Piattaforme Abilitanti Distribuite
- PAD -

Distributed Enabling Platforms

Nicola Tonellotto
(ISTI, CNR)
nicola.tonellotto@isti.cnr.it

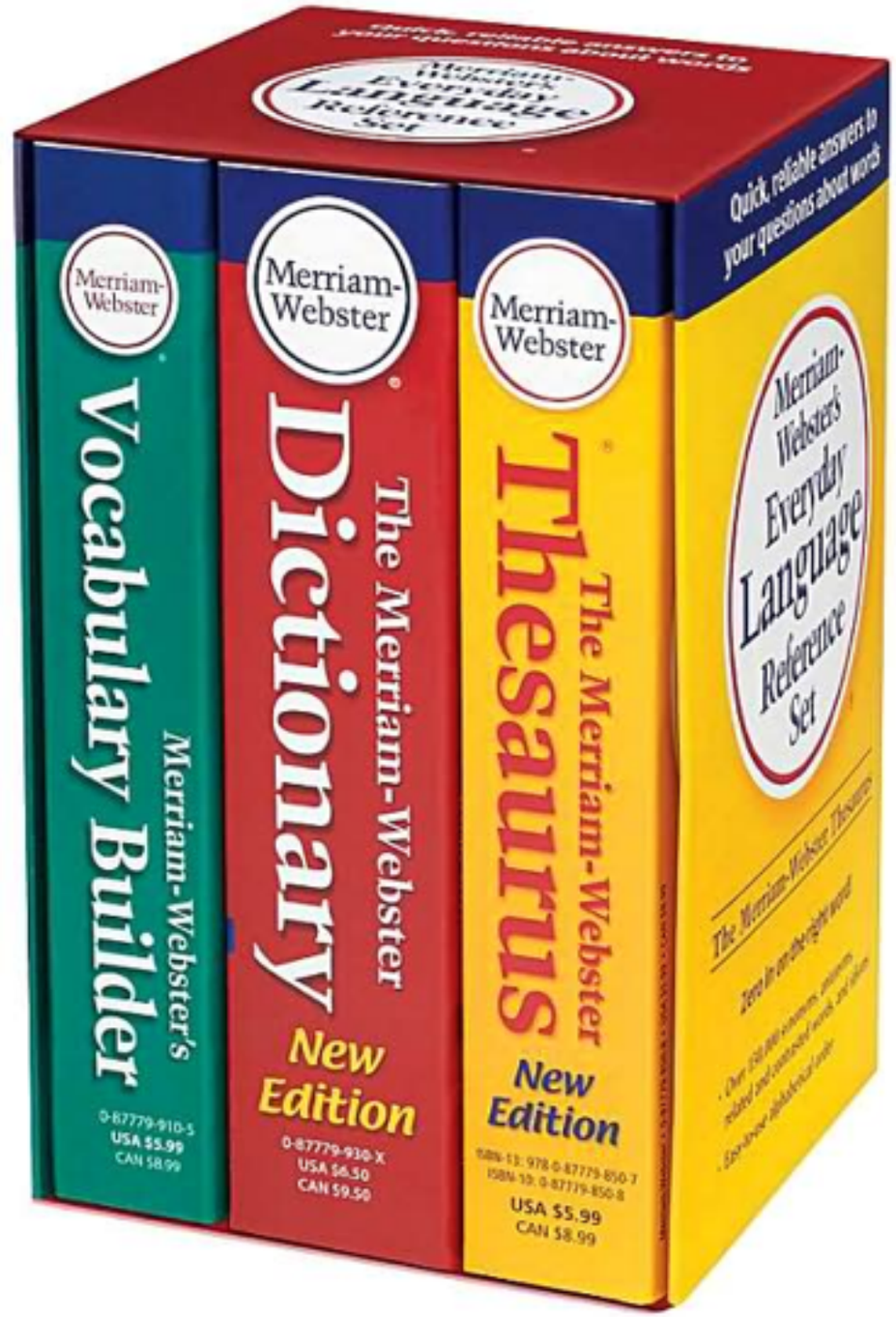
Today

Who?

What?

What is the meaning of words?

- Distributed…
  - relating to a **computer network** in which at least some of the processing is done by the **individual computers** and **information is shared** by and often stored at the computers
- Enabling…
  - to make possible, practical, or easy
- Platforms…
  - the computer architecture and equipment used for a particular purpose



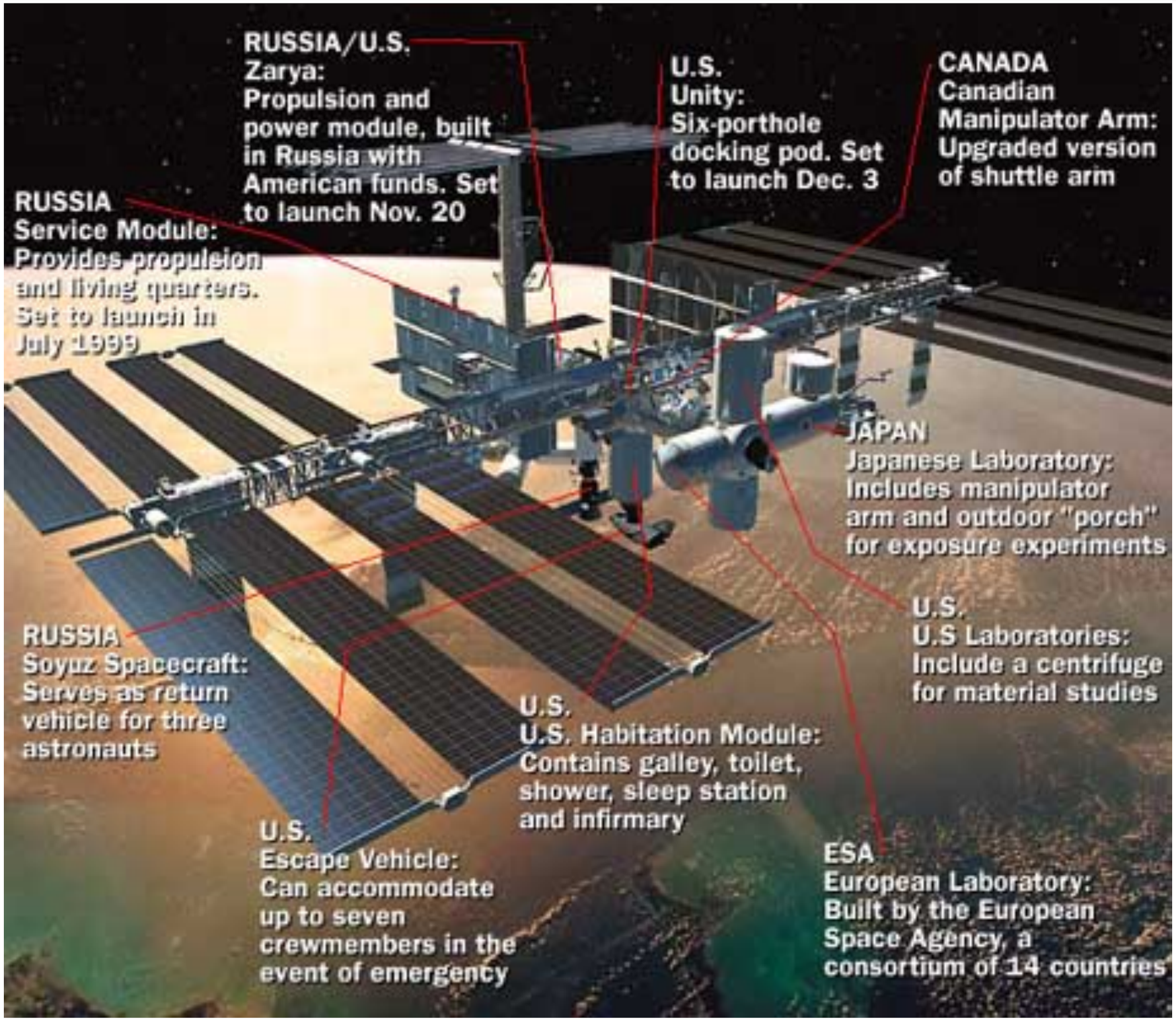
To do what?


$a^n + b^n = c^n$,
$n, a, b, c \in N$,
$n > 2$
$\varphi(x) = \sum_{k \in z} u(k)\sqrt{2}\varphi(2x-k)$
$\hat{\varphi}(\xi) = \prod_{j=1}^{\infty} m_0(\xi/2^j)$
$\langle \psi_{j,k} - \psi_{l,m} \rangle = \delta_{j,l} - \delta_{k,m}$
$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\xi) e^{ix\xi} d\xi$
©Walt Disney Co.
Don Rosa

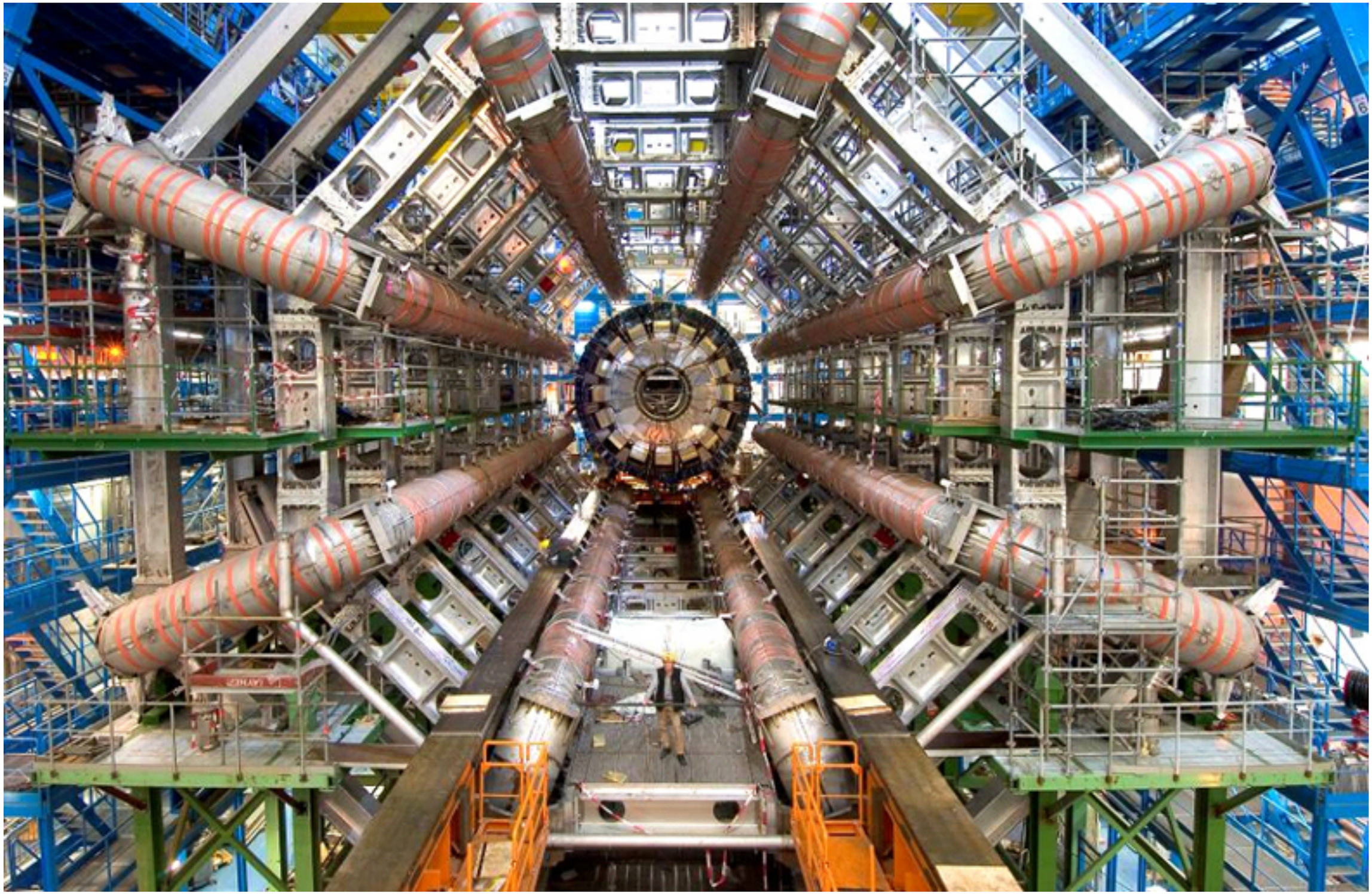
Solve large scale problems!

