Parallel and Distributed Data Series Processing on Modern and Emerging Hardware

H.F.R.I.
Hellenic Foundation for Research & Innovation

European Commission

Panagiota Fatourou, Professor

University of Crete and FORTH

MARIE CURIE ACTIONS

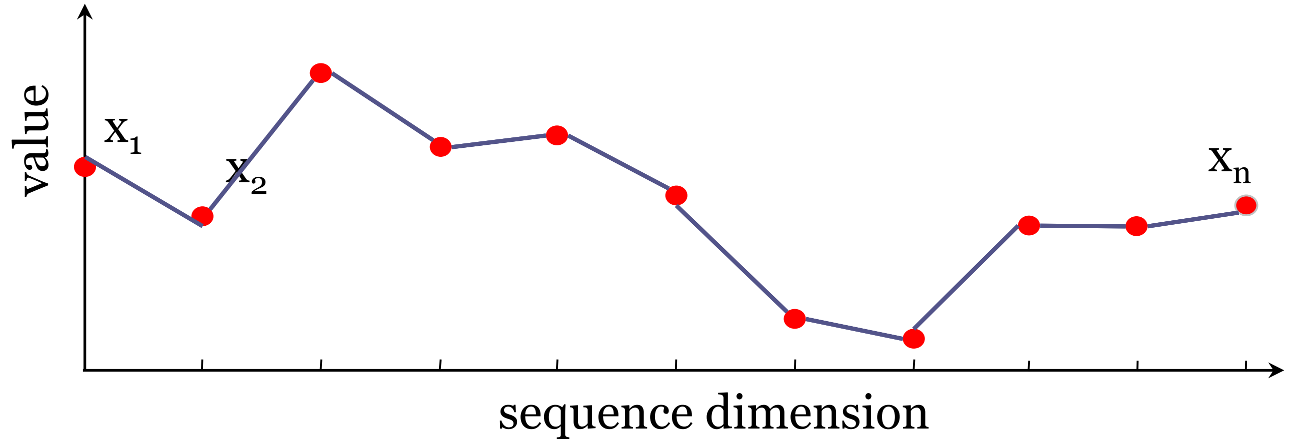
FORTH-ICS

data intelligence institute of Paris

Université de Paris

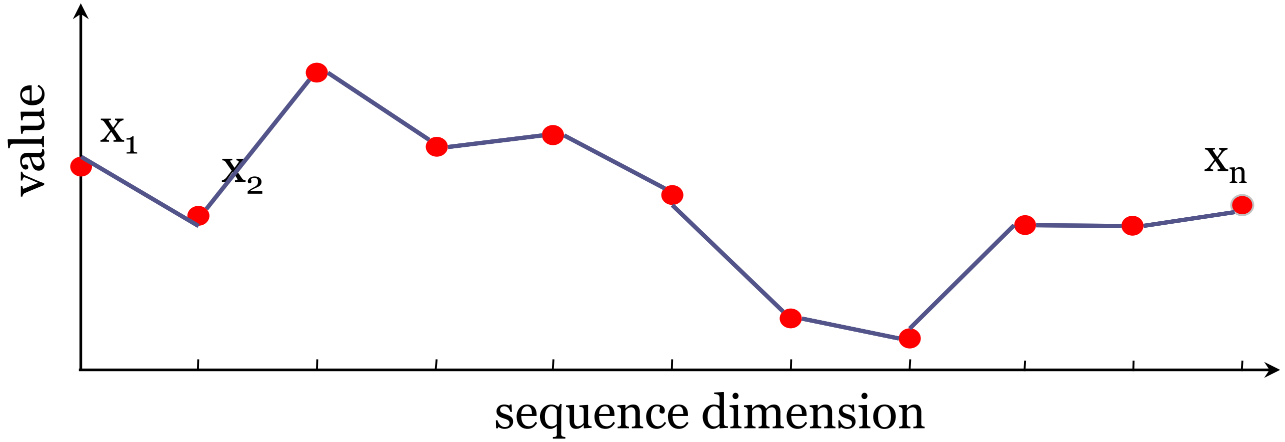
Data series

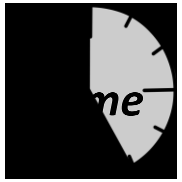
- Sequence of points ordered along some dimension



Data series

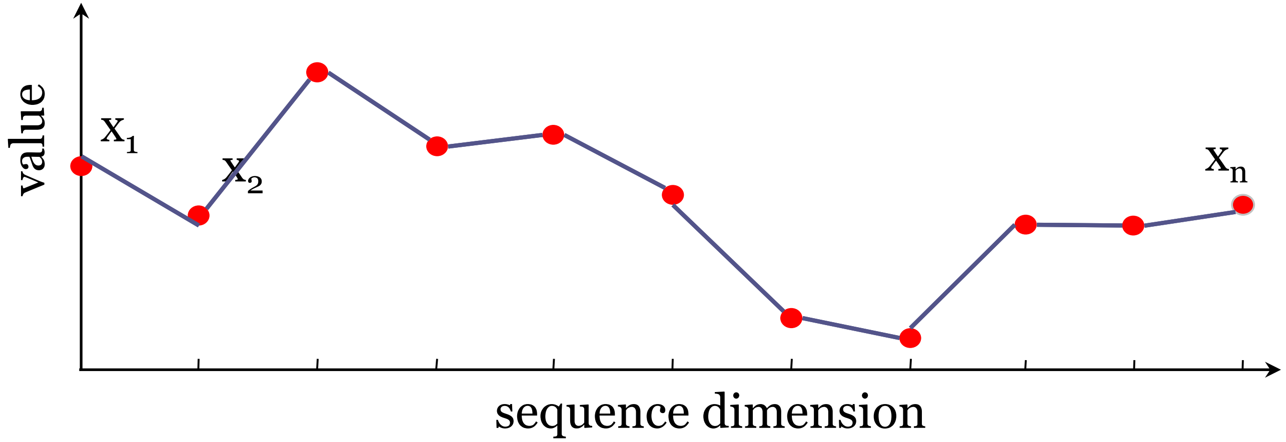
- Sequence of points ordered along some dimension

