RESEARCH DATA ALLIANCE

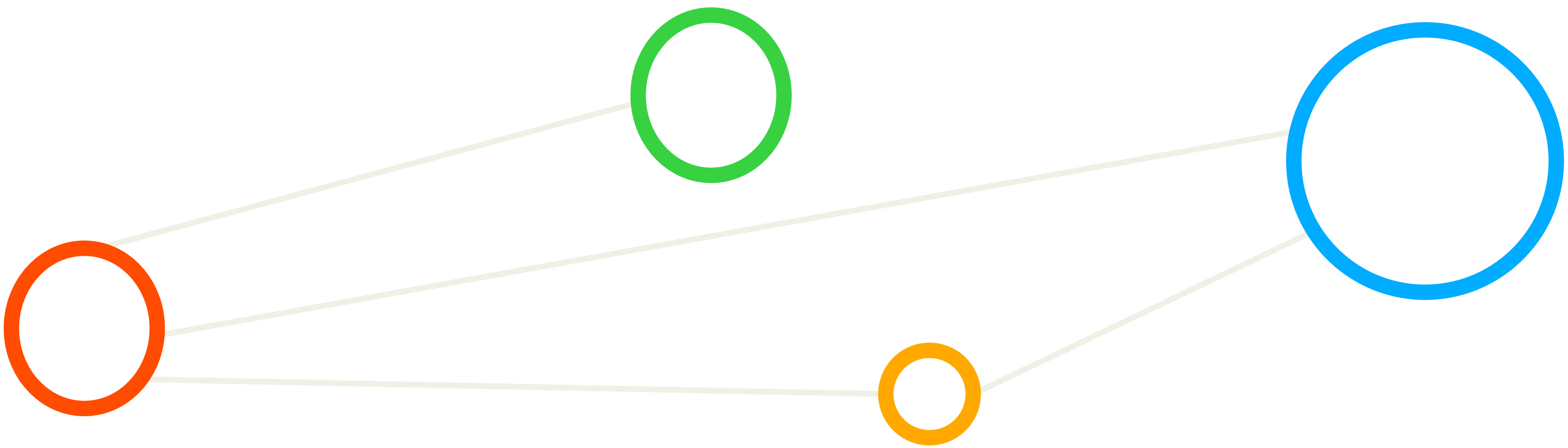
Understanding Big Data Analytics in the context of Scientific Computing

Morris Riedel et al., Juelich Supercomputing Centre, Germany
Co-Chair RDA Big Data Analytics Group

research data sharing without barriers
rd-alliance.org

2nd Annual CHANGES Workshop, 10. – 12. September, Chicago

Outline



RESEARCH DATA ALLIANCE

Research Data Sharing without barriers

[3] RDA Web Page

- Agricultural Data Interoperability IG
- **Big Data Analytics IG**
- Brokering IG
- Certification of Digital Repositories IG
- Community Capability Model WG
- Data Citation WG
- Data Foundation and Terminology WG
- Data in Context IG
- Data Type Registries WG
- Defining Urban Data Exchange for Science IG
- Digital Practices in History and Ethnography IG
- Engagement Group IG
- Legal Interoperability IG
- Long tail of research data IG
- Marine Data Harmonization IG
- Metadata IG
- Metadata Standards Directory WG
- PID Information Types WG
- Practical Policy WG
- Preservation e-Infrastructure IG
- Publishing Data IG
- Standardization of Data Categories and Codes IG
- Structural Biology IG
- Toxicogenomics Interoperability IG
- UPC Code for Data IG
- Wheat Data Interoperability WG

The New York Times

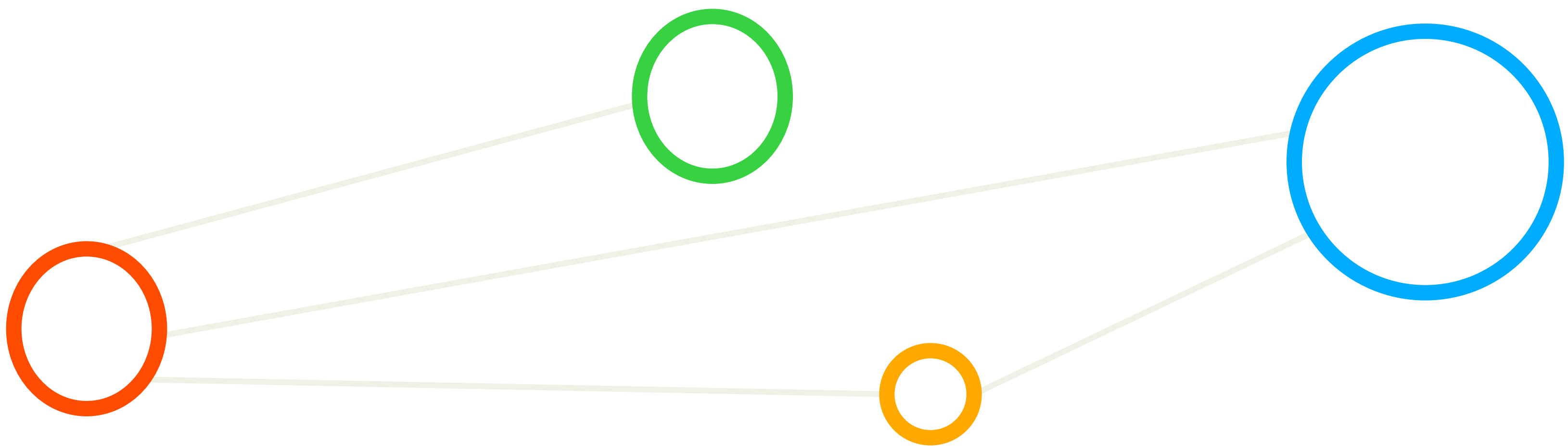
[10] 'How to share scientific data' Google & RDA New York Times Article

- ## Big Data Analytics IG
  - <u>Develops community based recommendations</u> on feasible data analytics approaches to address scientific community needs of utilizing large quantities of data.
  - <u>Analyzes different scientific domain applications</u> and their potential use of various big data analytics techniques.
  - <u>A systematic classification of feasible combinations</u> of analysis algorithms, analytical tools, data and resource characteristics and scientific queries will be covered in these recommendations.

Outline

- Motivation & Terms
  - Understand the focus of work
  - Discuss terms like Big Data, Data Analysis/Analytics, Causality/Correlation
- Shaping a 'Big Data Analytics landscape'
  - Analyze Scientific Application Use Cases
  - Investigate the different applied approaches & tools
  - (Closing the loop to long-term achieving & open data from talk yesterday)
  - Methodology of the group
- Summary
  - Concluding remarks
- References

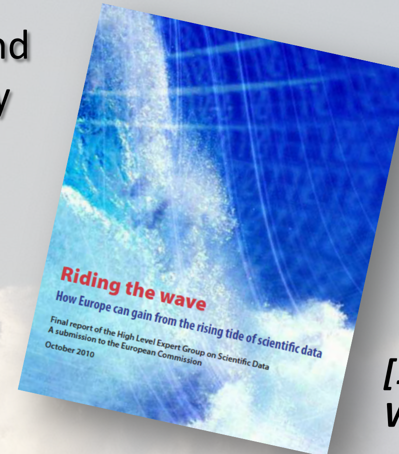
Motivation & Terms



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.'



[1] 'Riding the Wave' Report

'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.'