- In research
  - Frontier research in many different fields today requires world-wide collaborations
  - Online access to expensive scientific instrumentation
  - Scientists and engineers will be able to perform their work without regard to physical location
  - Simulations of world-scale mathematical models
  - Batch analysis of gazillion-bytes of experimental data
- In business
  - Crawling, indexing, searching the Web
  - Web 2.0 applications
  - Mining information
  - Highly interactive applications
  - Online analysis of gazillion-bytes of usage data

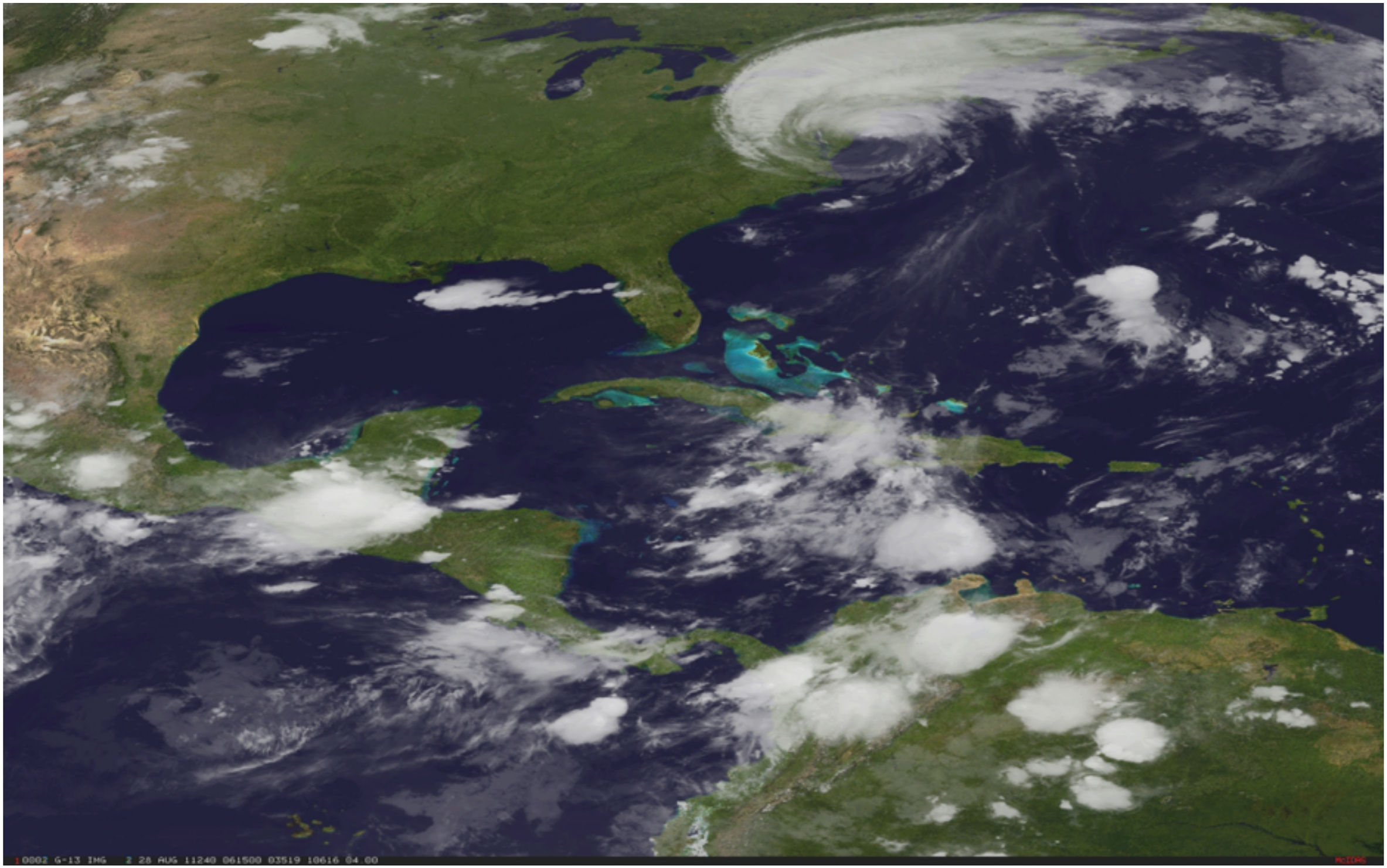
World-wide Collaborations



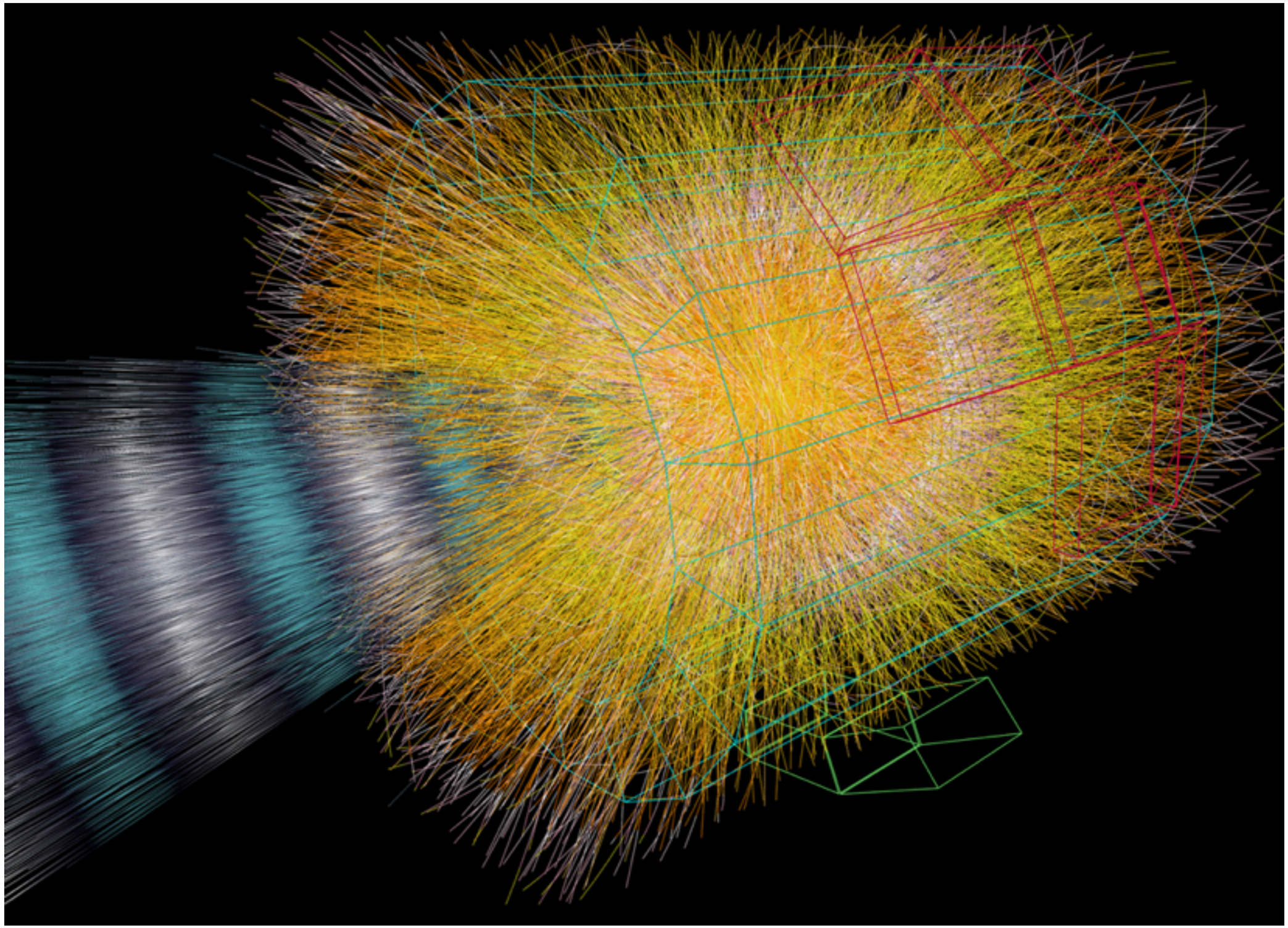
Expensive Scientific Instruments



World-scale Simulations



Batch analysis of huge data



Managing the Web

Google™

Web Images Groups News Froogle Local more »

Google Search I'm Feeling Lucky

Advanced Search
Preferences
Language Tools

Advertising Programs - Business Solutions - About Google

©2005 Google - Searching 8,058,044,651 web pages

Web 2.0

Online analysis of huge data

Some Numbers (I)

- 10,033 Tweets sent in 1 second
- 2,623 Instagram photos uploaded in 1 second
- 2,199 Tumblr posts in 1 second
- 1,862 Skype calls in 1 second
- 29,667 GB of Internet traffic in 1 second
- 50,512 Google searches in 1 second
- 107,401 YouTube videos viewed in 1 second
- 2,425,138 Emails sent in 1 second

source: http://www.internetlivestats.com

Some Numbers (II)