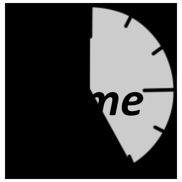


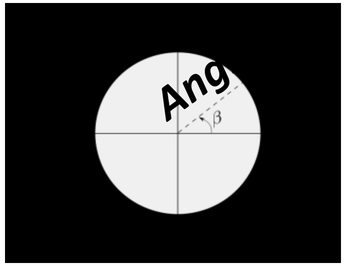


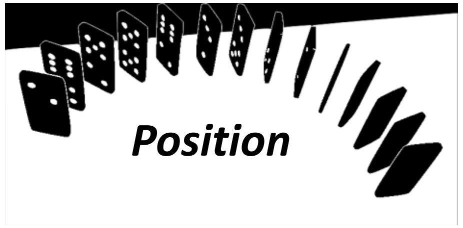
Data series

- Sequence of points ordered along some dimension

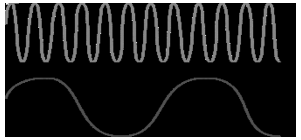








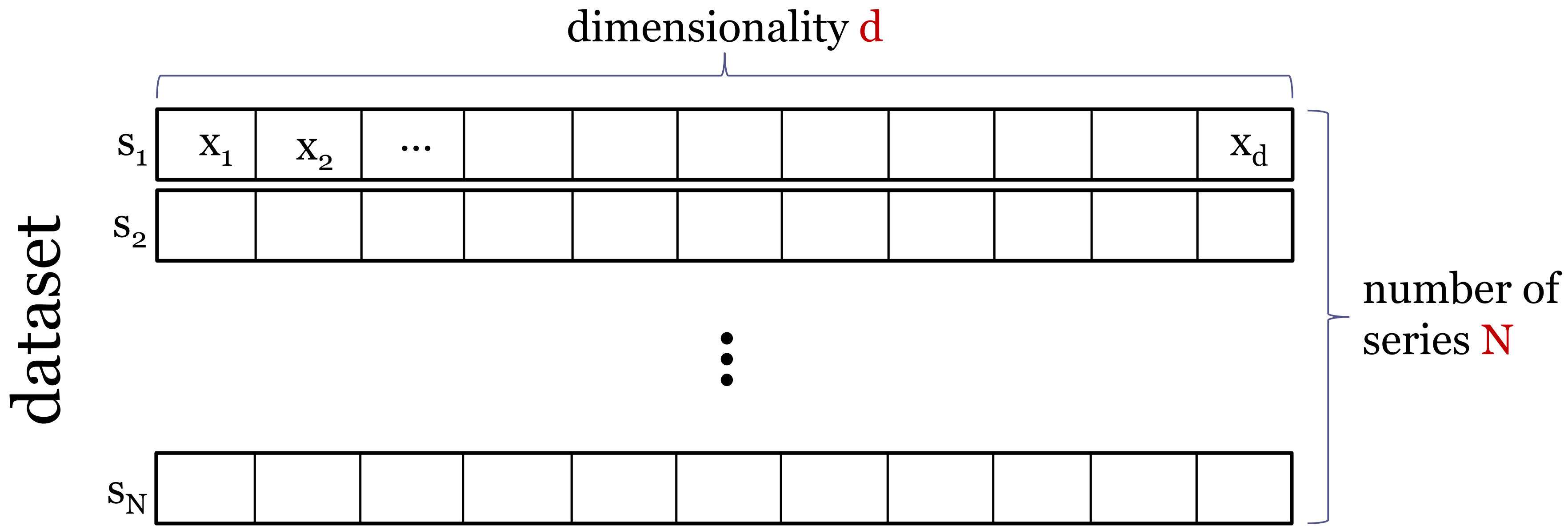
Mass



Frequency

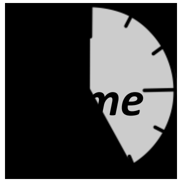
Data Series Collections

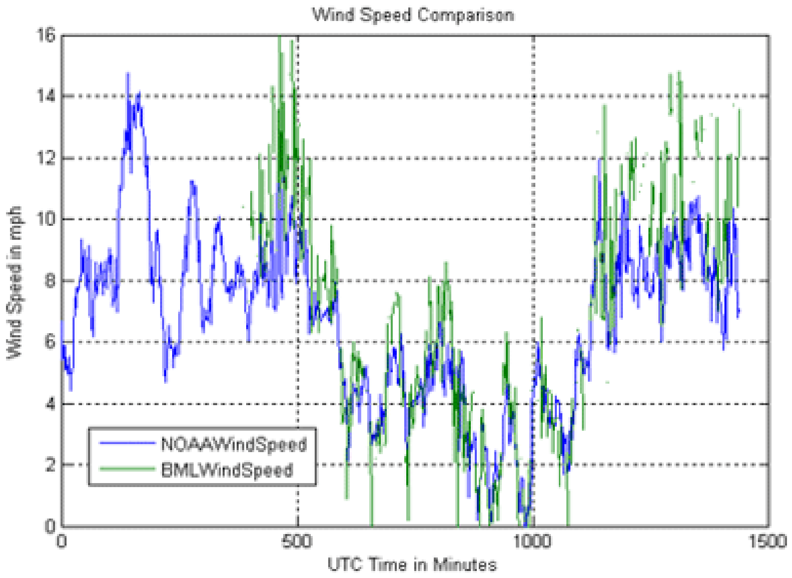
- represented as *N d*-dimensional vectors



Scientific Monitoring

meteorology, oceanography, volcanology, seismology, astronomy, finance, etc.





Wind speed

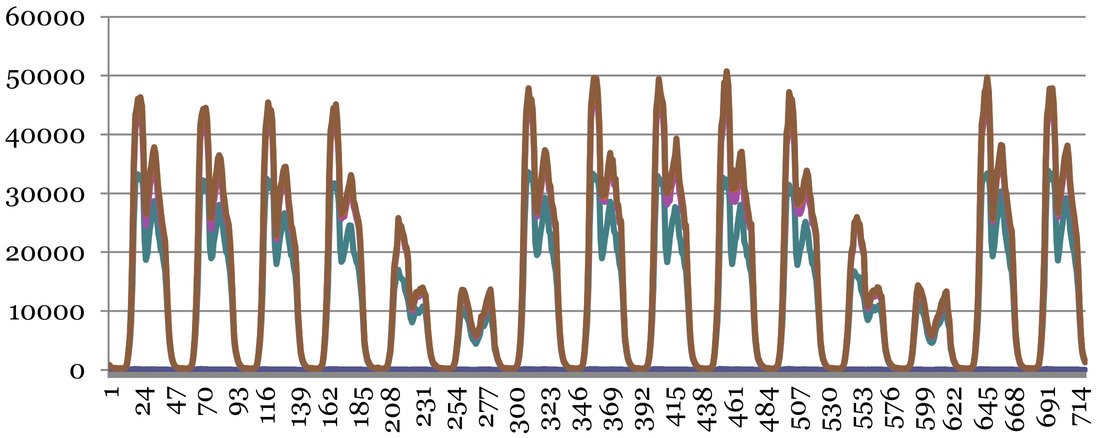
From ocean observing node project, http://bml.ucdavis.edu/boon/wind.html

Volcanic Activity Indicators

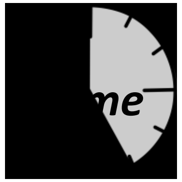
From British Geological Survey
https://www.bgs.ac.uk/geology-projects/volcanoes/