[2] 'A Surfboard for Riding the Wave' Report

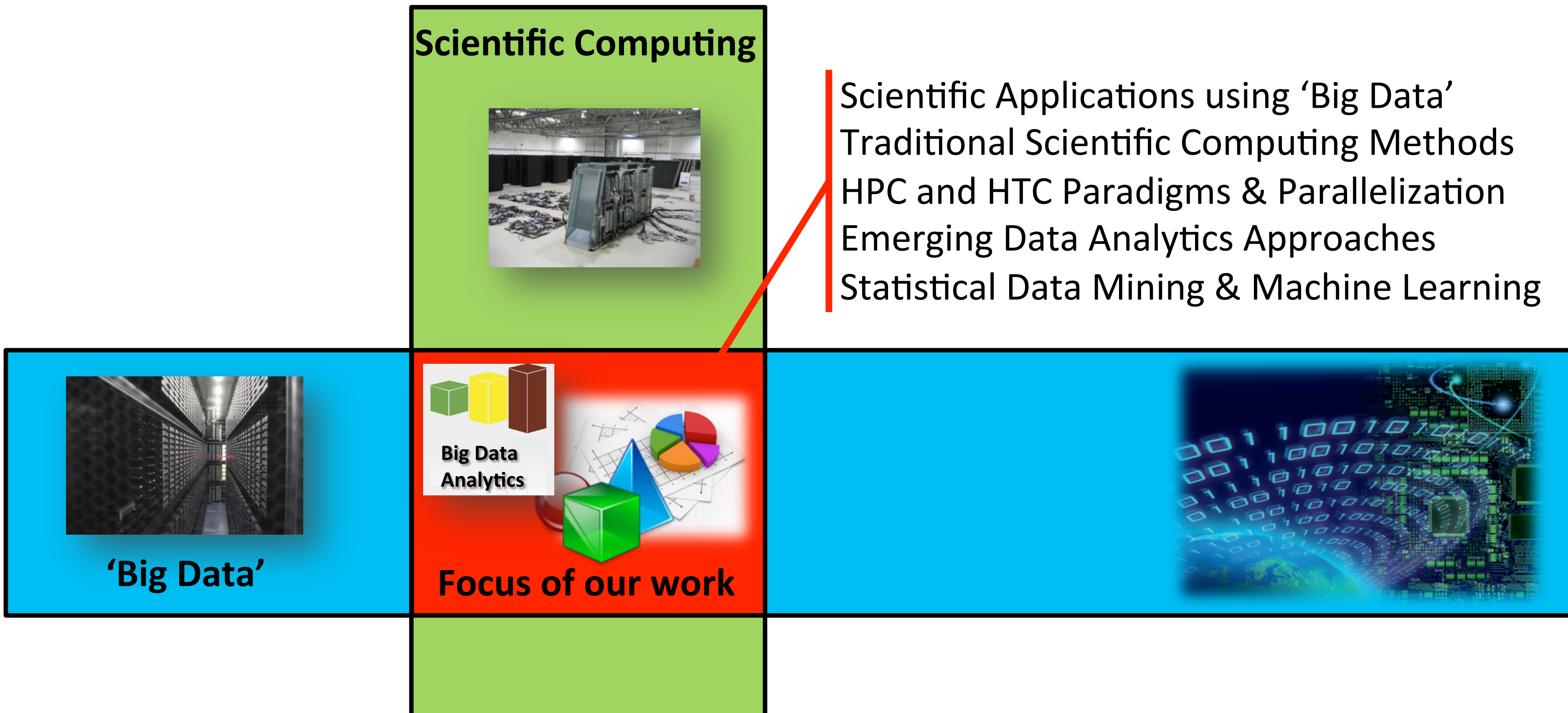
RESEARCH DATA ALLIANCE

How do we enable ,high productivity processing'?

How do we find ,a message in the bottle'?

Work on Intersection of two Broad Subjects

- Key goal: High Productivity Processing of Research Data

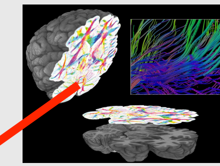


Searching 'Big Data' Evidence

Big Data Analytics

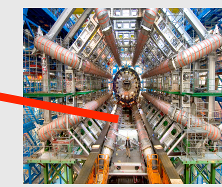
In Commercial environments the term 'Big Data' is often related to Volume – Variety – Velocity, but concrete 'numbers' are rarely given

Talk by Henning Gast

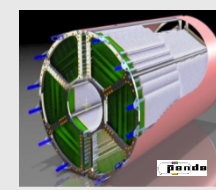


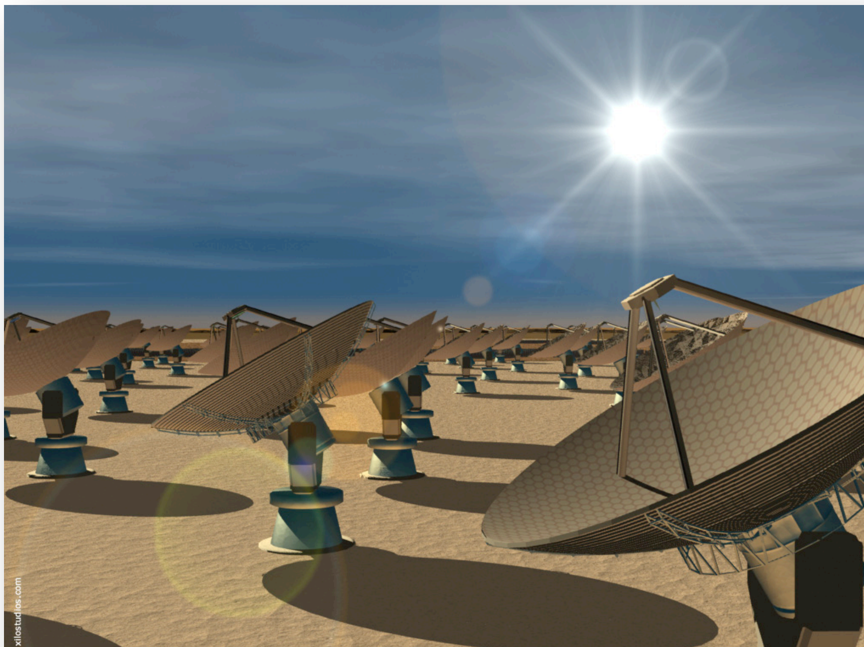
Talk by Katrin Amunts

Talk by Christopher Jung



Talk by Thomas Lippert





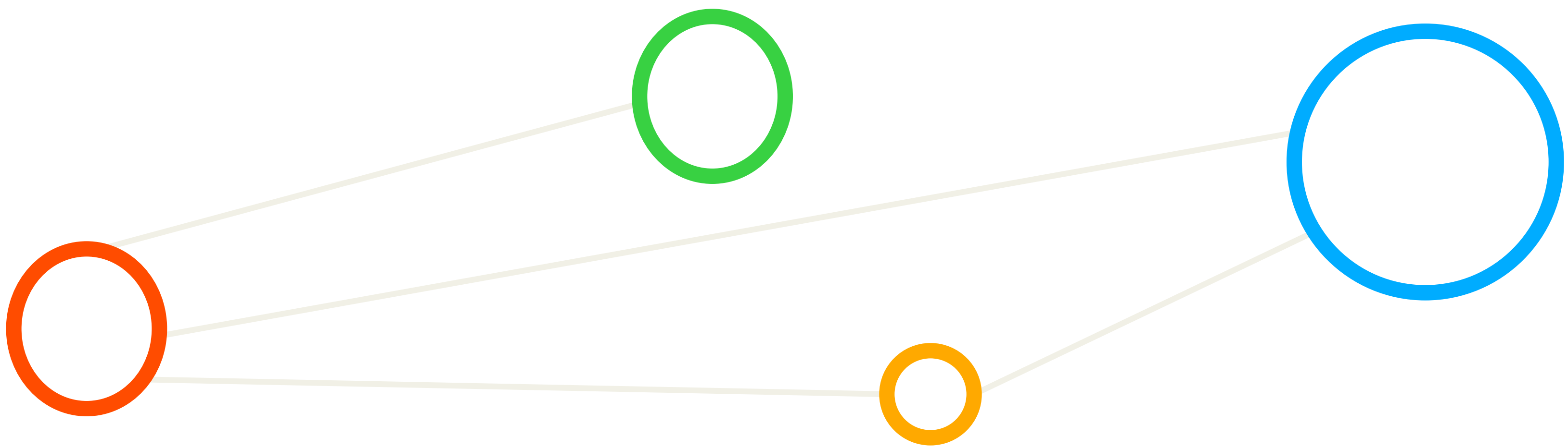
In Science environments the term 'Big Data' is often related to one concrete scientific experiment: e.g. square kilometre array → 1 PB / 20 seconds

Big Data Analysis vs. Big Data Analytics

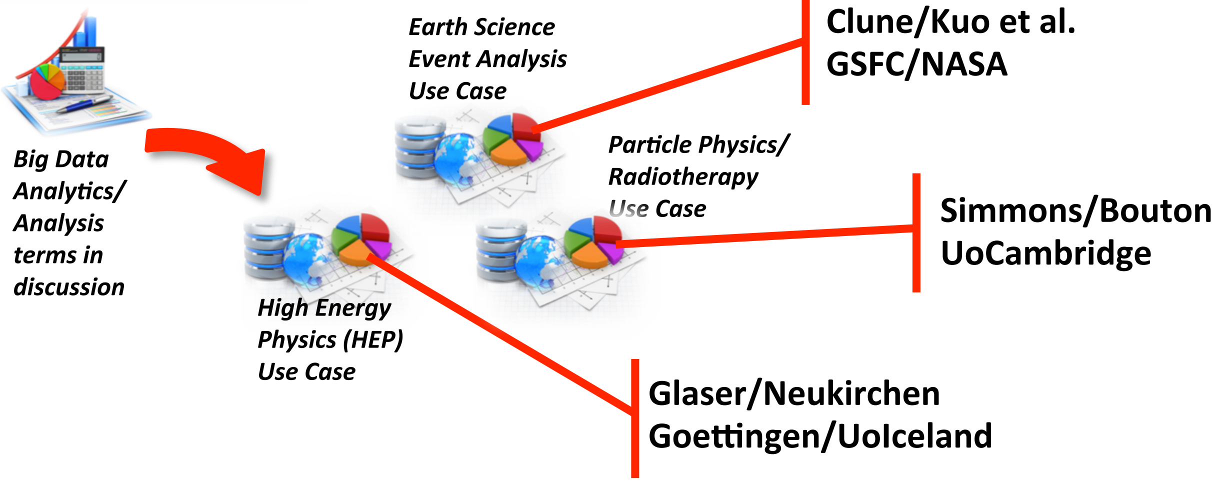
- ‘Data Analysis’ supports the search for ‘causality’
  - Describing exactly WHY something is happening → science
  - Understanding causality is hard and time-consuming, but is necessary
  - Searching it often leads us down the wrong paths…
- ‘Big Data Analytics’ is focussed on ‘correlation’
  - Not focussed on causality – enough THAT it is happening → money/events
  - Discover novel patterns and WHAT is happening more quickly
  - Using correlations for invaluable insights – often data speaks for itself

- **Analysis is the in-depth interpretation of ‚big data‘**
- **Analytics are powerful techniques to work on ‚big data‘**
- **Parameter/event space exploration may use (1) analytics, then (2) analysis**
- **Pre-/Post-Process data with (1) analytics for deeper/faster (2) data analysis processing**

