

ORACLE

# Oracle Converged DB for Development - 下一代OA系统技术方案分享

---

Johnson Kim (金永敏)

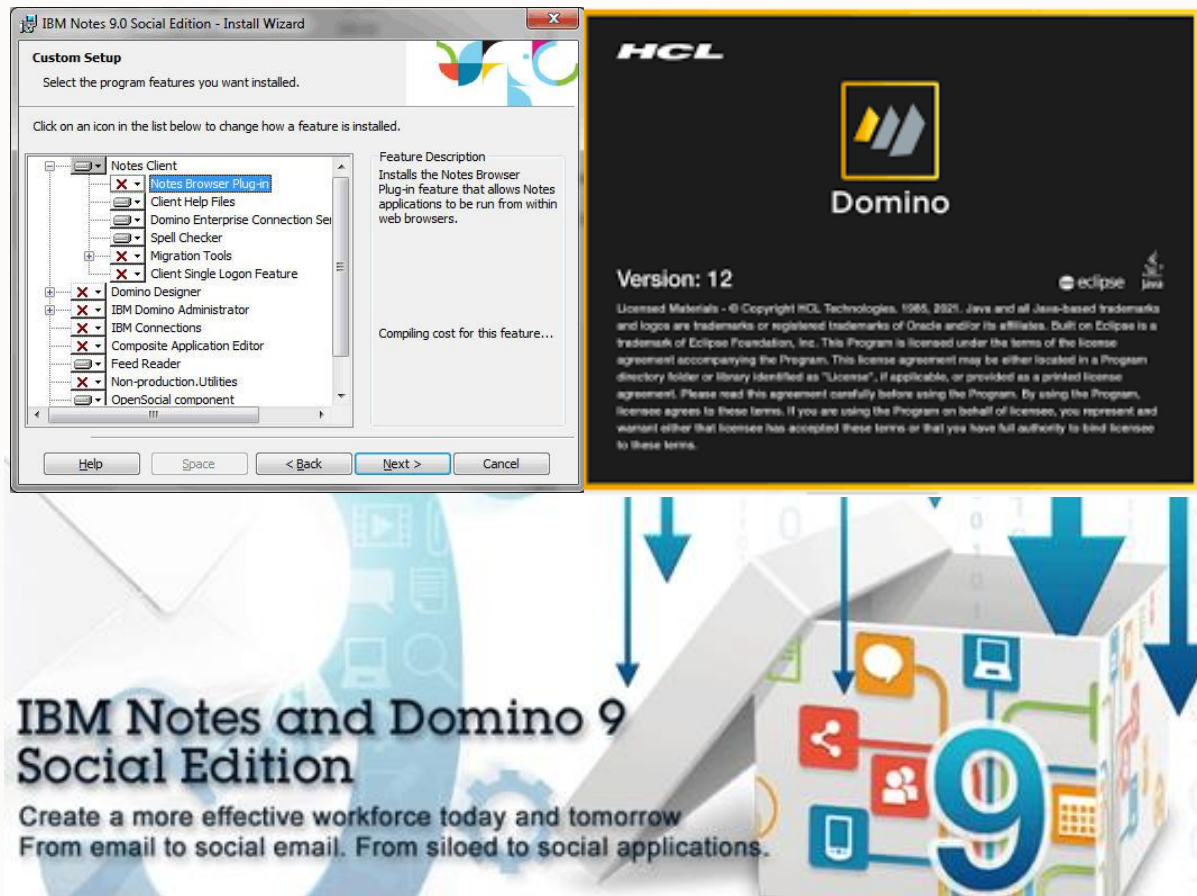


# Safe harbor statement

The following is intended to outline our general product direction. It is intended for information purposes only and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.

The development, release, timing and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

# 客户 OA 系统存在的问题



## 使用上的诸多限制

- 附件大小限制
  - 单个文件10MB
  - 总大小 20MB
- 响应速度很慢

## 陈旧的架构（版本陈旧）

- IBM 最新版停留在 9.0.1 /2013
- 需多台 Domino 服务器, 分离冷热数据
- 开发者社区不活跃



# 方案选择

客户当前 OA 系统 (Domino 9.01)



OA 系统选项

客户青睐的选项

1. Domino 技术

2. .Net + RDBMS

3. JAVA + RDBMS



# 方案背景

客户当前OA系统 (Domino 9.01)



	分类	条目	详细
1	使用软件	使用软件 数据库	Lotus Domino (版本9.01) Lotus Domino 文档数据库
2	用户数/并发量	用户数 并发数	用户总数 3万5 高峰期并发用户 3k ~5k
3	数据库部署情况	主库+归档库+备库	
4	数据库分布 (按功能分库)	发文库 收文库 (量最大) 呈批件管理库 情况通报库 公司公告库 辅助库	精益工作, 已办工作 (附件多, 但库不大) 待办, 已办条目数比较多10W 左右。 已办条目只保留一个月
5	数据量	5000 万条记录 40G/月	公文都带有附件 收文库只保留一个月 40G/月 (每天执行批处理, 除一个月以外的数据全部归档处理) 共 10 TB
6	限制	附件大小	10MB, 总共 20MB (下一代系统不限制附件大小)
7	字段	每个 Note 文档 200~300 字段	包括标签字段
8	性能	响应时间	高层领导抱怨系统慢 (下一代系统提供高性能)



