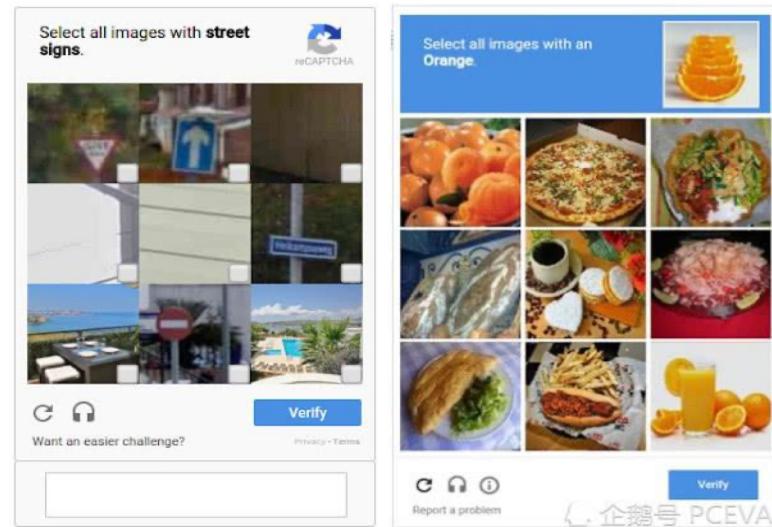


EECS 221A Project

Enze Zheng
Xiaoran Li
Xianpei Meng

CAPTCHA

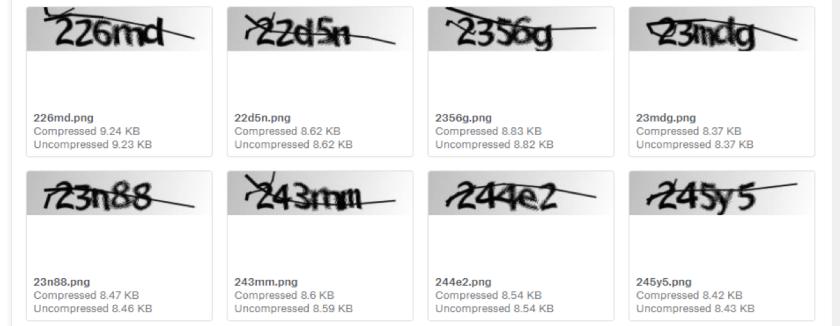
- Prevent crawler to access web content
- Sometimes, it is difficult for human to tell what it is



