JÜLICH FORSCHUNGSZENTRUM

RDA RESEARCH DATA ALLIANCE

UNIVERSITY OF ICELAND

High Productivity Processing Engaging in Big Data around Distributed Computing

Dr. – Ing. Morris Riedel et al.

Deputy Division Leader Federated Systems & Data, Juelich Supercomputing Centre, Germany

Adjunct Associate Professor, University of Iceland, Iceland

Co-Chair Research Data Alliance, Big Data Analytics Group

research data sharing without barriers

rd-alliance.org

mipro 2013

36th International Convention on Information and Communication Technology, Electronics and Microelectronics

Monday 20th - Friday 24th of May 2013 - Opatija/Abbazia, Adriatic Coast, Croatia

Grand Hotel Adriatic Convention Centre and Hotel Admiral

Outline

Terms & Motivations

Selected ‚Big Data Analytics‘

Classification & Hardware impacts

Summary & References

Big Data Waves – Surfboards – Breakwaters

How can we manage the rising tide of scientific data

High Level Expert Group on Scientific Data Report

- Lists unsolved questions
- Outlines challenges
- Provides visions

A Surfboard for Riding The Wave Report

- Lists 4 key action drivers
- Identifies 3 strategic goals
- Clarifies Data Scientists

' Concrete' Next Steps →

In commercial environments

Big Data is all about

Volume – Variety – Velocity

'Big Data is data that becomes large enough that it cannot be processed using <u>conventional methods</u>.'

[1] O'Reilly Radar Team, 'Big Data Now: Current Perspectives from O'Reilly Radar'

