

1.分布式系统概论

1.概念和特点

1.1 概念

一个分布式系统是若干个独立的计算机的集合, 但是对该系统的用户来说, 感觉该 系统就像一台计算机一样。

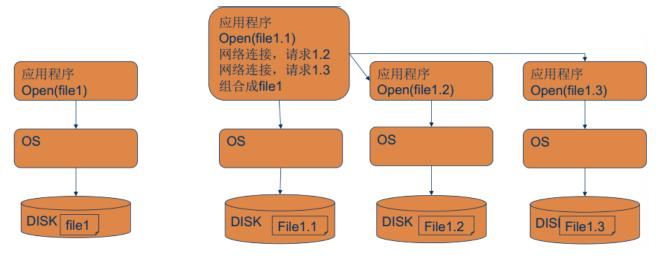
A distributed system is a collection of autonomous computing elements that appears to its users as a single coherent system.

1.2 需解决问题

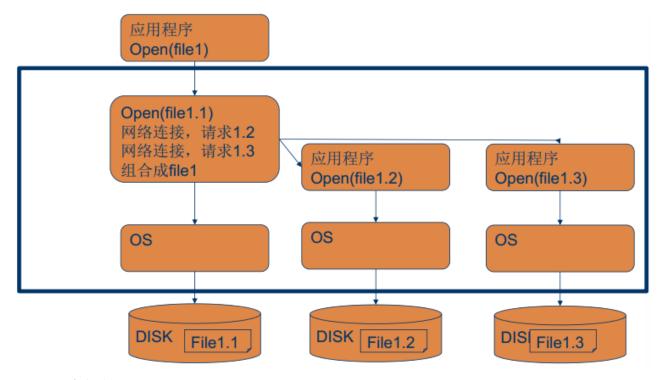
- 1. 时钟同步问题(射击游戏)
- 2. 容错性 (打印机)
- 3. 文件定位

1.3 中间件

无中间件系统



存在中间件系统



阿里云中间件:

- 分布式关系型数据库 DRDS
- 企业级分布式应用服务 EDAS
- 全局事物服务 GTS
- 性能测试服务 PTS
- 消息队列 MQ
- 云总线服务 CSB