Shaping a Big Data Analytics Landscape



Starting with Scientific Application Use Case



HEP Data Analysis Use Case Example

[4] Glaser and Neukirchen et al.,Using MapReduce for HEP data analysis' (submitted)

- Approach
  - LHC data: ~15 PB/year; reality work on: 10 – 200 TB per year (reduced data)
  - Performed HEP data analysis based on ROOT analysis framework
  - Based analysis on Monte-Carlo event data generated with PYTHIA
  - Assigned workloads to the 'data nodes' that keep data to be processed (compute → data or 'data locality' concept)

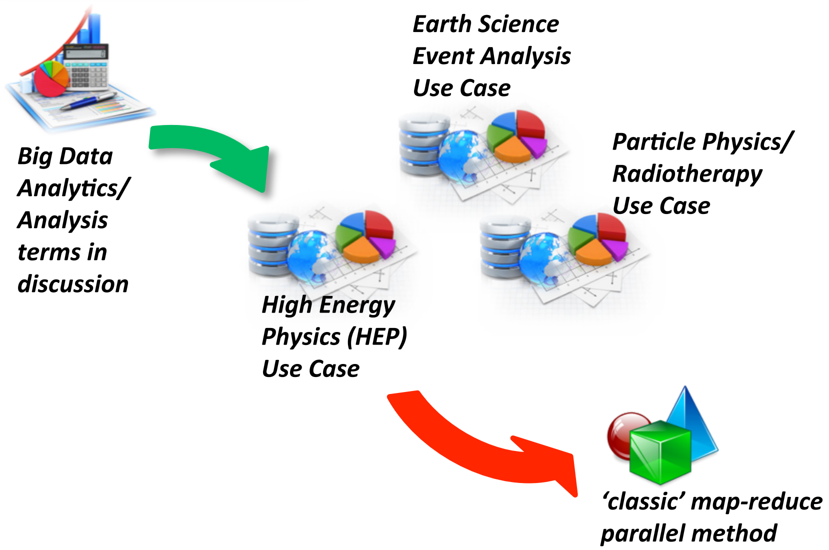
[5] Brun and Rademakers, 'ROOT'

[6] Sjostrand et al., 'PYTHIA'

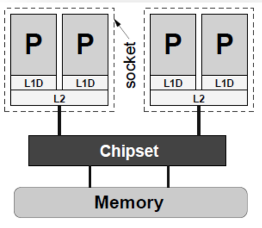
- **Event can be a single bunch crossing of two proton beams**
- **Events can be analyzed independently & in embarassingly parallel**
- **Use of the data analytics method ‚Map-Reduce' possible**

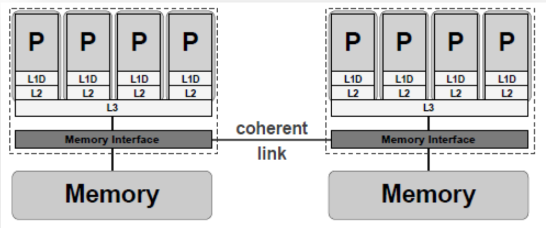
- Selected Conclusions
  - Reduced network load through 'data locality' (negligible vs. compute time)
  - Evaluation shows that using MapReduce for HEP data analysis is 'slower than using the normally applied traditional methods ('PROOF')'

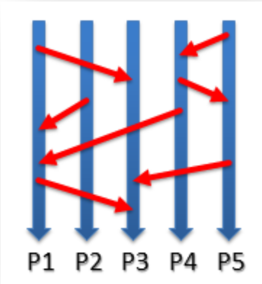
Adding the 'Classic' Map-Reduce Method

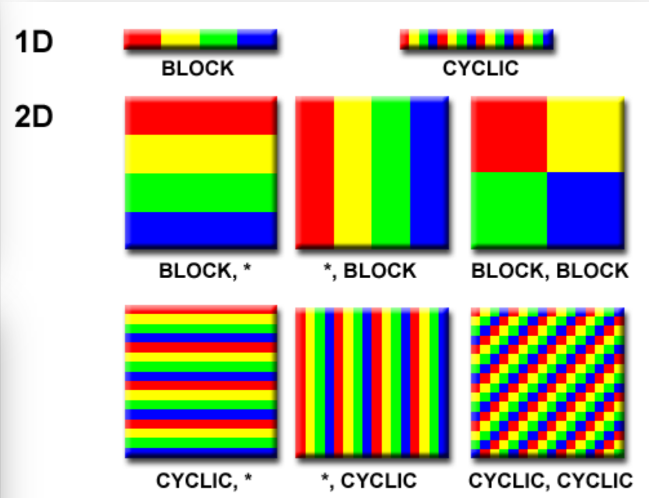


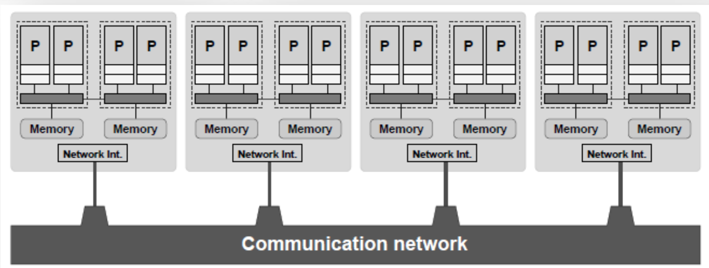
Motivation for Map-Reduce: HPC Complexity

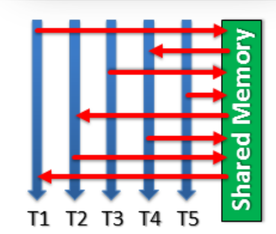










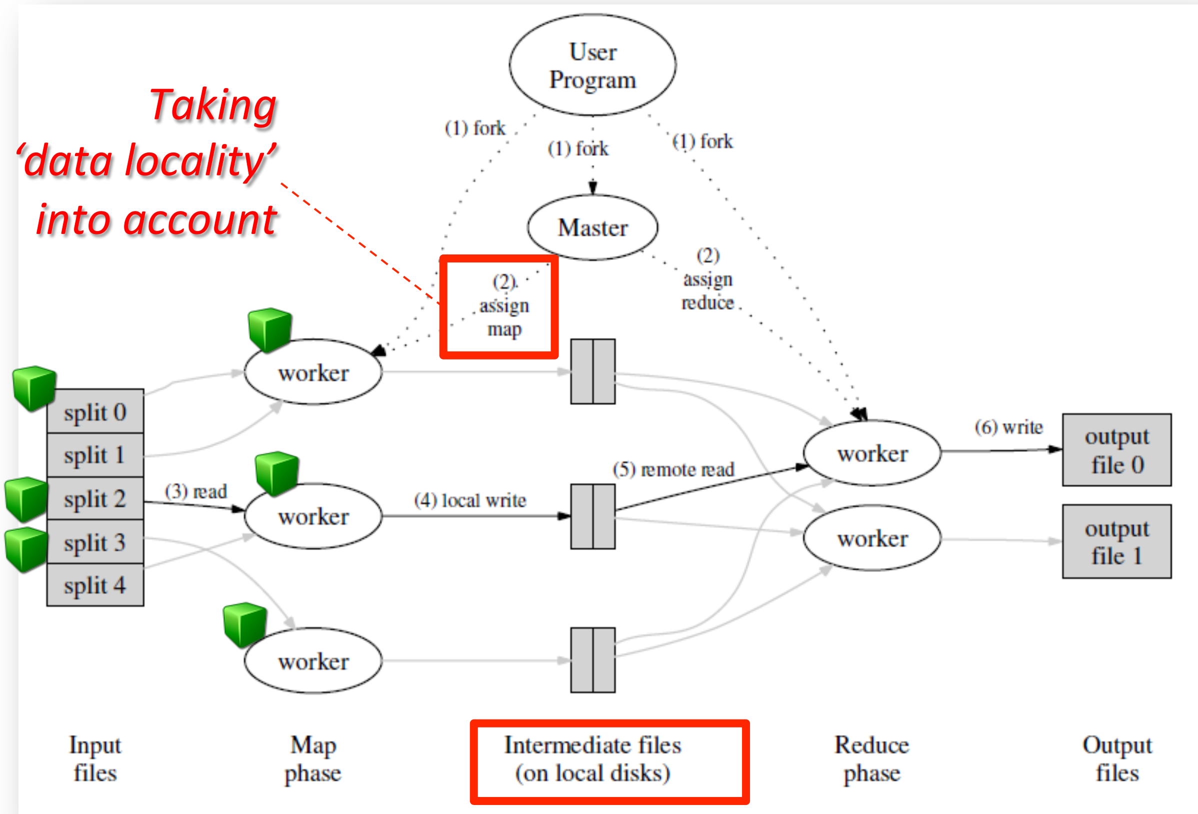


- Different HPC Programming elements (barriers, mutexes, shared-/distributed memory, etc.)
- Task distribution issues (scheduling, synchronization, inter-process-communication, etc.)
- Complex heterogenous architectures (UMA, NUMA, hybrid, various network topologies, etc.)
- Data/Functional parallelism approaches (SMPD, MPMD, domain decomposition, ghosts/halo, etc.)

