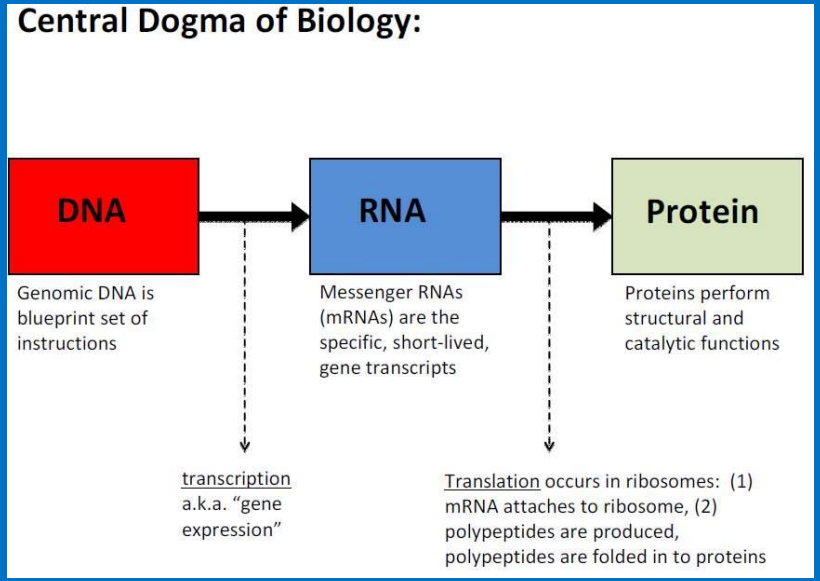
Personalized Medicine and IT

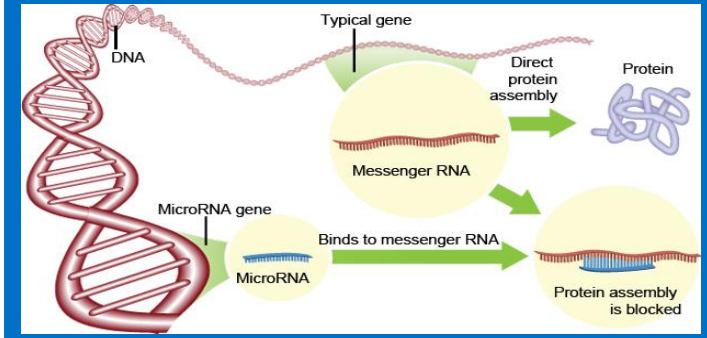
Data-driven Medicine in the Age of Genomics