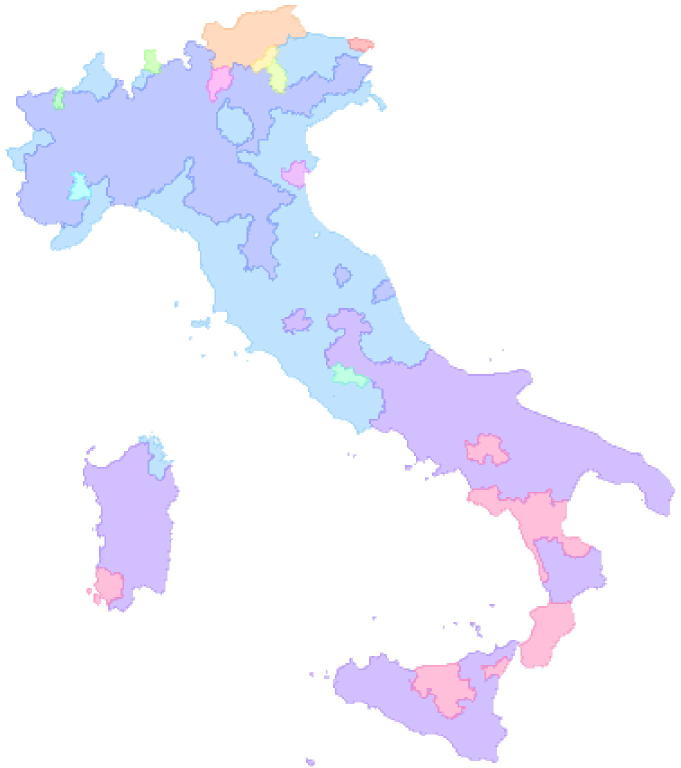
Telecommunications

- analysis of call activity patterns, Telecom Italia



average number of calls



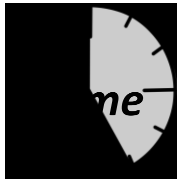


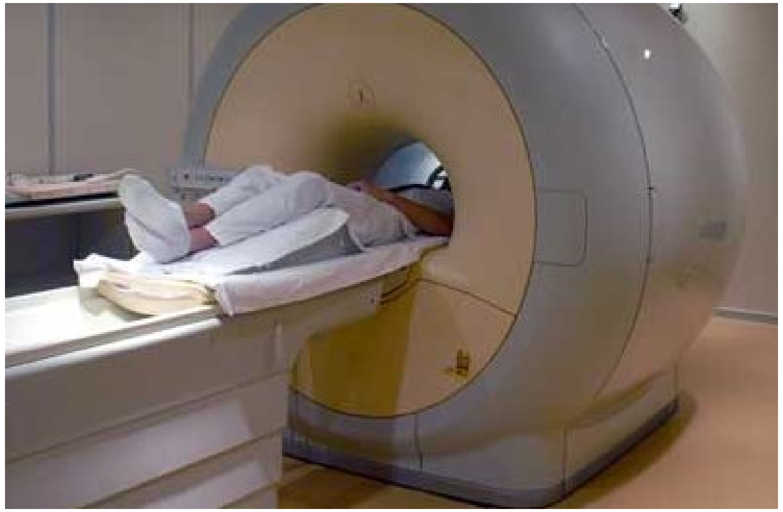
clustermap of incoming calls time series

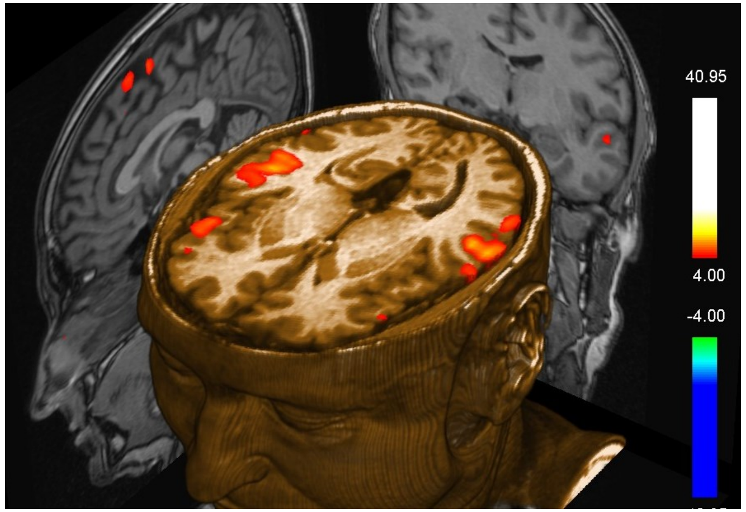
Paraskevopoulos et al. NetMob, Special Session on the Data for Development (D4D) Challenge, 2013.

Neuroscience

- functional Magnetic Resonance Imaging (fMRI) data
  - primary experimental tool of neuroscientists
  - reveal how different parts of brain respond to stimuli