‘Classic’ Map-Reduce Parallel Method

- Think map-[sort/shuffle/group]-reduce to understand it
  - Enables many ‘common parallel calculations easily’ on large-scale data
  - Performs well on commodity computing clusters (but security critics exists)
  - Offers system that is very tolerant of hardware failures in computation
- Many Implementations exist
  - Often referenced is Apache Hadoop (Java) framework, (mostly because of large ‘ecosystem’)
  - Strong relationship to an underlying ‘distributed file system’
- Key to understanding is the map-reduce runtime
  - **Takes care or the partitioning of input data and the communication**
  - **Manages parallel execution and <u>performs sort/shuffle/grouping</u>**
  - **Coordinates/schedules all tasks that either run Map and Reduce tasks**
  - **Handles faults/errors in execution and re-submit tasks**

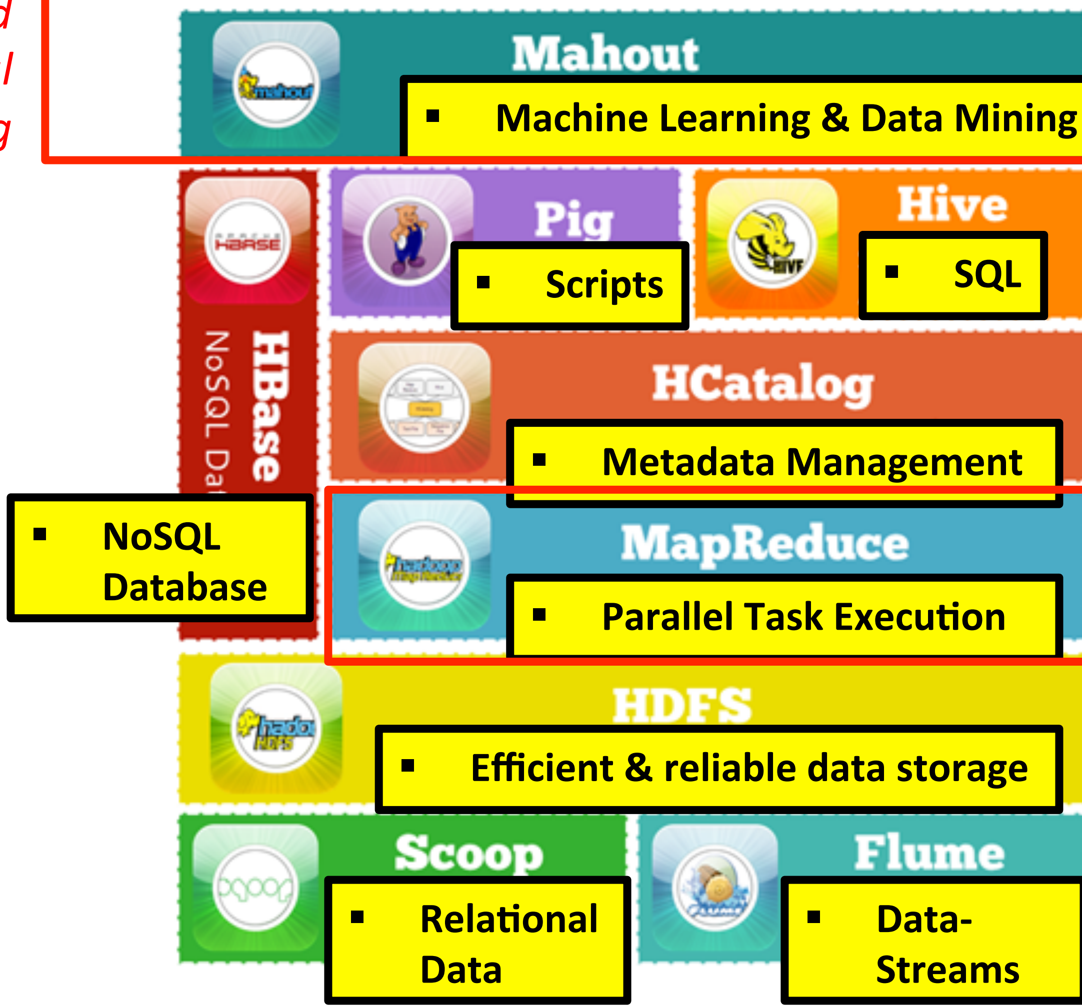
'Classic' Map-Reduce Parallel Method



Modified from [8] Dean & Ghemawatt et al., 'MapReduce: Simplified Dataset on Large Clusters'

Big ecosystem evolved around Map-Reduce

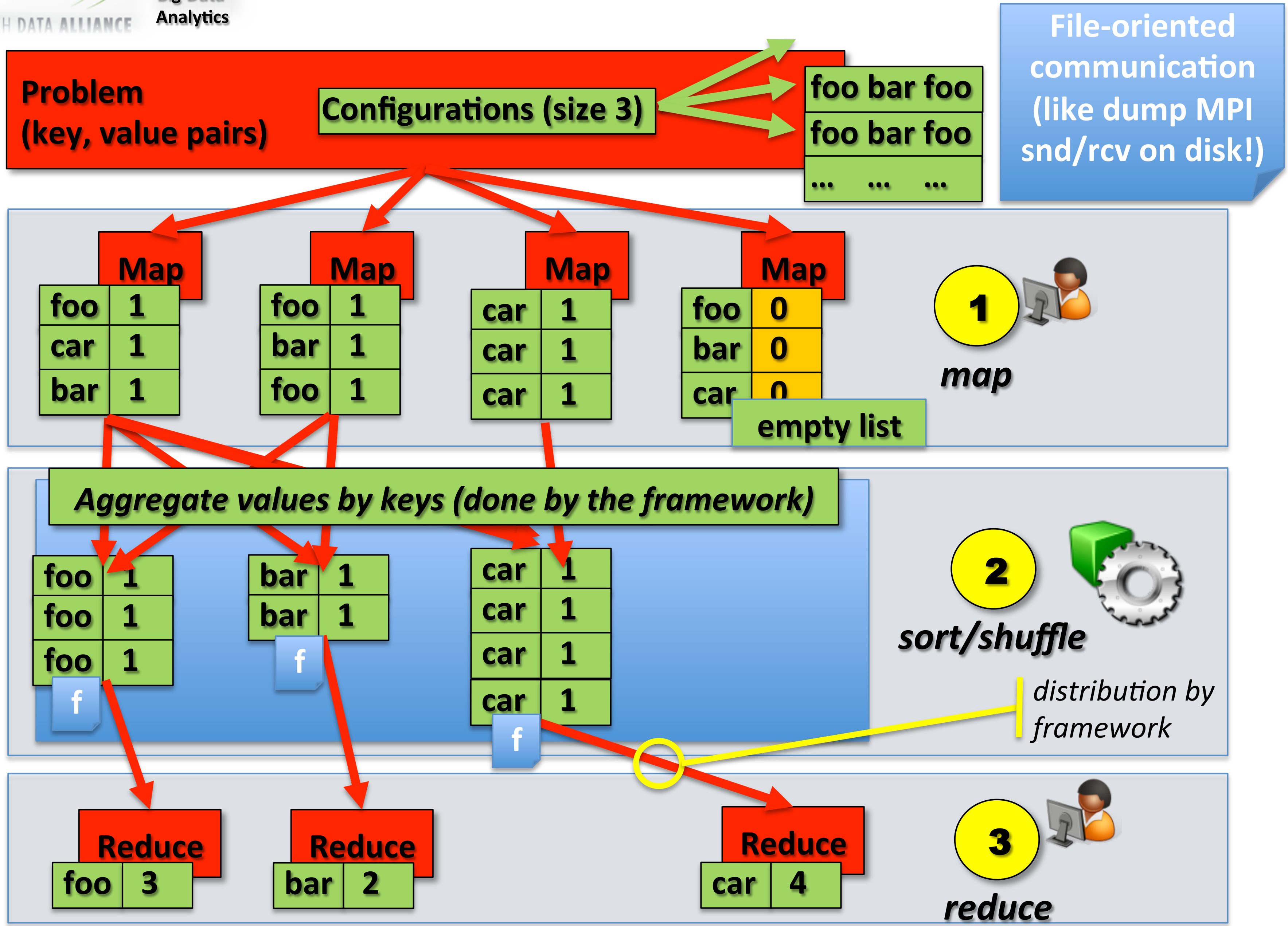
Applied Statistical Data Mining



'Classic' Map-Reduce with Hadoop

Modified from [7] Map-Reduce

‘Classic’ Map-Reduce Example – WordCount



HEP Application Use Case Revisited (1)

[4] Glaser and Neukirchen et al.,Using MapReduce for HEP data analysis' (submitted)

- More information about the data analysis
  - Analysis in two steps: (1) check events containing a specific signature; (2) deeper event/signature analysis
  - Use of Monte Carlo event generation program PYTHIA
  - PYTHIA output: traces of the involved particles/event from detectors
  - Point of closest approach: two particle tracks (positive/negatively charged)
  - Closest distance is point where particle of interest decayed into the two further particles
  - The outcome of the analysis is a histogram that depicts the reconstructed particle masses in a certain range.
- Use of the Map-Reduce framework Hadoop
  - Open-source, and broadly used

