Welcome to CCGrid-Life 2015

Sandra Gesing
Center for Research Computing
sandra.gesing@nd.edu

Jesús Carretero
University Carlos III o Madrid
Jesus.carretero@uc3m.es

4 May 2015

UNIVERSITY OF NOTRE DAME

CRC CENTER FOR RESEARCH COMPUTING

CCGrid-Life 2015

CRC
CENTER FOR RESEARCH COMPUTING

- Workshop on
  Clusters, Clouds and Grids for Life Sciences
- Chairs
  - Jesus Carretero, U. Carlos III of Madrid, Spain
  - Javier Garcia Blas, U. Carlos III of Madrid, Spain
  - Sandra Gesing, U. Notre Dame, IN, USA
  - Johan Montagnat, CNRS, France
- Merged event from CCGrid-Health 2014 and C4Bio 2014
- 6 accepted submissions out of 10

CCGrid-Life Program

CRC
CENTER FOR RESEARCH COMPUTING

Coffee Break 9:30 - 10:30 **Welcome and Keynotes**	
10:30 - 11:00	Welcome and Introduction by the Workshop Chairs
11:00 - 12:00	Keynote: Science gateways for medical imaging: VIP, CBRAIN, and their interoperability Tristan Glatard
Lunch Break 12:00 - 13:30 **Session 1**	
13:30 - 15:30	Analysing Cancer Genomics in the Elastic Cloud Christopher Smowton, Crispin Miller, Wei Xing, Andoena Balla, Demetris Antoniades, George Pallis and Marios D. Dikaiakos Scaling Machine Learning for Target Prediction in Drug Discovery using Apache Spark Dries Harnie, Alexander E Vapirev, Jörg Kurt Wegner, Andrey Gedich, Marvin Steijaert, Roel Wuyts and Wolfgang De Meuter A Comparative Analysis of Scheduling Mechanisms for Virtual Screening Workflow in a Shared Resource Environment Bui The Quang, Jik-Soo Kim, Seungwoo Rho, Seoyoung Kim, Sangwan Kim, Soonwook Hwang, Emmanuel Medernach and Vincent Breton SparkSW: scalable distributed computing system for large-scale biological sequence alignment Guoguang Zhao, Cheng Ling and Hongdong Sun
Coffee Break 15:30 - 16:00 **Session 2**	
16:00 - 17:00	Classifications of computing sites to handle numerical variability Tristan Glatard and Alan Evans Multicenter Data Sharing for Collaboration in Sleep Medicine Maximilian Beier, Christoph Jansen, Geert Mayer, Thomas Penzel, Andrea Rodenbeck, René Siewert, Jie Wu and Dagmar Krefting Conclusions and closing

CCGrid-Life Special Issue

CRC
CENTER FOR RESEARCH COMPUTING

- Special issue on Clusters, Clouds and Grids for Life Sciences
- Future Generation Computer Systems (FGCS), impact factor: 2.639
- Invited papers of CCGrid-Life and open call

Important Dates

Paper submission due: July 31, 2015
First-round acceptance notification: October 31, 2015
Revision submission: December 14, 2015
Notification of final decision: January 31, 2016
Submission of final paper: March 31, 2016
Publication date: June 30, 2016 (Planned)

Guest Editors

Sandra Gesing
University of Notre Dame, USA
sandra.gesing@nd.edu

Johan Montagnat
CNRS, France
johan.montagnat@cnrs.fr

Jesus Carretero
Universidad Carlos III de Madrid, Spain
jesus.carretero@uc3m.es

Javier Garcia Blas
Universidad Carlos III de Madrid, Spain
fjblas@inf.uc3m.es

http://lsgc.org/ccgrid-life/
http://www.journals.elsevier.com/future-generation-computer-systems/call-for-papers/special-issue-on-clusters-cloud-and-grid-on-life-sciences/

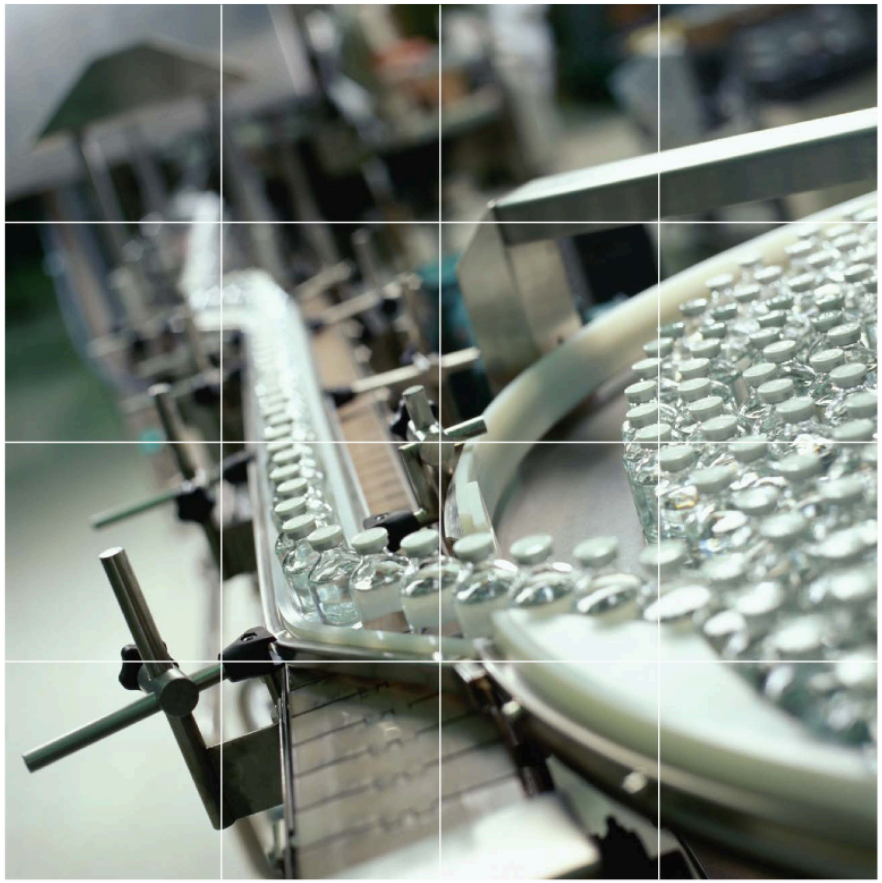
Life Sciences Cluster Report 2014

CRC
CENTER FOR RESEARCH COMPUTING

Life Sciences Cluster Report

Global | 2014

JLL



Global competition heats up as small and mid-tier companies dictate demand

- Innovative leadership is recalibrating as emerging countries, particularly those in Asia, report impressive growth figures in patent applications, R&D funding, labor productivity and science degrees.
- The demand environment has shifted from one driven by Big Pharma companies to one where mid-tier companies and specialty players are experiencing overnight growth and highly variable needs in response to promising products and treatments.

http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Roger's Global Watch List: 10 cities in the spotlight

London/Cambridge/Oxford, United Kingdom

Specialized projects aim to transform Greater South East England into a "golden triangle" of innovation

- A $6.4 million investment aims to connect and transform London, Cambridge and Oxford into a regional innovation powerhouse. Local governments hope this funding will support the commercialization of local innovation.