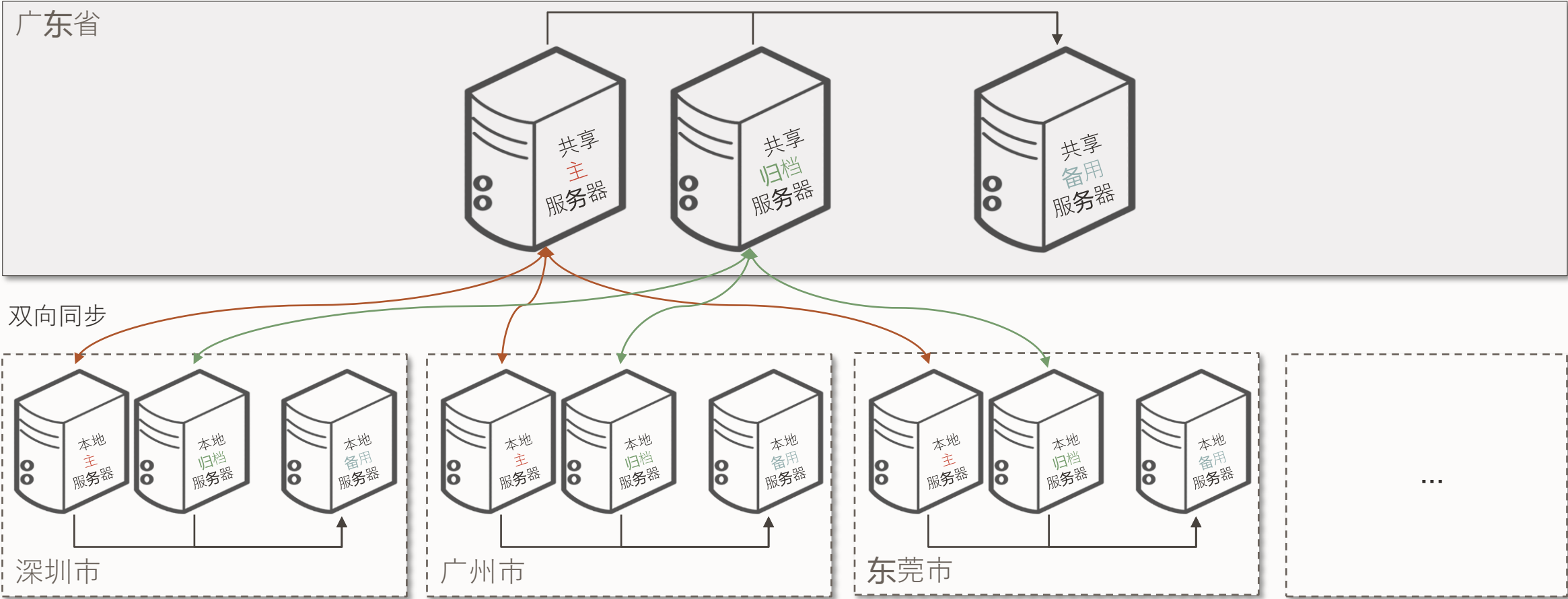
# Domino OA 系统架构



Domino 主服务器：保存热数据，如待处理、已批准的公文（1个月内）

Domino 归档服务器：保存冷数据，如已批准的公文（超过1个月）

Domino 备用服务器：保存所有公文





# OA 系统关注点







## 议程

- 1 业务敏捷性
- 2 用户体验
- 3 如何保存附件
- 4 数据安全性
- 5 高可靠性和灵活性
- 6 客户受益
- 7 Q&A

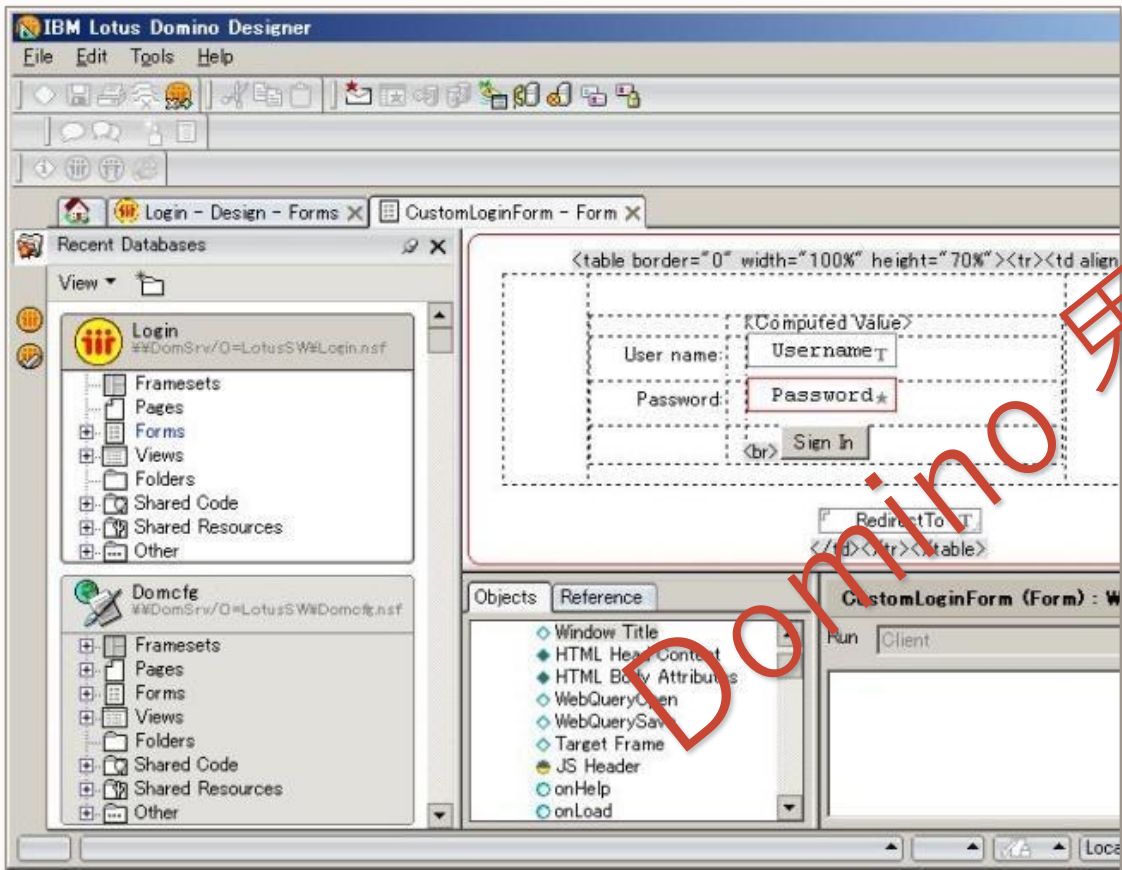




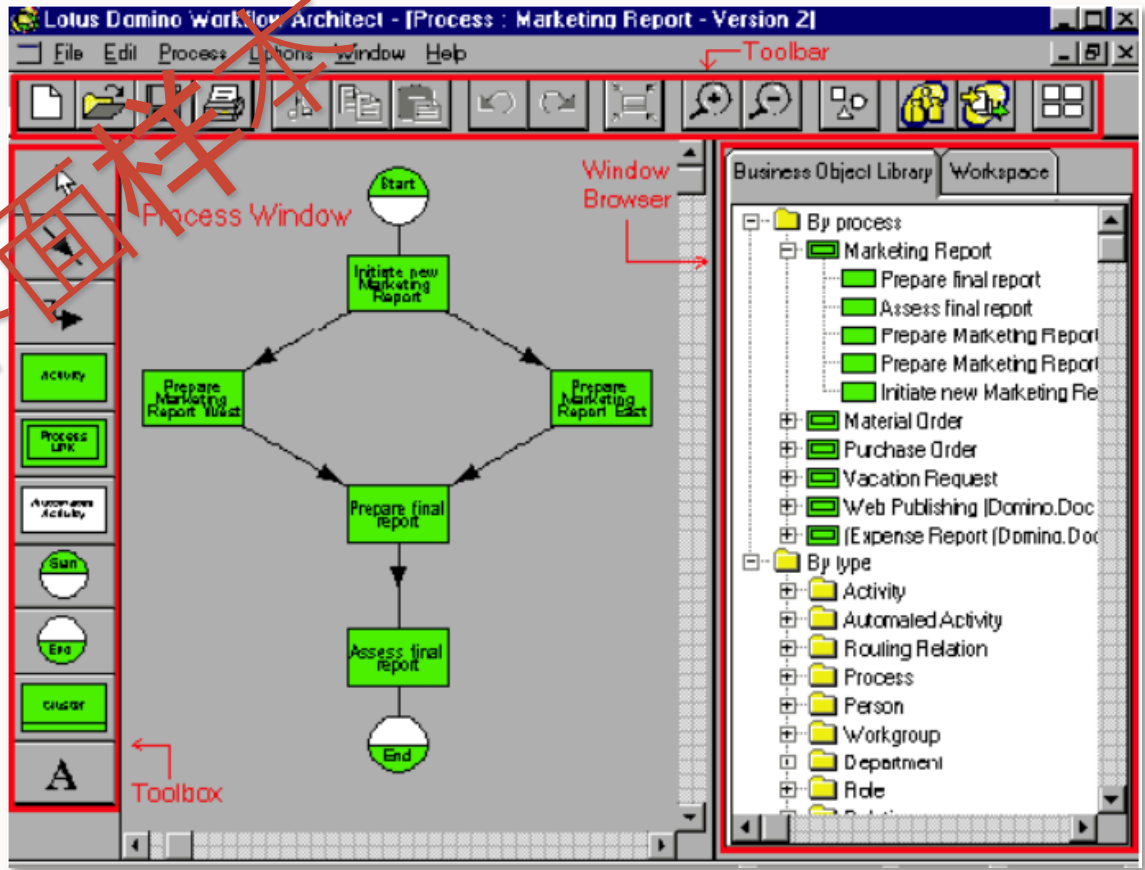
## 1. 业务敏捷性

# 表单设计器 & 工作流 (As-Is)

## 表单设计器



## 工作流管理







To

dalian@oracle.com

Subject

下一代OA系统架构

Normal

Normal

Sans Serif

B

I

U

☒

x<sub>2</sub>

x<sup>2</sup>

T<sub>x</sub>

A

☒

☰

☷

☶

☵

☱

”

</>

🔗

🖼️

📺

f<sub>x</sub>

[source]

⚙️

✕

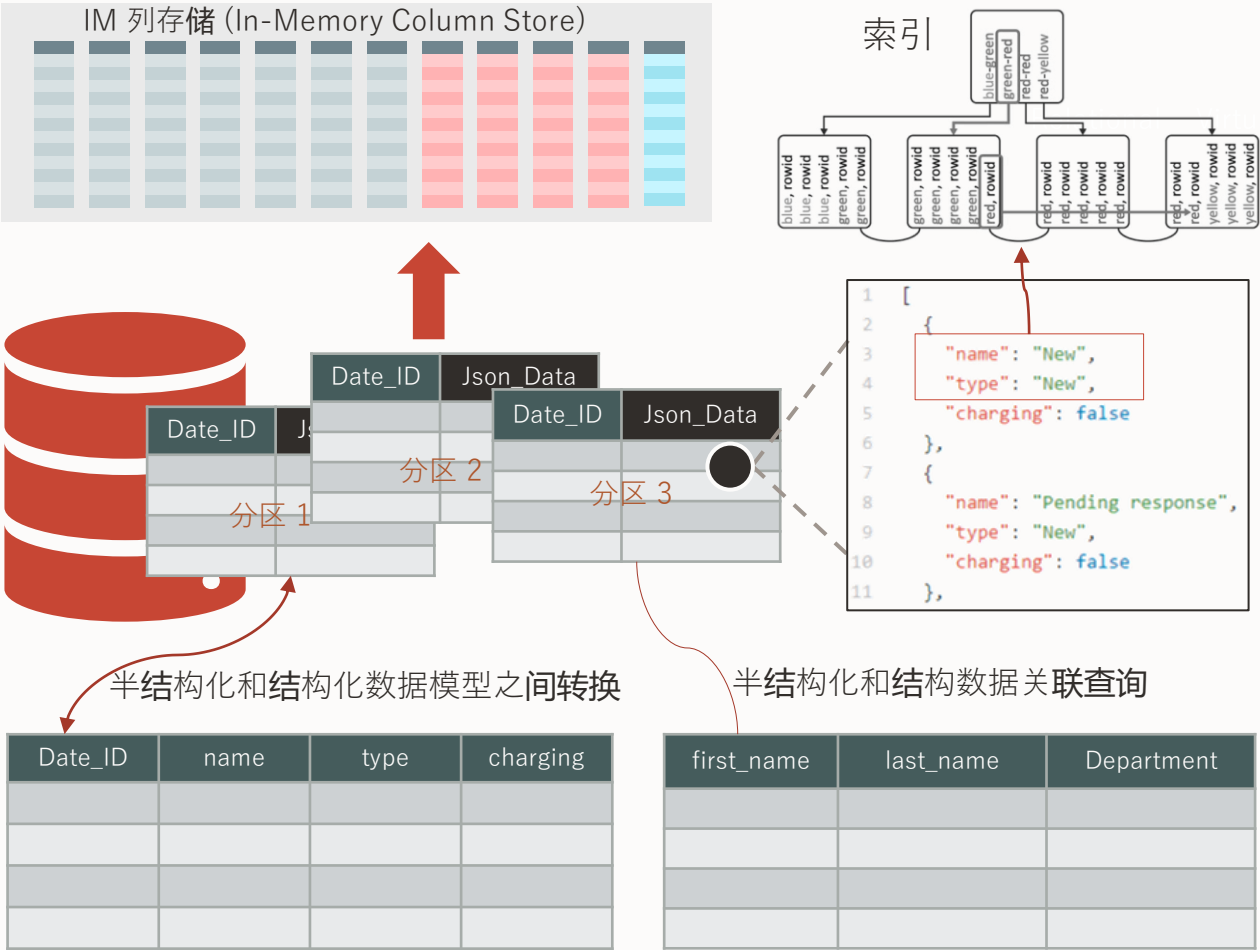
JAVA+RDBMS  
.....

Send





# Oracle 数据库中的 JSON



- 半结构化数据与结构化数据的关联查询
- 提供丰富的 JSON 操作函数
- 不同数据模型之间的相互转换
- 支持基于 Json 元素的表分区
- 支持多种索引，如 b-tree索引，复合索引、位图索引，以及搜索索引
- 支持完整的事务
- 20 - 60x 性能的 IM 列存储 (In-Memory Column Store)