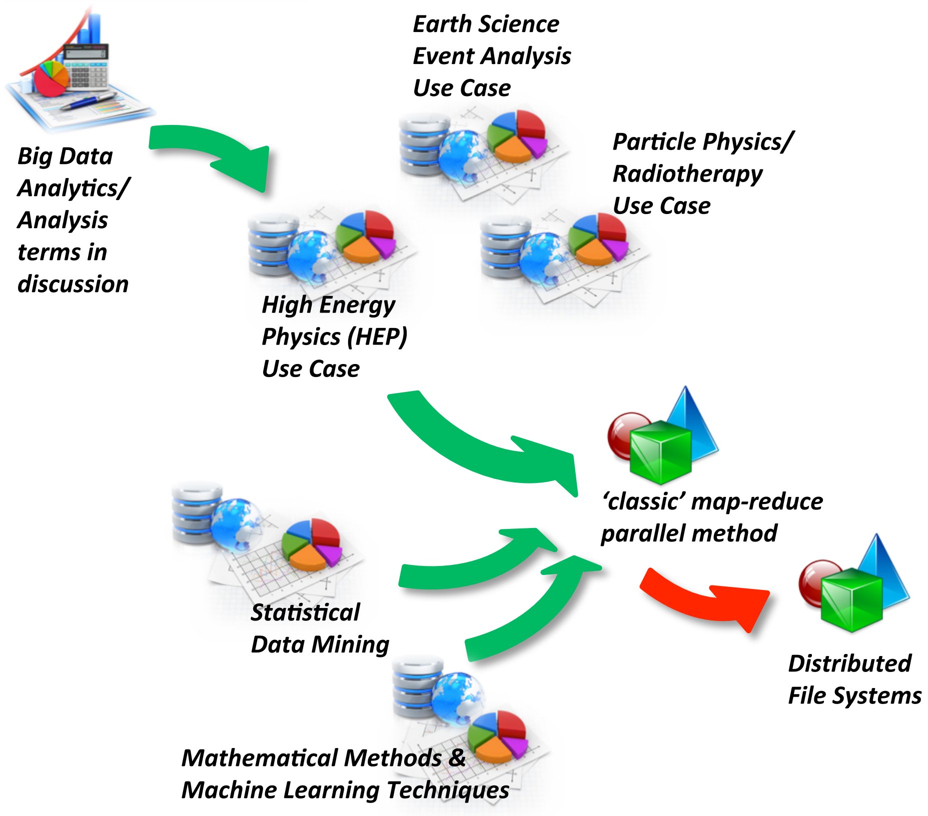
HEP Application Use Case Revisited (2)

[4] Glaser and Neukirchen et al.,Using MapReduce for HEP data analysis' (submitted)

- More information about Map-Reduce & Parallelization
  - Each event is checked whether it contains particles with certain characteristics (in the example use case data analysis: a certain mass)
  - Intermediate results for the matching particles are produced, then merged
- Map: Event-level analysis (afterwards sorted by framework)
  - Producing key/value pairs with mass of matching particle
  - Input: <path to event file, event # in file>
  - Output: <mass of the particle, # of the observed particles with this mass>
- Reduce: Statistical analysis on sorted data
  - Use data with mass of matching particles & produce histogram using ROOT

- **Results depend on the structure and organization of input files and the underlying filesystem**
- **Apache Hadoop uses the Hadoop Distributed File System (HDFS)**

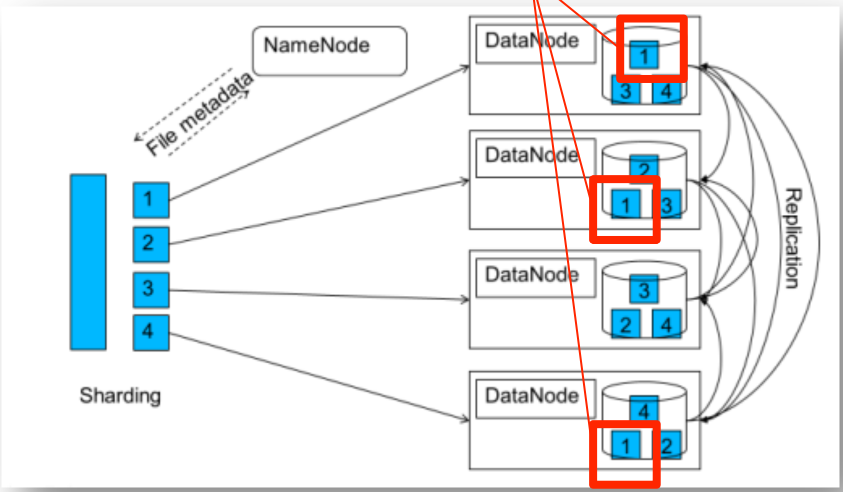
Adding strong relationships of Map-Reduce



Benefits via Distributed File System (DFS)

- File Handling designed for…
  - Files spanning multiple nodes (TBs, PBs), [better 'accumulated over time']
  - Files are rarely updated [optimized rather for read/append – not re-write]
  - Bring computation to the files, instead of the files to the computing resource ('data locality principle') [replication of big data sets n times?]
  - Assuming failures & enable reliability
  - Scalability of the whole system

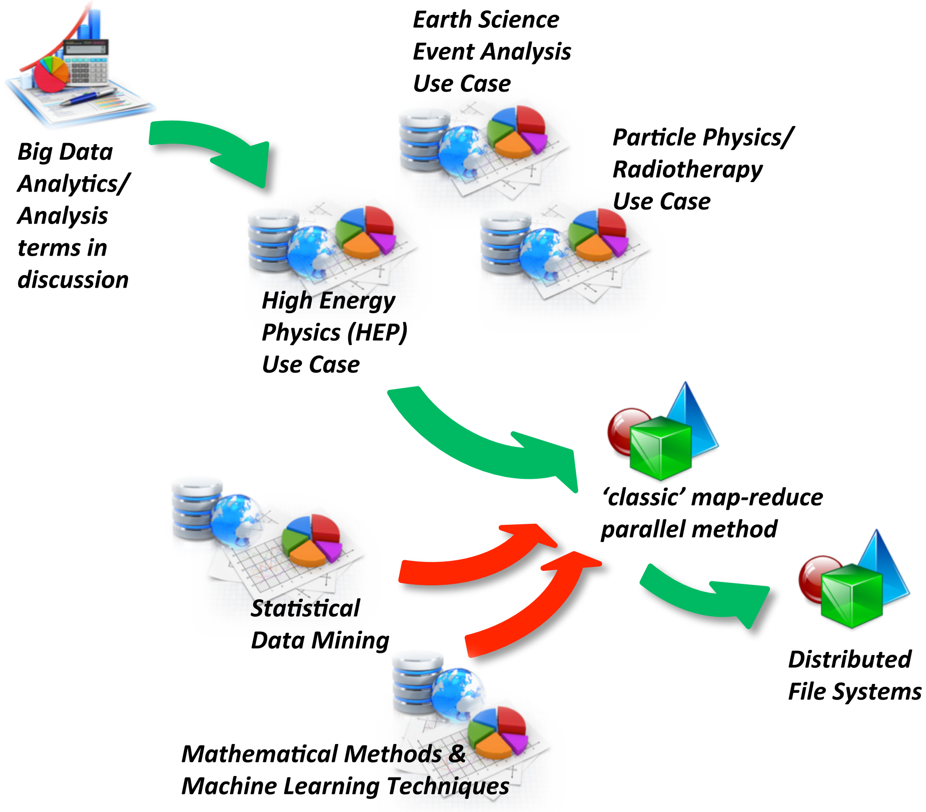
Distributed Filesystem



- **Files are divided into blocks – also named as 'chunks' (default 64MB)**
- **Blocks are replicated at different compute nodes (default 3x – realistic in science?)**
- **Blocks holding copy of one dataset are distributed across different racks**

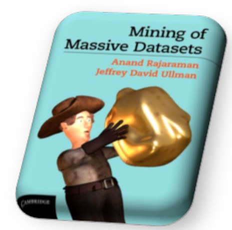
[9] Map-Reduce Virtual Workshop

Relationships to certain other fields



Statistical Data Mining & Machine Learning

- Machine Learning & Data Mining (related, but different fields)
  - Making predictions by learning from data or mining huge data sets
  - Use: (1) A pattern exists; (2) Not mathematically describable; (3) DATA!
  - Approaches: Classification, Clustering, Regression (concrete rankings)
  - Fields that are usually not applying techniques of ‘parallelization’
  - Tools used are R for ‘statistical computing & Rattle (GUI for R) *[17] R-Project*
- Training next generation ‘data scientists’ is a mix of various fields
  - Scientific computing skills combined with data analytics & applied statistics
  - Partnership Juelich-UoIceland formed in this area
  - PhDs, MSc: HPC-B(ig data) & Statistical Data Mining



UNIVERSITY OF ICELAND
SCHOOL OF ENGINEERING AND NATURAL SCIENCES
FACULTY OF INDUSTRIAL ENGINEERING, MECHANICAL ENGINEERING AND COMPUTER SCIENCE

JÜLICH FORSCHUNGSZENTRUM

[12] Book Mining of Massive Datasets Online

Scalable Machine Learning and Parallelization

- Benefit: e.g. process whole datasets, instead of N samples of data
- Idea: Use parallelization with statistical program R
  - R map-reduce plugin exists
  - R-MPI exists, pbdR emerges (hide MPI for R users)
- Apache Mahout is emerging using parallelization
  - Works on top of map-reduce
  - Collaborative Filtering
  - User and Item based recommenders
  - K-Means, Fuzzy K-Means clustering
  - …
  - Random forest decision tree based classifier

