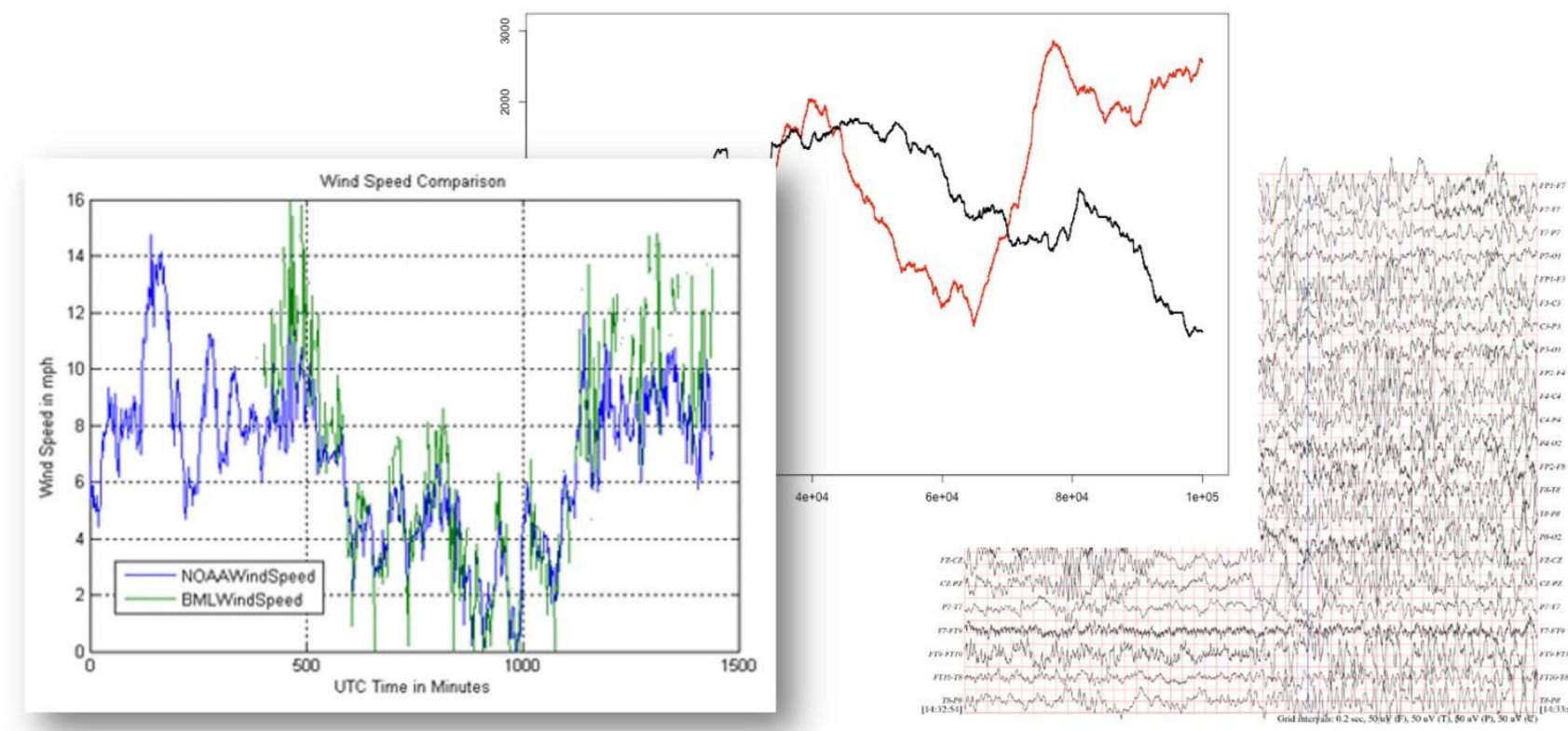


Motivation

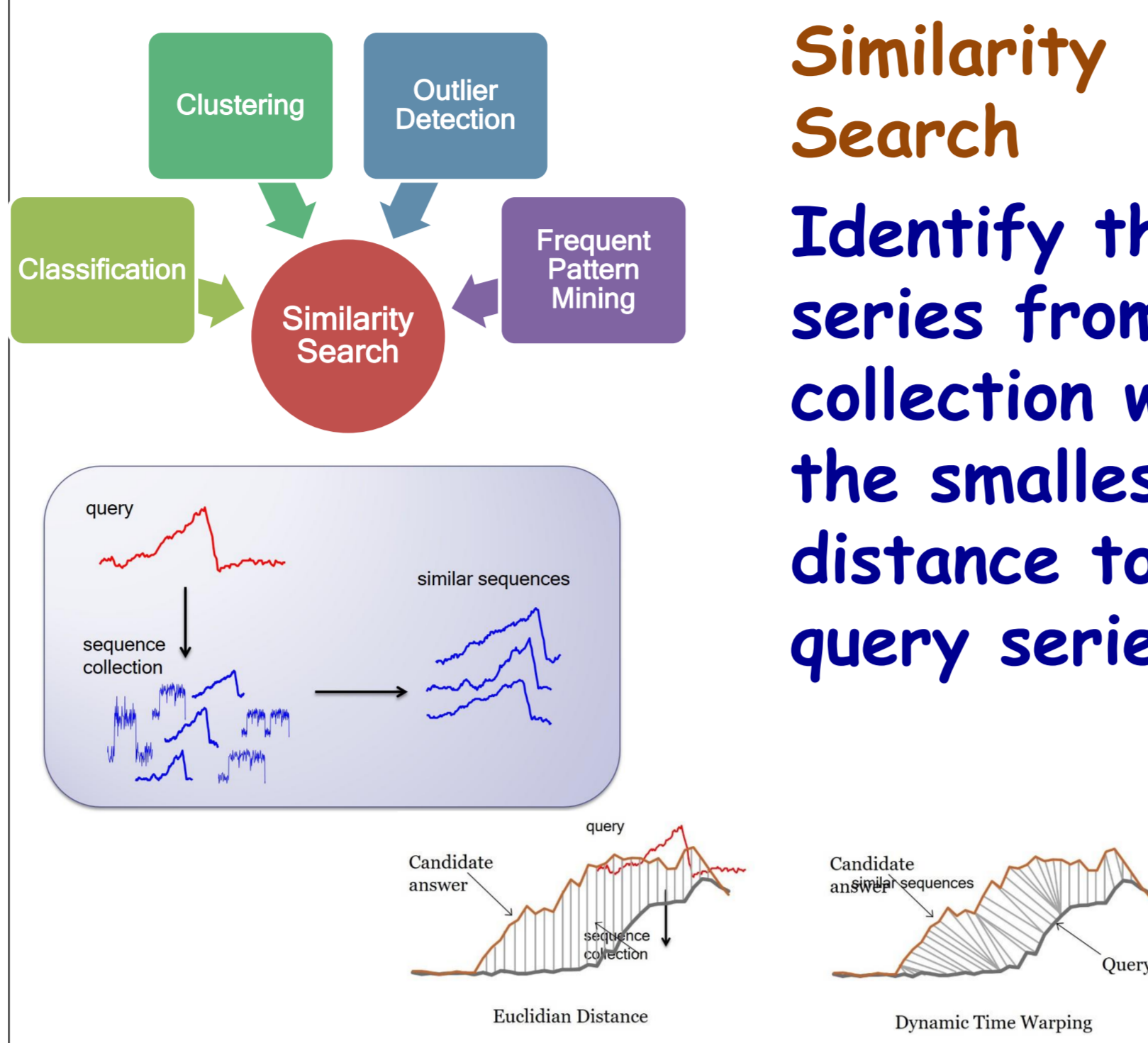
DATA SERIES



Appear in many scientific fields:

- seismology and earth sciences,
- finance
- astrophysics,
- neuroscience,
- engineering, etc.

