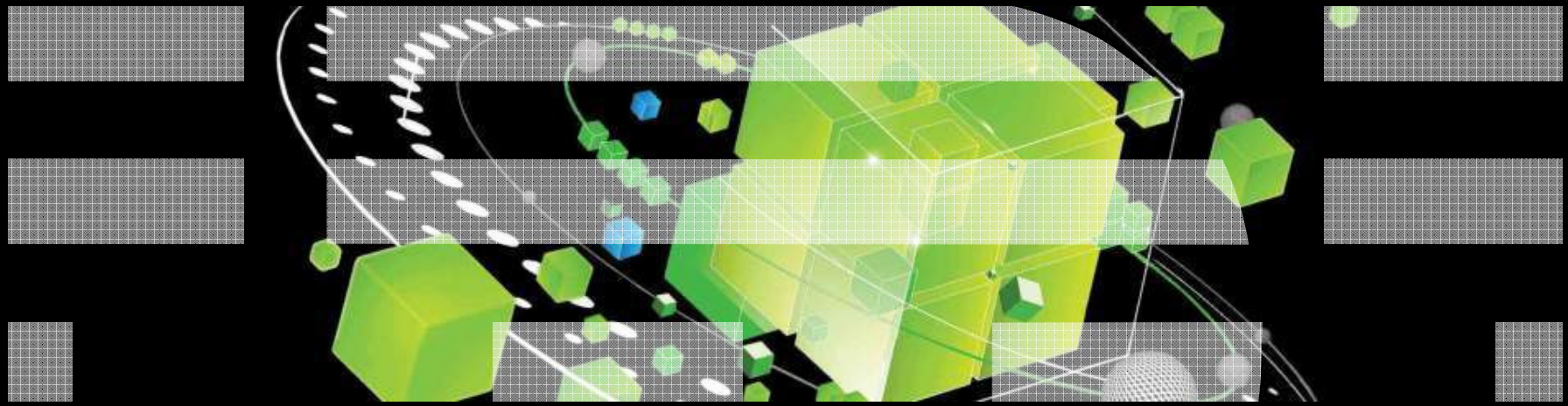
IBM Centennial
Getting ready for a Smarter Planet & Big Data

First Byte Symposium
1st Anniversary of LSDF for Life Sciences at Bioquant
Heidelberg 26 Mai 2011



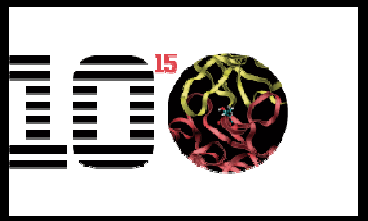
Dieter Münk
Vice President IBM WW Storage
Business Development, Client Care & Technical Support