# JSON 数据库用例

移动应用



IOT



360 度客户视图



用户信息管理



目录数据



内容管理



在线广告



实时大数据



社交网络



游戏





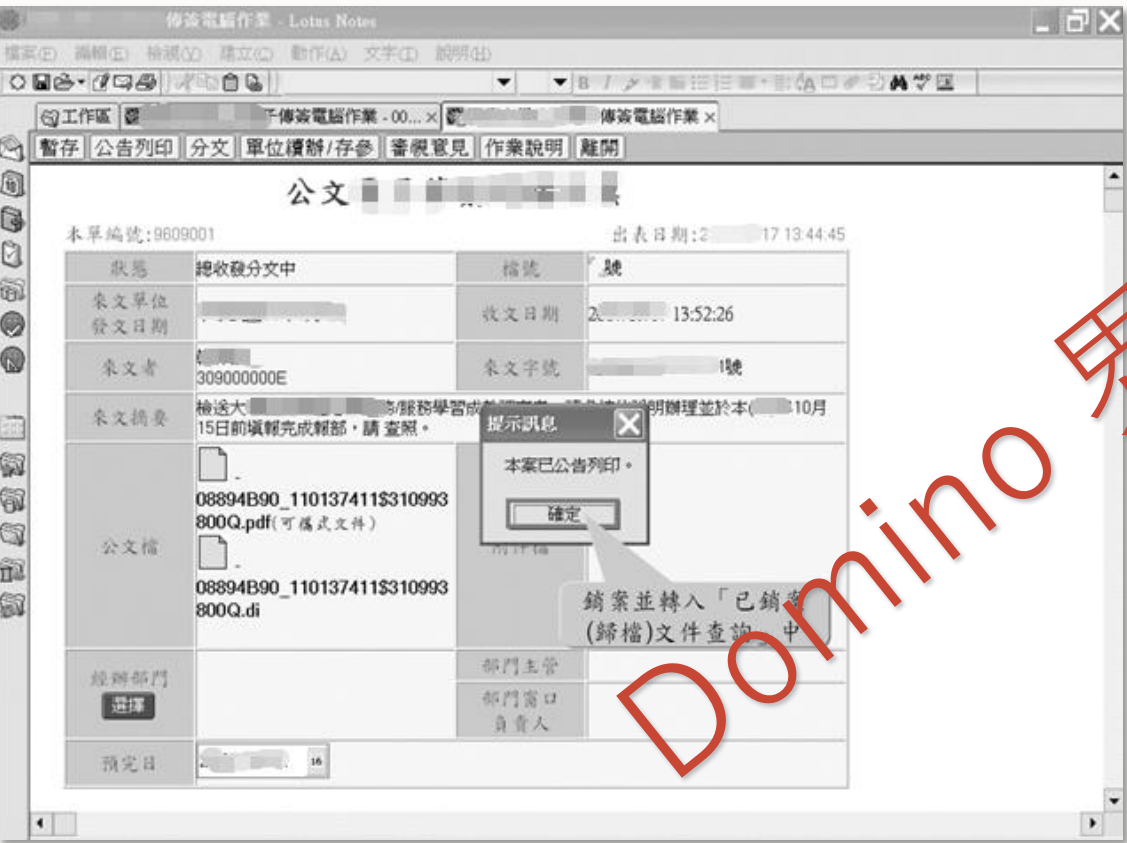
## 2. 用户体验



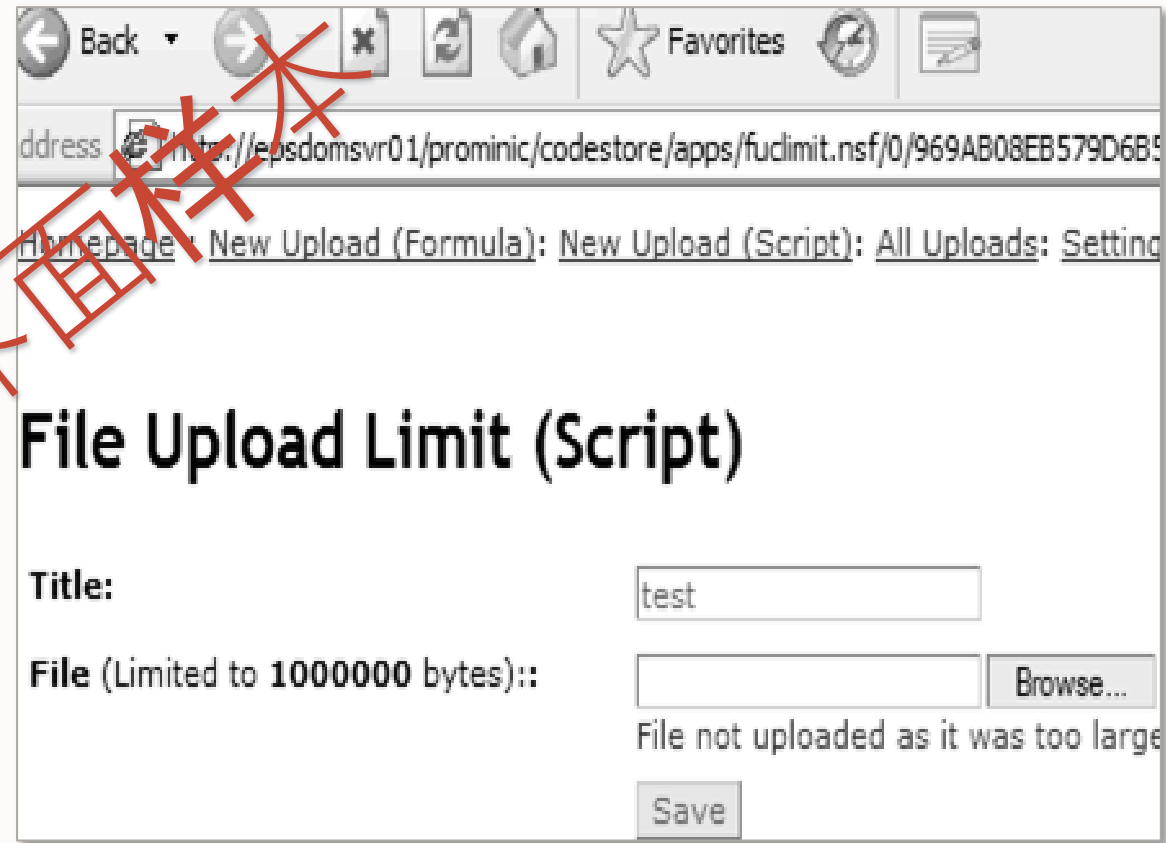
# Domino – 附件上传



C/S 模式



B/S 模式



# 附件上传模式

## 传统模式

<!DOCTYPE html>  
<html>  
<body>  
  
<h1>Show File-select Fields</h1>  
  
<h3>Show a file-select field which allows only one file to be chosen:</h3>  
<form action="/action\_page.php">  
  <label for="myfile">Select a file:</label>  
  <input type="file" id="myfile" name="myfile"><br>  
<br>  
  <input type="submit">  
</form>  
  
<h3>Show a file-select field which allows multiple files:</h3>  
<form action="/action\_page.php">  
  <label for="myfile">Select a file:</label>  
  <input type="file" id="myfile" name="myfile" multiple><br>  
<br>  
  <input type="submit">  
</form>  
  
</body>  
</html>

### Show File-select Fields

Show a file-select field which allows only one file to be chosen:

Select a file:  No file chosen

Show a file-select field which allows multiple files:

Select files:  No file chosen

Open

« Data ... » Technology Solution Engineering ...

Search Technology Solution Engineering ...

Organize New folder

Technology S

[Alvin Kim]

[Aspen Sang]

[Chulnam K]

[David Choi]

[Guang Xu]

[Hidemasa I]

[J.S. Park]

[JangWhan]

2020-06-12 Peer Coaching Hub Link...

2020-06-19 JAPAC TCE - OCI Certificat...

2020-11-19 Certification-Summar...

2021-01-08 China SE-Org [Tommy ...]

2021-02-00 FY21-Feb-S E Hub Solution ...

2021-03-19 JAPAC\_SEH UB\_Tech\_B angalore...

2021-04-30 Tech Hub Knowledge Sharing [J...

File name:

All Files (\*.\*)

Open Cancel

## 当前流行模式

JSON Form x HTML file u x PHP File Up x HTML input x Tryit Editor x Inbox (32,364) x