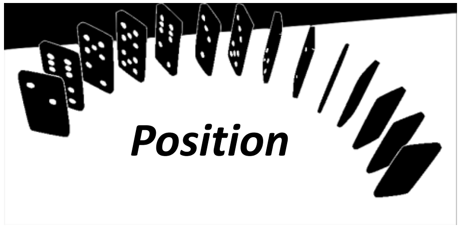


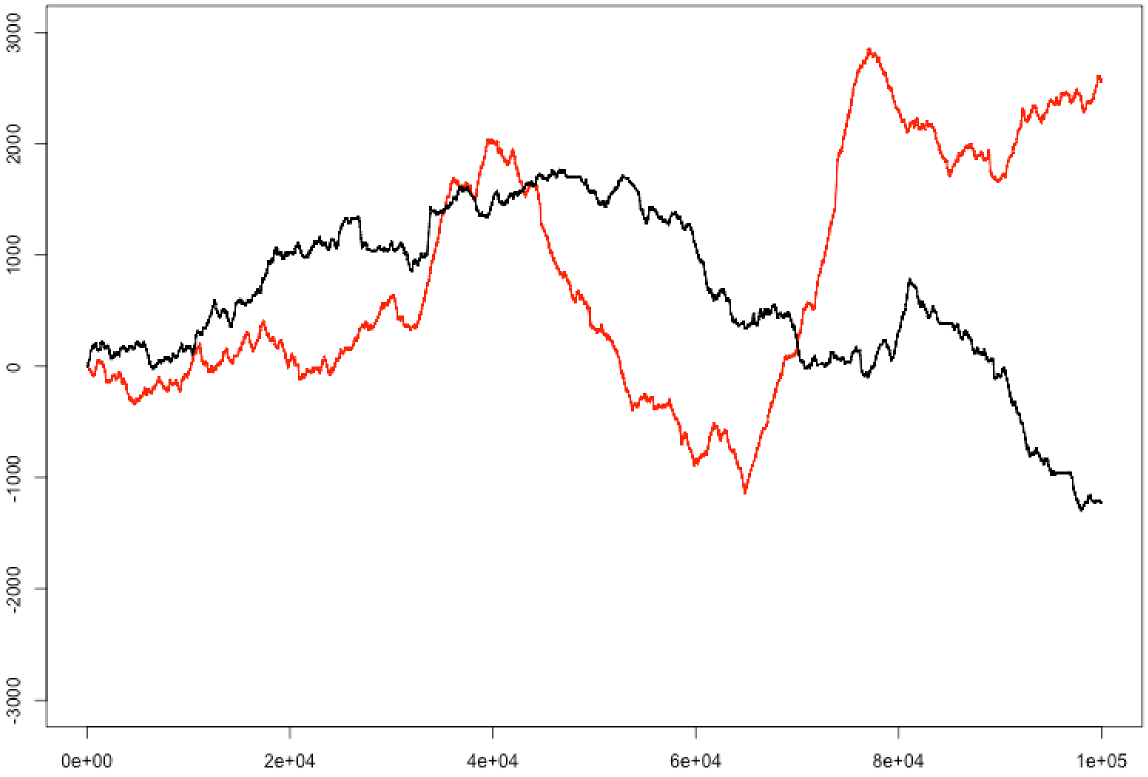




Biology

GTCAATGGCCAGGATATTAGAACAGTACTCTGTGAACCCTATTTATGGTGGCACCCCTTAG
ACTAAGAT GGAAGCCTCC
TGTAAAGA GGAAAGCCA
CATTTGAC AGATATTACA
AACTAAGA GAATAATCCC
TGGAGGTT CCGGTTGAA
ACTGAAAT ACTGGCAACT
GCCAACCA TCCCAAGGAT
CAACCTTC GGGCTTTGG
ACTGAAAT GCCAGTGACA
AATACAAC TTATGTCTTA
ACTTTGT GGCCAAGCAT
CCAAGGTT TGGTTATTTA
TTAGTCAC GAGCAGAATC
AATGAGCC CAAGATGAG
CATTGAAG AAGAGCCCA
GGTTGTTG ATGGCCAGG
ATATTAGAACAGTACTCTGTGAACCCTATTTATGGTAGCACCCCTTAGACTAAGATAACACA
GGGAGCAAGAGGTT





Data Series vs. high-d Vectors

Echihabi, Zoumpatianos, Palpanas. WIMS 2020.

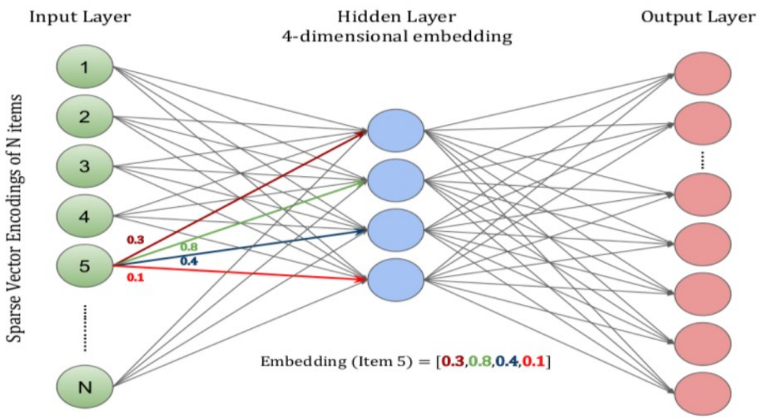
- two sides of the same coin
- data series techniques applicable to high-d vectors, too!

Data Series vs. high-d Vectors

Echihabi, Zoumpatianos, Palpanas. WIMS 2020.

- two sides of the same coin
- data series techniques applicable to high-d vectors, too!

molecules
chemical compounds
images
graphs
text
…

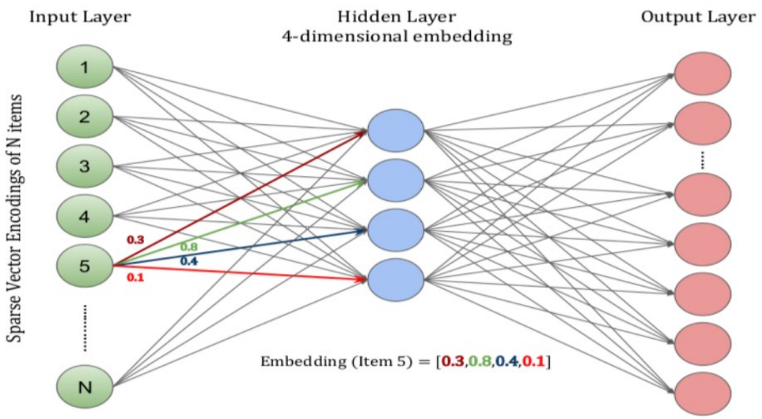


Data Series vs. high-d Vectors

Echihabi, Zoumpatianos, Palpanas. WIMS 2020.

- two sides of the same coin
- data series techniques applicable to high-d vectors, too!

molecules
chemical compounds
images
graphs
text
...