IBM Centennial

IBM 1401 – The Mainframe

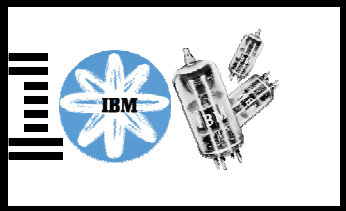
System 360

PC



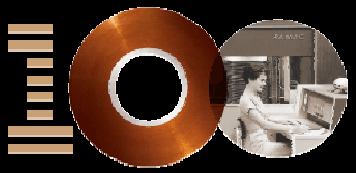
Blue Gene

Magnetic Tape



First Electronic Calculator

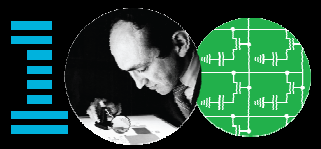
A Computer Called Watson



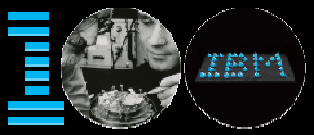
RAMAC

The Floppy Disk

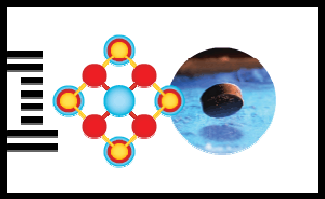
The Selectric Typewriter



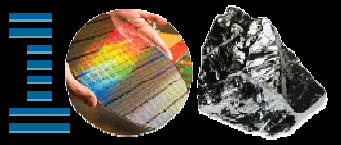
DRAM



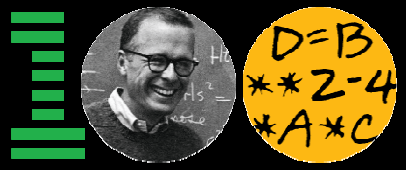
Scanning Tunneling Microscope



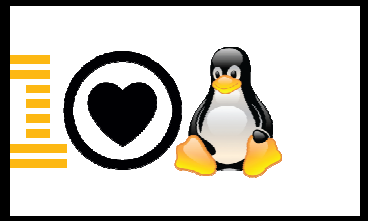
High Temperature Superconductor



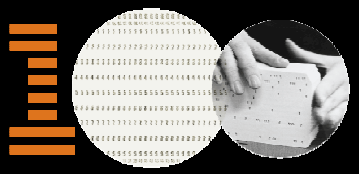
Silicon Germanium Chips



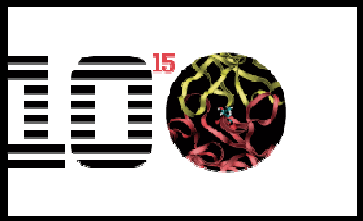
FORTRAN



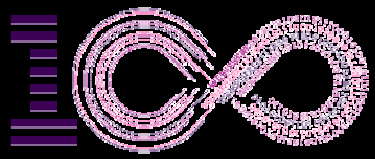
LINUX



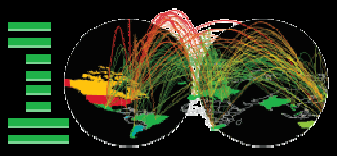
The IBM Punched Card



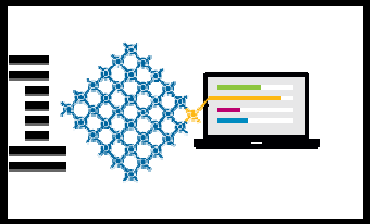
DB2 - Database



Webshpere



Rise of the Internet



World Community Grid

Selective Crime Fighting

Smart Water Management

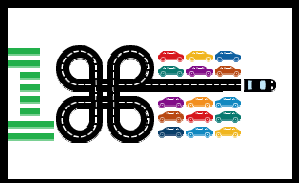
Optimizing Food Supply

Tracking Infectious Diseases

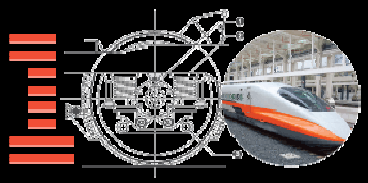
UPC

Magnetic Stripe Technology

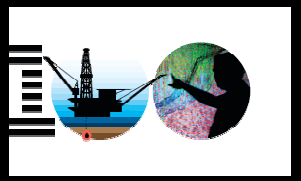
Sabre – Online Reservation



Manangement of
Transportation Flow

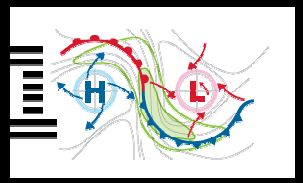


Optimization of Global Railways

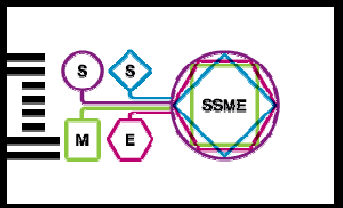


Optimization of
Oil Supplies

Mapping of Humanity's Family Tree

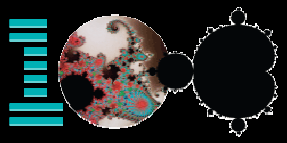


Deep Thunder

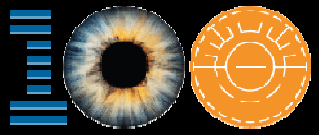


Service Science

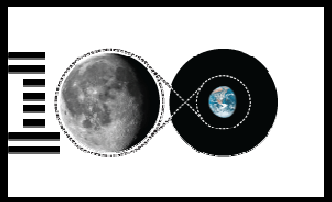
Machine-aided Translation



Fractual Geometry

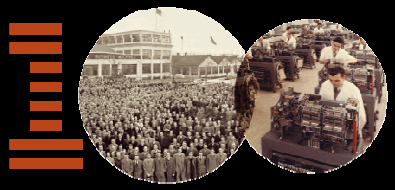


Excimer Laser Surgery



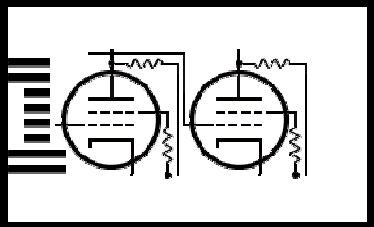
Apollo Mission

The Accessible Workforce



First Salaried Workforce

The Equal Opportunity Workforce



Patent & Innovation

A Culture of Think

Good Design is Good Business

Global Innovation Jam

Corporate Service Corps

The Making of IBM

Welcome to the decade of Smart

Every decade or so in computing, there is a chance to redefine the playing field.