COMPLEX ANALYTICS



Similarity Search
Identify the series from the collection with the smallest distance to the query series

MAIN CHALLENGE

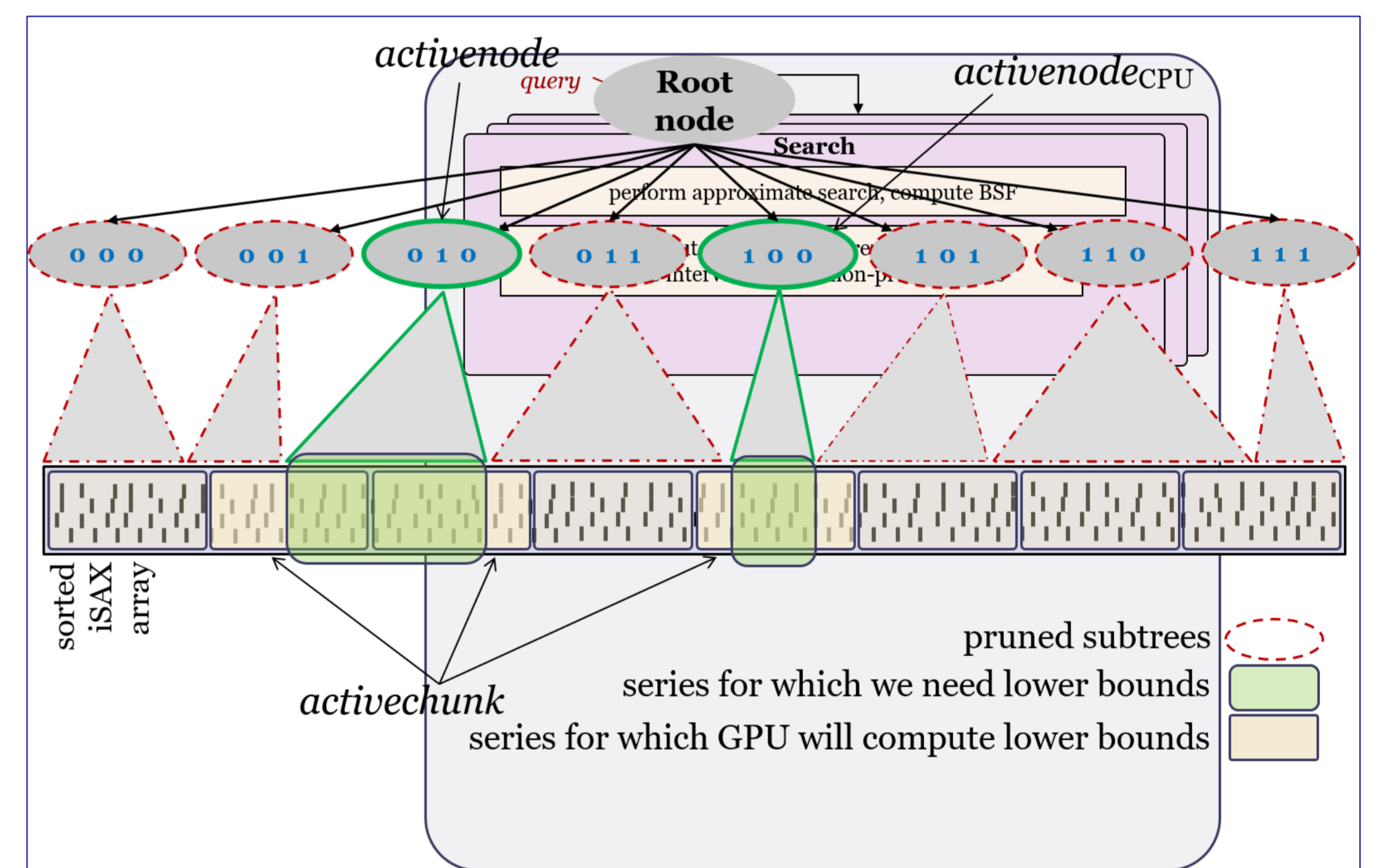
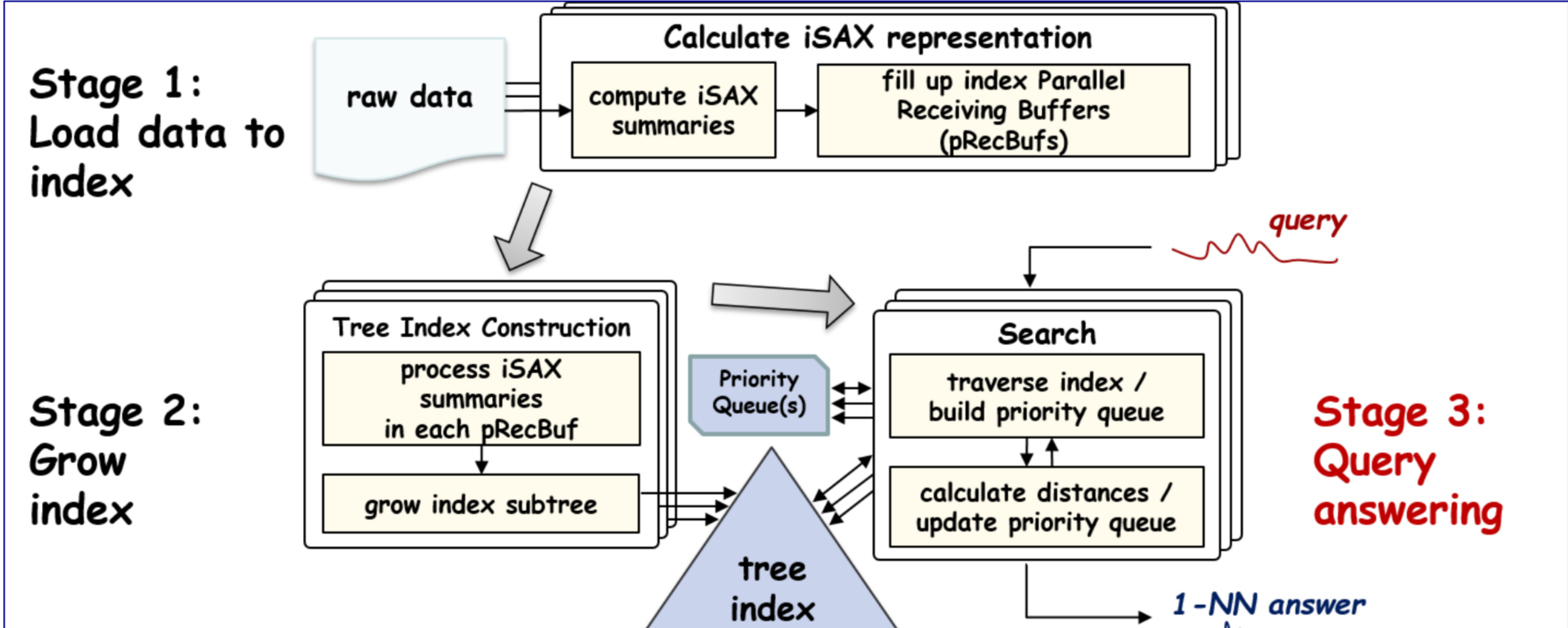
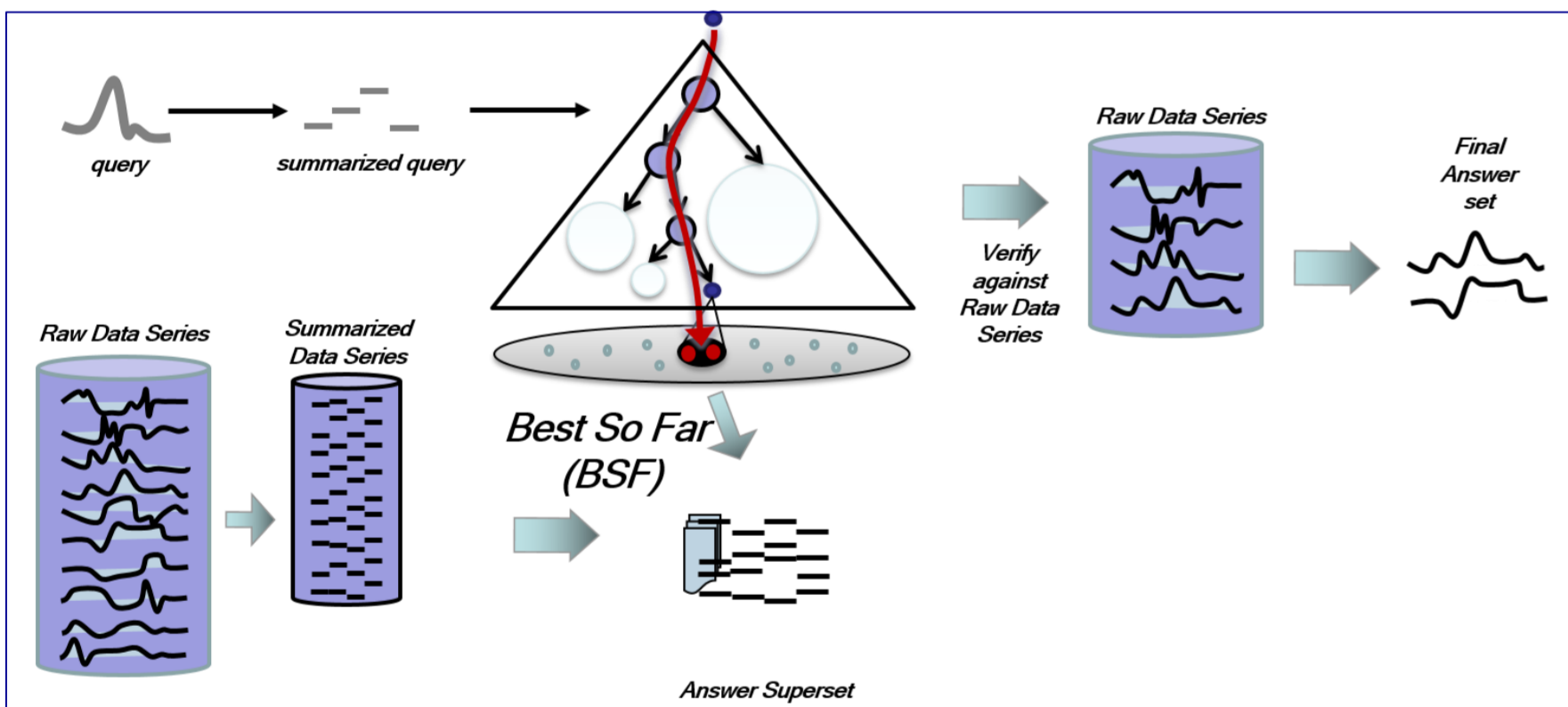
