RETHINK big

Roadmap for European Technologies
in Hardware and Networking for Big Data

RETHINK big Project

Adrián Cristal

Barcelona Supercomputing Center

20th, March 2014

www.rethinkbig-project.eu

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 619788.

Mission

Mission

- To identify and evaluate the existing competencies across European Big Data Hardware and Networking technology sectors and application domains

Mission

- To identify and evaluate the existing competencies across European Big Data Hardware and Networking technology sectors and application domains

- To prioritize the complementary interests and the shared opportunities to unlock the highest return on their respective investments

Mission

- To identify and evaluate the existing competencies across European Big Data Hardware and Networking technology sectors and application domains

- To prioritize the complementary interests and the shared opportunities to unlock the highest return on their respective investments

- Resulting in a roadmap that would be irrational not to follow.

Why Hardware matters?

Why Hardware matters?

- **All Software runs on hardware….**
  - MapReduce
  - CUDA
  - OpenCL
  - MPI
  - OpenMP
  - …

Why Hardware matters?

- **All Software runs on hardware….**
  - MapReduce
  - CUDA
  - OpenCL
  - MPI
  - OpenMP
  - …
- **You might not want to admit it but…**
  - You are constrained by hardware and the network

- In the next 10 years, the HW will change more dramatically than it has in the past 10 years

- HW will influence the products and services that you provide

The world in 3D (3D Stacking)

- Very large bandwidth
- Very low latency
- Large amounts of memory on a chip
- Locality will be extremely important
- Thermal problems

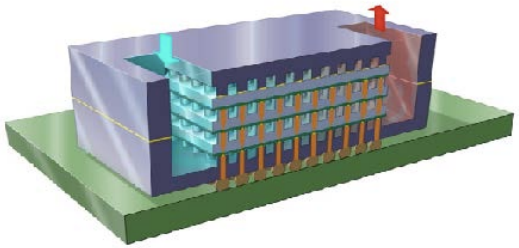


Figure from EPFL
http://esl.epfl.ch/page-58161-en.html

Non-volatile memory

- **New technologies (STT-RAM, CB-RAM, RRAM, …)**
- **More density**
- **Replacement for DRAMs**
  - Endurance problem
- **Large influence on software**
  - Data base systems
  - File systems

Dark Silicon Era

- **Thermal problems**
  - Not all cores will be able to be on all the time
- **Extensive use of Accelerators**
- **Reconfigurable computing**

What happens if HW does not consider SW

- **Many changes in HW architecture do not survive**
  - Cell processor (Playstation 3 processor)
  - Itanium processor

What happens if SW does not consider HW

- **Terasort contest: sorting 100TB data**
  - Number 1: Hadoop
    - 2100 nodes, 12 cores per node, 64 Gb per node
      - 24.000 cores
      - 134 Tb memory
    - Time: 4300 segs
  - Number 2: Tritonsort
    - 52 nodes, 8 cores per node, 24 Gb
      - 416 cores
      - 1,2 Tb memory
    - Time: 8300 segs and 6400 segs
  - Hadoop is easy to program, but needs 57X more cores, 100X more memory, and only gets 2X performance

New Technologies

- **Neural Networks**
  - Analog Memristor-based
- **Graphene transistors**
- **Quantum Computing**
- **DNA computing**
- **…**

Challenges

Challenges

- **Work with different areas**
  - Applications and end users
  - Software Tools
  - Systems
  - Network
  - Hardware

Challenges

- **Work with different areas**
  - Applications and end users
  - Software Tools
  - Systems
  - Network
  - Hardware
- **Work with different requirements**
  - Speed
  - Volume
  - Real Time
  - Sensors
  - Variability
  - Power consumption