- 3.2 billions of Internet users
- 928 millions of Web sites
- 168 billions of email sent during the day
- 3.2 billions of google searches during the day
- 2.9 millions blog posts written during the day
- 640 millions of tweets sent during the day
- 6.8 billions of videos viewed on Youtube during the day
- 167 millions of Photos uploaded on Instagram during the day
- 1.5 billions of Facebook active users
- 1.5 billions of Google+ active users
- 328 millions of Twitter active users
- 120 millions of Skype calls during the day
- 40 thousands of Web sites hacked during the day
- 503 thousand of computers sold today
- 4 millions of smartphones sold today
- 772 thousands of tablets sold today
- 1.9 billions of GB (1.9 EB) Internet traffic today

source: http://www.internetlivestats.com

The Changing Nature of Data

"traditional" data	BIG DATA
gigabytes to terabytes	PETABYTES TO EXABYTES
centralized	DISTRIBUTED
structured	SEMI-STRUCTURED AND UNSTRUCTURED
stable data model	FLAT SCHEMAS
known complex interrelationships	FEW COMPLEX INTERRELATIONSHIPS

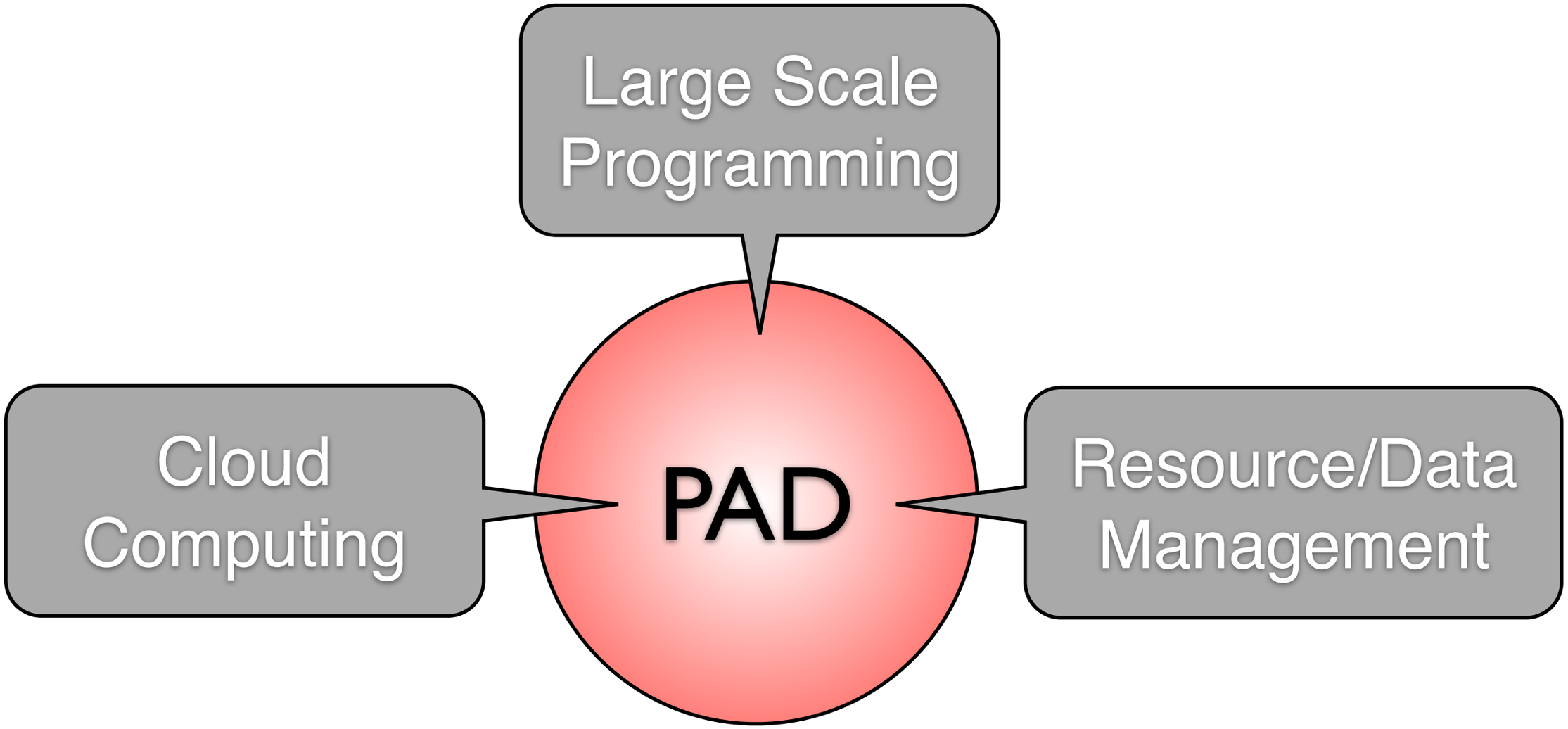
taken from: http://wikibon.org/

Modern Use Cases

- Recommendation Engine
- Sentiment Analysis
- Risk Modeling
- Fraud Detection
- Marketing Campaign Analysis
- Customer Churn Analysis
- Social Graph Analysis
- Customer Experience Analytics
- Network Monitoring
- Research And Development

How?

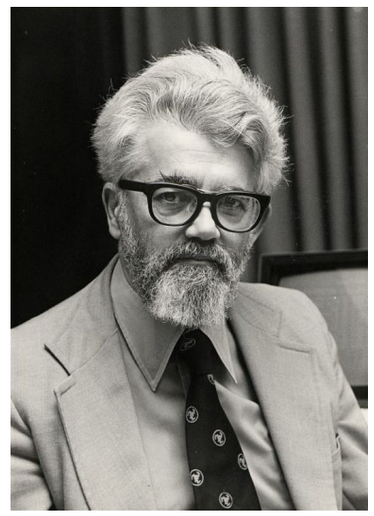
(not so?) Hot Technologies



Famous predictions

1961

[...] computing may someday be organized as a public utility just as telephone system is a public utility [...] the computer utility could become the basis of a new and important industry [...]



John McCarthy (1927-2011)
Turing Award (1971)
Artificial Intelligence

1969

As of now, computer networks are still in their infancy, but as they group up and become sophisticated, we will probably see the spread of computer utilities which, like present electric and telephone utilities, will service individual homes and offices across the country.

Leonard Kleinrock (1934)
Queueing Theory

The 5th Utility

Computing is being transformed to a model consisting of services that are commoditized and delivered in a manner similar to traditional utilities

Demand for more computing power

- There are three ways to improve performance:
  - Work smarter
  - Work harder
  - Get help

- In computing:
  - Using optimized algorithms and techniques
  - Using faster hardware
  - Using multiple computers

Cluster Computing

- A cluster is a type of parallel and distributed system, which consists of a collection of inter-connected stand-alone computers working together as a single integrated computing resource.

- Basic element is the node, a single or multiprocessor system with memory, I/O and OS
- Generally two or more nodes connected together
- In a single rack, or physically separated and connected via a LAN
- Appears as a single system to users and applications
- Specialized access, management and programming

Utility Computing History

1990

2010

Grid Computing

Solving large scale problems with parallel computing

Utility Computing

Offering computing resources as a metered service

Software as a Service

Network-based subscriptions to applications

Cloud Computing

Anytime anywhere access to resources delivered dynamically as a service

Once upon a time...

