Case Study 4: Collaborative Filtering

Graph-Parallel Problems

Synchronous v. Asynchronous Computation

Machine Learning/Statistics for Big Data
CSE599C1/STAT592, University of Washington

Carlos Guestrin

March 12th, 2013

Needless to Say, We Need Machine Learning for Big Data

flickr
6 Billion Flickr Photos

28 Million Wikipedia Pages

facebook
1 Billion Facebook Users

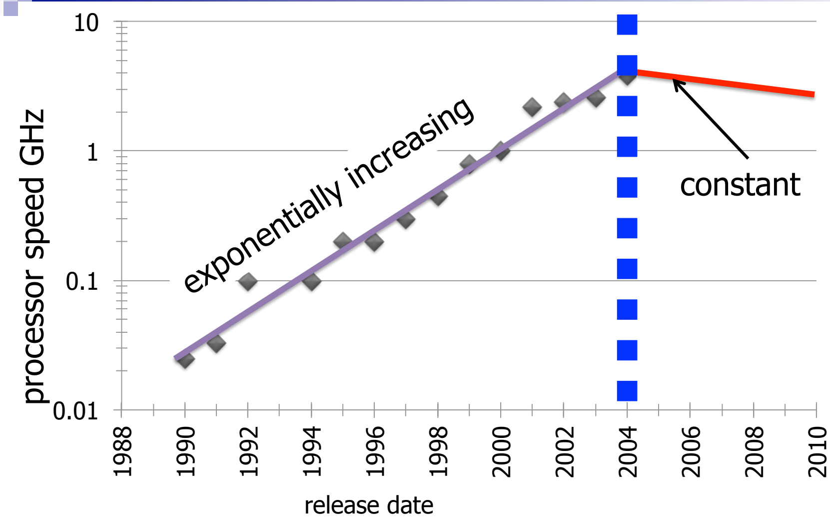
YouTube
72 Hours a Minute YouTube

The New York Times
SundayReview
WORLD U.S. N.Y. / REGION BUSINESS TEC

NEWS ANALYSIS
The Age of Big Data
By STEVE LOHR
Published: February 11, 2012

"… data a new class of economic asset, like currency or gold."

CPUs Stopped Getting Faster…



ML in the Context of Parallel Architectures

GPUs Multicore Clusters Clouds Supercomputers

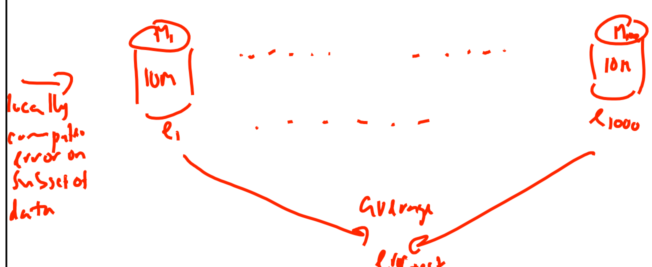
- But scalable ML in these systems is hard, especially in terms of:
  1. Programmability
  2. Data distribution
  3. Failures

Move Towards Higher-Level Abstraction

- Distributed computing challenges are hard and annoying!
  1. Programmability
  2. Data distribution
  3. Failures
- High-level abstractions try to simplify distributed programming by hiding challenges:
  - Provide different levels of robustness to failures, optimizing data movement and communication, protect against race conditions…
  - Generally, you are still on your own WRT designing parallel algorithms
- Some common parallel abstractions:
  - Lower-level:
    - Pthreads: abstraction for distributed threads on single machine
    - MPI: abstraction for distributed communication in a cluster of computers
  - Higher-level:
    - Map-Reduce (Hadoop: open-source version): mostly data-parallel problems
    - GraphLab: for graph-structured distributed problems

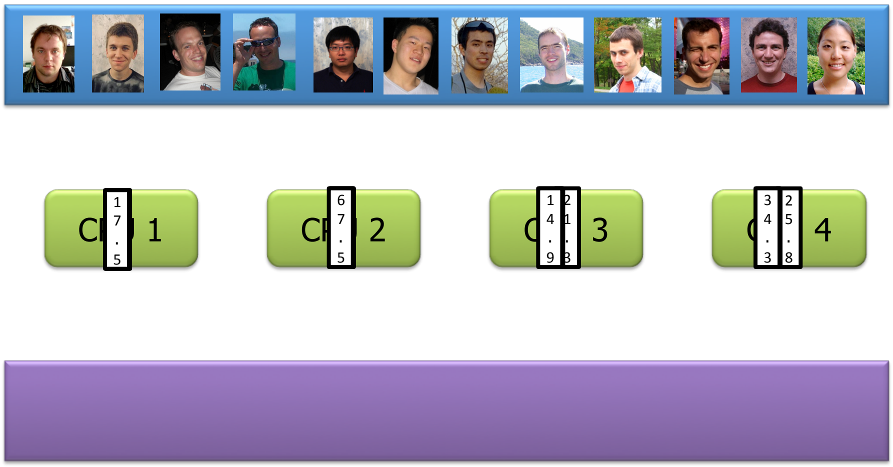
Simplest Type of Parallelism: Data Parallel Problems

- You have already learned a classifier w^*
  - What's the test error? $err = \frac{1}{N_{test}} \sum_i \frac{1}{2} \left| y^{(i)} - sign(w^* \cdot x^{(i)}) \right|$
- You have 10B labeled documents and 1000 machines



