

# Saguaro: An Edge Computing-Enabled Hierarchical Permissioned Blockchain

作者: Mohammad Javad Amiri, Ziliang Lai, Liana Patel, Boon Thau Loo, Eric Lo, Wenchao Zhou  
University of Pennsylvania, Chinese University of Hong Kong, Stanford University, Georgetown University

分享者: 李国泽

## 背景与贡献

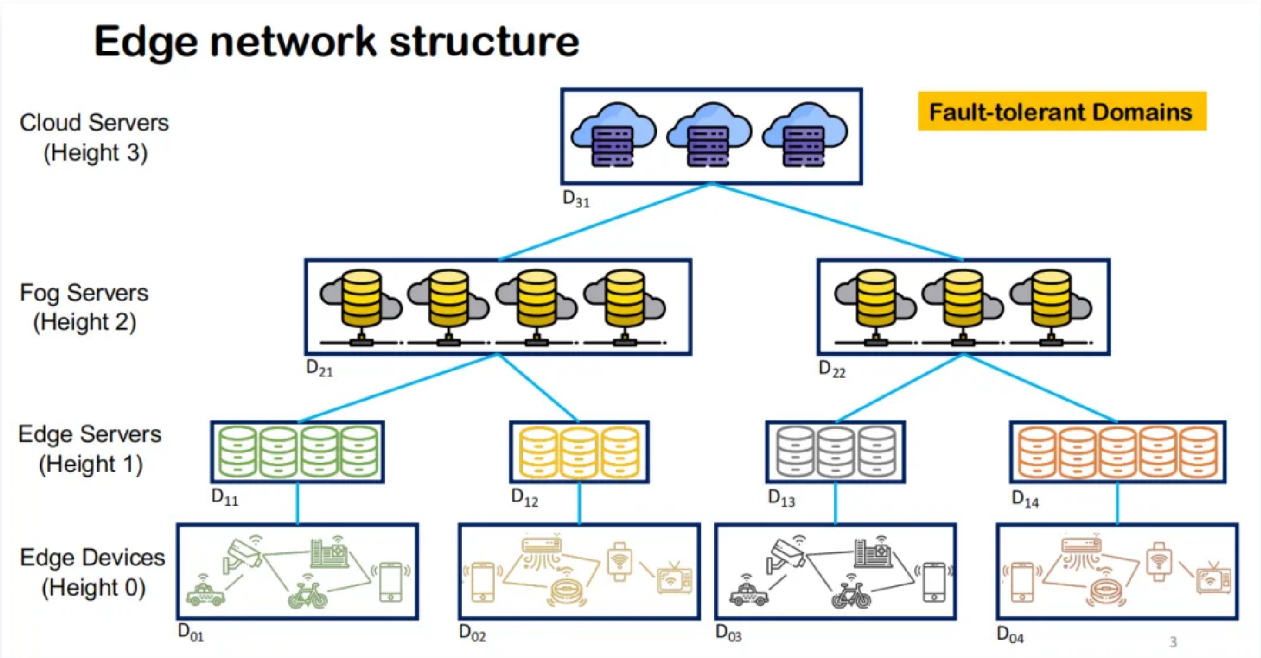
### 背景:

边缘计算场景下, 设备通常分布在多个地理区域, 它们需要在不同的域之间进行频繁的通信还可能涉及多个层次的服务器; 相对的, 高层服务器也经常有追踪各个边缘设备的某一状态的需求, 并作出应用级别的处理; 此外, 边缘设备的移动性也对跨域事务处理提出了更高的要求, 因为设备可能会临时移出本地域, 参与到远程域的操作中。然而这些需求都远非传统分布式应用所依赖于中心化的数据处理架构所能够处理的。

### 贡献:

- 基于边缘计算的分层特点, 设计了一种基于协调器的事务处理共识协议和一种基于乐观估计的事务处理共识协议。
- 基于边缘计算的分层特点, 设计了一种数据聚合协议, 允许高层服务器定期了解到事务处理情况。
- 引入了一种简单的设备管理机制, 使得设备移动成为可能。