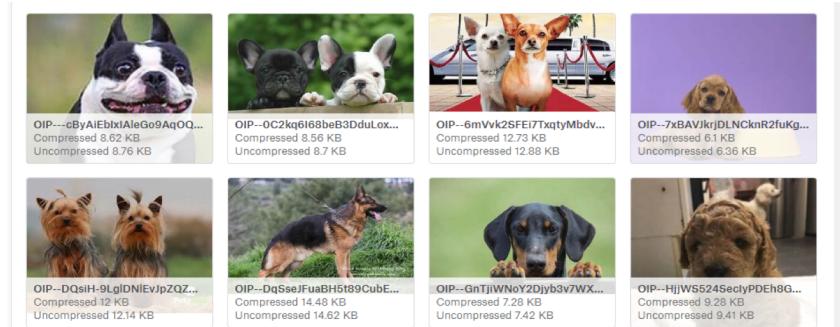
Machine Learning



TensorFlow



1070 files with answer for supervised training



10 different kinds of animals

More than 20 thousands files in total

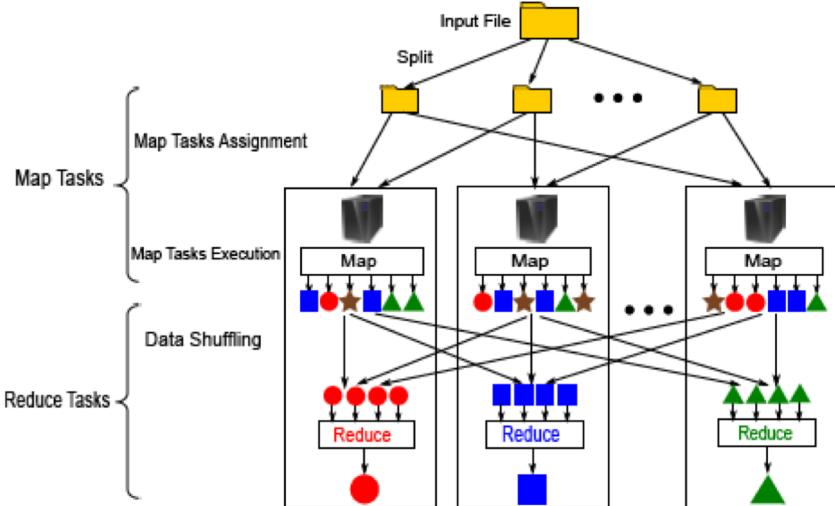
Related Work

Large-scale Artificial Neural Network: MapReduce-based Deep Learning(2015)

A Fundamental Tradeoff between Computation and Communication in Distributed Computing(2017)

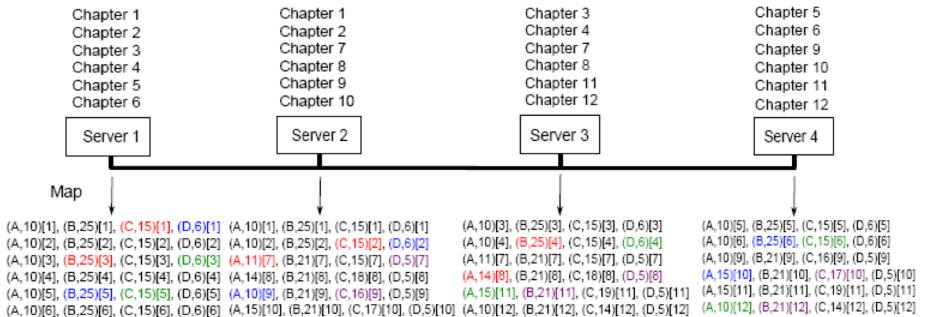
Gradient Coding: Avoiding Stragglers in Distributed Learning(2017)

Coded MapReduce

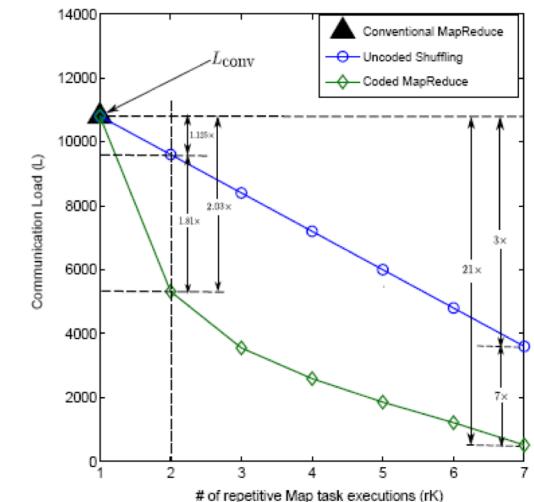


$$\max \left\{ QN \frac{1-r}{K-1}, \max_{s \in \{1, \dots, K\}} sQN \left(\frac{1}{K} - \frac{r}{\lfloor \frac{K}{s} \rfloor} \right) \right\} \leq L^*(r) \leq L_{CMR}(r) \triangleq \frac{QN}{K} \left(\frac{1}{r} - 1 \right) + o(N)$$