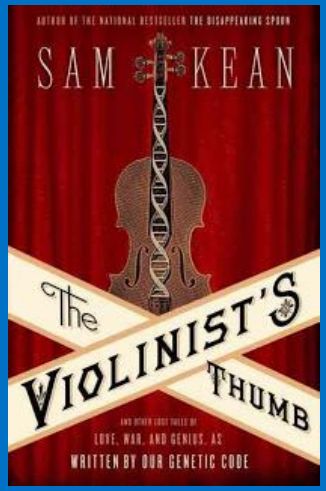
www.intel.com/healthcare/bigdata

Ketan Paranjape
General Manager, Life Sciences
Intel Corp.
@Portlandketan

The Central Dogma of Biology





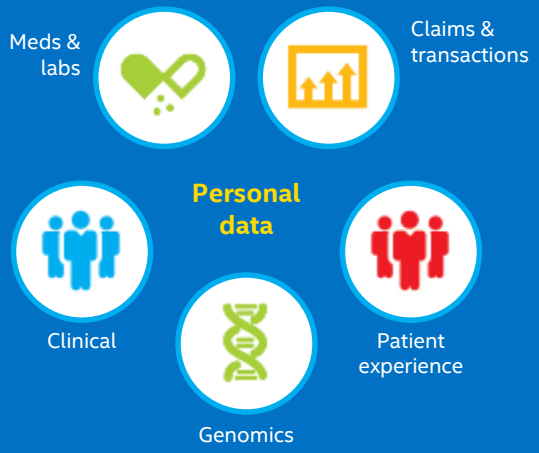


Acknowledgments

It Takes a Village …

Big Data Analytics in Health and Life Sciences

Today: Many disparate data types, streams…



Future: Integrated computing and data

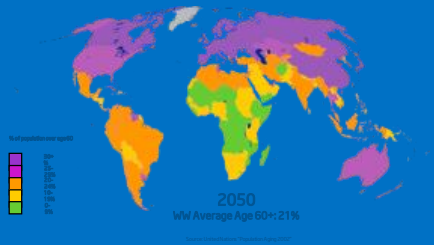
Leading to better decisions

- Improved patient experience
- Healthier population outcomes
- Reduced costs

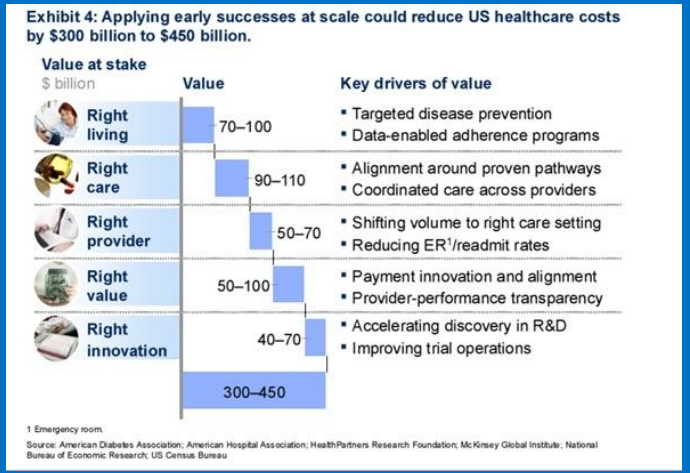
Health & Life Sciences at Intel
Where information and care meet

We are at an Inflection Point in Healthcare - TRENDS

Healthcare costs are RISING
Significant % of GDP



Global AGING
Average Age 60+: growing from 10% to 21% by 2050



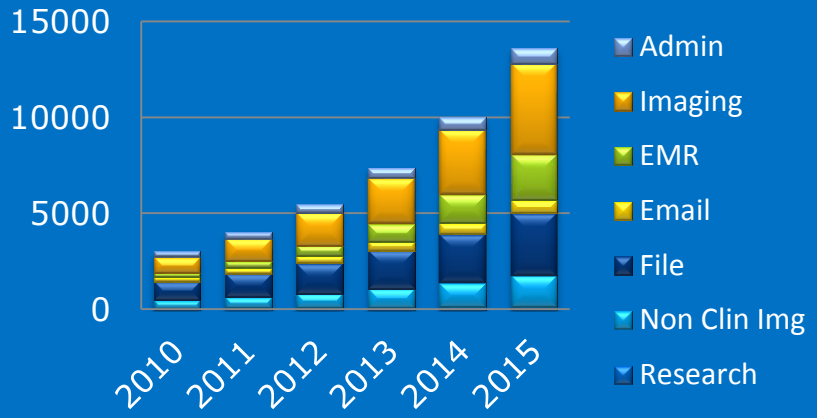
US Healthcare BIG DATA Value
$300 Billion in value/year
~ 0.7% annual productivity growth

Source: McKinsey Global Institute Analysis
ESG Research Report 2011 – North American Health Care Provider Market Size and Forecast

We are at an Inflection Point in Healthcare - TRENDS

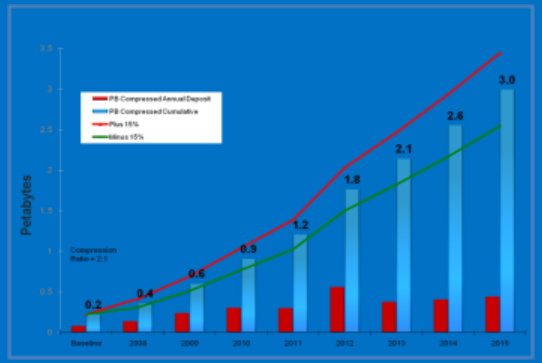
Storage Growth

Total Data Healthcare Providers (PB)