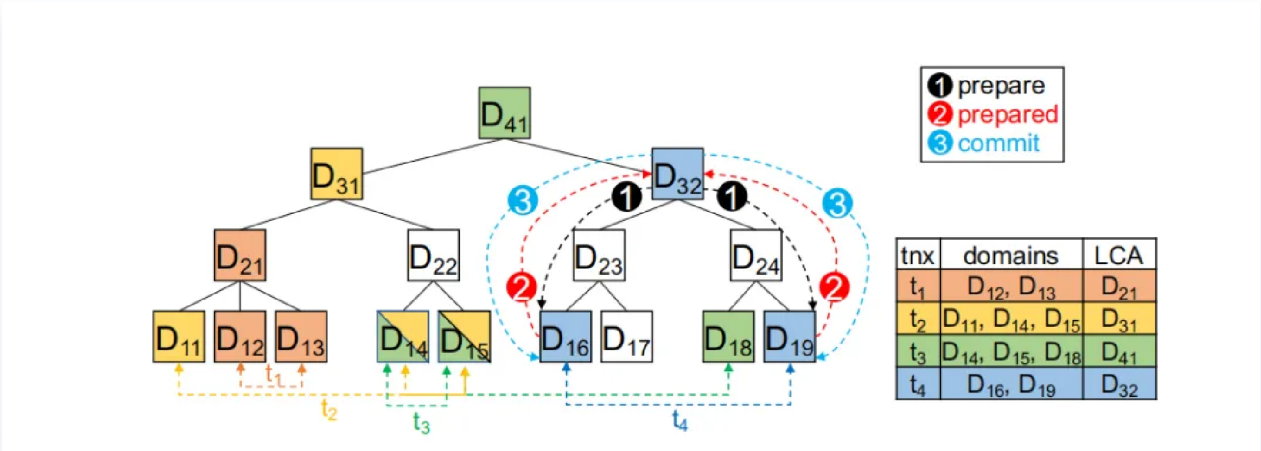
系统模型



网络由一系列层次化的容错域组成，分别由边缘设备、边缘服务器到雾计算服务器及云服务器组成。每个容错域内包含多个节点，每个域可以有这不同的容错模型：崩溃故障模型和拜占庭故障模型，并运行不同的容错协议（Paxos或PBFT）。

系统方法

基于协调器的跨域事务处理共识协议



基于协调器的共识协议核心是使用最近公共祖先容错域（LCA）作为当事务出现跨域时的排序决策者。其协议流程可以概述如下：

- Prepare阶段

当某个边缘域主节点接收到一个跨域事务，其转发给所涉及域的最低公共祖先（LCA）域下所有节点。LCA域的主节点接收到请求后进行检查：如果当前正在处理的事务与这个跨域事务所涉及的域存在冲突（存在交集），则等待直到前一个事务提交，以确保在交集域上交易的顺序一致。如果没有冲突，则LCA域的主节点将为该事务分配一个序列号，并将包含序列号、事务摘要和事务本身等信息通过签名作为Prepare信息发送给所有涉及域下的节点。

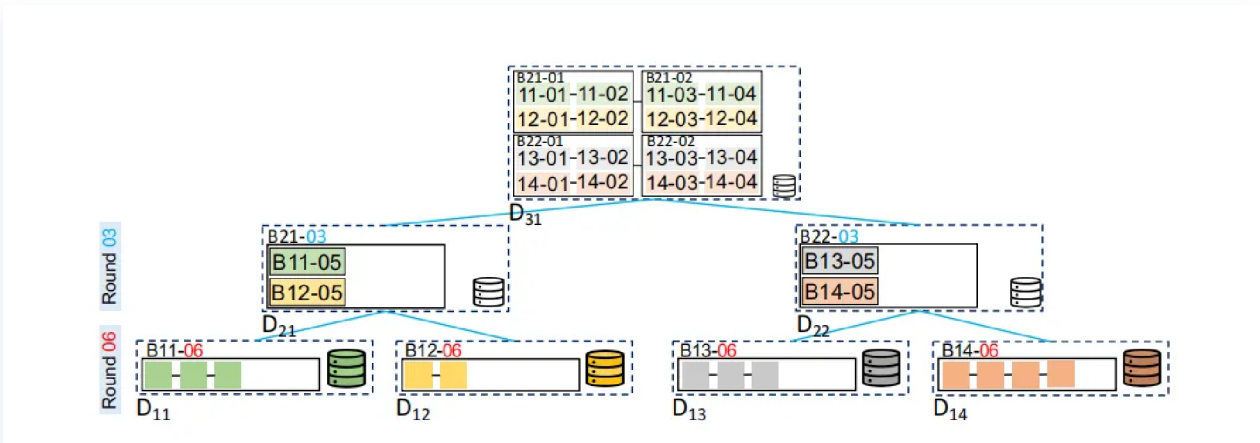
- Prepared阶段

涉及域的主节点收到Prepare信息后，若不存在冲突，则其同样为该请求分配一个序列号，并在其域下发起对该事务的共识。一旦共识达成，该域的主节点将共识结果通过签名形式作为Prepared信息发送给LCA域下的所有节点。