mail.google.com/mail... 32,364

Compose

Inbox 32,364

Starred

Snoozed

Important

Chats

Sent

Drafts 16

All Mail

Spam 22

Trash

Categories

Meet

New meeting

Join a meeting

Hangouts

J.S. v

Jiyong Cai Sent a message

Hunter Shin No one has joined your call

harrisonkang@gmail.com Missed video call

Renhua Piao Missed video call

Ann Kelli Chabshank

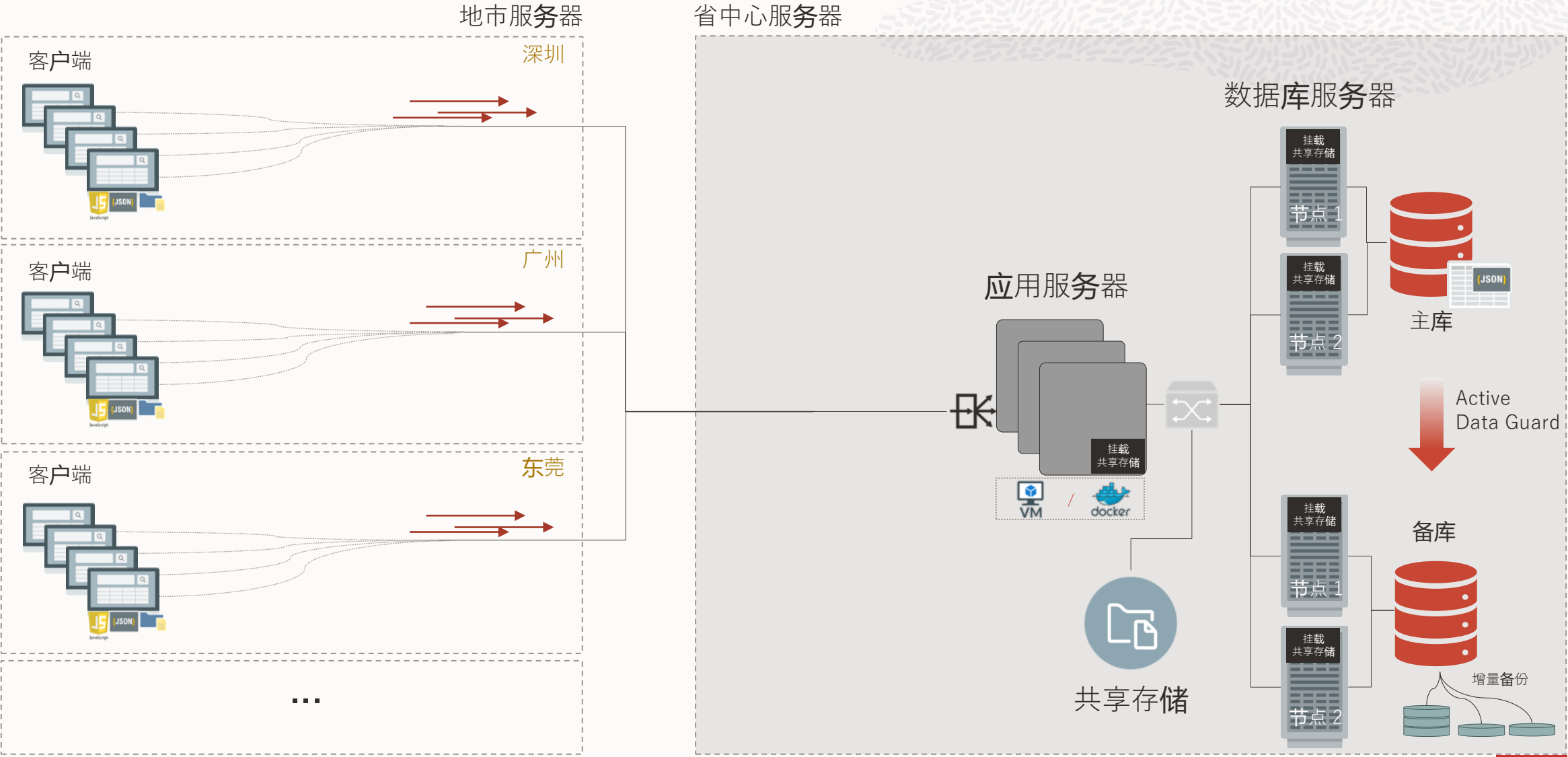
2020-04-16 Message from G... 129.5M

2021-02-00 FY21-Feb-SE Hub... 19.3M

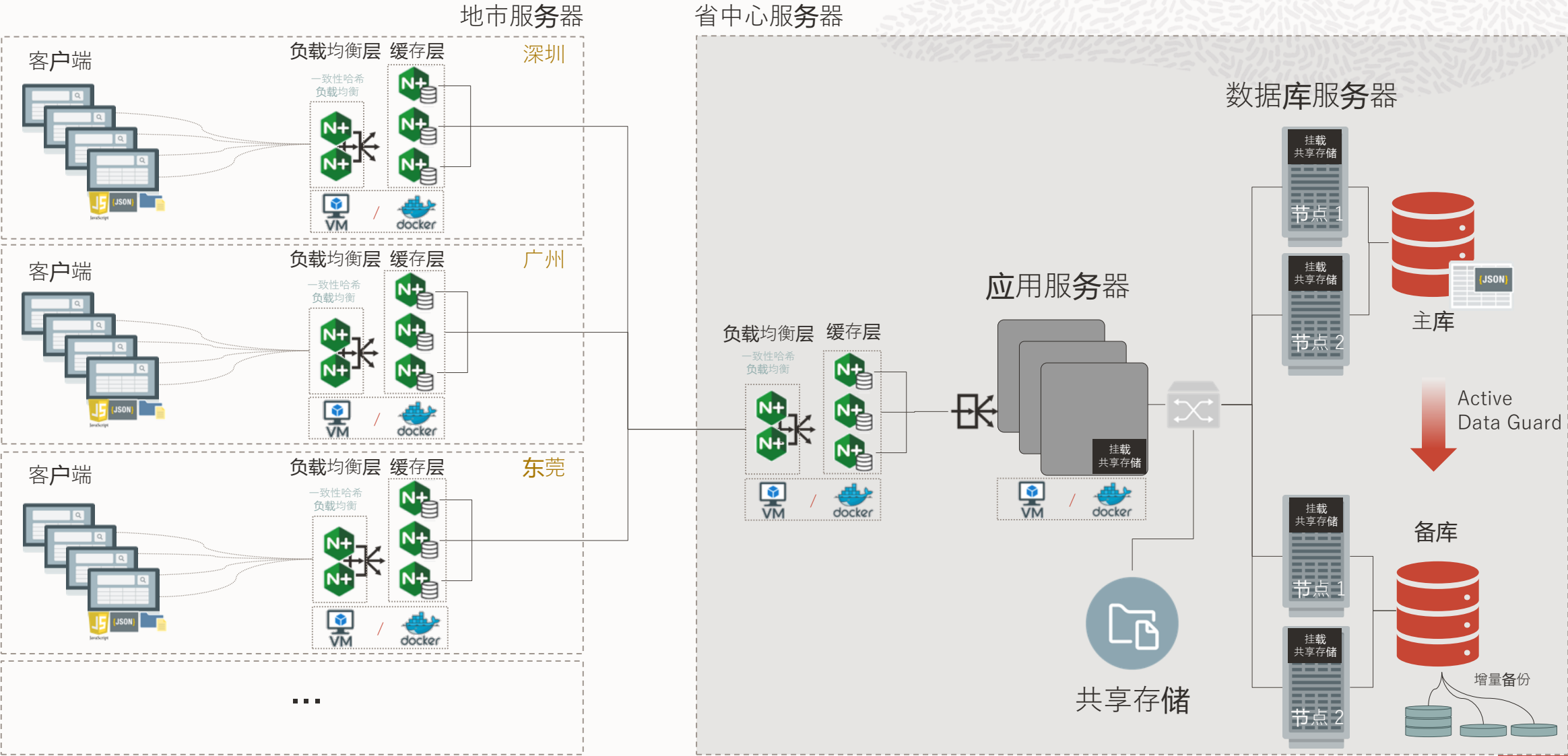
Cancel

18 Copyright © 2022 Oracle and/or its affiliates

# 架构图 1



# 架构图 2

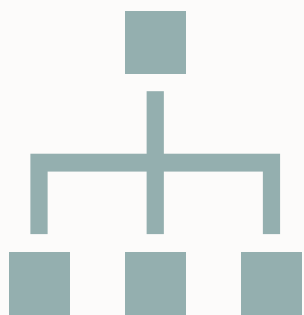






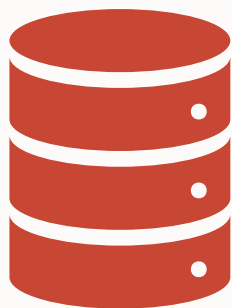
### 3. 如何保存附件

# 文件系统 Or 数据库



文件系统

Or



数据库系统

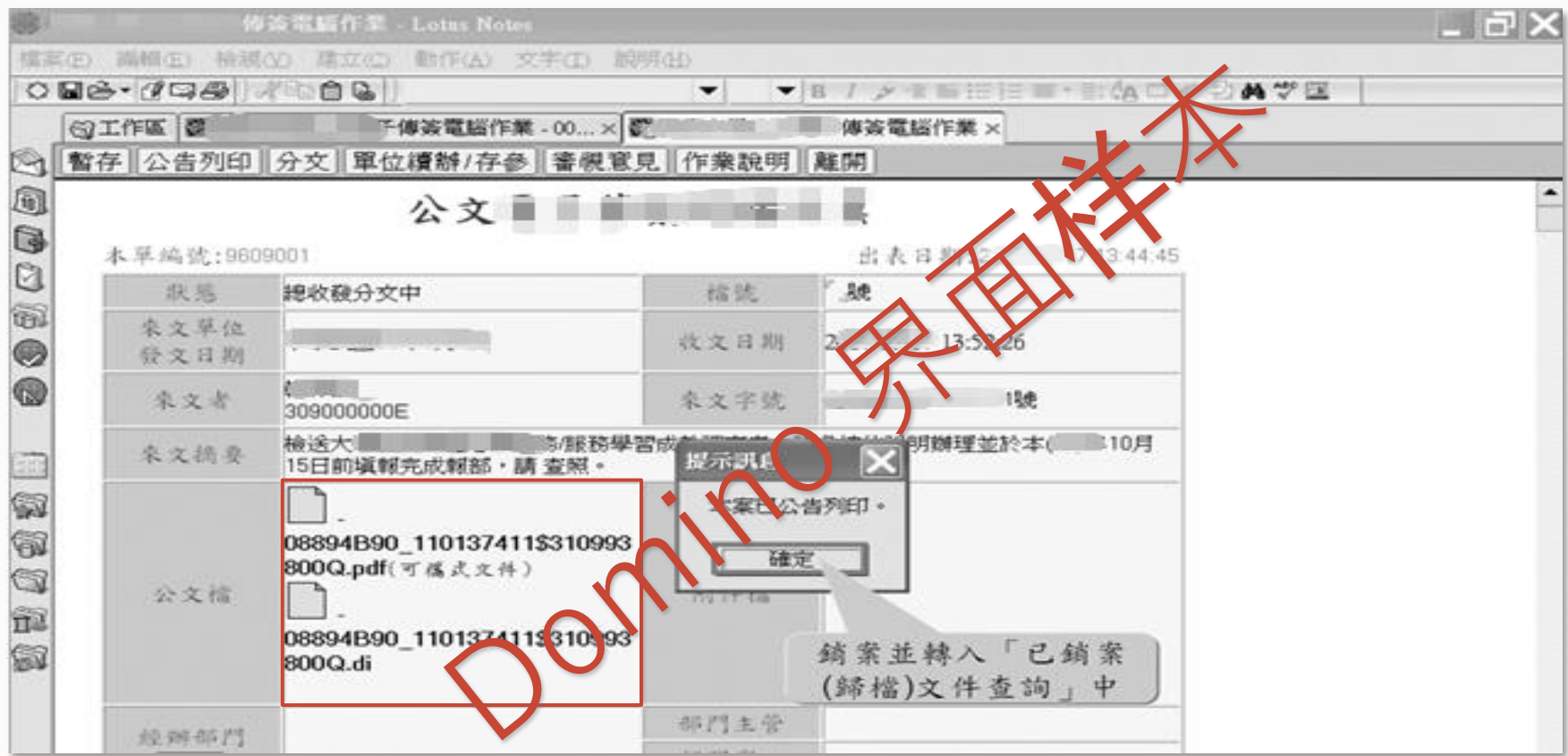