Medical Imaging Archive Projection

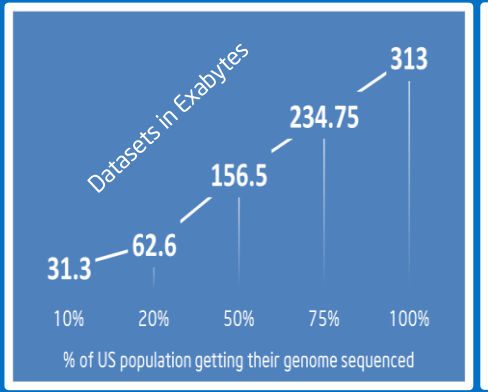
Case from just 1 healthcare system

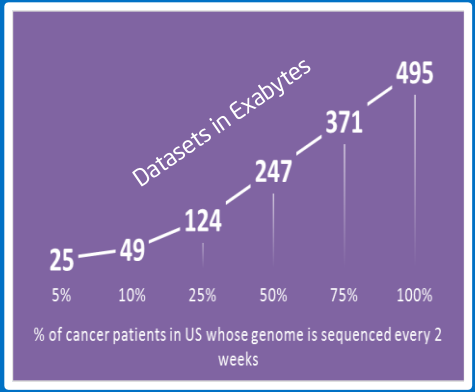


Data Explosion projected to reach 35 Zetabytes by 2020, with a 44-fold increase from 2009[5]

Source: McKinsey Global Institute Analysis
ESG Research Report 2011 – North American Health Care Provider Market Size and Forecast

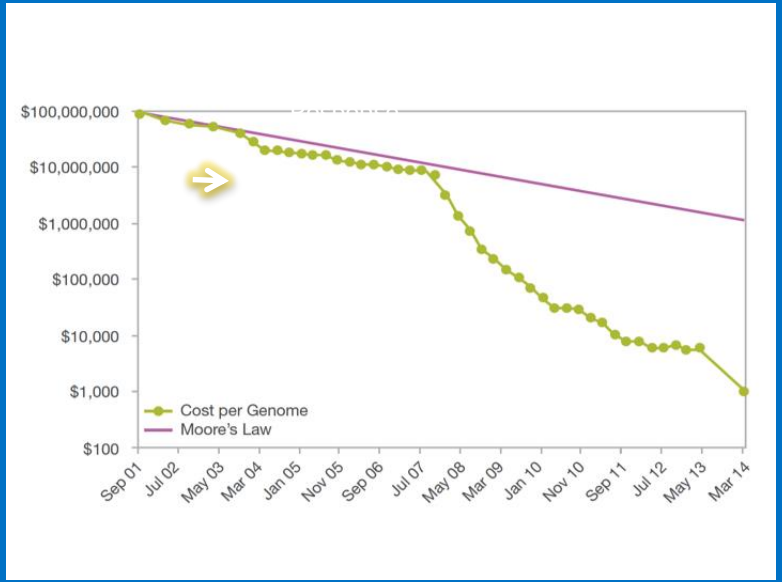
We are at an Inflection Point in Healthcare - TRENDS





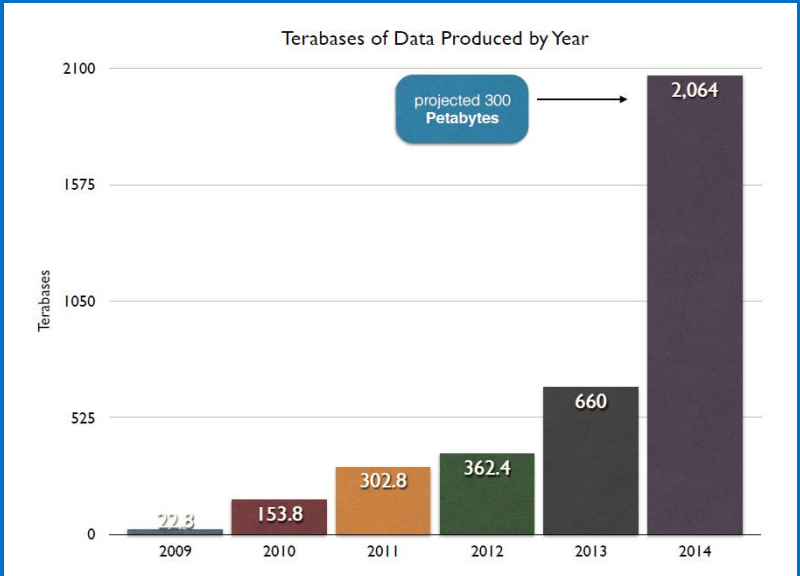
313 *Exabytes*
if everyone in the US has their genes sequenced