The next generation radio telescope for science…

The square kilometre array

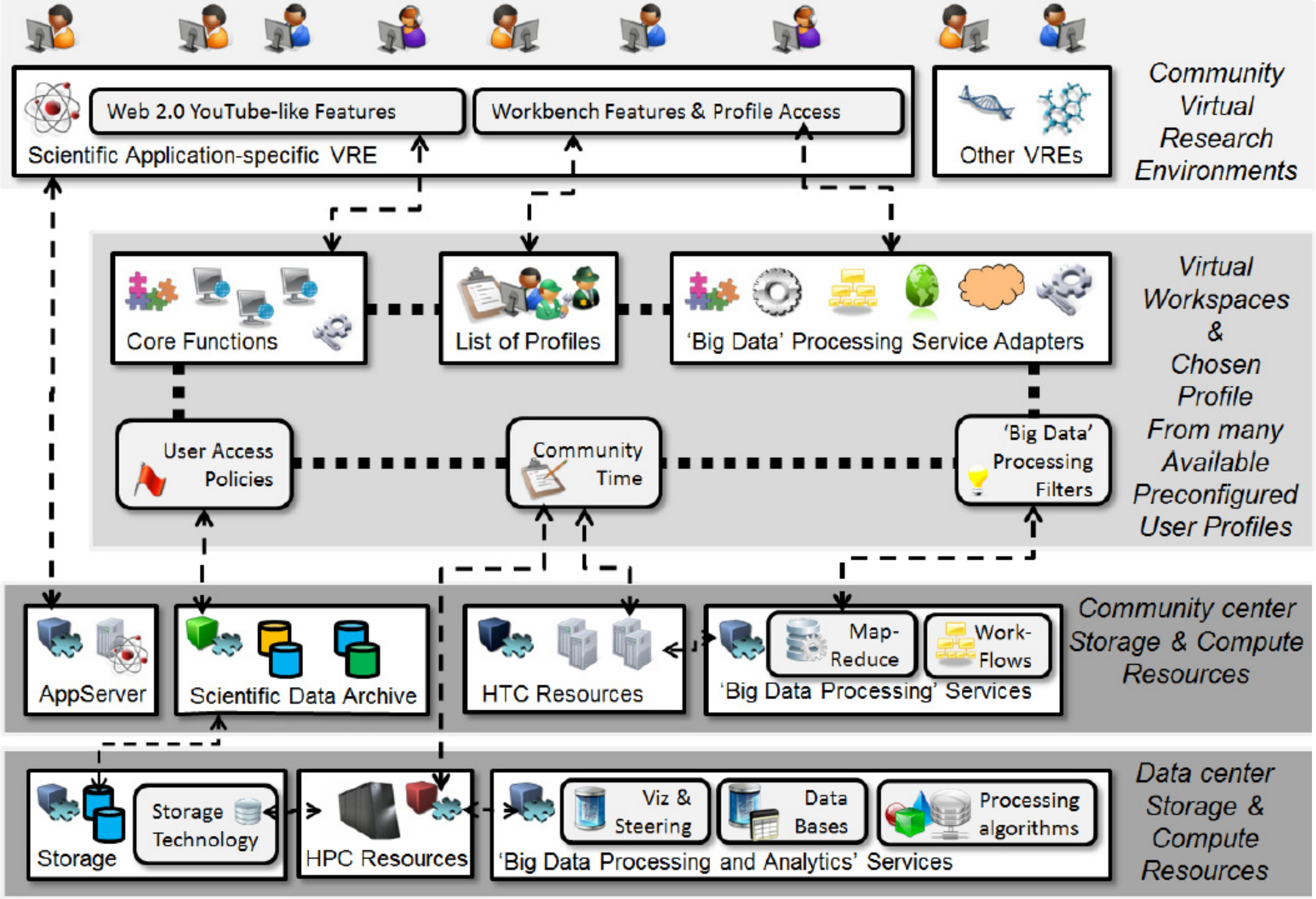
… **1 PB** in 20 seconds

LOFAR
test site Jülich

xilostudios.com

Collaborative Data Infrastructure EUDAT

Providing the foundations to engage in data-driven research



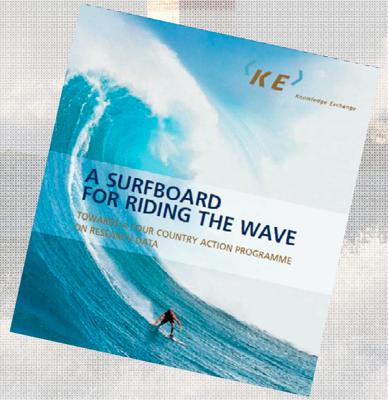
[5] M. Riedel and P. Wittenburg et al. 'A Data Infrastructure Reference Model with Applications: Towards Realization of a ScienceTube Vision with a Data Replication Service'

Analytics are Needed in Big Data-driven Scientific Research

The challenge is to understand which analytics make sense

'Understanding climate change, finding alternative energy sources, and preserving the health of an ageing population are all cross-disciplinary problems that require high-performance data storage, **smart analytics**, transmission and mining to solve.'

'In the data-intensive scientific world, **new skills are needed for** creating, handling, **manipulating, analysing,** and making available large amounts of data for re-use by others.'



Analytics are techniques to work on large data…

Data Analysis is the interpretation of research data

Shifts from Causality to Correlation

Challenging research with progress based on reason?

Traditional search for causality

Describing exactly WHY something is happening
Understanding causality is hard and time-consuming
Searching it often leads us down the wrong paths

Big Data Analytics

Not focussed on causality – enough THAT it is happening
Discover novel patterns and WHAT is happening
Using correlations for invaluable insights – data speaks for itself

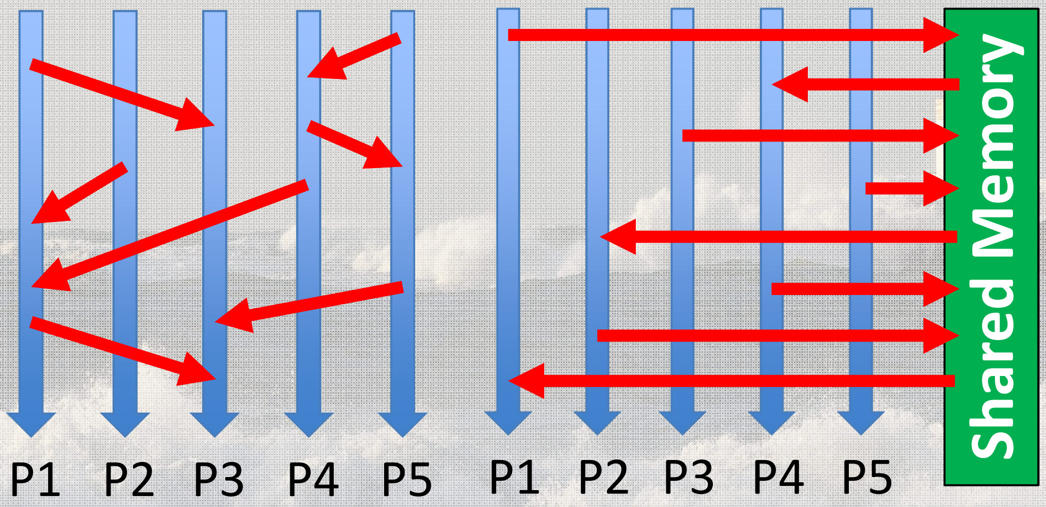
A complementary & alternative approach to scientific problems