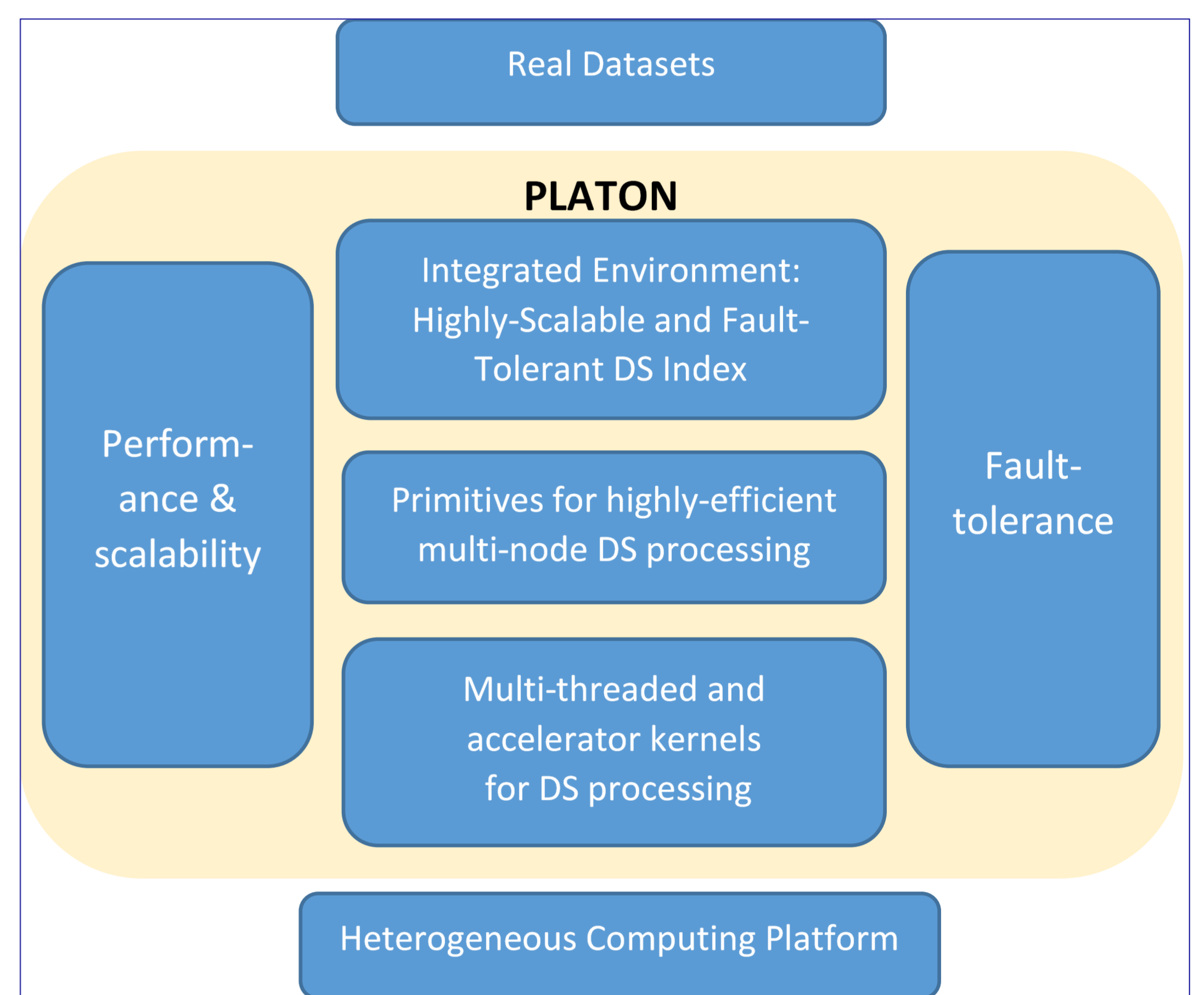
<p>NASA's Solar Observatory 1.5 TB per day</p>	<p>Human Genome project 130 TB</p>
<p>Large Synoptic Survey Telescope (2019) ~30 TB per night</p>	<p>data center and services monitoring 2B data series 4M points/sec</p>
<p>passenger aircraft 20 TB per hour</p>	