## 为什么在文件系统中存文件？

- 文件读写速度比较快
- 允许使用轻量级网络服务器/CDN/Amazon S3 服务
- 成本效益，文件服务器比数据库便宜得多
- 避免CRUD 操作带来的数据库开销
- 更容易与第三方共享文件（视频、音频等）

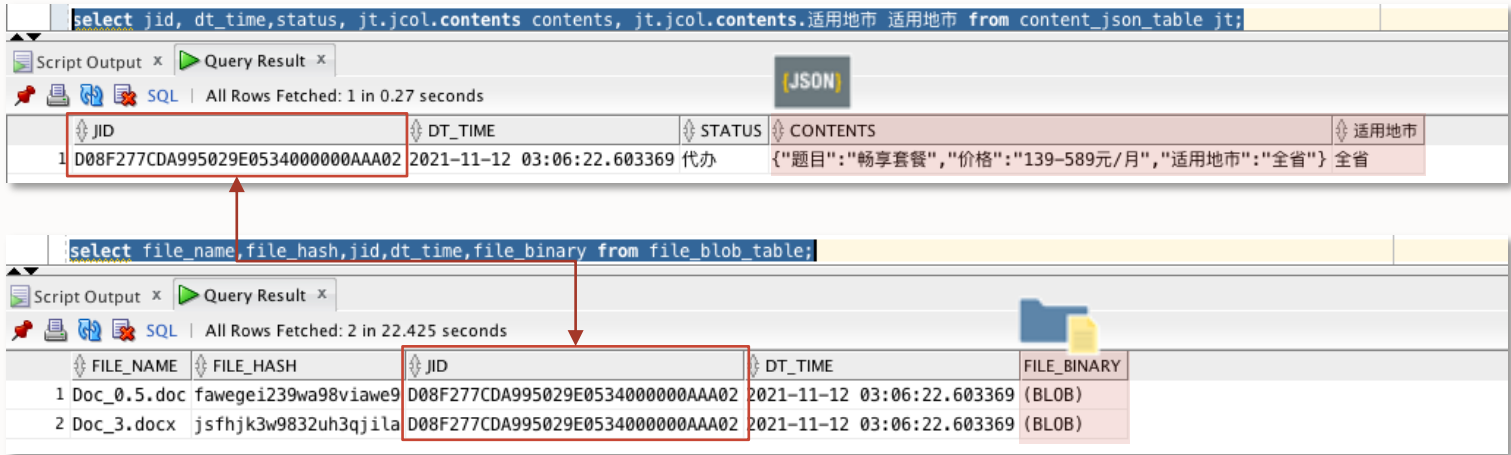
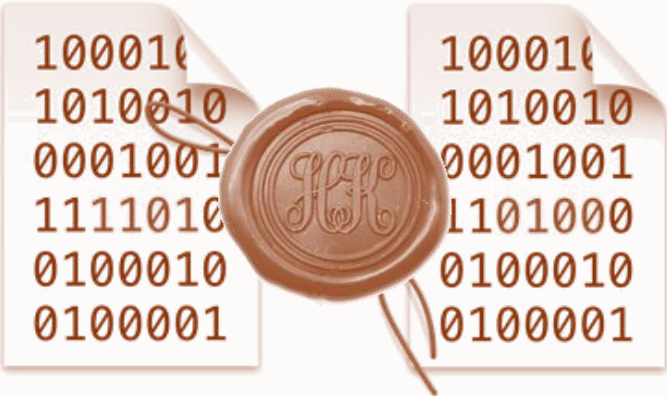
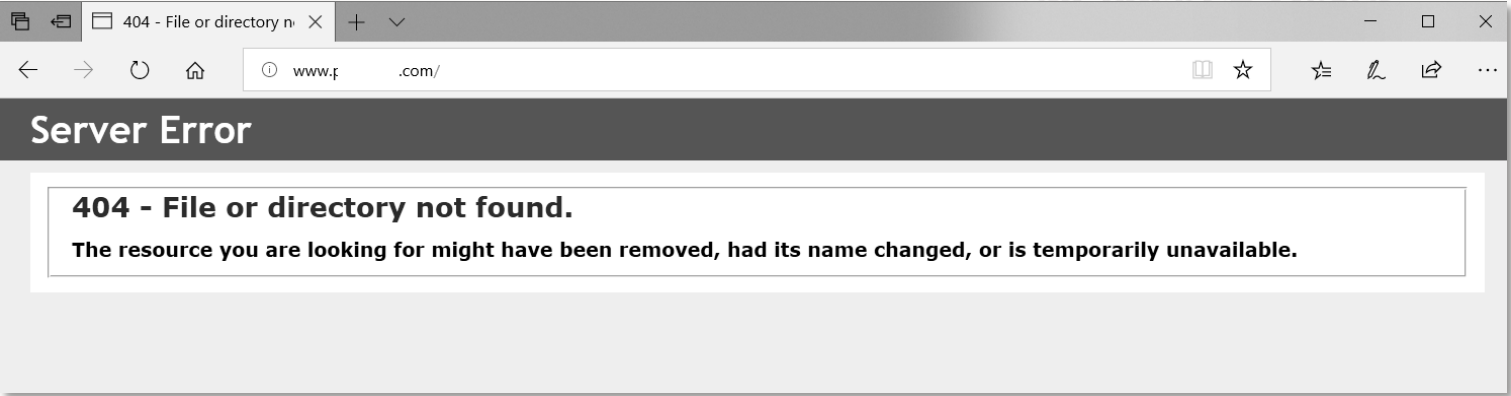
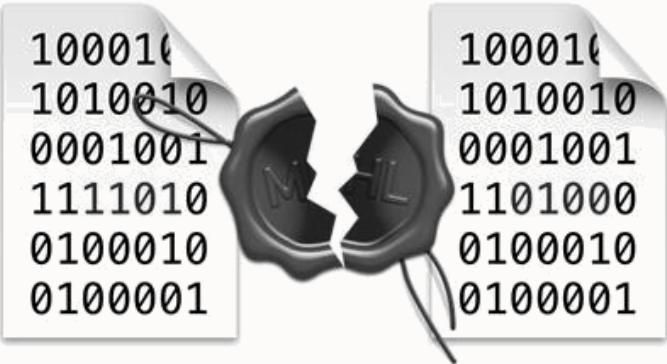
## 那为什么要在数据库中存文件？

- 数据库每一行遵循ACID规则(原子性、一致性、隔离性、持久性)
- 数据库提供文件与其元数据之间的数据完整性
- 默认提供数据库安全性
- 备份包含文件，不需要额外的文件系统管理
- 保存的数据较多时，数据库索引比文件系统的索引性能更好
- 文件删除/更新始终与行操作同步，不需要额外的维护

# Domino – 附件与正文



# 数据一致性





# Secure File Blob 字段全文检索



- 支持 TXT, Office Word, Excel, PowerPoint, PDF 等多种文件格式。
- 基于Text Search的索引
- 结构化数据或半结构化数据关联搜索
- 支持多种压缩模式, 可对历史数据压缩节省存储空间
- 压缩文件内容也能通过SQL直接查询
- 轻松自定义搜索文件内容的访问控制列表(ACL)
- 批量操作支持并行DML
- 通过机器学习获得洞察力
- 恢复或闪回数据(文件)到任何时间点