- Problems that can be broken into independent subproblems are called data-parallel (or embarrassingly parallel)
- Map-Reduce is a great tool for this…
  - Focus of today's lecture
  - but first a simple example

Data Parallelism (MapReduce)



*Solve a huge number of **independent** subproblems, e.g., extract features in images*

Map-Reduce Abstraction

- Map: Transforms a data element
  - Data-parallel over elements, e.g., documents
  - Generate (key,value) pairs
    - "value" can be any data type

('UW', 17)

in this example: ('Mary',1) ('UW',1) ('Mary',1)

- Reduce: Take all values associated with a key
  - Aggregate values for each key and aggregate
  - Must be commutative-associate operation
  - Data-parallel over keys
  - Generate (key,value) pairs

reduce ('UW', [1, 17, 0, 0, 12])
emit ('UW', 30)

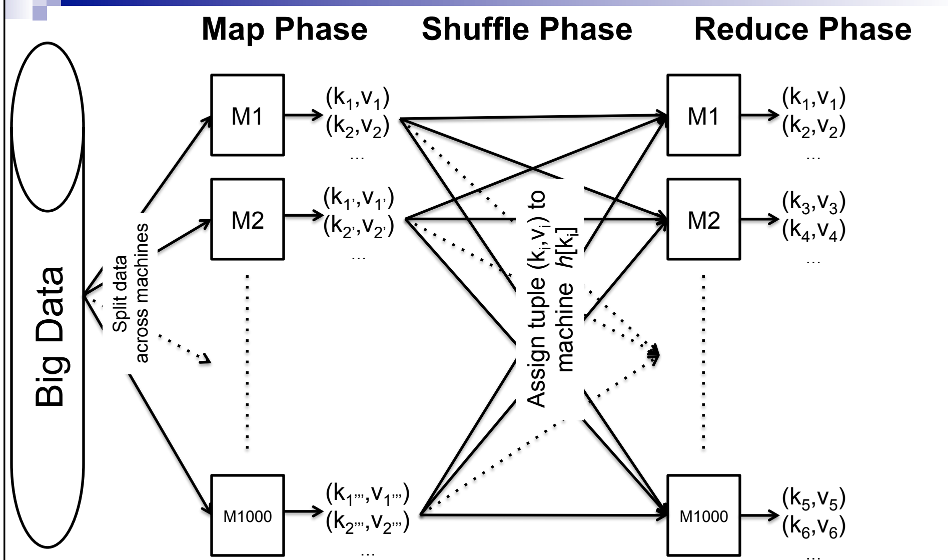
word count

map (document)
for word in doc
emit (word, 1)

Reduce (word, count: list(int))
c = 0
for i in count
c += count[i]
emit (word, c)

- Map-Reduce has long history in functional programming
  - But popularized by Google, and subsequently by open-source Hadoop implementation from Yahoo!

Map-Reduce – Execution Overview



Issues with Map-Reduce Abstraction

- Often all data gets moved around cluster
  - Very bad for iterative settings
- Definition of Map & Reduce functions can be unintuitive in many apps
  - Graphs are challenging
- Computation is synchronous

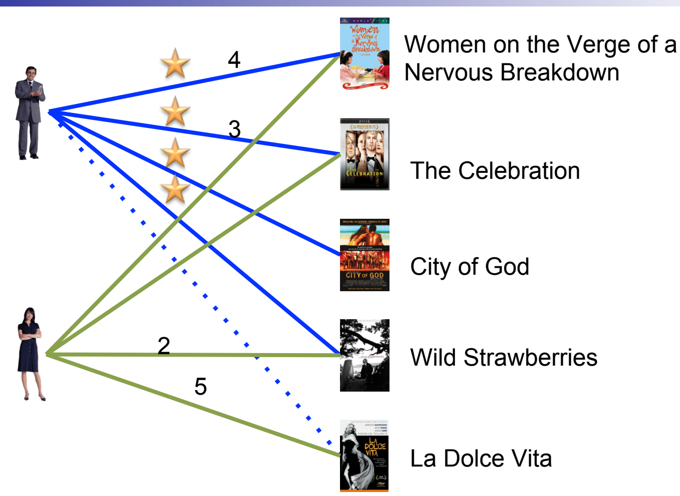
SGD for Matrix Factorization in Map-Reduce?

$$\epsilon_t = L_u^{(t)} \cdot R_v^{(t)} - r_{uv}$$

$$\begin{bmatrix} L_u^{(t+1)} \\ R_v^{(t+1)} \end{bmatrix} \leftarrow \begin{bmatrix} (1-\eta_t\lambda_u)L_u^{(t)} - \eta_t\epsilon_t R_v^{(t)} \\ (1-\eta_t\lambda_v)R_v^{(t)} - \eta_t\epsilon_t L_u^{(t)} \end{bmatrix}$$

- Map and Reduce functions???
- Map-Reduce:
  - Data-parallel over all mappers
  - Data-parallel over reducers with same key
- Here, one update at a time!

Matrix Factorization as a Graph



Flashback to 1998

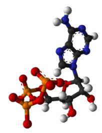
altavista SEARCH SOFTWARE

Why?

Google! BETA

First Google advantage:
a **Graph Algorithm** & a **System to Support** it!