Building an Ecosystem

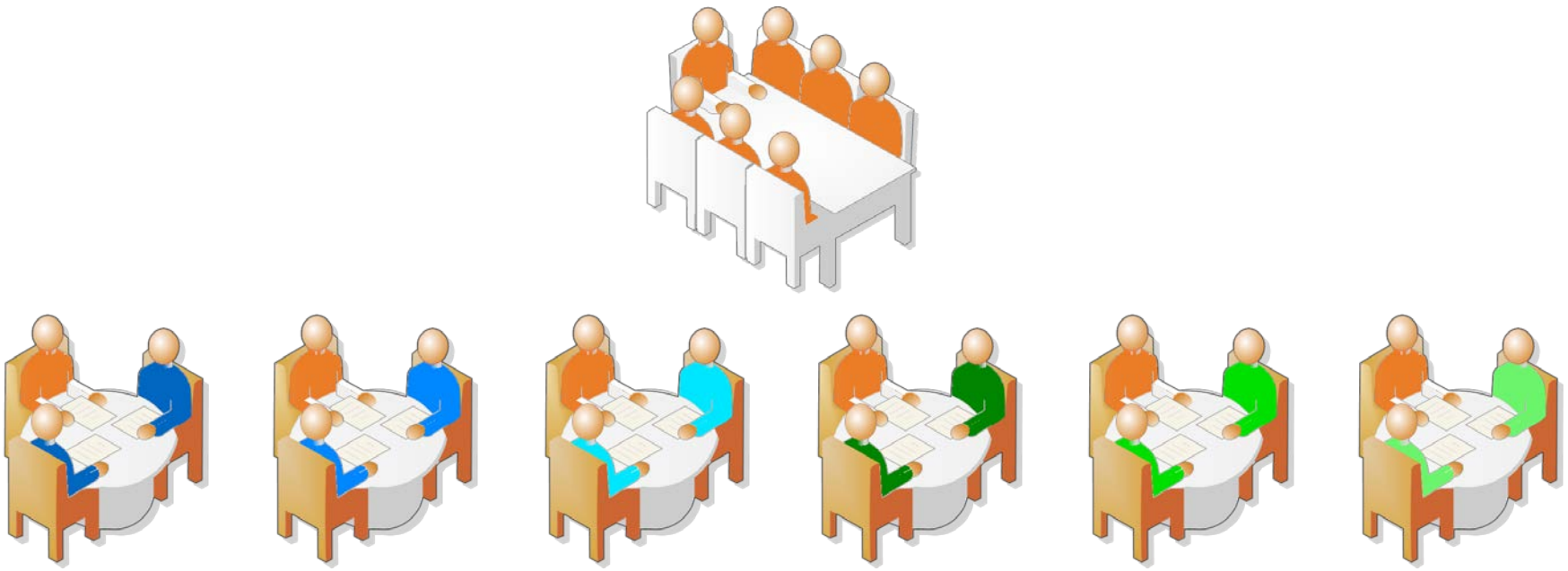
Building an Ecosystem

Consortium

Building an Ecosystem

Working group
Application Challeges

Building an Ecosystem

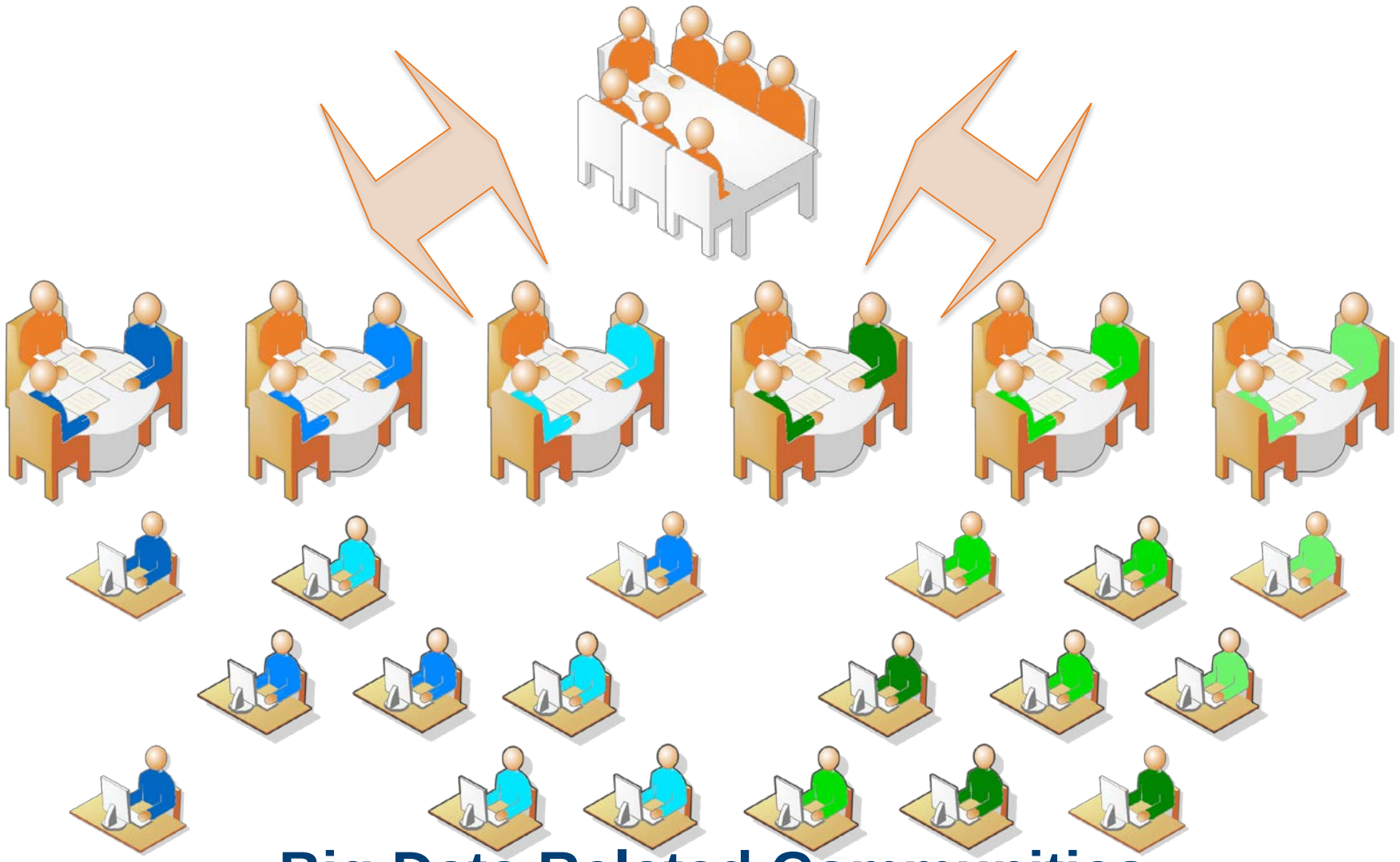


Working group
Enabling Technologies

Building an Ecosystem

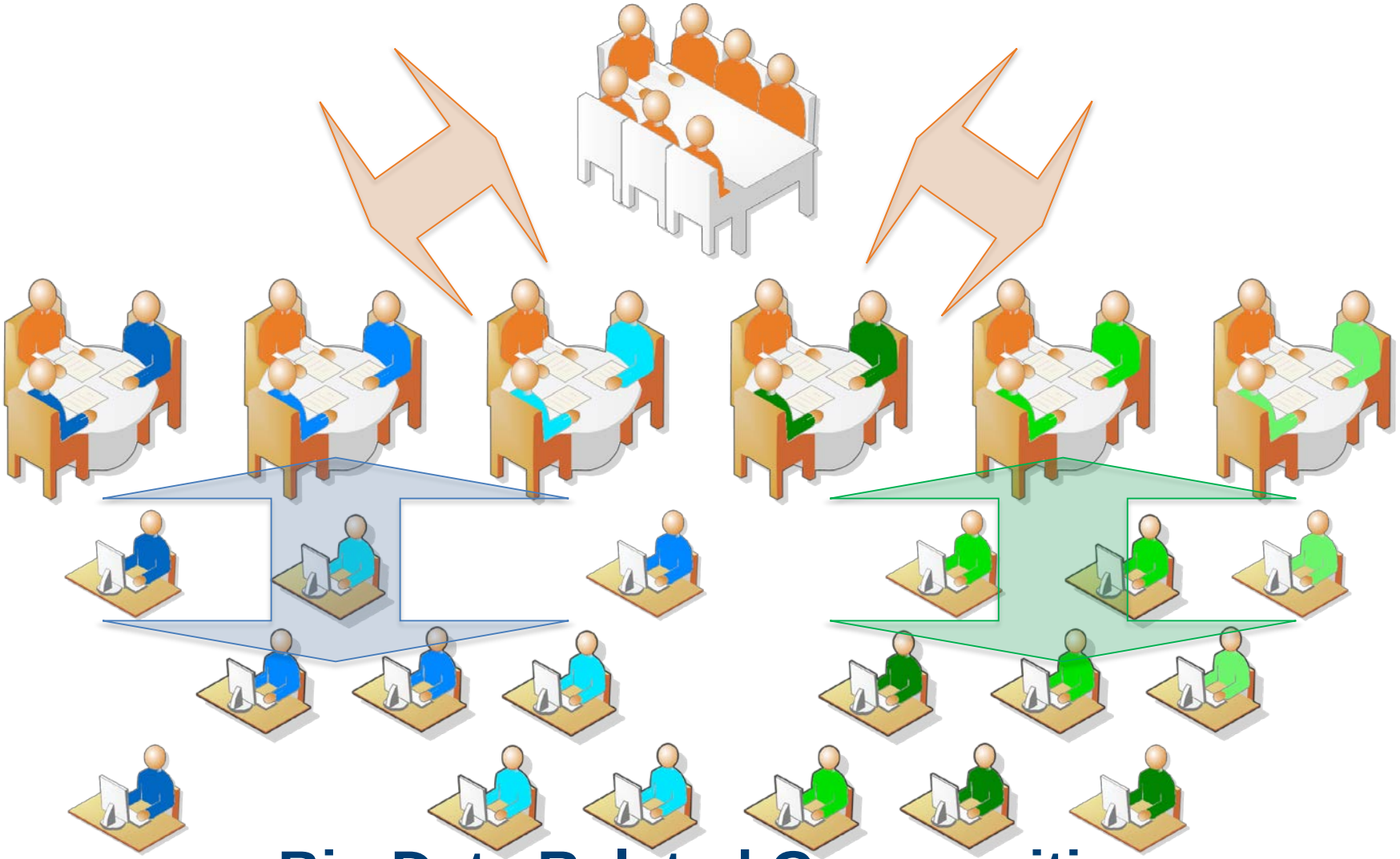
Big Data Related Communities

Building an Ecosystem



Big Data Related Communities

Building an Ecosystem



Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Building an Ecosystem

Big Data Related Communities

Ways to Participate

- **Web**
  - http://www.RETHINKbig-project.eu
- **Linkedin**
  - https://www.linkedin.com/groups/RETHINKbig-7457953
- **Twitter**
  - https://twitter.com/RETHINKbig

Thank you

Stop by our stand

http://www.RETHINKbig-project.eu