Microcomputer

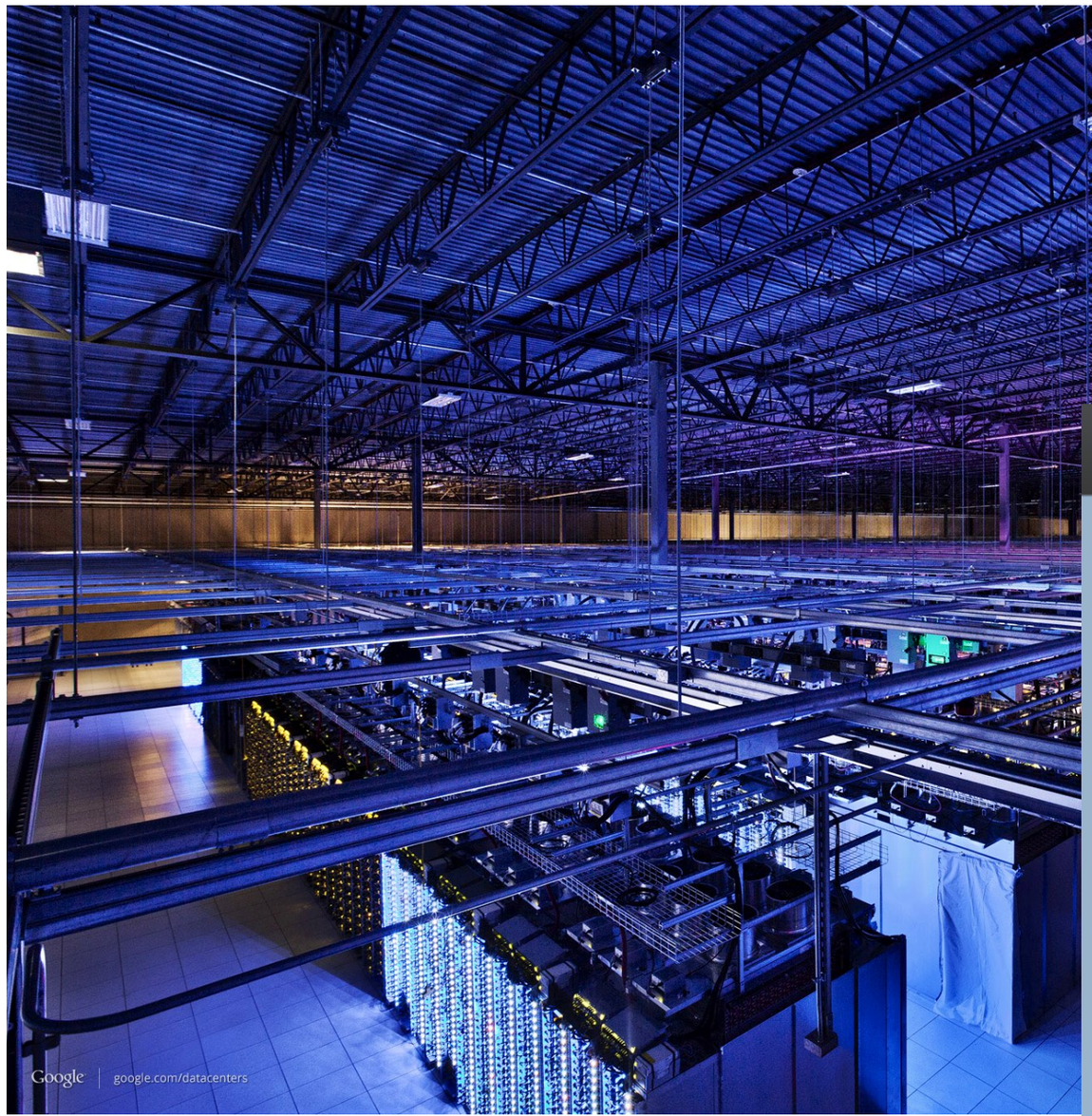
Minicomputer

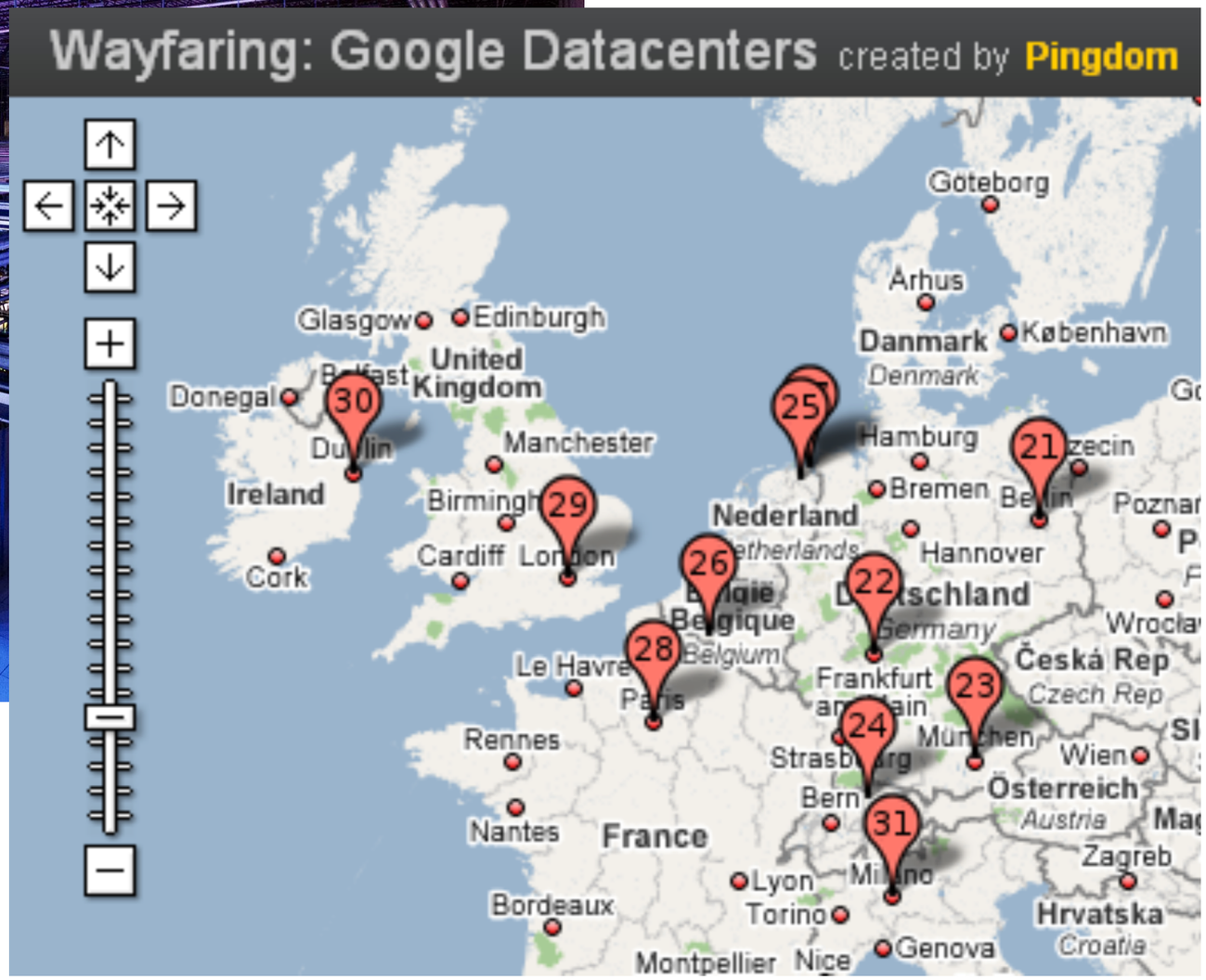
Cluster

Mainframe

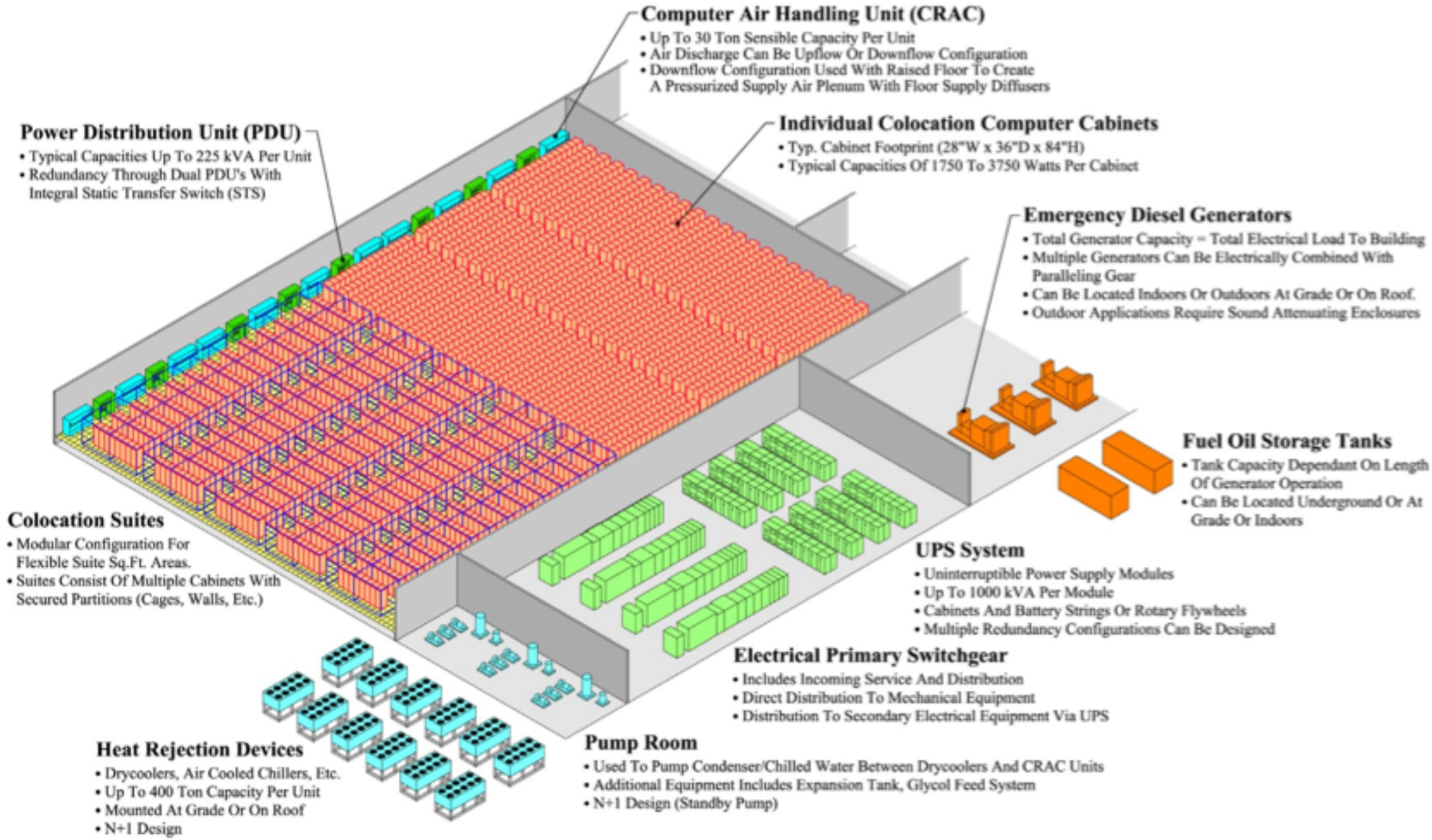
...up to the Grid...

...up to the Cloud





Datacenter


Computer Air Handling Unit (CRAC)
• Up To 30 Ton Sensible Capacity Per Unit
• Air Discharge Can Be Upflow Or Downflow Configuration
• Downflow Configuration Used With Raised Floor To Create A Pressurized Supply Air Plenum With Floor Supply Diffusers
Power Distribution Unit (PDU)
• Typical Capacities Up To 225 kVA Per Unit
• Redundancy Through Dual PDU's With Integral Static Transfer Switch (STS)
Individual Colocation Computer Cabinets
• Typ. Cabinet Footprint (28"W x 36"D x 84"H)
• Typical Capacities Of 1750 To 3750 Watts Per Cabinet
Emergency Diesel Generators
• Total Generator Capacity = Total Electrical Load To Building
• Multiple Generators Can Be Electrically Combined With Paralleling Gear
• Can Be Located Indoors Or Outdoors At Grade Or On Roof.
• Outdoor Applications Require Sound Attenuating Enclosures
Fuel Oil Storage Tanks
• Tank Capacity Dependant On Length Of Generator Operation
• Can Be Located Underground Or At Grade Or Indoors
Colocation Suites
• Modular Configuration For Flexible Suite Sq.Ft. Areas.
• Suites Consist Of Multiple Cabinets With Secured Partitions (Cages, Walls, Etc.)
UPS System
• Uninterruptible Power Supply Modules
• Up To 1000 kVA Per Module
• Cabinets And Battery Strings Or Rotary Flywheels
• Multiple Redundancy Configurations Can Be Designed
Electrical Primary Switchgear
• Includes Incoming Service And Distribution
• Direct Distribution To Mechanical Equipment
• Distribution To Secondary Electrical Equipment Via UPS
Heat Rejection Devices
• Drycoolers, Air Cooled Chillers, Etc.
• Up To 400 Ton Capacity Per Unit
• Mounted At Grade Or On Roof
• N+1 Design
Pump Room
• Used To Pump Condenser/Chilled Water Between Drycoolers And CRAC Units
• Additional Equipment Includes Expansion Tank, Glycol Feed System
• N+1 Design (Standby Pump)