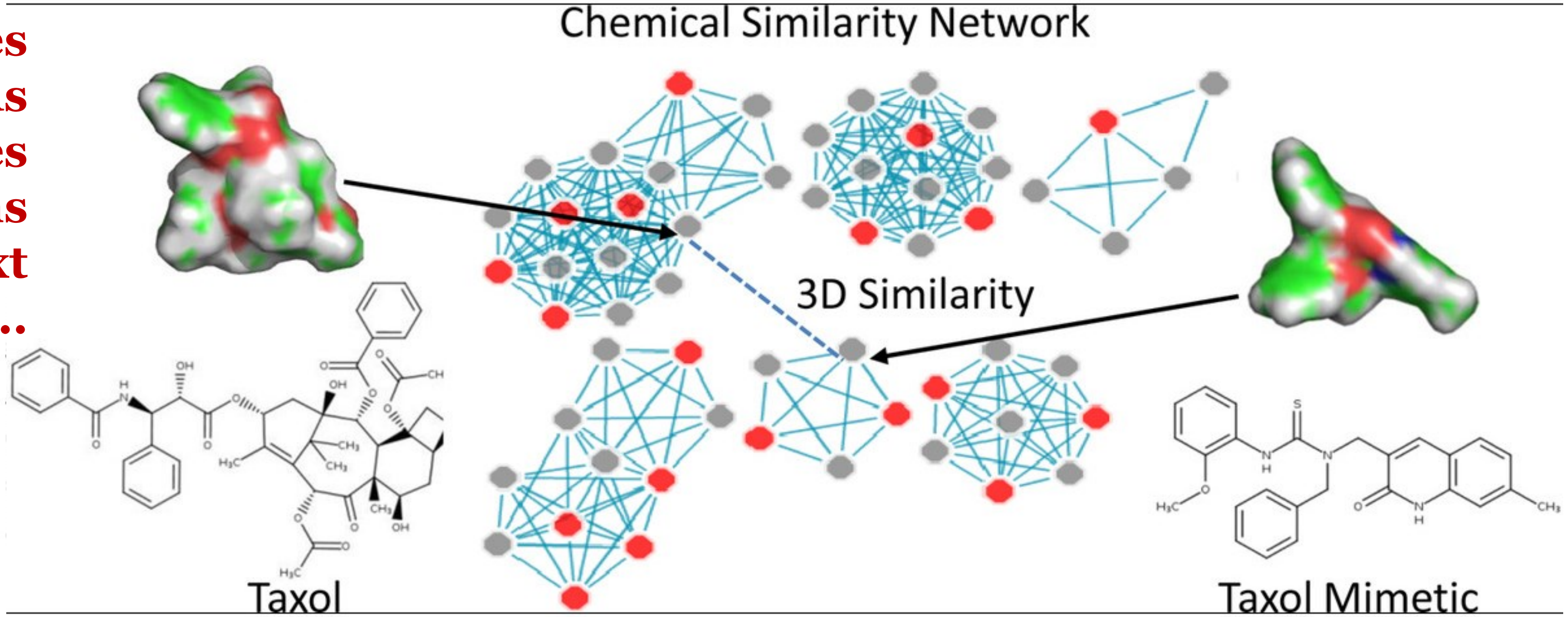
deep embeddings
high-d vectors learned using a DNN

Echihabi, Zoumpatianos, Palpanas. WIMS 2020.

Data Series vs. high-d Vectors

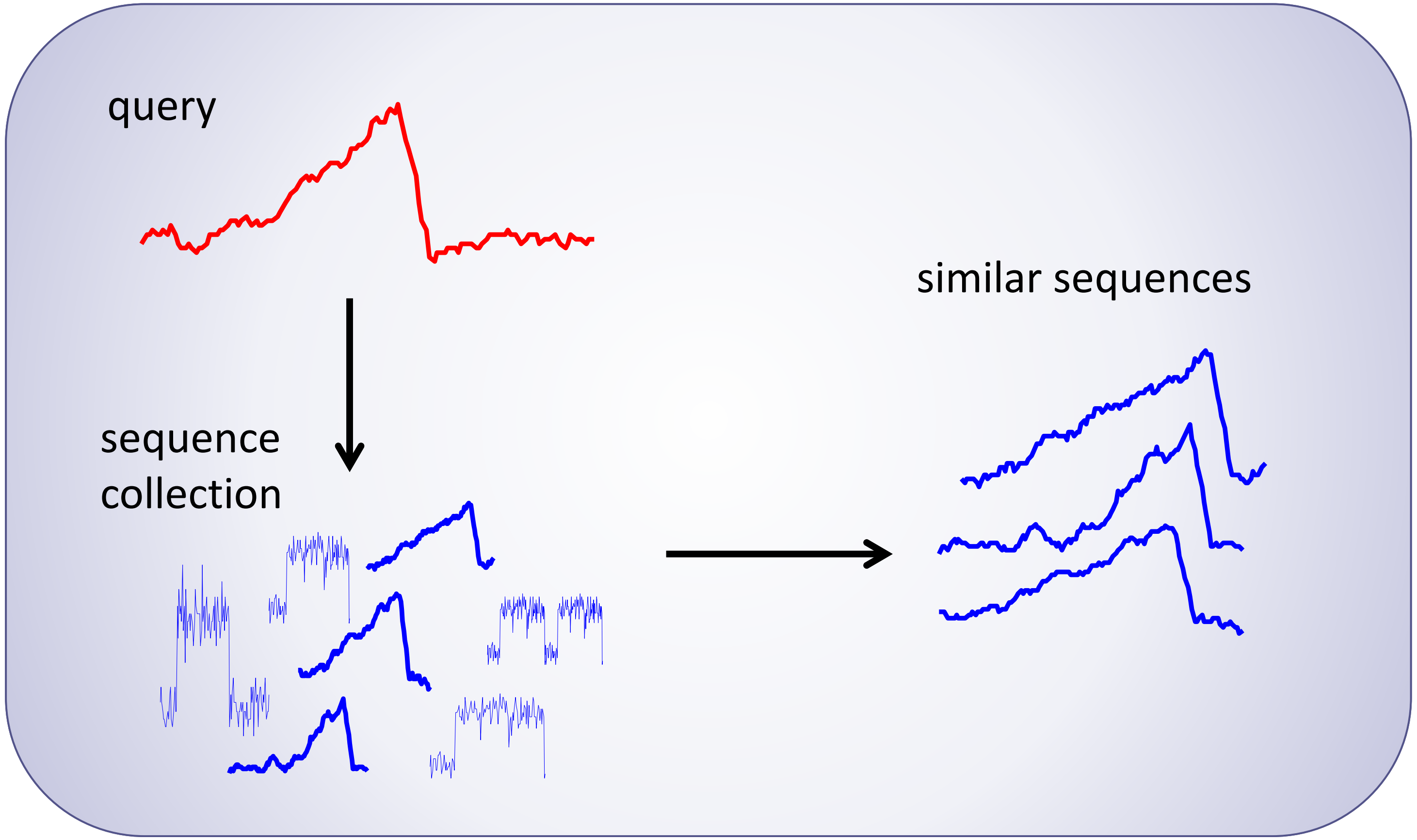
- two sides of the same coin
- data series techniques applicable to high-d vectors, too!

molecules
chemical compounds
images
graphs
text
…

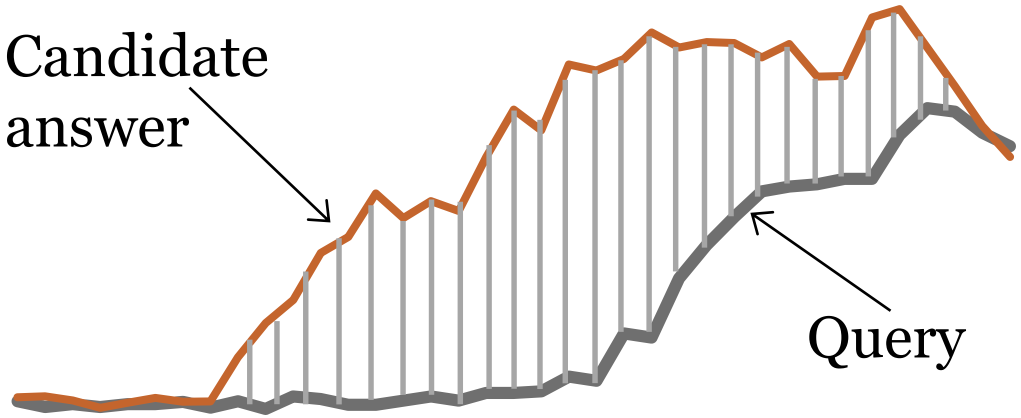


What do we want to do with data series?

