

Ideas about Mobile Learning

张海鹏
2017-04-23

BG(1/3)



FACEBOOK
DEVELOPER
CONFERENCE

April 18 + 19, 2017

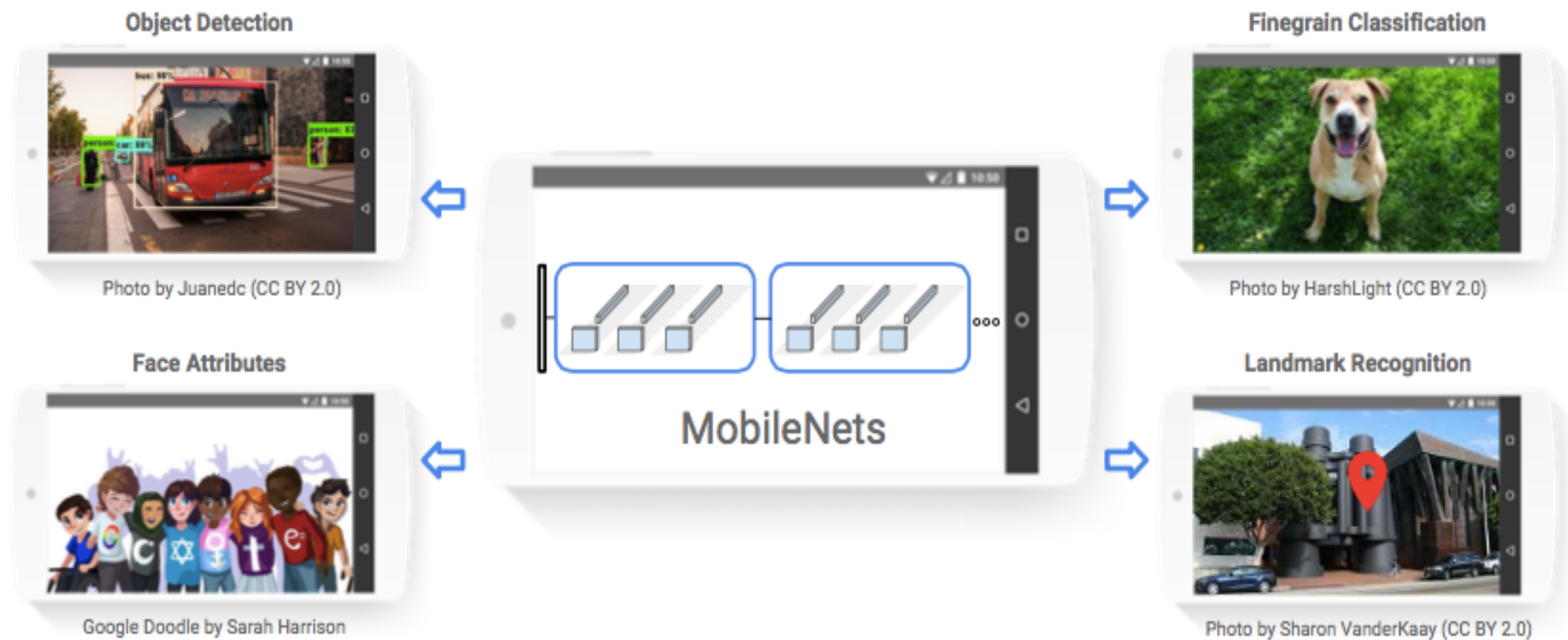
SAN JOSE, CALIFORNIA



Caffe2

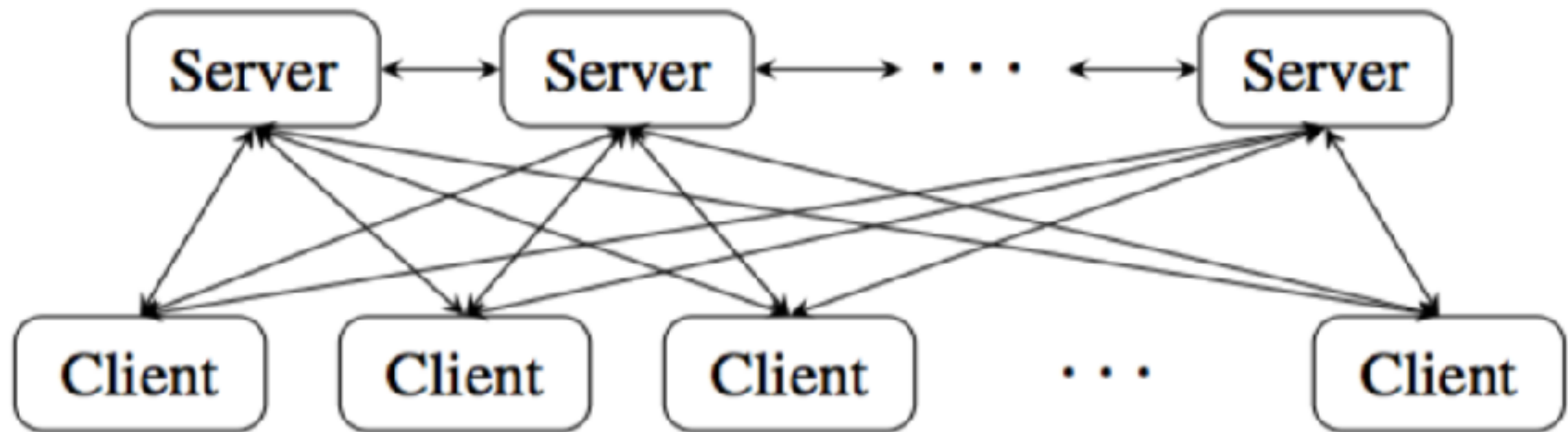
A New Lightweight, Modular, and Scalable Deep Learning Framework

BG(2/3)



«MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications»
(Andrew G. Howard etc. Google Inc)

BG(3/3)



«Scaling Distributed Machine Learning with the Parameter Server»
(M Li, OSDI 2014)

SGD(Distributed)

Algorithm 3 SimuParallelSGD(Examples $\{c^1, \dots, c^m\}$, Learning Rate η , Machines k)

Define $T = \lfloor m/k \rfloor$

Randomly partition the examples, giving T examples to each machine.

for all $i \in \{1, \dots, k\}$ **parallel do**

 Randomly shuffle the data on machine i .

 Initialize $w_{i,0} = 0$.

