MapReduce
- Developed as a research project out of Google.
- OSDI'04: *"MapReduce: Simplified Data Processing on Large Clusters"*
- **Big Idea:** Create a framework for processing data based on functions that can be "automatically parallelized".
  - Allows many nodes to contribute to processing the data without human design/programming.

MapReduce: Simplified Data Processing on Large Clusters

Jeffrey Dean and Sanjay Ghemawat

jeff@google.com, sanjay@google.com

Google, Inc.

https://static.googleusercontent.com/media/research.google.com/en//archive/mapreduce-osdi04.pdf

MapReduce: Map Functions
- Input:

- Output:

Reduce Function:
- Input:

- Output:

Example #1: Word Count

The	quick	brown	fox	jumps	over	the	lazy	dog
[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]

Map:

Reduce:

Example #2: Mutual Friends
Through asking about your friends about their friends, you have identified who are friends of whom (→ means *"is friends with"*):

```
A → B, C
B → A, C, D
C → A, B, D
D → B, C
```

You want to identify all **mutual friends** to any set of two people. For example: {A, B} → C, D.

Map:

Reduce:

Software as a Service (SaaS)

Q: What makes SaaS different from **just** software?

(1):

(2):

(3):

> Much like all of the other things talked about in the cloud, SaaS is a model of software that is **independent of programming language** and **independent of architectural design.**

Q: What features differentiate **SaaS** from **PaaS**?

(1):

(2):

(3):

> Some software tools have characteristics of both SaaS and PaaS – these are general terms to understand the level of abstraction that you, as a developer, are working with or providing.

Large Categories of SaaS Software
- Human Resource Management (HRM)
- Customer Relationship Management (CRM)
- Enterprise Resource Planning (ERP)
- Content Management Systems (CMS)
- Billing Platforms
- Project Management
- Bookkeeping/Accounting
- Ecomerence
- Communication

Examples of SaaS at Illinois

Advantages of SaaS

1)

2)

3)

4)

Legend:

	Abstracted by Cloud Provider		Customer Managed Unit of Scale		Customer Managed

IaaS Infrastructure as a Service	CaaS Containers as a Service	PaaS Platform as a Service	FaaS Functions as a Service	SaaS Software as a Service
Data	Data	Data	Data	Data
Functions	Functions	Functions	Functions	Functions
Applications	Applications	Applications	Applications	Applications
Runtime	Runtime	Runtime	Runtime	Runtime
Containers*	Containers	Containers*	Containers*	Containers*
OS	OS	OS	OS	OS
Virtualization	Virtualization	Virtualization	Virtualization	Virtualization
Hardware	Hardware	Hardware	Hardware	Hardware