



Conception Avancée de Bases de Données



Google

Google
Big Table



Traduction en cours



1996 1997



- Google started as a research project at Stanford University
- Created by Ph.D. candidates Larry Page and Sergey Brin when they were 24 years old and 23 years old respectively.
- BackRub became Google from « googol » 1 leading 100 zeros.



Time line

- 1998 : Susan Wojcicki garage
 - 232 Santa Margarita à Menlo Park.
- 2001 :
 - Eric Schmidt CEO
- 2004 : "Googleplex »"
 - 1600 Amphitheatre Parkway à Mountain View
- August 2004 :
 - Wall street NASDAQ 19 605 052 shares 85 \$ (1000 \$).



Alphabet Inc Class A

NASDAQ: GOOGL - 2 déc. à 08:53 UTC-5

764,33 USD 0,00 (0,00 %)

Avant l'ouverture: 761,00 **+0,44 %**

1 jour

5 jours

1 mois

3 mois

1 an

5 ans

max

1 000

800

600

400

200

0

2006

2008

2010

2012

2014

2016

Ouverture

+Haut

+Bas

Capitalis.

513,38 Md

Cours/bén.

27,47

Rend. div.

-

Alphabet Inc Class C

NASDAQ: GOOG

+ Follow



Overview

News

Compare

Financials

1,048.65 USD **+8.04 (0.77%) ↑**

Dec 14, 3:31 PM EST · Disclaimer

1 day

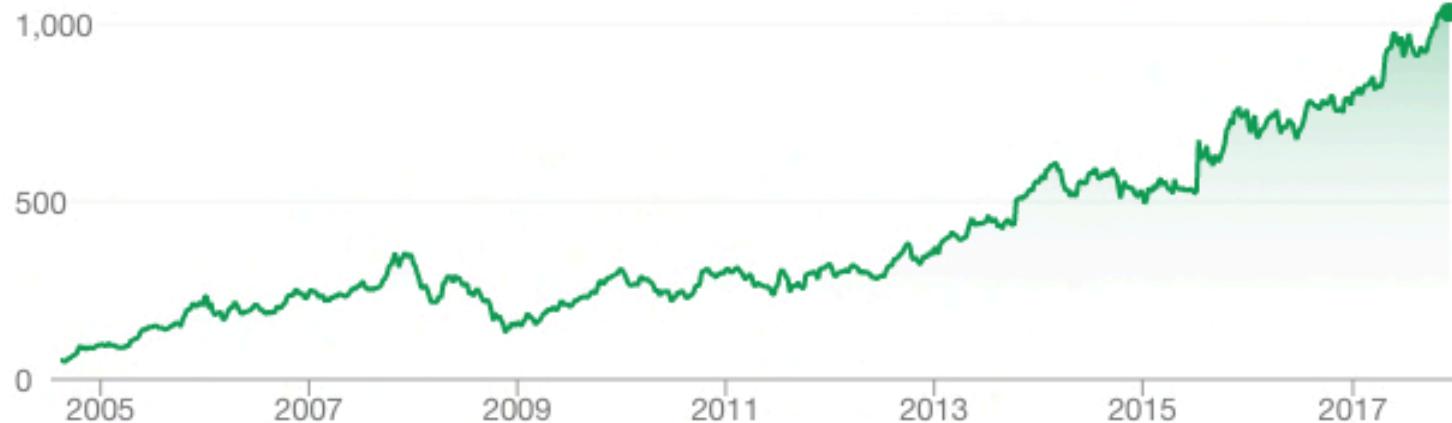
5 days

1 month

1 year

5 years

Max



Open

1,045.00

Div yield

-

High

1,058.50

Prev close

1,040.61

Low

1,043.11

52-wk high

1,062.38

Mkt cap

732.41B

52-wk low

770.41

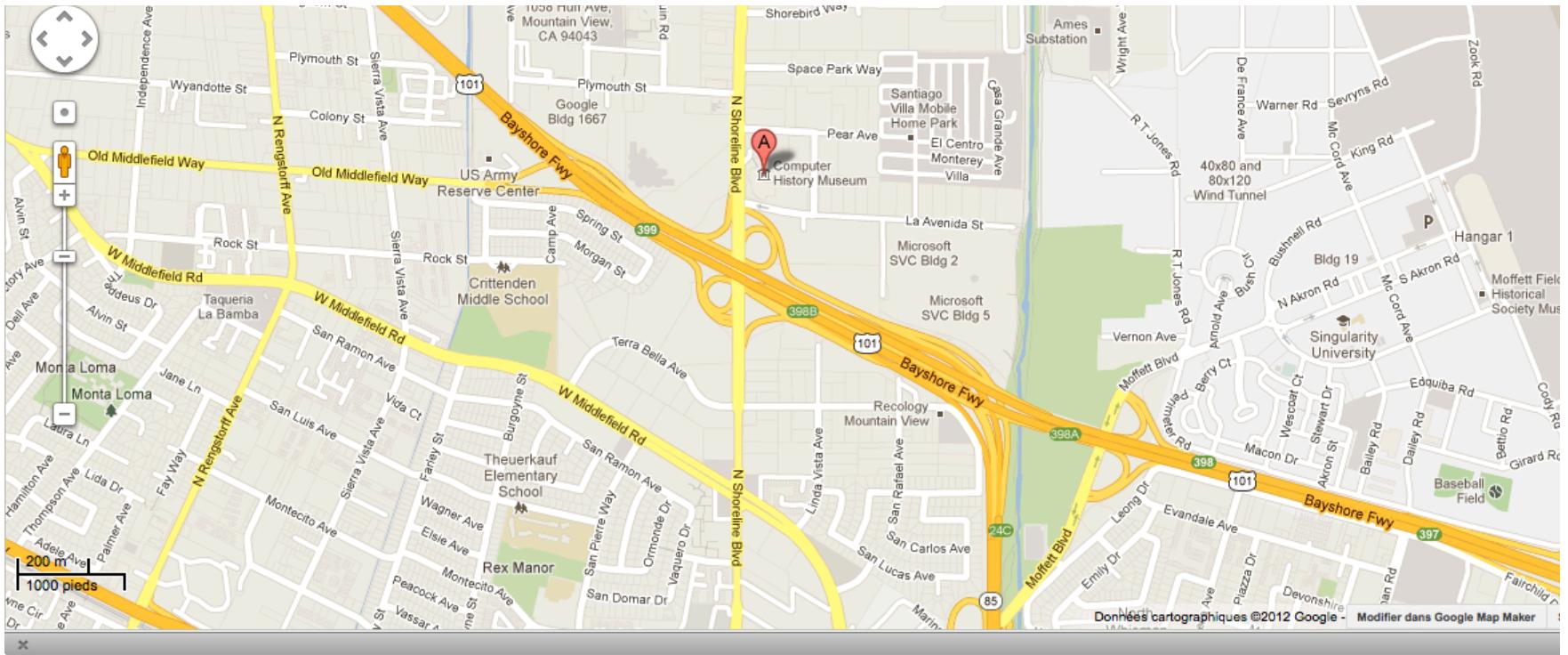
P/E ratio

34.94



Computer History Museum

- 1401 North Shoreline Boulevard Mountain View, CA 94043, United States





Sergey Brin & Larry Page



Susan Wojcicki : Google VP





Google Big Table



<http://actu.abondance.com/2008/04/google-pourrait-proposer-big-table-son.html>

Bigtable



- Bigtable is a distributed storage system for managing structured data that is designed to scale to a very large size: petabytes of data across thousands of commodity servers.
- Many projects at Google store data in Bigtable, including web indexing, Google Earth, and Google Finance.



