To Cloud or Not To.

An exploration of the economics of clouds.

radu
sion
sion@cs.stonybrook.edu

Systor 2012
Haifa, Israel June 4 – 6
The 5th Annual International Systems and Storage Conference
In cooperation with ACM, IEEE, USENIX, and TCE 2012

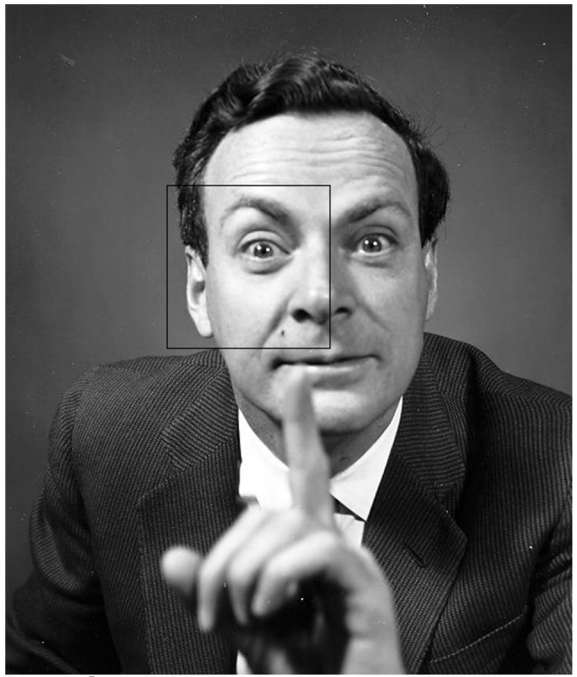
C3 Cloud Computing Center
Center of Excellence
Wireless and Information Technology

Stony Brook University

National Science Foundation
WHERE DISCOVERIES BEGIN

ver 2.5 widescreen

Feynman Moment



"I have experience only in teaching graduate students […] and as a result […] I know that I don't know how to teach."

please interrupt and engage!

The cloud

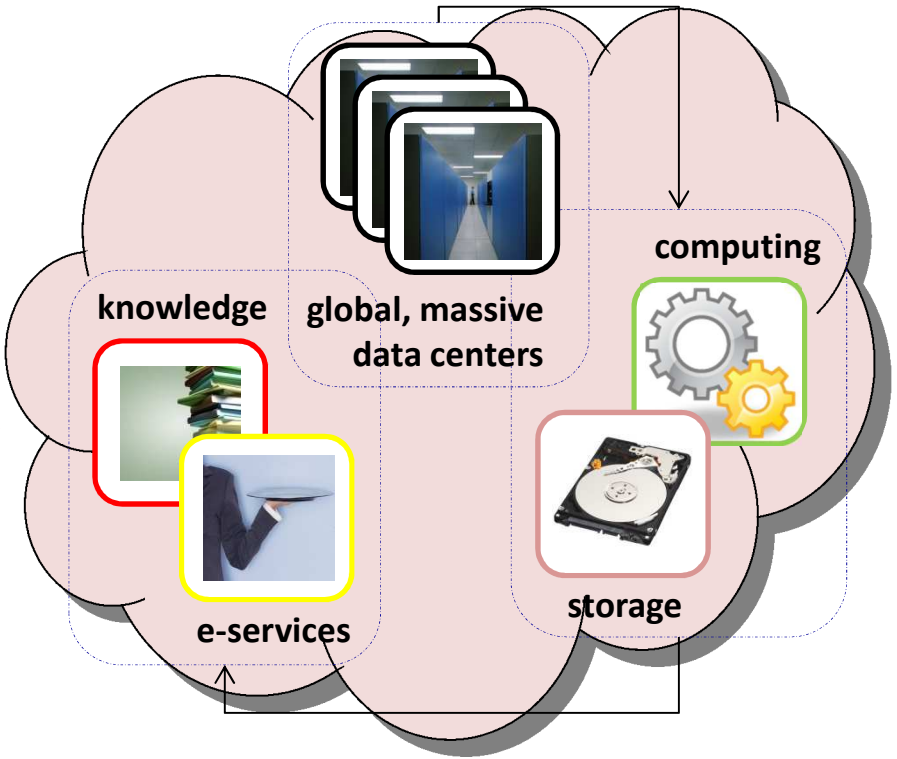
government
corporations
healthcare

consumers

mobile

"Utility" Computing
Transparency
Availability
Extremely cheap
Efficient
On-Demand

Online Services
Software
Platforms
Infrastructure
Knowledge



Clouds v. Grids v. …

+ Control Structure
+ Illusion of “Unlimited”
+ No up-front commitment (“pay as you go”)
+ On-demand
+ (Very) Short-term allocation
+ Close to 100% Transparency
+ Increased Platform Independence
+ It is actually here and happening!

Buzzword Bandwagon

On your marks, get set, GO

Race to results with the powerful Sun Grid Compute environment and our first class catalog of [illegible] applications!

» Learn More

ON DEMAND BUSINESS™

ORACLE 11g DATABASE

10g ORACLE DATABASE

Flavors

Traditional Outsourcing [(Semi)Private Clouds]
ACME Corp. manages servers for XYZ Financials

Clouds
Amazon EC2, Google Apps, MS Azure

Managed servers

Un-managed hardware

Should I buy it?

costs vs. benefits

costs
technology costs
cost of security
etc.

clients

benefits
availability
opportunity
consolidation
etc.

the “cloud”

Core costs of computing

+ Storage ($/MByte/year)
+ Computing ($/CPU Cycles)
+ Networking ($/bit)

Reality is way more mundane

Hardware

servers, disks, **network**, racks, power, cooling

Energy

power, cooling, infrastructure

People/Service

maintenance, development

Space

Size does matter

Home Users (1-10 CPUs)
"no" rent/cooling/administration

Small Enterprises (up to 1k)
no custom hardware, low utilization

Mid-size Enterprises (up to 20k)
better network service, better utilization

Large/Clouds (50k+)

Clouds

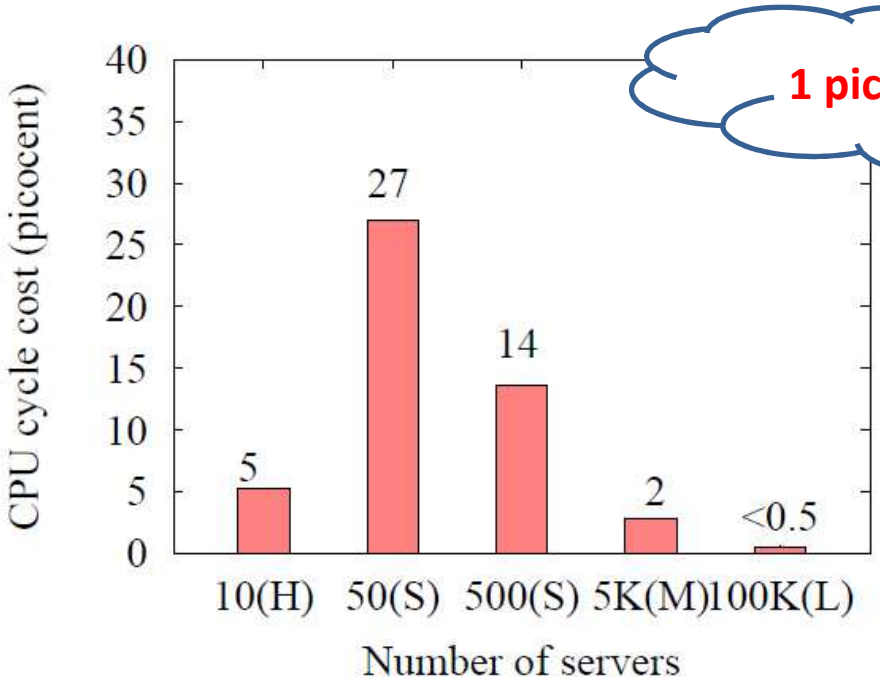
+ Custom hardware
+ Efficient cooling
+ Cross-timezone load shifting
+ High CPU utilization
+ Preferential network deals
+ High Power Usage Efficiency (PUE)