We are in that phase of redefinition now

The World Is Becoming Smarter

FLATTER

SMALLER

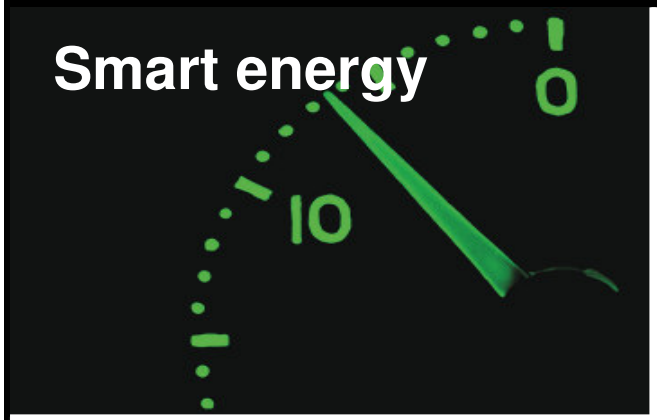
INSTRUMENTED

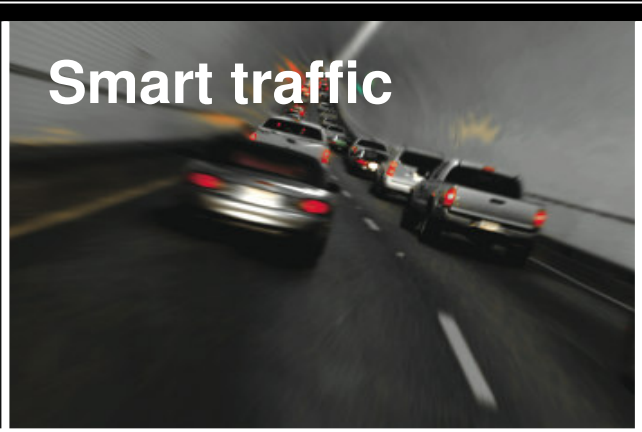
INTERCONNECTED

INTELLIGENT

The World is Becoming Smarter Every Day

- 30 billion RFID tags sold in 2010
- 4 billion camera phones sold through 2009
- 900 million GPS devices sold annually by 2013
- 76 million smart electric meters in 2009. 200M by 2014
- 2 billion people on the Web by 2011



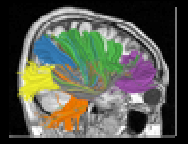


Four Technologies that Will Change Industries, Clients and the World

Compute+ Natural Language+ Analytics

Deep Q&A Computers

Program → Learn



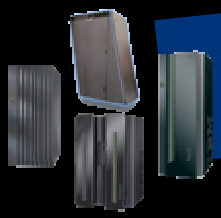
Cognitive Computing
- "Synapse" devices

Big Data

1,000 → 1,000,000 X
Smarter Planet
(Internet of Things + People)

BIG/*Fast*
- Data + analytics (zettabytes + milli / microseconds

Workload Optimized Systems



1000X

Exascale
(Datacenter-in-a-box)
- Massive parallelism
- Flexible system optimization