1.4 分布式系统特点

优点: 性价比高、可扩展性强、可以应对应用的固有的分布性、可靠性高

缺点: 复杂性、受到网络限制、安全性和可信性

2.设计目标

2.1 支持资源共享

2.2 透明性

• 访问透明性: 用户不必知道数据表示和访问方法

• 位置透明性: 用户不必知道资源位于何处

• 迁移透明性: 资源可以不改名, 而随意移动

• 重定位透明性: 资源可在使用中移动位置

• 复制透明性: 用户不必知道有多少个备份存在

• 并发透明性: 多个用户可以自动地共享资源

• 故障透明性: 用户不必了解故障恢复处理

2.3 开放性

通过一系列的接口来提供服务, 而这些接口的定义是符合某个标准的。

(1) Open System

- 1. 它可以提供组件为其他系统所用;
- 2. 它经常会用到其他系统提供的组件。

(2) 特点

• 互操作性: 不同系统依靠共同的标准, 可以依赖彼此的组件共同工作

• 可移植性: 应用程序可在具备相同接口的、不同分布式系统上运行

• 灵活性: 可配置不同开发者的组件

• 可扩展性: 可增减组件

2.4 可伸缩性

当系统的运行环境或者需求改变的时候, 能够不影响系统正常运行的能力。

(1) 规模可伸缩性

当用户数量、 资源数量改变的时候不影响系统的使用

(2) 地理位置可伸缩性

当用户、 资源距离很远的时候不影响使用。

• 原因 1: 在同步通信时, 由于距离带来的网络延时。

• 原因 2: 在广域网中通信可靠性不如局域网。

• 原因 3: 在广域网中多播手段较少。

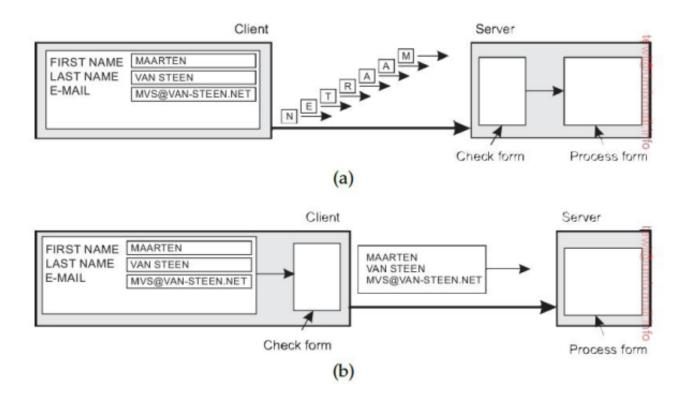
(3) 管理可伸缩性

使分布式系统能够方便的在多个管理域中使用。

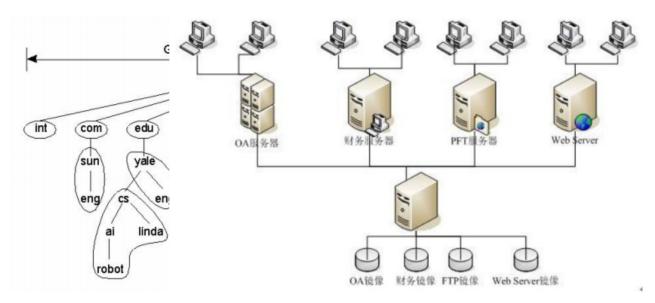
(4) 实现可伸缩性采用的技术

减少通信延迟

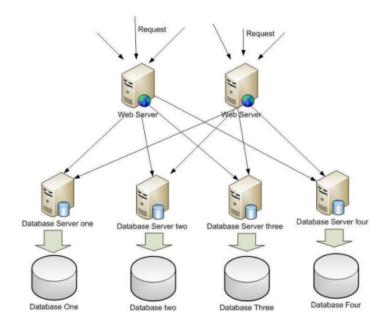
在 web 应用中,将表单校验的工作放在客户端来做。

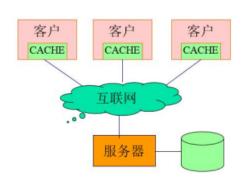


将服务分割并分布化部署



副本 (replication) 与 缓存 (caching)





(5) 易犯错误

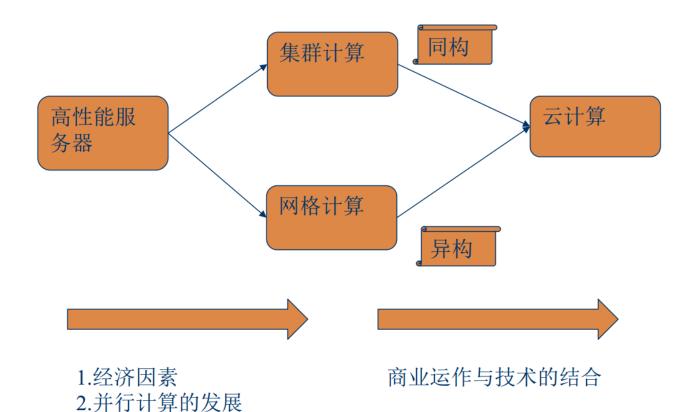
- 网络事可靠的
- 网络事安全的
- 网络是同构的
- 网络拓扑不会改变
- 没有网络延迟
- 带宽无限
- 没有复制成本
- 会有一个系统管理员

2.5 高性能

3.类型

3.1 高性能计算分布式系统

以提高性能为目标



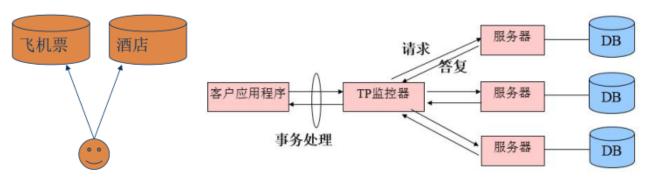
3.2 分布式信息管理系统

以满足某种特定的业务逻辑为目标 如事务管理系统、企业应用集成(EAI)

(1) 事务管理系统

事务的 ACID 性质: 原子性、一致性、隔离性和持久性

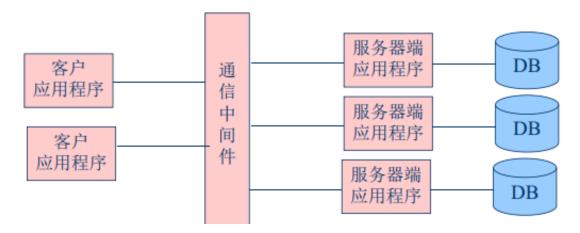
分布式或嵌套式事务: 一个全局事务加上多个子事务组成



(2) 企业应用集成(EAI)

对应用程序进行集成, 使其完成某一个业务

这个业务的完成, 需要各个程序进行协作



3.3 分布式普适计算系统

(1) 普适计算设备

移动设备, 如手机

(2) 普适应用系统的需求【2004】

- 接受上下文环境变化
- 鼓励即席(Ad hoc)组合
- 认可共享是缺省值

(3) 普适应用系统的需求【2009】

- Distribution 设备是分布的
- Interaction 人与设备的交互不张扬
- Context awareness 上下文感知
- Autonomy 高度自治
- Intelligence 能自动处理大量事务