Understand cost of CPU cycle

Parameters	H	S	M	L
CPU utilization	5-8%	10-12%	15-20%	40-56%
server:admin ratio	N.A.	100-140	140-200	800-1000
Space (sqft/month)	N.A.	$0.5	$0.5	$0.25
PUE	N.A.	2-2.5	1.6-2	1.2-1.5

$$\frac{\lambda_s \cdot N_s/\tau_s + (w_p \cdot \mu + w_i \cdot (1-\mu)) \cdot PUE \cdot \lambda_e + \frac{N_s}{\alpha} \cdot \lambda_p + \lambda_w \cdot N_w/\tau_w + \lambda_f \cdot \frac{(w_p \cdot \mu + w_i \cdot (1-\mu)) \cdot PUE}{\beta}}{\mu \cdot \nu \cdot N_s}$$

CPU cycle cost (circa 2009)

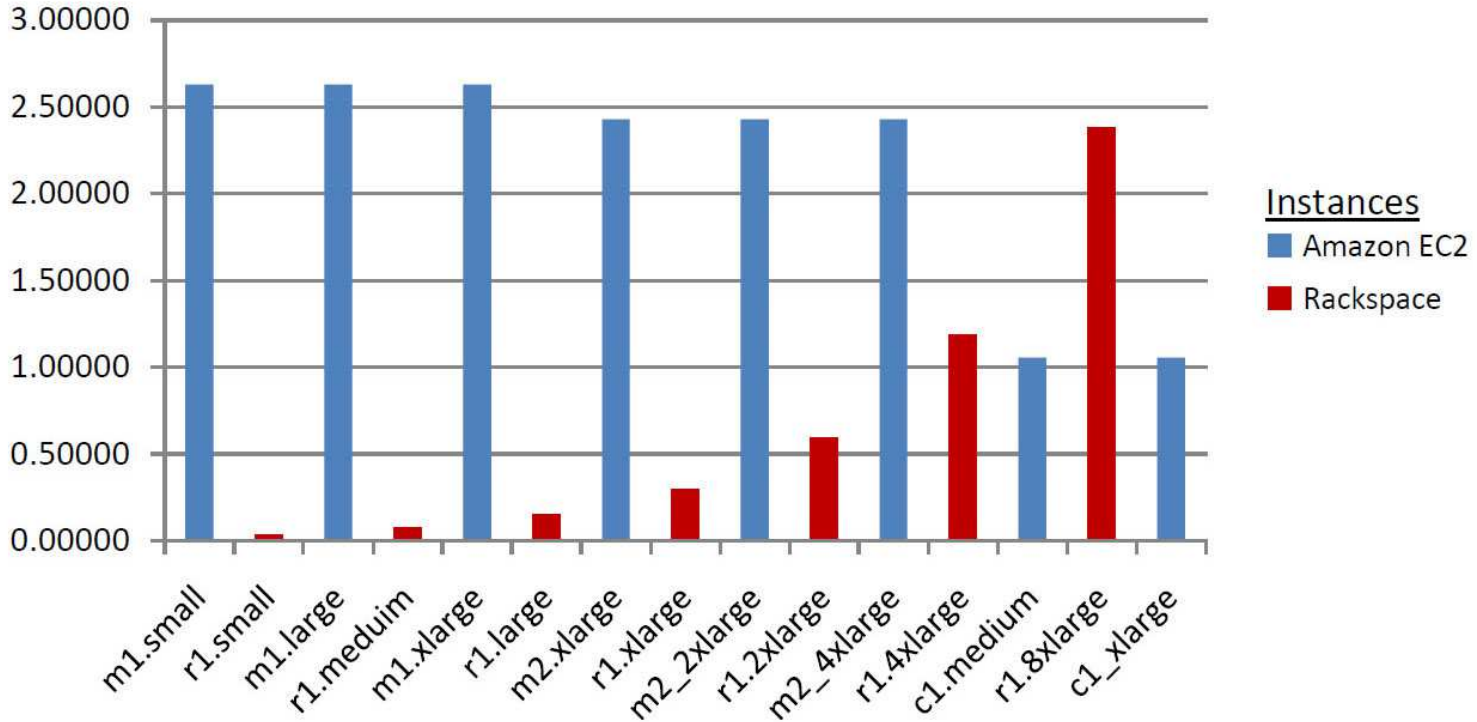


1 picocent = 10^{-14} USD

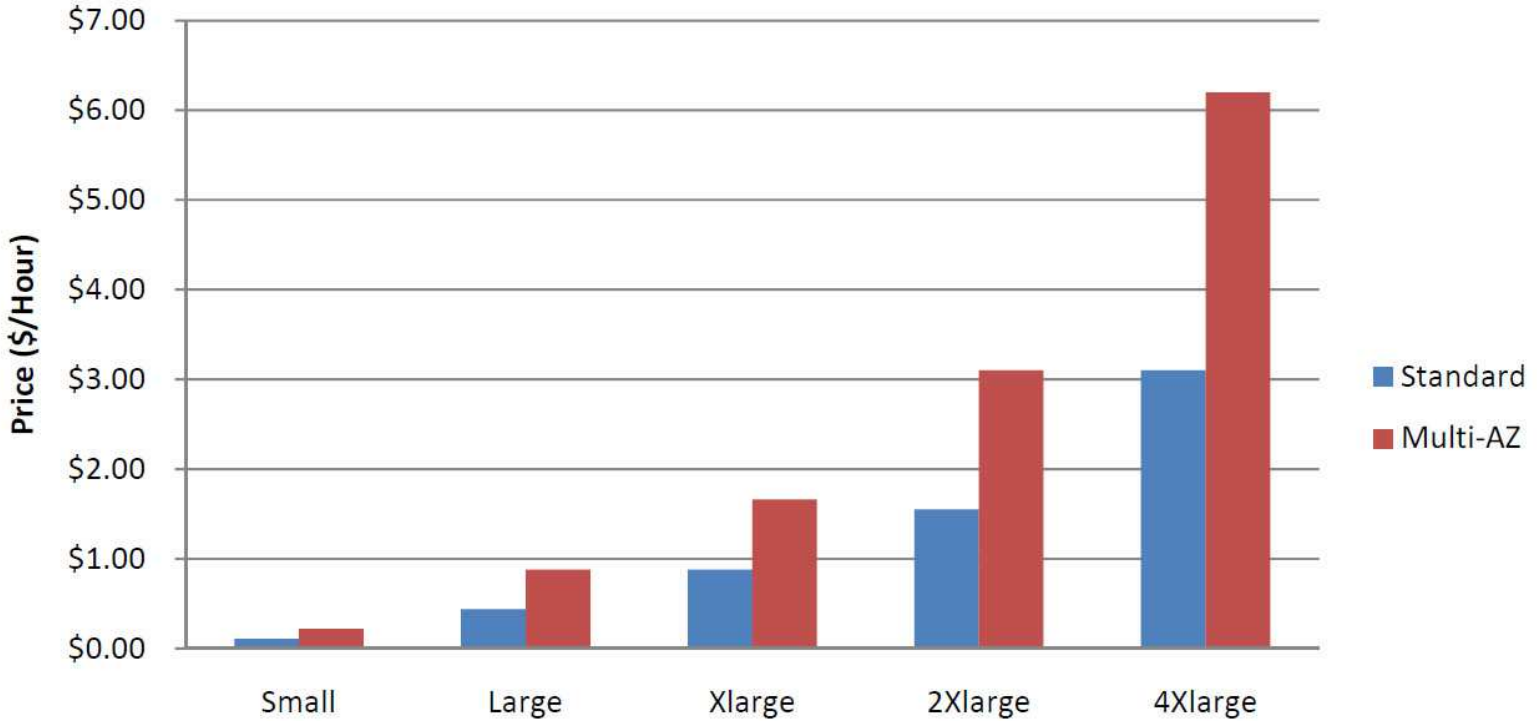
Consumer clouds today (cca. 2009)