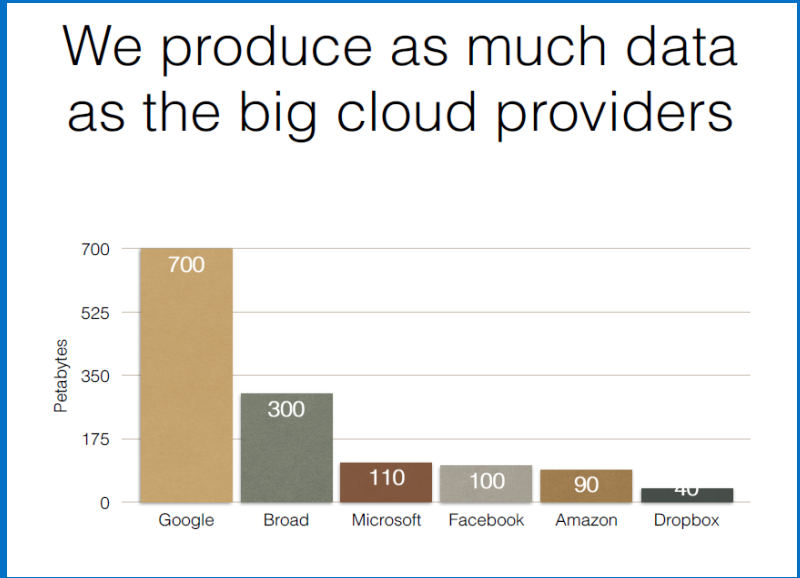
495 *Exabytes*
if every cancer patient in the US has their genes sequenced every 2 weeks