Search Engines Architecture

Components in a Search Engine

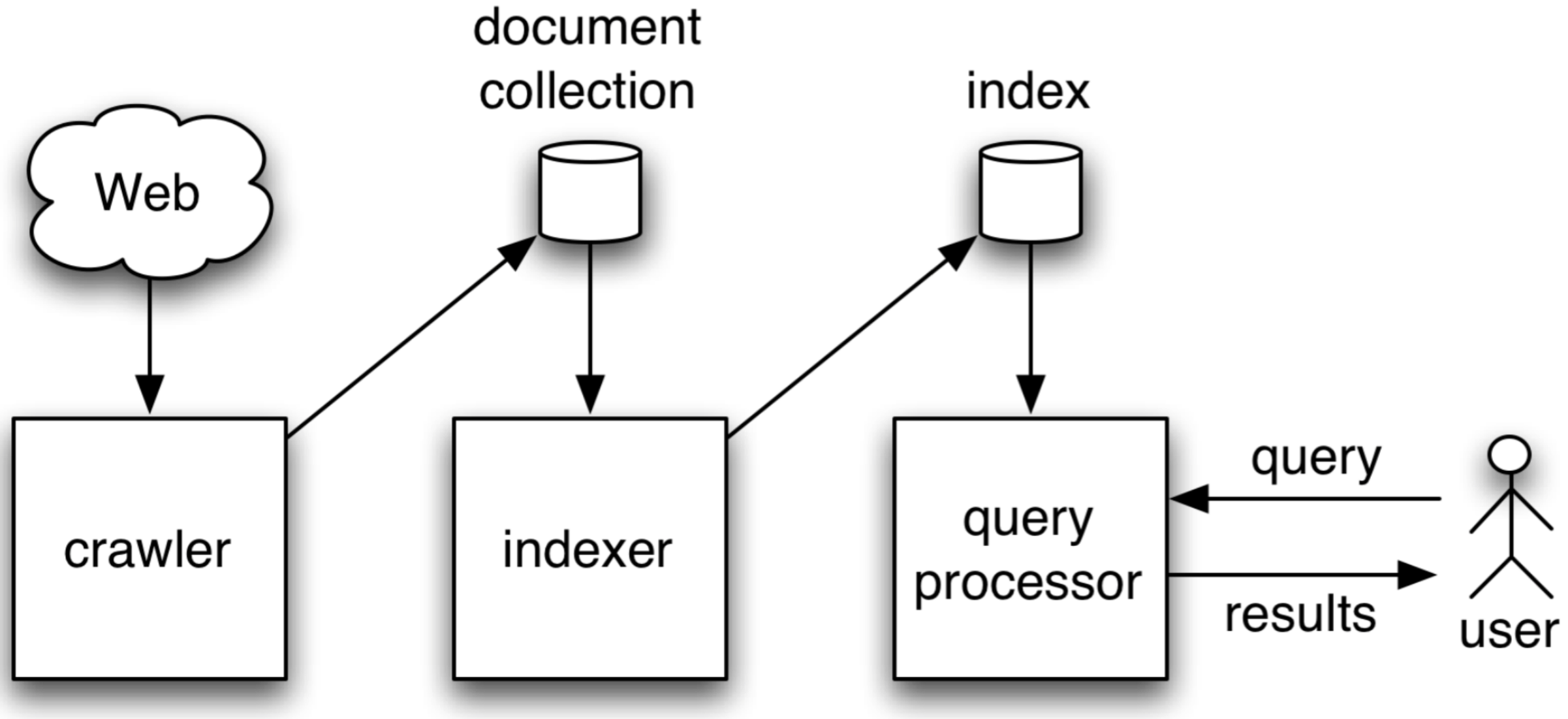


Image courtesy of B. B. Cambazoglu & R. Baeza-Yates, SIGIR 2016

Single Computer Setting

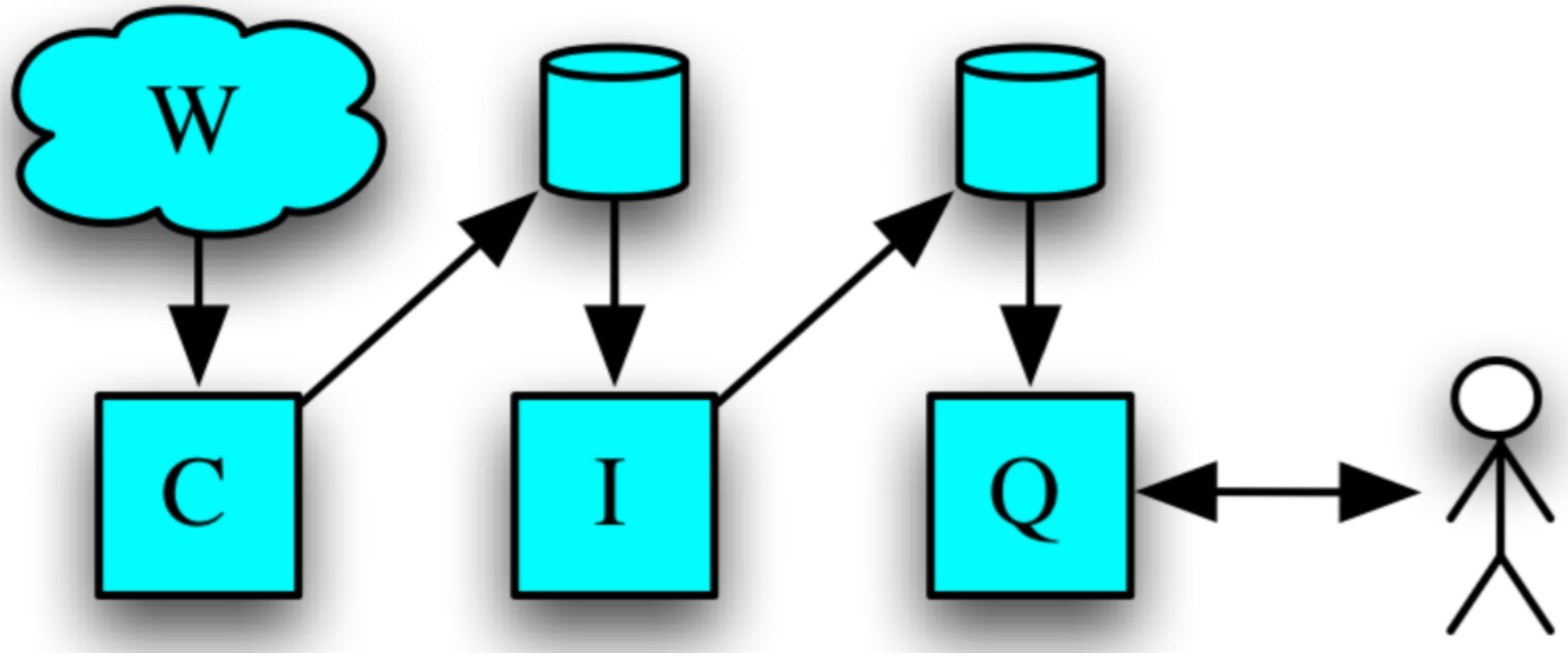


Image courtesy of B. B. Cambazoglu & R. Baeza-Yates, SIGIR 2016

Single Cluster Setting

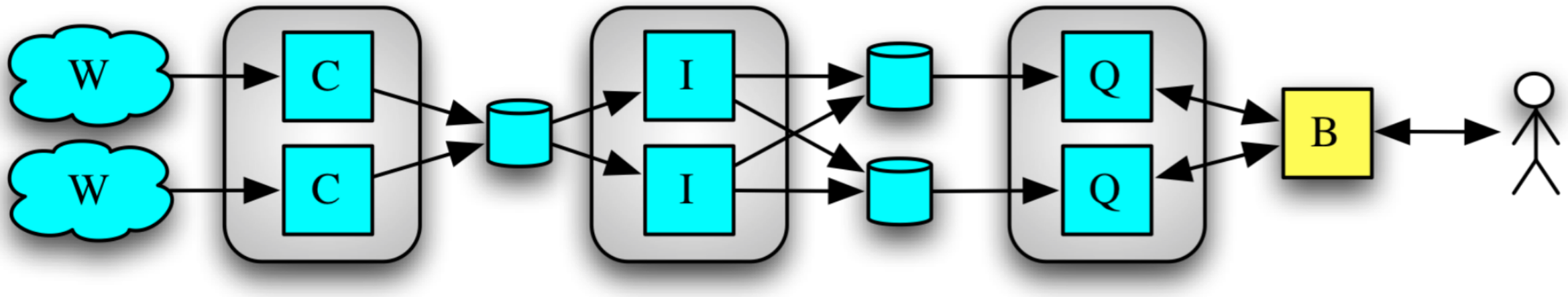


Image courtesy of B. B. Cambazoglu & R. Baeza-Yates, SIGIR 2016

Multiple Cluster Setting

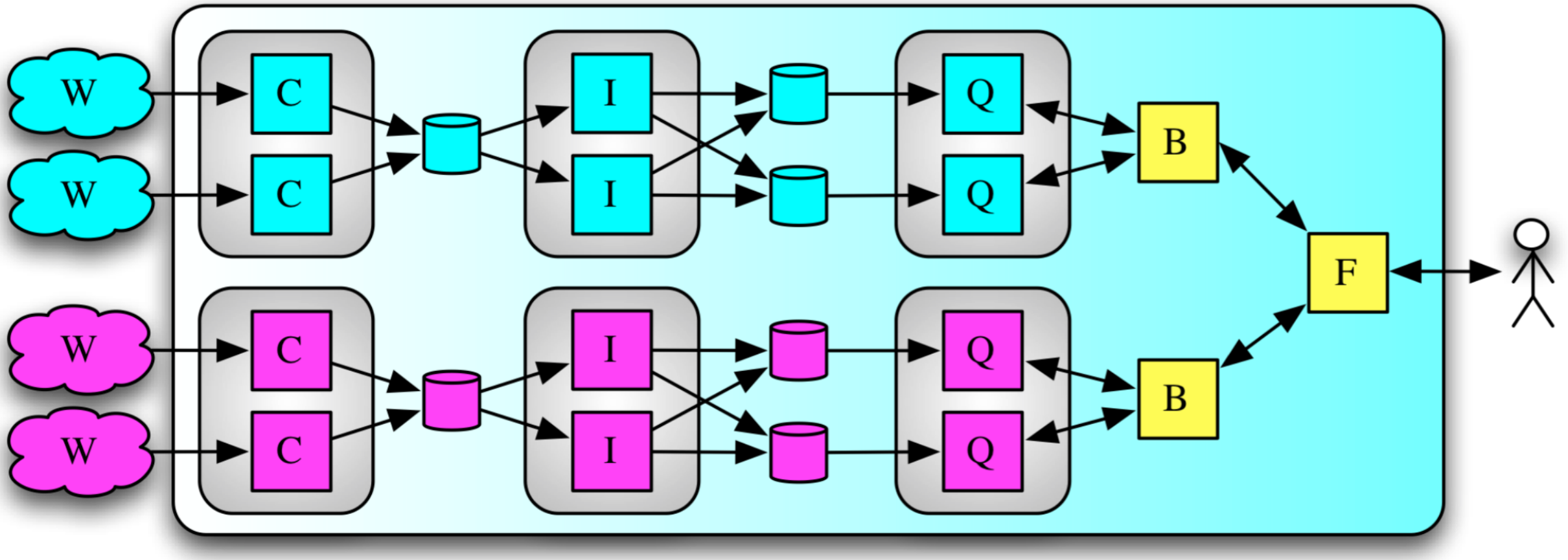


Image courtesy of B. B. Cambazoglu & R. Baeza-Yates, SIGIR 2016

Multisite Setting

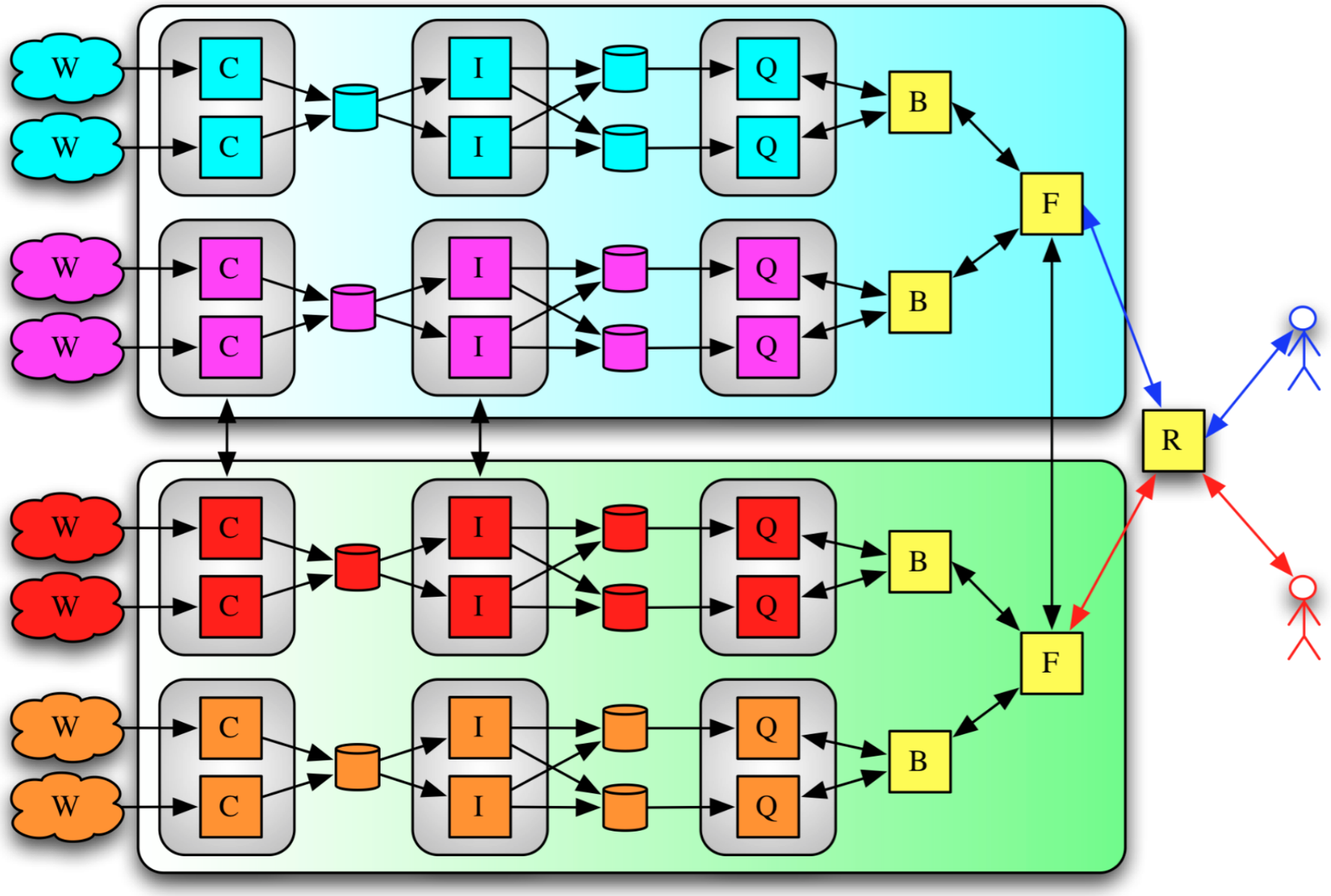


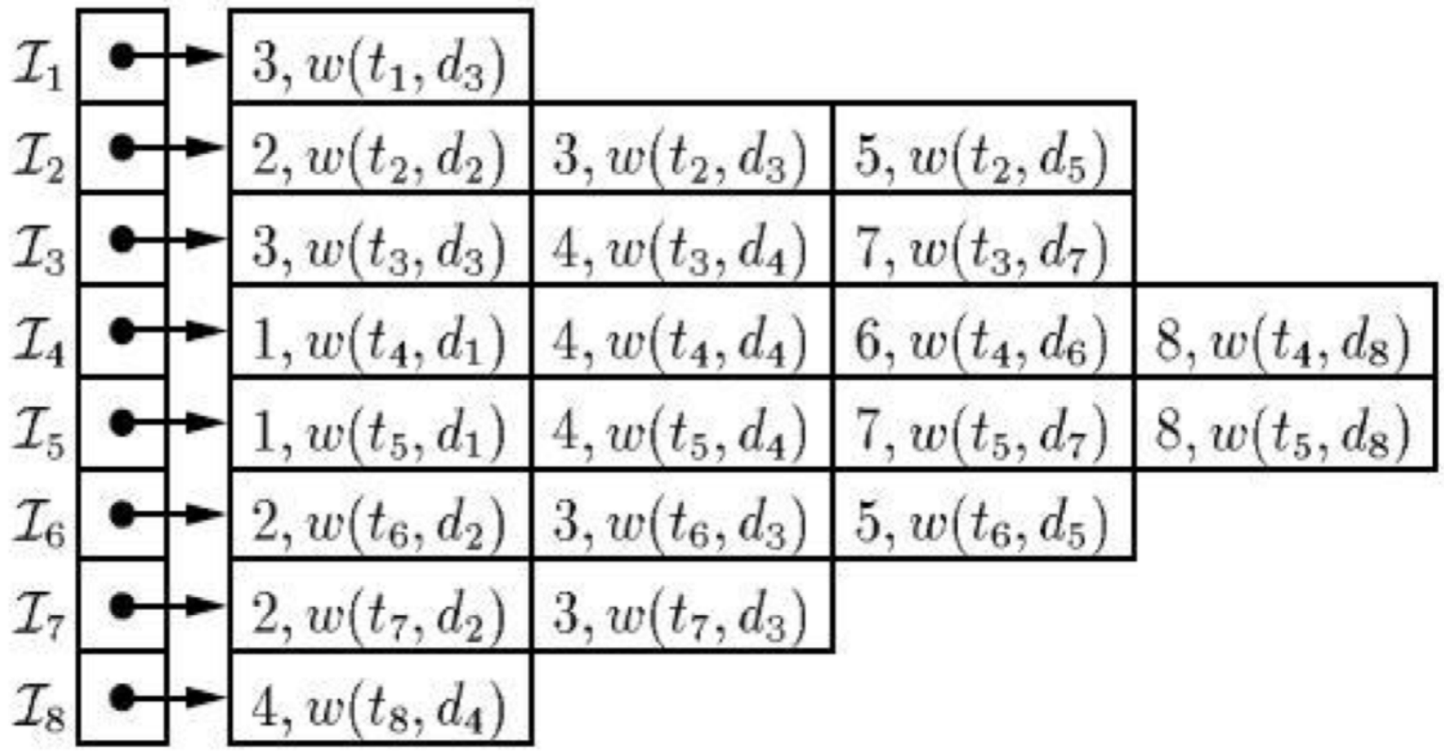
Image courtesy of B. B. Cambazoglu & R. Baeza-Yates, SIGIR 2016

Web Crawling

- Web crawling is the process of locating, fetching, and storing the pages on the Web