Social Media | **Science** | **Advertising** | **Web**



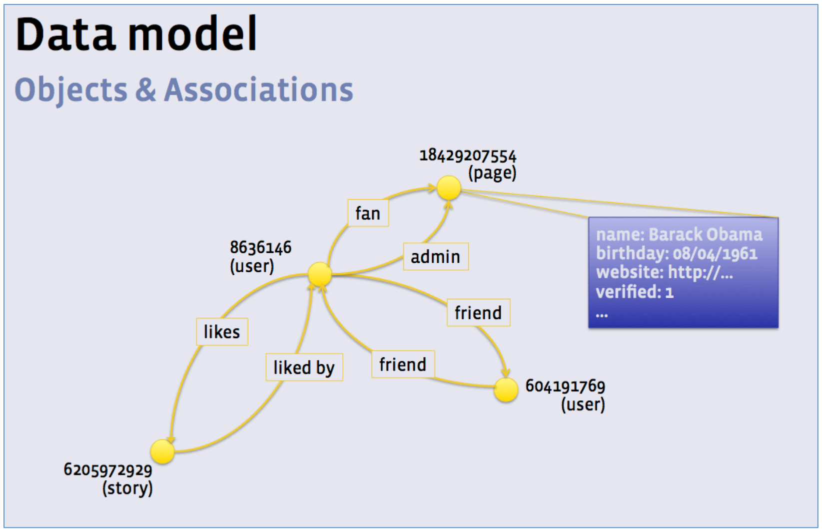
NETFLIX

- **Graphs** encode the **relationships** between:

People Products Ideas
Facts Interests

- **Big: 100 billions** of **vertices** and **edges** and rich metadata
  - Facebook (10/2012): 1B users, 144B friendships
  - Twitter (2011): 15B follower edges

Facebook Graph



Slide from Facebook Engineering presentation

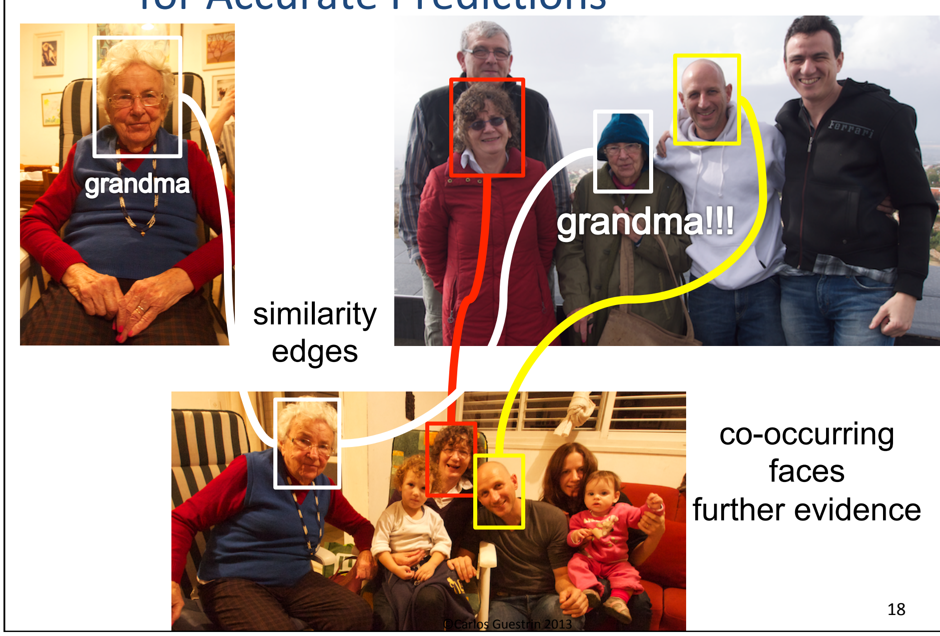
Label a Face and Propagate

Pairwise similarity not enough…

©Carlos Guestrin 2013

17

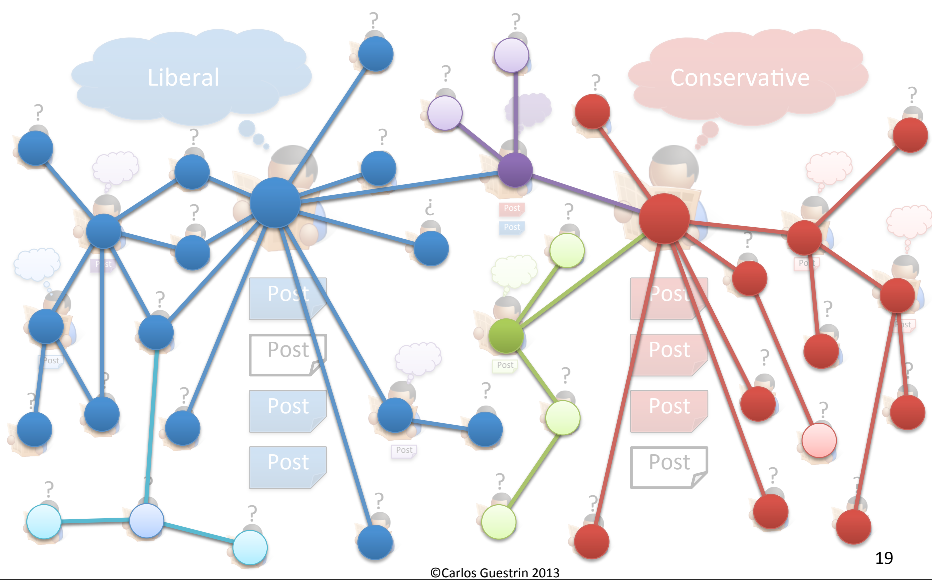
Propagate Similarities & Co-occurrences for Accurate Predictions



