

建立高可用集群邮件系统

沈朝阳

<http://www.mailer.com.cn/>

北京春笛信息技术有限公司

为什么要用E-Mail?

E-Mail已经成为人们生活的一部分

廉价、快捷、信息量大

E-Mail 方便易用

标准的协议：smtp, pop3, imap, mime, PGP, ldap

各种平台下都有相应的开发工具和API

大容量邮件集群系统可以保证大用户量下的高效率和7*24高可用性

目标

用普通的硬件(x86, IA64)设备构建大容量、分布式、可扩展邮件系统

从三个方面实现高度可扩展性...

- ▶ **性能:** 随着集群数量线性增长
- ▶ **易管理性:** 高度自动化、智能化
- ▶ **高可用性:** 平均无故障率 : 99.99%

常规的邮件系统集群方案

静态功能分割

性能问题:

不能实现动态负载均衡

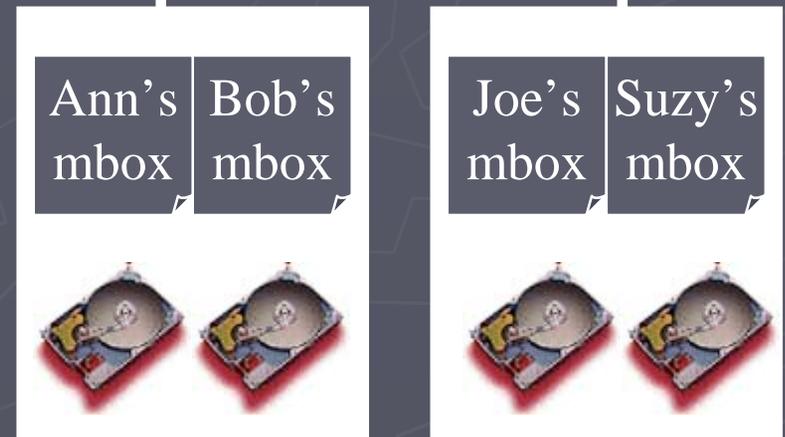
管理问题:

静态磁盘分区, 不能动态扩展

可用性问题:

有限的故障冗余措施

SMTP/IMAP/POP



NFS servers

金笛-高可用集群邮件系统

概要



技术架构

原理和技术要点

基本操作和数据结构

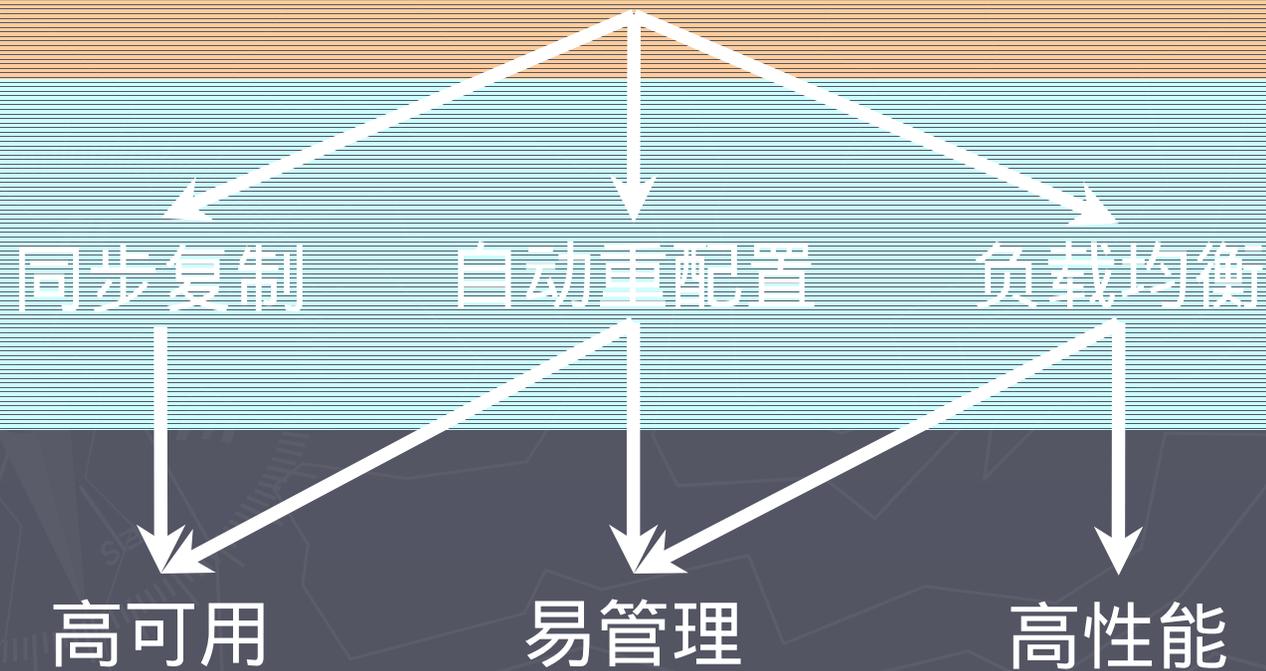
优势

解决方案

结论

关键技术

网格Grid、节点Node
“任意节点可完成任意功能”



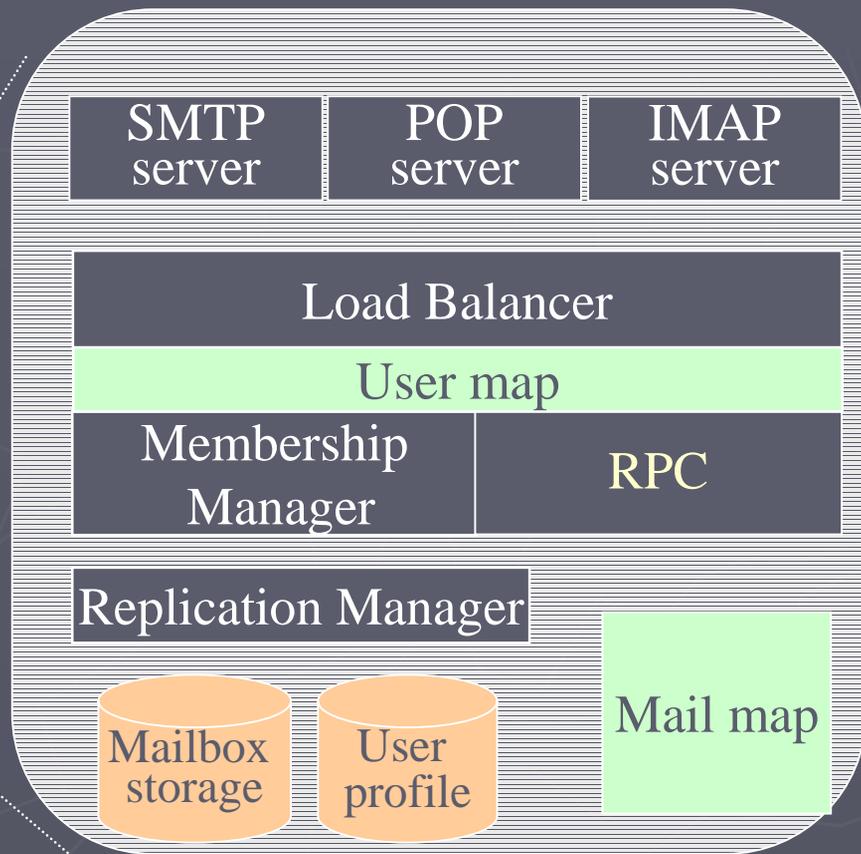
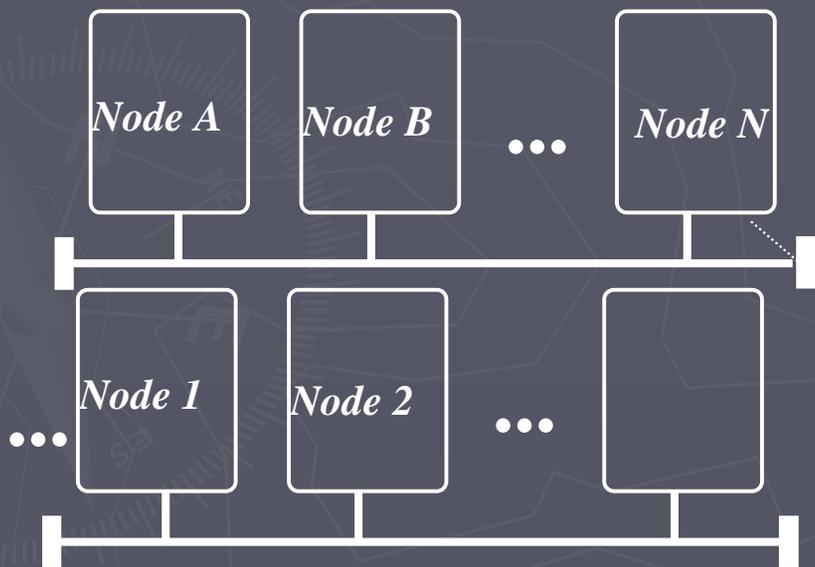
框架