- Commit 阶段

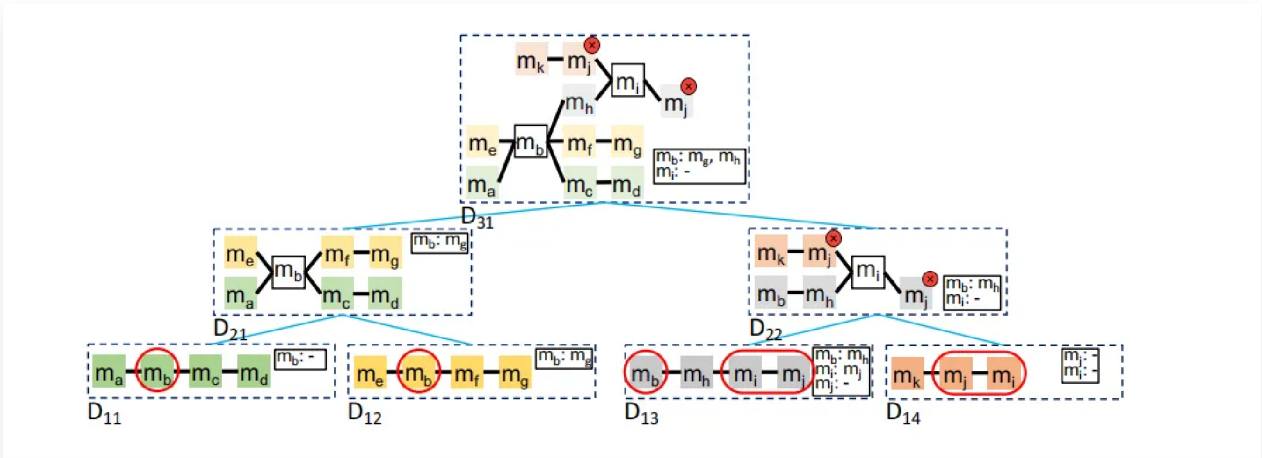
当LCA域的主节点收到所有涉及域的Prepared信息后，在域内开始对该事务进行共识并最终生成一个包括各涉及域的序列号的Commit信息。如果某个涉及域未同意该交易，LCA域会发送一个Abort信息。所有涉及域的节点在收到Commit信息后，将交易视为已提交，并根据顺序执行。

数据聚合协议



Saguaro使用一种Lazy形式的聚合协议，当在一个预定长度的回合后，每个边缘域的主节点打包本轮的事务为一个区块信息并进行共识，共识完成后主节点将区块信息和签名一起发送至父级域的所有节点。如果某个域没有收到任何事务，则其发送一个空消息。类似的，更高层级的域中对从子域接收到的区块消息进行共识并发送给更高父域的所有节点。如果父域的主节点在预定时间内没有收到来自子域的区块消息，它会向子域的所有节点发送查询消息并引起子域的换届机制。

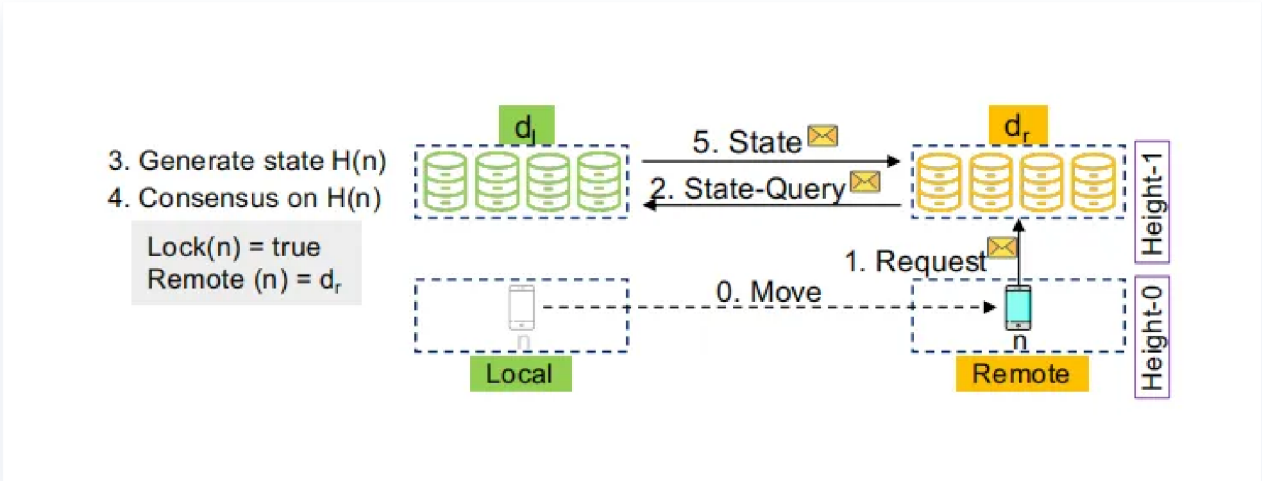
高层域从多个子域接收区块消息，并对收到的所有事务进行排序。如果不同子域的交易之间没有依赖关系，则任意一种交易顺序都是可行的。然而，跨域交易必须仅一次附加到父域的账本，从而使得生成的账本是一个有向无环图，以捕获交易的顺序依赖关系。最终，根域的账本将包含系统中处理的所有交易。