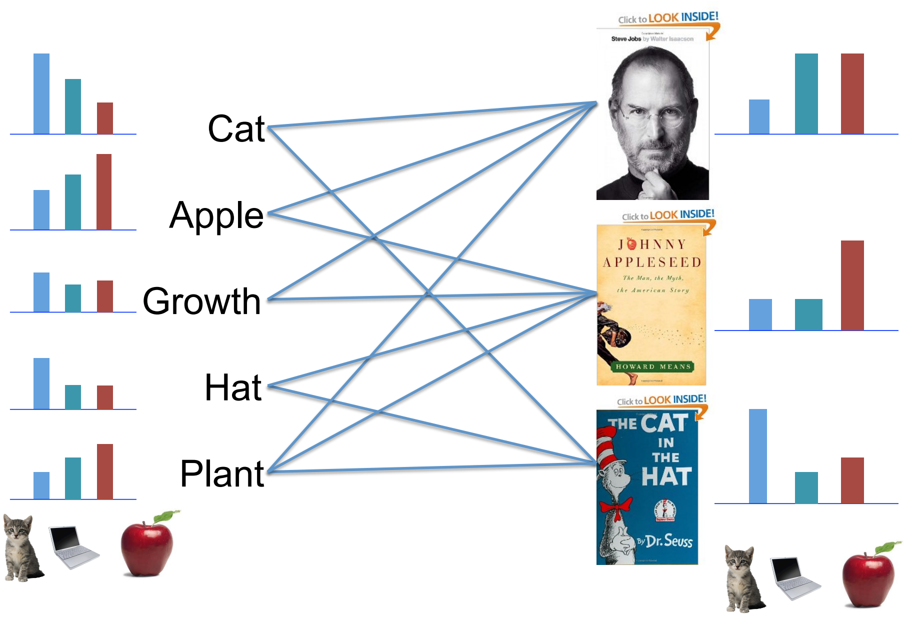
©Carlos Guestrin 2013

18

Example: *Estimate Political Bias*



Latent Topic Modeling (LDA)



ML Tasks Beyond Data-Parallelism

Map Reduce

Feature Extraction

Cross Validation

Computing Sufficient Statistics

Graphical Models
Gibbs Sampling
Belief Propagation
Variational Opt.

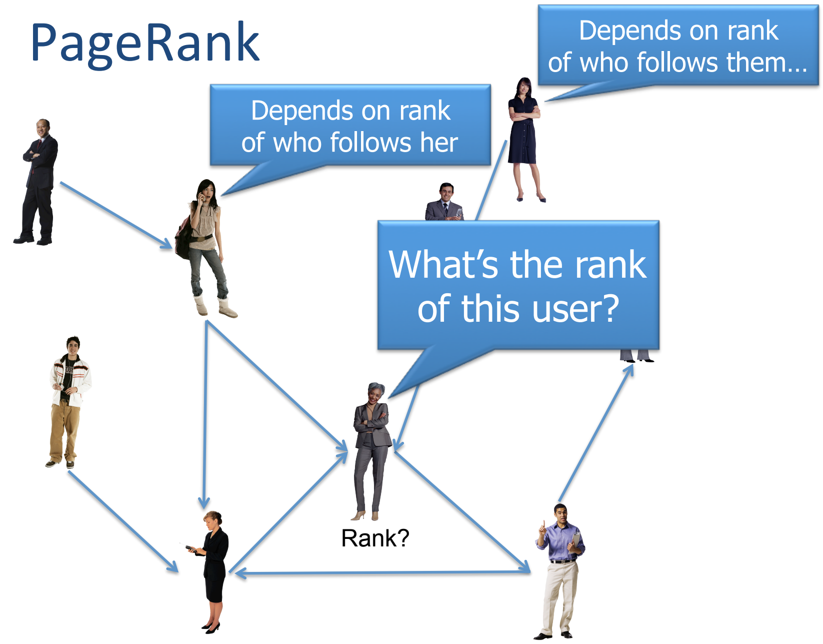
Semi-Supervised Learning
Label Propagation
CoEM

Collaborative Filtering
Tensor Factorization

Graph Analysis
PageRank
Triangle Counting

Example of a Graph-Parallel Algorithm

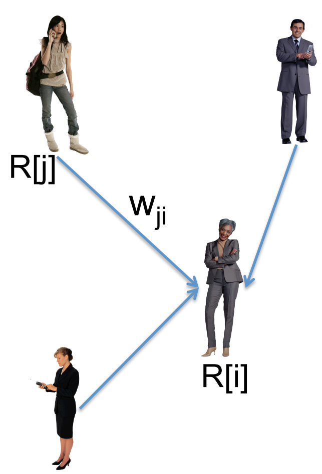
PageRank



Loops in graph → Must iterate!

23

PageRank Iteration



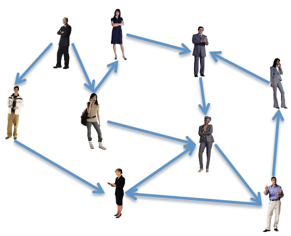
$$R[i] = \alpha + (1-\alpha) \sum_{(j,i) \in E} w_{ji} R[j]$$

- α is the random reset probability
- w_{ji} is the prob. transitioning (similarity) from j to i

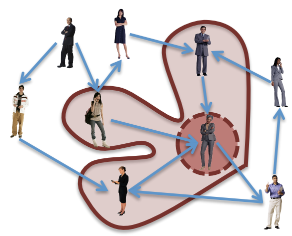
24

Properties of Graph Parallel Algorithms

Dependency Graph



Local Updates



Iterative Computation

Addressing Graph-Parallel ML

Data-Parallel ⟷ Graph-Parallel

Map Reduce

Feature Extraction

Cross Validation

Computing Sufficient Statistics

Graph-Parallel Abstraction

Graphical Models
Gibbs Sampling
Belief Propagation
Variational Opt.