Complexities in Conventional Methods

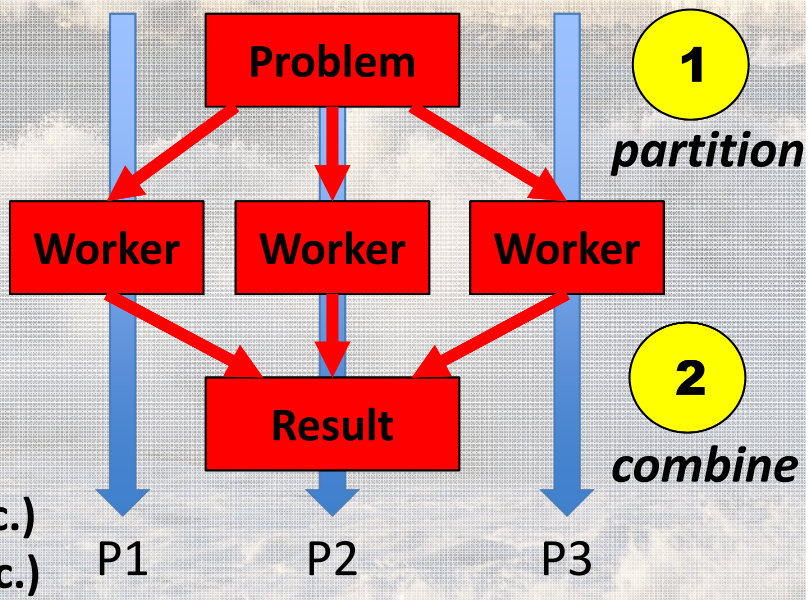
Traditional Programming Models

Distributed Architecture Issues

Message Passing **Shared Memory**



Divide & Conquer



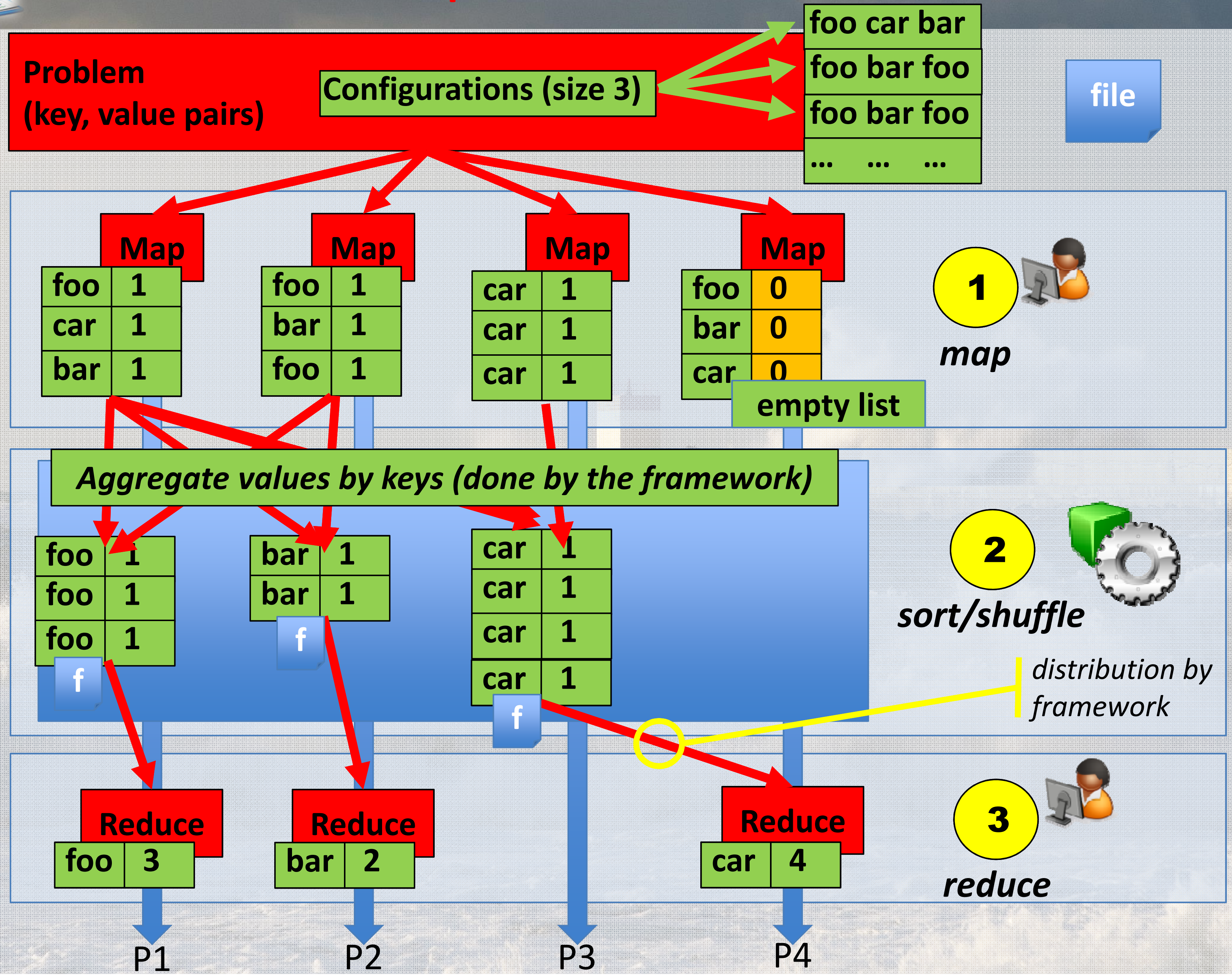
Increasing Complexities:

Different Programming elements (barriers, mutexes, etc.)

Distribution issues (scheduling, synchronization, IPC, etc.)

Architecture issues (UMA, NUMA, SIMD, MIMD, etc.)