Provider	Picocents
Google	0.5 – 2.31
Microsoft	0.7 – 1.96
Amazon	0.93 – 2.36
Rackspace	0.02 – 2.4

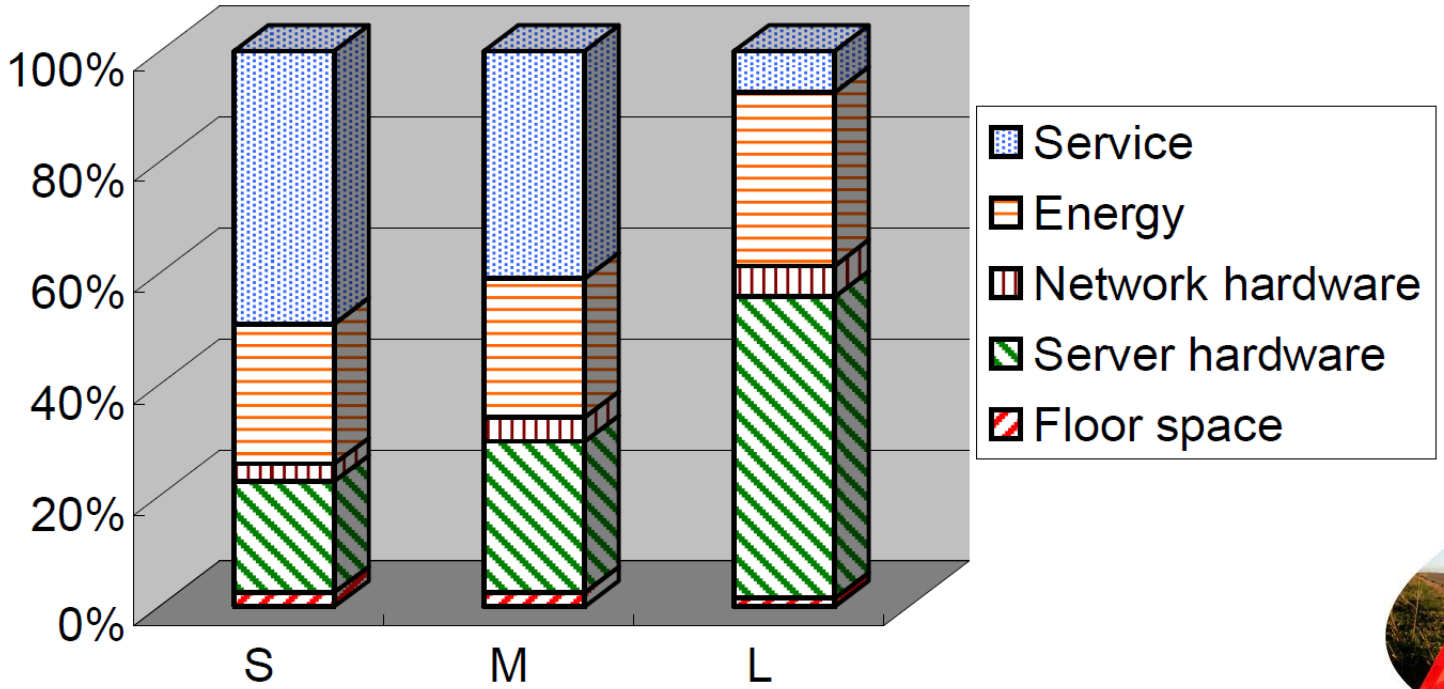
Rackspace vs. Amazon (2011)



Amazon RDS (Q4, 2010)



Breakdown


100%
80%
60%
40%
20%
0%
S
M
L
Service
Energy
Network hardware
Server hardware
Floor space

So: is it worth it?

Mostly yes …

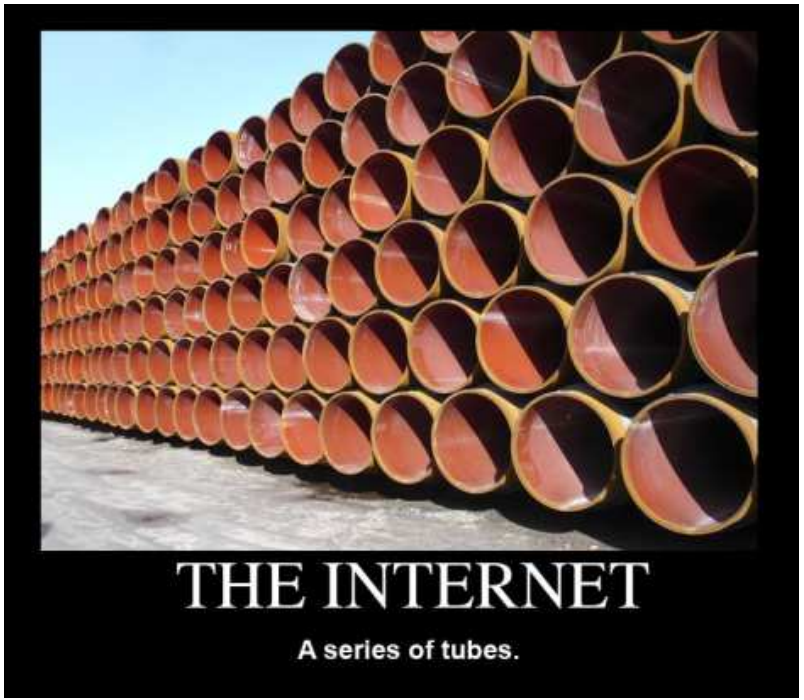
Why ?

1 client cycle
6-27 US picocents

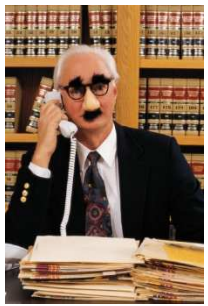
clients

1 cloud cycle
0.58 picocents

What about the tubes?



We are far!