Nano Devices

1B Transistors

1000X

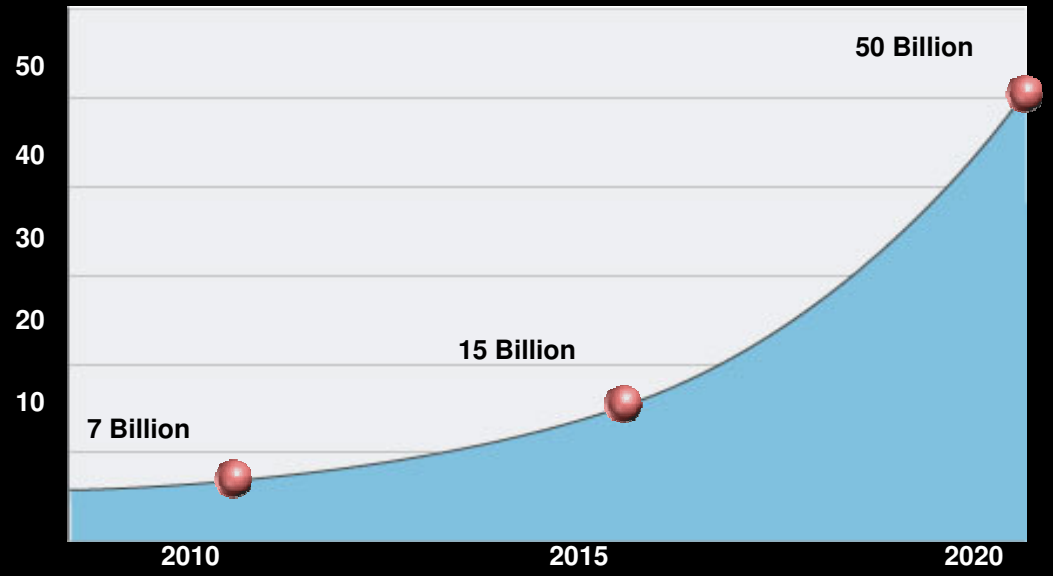
1T Devices

Nano Systems
(Systems-on-a-chip)
- Photonics
- DNA Transistor

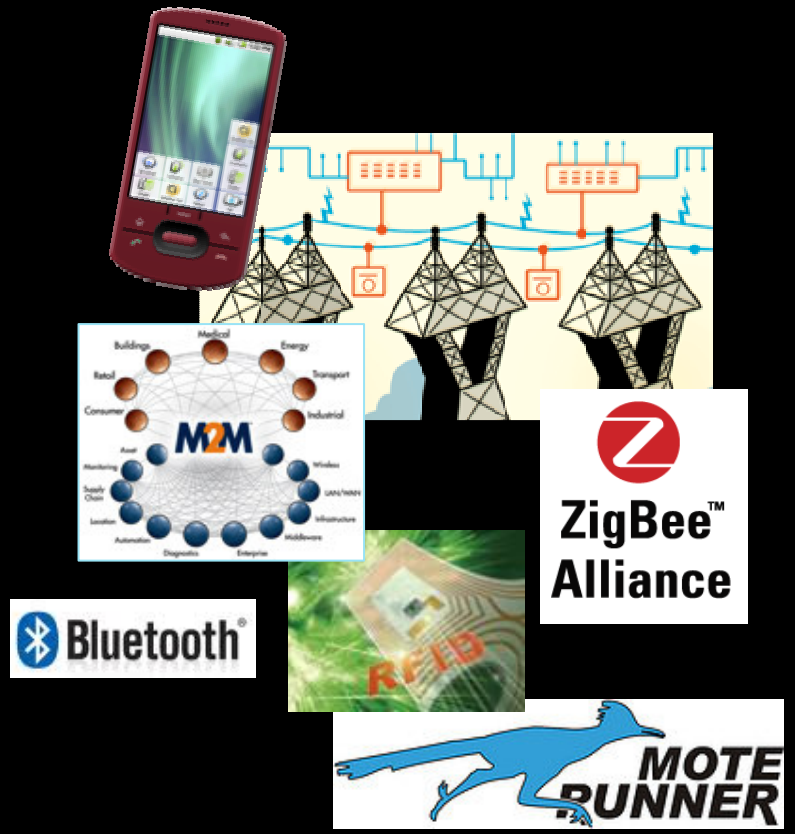
Power7 chip

Smarter Planet will Drive the Creation of **Big**/*Fast* Data

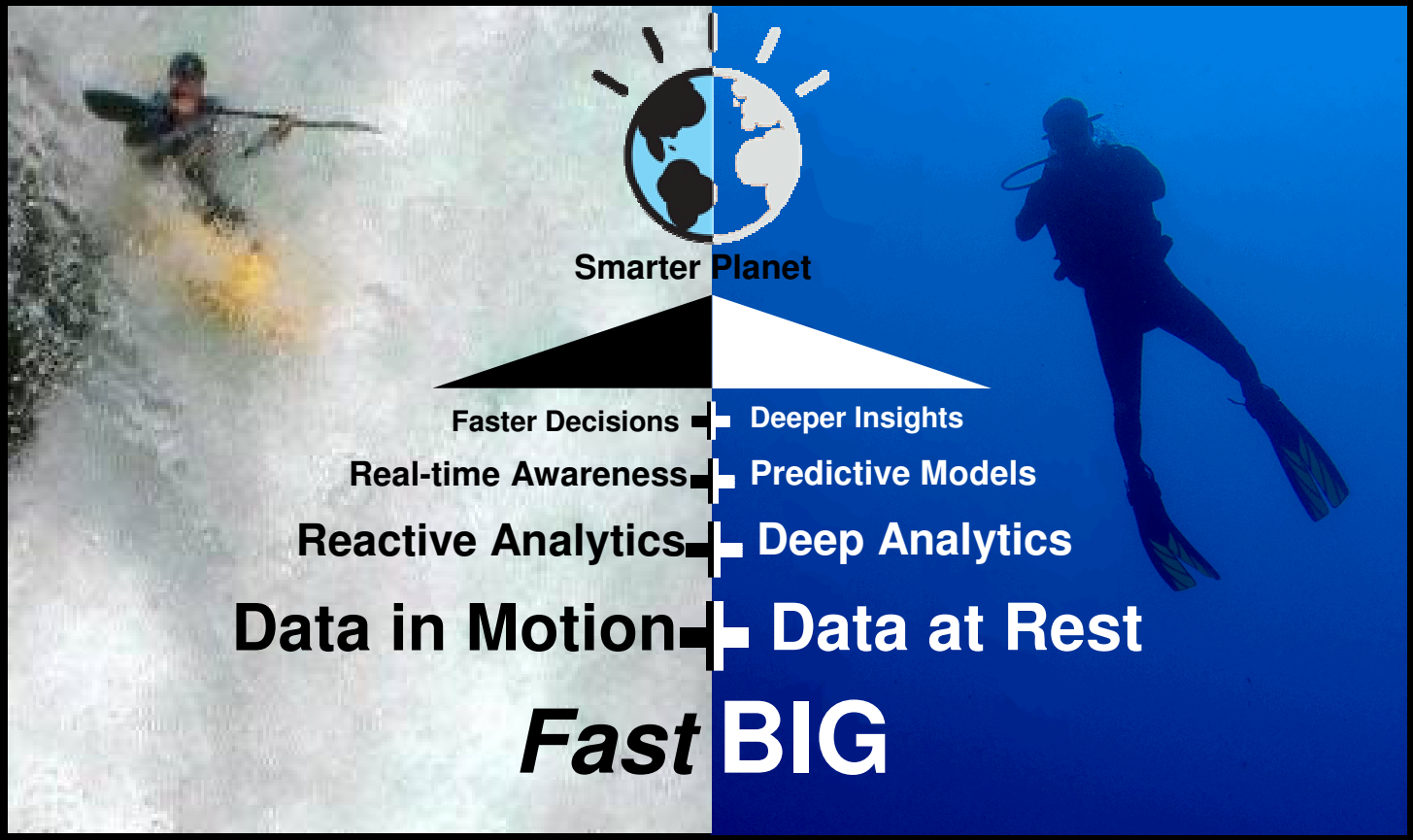
Number of Connected Devices



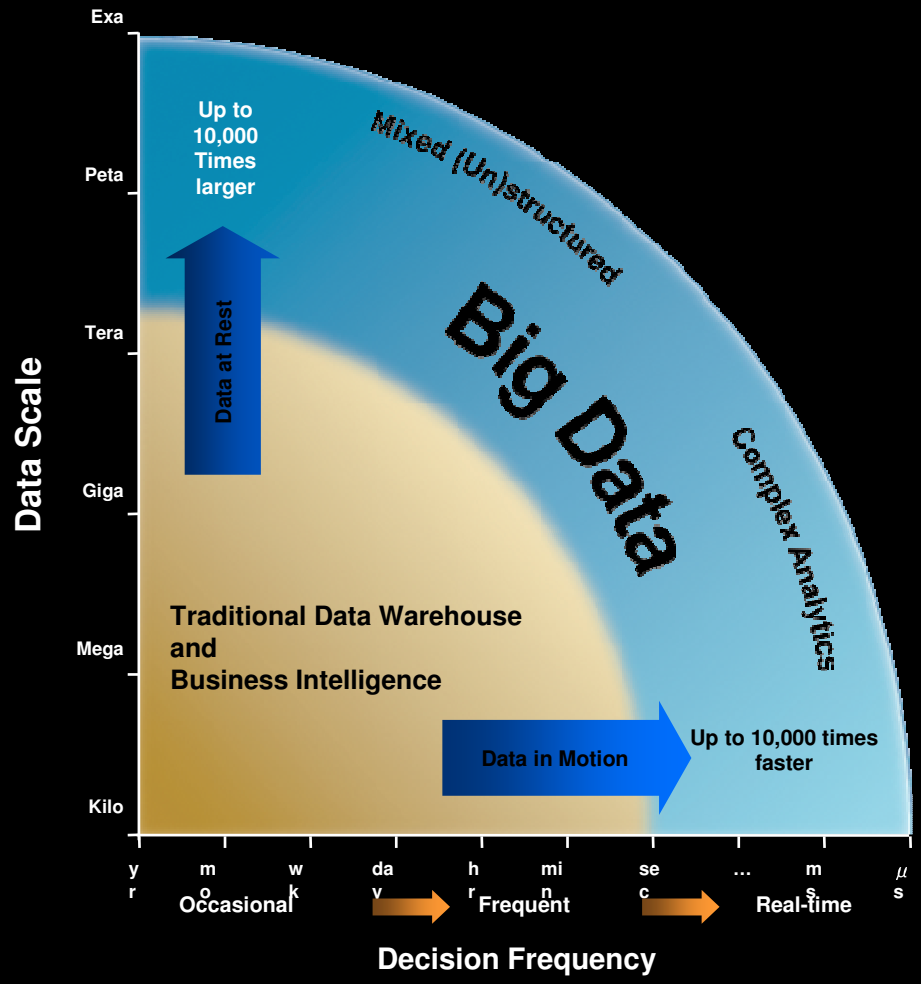
Multiple Sources: Intel, Ericsson, Gartner, etc.