Leiden, Netherlands

Business-friendly environment and new training center allures new companies & expansions

- Efforts by The WestHolland Foreign Investment Agency (WFIA) aided 42 company relocations/expansions during 2013, the most commitments in 13 years. One such company, medical diagnostic equipment manufacturer, Welch Allyn, plans to consolidate European operations into a single site in Leiden and noted business-friendly taxes in its decision.

Beijing, China

Continued investment and development planned for Zhongguancun

- Nicknamed "China's Silicon Valley", Zhongguancun is the one of the densest scientific, educational and talent resource bases in China. Through 2015, the hub will further improve its science & technology cities, promote the development of the northern R&D and industrial belt, and the southern high-tech manufacturing and emerging industrial belt.

Tsukuba, Japan

Tsukuba Science City aims to become Japan's flagship science and technology hub

- Home of the Tsukuba Science City, which is comprised of dozens of national research institutes and hundreds of private research facilities. The city is applying for 'Strategic Global Innovation Center' status to secure tax benefits for small and medium sized businesses, infrastructure development and training for overseas personnel.

Osong Bio Valley, Chungcheong Province, Korea

The geographical focus of national government efforts to establish a life sciences hub in Korea

- In total, the government plans to invest approximately $5.3 billion in its development over the next 30 years and six governmental institutions will relocate to Osong including the Ministry of Food and Drug Safety and Korea National Institute of Health.

Moscow, Russia

Public and private projects aim to grow Moscow's infrastructure and capabilities

- Russia's spend on medicines has grown dramatically over the past five years and is expected to grow anywhere from 8.0 to 13.0 percent by 2020. Private medicine is on the rise and the Russian government, currently the biggest spender, is also focused on growing domestic capabilities.

Tel Aviv, Israel

A start-up friendly infrastructure further bolstered by new R&D park

- Country-wide, industry companies have grown more than 400.0 percent since 1996. Several incubators have been created to encourage new talent and foster mentoring, which continually attract multi-national interest. A sizable high-tech presence and infrastructure to support start-ups provide additional resources.

Seattle, United States

Poised to join the ranks of Boston, San Diego and the Bay Area with strong R&D capabilities

- The industry and lifestyle dynamics of Seattle strongly mirror those of well-established life sciences hubs in the United States. Specifically nearly half of life sciences employment is based in R&D functions and it's coastal geography and urban environment are highly sought after by the younger generation of workers.

Vancouver, Canada

A rich infrastructure for biotech and R&D bodes well as the industry favors smaller, nimbler organizations

- Among the major Canadian cities with active life sciences industries, Vancouver has the greatest share of R&D jobs and tends to feature more start-ups and small to medium-sized companies, as opposed to large pharmaceutical campuses.

São Paulo, Brazil

A stable national economy & improved industry practices bode well for the city's critical mass of companies

- Brazil currently ranks second among global nations for the production of biotech crops and has the largest medical equipment market in South America. Roughly 68.0 percent of medical device manufacturers are located in São Paulo and the largest share of biotech companies among Brazilian cities.

http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Life Sciences Cluster Report 2014

CRC
CENTER FOR RESEARCH COMPUTING

Geographic shifts in innovation

Globally innovation is shifting to emerging markets.

1 Innovation around the globe

Asia has surpassed both North America and Europe in overall Patent Cooperation Treaty (PCT) applications, a measure of innovative output.

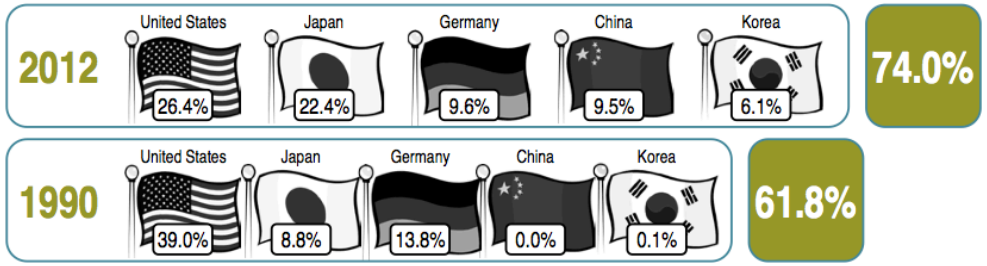
North America: 54,398 applications

Europe: 58,141 applications

Asia: 78,990 applications

2 Global innovation share shifts

As of 2012, the **top five countries accounted for 74.0% of all PCT applications**. Since 1990, China and South Korea are new to the top five, and the United States and Germany have lost share.



3 Innovation growth

Life sciences PCT applications witnessed **3.7% growth from 2011 to 2012.**

Leading growth clusters include:

China: **42.8%**
Russia: **36.8%**
Switzerland: **20.8%**
Japan: **11.2%**
Germany: **7.8%**
Mexico: **7.1%**
The Netherlands: **6.2%**
Brazil: **5.6%**

4 Innovation without collaboration

Emerging clusters showing the highest levels of PCT growth are also the least collaborative, reporting the least amount of collaboration among total applications.

Share of foreign inventors on PCT applications:

Israel: **9.1%**
China: **4.3%**
India: **3.9%**
Korea: **3.6%**
Japan: **3.6%**

5 Innovation with collaboration

Mature North American and Western European countries among the most collaborative nations on PCT applications.

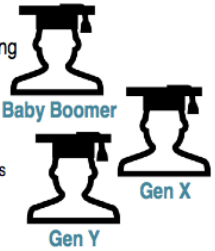
Share of foreign inventors on PCT applications:

Switzerland: **74.5%**
The Netherlands: **47.1%**
Canada: **29.1%**
United Kingdom: **27.2%**
United States: **27.1%**

6 Global education levels

While the United States maintains the highest percentage of the working population (25-64 years old) with bachelor's-type degrees, it loses its lead when only the younger generations (ages 25-34) are considered.