Cost of sequencing is rapidly falling…

Genomics is a Big Data Problem



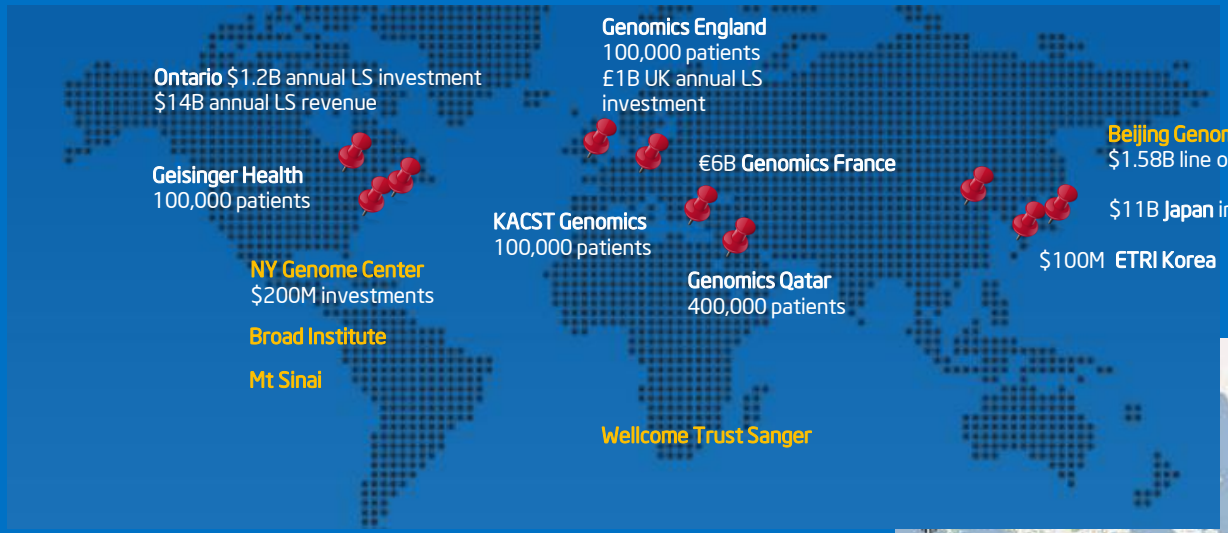


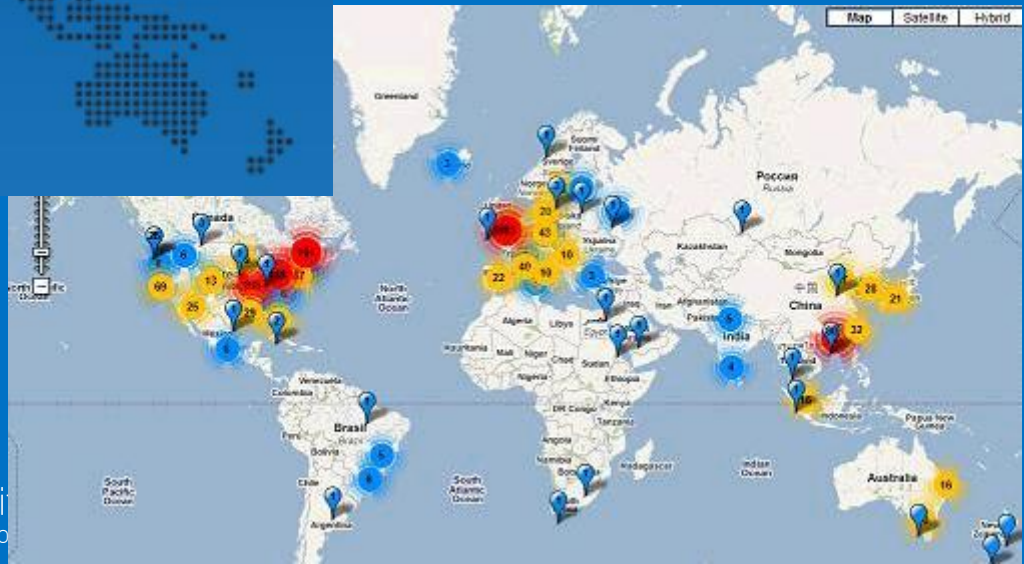
The Broad Institute will produce more data than Microsoft, Facebook and Amazon combined by 2015 ....

BROAD INSTITUTE

The Challenges of analyzing hundreds of thousands of genomes; Mauricio Carneiro, PhD, Broad Institute

Life Sciences World Map



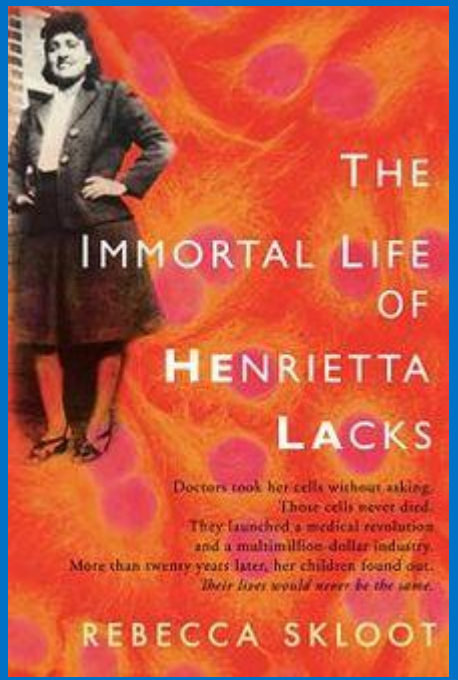


We are barely scratching the surface …

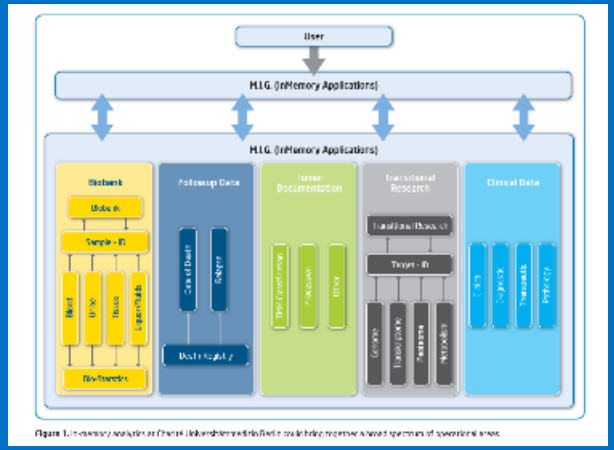
Health & Li
Where info

Barriers

- Clinical
  - Lab, Clinical, CLIA, Direct to Test (DTC), Clinical efficacy
- Literacy and societal challenges
  - Education, Clinical pathway/guidelines, Transparency, Trust, Accountability
- Technical
  - Data Generation, Management, Interpretation, Storage
- Ethical
  - Clinical research, Privacy, Discrimination (GINA)
- Economic and Commercial
  - Insurance, Reimbursement, Preventative Shift, Ownership Sustainability