for all $t \in \{1, \dots, T\}$: **do**

 Get the t th example on the i th machine (this machine), $c^{i,t}$

$w_{i,t} \leftarrow w_{i,t-1} - \eta \partial_w c^i(w_{i,t-1})$

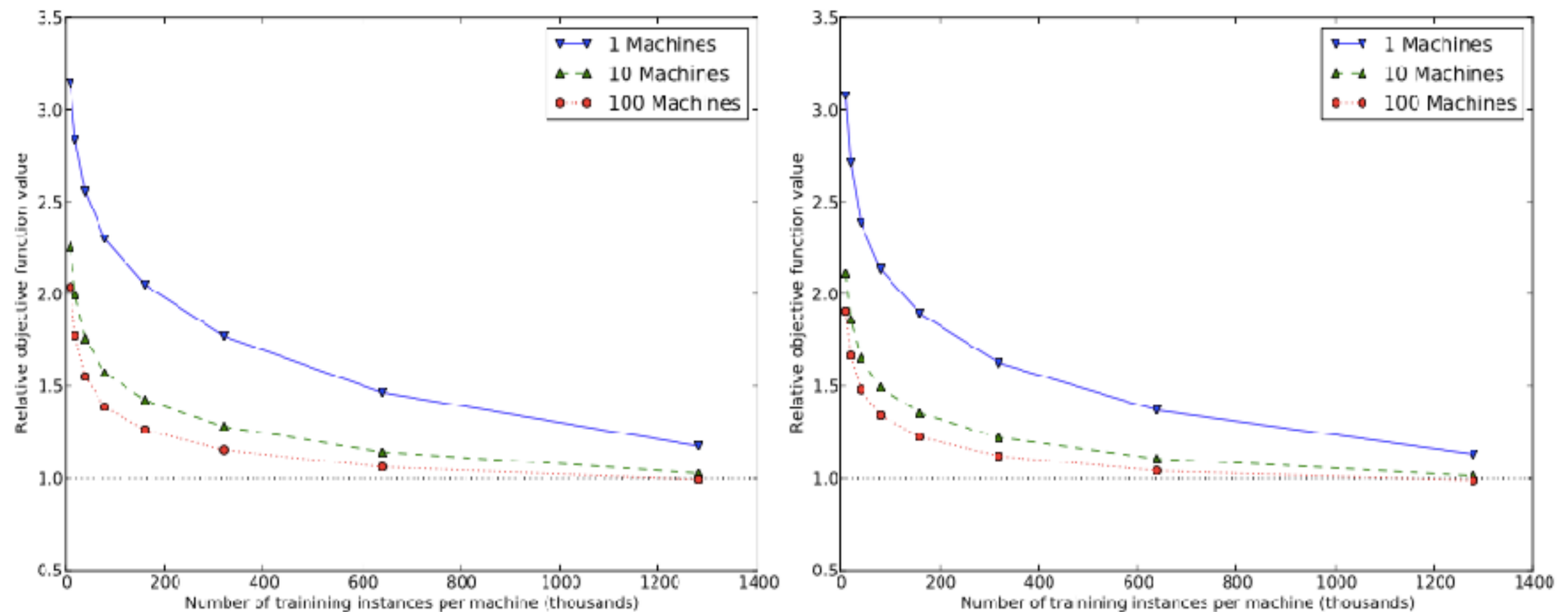
end for

end for

Aggregate from all computers $v = \frac{1}{k} \sum_{i=1}^k w_{i,t}$ and **return** v .

«Parallelized Stochastic Gradient Descent » (Martin A. Zinkevich etc.NIPS 2010)

SGD(Distributed)



Cs v.s. Ms

总体配置

| 项目 | 描述 | 数目 |
|------|---------------|----|
| 管理节点 | 2U服务器 | 一台 |
| 计算节点 | 2U服务器 | 四台 |
| 网络连接 | 千兆交换机 | 一台 |
| 计算网络 | Infiniband交换机 | 一台 |
| 供电系统 | 32A PDU | 两个 |

Cs v.s. Ms

管理节点配置

| 项目 | 描述 | 数目 |
|---------|---|----|
| CPU | Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz | 2 |
| 内存 | 16G ECC REG | 4 |
| 硬盘（系统盘） | Intel 240GB SSD | 1 |
| 硬盘（存储盘） | 2TB SATA HDD | 3 |

Cs v.s. Ms

计算节点配置

| 项目 | 描述 | 数目 |
|--------|---|----|
| CPU | Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz | 2 |
| 内存 | 32G ECC REG | 8 |
| 硬盘 | 1TB SATA HDD | 1 |
| GPU计算卡 | NVIDIA Tesla K80 | 2 |

Cs v.s. Ms

Cons for Cs:

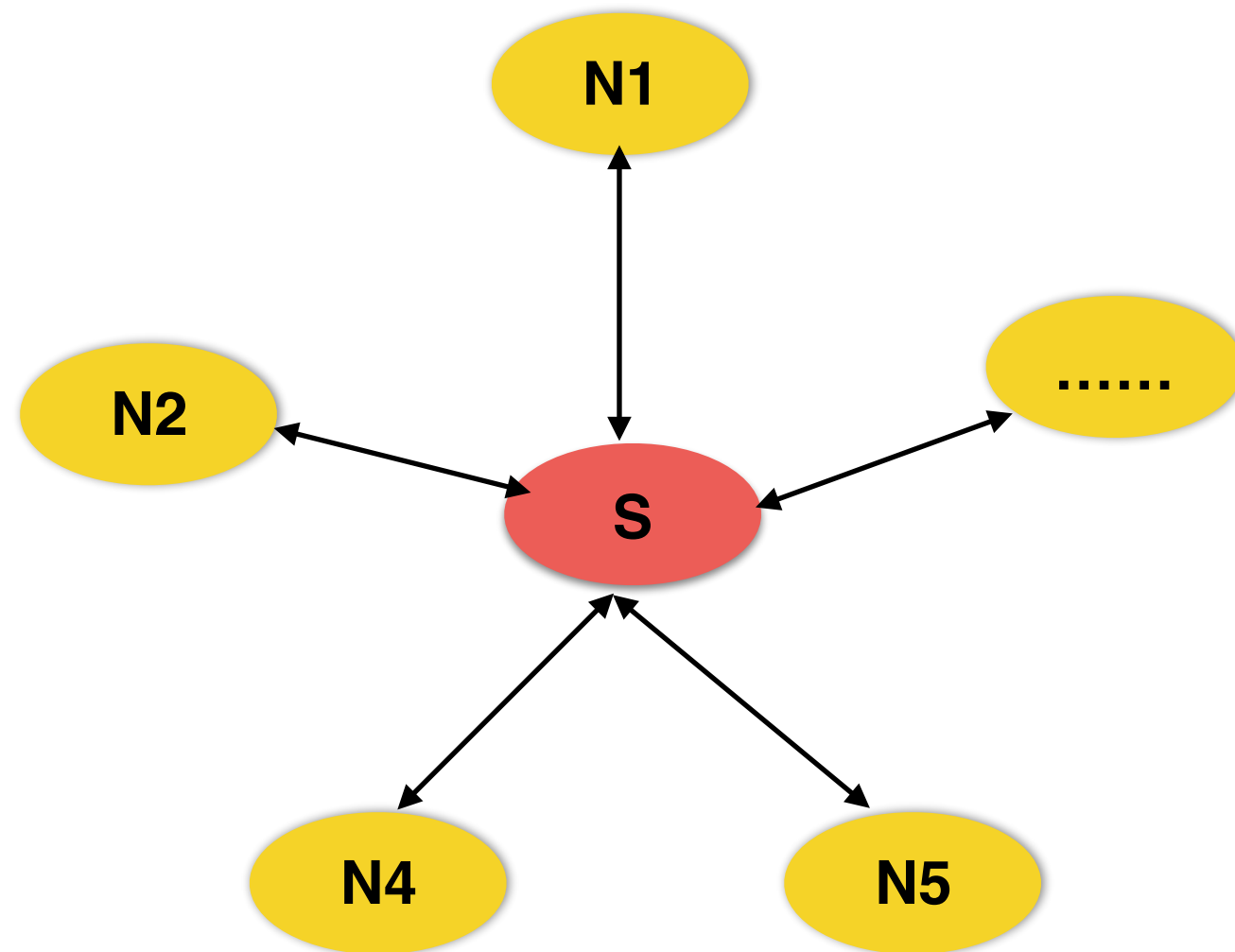
1. 数据在**master**节点，**worker**节点通过分发获取数据
 2. **worker**节点数目扩展不易
 3. **master**和**worker**节点成本高
 4. 通信成本高(**infiniband**)
 5. 物理移动性差
-