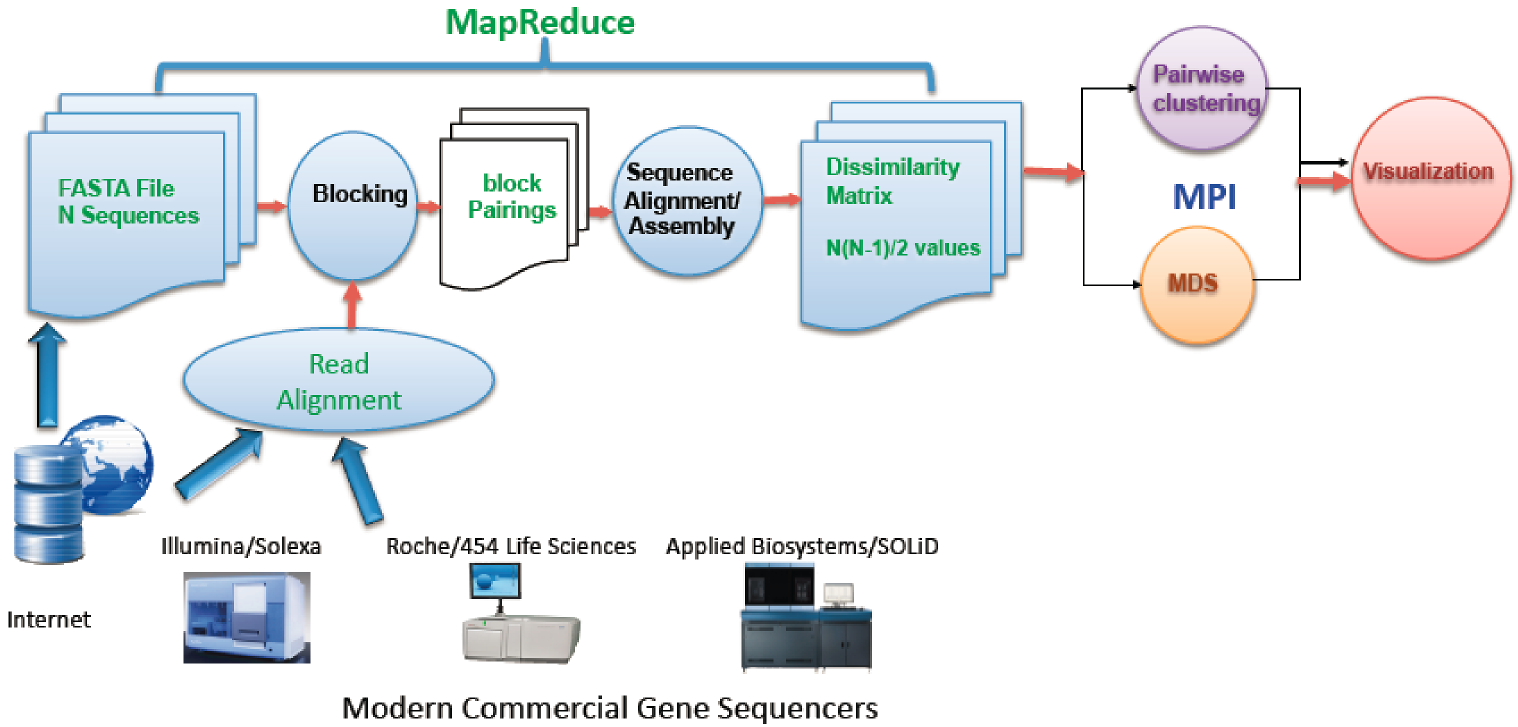
Classic Map-Reduce



Classic Map-Reduce Example with Limits

Typical Application Challenge:

DNA Sequencing Pipeline



Linear Algebra or Expectation Maximization based data mining poor on MapReduce – equivalent to using MPI writing messages to disk and restarting processes each step/iteration of algorithm