技术

目标

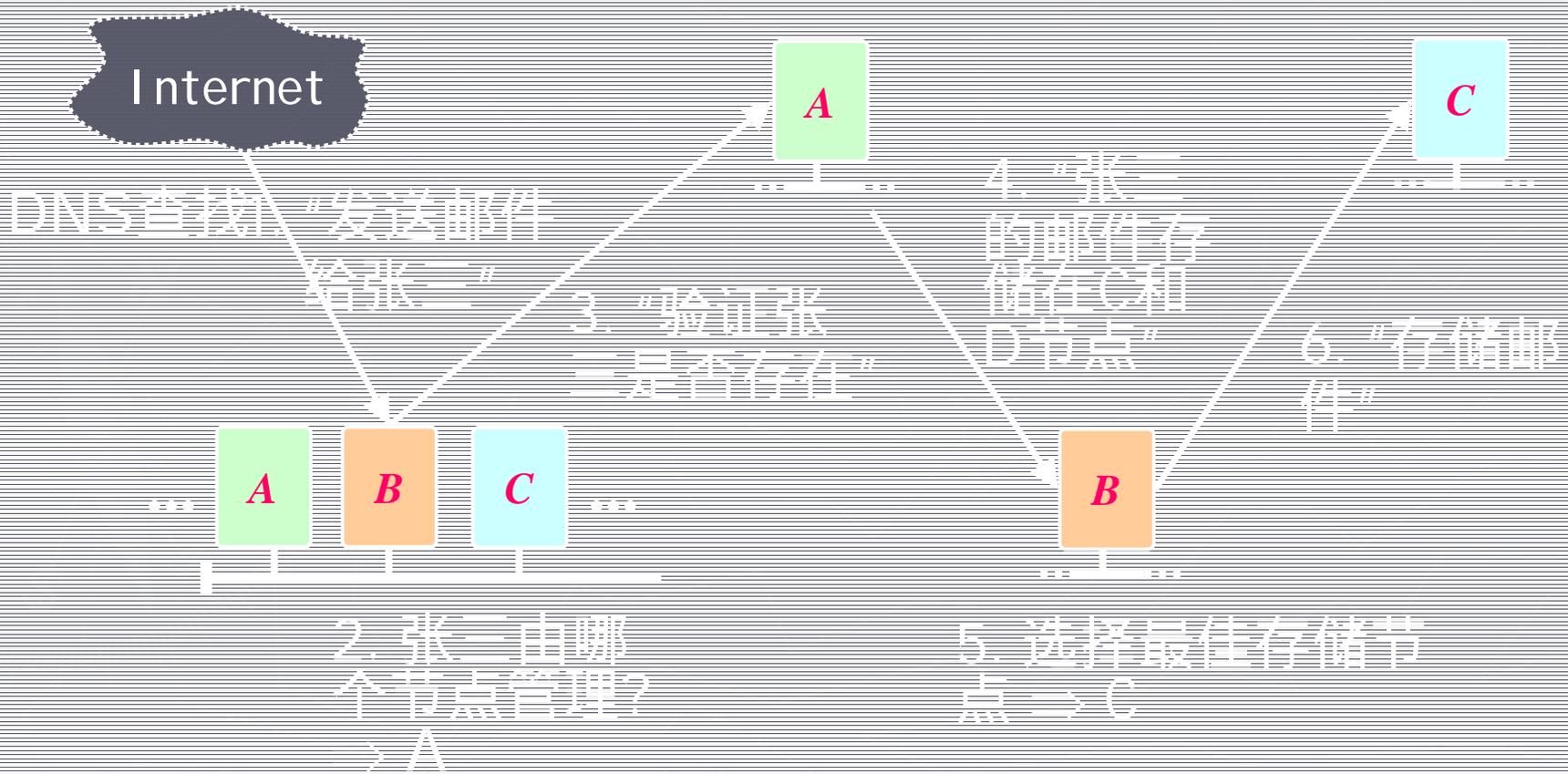
技术架构

每一个节点都是完整的
邮件系统单元，所有的
单元构成**邮件阵列**

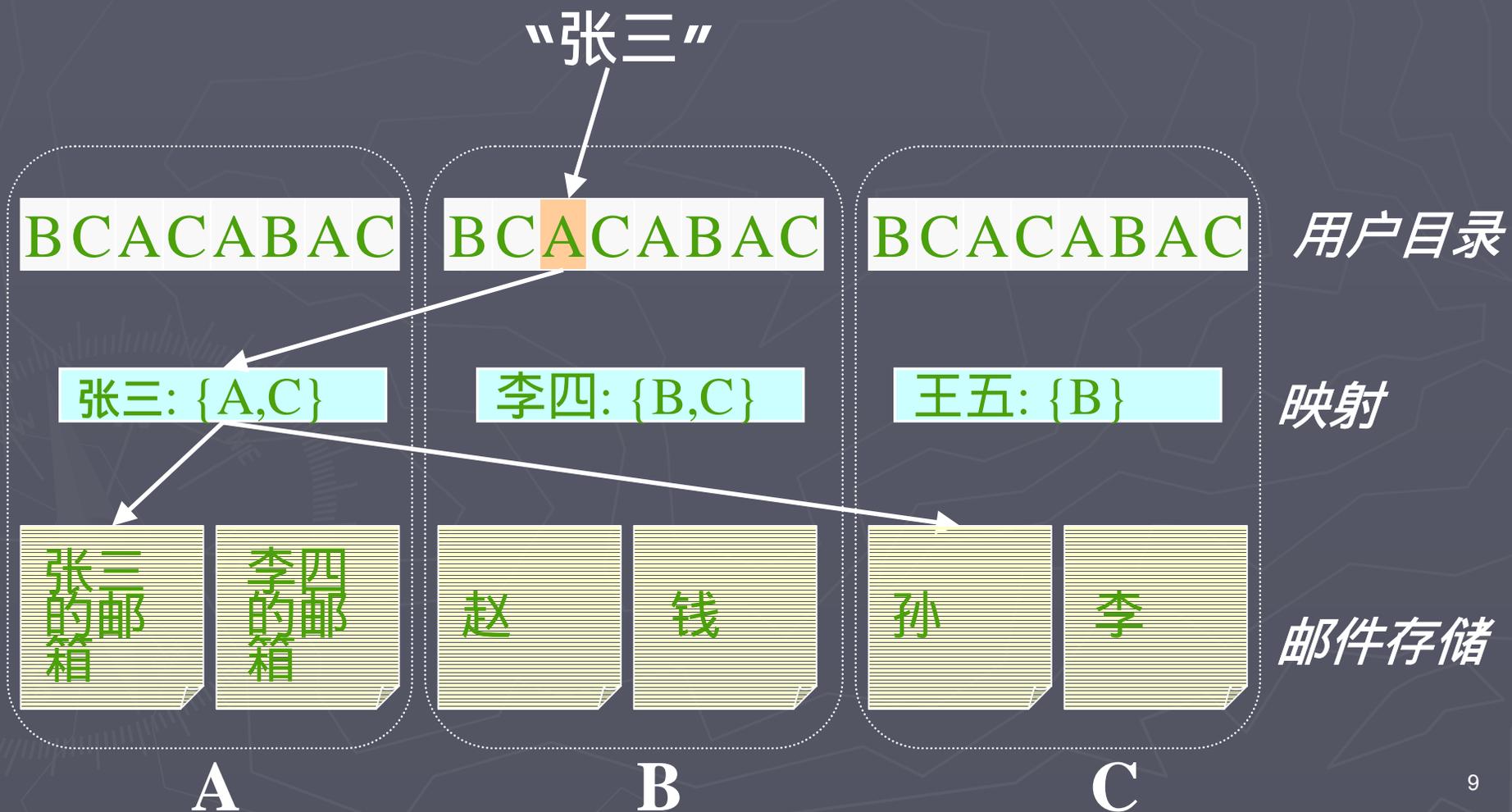


邮件投递流程

协议层数据接收 \Rightarrow 查找用户 \Rightarrow 均衡负载 \Rightarrow 消息存储



基本数据结构



技术架构优势

优势:

最佳资源优化、利用

自动配置

在节点失效/恢复基础上的任务再分配

动态均衡负载

结果:

高可用性

易管理

高性能

性能

目标:

集群节点数量增加，性能也线性增长

策略:

避免出现数据流量瓶颈

数据分区在节点之间统一部署

数据在动态增长时动态扩展

运行环境

30 节点构成的 PC 集群

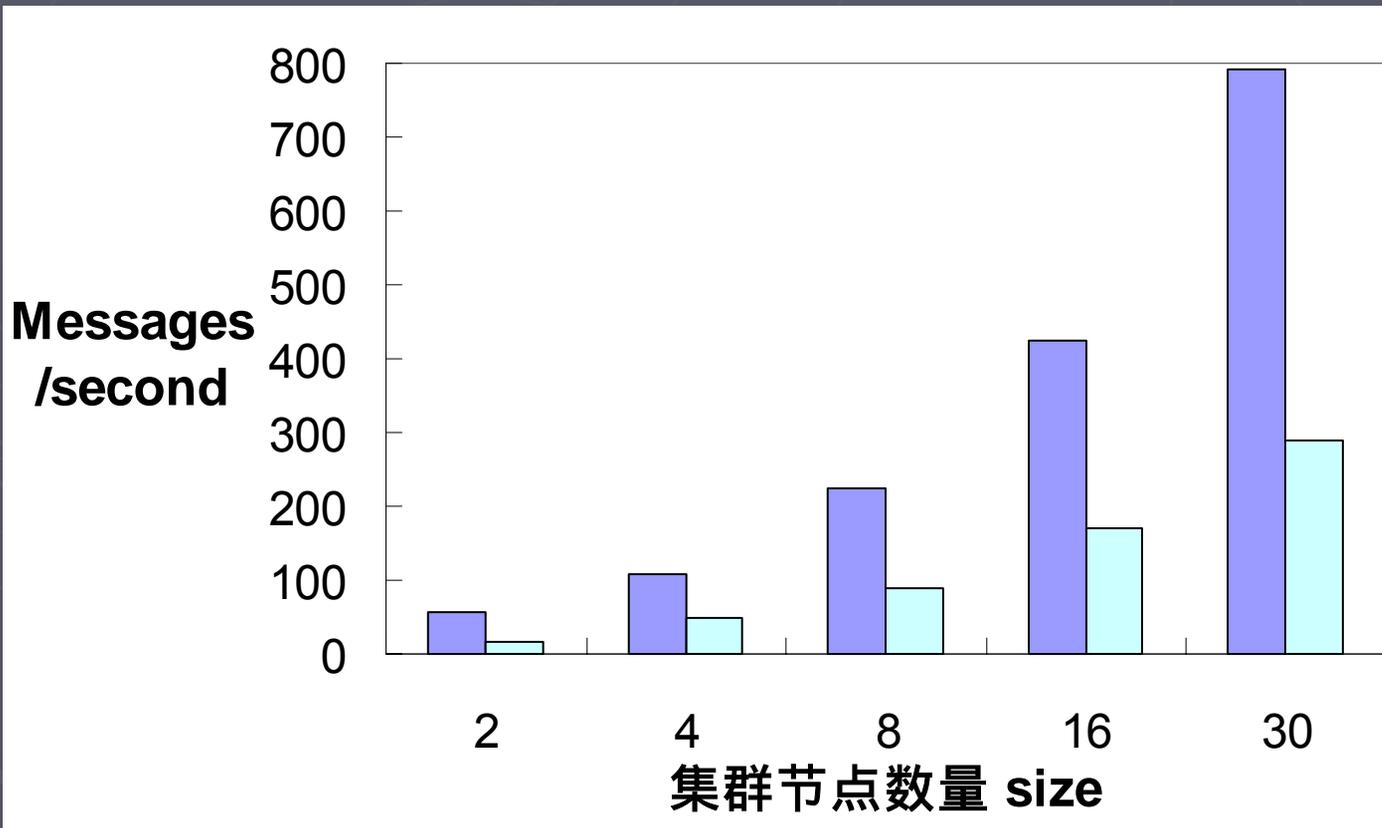
100Mb/s Ethernet + 1Gb/s hubs

Linux 2.4.10

综合比较

与 sendmail+popd 性能比较

金笛集群邮件系统性能如何?