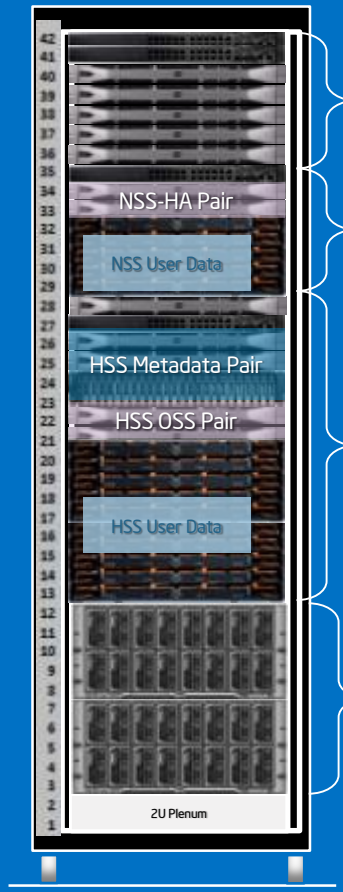
Charite **"Real-time" Cancer Analysis** – Matching proper therapies to patients using **in-memory techniques**



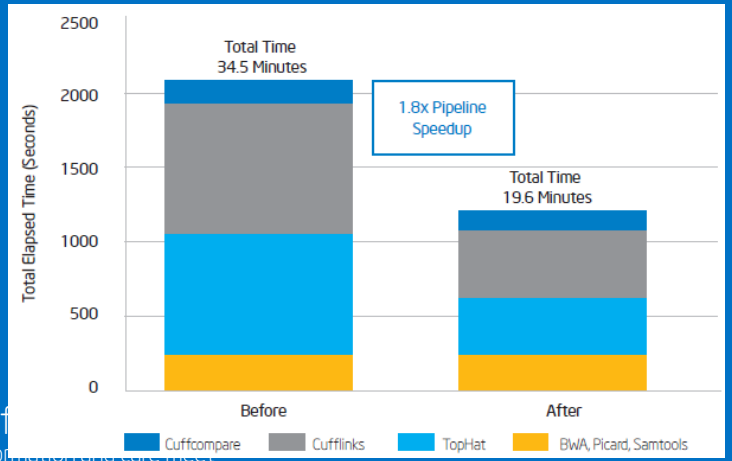
- **Challenge**: Real-time analysis of cancer patients using in-memory SAP HANA Oncolyzer database running on Intel® Xeon® family infrastructure. (3.5M Data points per Patient, Up to 20 TB of data/patient)
- **Solution**: Using structured and unstructured data to collect and analyze tables used to take up to **two days -- now takes seconds**
- **Benefits**: Improves medical quality in disruptive way for Patient, Doctor, Hospital, Research

http://moss.ger.ith.intel.com/sites/SAP/SAP%20account%20team%20documents/Marketing/SAP%20HANA/SAPHANA_Charite_case_study_HI.PDF

HPC **Appliances** for Life Sciences



- **Challenge:** Experiment processing takes 7 days with current infrastructure. Delays treatment for sick patients
- **Solution:** Dell Next Generation Sequencing Appliance
  - Single Rack Solution; 9 Teraflops, Lustre File Storage; Intel SW tools
- **Benefits:** RNA-Seq processing reduced to **4 hour**
- Includes everything you need for NGS - compute, storage, software, networking, infrastructure, installation, deployment, training, service & support



Health & Lif
Where infor

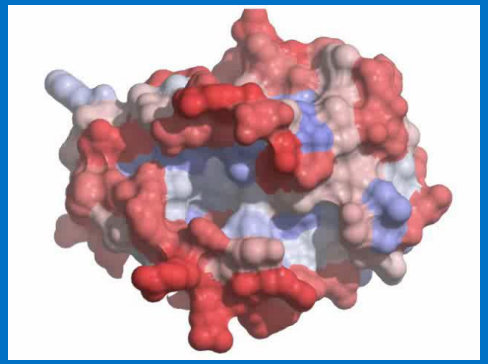
*Other names and brands may be claimed as the property of others.

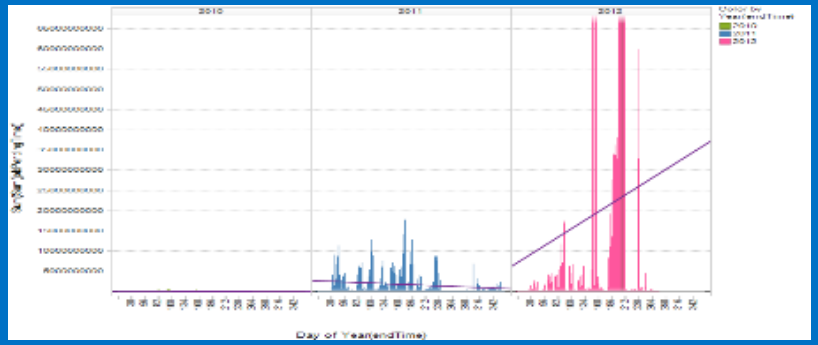
Actual placement in racks may vary.