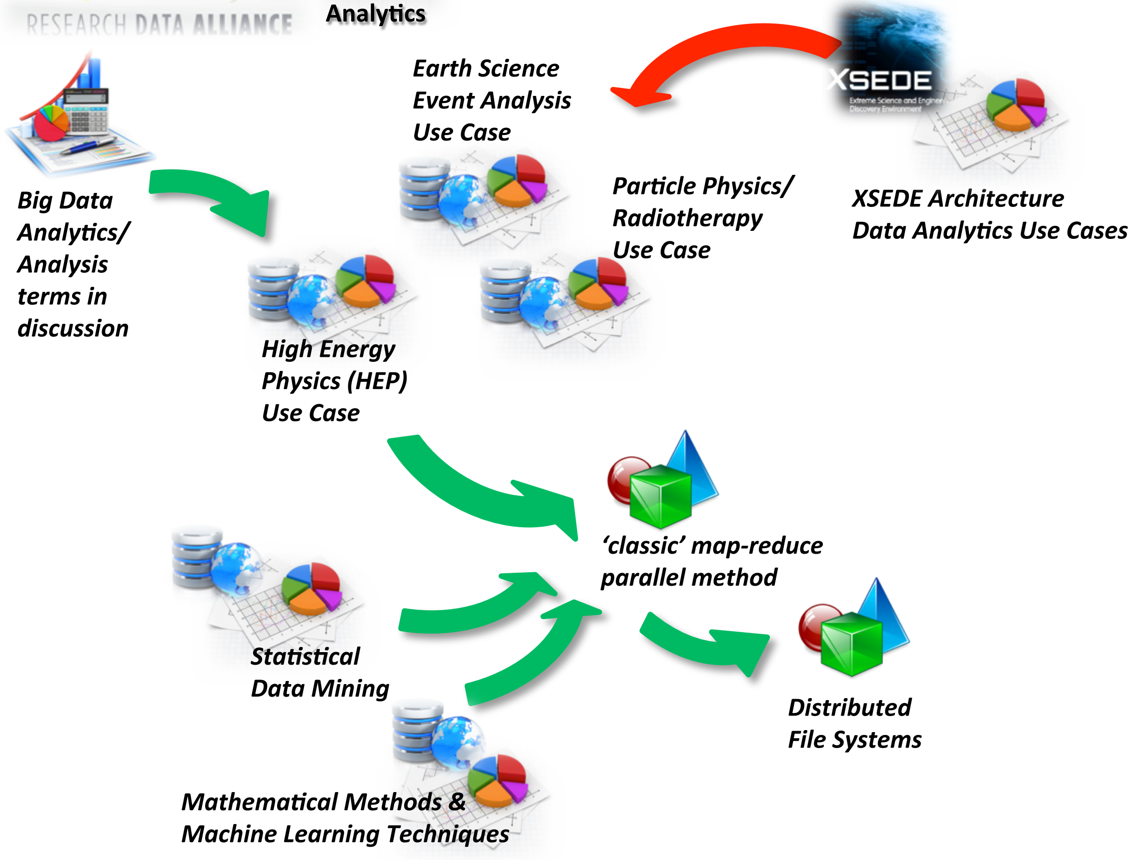
[13] XSEDE 13 Turorial on pdbR

[13] XSEDE 13 Turorial on Apache Hadoop & Mahout on Gordon

[14] Apache Mahout Webpage

Big Data Analytics

CHANGES: XSEDE data analytics use cases



XSEDE architecture team collects use cases

XSEDE Data Analytics Use Cases

14th Jun 2013

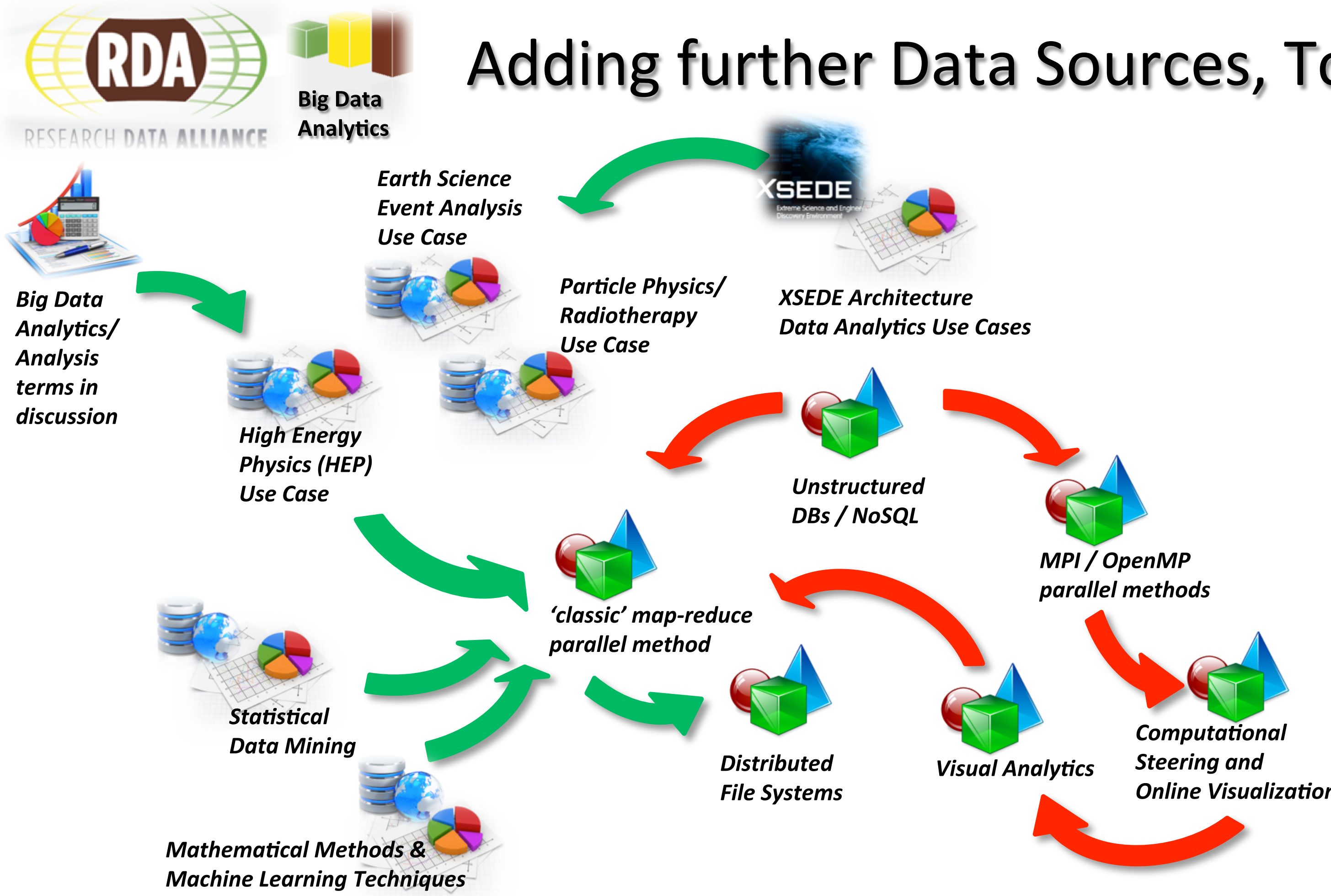
Version 0.3

Shawn Strande et al.

XSEDE Architecture Data Analysis Use Cases

- Data Analytics use cases from stakeholders
  - (Hard to reference since on google docs and not completely open yet)
- Includes software requirements
  - Visualization, data mining, and statistical tools need to be available
  - Examples include: R, Hadoop, etc.
- Includes HPC Simulation Data Analysis
  - Postprocessing and visulization (interesting might be also pre-processing)
- Interactive Computational steering
  - Can be roughly considered as a form of visual analytics