Semi-Supervised Learning
Label Propagation
CoEM

Collaborative Filtering
Tensor Factorization

Data-Mining
PageRank
Triangle Counting

Graph Computation:

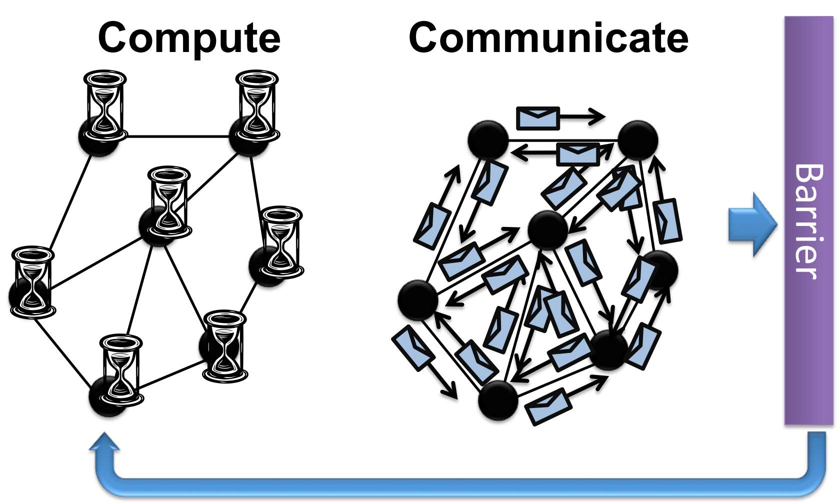
Synchronous

v.

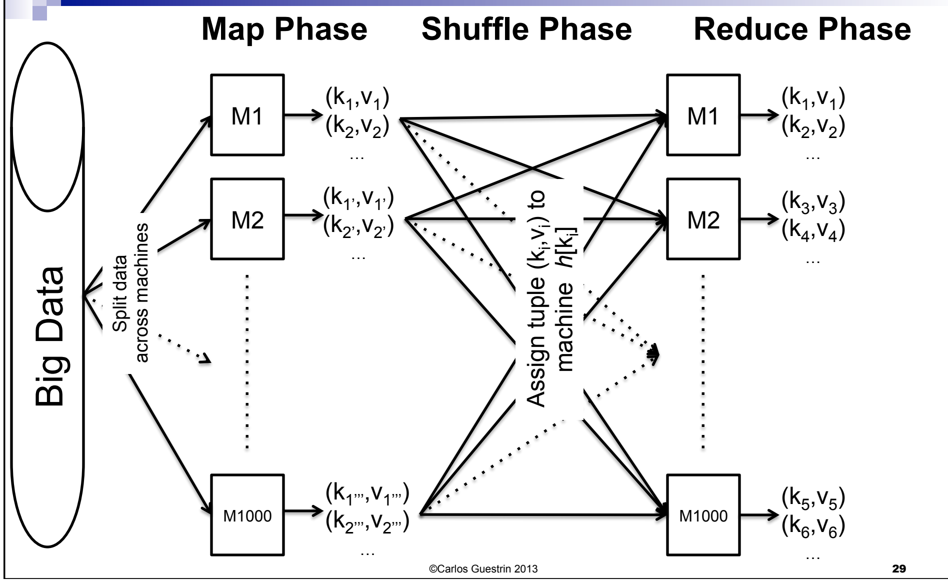
Asynchronous

Bulk Synchronous Parallel Model: Pregel (Giraph)

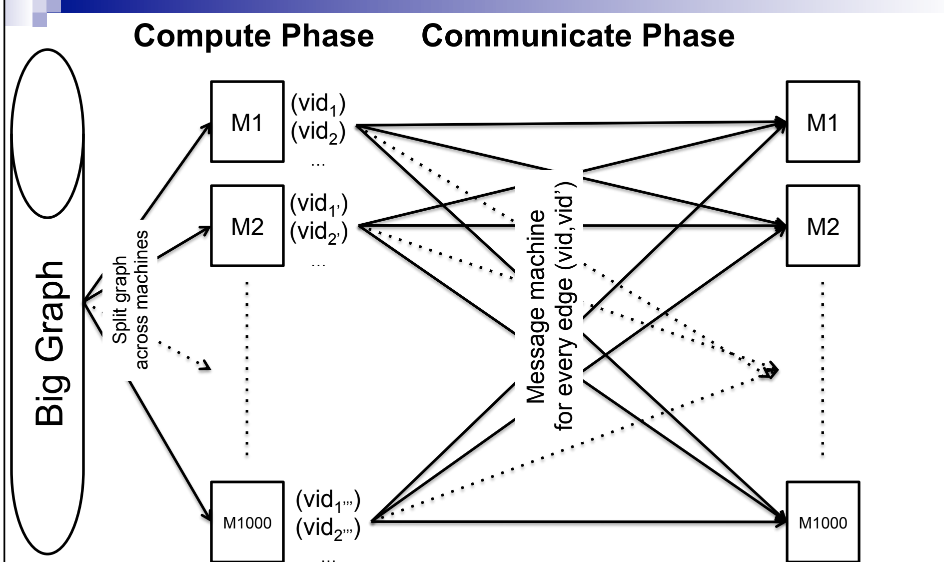
[Valiant '90]