G. Fox et al.

Summary ‚Classic Map-Reduce‘

Map-Reduce

Classic Map-Reduce

Lousely-Coupled Commun-ication

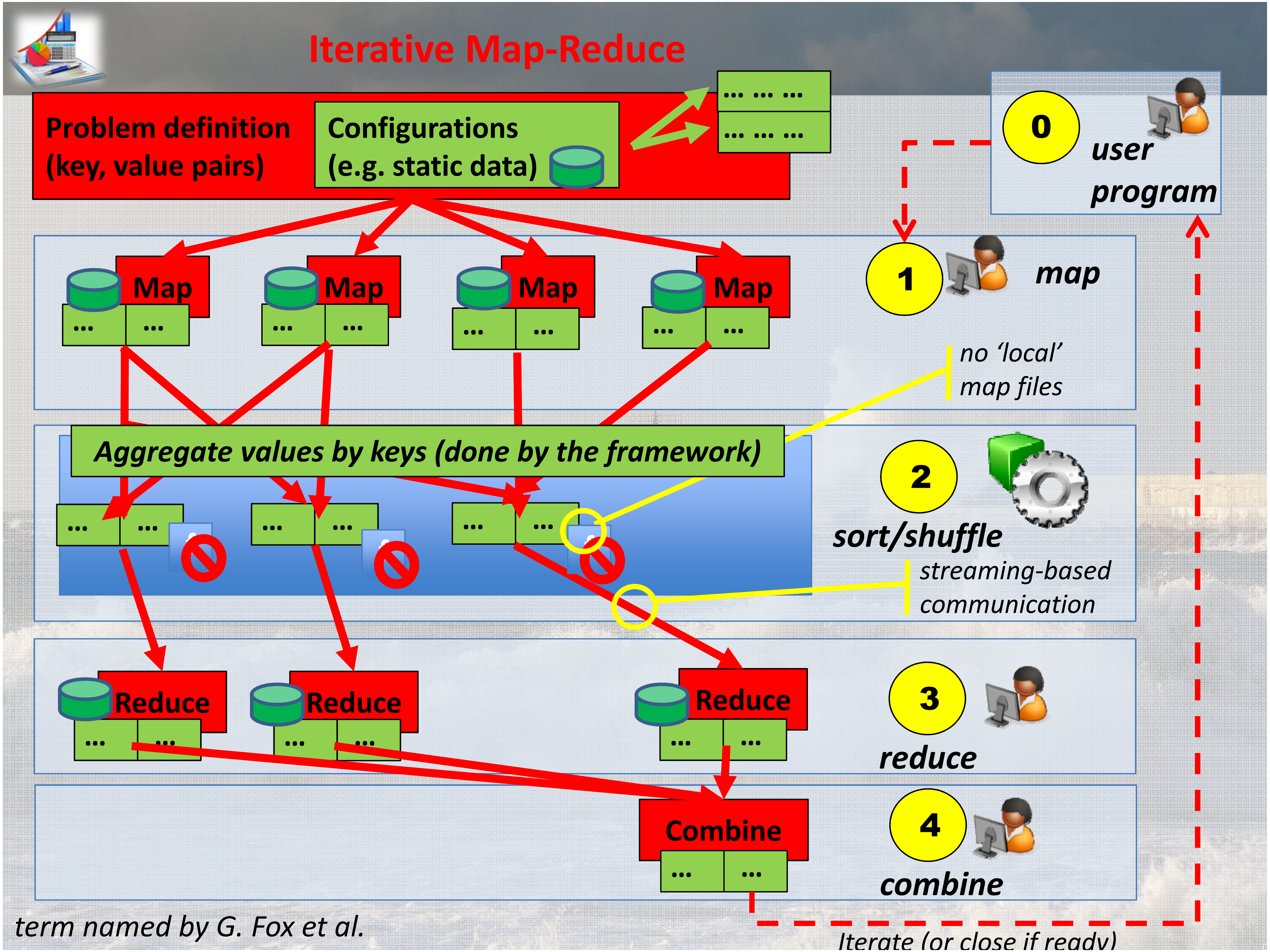
BLAST, Matlab Parameter Sweeps, Ensemble Runs, Distributed Search

Mostly HTC, Apps

Classic Map – Reduce is not for all problems:

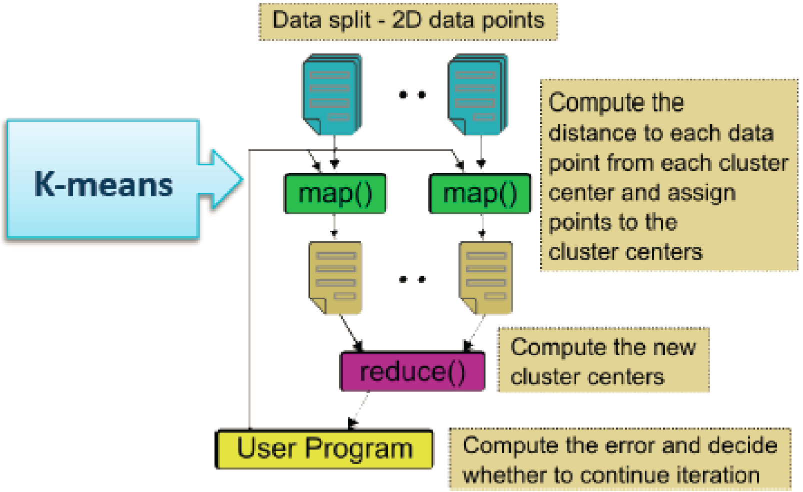
Frameworks take care of sorting/shuffling & distribution
Slow communication (like MPI snd/rcv via files write/read)
Programming is relatively simple, but also needs thinking

Apache Hadoop – Java
Dryad – Windows


Iterative Map-Reduce
Problem definition
(key, value pairs)
Configurations
(e.g. static data)
... ...
... ...
0
user
program
Map
Map
Map
Map
1
map
no 'local'
map files
Aggregate values by keys (done by the framework)
2
sort/shuffle
streaming-based
communication
Reduce
Reduce
Reduce
3
reduce
Combine
4
combine
term named by G. Fox et al.
Iterate (or close if ready)

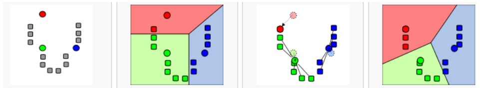
Iterative Map-Reduce Example

Iterative and non-Iterative Computations



K-Means Clustering (NP-hard)

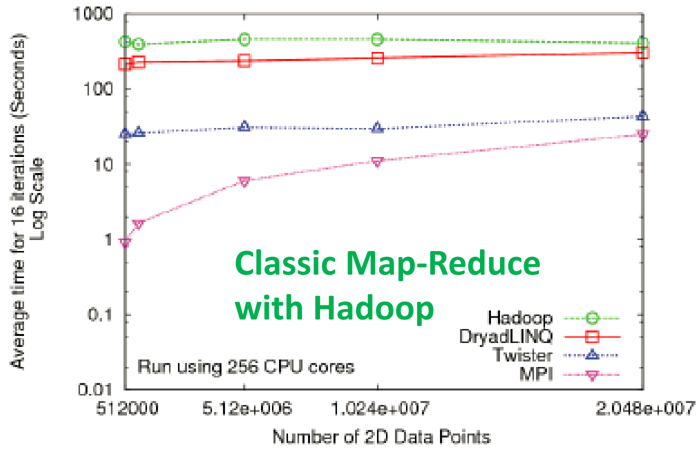
Partition n observations into n clusters → Voronoi cells