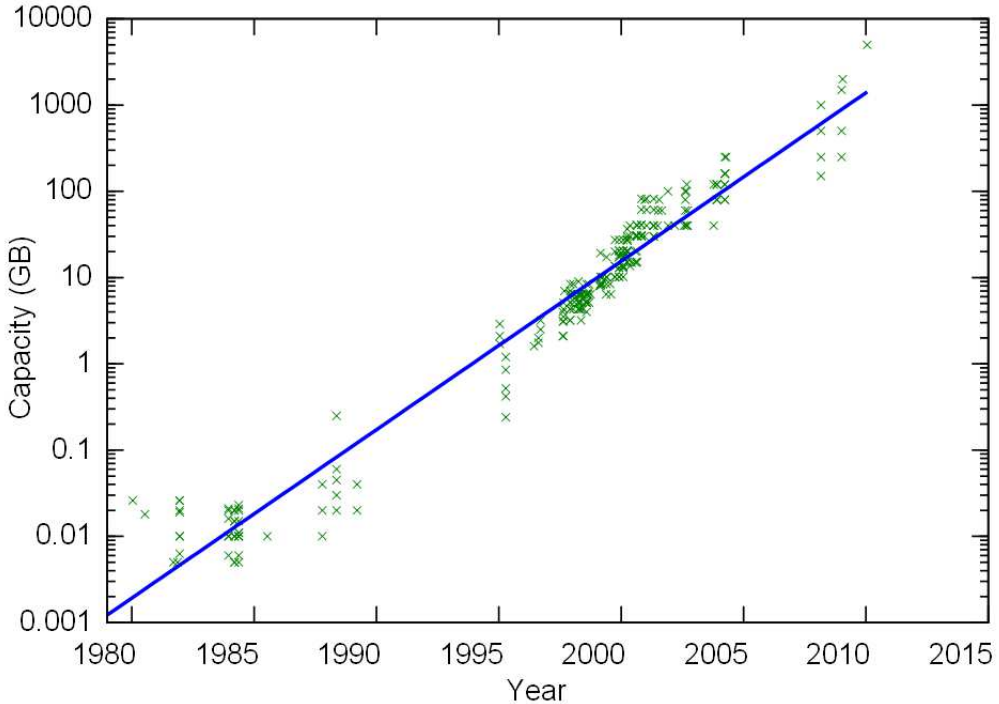
provider	monthly	bandwidth (d/u)	picocent/bit
	$29.95	15 Mbps /5 Mbps	77/231
	$44.9	30 Mbps /5 Mbps	58/346
	>$1000	5-1000 Mbps	5000 (est.)
	$19.99	1 Mbps/384 Kbps	771/2008
	$29.99	3 Mbps/768 Kbps	386/1506
	$42.99	7.1 Mbps/768 Kbps	233/2160
Mid-size	$95 (est.)	1 Mbps (dedicated)	3665 (est.)
Large/cloud	$13 (est.)	1 Mbps (dedicated)	500 (est.)

Additional ammunition?

Disk	cap. (GB)	price (USD)	Adj. MTBF (mil.hrs)	amort. acq. (pcent/bit/yr)	power seek (W)	power2 idle (W)	power3 (W)	power cost (pcent/bit/yr)	total cost (pcent/bit/yr)	acq. %	avg. seek time (ms)	avg. seek4 cost (pcents)	power5 read (W)	read cost (pcent/bit)
Maxtor Diamond Max	500	53	0.35	32.89	13.6	8.10	10.85	237.62	270.50	12.16	9.00	377542	11.16	0.03
Hitachi Deskstar 7k500	500	67	0.29	49.89	15	9.60	12.30	269.37	319.26	15.63	8.50	407953		
Hitachi Ultrastar A7K1000	1024	153	0.35	46.36	14	9.00	11.50	122.97	169.33	27.38	8.20	417631		
WD Caviar GP Low Power	1024	103	0.29	37.43	7.5	4.00	5.75	61.49	98.93	37.85	8.90	271994	7.40	0.02
Seagate Barracuda 7200.10	750	63	0.35	26.06	12.6	9.30	10.95	159.87	185.93	14.02	9.25	369615	13.00	0.06
WD Caviar SE16	500	62	N/A		8.77	8.40	8.59	188.01			9.90		8.77	0.04
Samsung SSD	32	269	0.29	3129.65	1	1.00	1.00	342.19	3471.83	90.14	1.70	47912	0.5	0.0017
Intel SSD X18-M	80	389	0.35	1508.59	0.15	0.06	0.11	14.37	1522.96	99.06			0.15	0.0002
Intel SSD X25-M	160	765	0.35	1483.38	0.15	0.06	0.11	7.19	1490.57	99.52			0.15	0.0002

Up to 350 for 3 year lifetime!

Storage capacity over time



So: should I buy a piece of sky?

… not always.

CPU Cycle
6-27 picocents

1 bit storage/year
6 picocents

clients

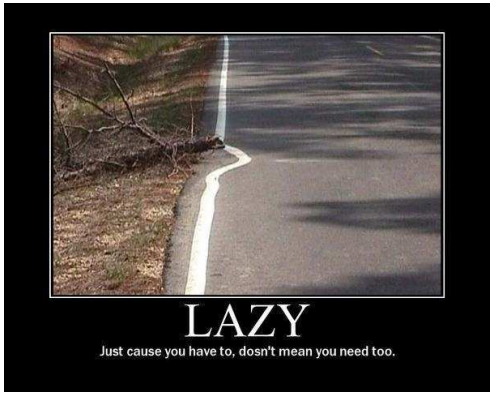
1 bit network transfer
800-6000 picocents

CPU Cycle
0.58 picocents

1 bit storage/year
5.3-6 picocents

So when is it clearly worth it?

Q: is the application doing enough computation work (cheaper) to offset the distance cost to the cloud?



First Principle of Cloud Viability

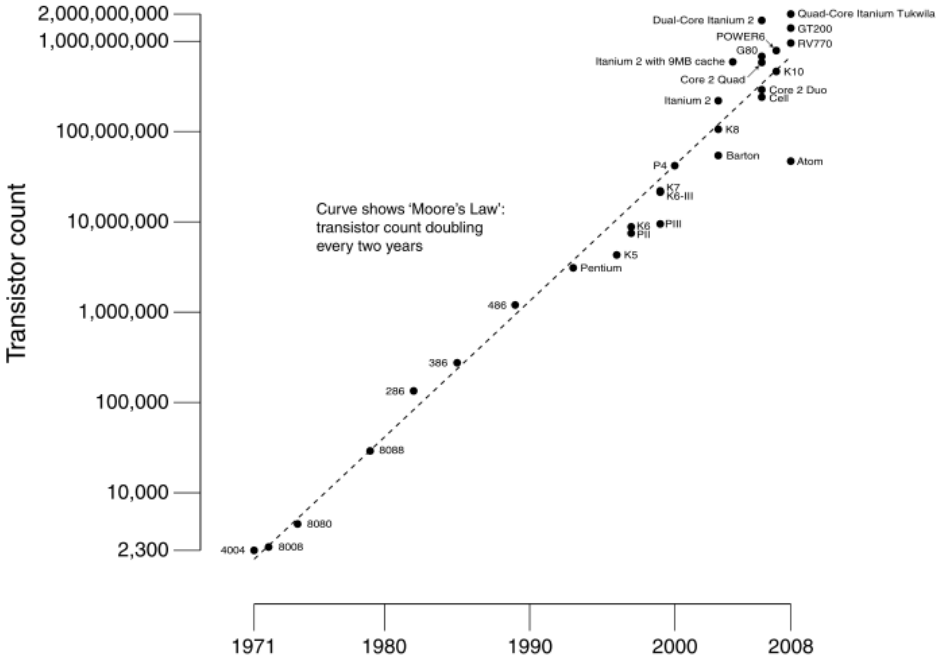
It is not worth outsourcing any task of less than 4000 CPU cycles per transferred 32-bit input.

Why should this hold tomorrow?