Massive data series collections
Harness the difficulties of large-scale data series processing.

PLATON Objectives and Methodology

OBJECTIVES

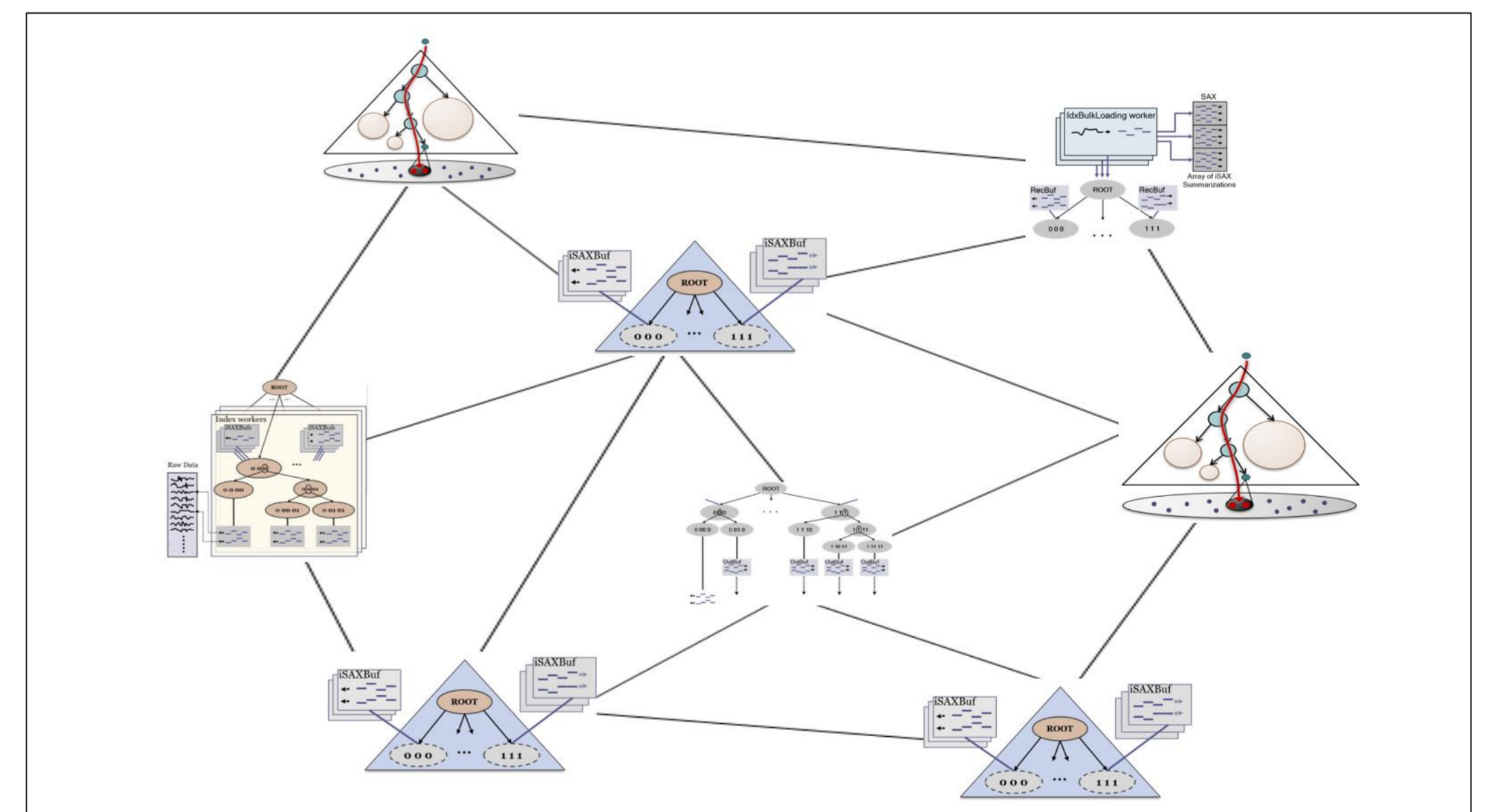
1. Design new algorithms and techniques for highly-efficient data series processing in a multi-node setting:
 - a. low-cost, **query-aware** data partitioning and mapping techniques for heterogeneous computing platforms
 - b. load balancing and communication primitives for **highly scalable and efficient multi-node query processing**
2. Develop new **fault-tolerant** multi-threading indexes and query processing schemes
3. Combine the power of general purpose CPUs with accelerators, such as **Graphical Processing Units (GPUs)**
4. Compile the different components to get a **new powerful index**:
 - a. highly-efficient and scalable
 - b. facilitate processing of orders of magnitude more data series than those tested by state-of-the-art indexes.



PLATON Innovation & Impact

BEYOND STATE-OF-THE-ART

1. **Advanced time-series processing index.** Exploiting (for the first time) the combined computational power of heterogeneous computing platforms (multiple nodes, CPUs, GPUs, SIMD) to build a revolutionary time-series index in terms of many dimensions (performance, scalability, fault-tolerance, etc.).
2. **Fault-tolerant processing technologies.** Design and implementation of novel mechanisms and algorithms for NUMA-aware and fault-tolerant data series indexing.
3. **Bridging the gap between technology and service creation.** Demonstrate the impact of developed software using real datasets.



ECONOMIC IMPACT

Domains that are aggressive data series producers:

- Construction of EU E-ELT telescope3: **1 billion €**
- EU Human Brain Project5: estimated **1.19 billion €**
- US Human Genome Project has a **\$796 billion** impact and created 310,000 jobs)

CONTACT

Prof. Panagiota Fatourou
University of Paris,
LIPADE
Tel.: +30 6973 991277
Email: faturu@csd.uoc.gr

MORE INFORMATION

www.ics.forth.gr/~faturu/platon/
www.facebook.com/Platon-Project-Marie-Sklodowska-Curie-Action-11653541305219
www.linkedin.com/groups/14013902/
twitter.com/PlatonProject