** 2-socket Intel(R) Xeon(R) CPU E5-2687W / 3.1 GHz

intel Look Inside.

High Throughput Science: Embracing **Cloud-based Analytics** for Computational Chemistry Simulation





- **Challenge:** Sustaining 50000+ compute cores for large scale simulations, for less than a week; CapEX v. OpX
- **Solution:** Novartis leveraged software from AWS partner, Cycle Computing, and MolSoft to provision a fully secured cluster of 30,000 CPUs, powered by the Intel® Xeon® processor E5 family.
  - **Completed screening of 3.2 million compounds in approximately 9 hrs, compared to 4 -14 days on existing resources.**

NOVARTIS INSTITUTES FOR BIOMEDICAL RESEARCH

amazon web services

CYCLECOMPUTING

MolSoft molecules in silico

Regional Health Information Network
RHIN – China (Jinzhou, Pop 3M)

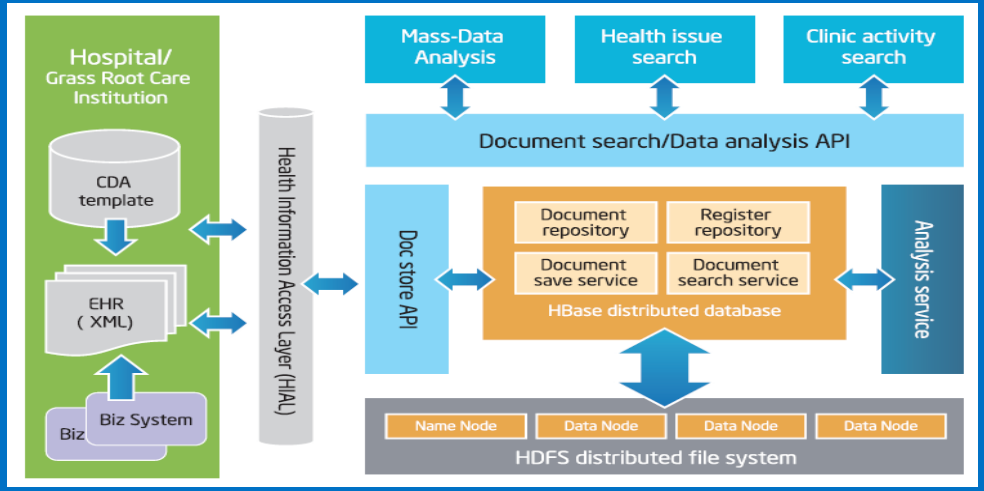
Challenge: RHIN has challenges with scalability, performance and maintenance. Data storage is expensive

Solution: EMR data and healthcare services running on Apache Hadoop* and Intel® Xeon® E5 servers

Benefits: High performance and scalability demonstrated via POC and stress testing. Significantly reduced storage cost

1/5 Reduction in Response Time;

5x Concurrent Users



Data processing flow of RHIN platform

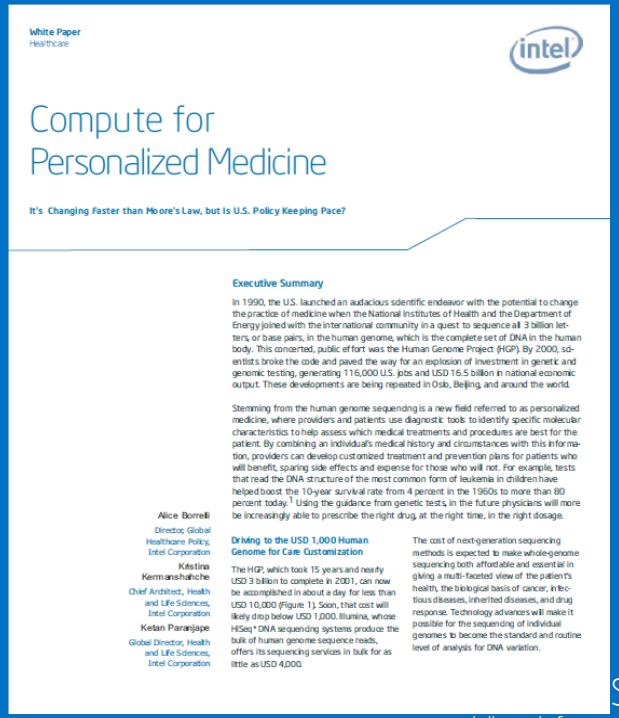
http://hadoop.intel.com/pdfs/IntelChinaHealthyCityAnalyticsCaseStudy.pdf