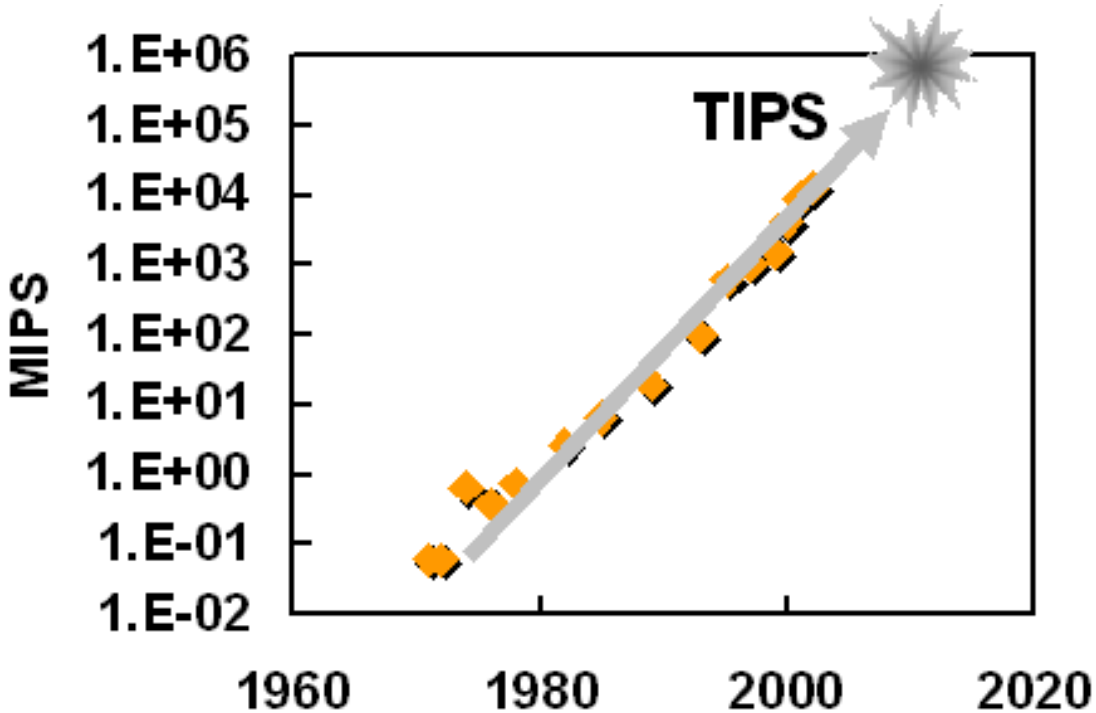
Ratio of exponentials is exponential ☺

Moore vs. Nielsen

Density (or cycles/$)

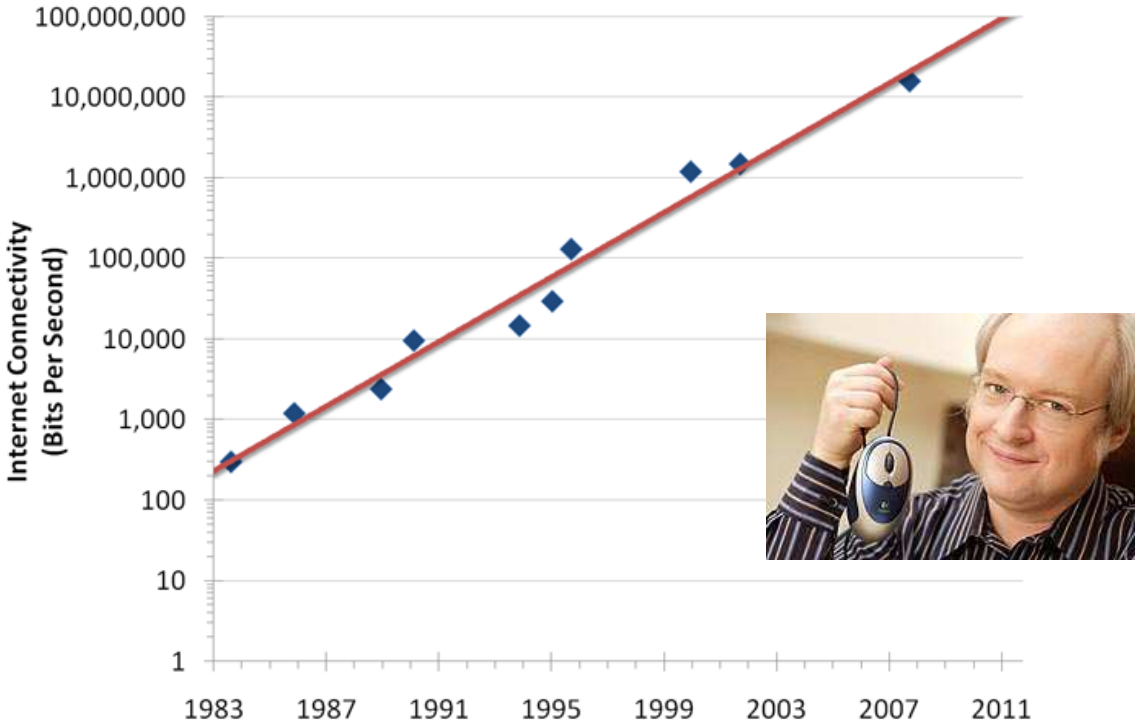


Speed



Source: "Gigascale Integration-Challenges and Opportunities", Shekhar Borkar, Director, Microprocessor Technology, Intel