- The computer programs that perform this task are referred to as Web crawlers or spiders
- A typical Web crawler
  - starts from a set of seed pages,
  - locates new pages by parsing the downloaded seed pages,
  - extracts the hyperlinks within,
  - stores the extracted links in a fetch queue for retrieval,
  - continues downloading until the fetch queue gets empty or a satisfactory number of pages are downloaded.

Indexing

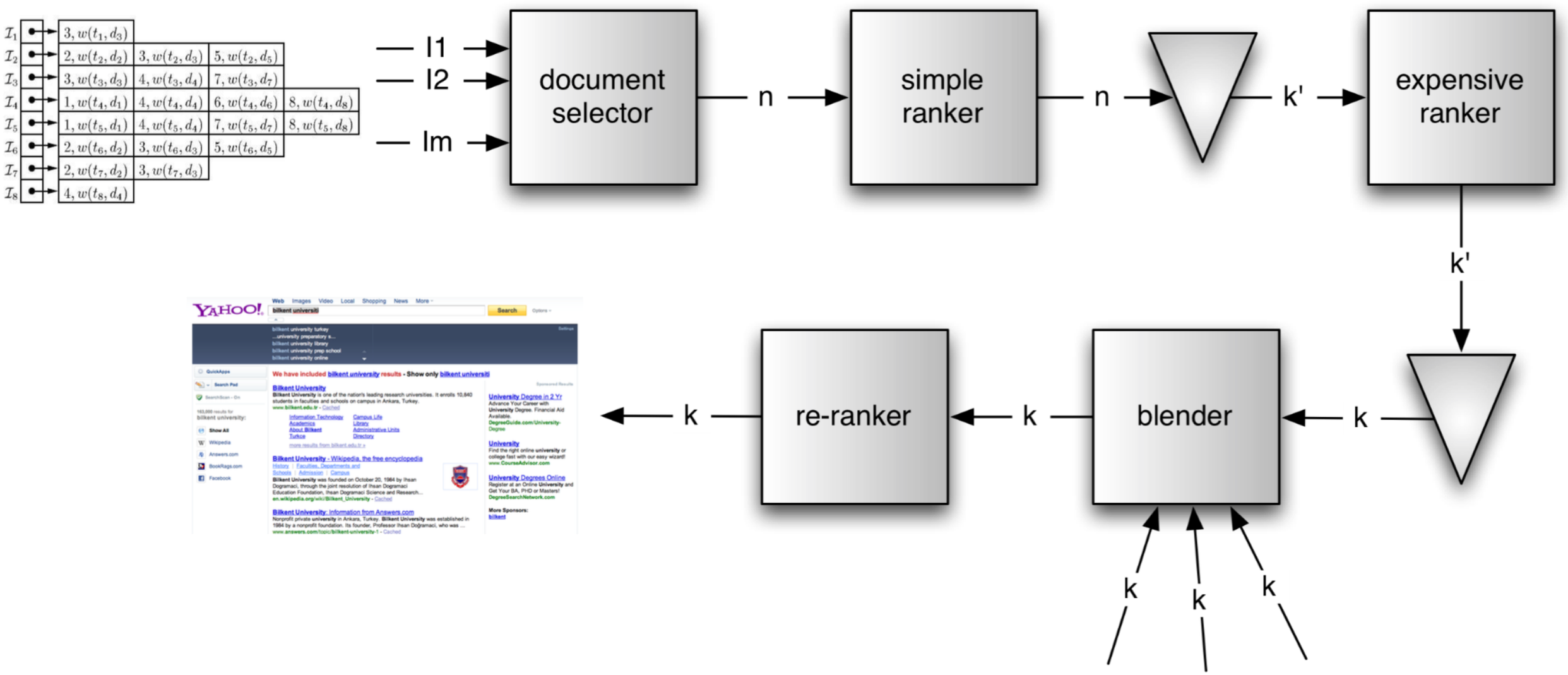
- Indexing is the process of converting crawled web documents into an efficiently searchable form
- An inverted index is a representation for the document collection over which user queries will be evaluated
- Alternatives
  - signature files
  - suffix arrays
- An index has two parts
  - a set of inverted lists
    - a set posting entries
      - document id
      - word score
      - word positions
  - an index into these lists