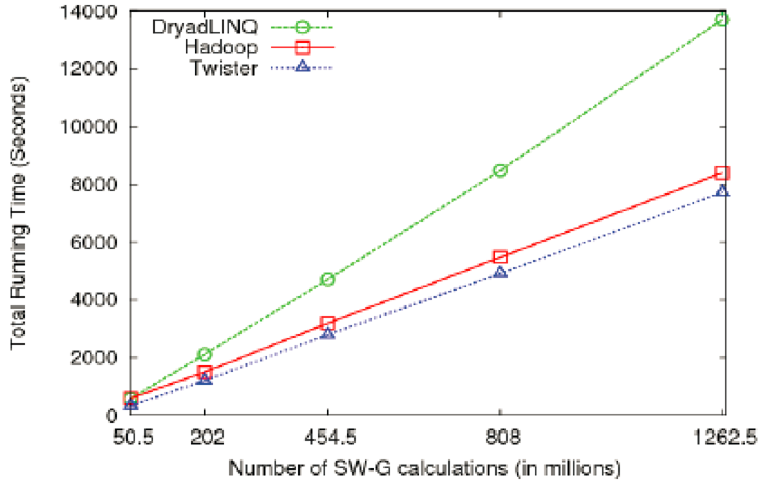


www.wikipedia.org

Smith Waterman is a non iterative case and of course runs fine → used in bio-informatics

Performance of K-Means





(modified from G. Fox et al.)

Summary ‚Iterative Map-Reduce'

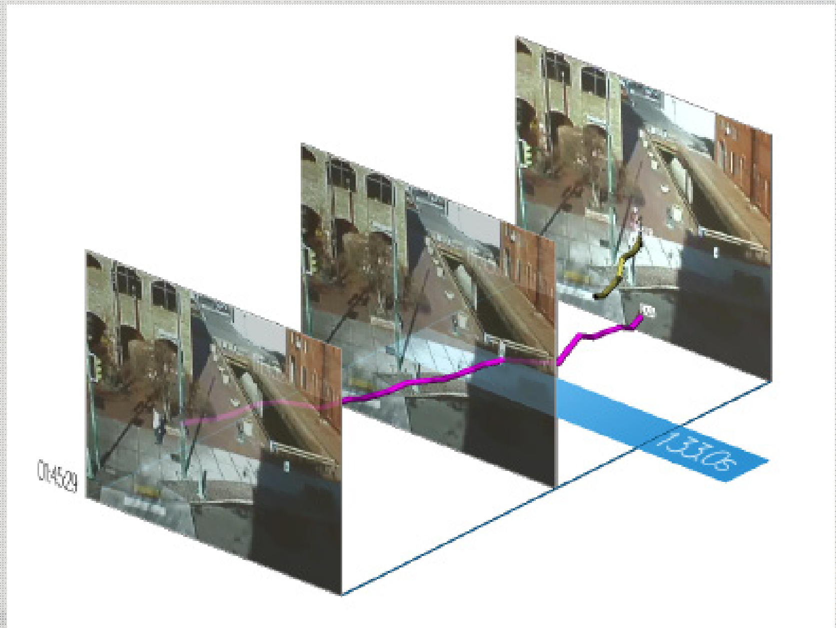
Map-Reduce	
Classic Map-Reduce	**Iterative Map-Reduce**
Lousely-Coupled Commun-ication	Iterative loosely coupled, Pub-Sub Commun-ication
BLAST, Matlab Parameter Sweeps, Ensemble Runs, Distributed Search	Linear-algebra, Step-wise algorithms and iterative scientific problems, Page rank, K-Means
Mostly HTC, Apps	HTC towards HPC, Apps

Iterative Map – Reduce closer to scientific computing

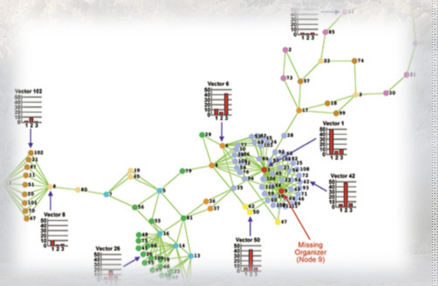
Enables scientific programs that are iterative in nature
Communication improved: in-memory, no ‚local' map files
Programming model getting more complex (‚tunings')

Twister (Map-Reduce++)
Dryad Language Integrated Query (DryadLINQ)