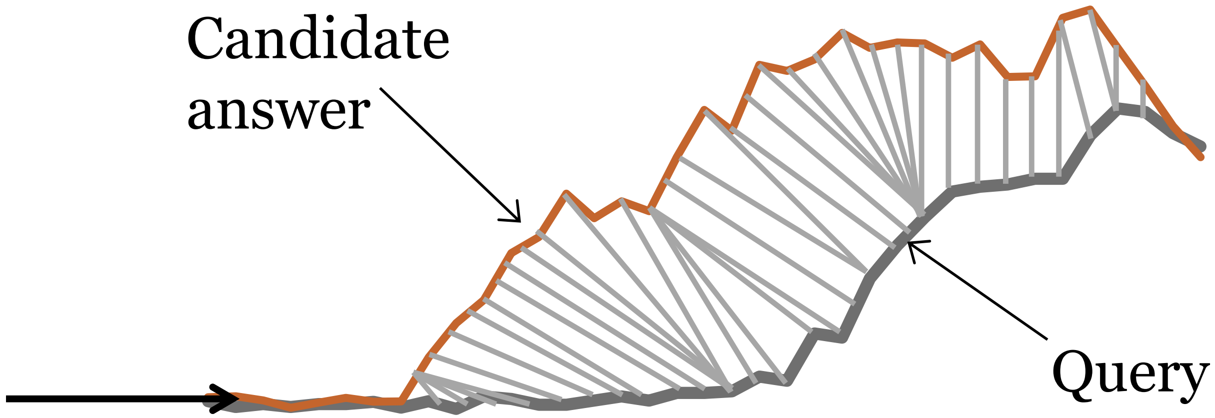
Similarity Search



What do we want to do with data series? Complex analytics



Euclidean Distance



Dynamic Time Warping

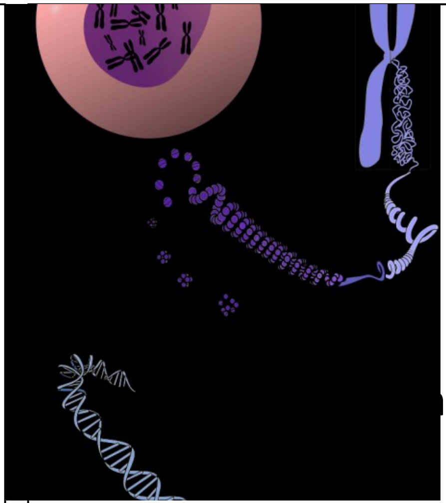
Challenge - Massive data Series collections

NASA's Solar Observatory

1.5 TB per day

Large Synoptic Survey Telescope

~30 TB per night



an Genome project

130 TB

passenger aircrafts

20 TB per hour

data center and services monitoring

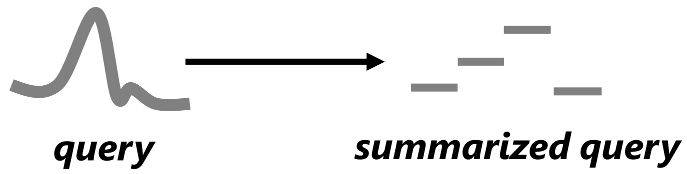
2B data series

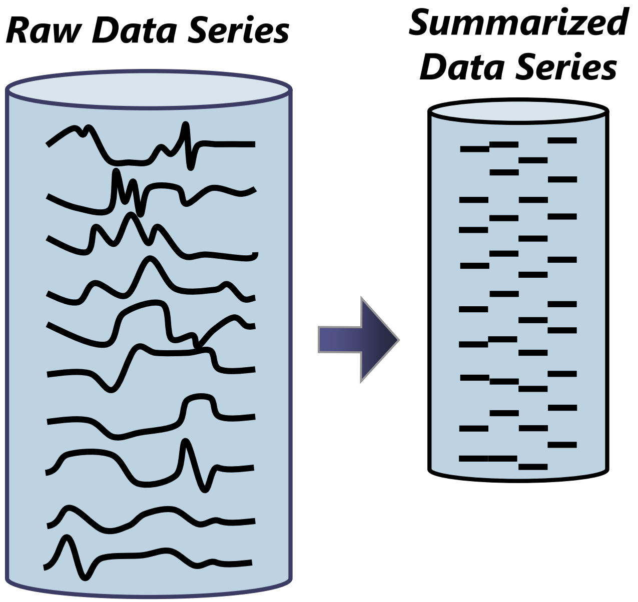
4M points/sec

HARD, because of very high dimensionality: each data series has several hundreds to several thousands of points!

ADS, Adaptive Data Series Index (State-of-the-Art Sequential)