deadlyphoto.com

Google Servers architecture

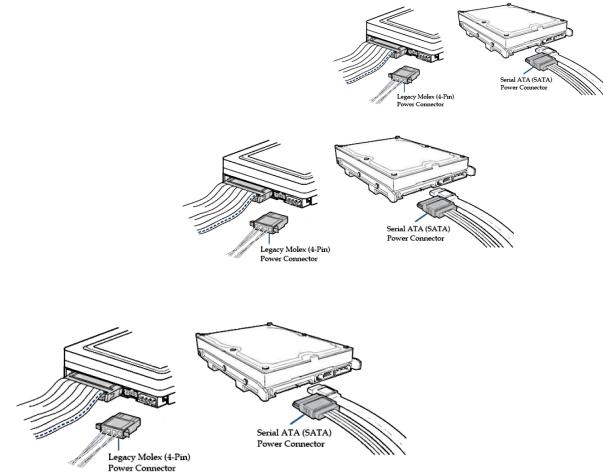
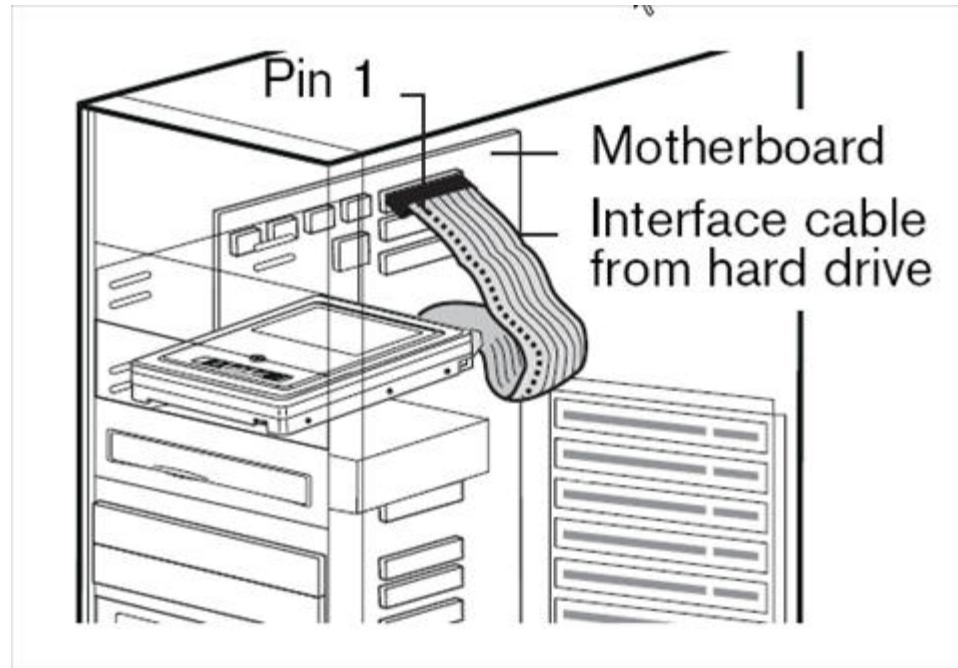


- PC mother boards and disks

Google rack



Low cost PC



FRYS.COM

SURCOUF

FRYS Mountain View, CA 94043



Fry's Electronics, the first electronic superstore: 1177 Kern Ave, Sunnyvale





fry's
electronics

http://www.yelp.com/biz_photos/CsFTB9cErCkOkOqWGIqzHA?select=4RHMVug11I7cMSAaHHRLuA#4RHMVug11I7cMSAaHHRLuA

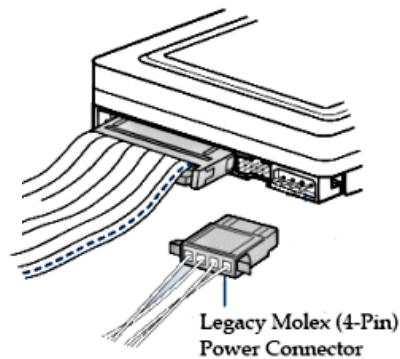
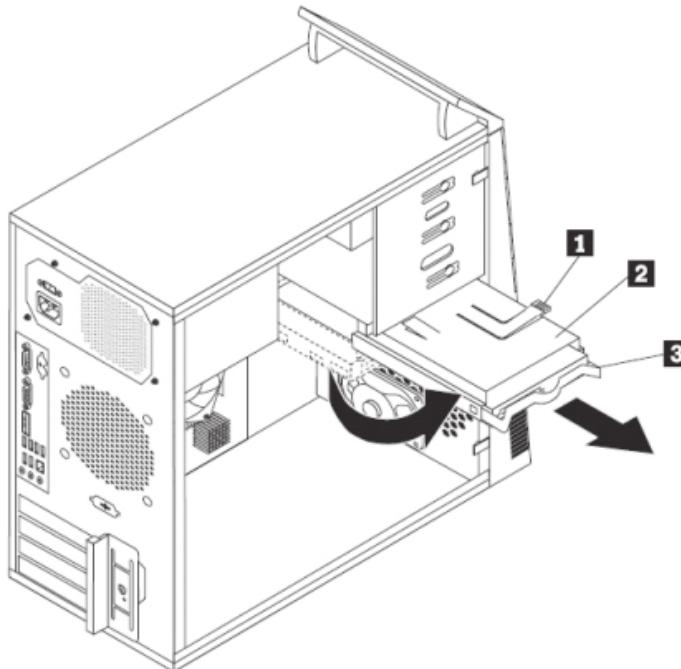