Every Smarter Planet Solution Has **Big**/*Fast* Data and Needs **Big**/*Fast* Analytics



New **Big**/*Fast* Data Brings New Opportunities, Requires New Analytics



Homeland Security
600,000 records/sec, 50B/day
1-2 ms/decision
320TB for Deep Analytics

bharti

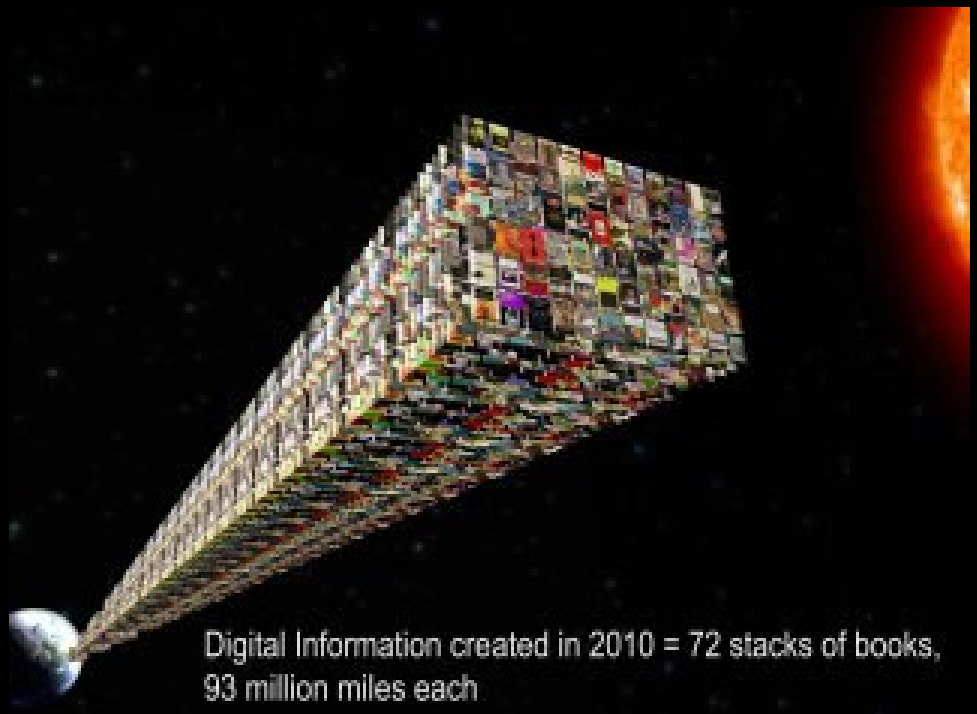
Telco Promotions
100,000 records/sec, 6B/day
10 ms/decision
270TB for Deep Analytics