Visual Analytics



http://www.visual-analytics.eu/



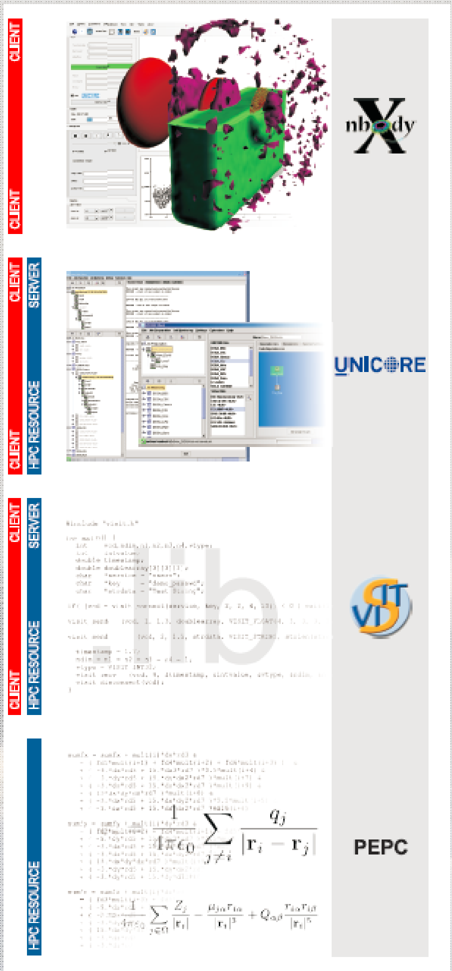
Putting the human in the analysis loop for analytical reasoning

Based on visual inter-linked data, interactive approaches, and interfaces

Visual representation of analytical reasoning and data transformations

Enable conclusions through a combination of evidence and assumptions

Visual Analytics – Example



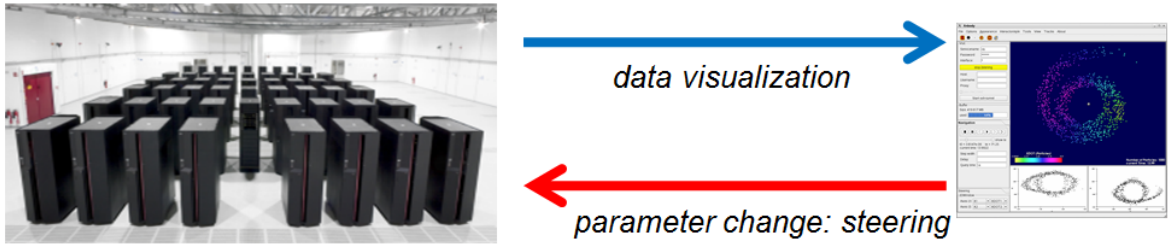
Collaborative Online Visualization and Steering (COVS)

Enables incorporation of human judgements from distributed teams

Provides mechanisms to influence HPC simulation during run-time

Topics: .e.g. extend scalability, combine with iterative map-reduce

- Save compute time by focusing on special areas
  - *E.g. Parameter space exploration of an application*
- Switch 'options' in a parallel HPC application during runtime
  - *E.g. Heating-Ventilating-AirConditioning (HVAC) simulations in engineering, Computational Fluid Dynamics (CFD),…*



M. Riedel, Th. Eickermann, S. Habbinga, W. Frings, P. Gibbon et al., ***Computational Steering and Online Visualization of Scientific Applications on Large-Scale HPC Systems within e-Science Infrastructures***

Summary ‚Visual Analytics'

Visual Analytics	
Online/real-time Visualization	Computational Steering
Communication from data generator to visualizer	Communication from visualizer to steered process
Data streaming applications for thousands of data elements, interlinked data mesh	Iterative problems and step-wise approaches, nbody simulations, CFD codes
HTC and HPC, viz cluster, Apps	HTC, rather HPC, Apps, BGAS

Visual Analytics relies on new and old approaches

Enables view on research data (interlinked and interactive)

Provides mechanisms to filter/reduce big data streams

Well suited for iterative scientific applications

VISIT Toolkit (UoBerkeley)

COVS Framework (JSC)

many others…

Context: ‚Crowd Sourcing‘

Getting a ‚crowd of people‘ to help with data gathering not possible before

Allows experimentalists and observers to collect huge amount of data from citizens

Often experiments and observation studies talk about 10-100 of subjects, soon millions

Requires the methodology of scientific fields having much more ‚balanced subject data‘

Change of mass and inherent complexity of the data that need to be processed, stored, managed, and analyzed to extract salient patterns proven by large statistics

Summary ‚Crowd Sourcing'

Extreme Data Sources
Crowd Sourcing
Massive amount of parallel commun-ication streams
Data gatherings, Cor-relations, ranking, community reviews, localized data
Apps, HTC, DDN Web Scaler

Crowd sourcing offers insights – data speaks for itself

Enables statistics across a massive amount of participants

Often bottom-up and only slightly coordinated → realistic?

Typically includes ‚creation context' data (e.g. location)

Apps of mobile devices

Social network plugins

Context: ‚Fast Databases‘

Key Benefits of NoSQL DBs

Easy to deploy, implement
Relatively cheap to operate
Easy to geographically distribute
Designed with ‚no schemas‘
Scalability inherent in the DBs
Quickly process extremely large datasets
Low data consistency requirements

NoSQL DBs can handle Web Scale Data - whereas Relational DBs can not

www.fatcloud.com

Context: ‚Fast Data(base) Access‘

	server	core	mem[GB]	disk[TB]	Count
Tier 1	2950	8	16	22.50	40
Tier 2	R900	16	64	33.75	4
Tier 3	R900	16	128	11.25	2
total		**416**	**1152**	**1057.50**	**46**

A. Szalay et al., ‘GrayWulf: Scalable Clustered Architecture for Data Intensive Computing’

Hardware impact for fast database access must be taken into account, combination of DB methods useful in some cases

SQL for structured information (Oracle, etc.) – read fast through use of index approaches

NoSQL for unstructured information (MongoDB, etc.) – read/write fast through less validation

General community trend to tackle ‘big data’ towards ‘SMAQ stack’

- Storage MapReduce and Query (SMAQ)→High Throughput Computing (HTC)

Query (new forms of insights derived from powerful queries)

MapReduce (distributes computation over many servers)

Storage (distributed, non-relational or unstructured)

…but is ‘SMAQ’ indeed the full answer for ‘data-intensive science’ ?