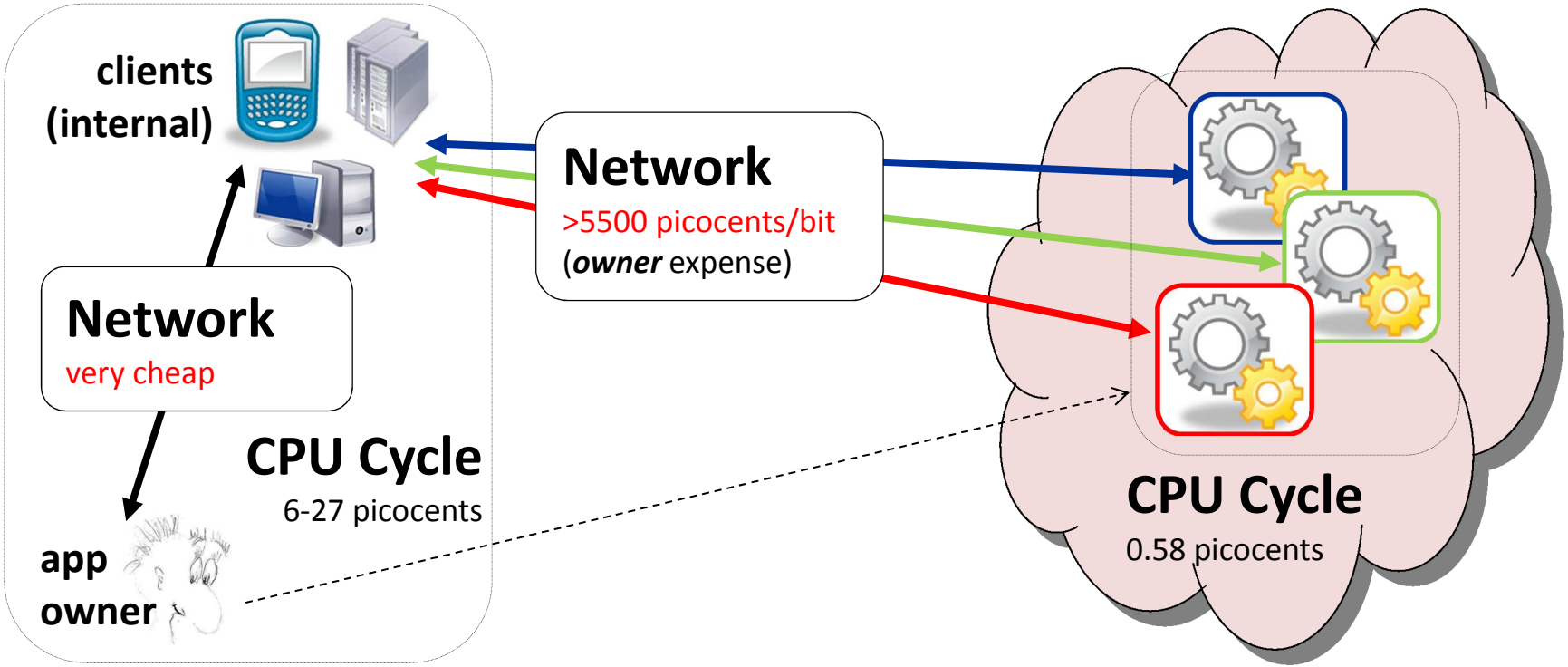
Networks



"high end connection speed grows 50% per year"


LIAR LIAR

App Owner = Sole Client

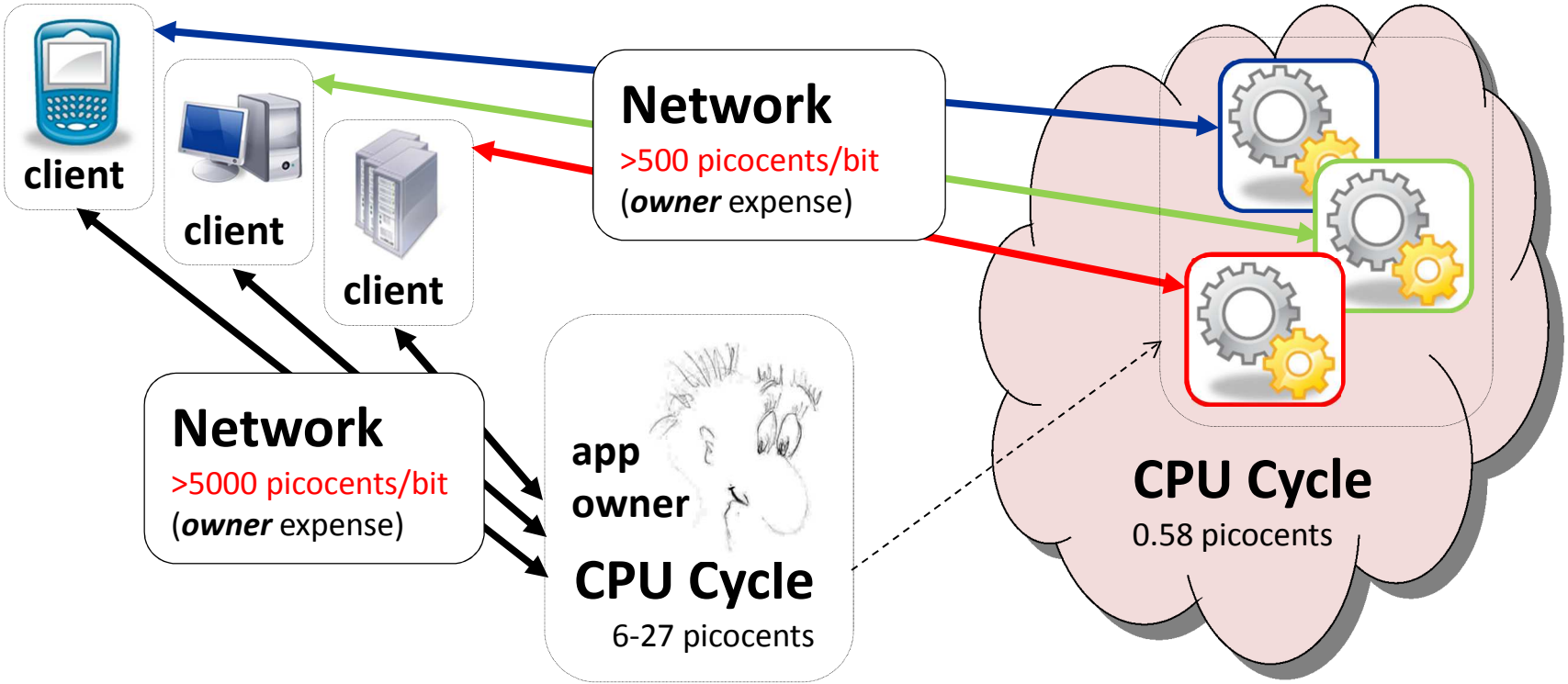


But is this the nominal case?

actual question to ask

what is the overall application profile?

App Owner != Client(s)



Insight: we had only partial view!

Second Principle of Cloud Viability

"It is almost always worth outsourcing"

Thus

cloud deployment saves

+ >4500 picocents per client-to-app traffic bit

+ tens of picocents per CPU cycle.


Hmmm …

But … it seems sooo expensive!!!

Computing in cloud

8c/hour = $1.92/day = $700/yr ≡ $2100/3yr

Instance utilization is still low! (<12%)

Computing "at home"

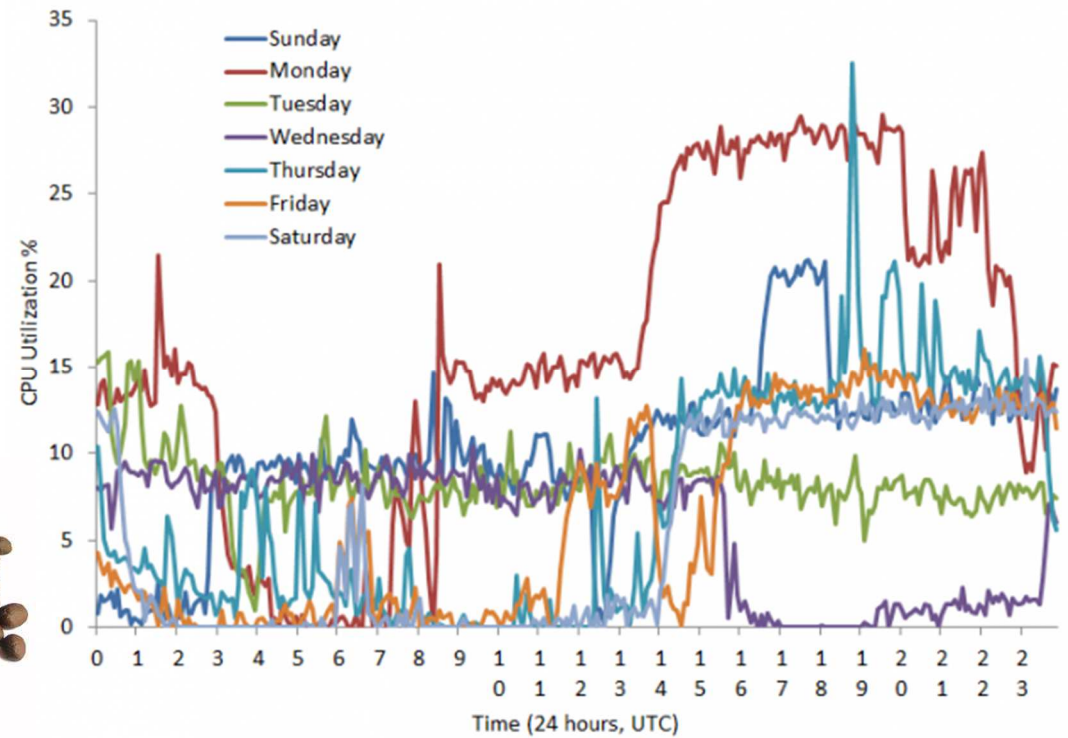
energy = 10c/kWh @ 150W ≡ $394/3yr

acquisition = $500

Networking in cloud

5-12c/GB = 582-1397 picocents/bit

Cloud CPU utilization (temp. based)




But wait theres more!
ICANHASCHEEZBURGER.COM

What about other goodies?