DeepQA
100s GB for Deep Analytics
3 sec/decision

Smart Traffic
250K GPS probes/sec
630K segments/sec
2 ms/decision, 4K vehicles

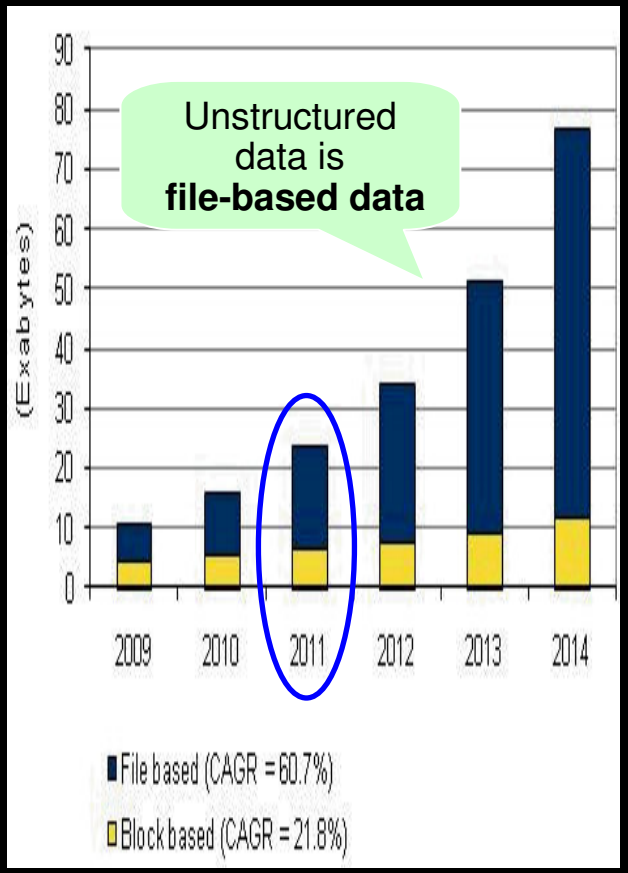
Observe …. the *nature* of workloads and data is rapidly shifting……

Rapid *file storage* growth



=

Worldwide File-Based vs Block-Based Storage Capacity Shipments, 2009–2014



Source: IDC's 2010 Enterprise Disk Storage Consumption Model

Space-Time-Travel

6 billion mobile phones

6.8 billion people

Source - blog by: Jeff Jonas/Las Vegas/IBM, Chief Scientist, IBM SWG Entity Analytics http://jeffjonas.typepad.com/

Space-Time-Travel

Geo-location data

Mobile Phones
600B transactions / day
(in US)

De-Identify

in volume
in real-time
share with third parties

6 billion mobile phones

6.8 billion people

Reveal

Where you spend time
Who with (e.g., friends)

Re-Identify

(figuring who is who)
is somewhat trivial

Source - blog by: Jeff Jonas/Las Vegas/IBM, Chief Scientist, IBM SWG Entity Analytics http://jeffjonas.typepad.com/

Space-Time-Travel

Absolute identification

Ultimate *biometric*

Unravel

All of one's *secrets*

Enormous Opportunity

Challenge all notions of privacy

Reshape

Tough problems
Image classification
Identification

Here Now

More to come

6 billion mobile phones

6.8 billion people

Source - blog by: Jeff Jonas/Las Vegas/IBM, Chief Scientist, IBM SWG Entity Analytics http://jeffjonas.typepad.com/

Possible….. Like Magic …

87% certainty **where you will be this Thursday at 5pm**

Top 10 people **you co-locate with (home / work)**

High quality traffic-avoid predictions ***pushed to you real-time***

Transactions not consistent with your pattern = ***reduce credit card theft 90%***

Political opponent crushed, resigns two days after announcing candidacy

Governments change

Due to mass online social networking

Cannot truly be turned off

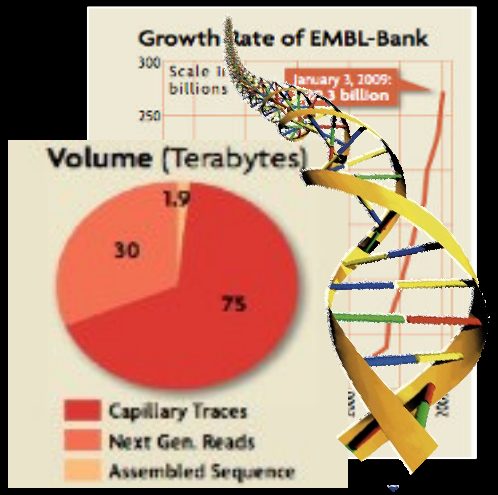
6 billion mobile phones

6.8 billion people

Source - blog by: Jeff Jonas/Las Vegas/IBM, Chief Scientist, IBM SWG Entity Analytics http://jeffjonas.typepad.com/