Attainment of bachelor's degree among total population **aged 25-64**

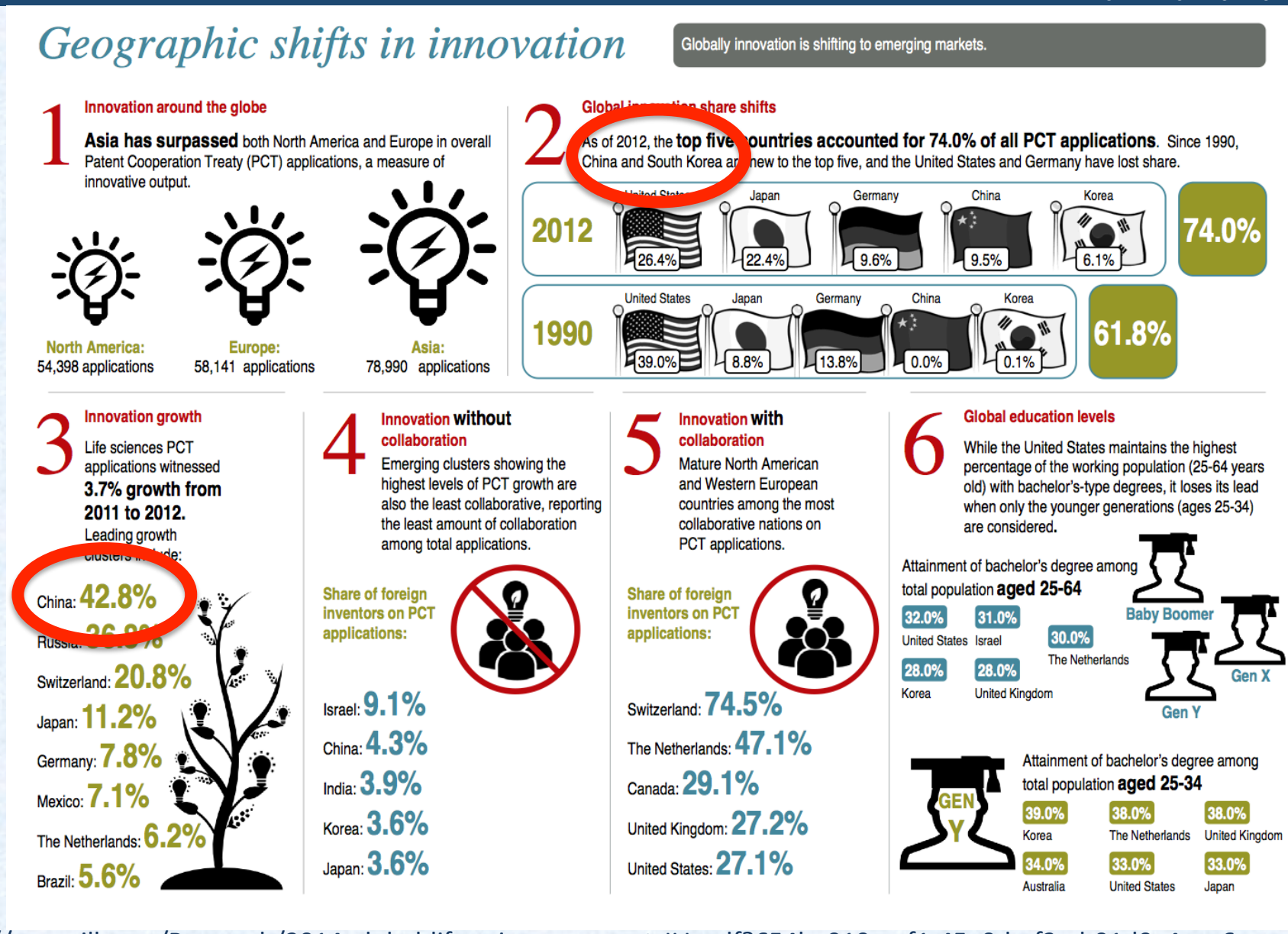


32.0% United States
31.0% Israel
30.0% The Netherlands
28.0% Korea
28.0% United Kingdom

Attainment of bachelor's degree among total population **aged 25-34**

39.0% Korea
38.0% The Netherlands
38.0% United Kingdom
34.0% Australia
33.0% United States
33.0% Japan

http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6



http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Life Sciences Cluster Report 2014

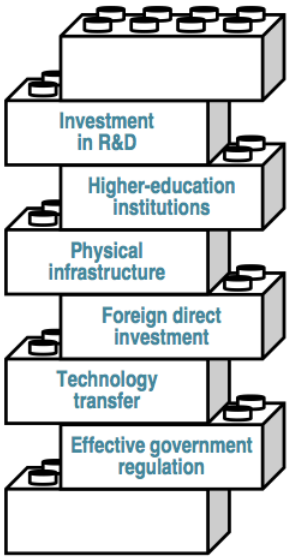
CRC
CENTER FOR RESEARCH COMPUTING

Technology transfer from bench to marketplace

Getting a promising new drug or treatment to market is reliant on factors beyond R&D funding and innovation.

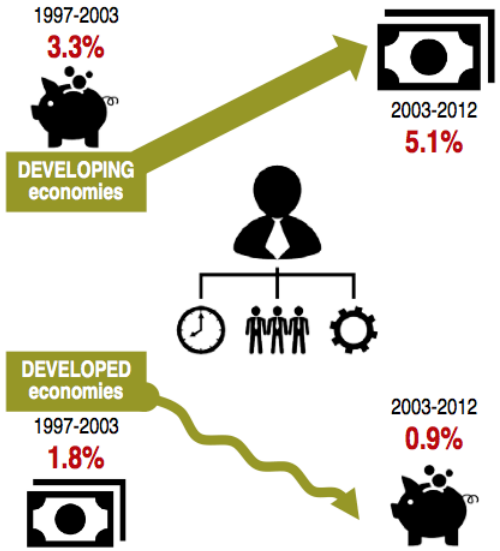
1 Building blocks of development

These foundational components of cluster development are also critical to labor productivity