Adding further Data Sources, Tools, Methods

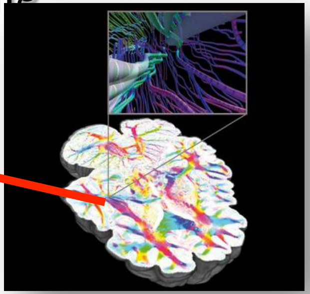


Supporting tools for Data Analytics

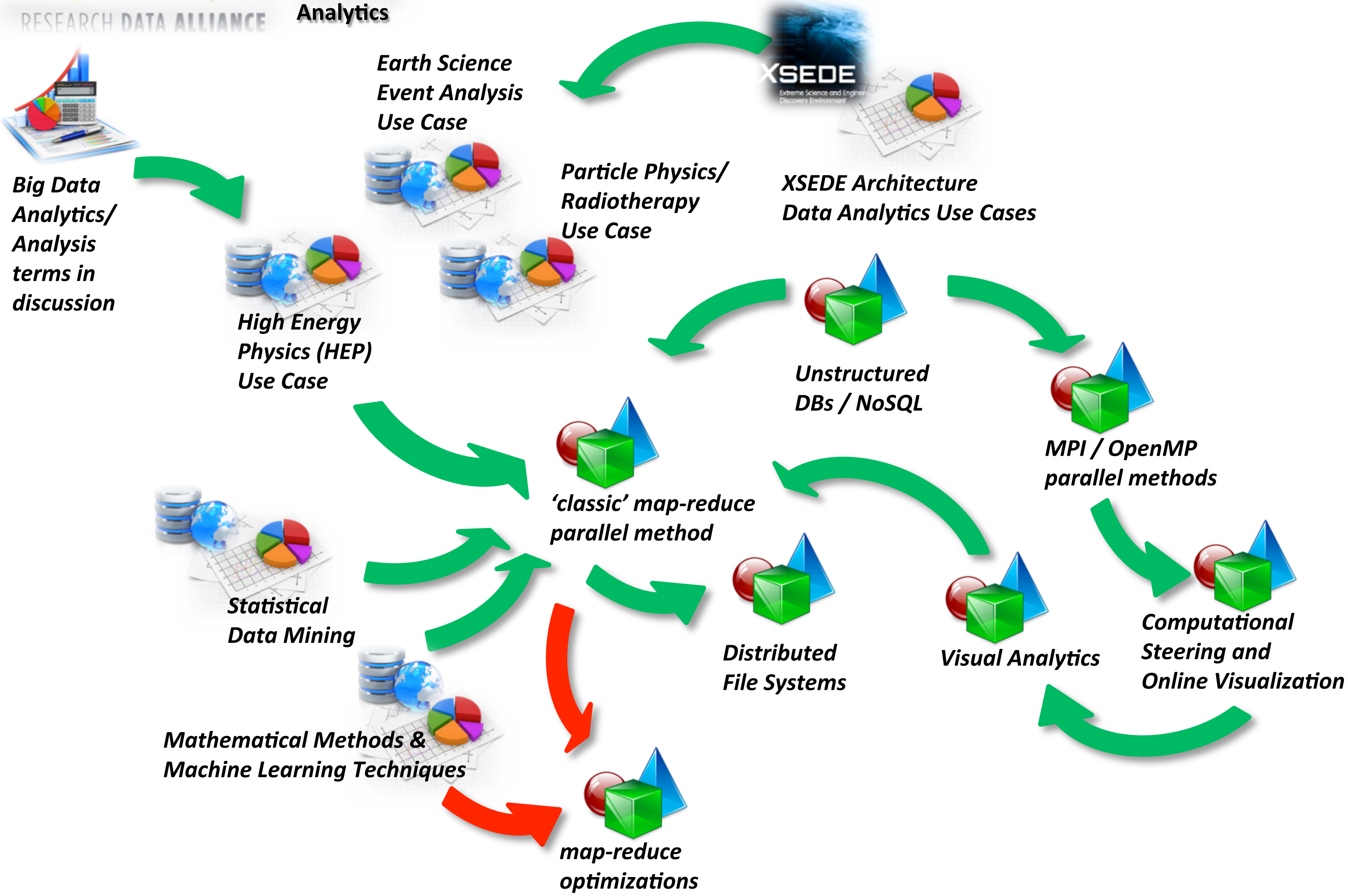
[15] NoSQL DBs

- Unstructured Databases / NoSQL DBs
  - Easy to deploy, implement and relatively cheap to operate
  - Easy to geographically distribute withs calability inherent in the DBs
  - Designed with ‚no schemas‘ → low data consistency requirements
  - Optimized for quickly processing extremely large datasets (streams)
- Extreme Data Sources (e.g. sensors, etc.)
  - Crowd Sourcing: Massive data streams (e.g. ratings from citizen scientists)
- ‘Visual Analytics’ → combination of evidence & assumptions
  - 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

Talk by Torsten Kuhlen

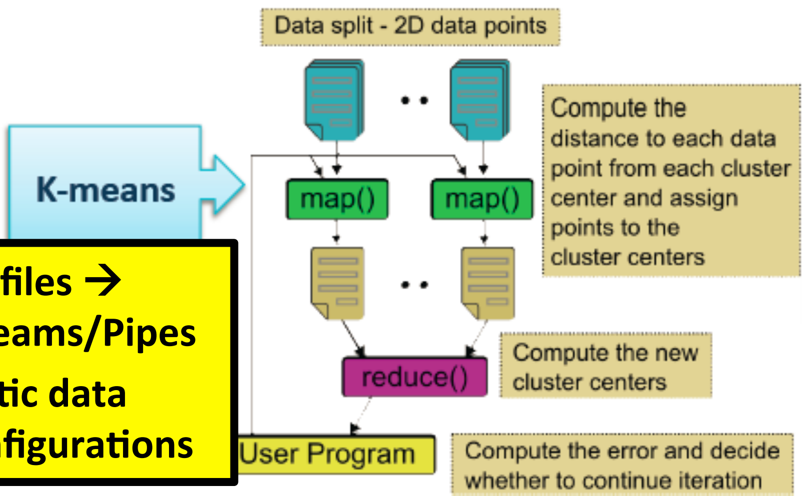


Using Optimizations of Map-Reduce



Twister: Iterative Map-Reduce as Example

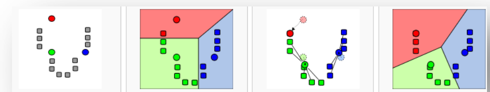
Iterative and non-Iterative Computations



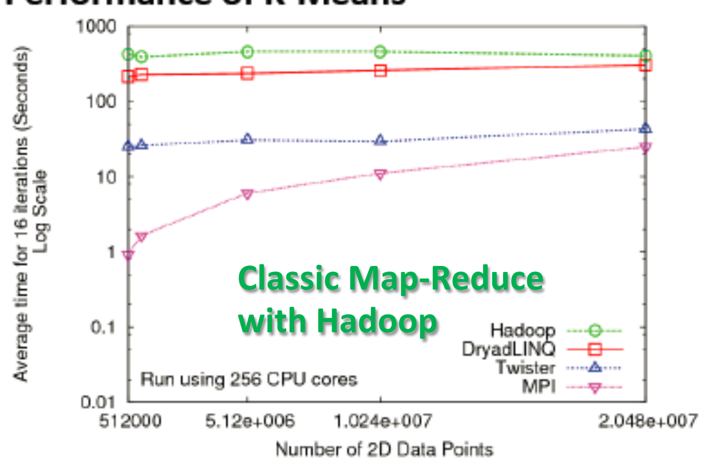
- No files → Streams/Pipes
- Static data configurations

K-Means Clustering (NP-hard)

Partition n observations into n clusters → Voronoi cells

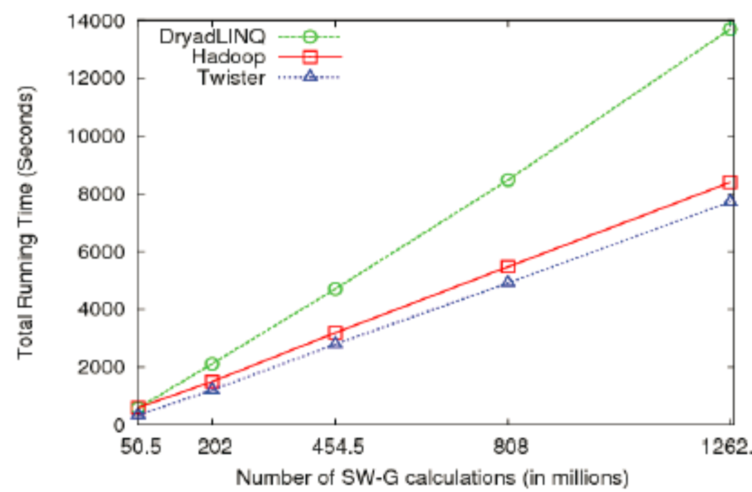


Performance of K-Means



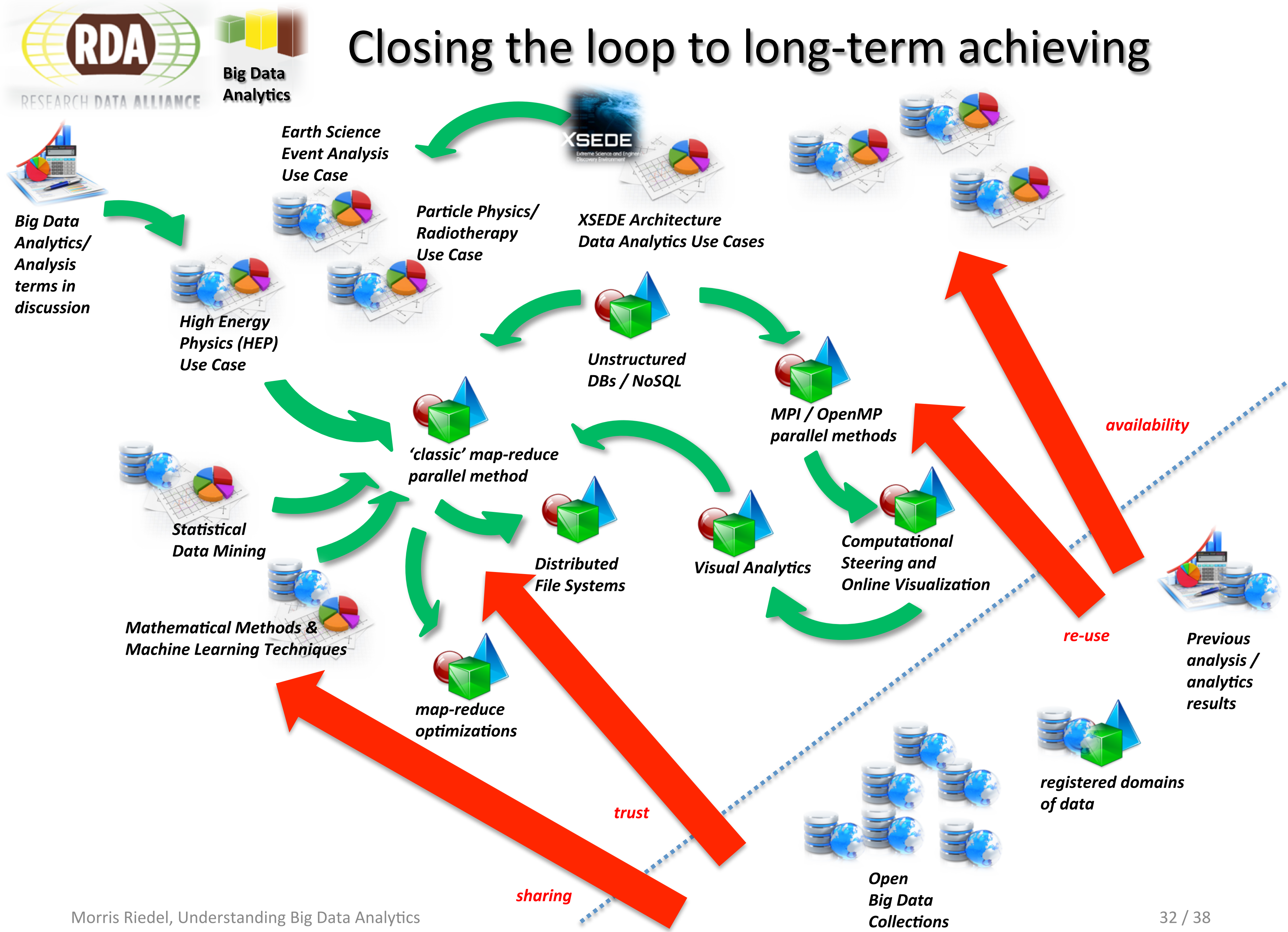
Smith Waterman is a non iterative case and of course runs fine

