

# 分布共享数据服务理论与技术研究

魏恒峰

导师: 吕建 黄宇

南京大学软件所

July 12, 2016

# 分布共享数据服务理论与技术研究

1 研究背景

2 研究问题

3 研究方法

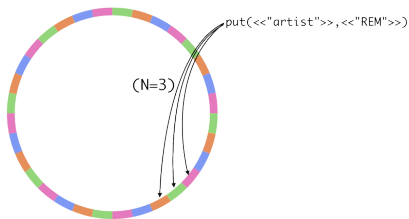
# 分布式应用



TODO: 动画: 分布部署

# 分布数据

(**TODO:** 动画: partition + replication)



distributed data : partition & replication

# 分布数据典型应用 (I)



图: 分布式存储系统 (开源 [左] & 商用 [右]).

应用需求 [Facebook@OSDI'10] vs. “分布数据”:

低延迟: 就近访问副本数据

高可用性, 高容错性: 备份容灾

# 分布数据典型应用 (II)

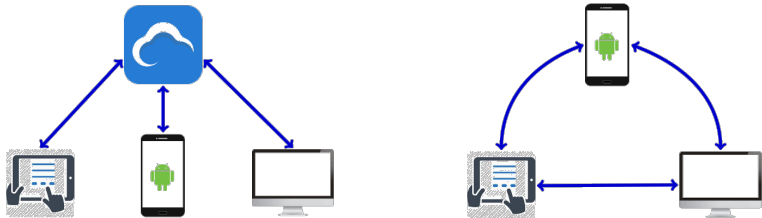


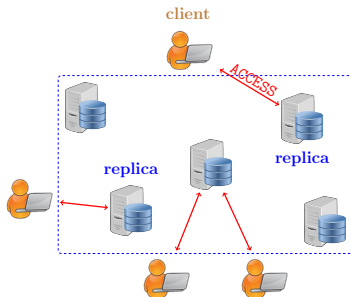
图: 个人多设备文件共享 ([基于云] C/S 结构 [左] & P2P 结构 [右]).

应用需求与特点 [Strauss@MIT Thesis'10] vs. “分布数据”:

功能需求: 文件副本

网络断连: 备份容灾; 离线可用

# 分布式应用访问分布数据



TODO: 重绘: 显示 partition & replication

# 分布共享数据服务

- ▶ 简化上层应用的开发

**TODO:** 图: 分布共享数据服务



# 分布共享数据服务

TODO: 图: 分布共享数据服务

- ▶ 简化上层应用的开发
- ▶ 提供共享数据的抽象

# 分布共享数据服务

TODO: 图: 分布共享数据服务

- ▶ 简化上层应用的开发
- ▶ 提供共享数据的抽象
- ▶ 屏蔽底层数据的分布性

# 分布共享数据服务理论与技术研究

1 研究背景

2 研究问题

3 研究方法

# 数据一致性问题

**TODO:** 图: 从分布到共享

读操作语义问题: 在分布数据环境下, 读操作允许返回什么值?

# 数据一致性问题

**TODO:** 图: 从分布到共享

读操作语义问题: 在分布数据环境下, 读操作允许返回什么值?

## 数据一致性问题

# 数据一致性问题

数据一致与否是相对于应用逻辑而言的:

# 数据一致性问题

数据一致与否是相对于应用逻辑而言的:

- ▶ 数据一致性模型多样

# 数据一致性问题

数据一致与否是相对于应用逻辑而言的:

- ▶ 数据一致性模型多样
- ▶ 数据一致性模型有强弱之分



# 数据一致性问题

数据一致与否是相对于应用逻辑而言的:

- ▶ 数据一致性模型多样
- ▶ 数据一致性模型有强弱之分
- ▶ 应用规约数据一致性需求

# 数据一致性问题

数据一致与否是相对于应用逻辑而言的:

- ▶ 数据一致性模型多样
- ▶ 数据一致性模型有强弱之分
- ▶ 应用规约数据一致性需求
- ▶ 数据不一致导致应用异常 (anomalies)

# 数据一致性问题举例 (I)

Alice: I've **lost** my ring.

Alice: I **found** it upstairs.

Bob: **Glad** to hear that.

Alice: I've **lost** my ring.

Bob: **Glad** to hear that.

