时间。

# 三 测试与案例

**测试一**：

配置信息



测试结果

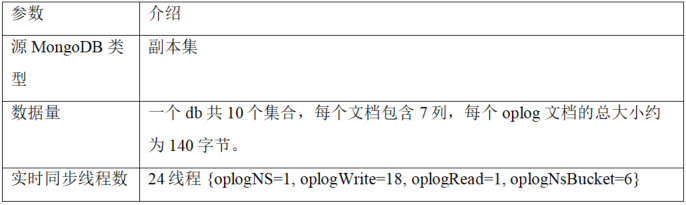


**结论：**在40核CPU，128GB内存，1G带宽的服务器中，全量同步模式下 配置8读取线程，16写入线程，QPS达到25w条每秒。

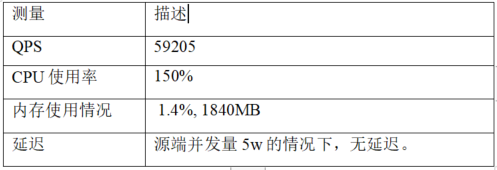
当然在全量同步模式下，原则上来说提供资源数越多，QPS越高。

**测试二**：

配置信息

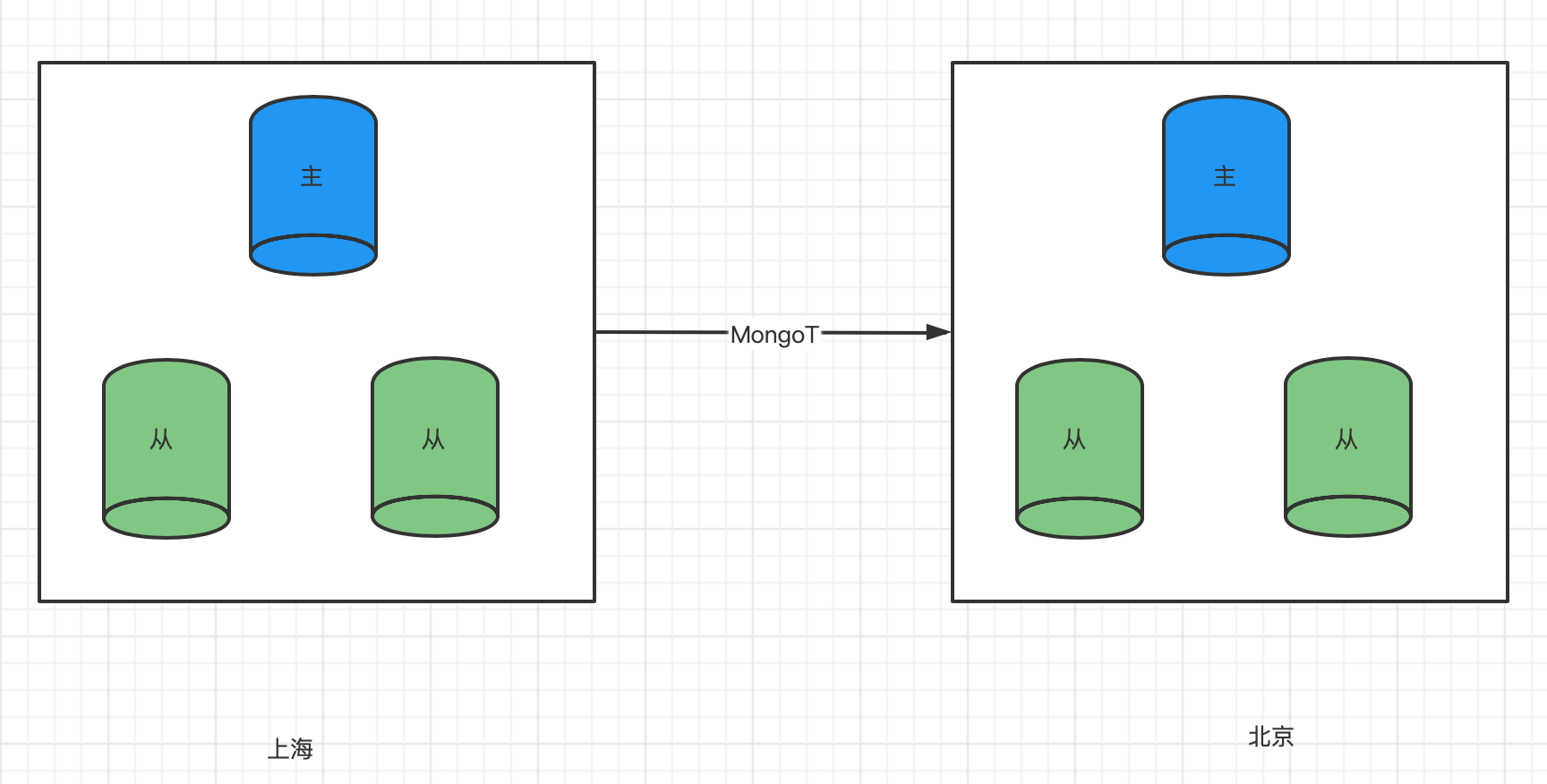


测试结果



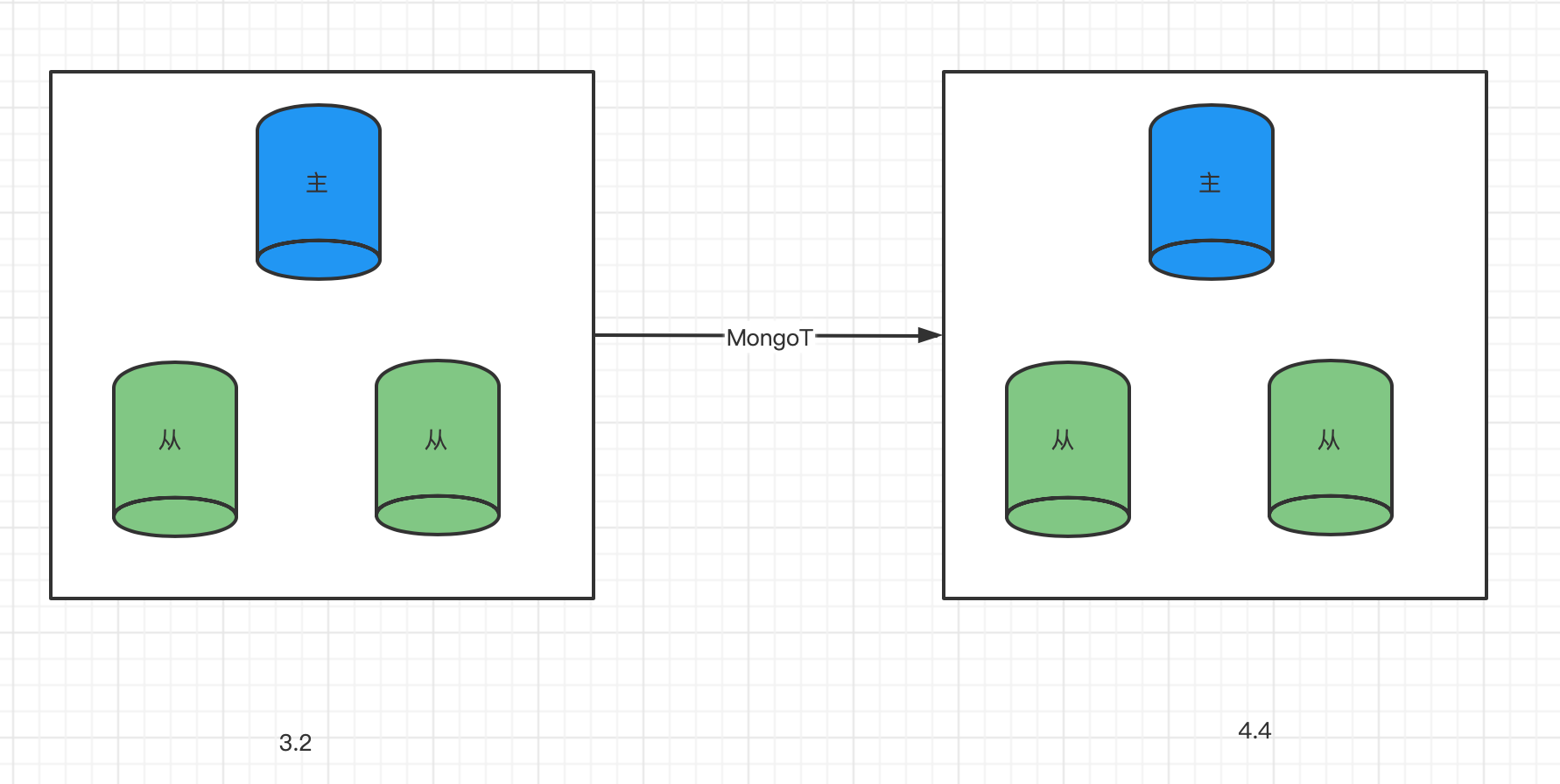
**结论：**在实时同步时，QPS可达6W条每秒，一般是很高的 可以满足绝 大多少业务场景。

**案例一**：某证券公司



容灾的好处：生产中心的业务除了在本地备份外，还可以在灾备中心进行备份，双活架构下支持双中心互备，提升业务韧性，实现业务的双保险。利用DDT同步工具，把远端数据实时写入目标端。

**案例二**：某航空公司



某航空公司，要进行MongoDB复制集集群跨大版本升级3.2至4.4版本。由于应用端需要快速升级变更，传统的MOngoDB复制集需要进行逐级版本升级，此过程比较耗时，且出现异常情况，不能及时切换原正确状态。

我司给航空公司的的解决方案是，搭建新4.4版本数据库，利用DDT实时迁移旧数据到新集群，当新旧集群无延迟时，修改应用端数据库地址。

在此案例中，原端数据为700GB，实时数据1w条每秒，中间穿插出现建标，删表等DDL操作。DDT一共用时6小时传输完成，其中全量用时5小时，实时用时1小时。

# 四 公司简介

上海锦木信息技术有限公司是一家专业的IT数据咨询与服务提供商，我们致力于向用户提供优质的信息化产品、咨询及服务，锦木于2015年在上海成立，并于北京、深圳、广州设立分支机构。

锦木信息是MongoDB北亚区核心合作伙伴，Akamai、Vonage的中国区核心合作伙伴，为客户提供专业技术服务、咨询、应用开发。

作为以技术创新、客户为先驱动的IT服务提供商，锦木信息的产品及服务已被国内领先企业所认可，拥有 50+ 核心客户，在金融、保险、券商、游戏、电商等行业推出了优质服务和创新产品解决方案，销售业务涵盖中国大陆及中国香港。

锦木信息官网：www.jinmuinfo.com

咨询邮箱：support@jinmuinfo.com

联系电话：021-58870038

021-66696778