# Secure File Blob 经验证的 I/O 性能

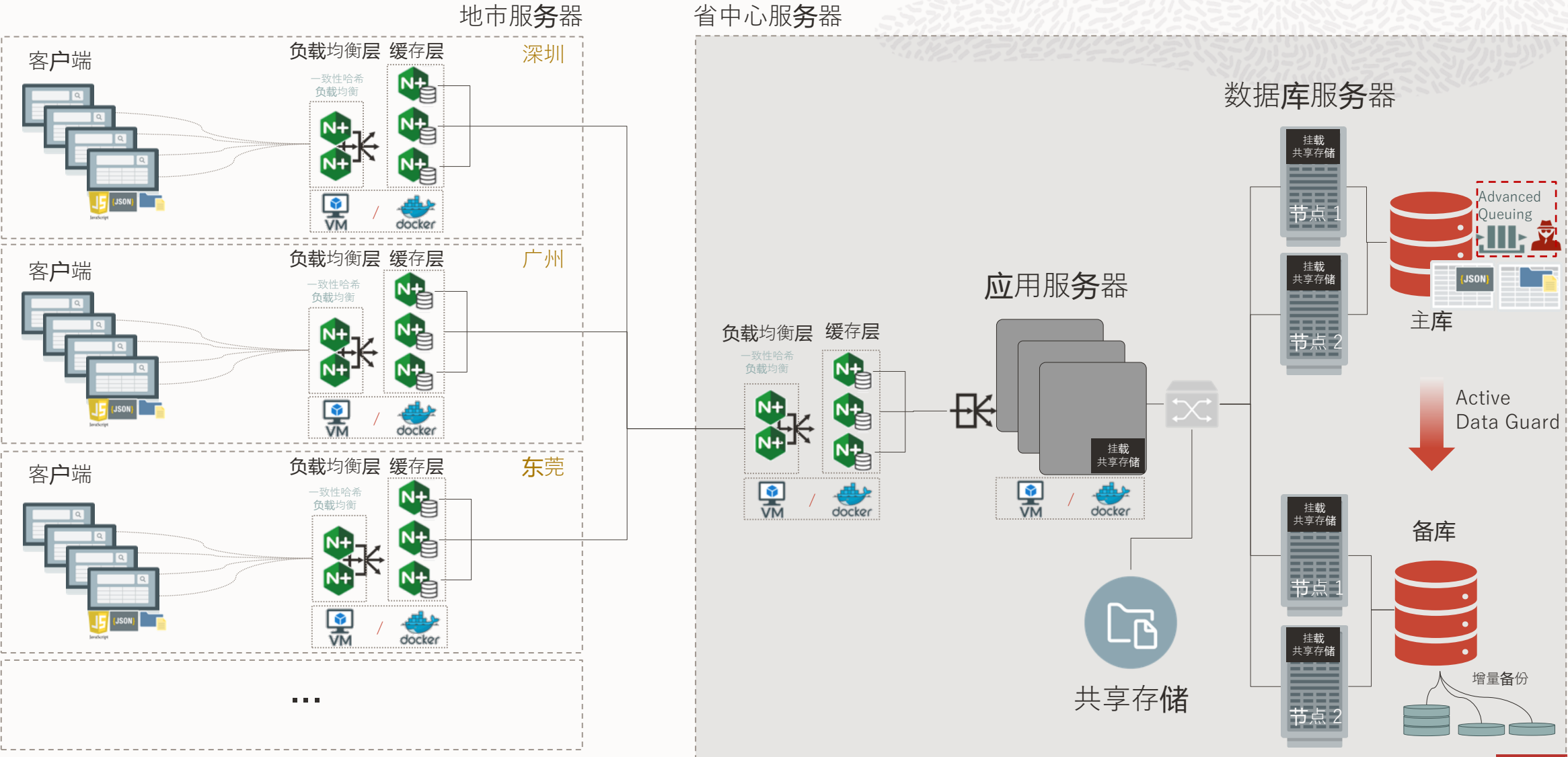
Description	CPU	Memory	Network
Compute Node	8 x Eighteen-Core Xeon® E7-8895 v3 Processors (2.80 GHz)	4 TB	refer X5-8 data sheet
Storage (full rack)	14 storage cells EF (Extreme Flash - all disks are flash cache)	Data sheet	Data Sheet

Total data inserted or selected is about **1 TB**

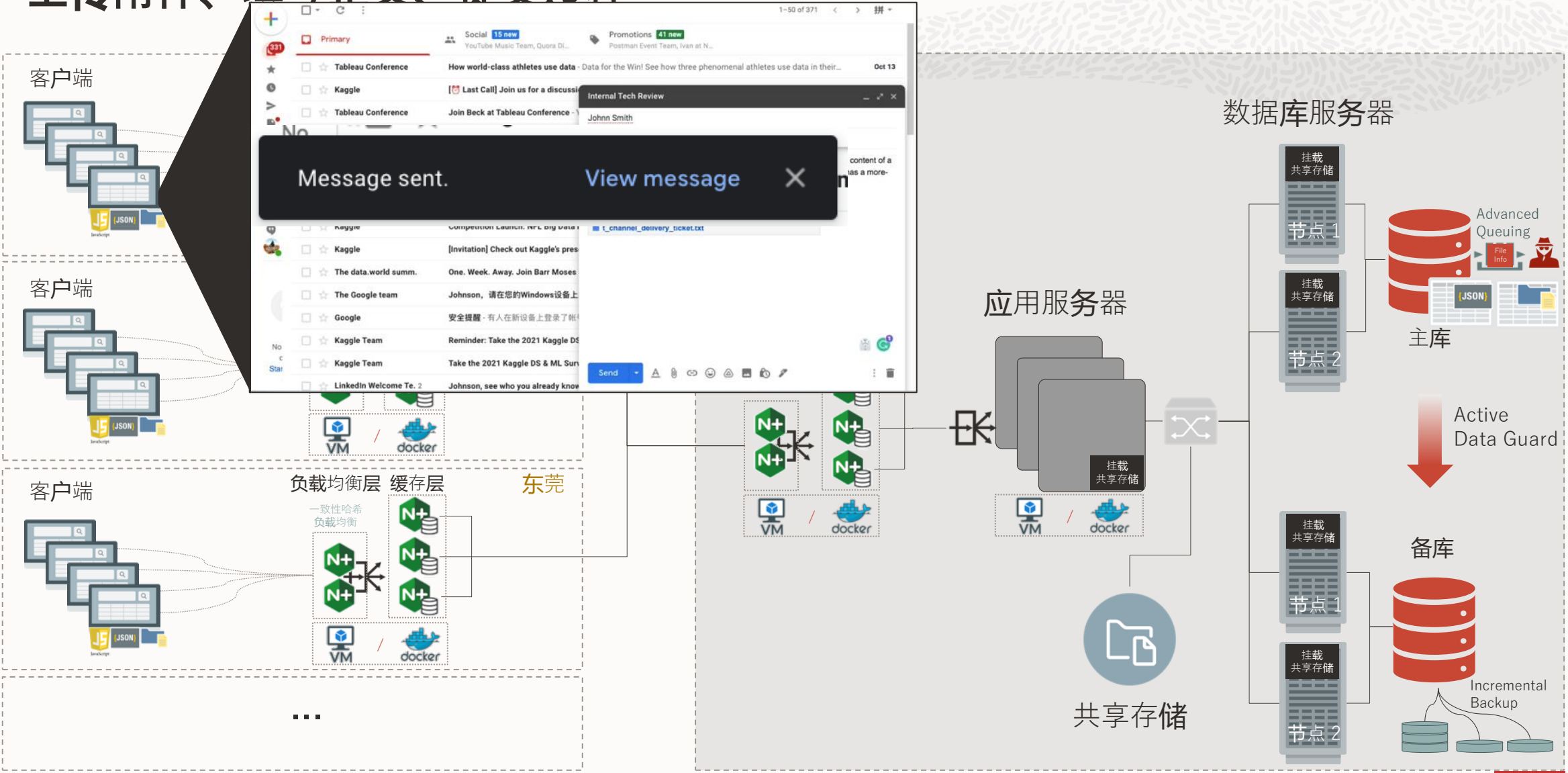
Test#	Description	LOB Size						Comment
		500K	1M	3M	5M	10M	50M	
64u-ins-1	<b>64 users - insert</b> rate GB/s	<a href="#">4</a>	<a href="#">4</a>	<a href="#">5</a>	<a href="#">5</a>	<a href="#">6</a>	<a href="#">6.5</a>	<b>No ARCH</b> logging
64u-sel-1	64 users - <b>select</b> rate GB/s	<a href="#">8.5</a>	<a href="#">13</a>	<a href="#">14</a>	<a href="#">14</a>	<a href="#">14</a>	<a href="#">15</a>	No ARCH logging
128u-ins-1	<b>128 users - insert</b> rate GB/s	<a href="#">4.5</a>	<a href="#">5</a>	<a href="#">6</a>	<a href="#">7</a>	<a href="#">7.5</a>	<a href="#">7.5</a>	No ARCH logging
128u-sel-1	128 users - <b>select</b> rate GB/s	<a href="#">8</a>	<a href="#">14</a>	<a href="#">15</a>	<a href="#">15</a>	<a href="#">15</a>	<a href="#">14</a>	No ARCH logging
128u-ins-2	<b>128 users - insert</b> rate GB/s	<a href="#">1.5</a>	<a href="#">1.5</a>	<a href="#">1.5</a>	<a href="#">1.5</a>	<a href="#">1.5</a>	<a href="#">1.5</a>	<b>ARCH</b> logging
128u-sel-2	128 users - <b>select</b> rate GB/s	<a href="#">8.5</a>	<a href="#">15</a>	<a href="#">15</a>	<a href="#">15</a>	<a href="#">15</a>	<a href="#">14</a>	ARCH logging



# 架构图 3

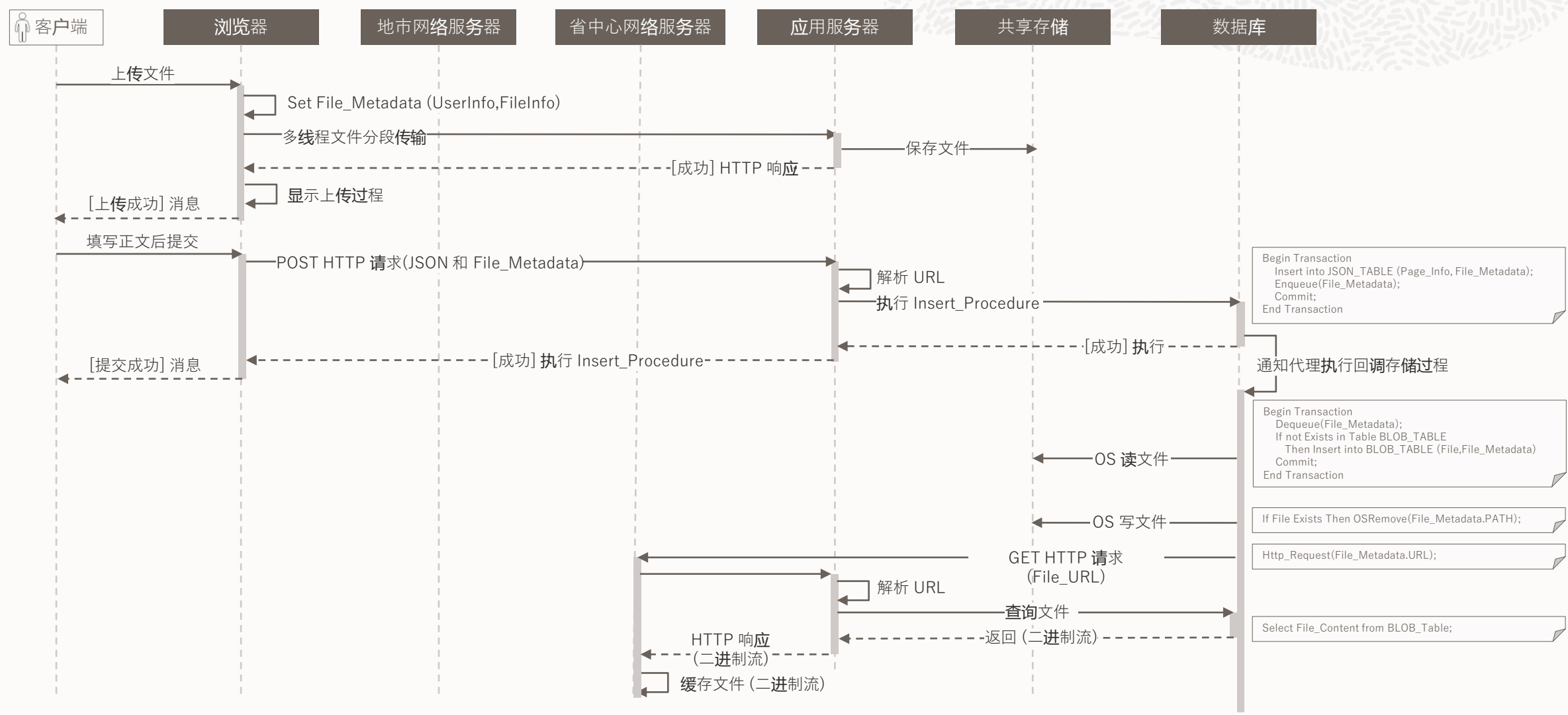


# 上传附件、填写正文、提交过程





# 上传附件、填写正文、提交过程



显示正文

Internal Tech Review

Inbox x



Johnson Kim <johnson.kim@...>

Oct 14, 2021, 10:49 PM (11 hours ago)