Map-Reduce – Execution Overview



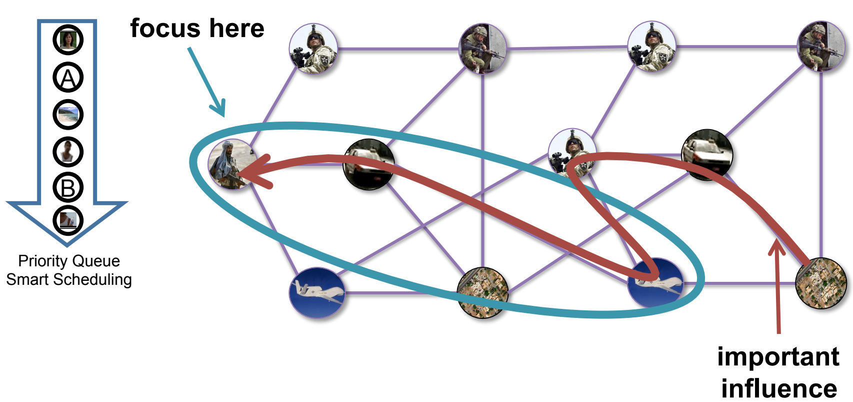
BSP – Execution Overview



Bulk synchronous parallel model ***provably inefficient*** *for some ML tasks*

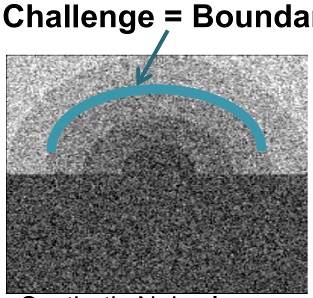
Analyzing Belief Propagation

[Gonzalez, Low, G. '09]

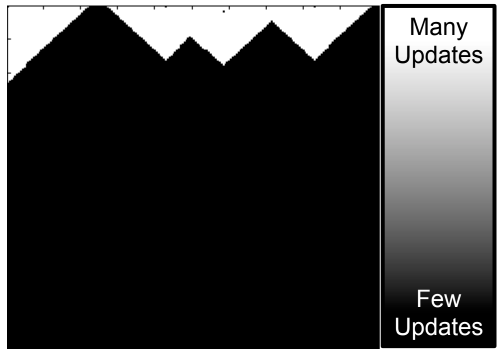


Asynchronous Belief Propagation

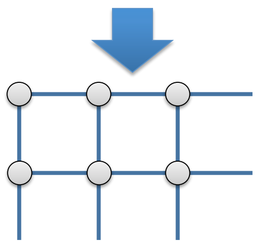
Challenge = Boundaries



Synthetic Noisy Image



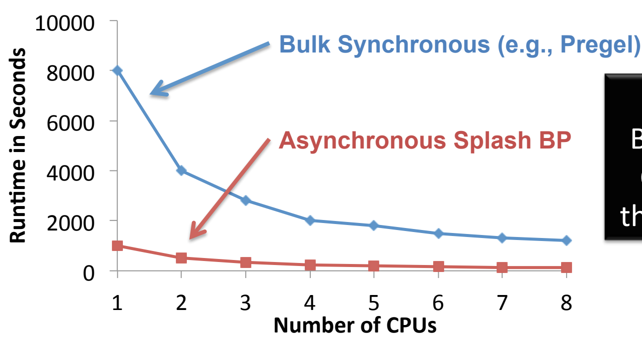
Cumulative Vertex Updates



Graphical Model

Algorithm identifies and focuses on hidden sequential structure

BSP ML Problem: Synchronous Algorithms can be **Inefficient**



Theorem:
Bulk Synchronous BP O(#vertices) slower than Asynchronous BP

Synchronous v. Asynchronous

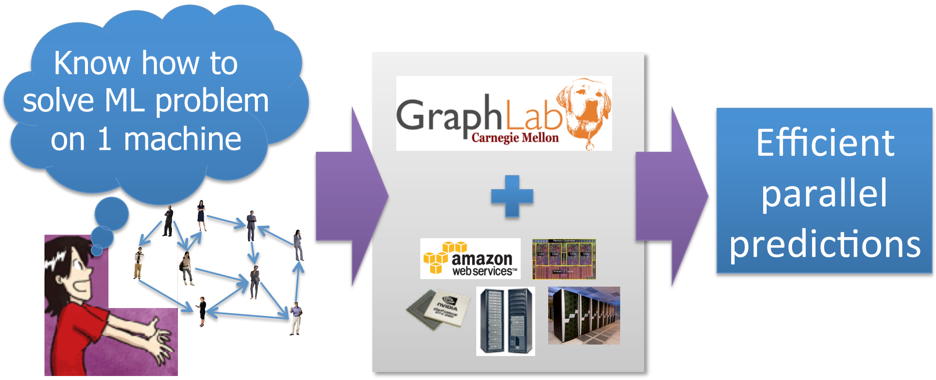
- Bulk synchronous processing:
  - Computation in phases
    - All vertices participate in a phase
      - Though OK to say no-op
    - All messages are sent
  - Simpler to build, like Map-Reduce
    - No worries about race conditions, barrier guarantees data consistency
    - Simpler to make fault-tolerant, save data on barrier
  - Slower convergence for many ML problems
  - In matrix-land, called Jacobi Iteration
  - Implemented by Google Pregel 2010
- Asynchronous processing:
  - Vertices see latest information from neighbors
    - Most closely related to sequential execution
  - Harder to build:
    - Race conditions can happen all the time
      - Must protect against this issue
    - More complex fault tolerance
    - When are you done?
    - Must implement scheduler over vertices
  - Faster convergence for many ML problems
  - In matrix-land, called Gauss-Seidel Iteration
  - Implemented by GraphLab 2010, 2012