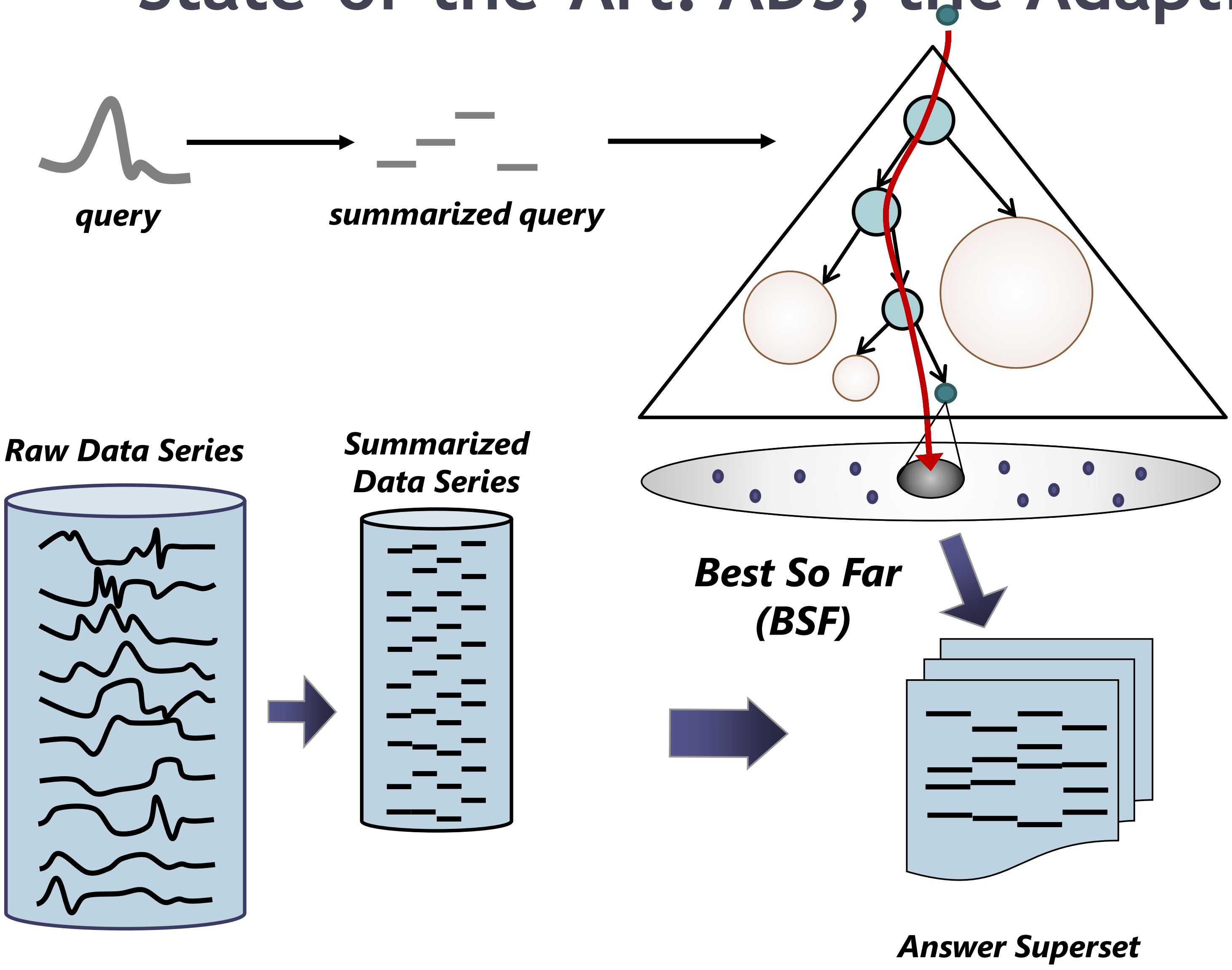
Zoumpatianos, Idreos, Palpanas, VLDB Journal, 2016





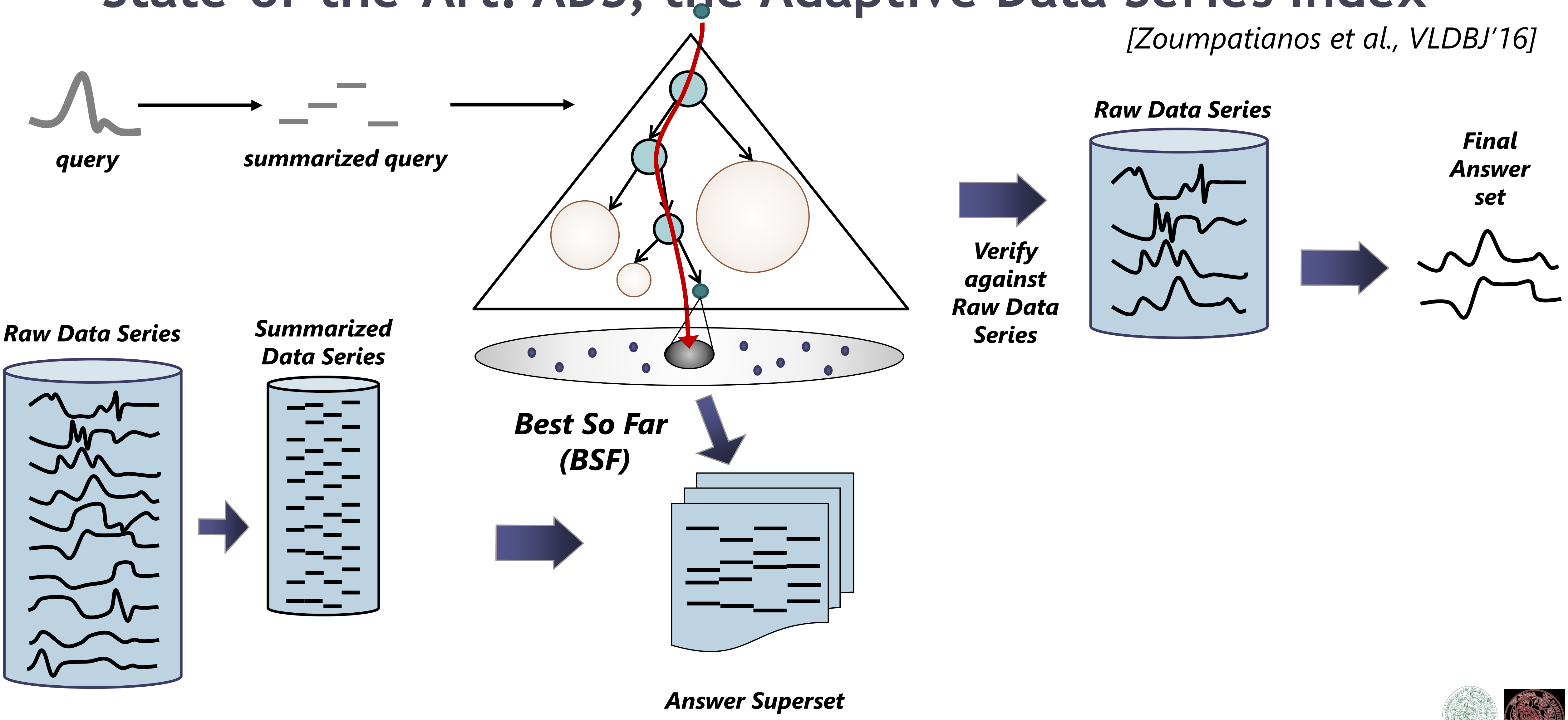
State-of-the-Art: ADS, the Adaptive Data Series Index

[Zoumpatianos et al., VLDBJ'16]



State-of-the-Art: ADS, the Adaptive Data Series Index

[Zoumpatianos et al., VLDBJ'16]



Symbolic Aggregate approXimation (SAX)