Jindi-cluster

Sendmail+pop

高可用性

目标:

故障自动恢复、自动隔离
集群节点部署快捷方便
性能平滑提升

策略: 从硬件、软件两方面保证

硬件: 邮件数据email messages,
用户数据user profiles,

⇒ *Raid, SAN, NAS*

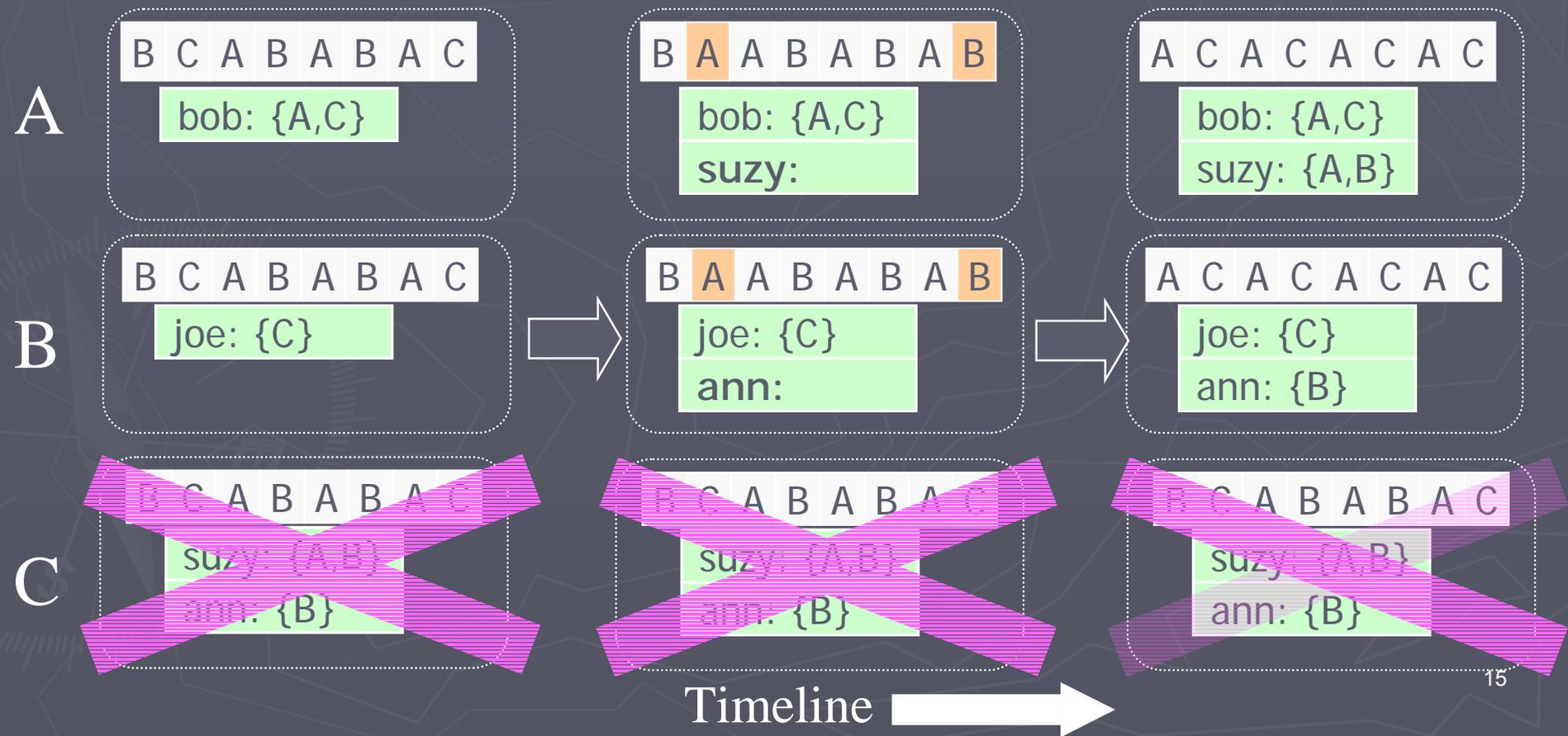
软件: 用户映射, 邮件映射

⇒ 自动重构

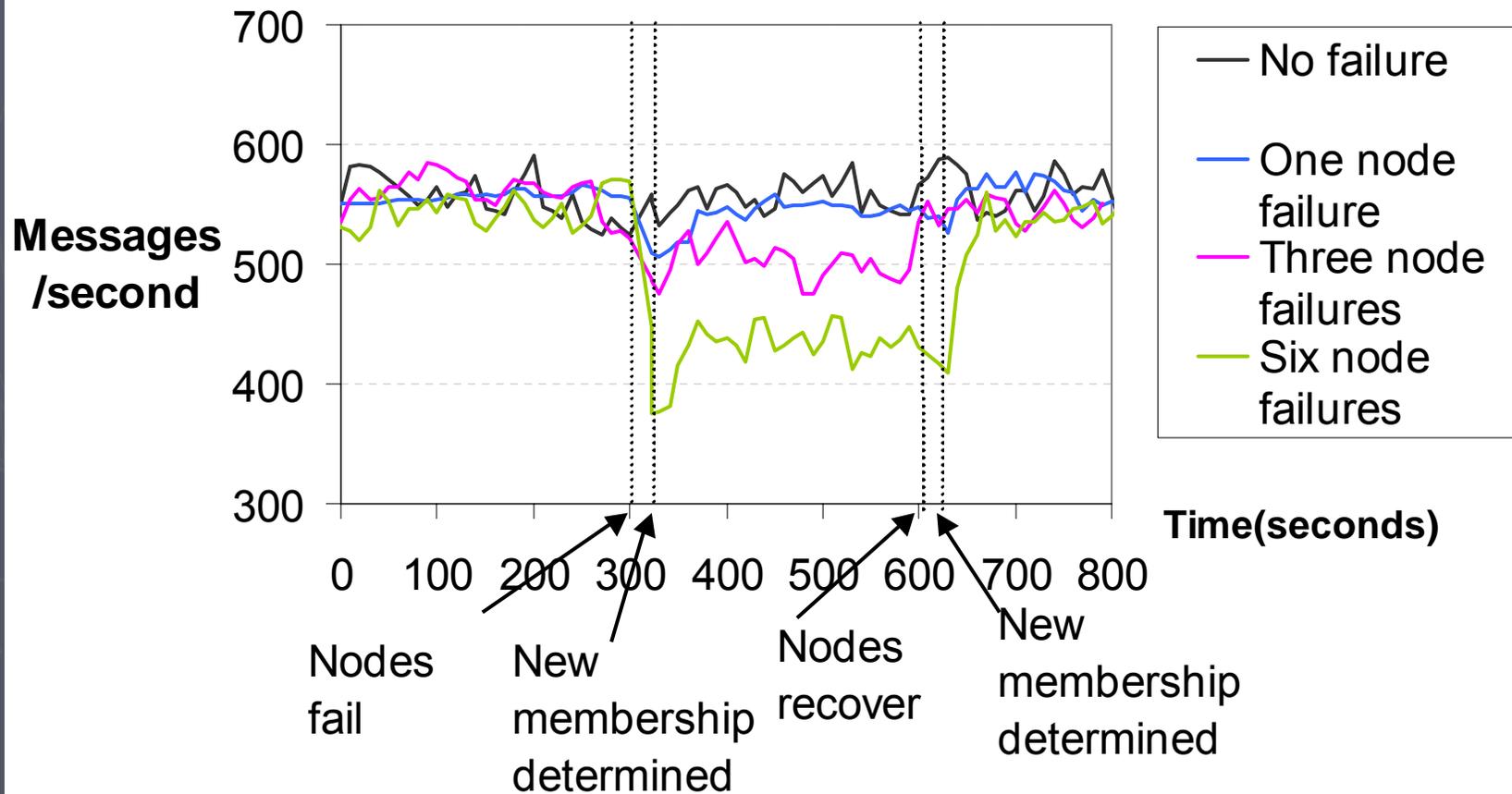
故障及恢复

1. 用户数据映射自动转移

2. 用户数据扫描



邮件系统监控



数据同步镜像

目标:

发生硬件故障数据很容易恢复
应付突发事件

策略: 充分利用互联网技术手段

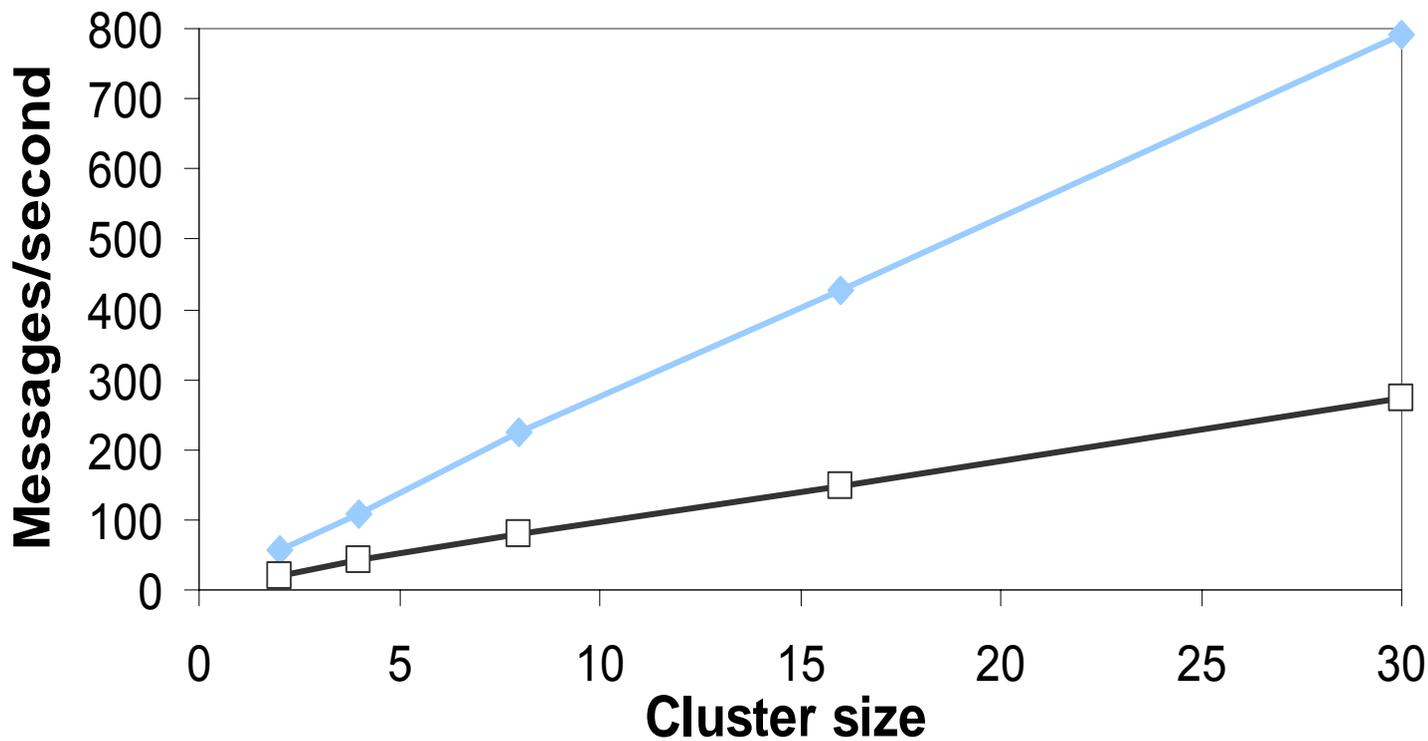
所有数据都要同步复制

每个邮件, 每个用户信息复制

不需要管理员干预

不需要太多的专业知识

数据同步复制的效率



Jindi-mail

负载均衡:目标命中策略

目标:

负载平均分配

支持异构硬件

无须复杂的配置

结论

高性能, 高可用, 易管理是构建集群邮件系统主要目标

构建集群邮件系统需要遵循以下几个原则：

计算节点功能最小化、一致化

自动动态重构

同步数据复制

动态负载均衡

谢谢！