Training Programs

Bioinformatics, Life Sciences, Computer Sciences, Clinicians

Policy – United States, European Union

Snapshot of US, EU Recommendations



White Paper
Healthcare

intel

Compute for Personalized Medicine

It's Changing Faster than Moore's Law, but Is U.S. Policy Keeping Pace?

Executive Summary

In 1990, the U.S. launched an audacious scientific endeavor with the potential to change the practice of medicine when the National Institutes of Health and the Department of Energy joined with the international community in a quest to sequence all 3 billion letters, or base pairs, in the human genome, which is the complete set of DNA in the human body. This concerted, public effort was the Human Genome Project (HGP). By 2000, scientists broke the code and paved the way for an explosion of investment in genetic and genomic testing, generating 116,000 U.S. jobs and USD 16.5 billion in national economic output. These developments are being repeated in Oslo, Beijing, and around the world.

Stemming from the human genome sequencing is a new field referred to as personalized medicine, where providers and patients use diagnostic tools to identify specific molecular characteristics to help assess which medical treatments and procedures are best for the patient. By combining an individual's medical history and circumstances with this information, providers can develop customized treatment and prevention plans for patients who will benefit, sparing side effects and expense for those who will not. For example, tests that read the DNA structure of the most common form of leukemia in children have helped boost the 10-year survival rate from 4 percent in the 1960s to more than 80 percent today.[1] Using the guidance from genetic tests, in the future physicians will more be increasingly able to prescribe the right drug, at the right time, in the right dosage.

Alice Borrelli
Director, Global Healthcare Policy, Intel Corporation

Kristina Kermanshahche
Chief Architect, Health and Life Sciences, Intel Corporation

Ketan Paranjape
Global Director, Health and Life Sciences, Intel Corporation

Driving to the USD 1,000 Human Genome for Care Customization

The HGP, which took 15 years and nearly USD 3 billion to complete in 2001, can now be accomplished in about a day for less than USD 10,000 (Figure 1). Soon, that cost will likely drop below USD 1,000. Illumina, whose HiSeq® DNA sequencing systems produce the bulk of human genome sequence reads, offers its sequencing services in bulk for as little as USD 4,000.

The cost of next-generation sequencing methods is expected to make whole-genome sequencing both affordable and essential in giving a multi-faceted view of the patient's health, the biological basis of cancer, infectious diseases, inherited diseases, and drug response. Technology advances will make it possible for the sequencing of individual genomes to become the standard and routine level of analysis for DNA variation.

Develop an ICT-enabled European Strategy for Personalised Medicine

2014-2020

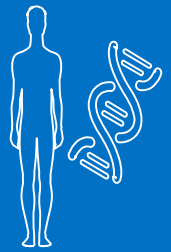
Driving research to unleash the potential of ICT at the point-of-care

EU R&D initiatives must address:

- Interoperability of technical standards for managing and sharing sequence data in research and clinical samples;
- Development of hardware, software and workflow algorithms to accelerate cost efficient analysis of genetic abnormalities that cause cancer and other complex diseases;
- Research to ensure convergence of Big Data and Cloud Computing infrastructure to meet the requirements of High Performance Computing and data throughout the life sciences and healthcare value chains

The eHealth Action Plan 2020 should include Personalised Medicine as a priority

- Gain knowledge of the challenges and barriers (technical, organizational, legal and political) to the adoption of ICT in support of Personalised Medicine leveraged by genomic information;
- Evaluate how to change workflows and education requirements to facilitate adoption of ICT mediated personalized medicine in clinical practice;
- Expand collaboration with other regions of the world in matters of common interest, e.g. by leveraging the eHealth MoU with the United States of America;
- Study, evaluate and disseminate technology neutral risk assessment frameworks for data privacy and security, covering the entire ICT enabled Personalised Medicine delivery chain;
- Develop effective methods for enabling the use of medical information for public health and research



Genes causing it identified & disease pathways determined

Precision medicine regime

Personalized Medicine: All in a day by ~~2050~~ 2020


intel®
Look Inside.™