fELECTRONICS
fry's

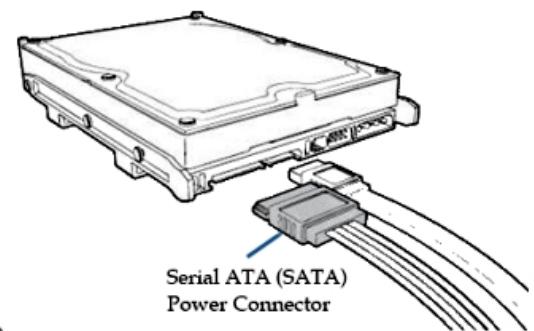


Google Story

Google™



Legacy Molex (4-Pin)
Power Connector

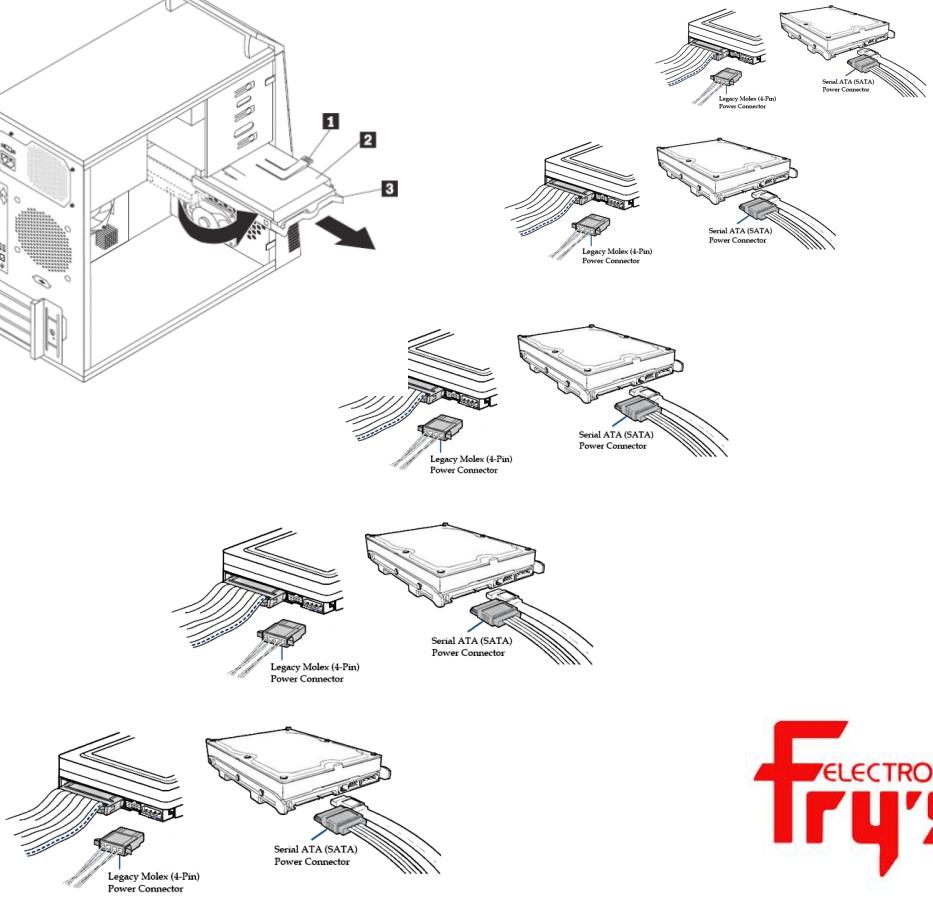
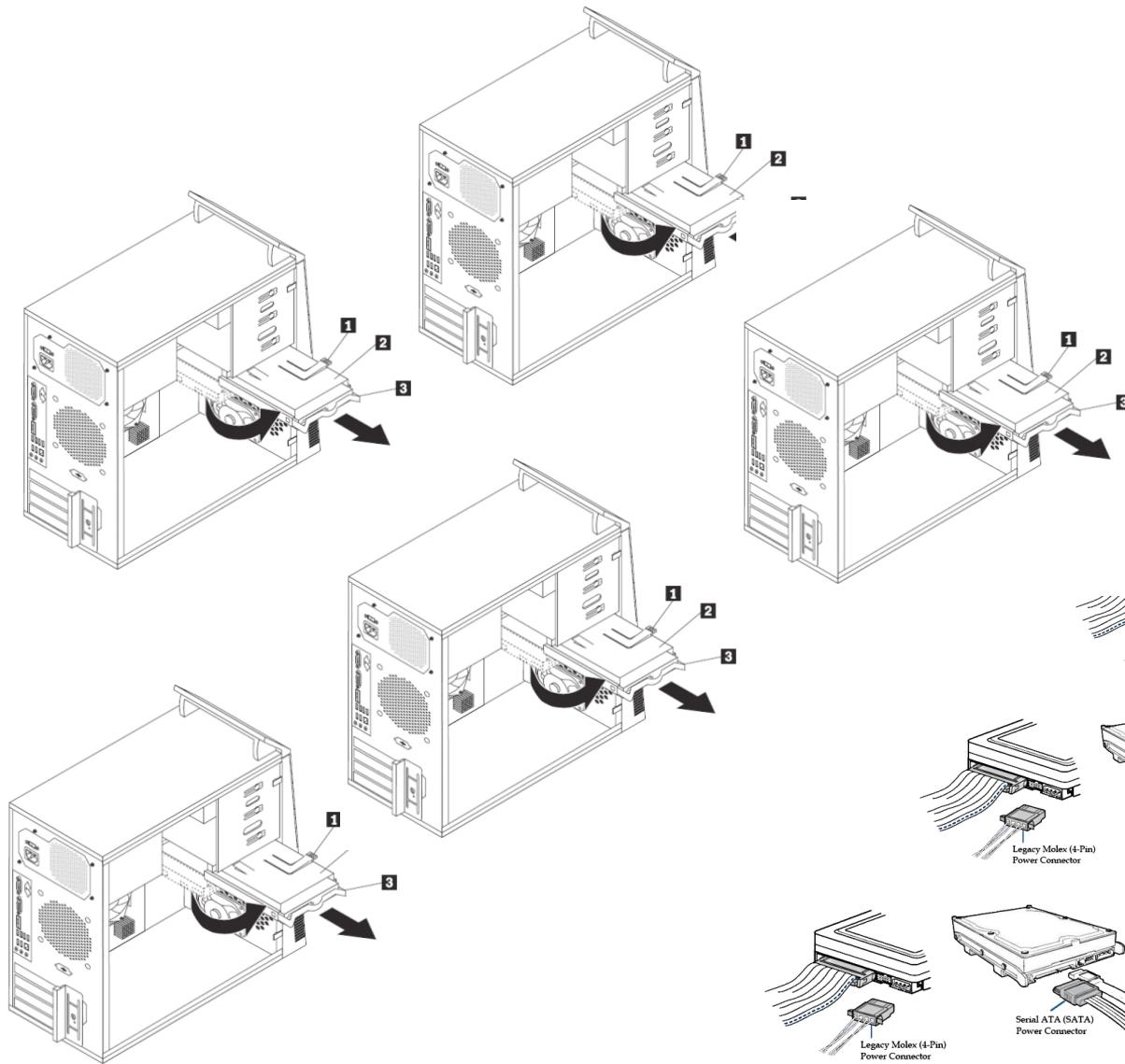