由于数据聚合协议下事务会在回合结束时逐层上传，所以Saguaro也可以采用一种乐观估计的形式进行事务提交和执行，并最终在LCA域进行一致性检查。

当某边缘服务器收到一条跨域事务，其将事务广播至所有涉及域节点，每个涉及域自行使用内部共识协议提交和执行该事务并维持一个依赖表，记录之后所有直接或间接依赖该事务的记录。当进行数据聚合，最终LCA域能够收到该跨域事务涉及域的区块信息并以此检查这些域是否对事务进行了一致排序，若检查失败，则其将该事务以及所有直接或者间接依赖其的事务标记为Aborted，并将相关的证明以签名的形式发送给其所涉及域。涉及域下节点进行回滚。

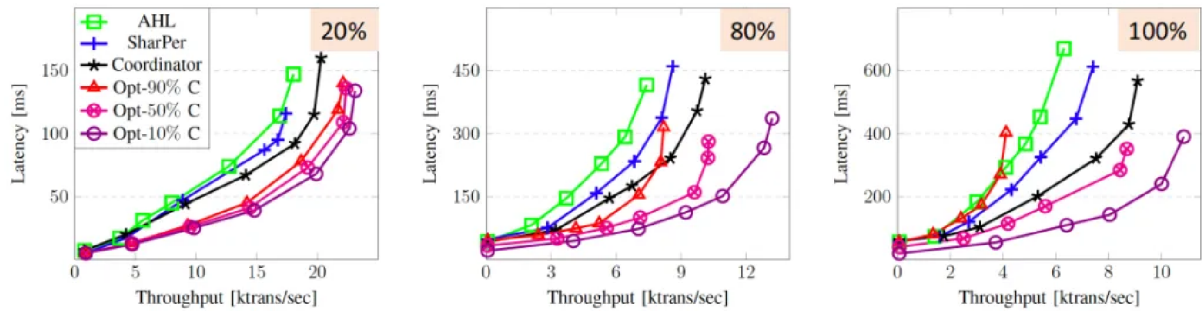
移动共识协议



为了支持设备移动，Saguaro以本地域充当中介角色，通过向新域传递最新状态，确保设备在新的远程域能够继续进行交易处理。每个域对其下设备维持一个状态H，一个锁位来跟踪设备的移动状态和一个remote变量。当一个设备在远端域发起事务，远程域主节点向其本地域发起State-Query信息以询问状态。当本地域的主节点检查设备的锁位：若锁位为TRUE，表示设备的本地状态是最新的，本地域对其状态进行共识并发送状态信息至远端域。如果设备的锁位为FALSE，表示本地域的状态已经过时，本地域根据remote变量向相关其他远程域请求设备的最新状态并返回给发起请求的远程域。

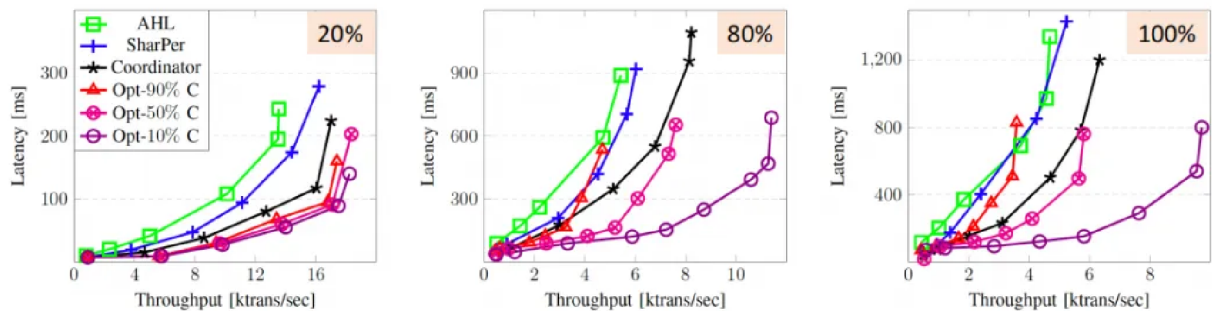
## Cross-domain transactions (crash-only)