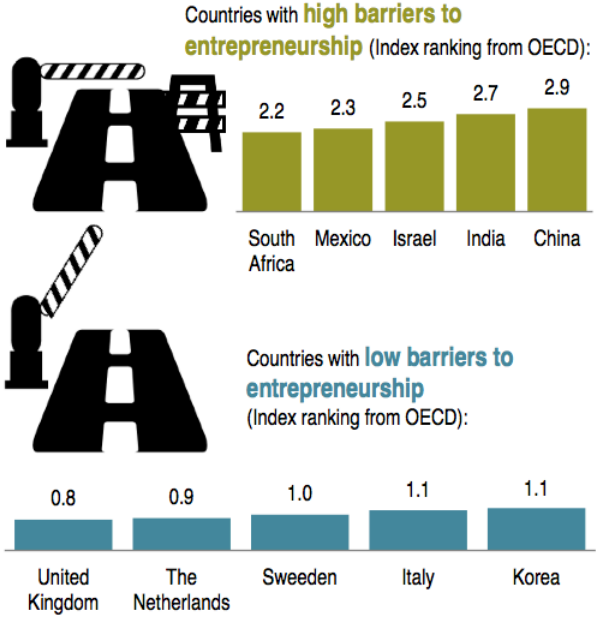
2 Labor productivity differences

Changes in **labor productivity differ greatly** between developing and developed economies



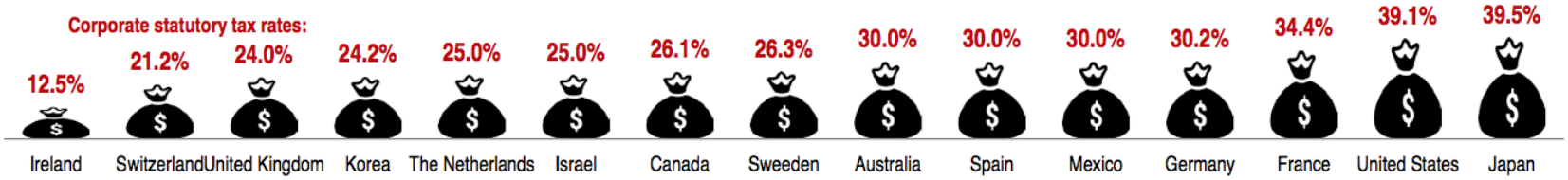
3 Declining entrepreneurship barriers

On the whole, barriers to entrepreneurship have declined in most countries and more established global clusters; however they remain high among emerging life sciences clusters.



4 Taxes and regulation

Corporate statutory tax burdens are **highest among Japan and the United States** at 39.5% and 39.1%, respectively; however these countries also have some of the best regulatory systems and high political and transparency rates.



http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Life Sciences Cluster Report 2014

CRC
CENTER FOR RESEARCH COMPUTING

Efficiency, compliance & facility trends

"The need to streamline costs and fill talent shortages leads to increased scope in IFM outsourcing."

1 To streamline costs, open new doors

Real estate and facility expenses are among the industry's biggest costs. So for the first time, many industry decision makers are now **open to outsourcing** functions in critical and highly regulated space.

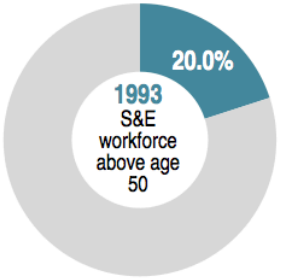
2 Move beyond the yellow line

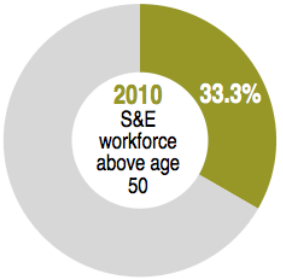
For decades, the yellow line designating a critical/regulated area has served as an actual and symbolic boundary for IFM providers. Advancements in the IFM world, are **breaking down some of the barriers** to entry.

- Technical development of 3rd party providers
- Efficiency gains
- Cost savings
- Synergy between critical & non-critical operations
- Integration & standardization

3 Ageing U.S. workforce

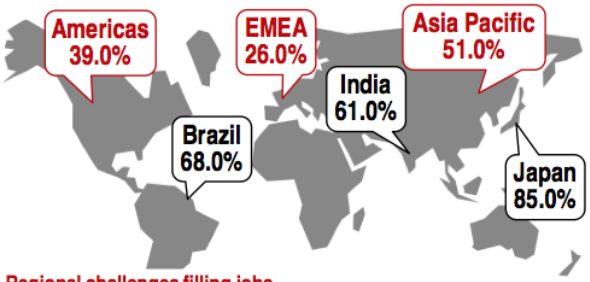
From 2010 to 2020 nearly **2 million life scientist and engineering jobs are expected to open.** A large share of that projection is attributed to a retiring baby boomer S&E workforce.





4 Global talent shortages

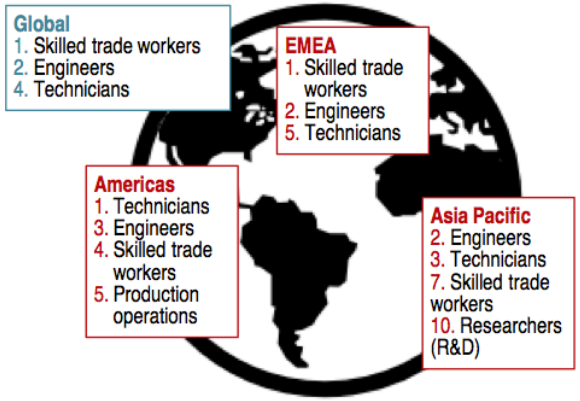
The ManPowerGroup's annual global survey of 38,000 employers in 42 countries revealed that **35.0% report difficulty in filling jobs**, of those 54.0% report this difficulty has a medium or high impact on their ability to meet client needs (up 12 percentage points over 2012).



Regional challenges filling jobs
Top 3 countries reporting difficulty in filling jobs

5 Difficult jobs to fill

The difficulty filling Life Sciences jobs may potentially **lead to rising costs and prevent getting the job done in time and on budget.** As the industry builds more complex and sophisticated facilities, the need for highly skilled workers will only exacerbate this problem.



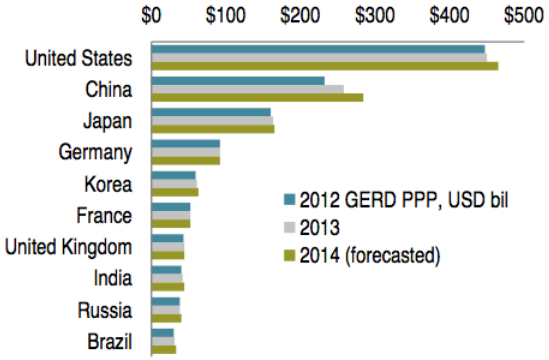
http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Global R&D funding perspective

Globally, the US still leads the world in R&D funding, but developing markets are emerging as new leaders.