Serial ATA (SATA)
Power Connector

fry's
ELECTRONICS

Google Story

Google™



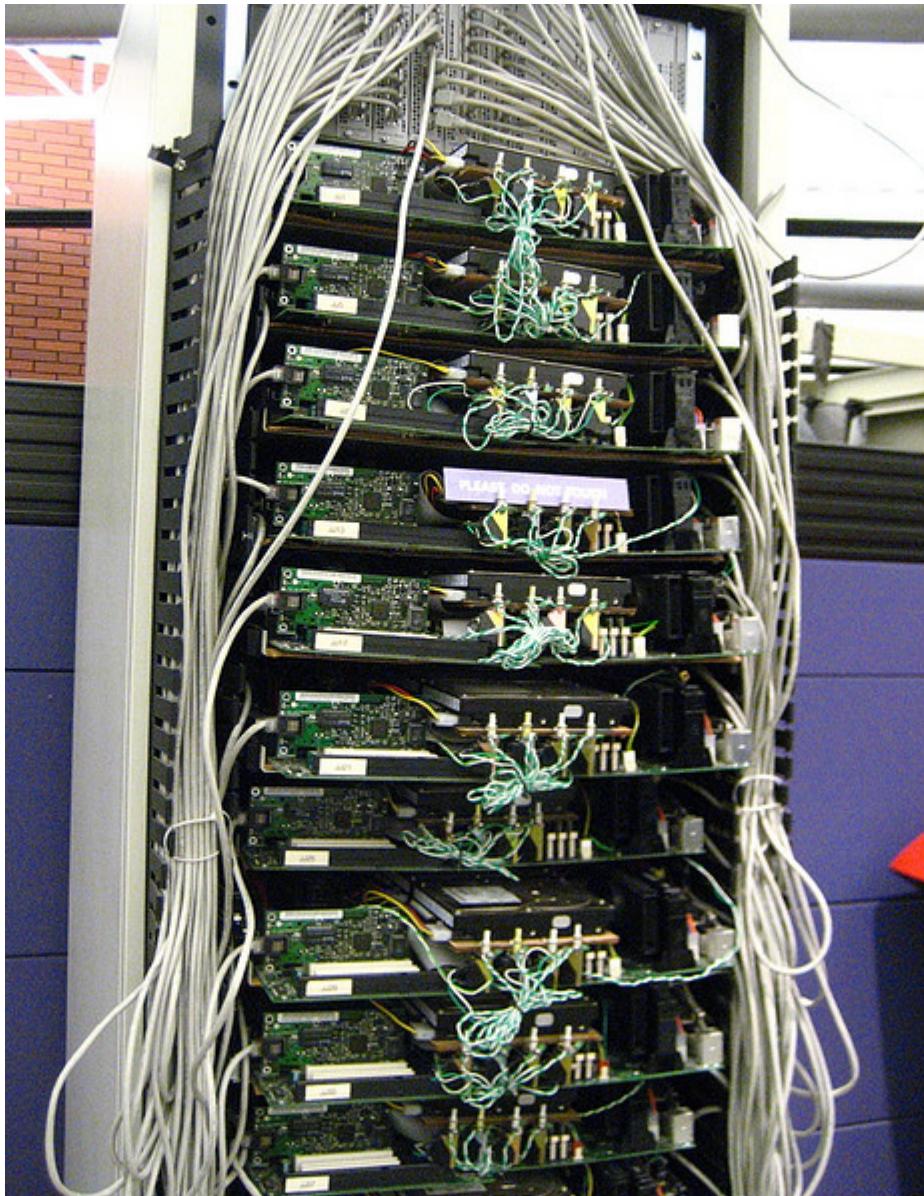
fry's ELECTRONICS

Google Story

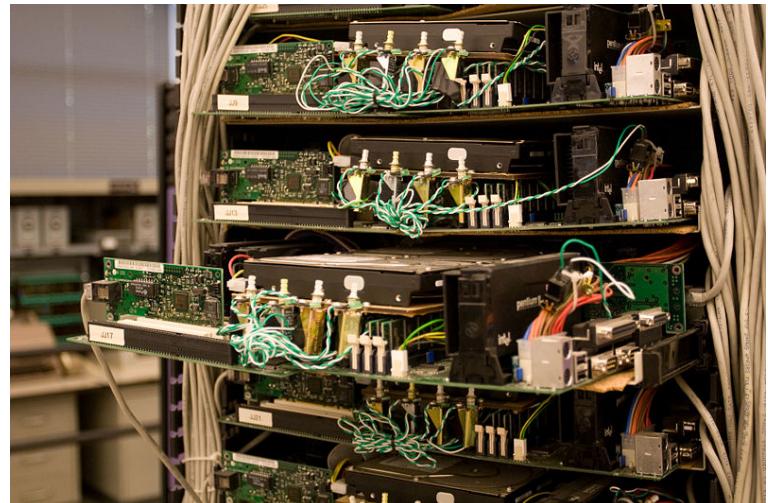


fry's ELECTRONICS

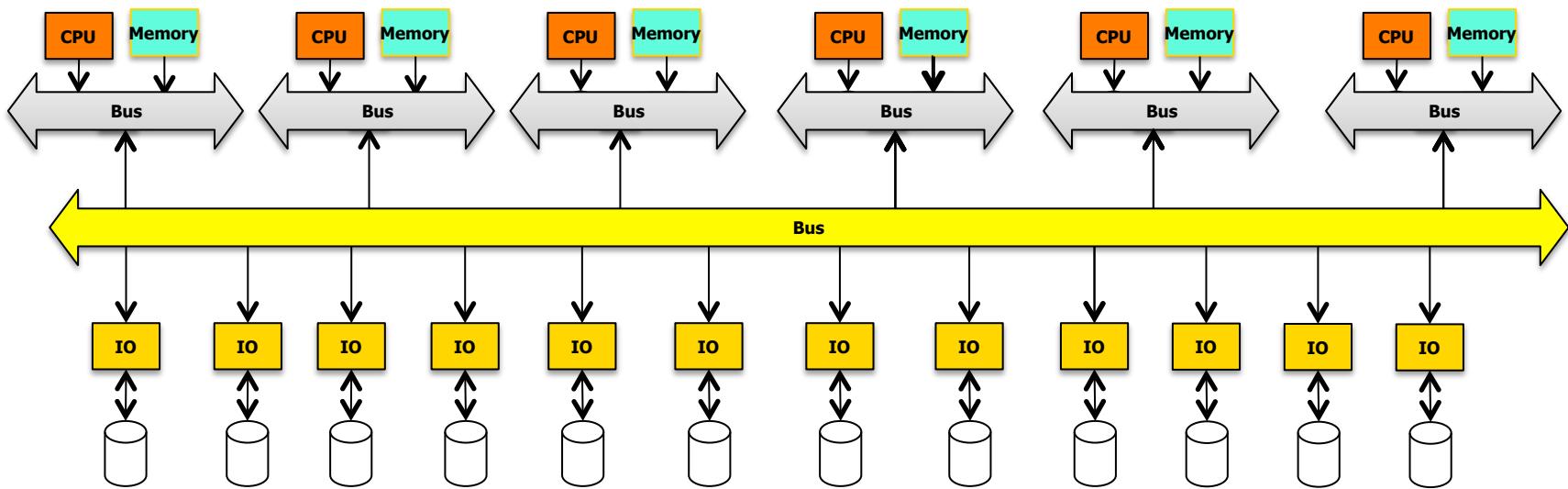
Google Rack



Google™



Horizontal Scalability : Shared Disk



low-cost low performance servers



- Commodity computer
- Many low-cost low performance servers.
- Servers are aggregated in clusters with more than a thousand servers each
- A cluster consists of tens of racks.
- The nodes inside a rack (about 80 per rack) are connected by a switch with something like a 1gb/s data rate between any two nodes.
- High degree of parallelism

Bigtable



- A Bigtable is a sparse, distributed, persistent multidimensional sorted map.
- The map is indexed by a row key, column key, and a timestamp; each value in the map is an uninterpreted array of bytes.