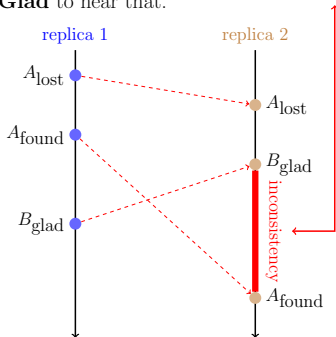


图: 社交网络中, 消息-评论乱序 [Lloyd@CACM'14].

# 数据一致性问题举例 (II)



图: 多设备文件共享时, 更新丢失 ( $\#N = 3, \#W = 2, \#R = 1$ ).

# 数据一致性问题举例 (II)

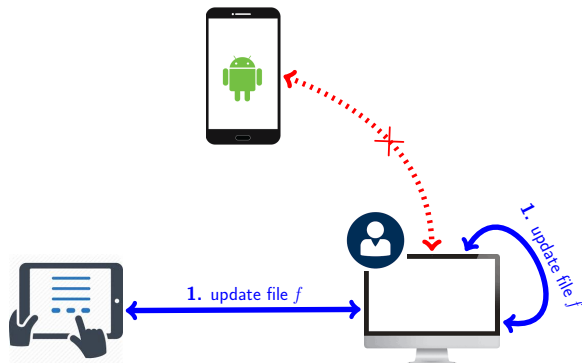


图: 多设备文件共享时, 更新丢失 ( $\#N = 3, \#W = 2, \#R = 1$ ).

# 数据一致性问题举例 (II)

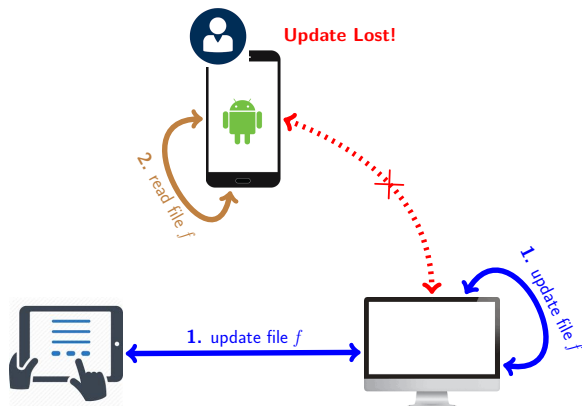
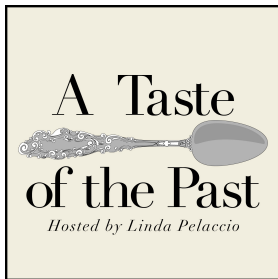


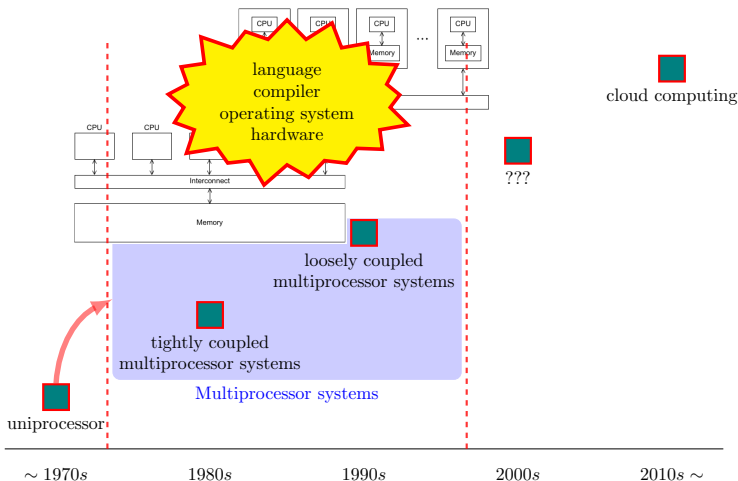
图: 多设备文件共享时, 更新丢失 ( $\#N = 3, \#W = 2, \#R = 1$ ).

# 数据一致性问题研究的历史阶段



TODO: 图: ps

# 数据一致性问题研究的历史阶段





# 数据一致性问题研究的发展趋势

## 云计算凸显应用价值观

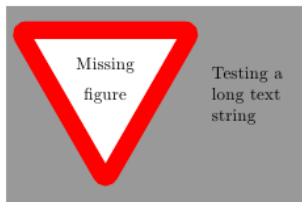


# 数据一致性问题研究的发展趋势



数据一致性“精细化”：从二元到连续谱.

# 数据一致性问题研究的发展趋势



数据一致性“多样化”：从单一到融合.