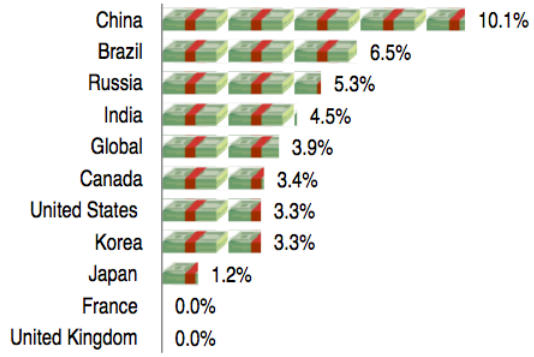
1 Global R&D leaders (3-year view)

European nations post flat growth while developing nations continually post year-over year increases in **Gross Expenditures on Research and Development (GERD)**



2 Emerging R&D leaders

The BRICK nations are emerging as global R&D leaders. China, Brazil, Russia, and India are expected to outpace the overall global total for **year-over-year increases in GERD.**



3 Concentration of R&D spend

10 countries account for **80.0%** of global R&D spend.

4 Source of R&D funding

While businesses contribute the most to R&D spend for most countries, the contribution ratios of this, and other funding sources, vary widely.
(The following is based off an industry-representative sample set of countries)
Highest percentages of:

Business R&D funds:
1. Israel
2. Japan
3. Korea

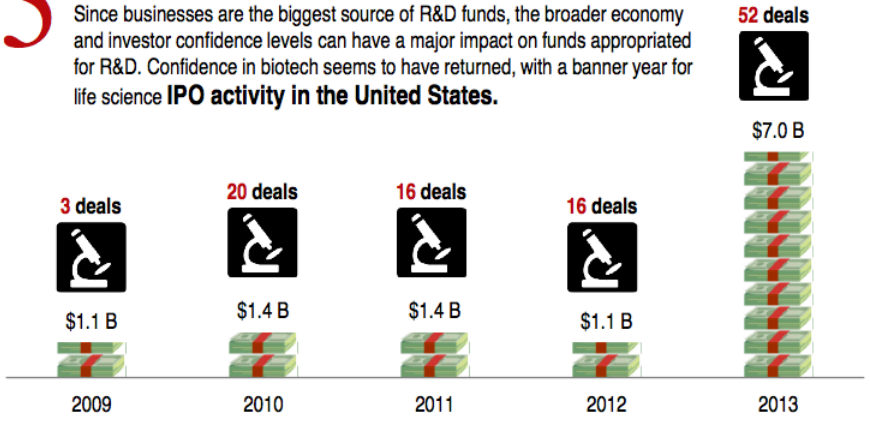
Higher education R&D funds:
1. The Netherlands
2. Canada
3. Mexico

Government R&D funds:
1. Mexico
2. Russia
3. South Africa

Public non-profit R&D funds:
1. United States
2. Israel
3. Australia

5 Economic impact to R&D funding

Since businesses are the biggest source of R&D funds, the broader economy and investor confidence levels can have a major impact on funds appropriated for R&D. Confidence in biotech seems to have returned, with a banner year for life science **IPO activity in the United States.**



http://www.jll.com/Research/2014-global-life-sciences-report-JLL.pdf?654be919-aef1-45a0-bef3-ab01d0a4ece6

Funding

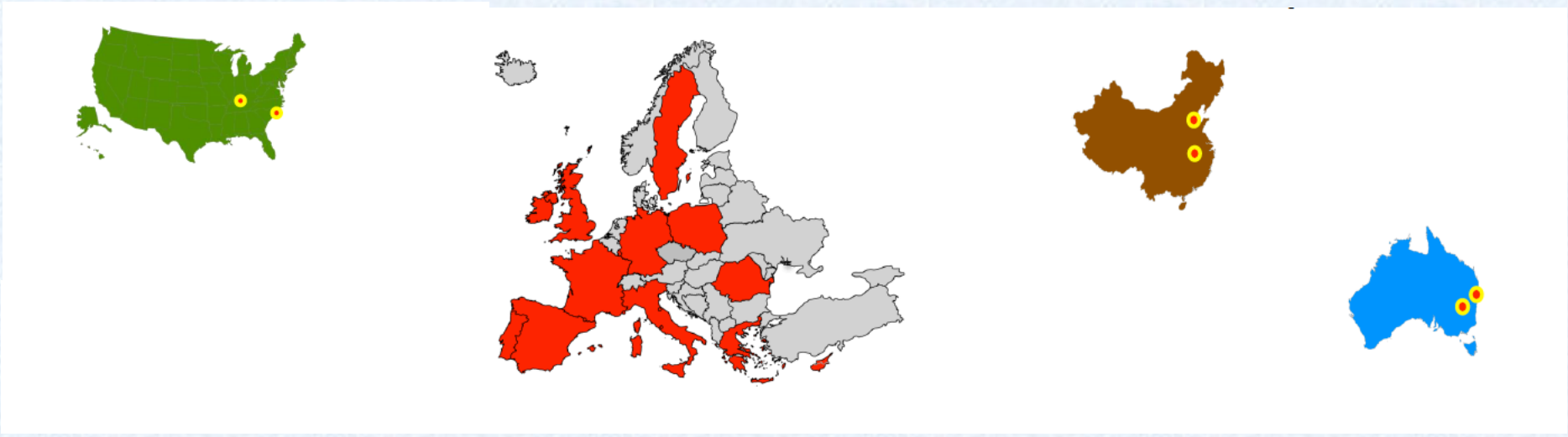

Grant Season
Congratulations on getting that summer research grant, Dr. D! What are you going to do with it?
Thanks. I'm going to spend it researching more grants!
#1 DAD
AS '11

EU COST Action cHiPSet (IC1406)