Domains: Frankfurt, Milan, London, and Paris (RTT: 9-25 ms)



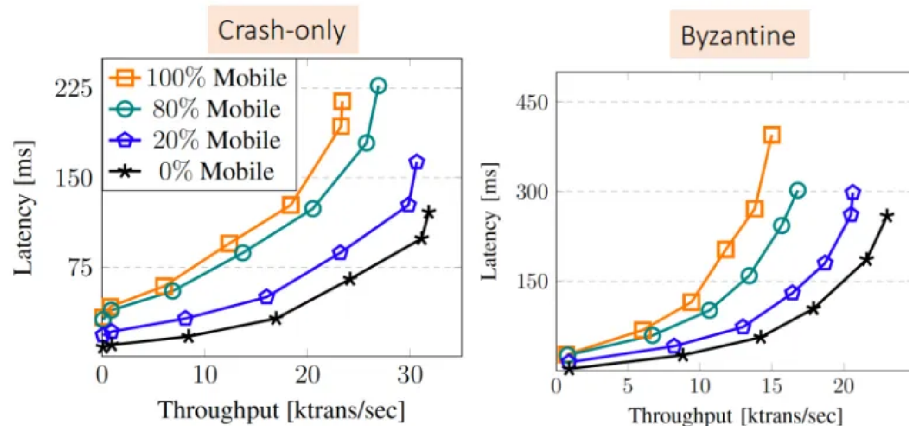
## Cross-domain transactions (Byzantine)

Domains: Frankfurt, Milan, London, and Paris (RTT: 9-25 ms)



## Mobile devices

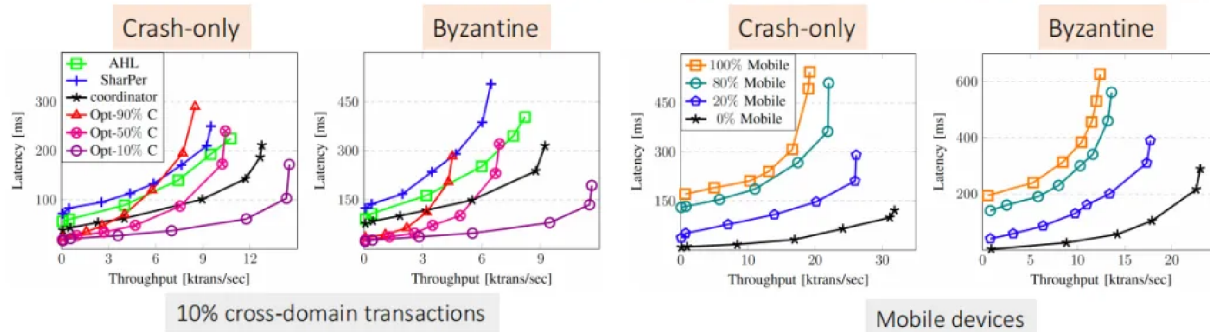
A mobile node initiates 10 transactions within the remote domain before moving back to its local domain.





# Wide-area networks

Domains: California, Oregon, Virginia, Ohio, Tokyo, Seoul, and Hong Kong



作者在 Amazon EC2 上以 VM 形式模拟服务器实现了一个四层满二叉树边缘计算网络拓扑，并分别在跨域事务，移动设备以及广域网络场景下与最新协议 AHL 和 SharPer 做了对比。基于协调器的协议在所有场景下均优于 SharPer 和 AHL，展示了广域网环境中可扩展的优势。基于乐观估计的协议避免了跨域通信高效处理事务，使得其在低冲突工作负载下表现优异。在移动场景下协议性能损失并不明显，证明了其在动态环境中的有效性。

## 评价

- 作者在巧妙利用了边缘计算的分布式特点，将共识协议零成本的融入到了网络中，这一叙事十分出色。
- 本篇文章的实验设置恰到好处，结构简单却目标明确。
- 此篇论文以特定场景设计了三个共识模块，将原本三个独立的内容结合到了一起。