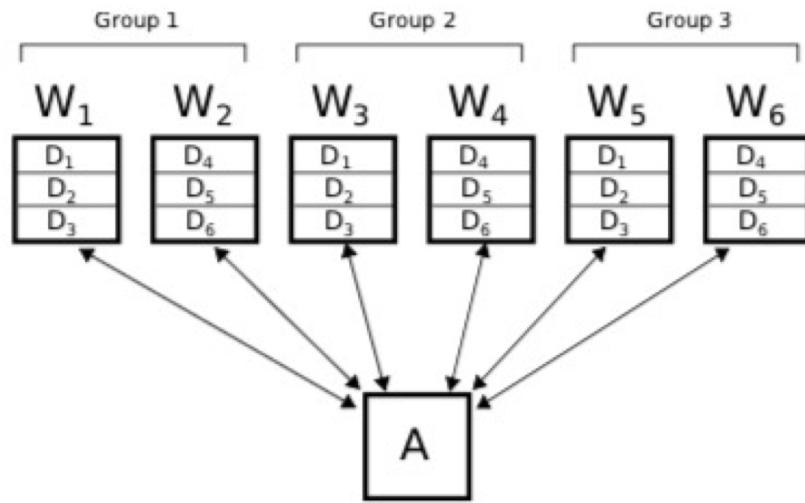
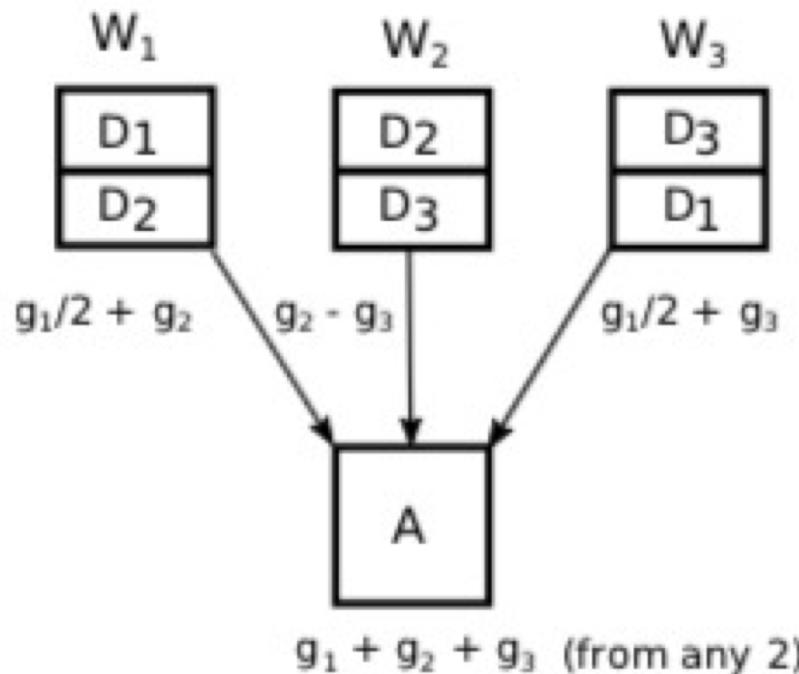
Q:keys
K:server
N:subfile
pK:servers



$(B, 25)[3]$	$\rightarrow (BC, 40)[3, 1]$	$(A, 11)[7]$	$\rightarrow (AC, 26)[7, 2]$	$(A, 14)[8]$	$\rightarrow (AB, 39)[8, 4]$	$(A, 15)[10]$	$\rightarrow (AB, 40)[10, 6]$
$(C, 15)[1]$		$(C, 15)[2]$		$(B, 25)[4]$		$(B, 25)[6]$	
$(B, 25)[5]$	$\rightarrow (BD, 31)[5, 1]$	$(A, 10)[9]$	$\rightarrow (AD, 16)[9, 2]$	$(A, 15)[11]$	$\rightarrow (AD, 21)[11, 4]$	$(A, 10)[12]$	$\rightarrow (AC, 25)[12, 6]$
$(D, 6)[1]$		$(D, 6)[2]$		$(B, 6)[4]$		$(B, 21)[12]$	$\rightarrow (BC, 38)[12, 10]$
$(C, 15)[5]$	$\rightarrow (CD, 21)[5, 3]$	$(C, 16)[9]$	$\rightarrow (CD, 21)[9, 7]$	$(D, 5)[8]$		$(C, 17)[10]$	
$(D, 6)[3]$							



Gradient Coding



$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1/2 & 1 & 0 \\ 0 & 1 & -1 \\ 1/2 & 0 & 1 \end{pmatrix}$$