→ used in bio-informatics



modified from [16] G. Fox et al., ‚MPI and Map-Reduce

Closing the loop to long-term achieving



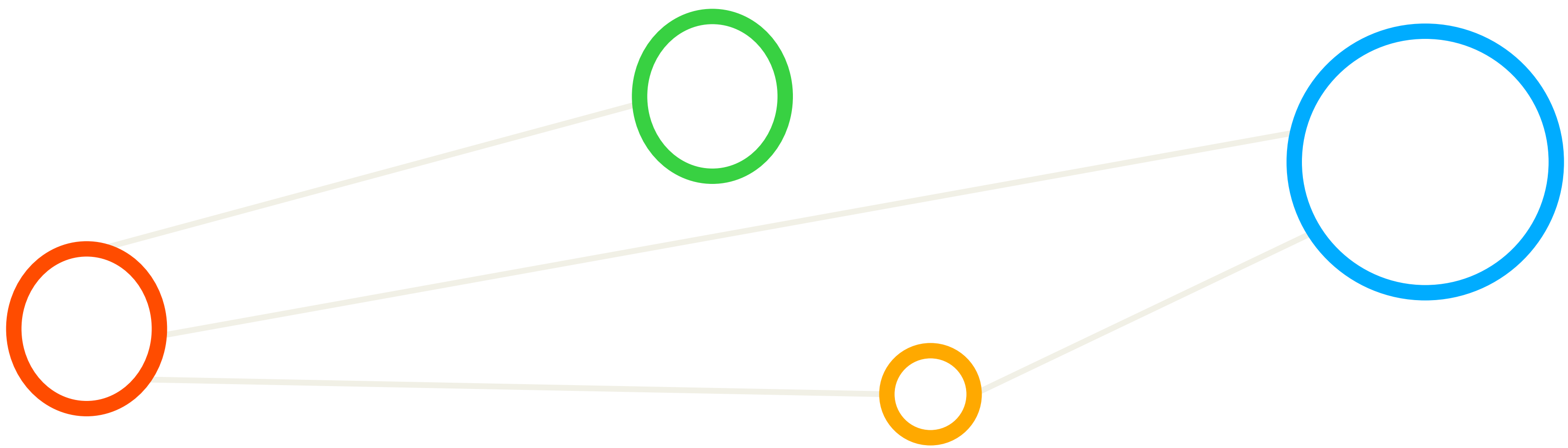
Current Systematic Analysis – Methodology

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'

Need more granularity and concrete ‚application databases' underpinned with evidence data

extended from [16] G. Fox et al., ‚MPI and Map-Reduce

Summary



RESEARCH DATA ALLIANCE

Big Data Analytics

Summary

- 'Rough concensus' on Big Data Analytics related terms exists
  - Discuss terms further in RDA BDA group and give some 'guidelines'
- Towards understanding 'Big Data Analytics in Research'
  - Based on concrete scientific computing application use cases
  - Increased evidence over time brings a clearer picture & more granularity
  - Combine different methods for different parts of data analysis/analytics
- Analyse methods & tools applied to 'big dataset' → Methodology
  - Initial results indicate that map-reduce is like a 'reliable HTC of the past'
  - Massive parallel HPC applications not going to vanish nor use map-reduce
  - More thinking about data storage with file locations required
  - Continue to list evidence with data-driven HPC use cases
  - Add dimension of GreenIT and consider 'low energy footprint'

Big Data Analytics @ 2nd RDA Plenary Meeting

RDA RESEARCH DATA ALLIANCE

Big Data Analytics

Building Global Partnerships - RDA Second Plenary Meeting

Date: 16/09/2013 to 18/09/2013

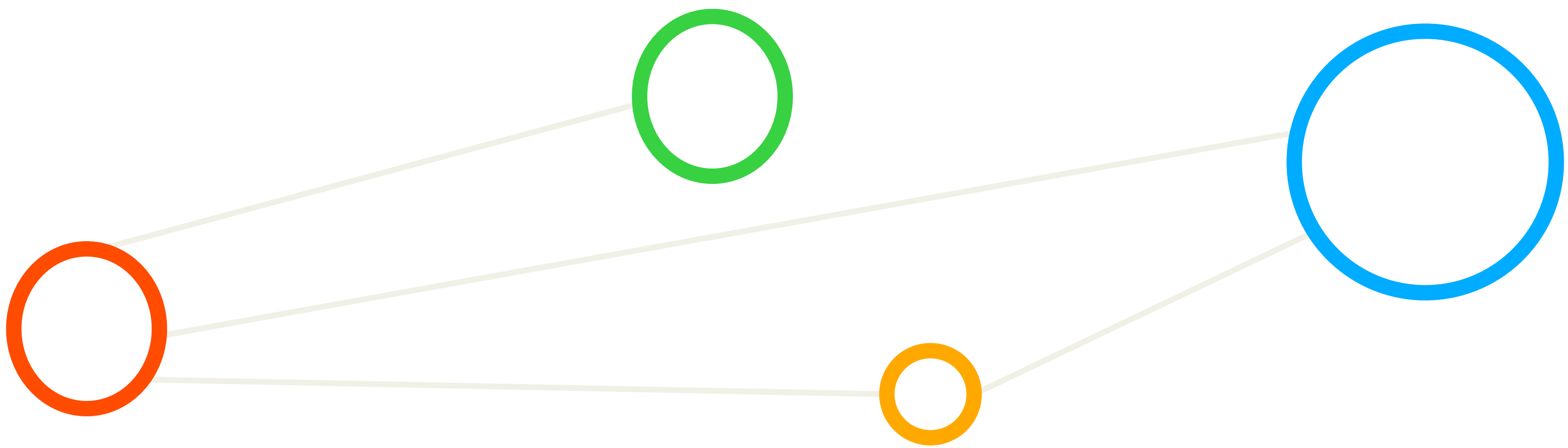
16-18 September 2013, National Academy of Sciences, Washington DC, US

[3] RDA Web Page

Tuesday 17th September | Washington Marriott

14:00 - 16:00	Break-out Session III						
Room	**Salon A**	**Salon B**	**Salon C**	**Salon D**	**Salon E**	**Salon F**	**Salon G**
14:00 - 15:00	Practical Policy	RDA/WDS Publishing Data	PID Types	Metadata	Brokering	Big Data Analytics	Organizational Members business meeting
15:00 - 16:00	Practical Policy	RDA/WDS Publishing Data	PID Types	Metadata	Brokering	Big Data Analytics	Long Tail of Research Data IG
16:00 - 16:15	Coffee break						

References



References

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

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

[3] Research Data Alliance (RDA) Web Page, Online: https://rd-alliance.org/node

[4] F. Glaser and H. Neukirchen et al.,Using MapReduce for HEP data analysis', submitted

[5] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 389, no. 1, pp. 81–86, 1997, see also http://root.cern.ch.

[6] T. Sjostrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1," Computer Physics Communications, vol. 178, no. 11, pp. 852–867, 2008.

[7] Einführung in Hadoop (German language, sorry),
online: http://blog.codecentric.de/2013/08/einfuhrung-in-hadoop-die-wichtigsten-komponenten-von-hadoop-teil-3-von-5/

[8] MapReduce: Simplified Dataset on Large Clusters, J. Dean and S. Ghemawat, 6th Symposium on Operating Systems Design and Implementation,
Online: https://www.usenix.org/legacy/event/osdi04/tech/full_papers/dean/dean.pdf

[9] Map-Reduce Virtual Workshop, Stampede,
Online: http://www.cac.cornell.edu/Ranger/MapReduce/dfs.aspx

[10] RDA New York Times Article, Online:
http://www.nytimes.com/2013/08/13/science/how-to-share-scientific-data.html?emc=edit_tnt_20130812&tntemail0=y&_r=2&

[11] Apache Hadoop, Online: http://hadoop.apache.org/

[12] Mining of Massive Datasets, Online: http://infolab.stanford.edu/~ullman/mmds/book.pdf

[13] XSEDE13 Tutorials, Online: https://www.xsede.org/web/xsede13/tutorials

[14] Apache Mahout, Online: http://mahout.apache.org/

[15] FatCloud Overview, Online: http://www.fatcloud.com

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

[17] R-Project for Statistical Computing, Online: http://www.r-project.org/