Cs v.s. Ms

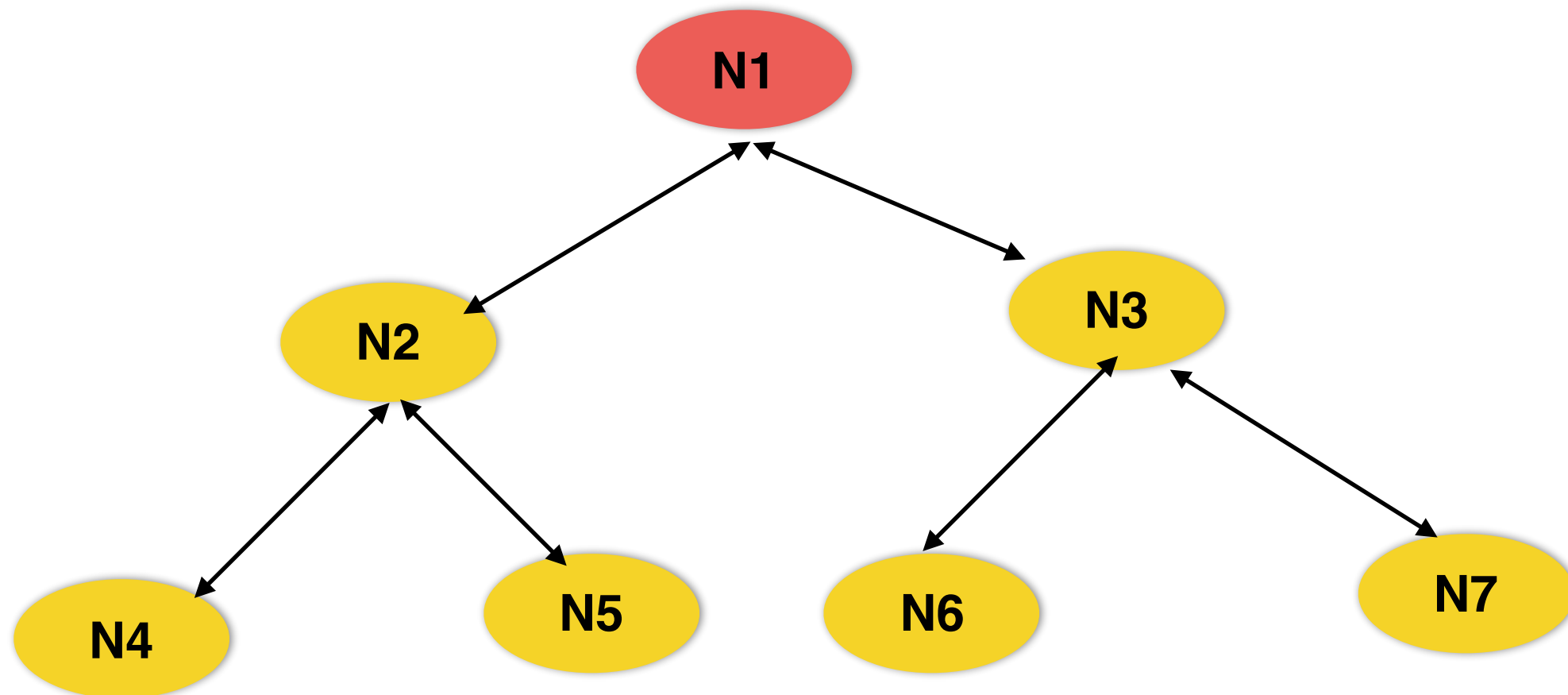
Pros for Ms:

1. 数据天然在worker上
 2. 天然的增量学习
 3. worker扩展容易
 4. worker成本较低
 5. 通信成本低（避免网络延迟和带宽问题）
 6. 物理移动性好
 7. 隐私性
 8. 用户体验
 9. 省电和水
-