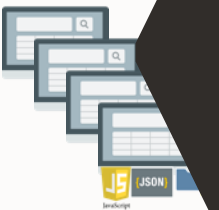
to me

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here,

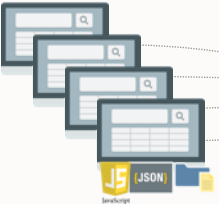
[Big Data & Analytics INTERNAL Stories.pptx](#)

[t\\_channel\\_delivery\\_ticket.txt](#)

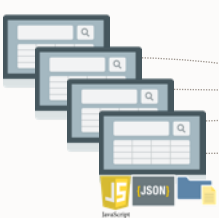
客户端



客户端

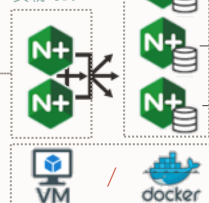


客户端



负载均衡层 缓存层

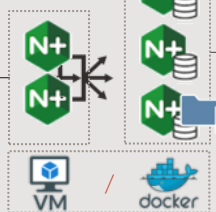
一致性哈希  
负载均衡



东莞

...

负载均衡

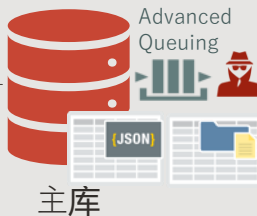


服务器



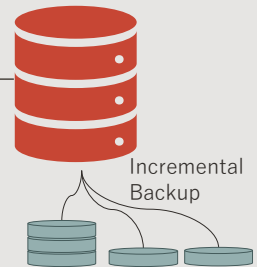
共享存储

数据库服务器

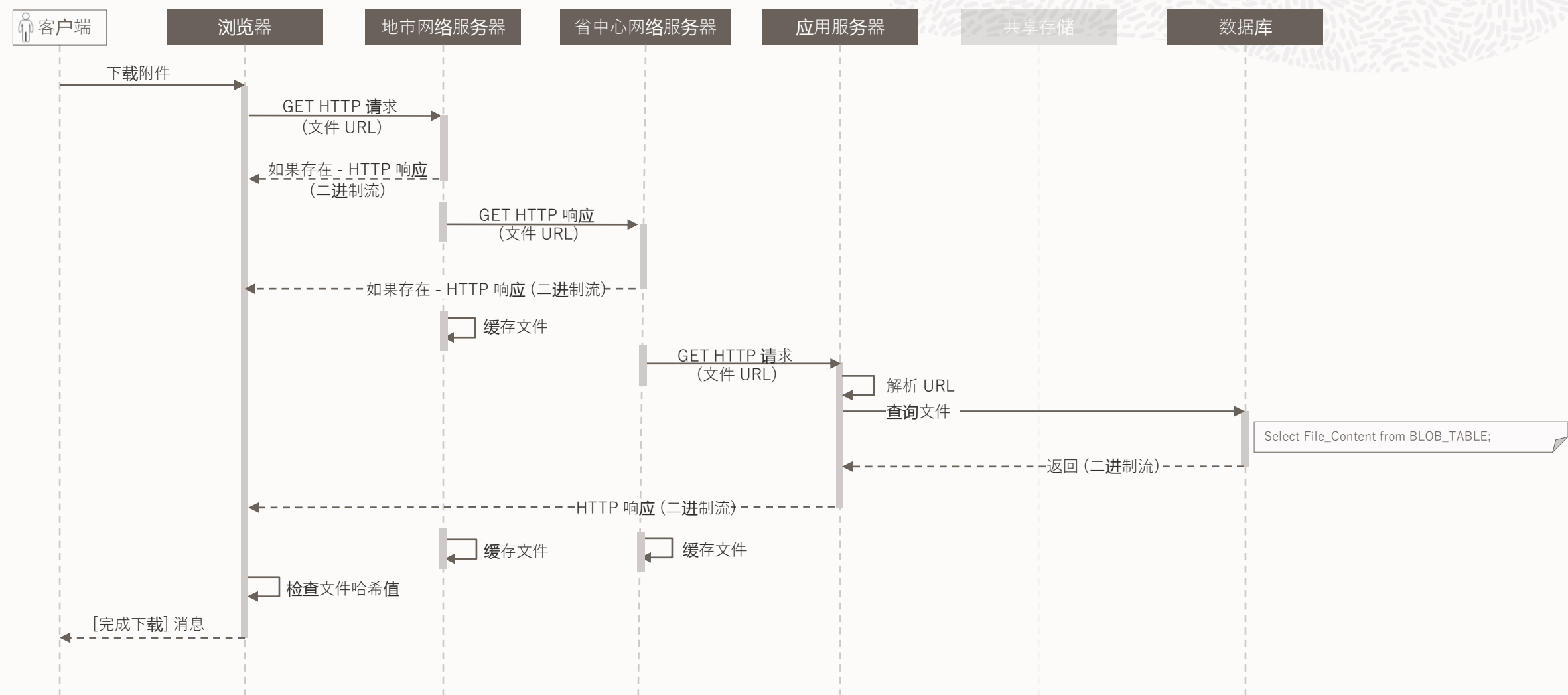


Active Data Guard

备库



# 下载附件过程

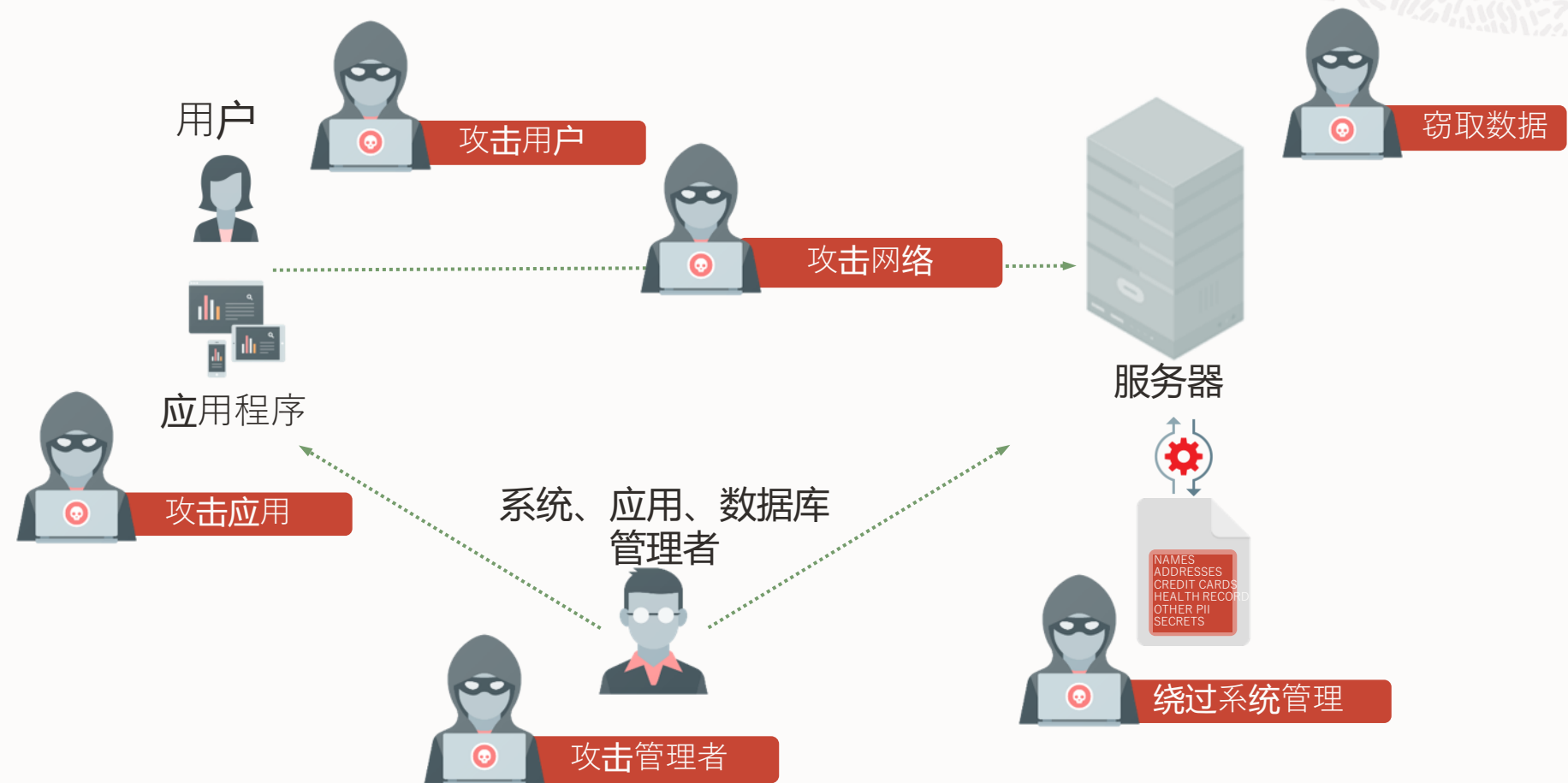




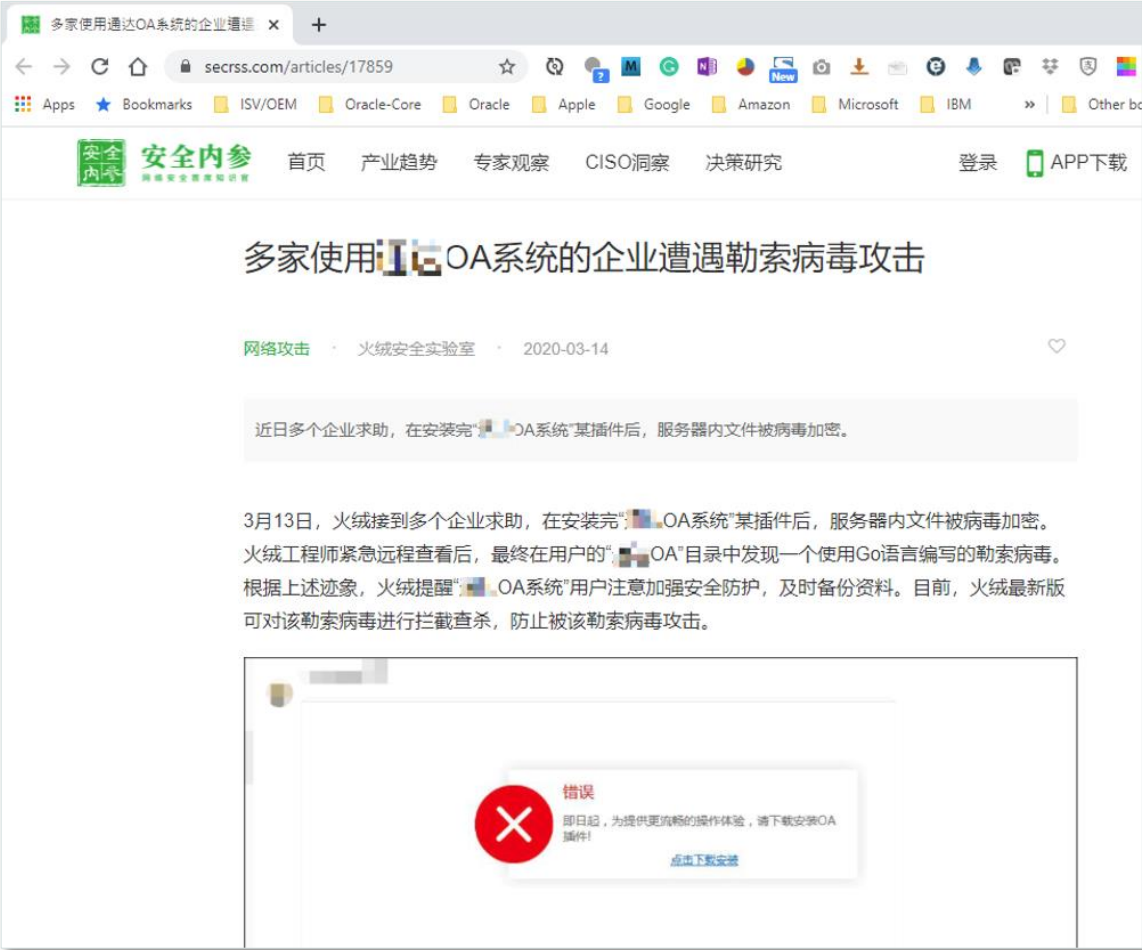
## 4. 数据安全性



# 黑客如何攻击？



# 真实案例 1 - 恶意软件、勒索软件

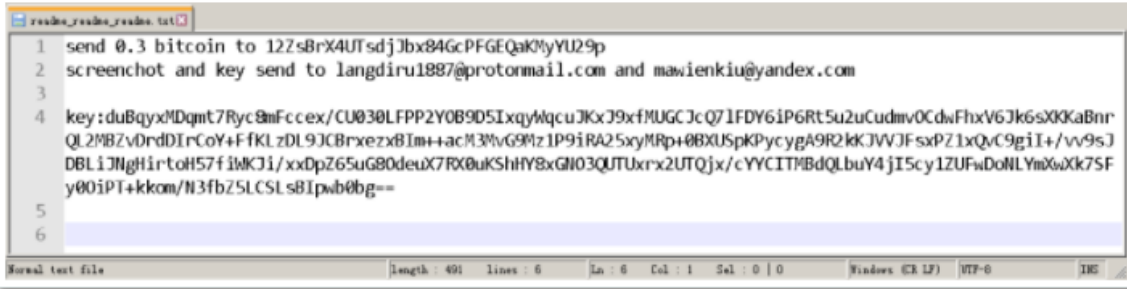


根据分析, 该勒索病毒在进入用户系统后会自动运行, 并会尝试结束mysql.exe进程, 再对.mdb、.sqlitedb、.doc、.docx、.xls、.xlsx、.ppt、.pptx等180种数据文件进行加密。

文件被加密后末尾被添加"1", 并会在桌面生成文件名为"readme\_readme\_readme.txt"的勒索信, 并索要0.3个比特币。

Chrysanthemum.jpg1	2020/3/13 13:58	1,146 KB	☆☆☆☆☆
Desert.jpg1	2020/3/13 13:58	1,102 KB	☆☆☆☆☆
Hydrangeas.jpg1	2020/3/13 13:58	776 KB	☆☆☆☆☆
Jellyfish.jpg1	2020/3/13 13:58	1,011 KB	☆☆☆☆☆
Koala.jpg1	2020/3/13 13:58	1,017 KB	☆☆☆☆☆
Lighthouse.jpg1	2020/3/13 13:58	731 KB	☆☆☆☆☆
Penguins.jpg1	2020/3/13 13:58	1,013 KB	☆☆☆☆☆
Tulips.jpg1	2020/3/13 13:58	809 KB	☆☆☆☆☆

图: 被加密的文件后缀名



<https://www.secrss.com/articles/17859>



# 真实案例 2 - 恶意脚本 (Web Shell)

图4的效果就是我们在消息内容中插入了一段简单的脚本代码：<iframe src=xx>。这段代码可以引入任意网址的内容到当前页面中，如果是网页木马的网址，那么后果可想而知。

我在一个办公室，给你发个消息，你只要一看该消息，你的电脑就莫名其妙的中毒了，这对于财务、人事这样的科室来说，可是巨大的损失。问题似乎不仅仅限于以上的这个漏洞，当我们翻看到绿叶OA系统提供的“网络硬盘”功能时，我们真的是吃惊了，如图5所示。



图5

请问各位这像不像是一个WebShell提供的文件管理界面，随意点击其中一个文件夹，如图6所示。

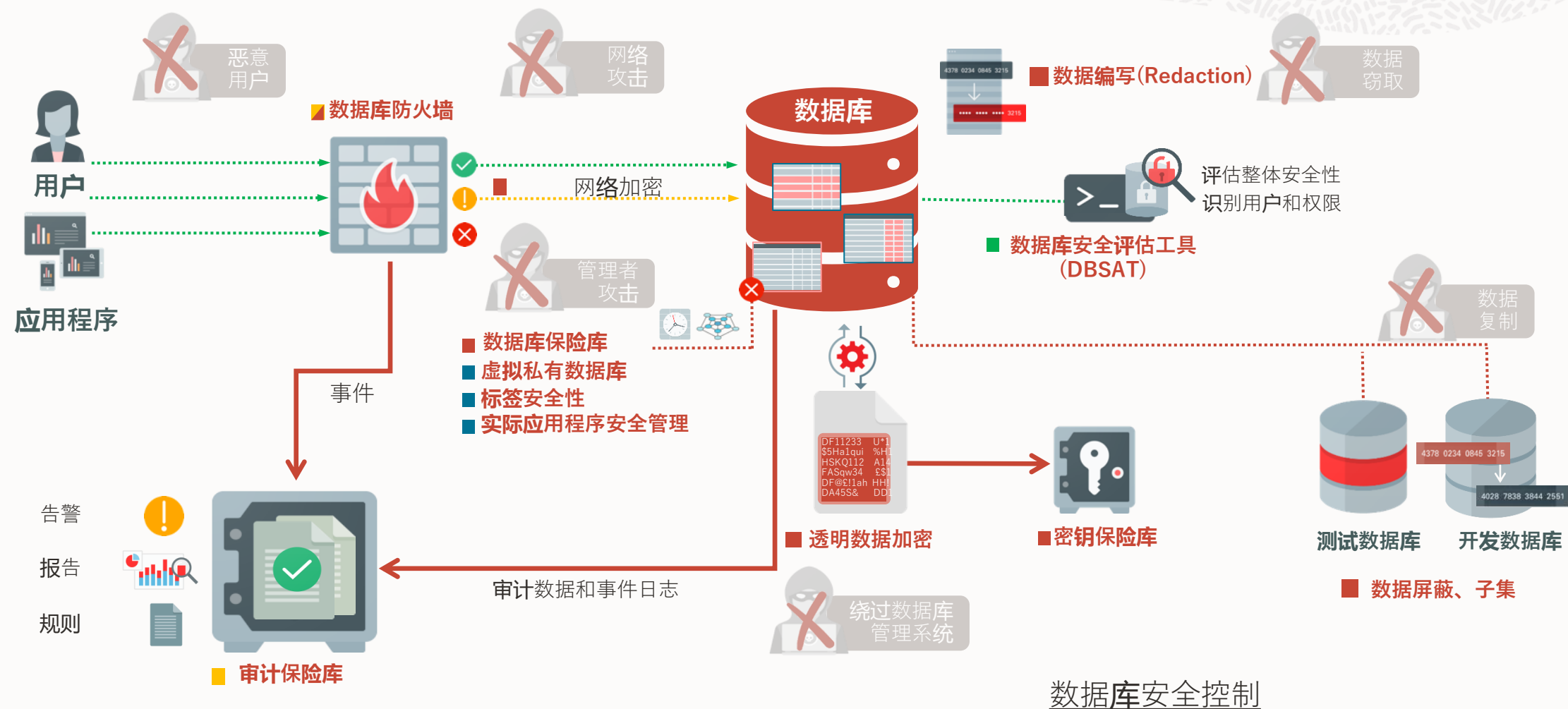
OA的其它漏洞这里就不再一一描述了，有兴趣的读者可以自己动手练习一下漏洞挖掘。本文旨在讨论技术，任何违法行为与作者和杂志无关。最后，在OA的官方网站上做了一下测试，效果如图7所示。



[https://www.sohu.com/a/201503388\\_765820](https://www.sohu.com/a/201503388_765820)