Case Study 4: Collaborative Filtering

GraphLab

Machine Learning/Statistics for Big Data
CSE599C1/STAT592, University of Washington
Carlos Guestrin
March 12th, 2013

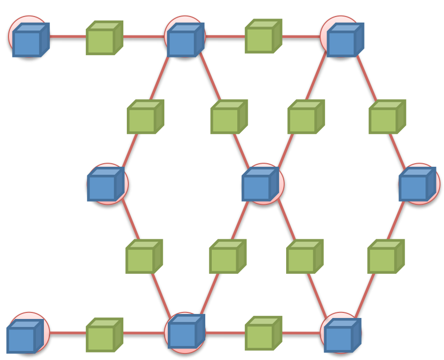
The **GraphLab** Goals



 37

Data Graph

Data associated with vertices and edges



Graph:
- Social Network

Vertex Data:
- User profile text
- Current interests estimates

Edge Data:
- Similarity weights

 38

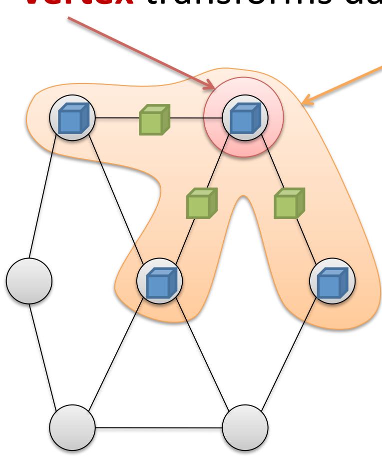
How do we *program* **graph** computation?

"Think like a Vertex."

-Malewicz et al. [SIGMOD'10]

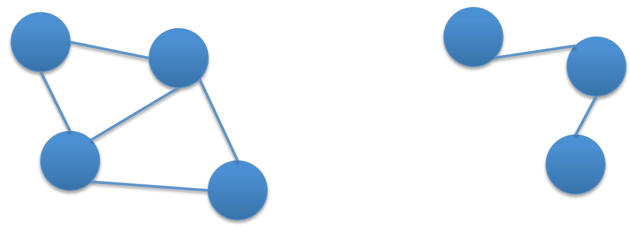
Update Functions

User-defined program: applied to **vertex** transforms data in **scope** of vertex



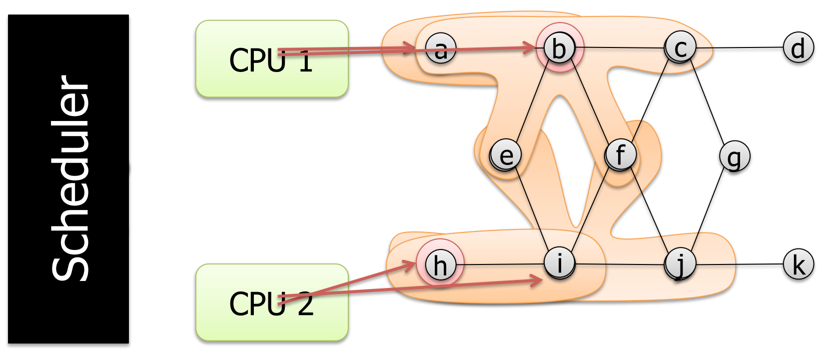
```
pagerank(i, scope){

}
```

Update Function Example: Connected Components



The Scheduler

The **scheduler** determines order vertices are updated

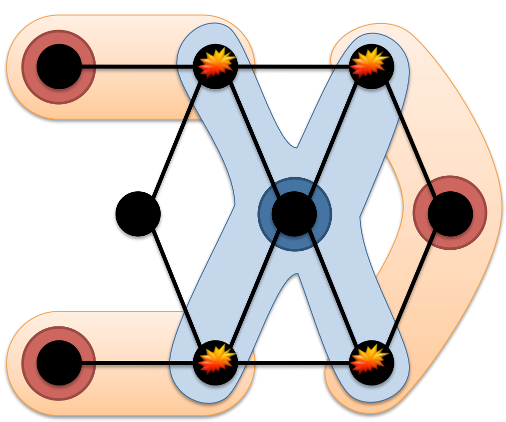


Example Schedulers

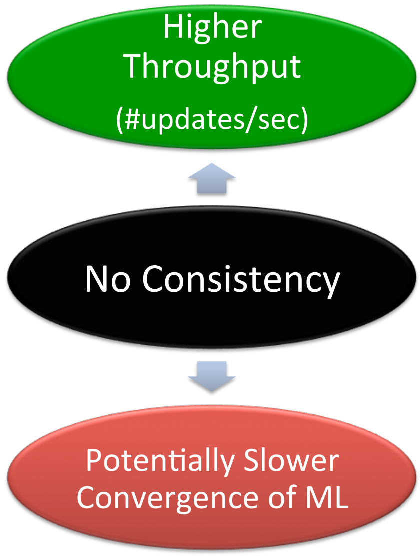
- Round-robin
- Selective scheduling (skipping):
  - round robin but jump over un-scheduled vertice
- FIFO
- Prioritize scheduling
  - Hard to implement in a distributed fashion
    - Approximations used (each machine has its own priority queue)

Ensuring Race-Free Code

How much can computation **overlap**?