# 数据一致性问题研究的发展趋势

1. 精细化, 可度量 [Bailis@VLDB'12]



2. 多样化, 可定制 [Terry@CACM'13]



# 数据一致性问题研究的发展趋势

## 1. 精细化, 可度量 [Bailis@VLDB'12]

**精细化:** “在大多数情况下, 访问到一致数据”

**可度量:** 量化系统执行, 后验系统对一致性的满足程度

## 2. 多样化, 可定制 [Terry@CACM'13]



# 数据一致性问题研究的发展趋势



## 1. 精细化, 可度量 [Bailis@VLDB'12]

**精细化:** “在大多数情况下, 访问到一致数据”

**可度量:** 量化系统执行, 后验系统对一致性的满足程度

## 2. 多样化, 可定制 [Terry@CACM'13]

**多样化:**

- ▶ 一致性族: causality; read-your-writes (RYW)
- ▶ 参数调节: 提供“有限度”的不一致 [Yu@TOCS'02]

**可定制:** 混合使用, 运行时可变



# 分布共享数据服务理论与技术研究

1 研究背景

2 研究问题

3 研究方法

- 理论模型: 分布共享数据
- 技术途径: 三维框架

# 分布共享数据服务理论与技术研究

1 研究背景

2 研究问题

3 研究方法

- 理论模型: 分布共享数据
- 技术途径: 三维框架



# 分布共享数据 (I)

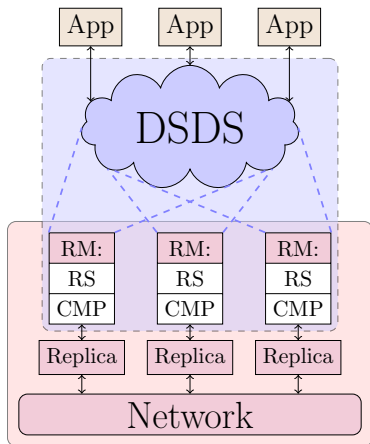
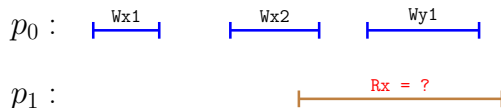


图: 分布数据共享服务.

# 分布共享数据 (II)

$x, y$  : 共享变量     $p_0, p_1$  : 客户进程

多进程并发提交 (读/写) 操作:



问题: 读操作允许返回什么值?

不同一致性  $\xrightleftharpoons[\text{定义}]{\text{规定}}$  不同合法返回值

# 分布共享数据 (III)

基本定位: 传统概念应用于新型平台

分布共享数据服务: 分布共享内存模型 + 分布数据系统

# 分布共享数据服务理论与技术研究

1 研究背景

2 研究问题

3 研究方法

- 理论模型: 分布共享数据
- 技术途径: 三维框架

# 分布共享内存中的数据一致性问题

数据一致性问题的三个层面:

- 1. 虚拟共享数据有什么?   ▶ 数据类型
- 2. 上层接口语义是什么?   ▶ 一致性模型
- 3. 底层消息传递为什么?   ▶ 一致性保障

# 研究框架

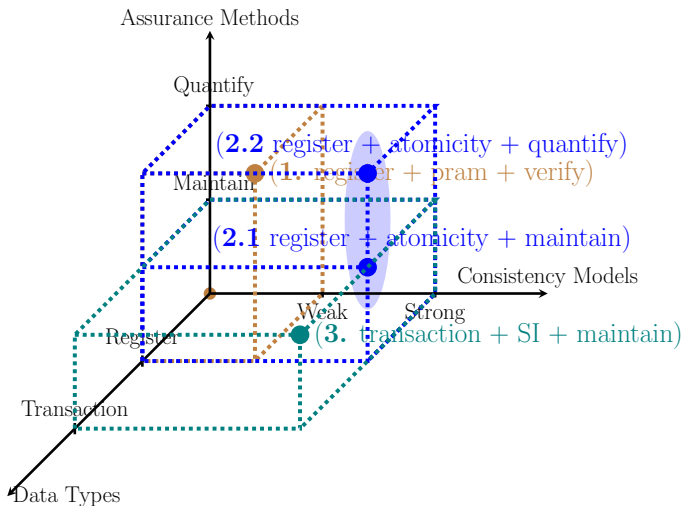


图: 数据一致性及保障技术研究框架