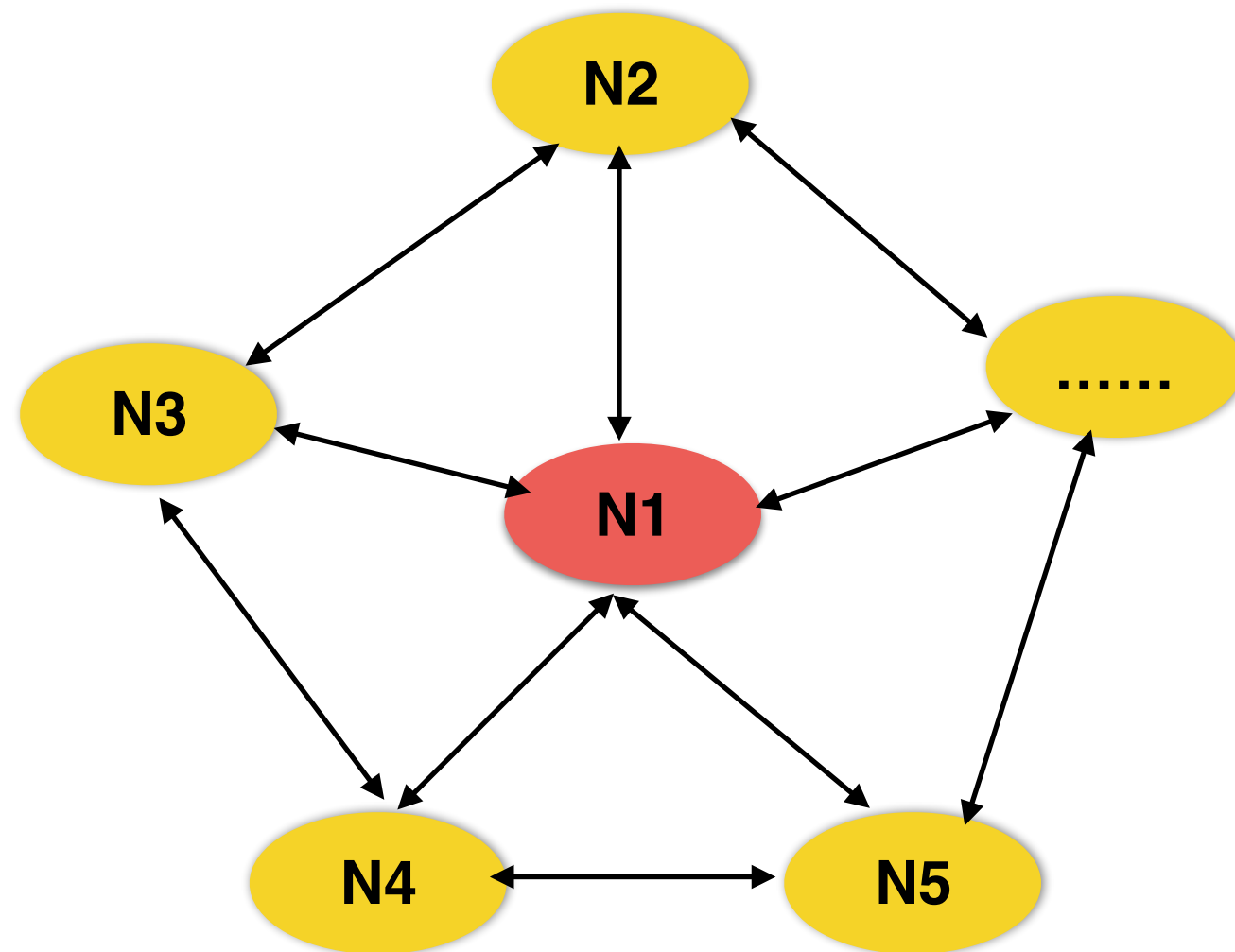
Ms(with S)



Ms(without S)



Ms(without S)



Ms(without S)

Pros:

1. 未来的移动社交应用场景，几部手机共同做一件事情
 2. 实现**worker**节点的通信，降低**master**的通信负载
 3. 天然的容错（以通信换容错），而且手机容错的能力更加强大
 4. 树形结构天然的优点，如果能够和模型结合更加的**OK**，树形结构本身的优良性质
 5. **N1**本身是信息的管理单元，不是计算单元
-

Implementation

研究意义: **Smart Phone for Smart Cities**

结构研究: 中心和无中心

问题关键: **Convergence, Complexity**

研究问题: **Logistic/Linear Regression Based On Parallel SGD**

潜在问题: 冷启动, 节点角色等

Q&R

TKS

创造始于问题，有了问题才会思考，有了思考，才有解决问题的方法，才有找到独立思路的可能。

——陶行知

LIVE FOR IDEAS.

——张海鹏