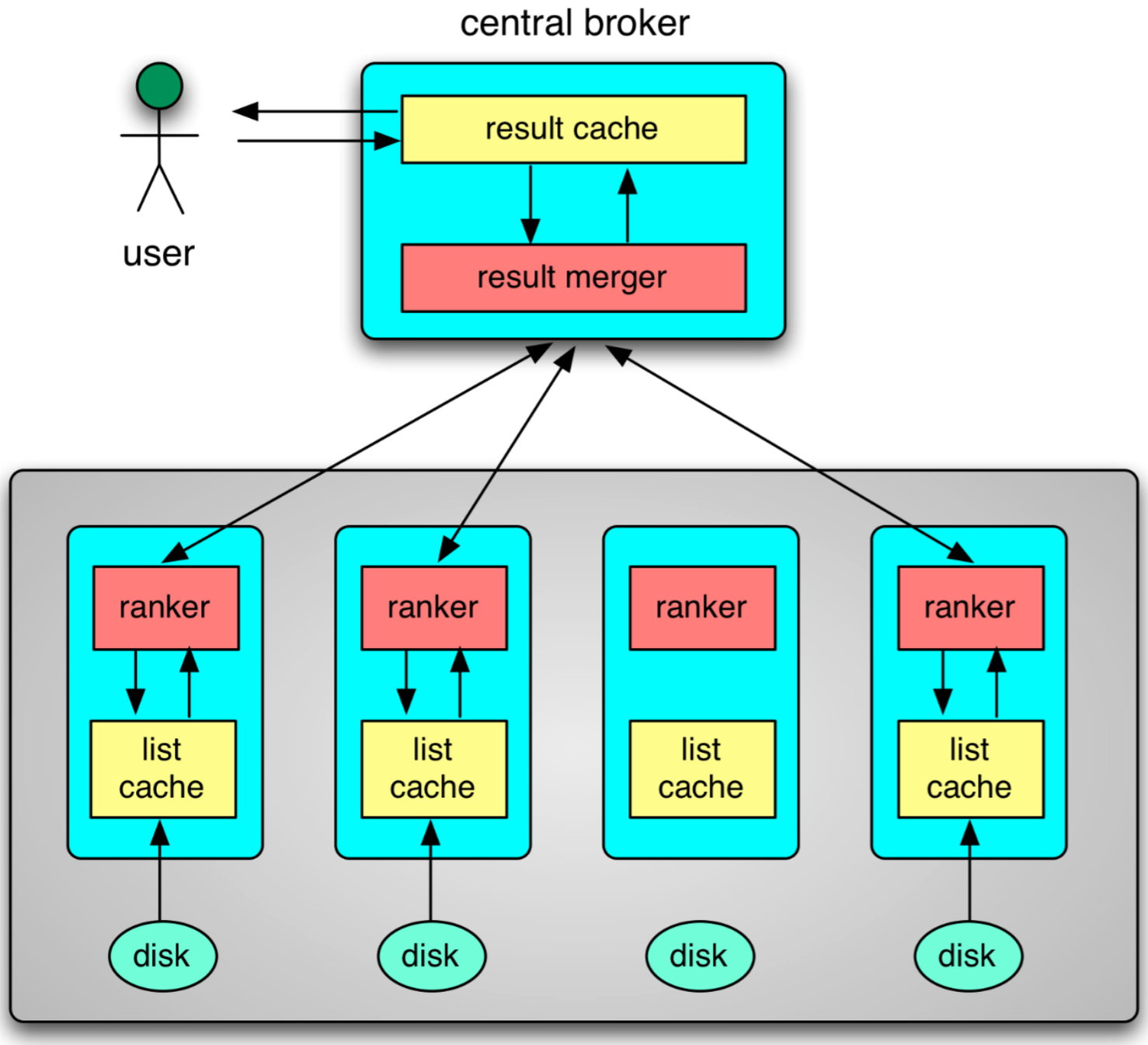
Query Processing...

- Query processing is the problem of generating the best- matching answers (typically, top 10 documents) to a given user query, spending the least amount of time
- Many features
  - term statistics (e.g., BM25)
  - term proximity
  - link analysis (e.g., PageRank) – spam detection
  - click data
  - search session analysis
  - social media

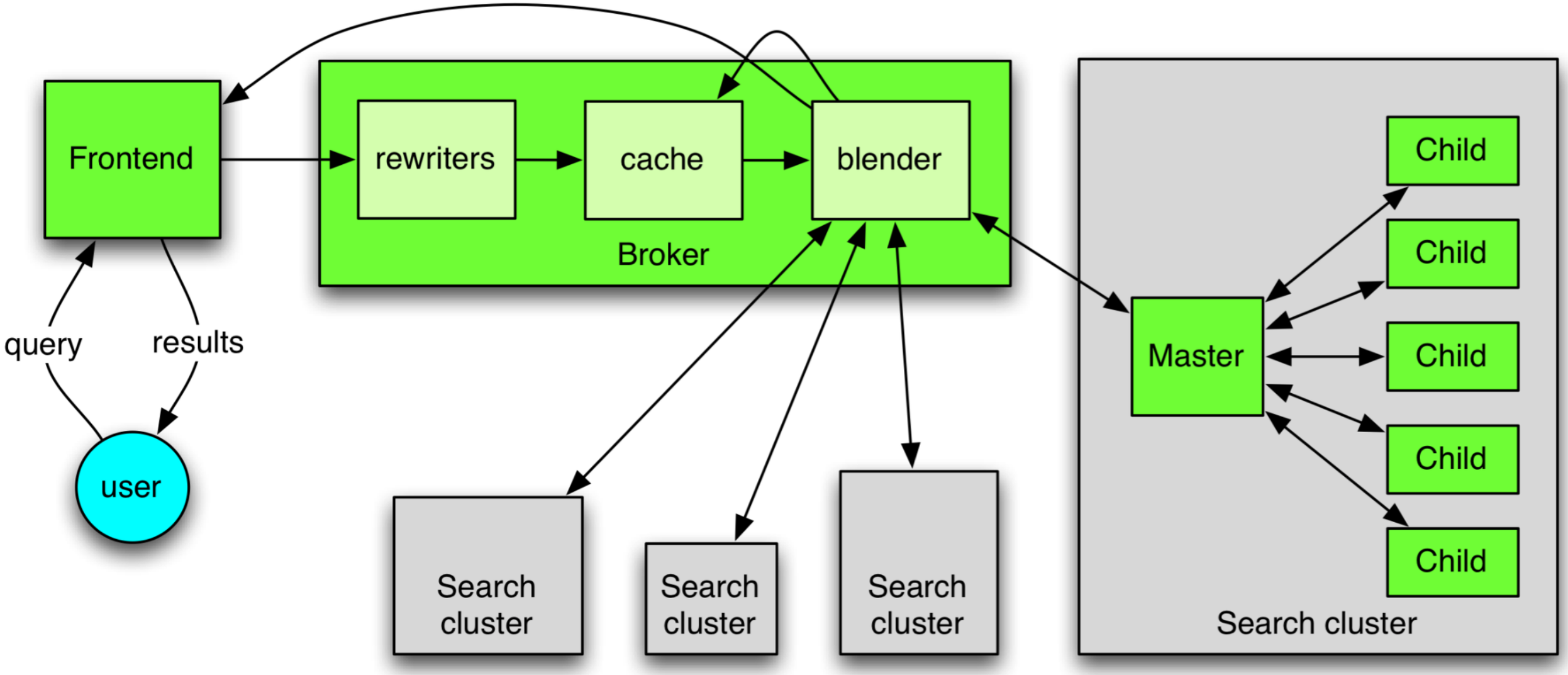
...in a single computer...