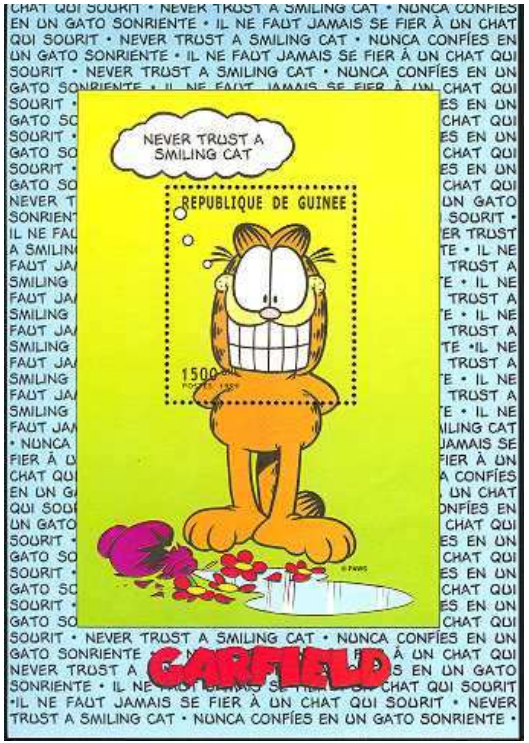
Primitive	Picocents
CPU Cycle	0.58 - 26
Disk Access /bit	0.02 - 0.06
Disk Access+DMA /bit	0.023- 0.11
Disk Seek	**270,000 - 417,000**
Disk Store /bit/hr.	0.011 - 0.036
Disk am. acq. /bit/hr.	0.003 - 0.0057
SDRAM am. acq. /bit/hr.	5.96 - 32.96
SDRAM Access /bit	0.003 - 0.05

Are clouds more or less secure?

+ Yes
+ But what is security?!

Trusting stuff …

"behave in the *expected manner* for the *intended purpose*"