Let's examine what IS being done <u>today</u> with all this "Big Data"……

Transactions: 46 Terabytes per year



Genomes: Petabytes per year

facebook

10 Terabytes per day

twitter

7 Terabytes per day

http://data.gov

AUSTIN ENERGY

100 Terabytes per year

bharti

Call Records: 3 Terabytes per day

Blogs: 10 Terabytes per year

“All the Data” Big Data makes possible paradigm shifts. Imagine:

Fraud detection = Rules

Fraud detection = “All The Data”

Existing “state of the art” fraud technology

10,000 rules for fraud detection

As new information comes in, add a rule

As transaction analytics determines patterns, add a rule

Need to understand and modify total set of rules

Not real time

- in terms of incoming information
- i.e. not a ‘tripwire’ approach

So in 2 years, you’ll have 20,000 rules

- Doesn’t scale as well

Future state of the art credit card fraud detection

Process **all the data** for

- That individual person
- Everything they’ve bought for the past 4 years
- Every place they’ve bought it

Evaluate each transaction in real time (4 seconds)

- Translation: **global indexes in memory**

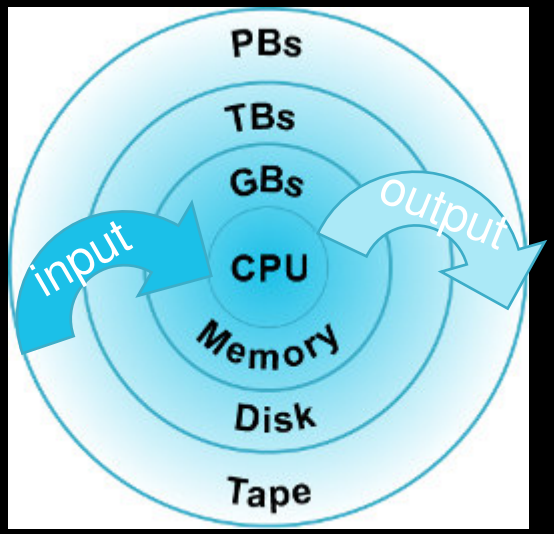
Transforms the nature of Fraud Detection

“Watson”

Source - blog by: Jeff Jonas/Las Vegas/IBM, Chief Scientist, IBM SWG Entity Analytics http://jeffjonas.typepad.com/

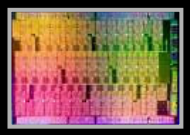
Educate yourself on Big Data-centric Architectures for Performance