Summary ‚Fast Data(base) Access‘

Fast Data Base Access
NoSQL Databases
In-memory access & commun-cation
Keeping data and un-structured information for quick processing and storage
Un-structured DBs, ‘In-memory’

Crowd sourcing offers insights – data speaks for itself

Enables statistics across a massive amount of participants

Often bottom-up and only slightly coordinated → realistic?

Typically includes ‚creation context‘ data (e.g. location)

NoSQL Databases (e.g. MongoDB)

‘In-memory’ databases

Classification Summary → building blocks to understanding

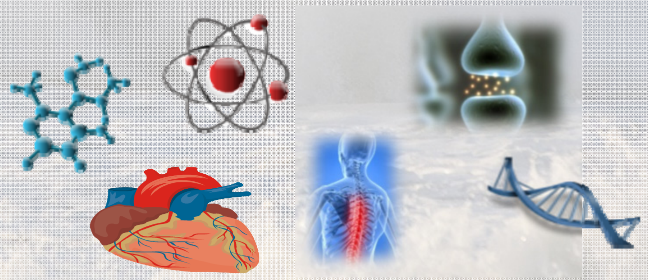
Map-Reduce		Visual Analytics		Algorithms for Large-scale Data Analysis	Extreme Data Sources	Fast Data Base Access
Classic Map-Reduce	Iterative Map-Reduce	Online/real-time Visualization	Computational Steering	Parallel algorithms, libraries, tools	Crowd Sourcing	NoSQL Databases
Loosely-Coupled Communication	Iterative loosely coupled, Pub-Sub Commun-ication	Commun-ication from data generator to visualizer	Commun-ication from visualizer to steered process	Massively parallel communication with synchronization, communicators, shared memory programming	Massive amount of parallel commun-ication streams	In-memory access & commun-cation
BLAST, Matlab Parameter Sweeps, Ensemble Runs, Distributed Search& Sorting	Linear-algebra, Step-wise algorithms and iterative scientific problems, Page rank	Data streaming applications for thousands of data elements, interlinked data mesh	Iterative problems and step-wise approaches, nbody simulations, CFD codes	MPI-programs, openmp, FFT algorithms, PDE solvers, particle dynamics, MD codes Reliability studies Using new hardware features such as virtualized networks	Data gatherings, Cor-relations, ranking, community reviews, localized data	Keeping data and un-structured information for quick processing and storage
Mostly HTC, Apps	HTC towards HPC, Apps	HTC and HPC, viz cluster, Apps *combination*	HTC, rather HPC, Apps, BGAS	HPC, JUROPA3, DDN, GPGPUs, small clusters, etc.	Apps, HTC, DDN Web Scaler	Un-structured DBs, 'In-memory'

Needs more in-depth studies with research data:

More granularity in categories and combinations

Accurate performance measurements & application DB

Impacts on energy-efficiency and programming efforts



Overall Summary

A better understanding of ‚analytics for research data' is still needed

Classic Map-Reduce only suited for rather embarassingly parallel computing jobs (i.e. HTC)

Iterative Map-Reduce less performance than tradional parallel computing (i.e. MPI/OpenMP)

Visual Analytics combine human judgements with automatic analysis/filtering

New innovative ‚crowd sourcing techniques' harness the power of the masses for statistics

NoSQL and in-memory DB approaches support the analytics process for Big Data (Streams)

What we can achieve by combining them in research?

Initial classification represents the ‚lighthouse' of the RDA group ‚Big Data Analytics' to perform systematic research

References

[1] V. Mayer-Schoenberger and Kenneth Cukier,‘Big Data – A Revolution That Will Transform How We Live, Work and Think‘, Book, ISBN 978-1-84854-791-9, John Murray Publisher, 2013

[2] G. Fox, ‘MPI and Map-Reduce‘, Talk at CCGSC 2010 Flat Rock,NC, 2010

[3] J. Wood et al., ‘Riding the Wave – How Europe can gain from the rising tide of scientific data’, report to the European Commission, 2010

[4] Knowledge Exchange Partner, ‘A Surfboard for Riding the Wave – Towards a Four Action Country Programme on Research Data’, 2011

[5] M. Riedel and P. Wittenburg et al. ‘A Data Infrastructure Reference Model with Applications: Towards Realization of a ScienceTube Vision with a Data Replication Service’, Internet Journal

Thanks for your attention

Engage in RDA with your experience!

http://rd-alliance.org/

Group Mailing List: http://lists.lists.rd-alliance.org/mailman/listinfo/rda-cwg-bda