Usually the monkey gets you

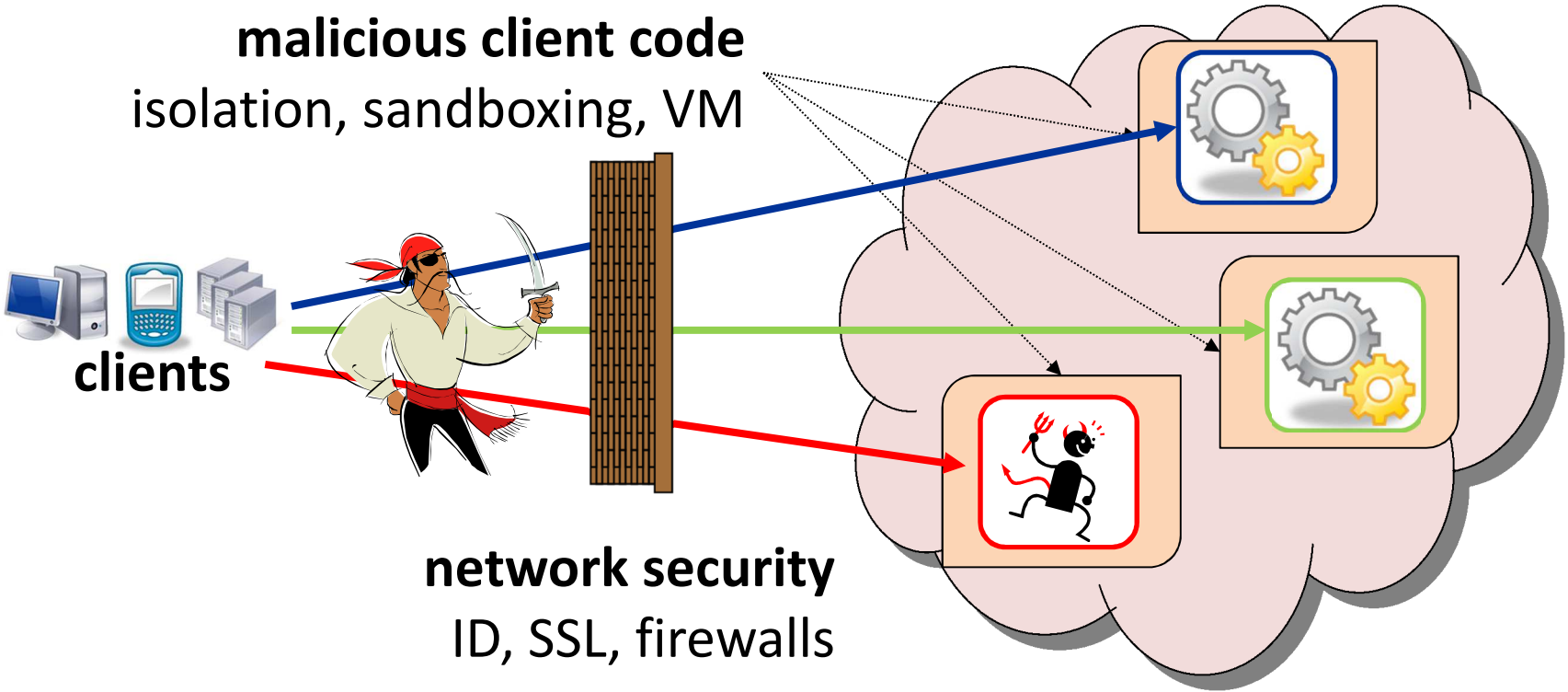
_____ Voting Machine

online public picture of ***actual*** key

Usual suspects

malicious client code

isolation, sandboxing, VM

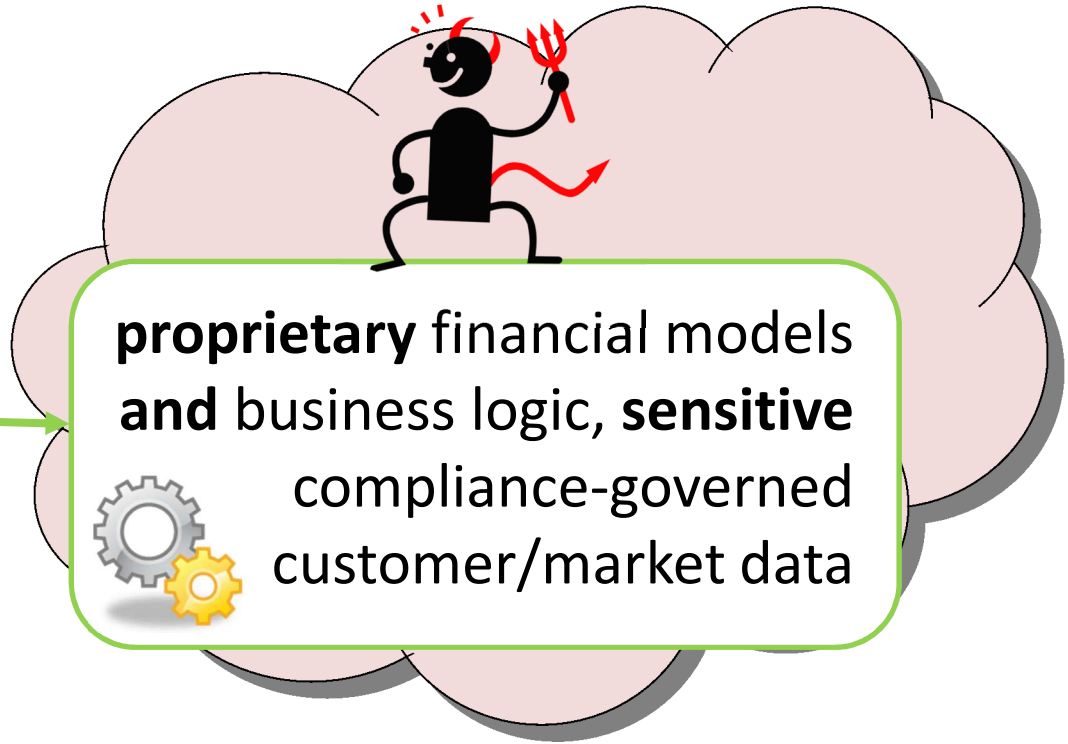


network security

ID, SSL, firewalls

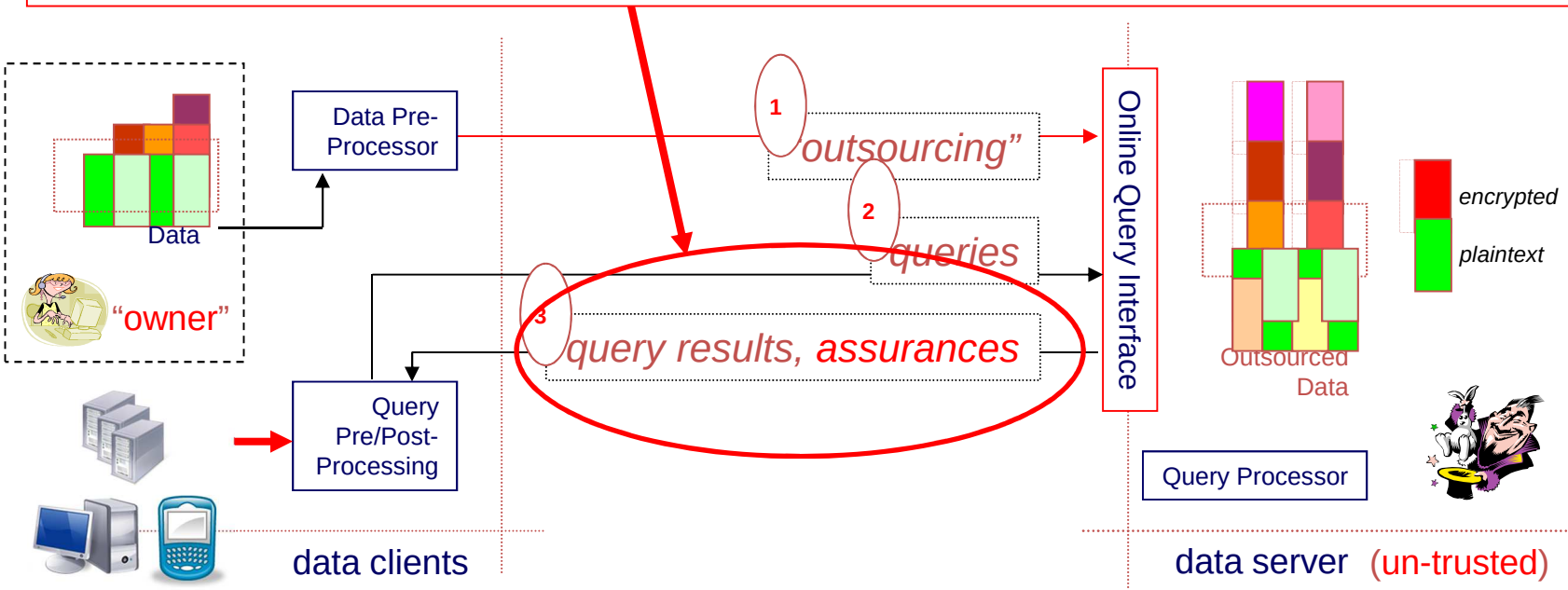
Secure Outsourcing

Finance Inc.

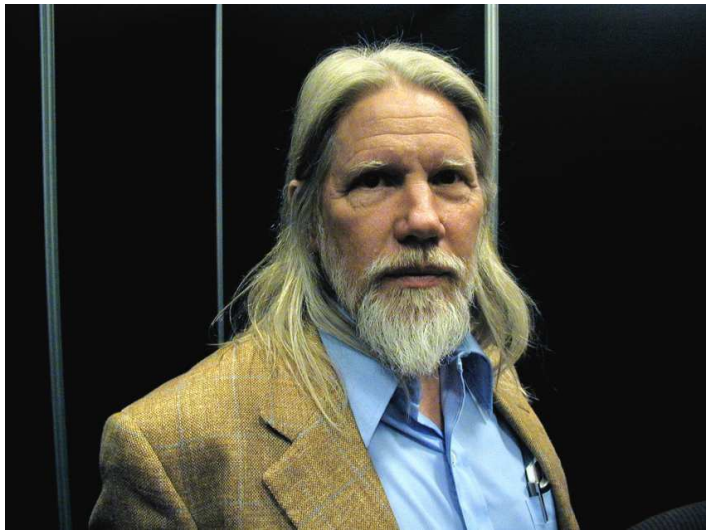


Ideas

assurances ⊆ {query correctness, data confidentiality, access privacy}



Diffie Moment



"Whit" Diffie

"while it is possible in principle for computation to be done on encrypted data, [...] current techniques would more than undo the economy gained by the outsourcing and show little sign of becoming practical".

So … do they work?

Unfortunately, not!

Why not ?

peanut counting is (too) **cheap.**

clients

we don't know how to *practically* "secure" anything more complex that peanut counting.

Peanut counting

Data Storage

700+ picocents/bit un-amortized extra costs (even in unsecured case!)

PIR (Private Information Retrieval)

2-3 orders of magnitude more expensive

Keyword Searches

4-5 orders of magnitude more expensive

Range Queries

2-3 orders of magnitude costlier even in unsecured case
some crypto (signature aggregation) would add another 2+ orders

Simple Aggregators

using homomorphisms (e.g., VLDB 2007) – would take 12 days/query

e.g., storage + data confidentiality

<0.1%

Decryption cost!

900 picocents per bit

900 picocents per bit

Doc

E[kw]

Practical Techniques for Searches on En
D. Song, D. Wagner, and A. Perrig.

Hundreds of picocents

e Public Key Encryption
D. Boneh, G. Crescenzo, R. Ostrovsky, G.Persiano

A few cpu cycles per word

Sequential search

Docs

Building index

Secure Indexes for Searching Efficiently on Encrypted compressed Data
E.-J. Goh

A few cpu cycles per word

It’s broken

Existing “secure” data outsourcing mechanisms are **2-5 orders of magnitude more expensive** than local execution.

Brute-forcing 80 bit key?

Oracle costs ~ 1 picocent/bit.

$2^{80} \times 80 / 2 = 5 \times 2^{83}$ picocents

~ **$483.5 billion**

for 64 bits … **$5 million** ☺

What can you buy with $1 ?

500,000 2048-bit DSA sigs

(in the comfort of your home)