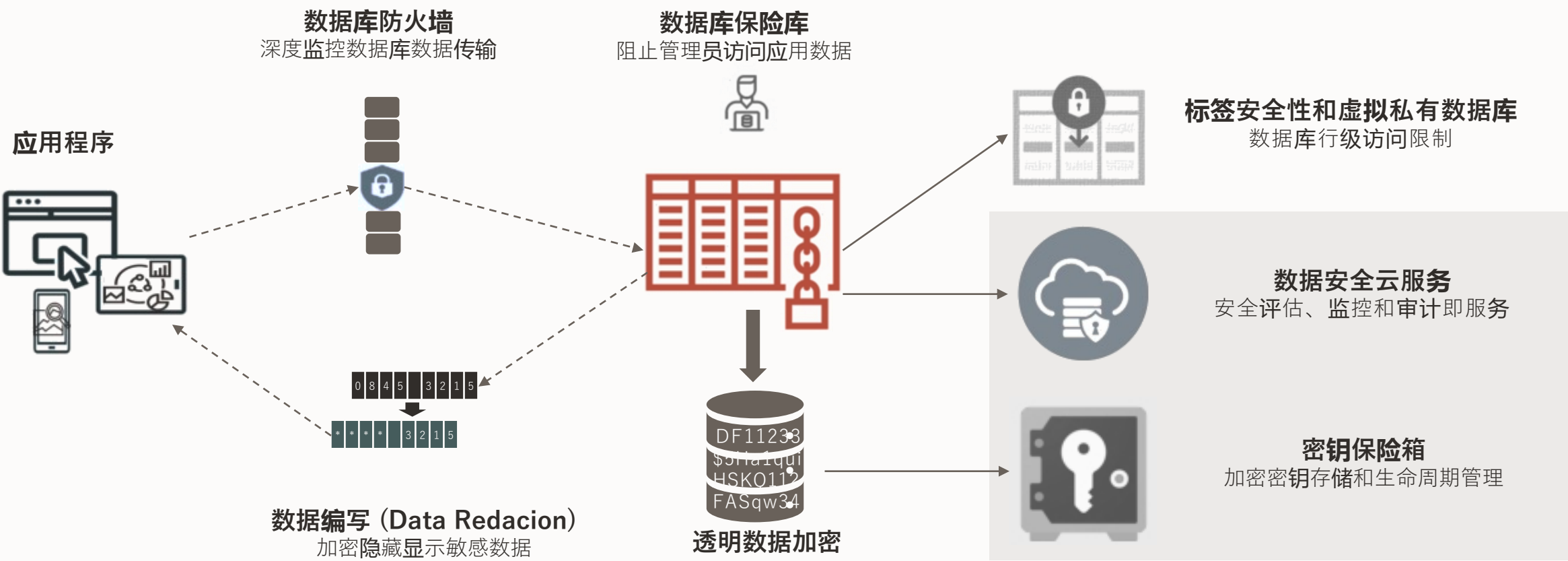
# Oracle MSA 架构





# 区块链增加到 Oracle MSA 架构

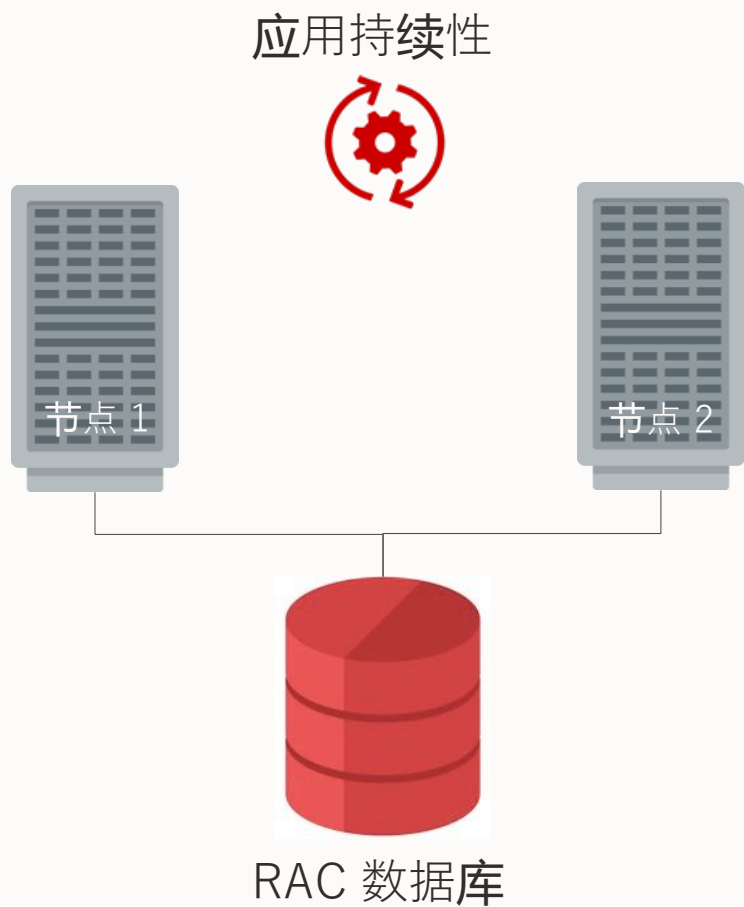
不可抵赖





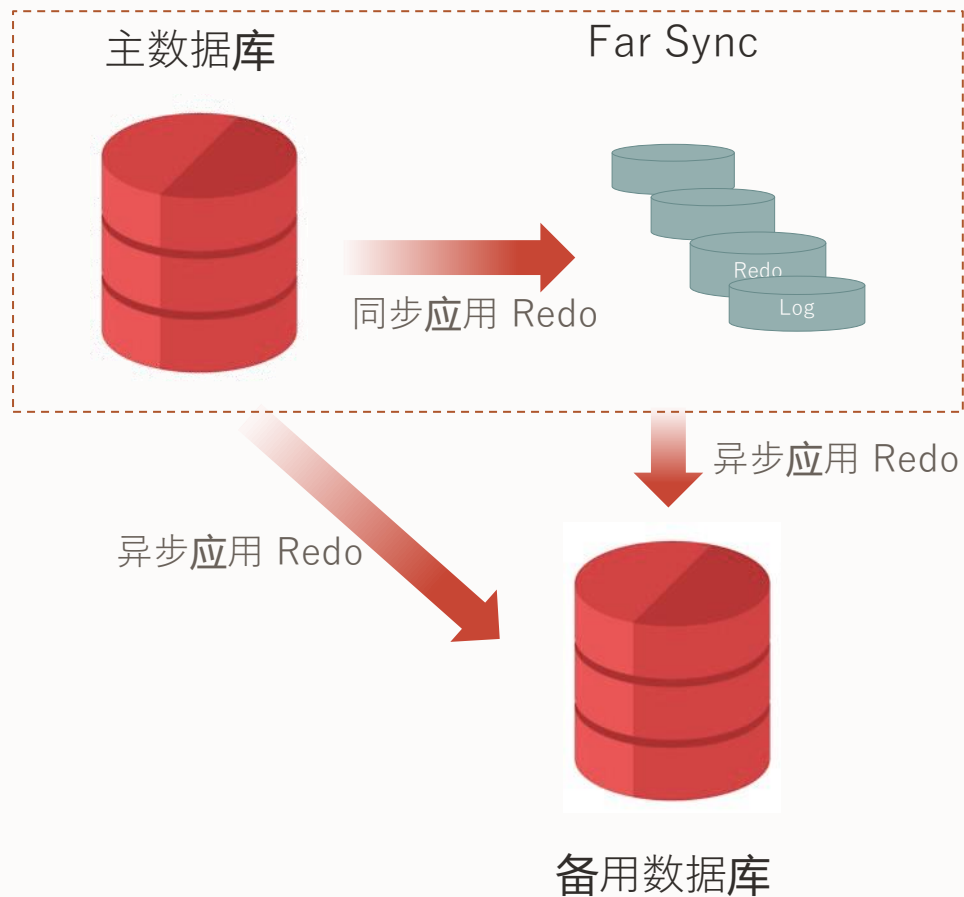
## 6. 高可靠性和灵活性

# Real Application Cluster (RAC) 解决方案



- 提供高可用架构，故障容错和无缝切换，将硬件和软件错误造成的影响最小化
- 多节点之间负载均衡，提高每秒交易数和连接数
- 通过多节点并行处理技术，减少事务响应时间，提高整体数据库服务器性能
- 滚动升级以及在线补丁，最大限度地减少或消除停机时间
- 应用程序连续性从应用程序的角度恢复不完整的请求，减少用户系统、通信、硬件故障引起的操作失败

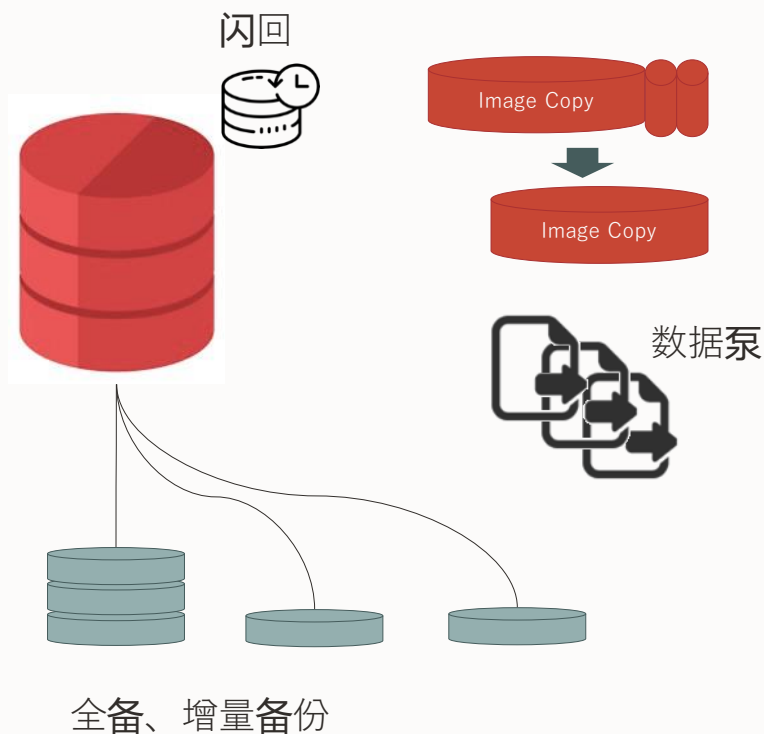
# Active Data Guard (ADG) 解决方案



- 高效、全面的灾难恢复和高可用解决方案
- 快速故障自动切换，减少计划中和计划外的中断所导致的停机时间
- 数据库备份负载卸载到备用数据库，降低生产库读写负载
- 将报告生成或部分查询业务放在备用数据库执行，降低生产库负载
- 利用备用数据库测试新的或现有的功能，不影响生产环境
- 利用 Far Sync 以极低的生产库负载实现零数据丢失保护
- 滚动升级打补丁最大限度地减少停机时间



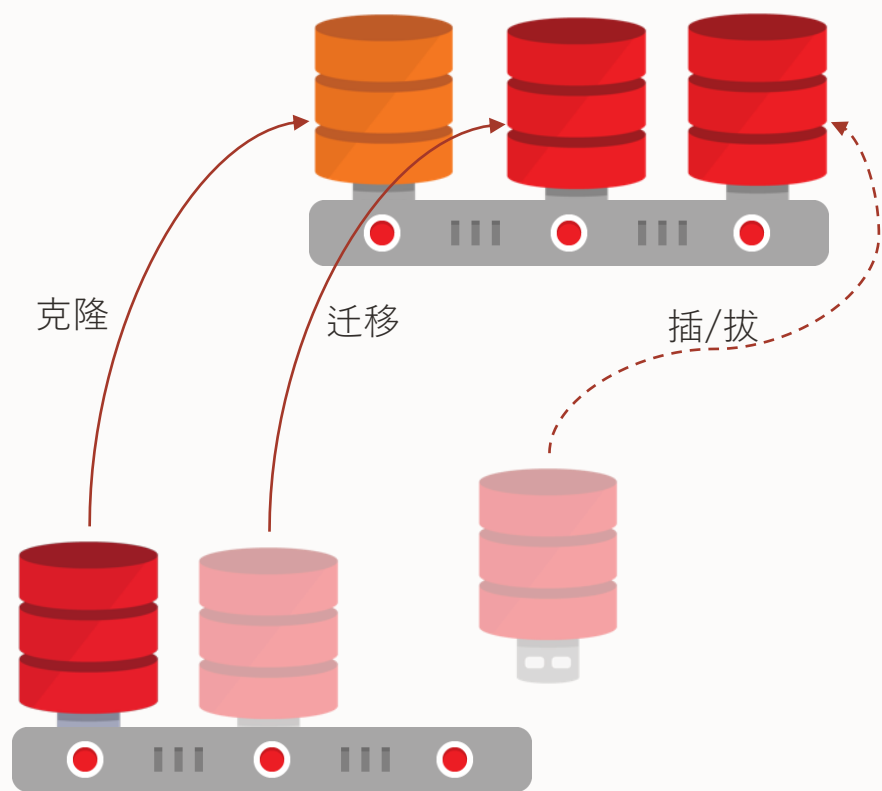
# 数据库备份恢复解决方案



- 支持物理备份和逻辑备份（RMAN, 闪回, 数据泵等）
- RMAN 备份加密, 防止数据盗窃
- RMAN 备份压缩减小备份集的大小, 节省存储空间
- RMAN 并行执行提高备份和恢复性能, 减少操作时间
- 块级、表级以及数据库级等多种级别恢复模式, 减少恢复成本
- 通过增量更新备份前滚数据文件镜像副本, 快速生成最新镜像副本
- 灵活的闪回查询、表数据、数据库, 无需宕机服务器



# 多租户架构解决方案



- 无需修改应用，传统架构经简单命令轻松转换为多租户架构
- 基于克隆快照，可快速部署开发测试环境
- 高效共享CPU、内存等资源，减少服务器数量降低硬件支出
- 通过资源管理器统一分配管理资源，方便优化性能、打补丁较少运维成本
- 支持多种备份模式，支持CDB级别和PDB级别单独备份恢复
- 可插拔特性实现数据库迁移更加高效便捷



## 7. 客户受益

# 经验证的 Oracle 产品路线图和解决方案



## • 业务

- I. 改善用户体验和提高业务效率
- II. 基于灵活工作流的业务敏捷性
- III. 节约成本

## • 技术

- I. 基于主流技术
- II. 易于理解, 实现简单, 可伸缩
- III. 高可用, 高安全性
- IV. 方便维护、测试、备份
- V. Exa路线图中已验证的通用x86服务器特性
- VI. 全文检索, 附件内容挖掘
- VII. 任务、文本和附件的高一致性





# Q & A



Our mission is to help people see  
data in new ways, discover  
insights,  
unlock endless possibilities.