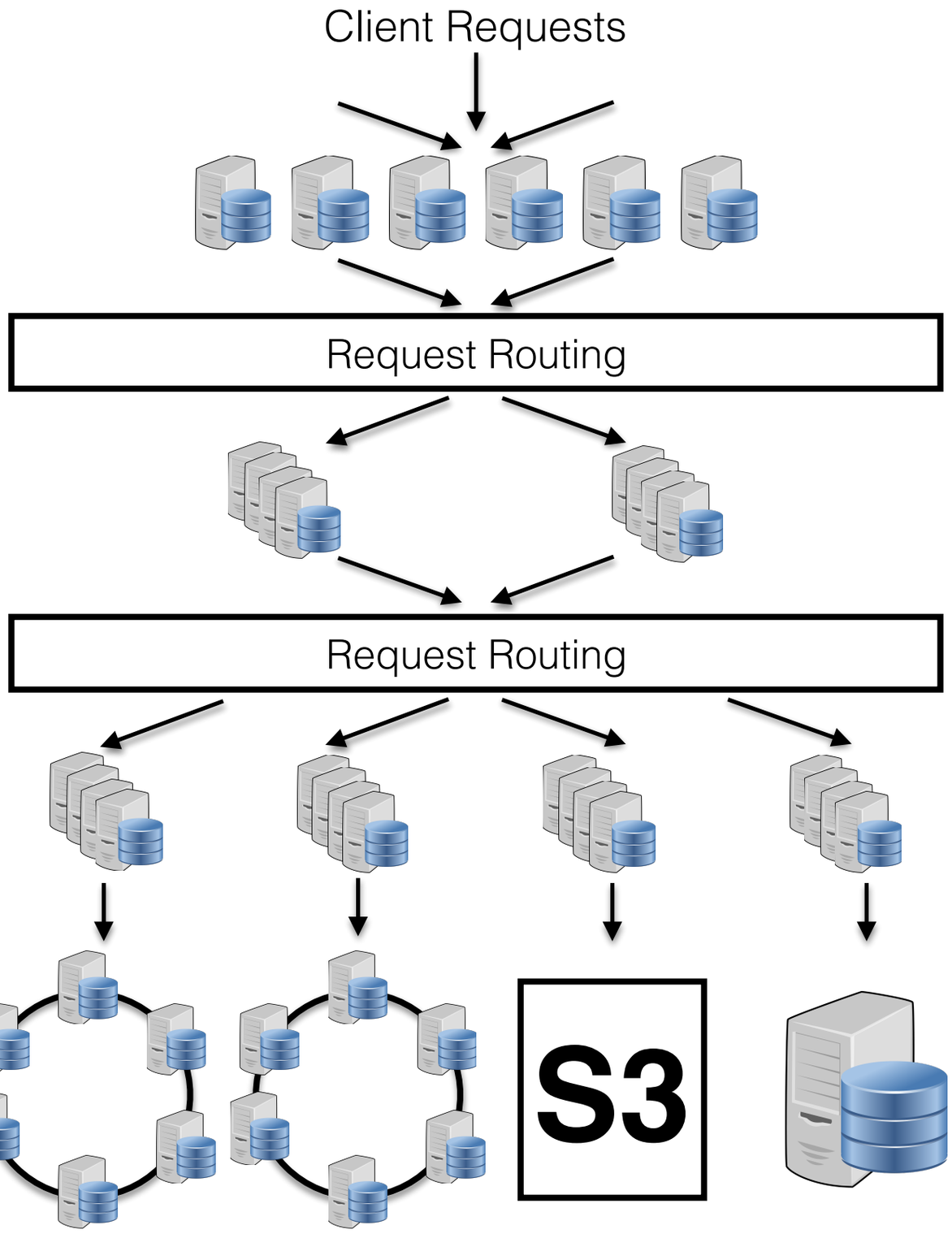
...in a single cluster...



...in a multiple cluster/datacenter...



Distributed Data Storage Architecture



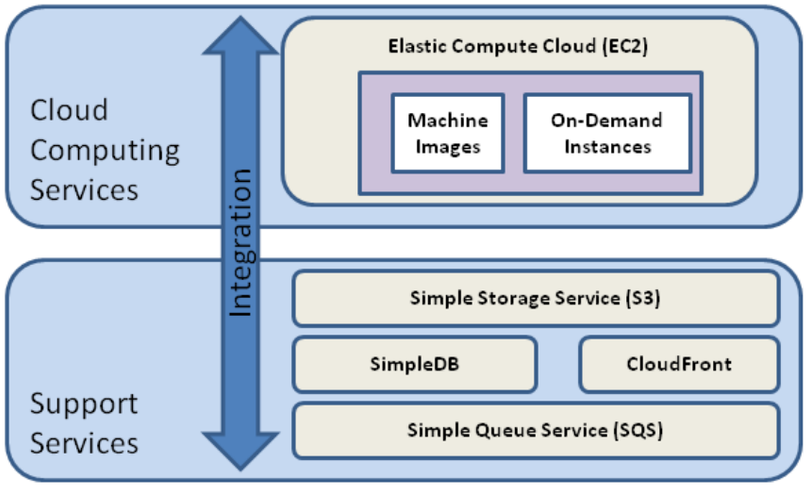
Page Rendering Services

Stateless Aggregator Services

Stateful Services

Datastores

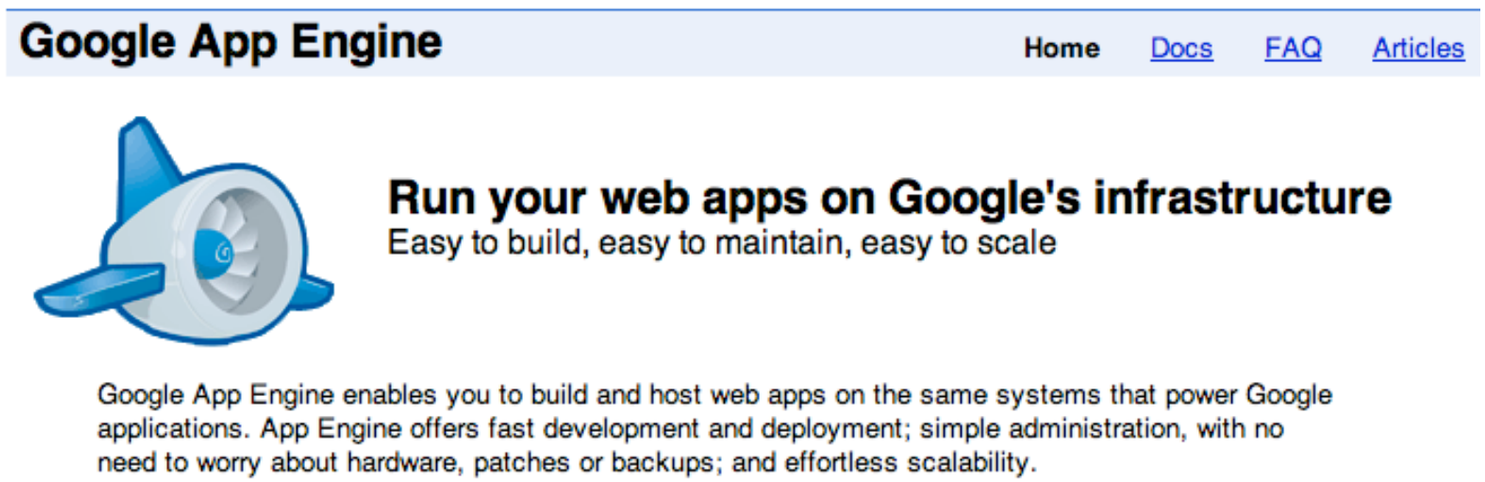
Everything as a Service



Infrastructure

Software

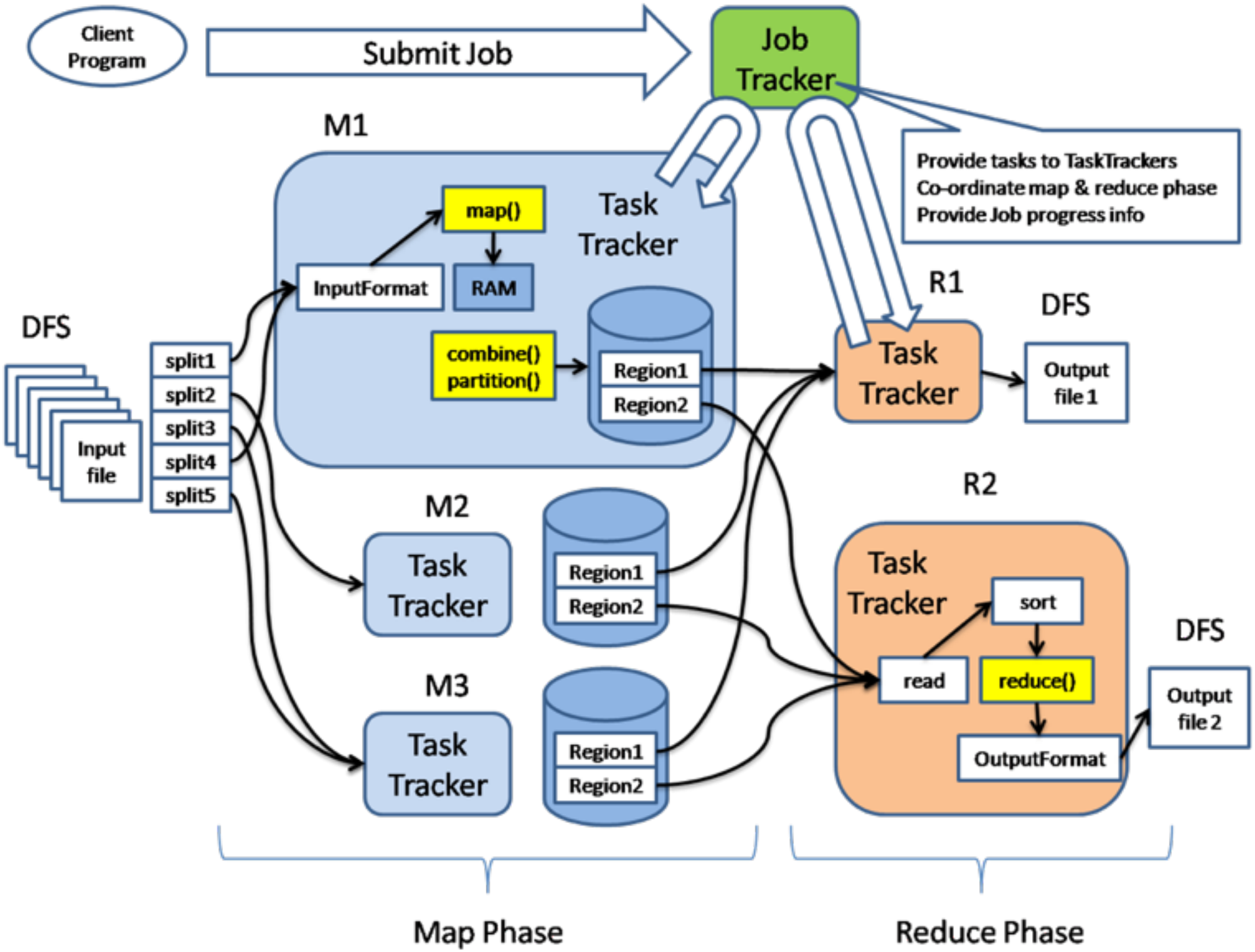
Platform



Large Scale Programming

HDFS	Google File System
Hadoop	Google MapReduce
HBase	Google BigTable

Map Reduce


Client Program
Submit Job
Job Tracker
M1
Task Tracker
map()
InputFormat
RAM
combine()
partition()
Region1
Region2
Provide tasks to TaskTrackers
Co-ordinate map & reduce phase
Provide Job progress info
R1
DFS
Task Tracker
Output file 1
DFS
Input file
split1
split2
split3
split4
split5
M2
Task Tracker
Region1
Region2
M3
Task Tracker
Region1
Region2
R2
Task Tracker
sort
read
reduce()
OutputFormat
DFS
Output file 2
Map Phase
Reduce Phase

Where? & When?

Course Organization

- 44 hours: ~30 lessons, ~14 laboratory
- Agreement on room and timetable
  - Currently: Thu 9-11 (room L1), Fri 9-11 (room N1)
  - Depending on availability
- Highly interactive lectures
- Laboratory
  - Java programming skills required
  - Bring your own laptop (don't forget plugs!)
- Slides and references available online
  - Updated in real time on the course wiki
- Final examination: project + oral session
  - To be agreed with teacher