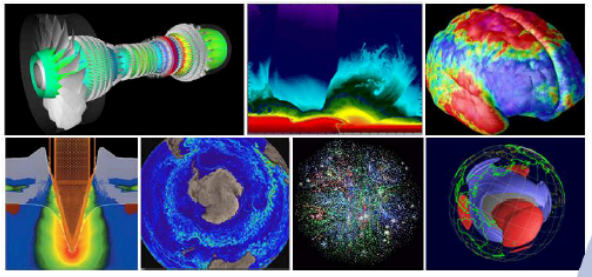
CRC
CENTER FOR RESEARCH COMPUTING

cHiPSet – High Performance Modelling and Simulation for Big Data Applications

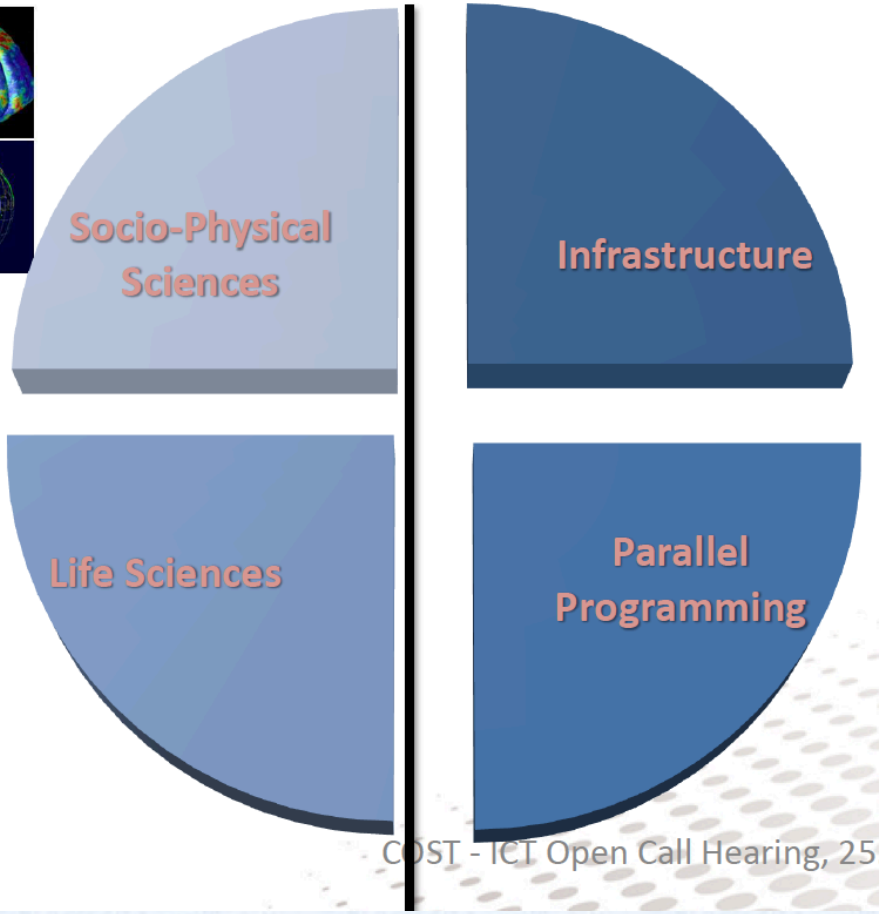
- April 2015 – April 2019
- 15 countries - 12 COST, 3 non-COST (US, China, Australia)
- 37 reseach organizations/companies (31 COST, 6 non-COST)

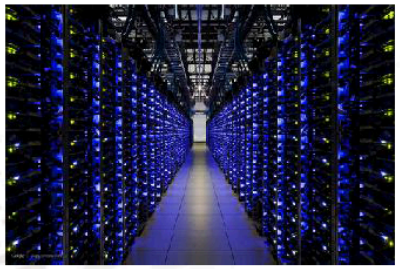


Structure & Coordinate Research on **HPC**-enabled **Modelling and Simulation (MS) for Big Data** in Europe.



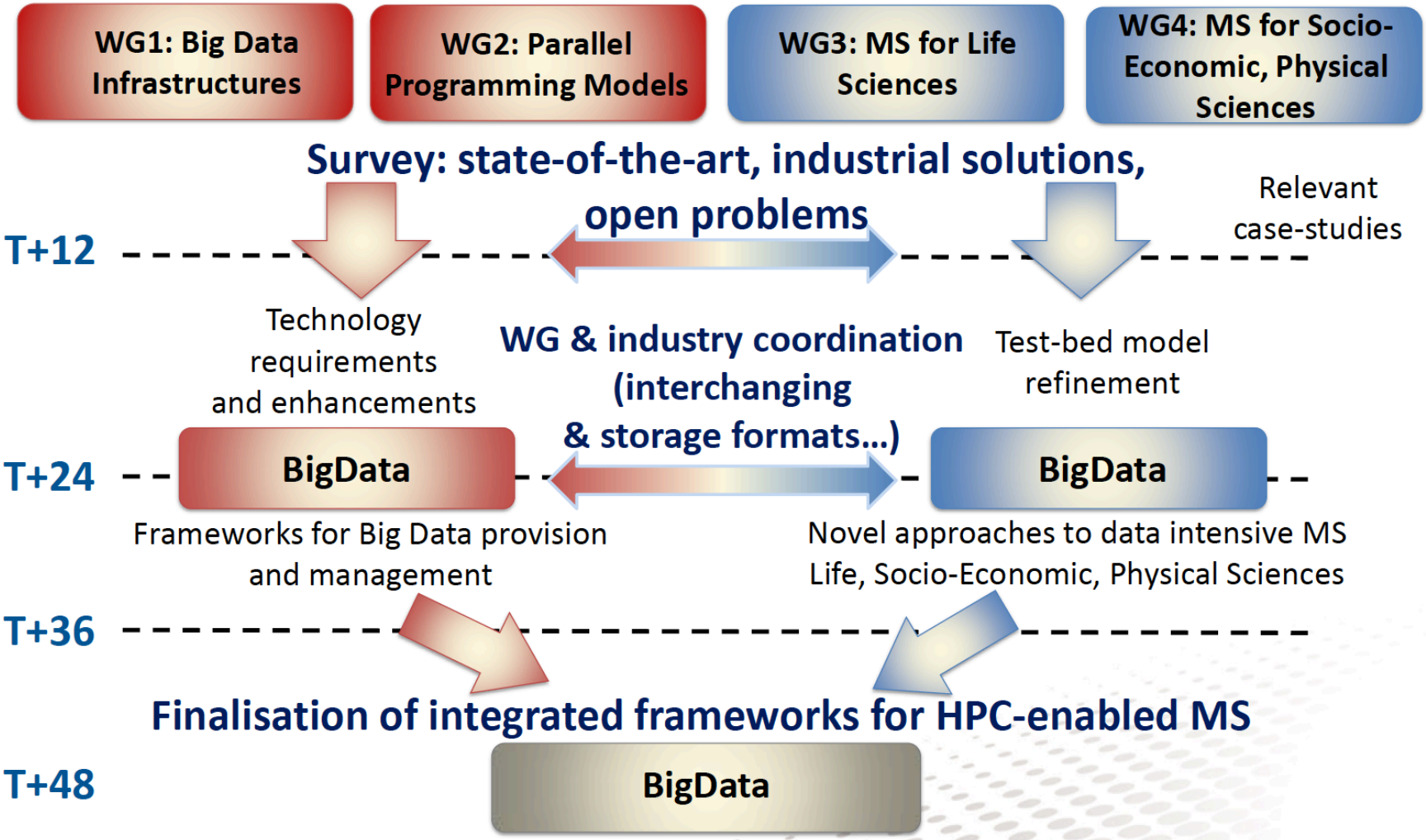
Domain-specific MS





HPC Support

Working Groups (WGs)