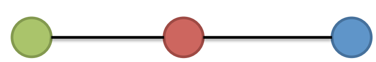


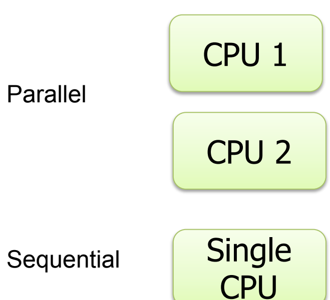
Need for Consistency?



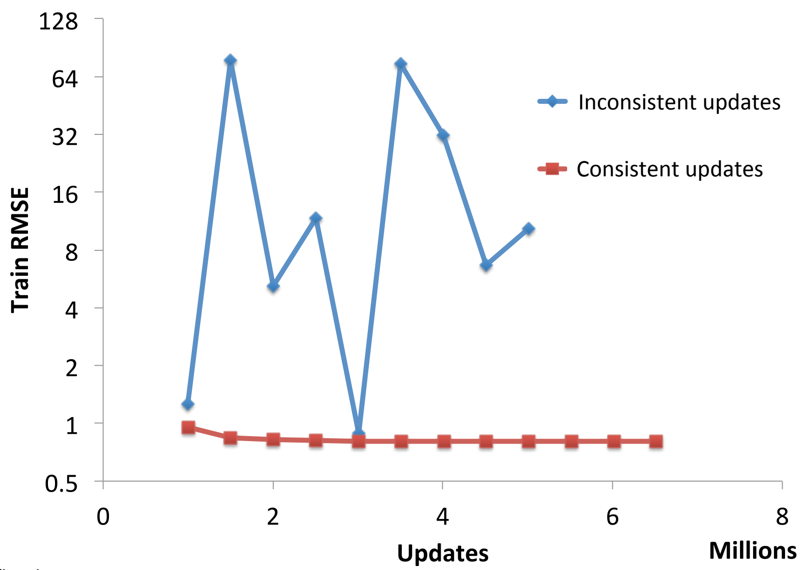
GraphLab Ensures **Sequential Consistency**

For **each parallel execution**, there exists a **sequential execution** of update functions which produces the same result





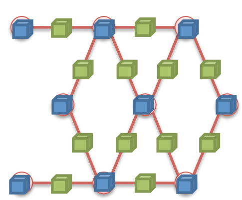
Consistency in Collaborative Filtering



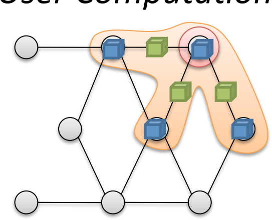
Netflix data, 8 cores

The GraphLab Framework

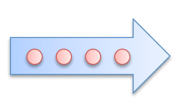
Graph Based
Data Representation



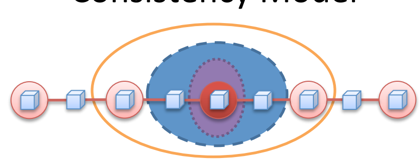
Update Functions
User Computation



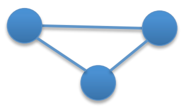
Scheduler



Consistency Model

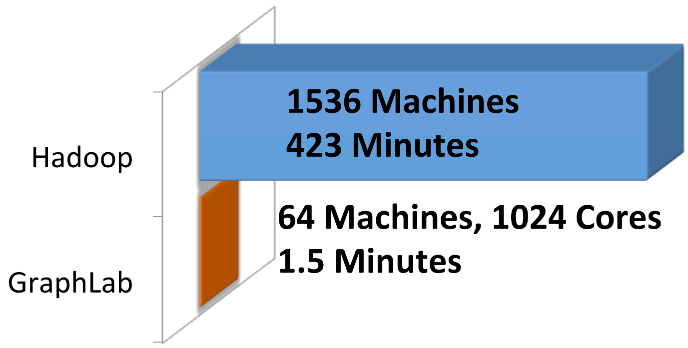


Triangle Counting in Twitter Graph



40M Users
1.2B Edges

Total:
34.8 Billion Triangles



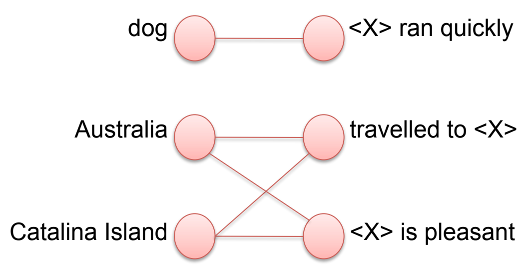
Hadoop results from [Suri & Vassilvitskii '11]

CoEM (Jones et al., 2005)

Named Entity Recognition Task

Is "Dog" an animal?
Is "Catalina" a place?

Vertices: 2 Million
Edges: 200 Million



Never Ending Learner Project (CoEM)

Hadoop	95 Cores	7.5 hrs
Distributed GraphLab	**32 EC2 machines**	**80 secs**

What you need to know…

- Data-parallel versus graph-parallel computation
- Bulk synchronous processing versus asynchronous processing
- GraphLab system for graph-parallel computation
  - Data representation
  - Update functions
  - Scheduling
  - Consistency model