Old Compute-centric Model

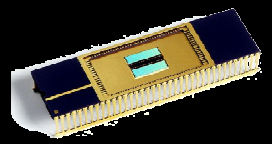


Data lives on disk and tape
Move data to CPU as needed
Deep Storage Hierarchy

Manycore **FPGA**

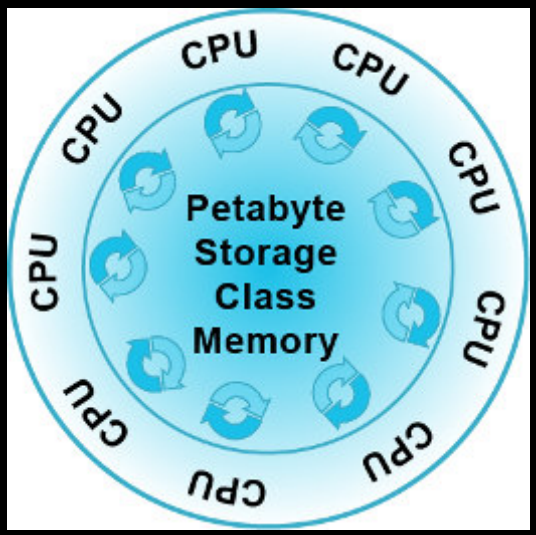


Massive Parallelism
Persistent Memory



Flash **Phase Change**

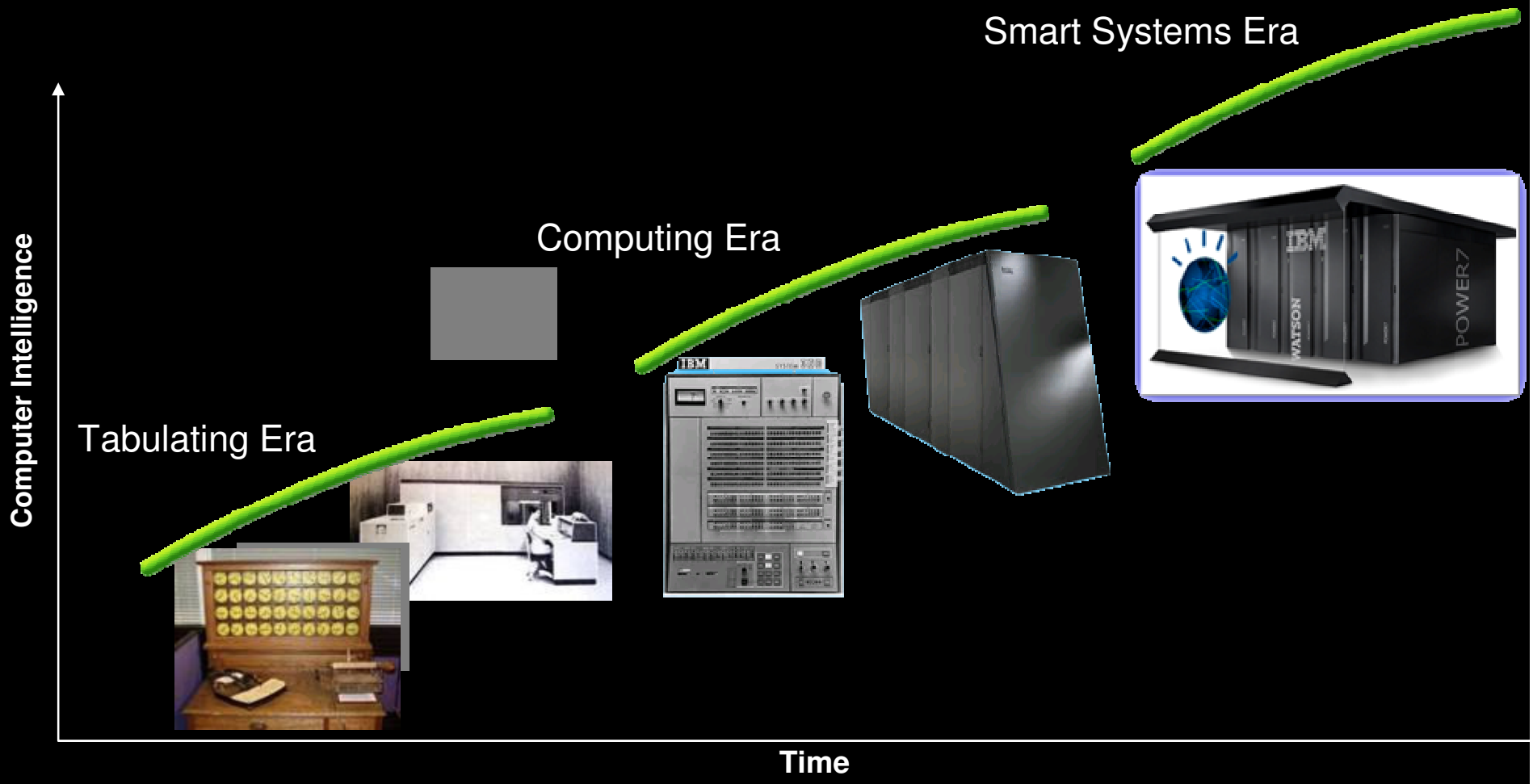
New Data-centric Model



Data lives in persistent memory
Many CPU's surround and use
Shallow/Flat Storage Hierarchy

Largest change in system architecture since the System 360
Having a huge impact on hardware, systems software, and application design

We Are Entering a New Era



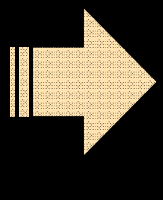
IBM Storage, Servers, Appliance, Software = Solutions for *Monetizing* Data

Change Big Data to *Smart Data*

Peta² Data-centric System, Servers

\+

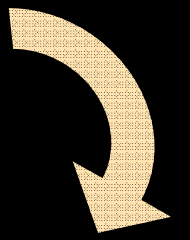
Reactive + Deep Analytics Platform



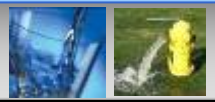
Peta² Analytics Appliance

\+

Big Analytics *Ecosystem*

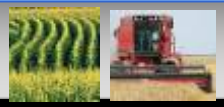


DeepWater
Water management



DeepCurrent
Power Delivery

DeepSoil
Farm Prediction

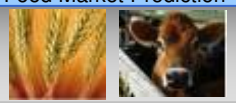


DeepPulse
Political Polling

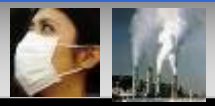
DeepEyes
Webcam Fusion

DeepTraffic
Area Traffic Prediction

DeepBasket
Food Market Prediction



DeepBreath
Air Quality Control



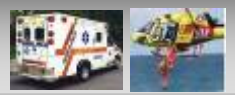
DeepSafety
Police/Security

DeepFriends
Social Network Monitor

Let's build a smarter planet

DeepThunder
Local Weather Prediction

DeepResponse
Emergency Coordination



More than IT ……. It's societal change on a global scale

Social networking = *find collaboration*

Implies re-tooling education for young people

Implies re-tooling our workforce

New Skill Sets:

Mobile device enablement

New technologies

Hadoop, Pig, Hive, Cascading, CR-X, **Netezza, eXFlash, SONAS, GPFS**, etc.

Scale out analytics

Automated modeling

Parallel processing

Parallel file systems

IBMers Value

Dedication to every client's success.

Innovation that matters—for our company and for the world.

Trust and personal responsibility in all relationships.

Thank You