cHiPSet - Collaboration

CRC
CENTER FOR RESEARCH COMPUTING

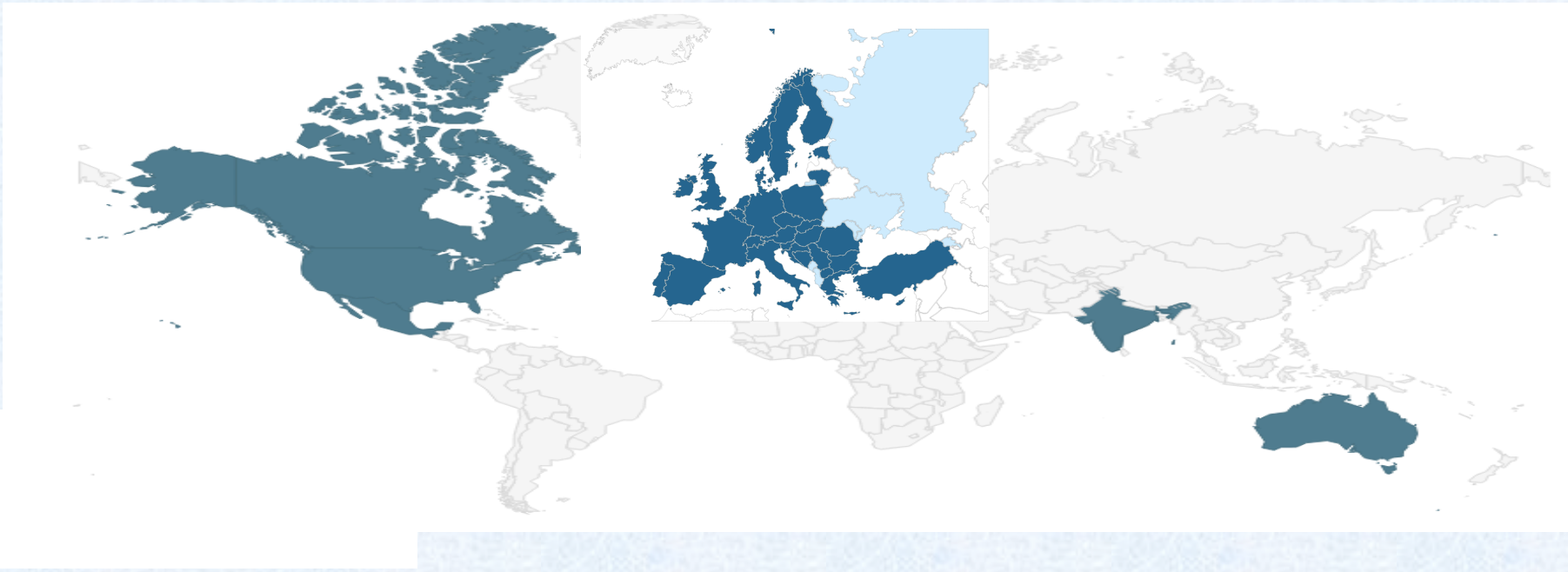
Projects already declared interest for collaboration

- NESUS (Network for Sustainable Ultrascale Computing)
- KEYSTONE (Semantic keyword-based search on structured data sources)
- AAPELE (Algorithms, Architectures and Platforms for Enhanced Living Environment)

And maybe YOU?

EU COST Action NESUS (IC1305)

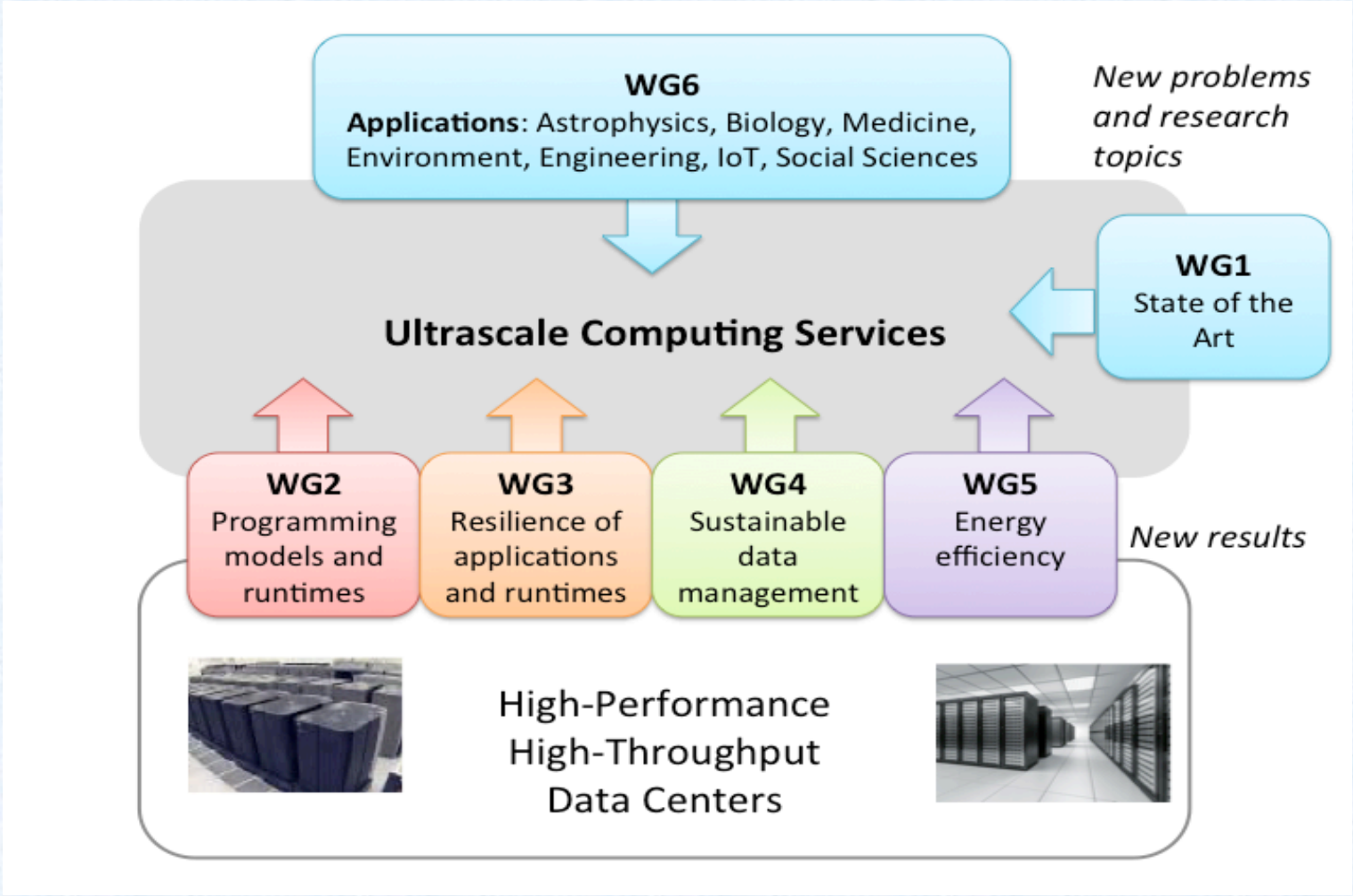
- Network for sustainable ultrascale computing systems
- April 2014 – April 2018
- 45 countries - 39 COST, 6 non-COST (US, China, Australia)
- 92 reseach organizations/companies
- More than 240 members