KEY

VALUE

(row:string, column:string, time:int64) → string



Google Search

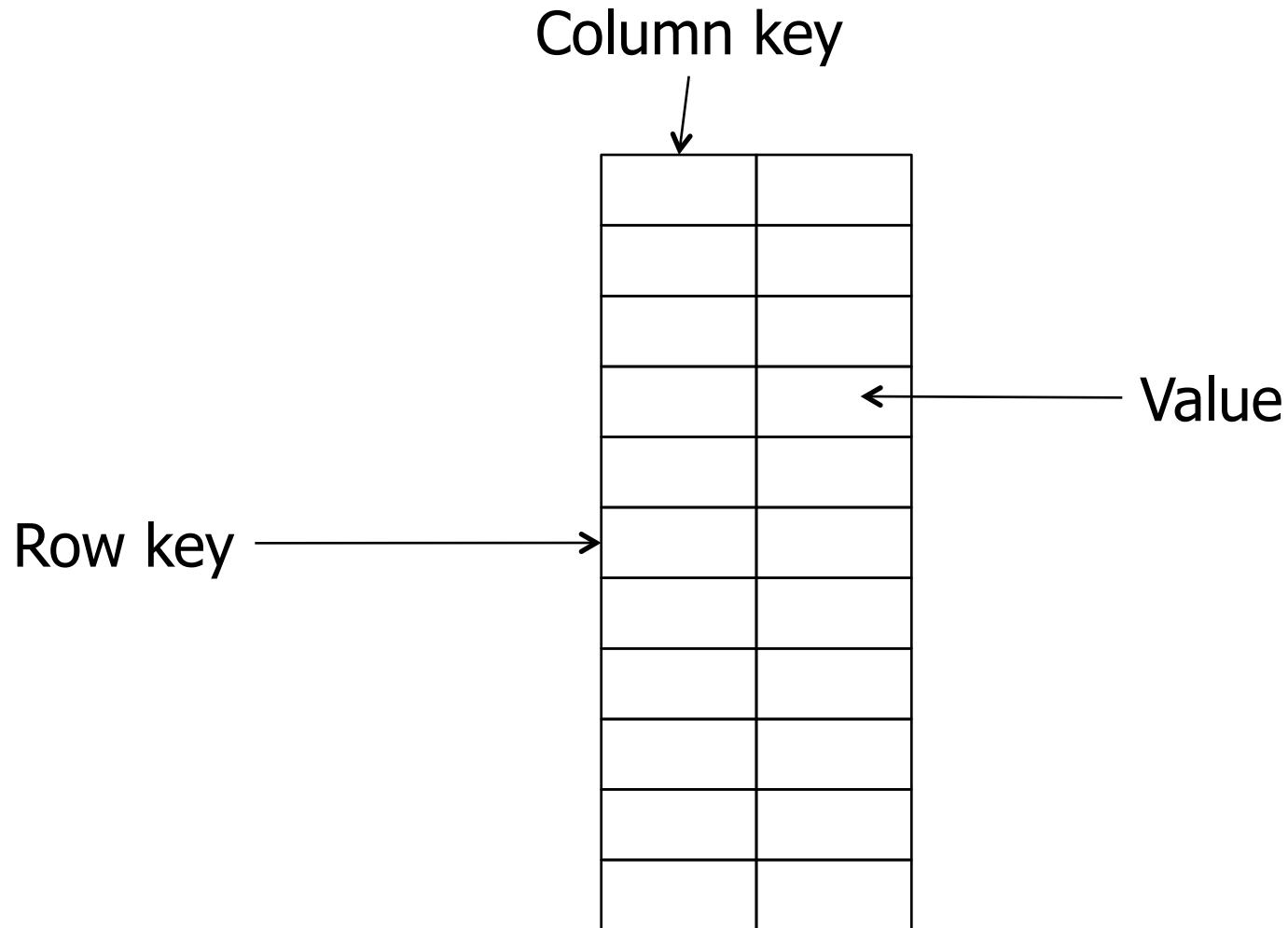
- The key is every possible search term.
- The value is the table of web sites to return for this search term.
- Index is replaced by an algorithm called MapReduce
- MapReduce is a parallel algorithm, the search can be done by multiple servers in parallel.



Hash Table

- Sparse 3D hash table :
 - row names,
 - column names,
 - versions (timestamps).
- Everything's a string (sequences of characters) :
 - row names,
 - column names,
 - data items

Keys value





Sort of 1 PB (10^{15}) in 6 hours

- One petabyte is a thousand terabytes,
- 10 trillion (10^{12}) 100-byte records
- on 4,000 computers
- It took six hours and two minutes



Google™

Refroidissement Google