Challenges in UCS

Programming environment	New programming models: Hierarchical models Many-task models Global memory models Data distribution and locality Awareness of data-movement cost	Emergence of new algorithms Energy-efficient algorithms Fault-tolerant algorithms Application code migration and re-writing High-productivity methods
System Software	Increased system heterogeneity Capability for virtualization Standardized APIs Scalability, modularity, robustness	Explosion of data volumes Awareness of data-movement cost Extensive system monitoring Performance models
System architecture	Heterogeneous platform architectures Power consumption I/O latency and bandwidth Concurrency and data locality Storage capacity	Extreme scale from sub-component to total system Resiliency , Reliability, Availability, Serviceability (RAS) Software-defined systems

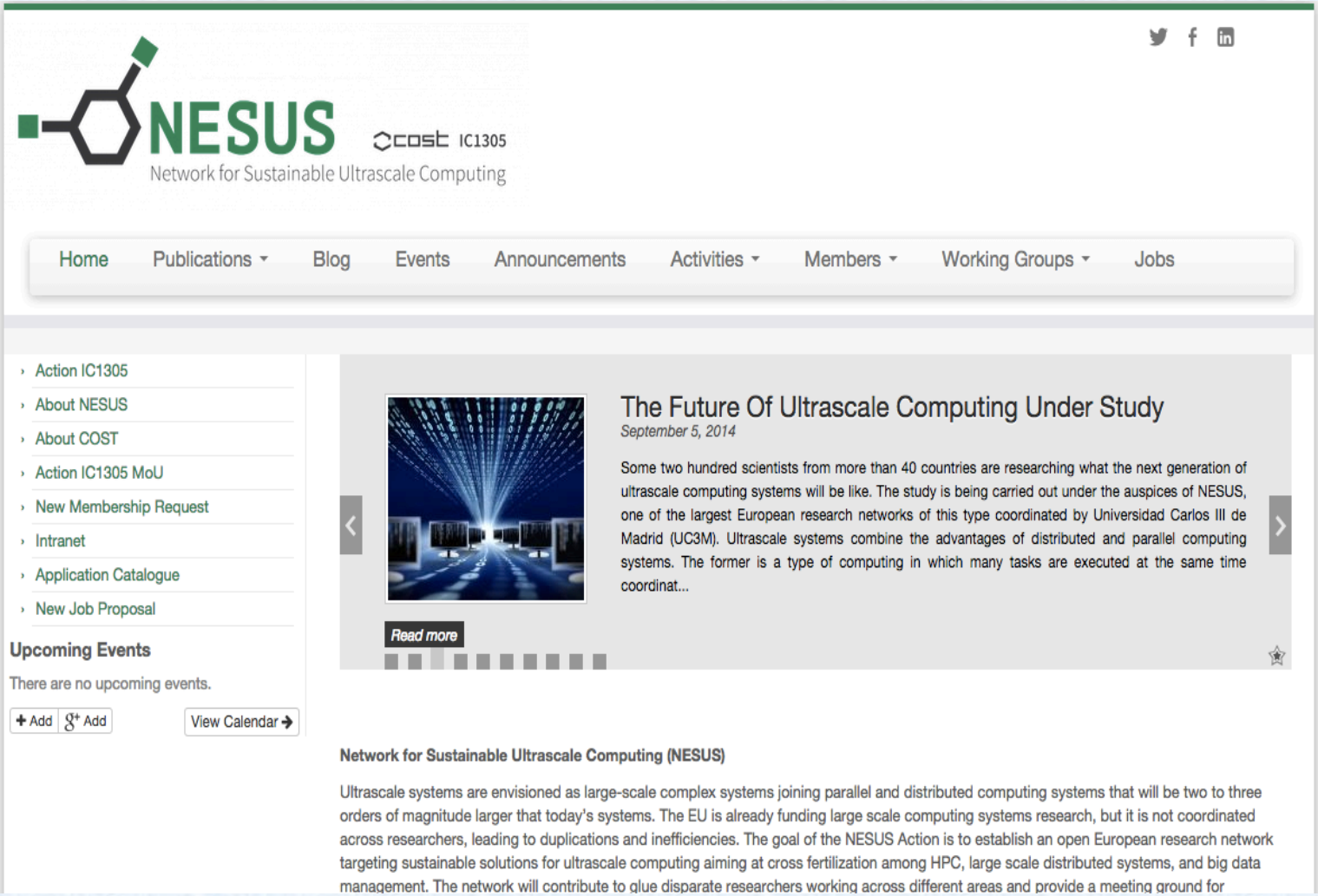
Workplan


WG6
Applications: Astrophysics, Biology, Medicine, Environment, Engineering, IoT, Social Sciences
New problems and research topics
WG1
State of the Art
Ultrascale Computing Services
WG2
Programming models and runtimes
WG3
Resilience of applications and runtimes
WG4
Sustainable data management
WG5
Energy efficiency
New results
High-Performance
High-Throughput
Data Centers

NESUS Activities

- Main activities
  - Research coordination
  - Working Group meetings
  - Research stays
  - Yearly workshop
  - Training school
  - PhD symposium
- Strong emphasis in cooperation
  - Join publications, tools, applications, …
  - With industry to solve real-world cases

NESUS Web portal (nesus.eu)



sandra.gesing@nd.edu
Jesus.carretero@uc3m.es