Shieh, Keogh. ISAX: Indexing and Mining Terabyte Sized Time Series. KDD 2008

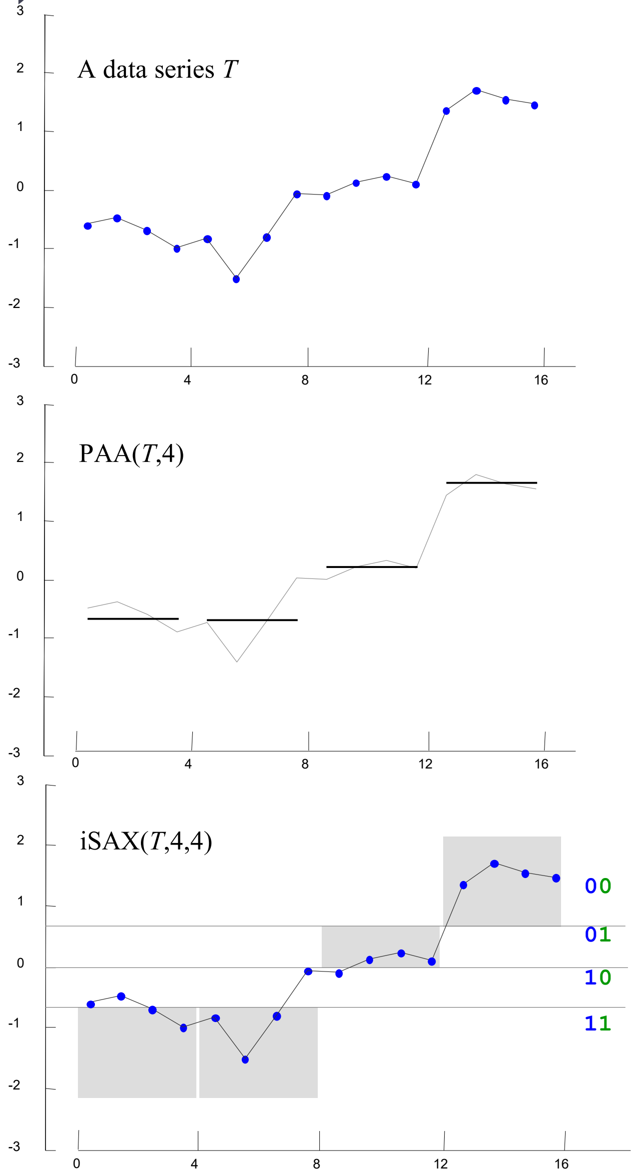
Represent data series T of length $\boldsymbol{n}$ with $\boldsymbol{w}$ segments using Piecewise Aggregate Approximation (PAA)

- PAA(T,w) = $\bar{T} = \bar{t}_1, \ldots, \bar{t}_w$

  where $\bar{t}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} T_j$

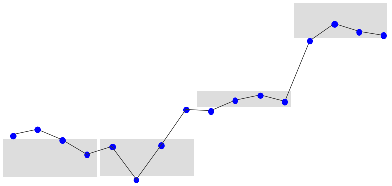
Discretize into a vector of symbols

- Breakpoints map to small alphabet $\boldsymbol{a}$ of symbols



iSAX Representation

- iSAX offers a bit-aware, quantized, multi-resolution representation with variable granularity



= { 6, 6, 3, 0} = {110 ,110 ,0111 ,000}

The number of bits can be different for each region

- Enables the creation of an hierarchical index tree (iSAX-based index tree)

iSAX Representation

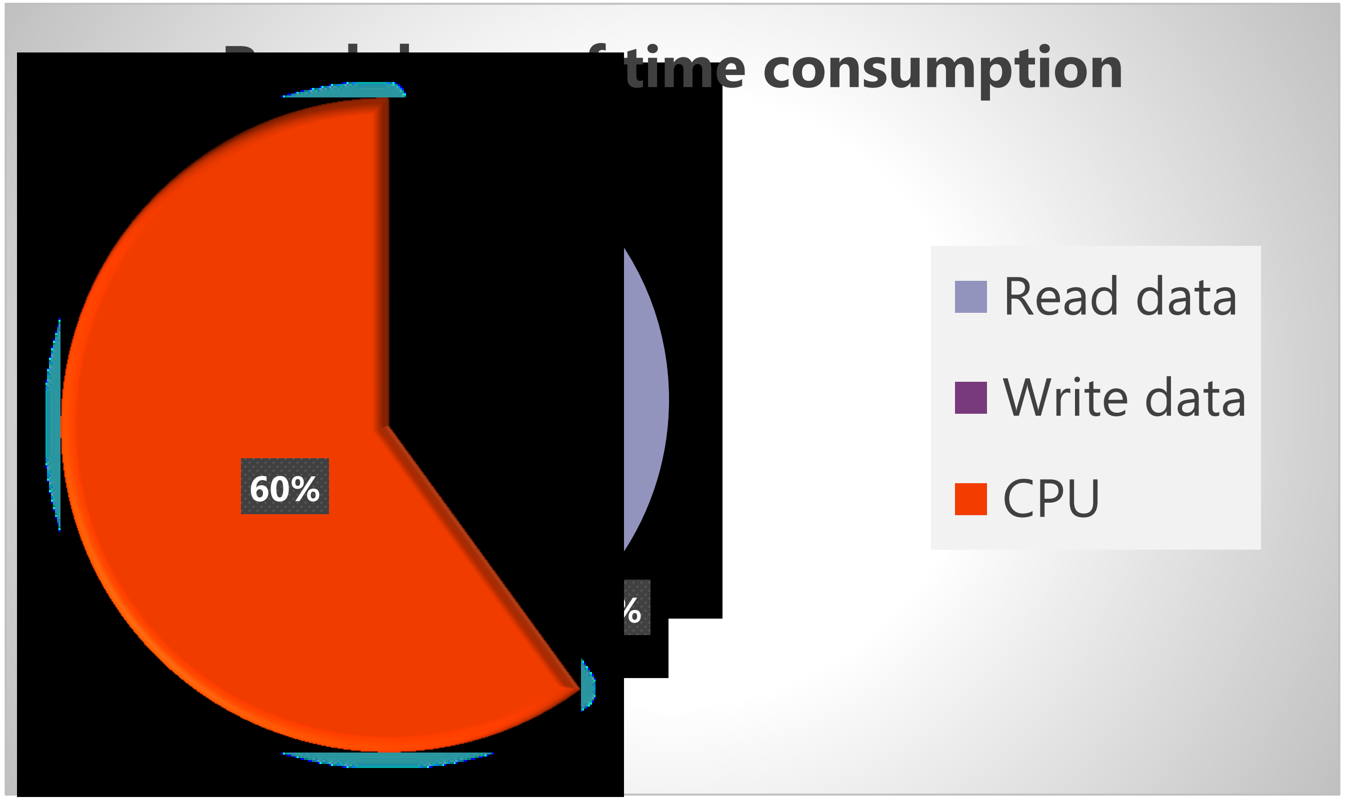
Lower Bound distance Calculation: Calculate distance between the iSAX summary of a data series and the query PAA

Real Distance Calculation: Calculate real Euclidean distance between the query and a data series

Lower-Bound Property of iSAX summaries: If the lower bound distance between a query Q and the data series DS is higher than a value v, then the real distance between Q and DS is also higher than v.

Disk-Based Indexes

ADS Index Creation



~60% of time spent in CPU: potential for improvement!

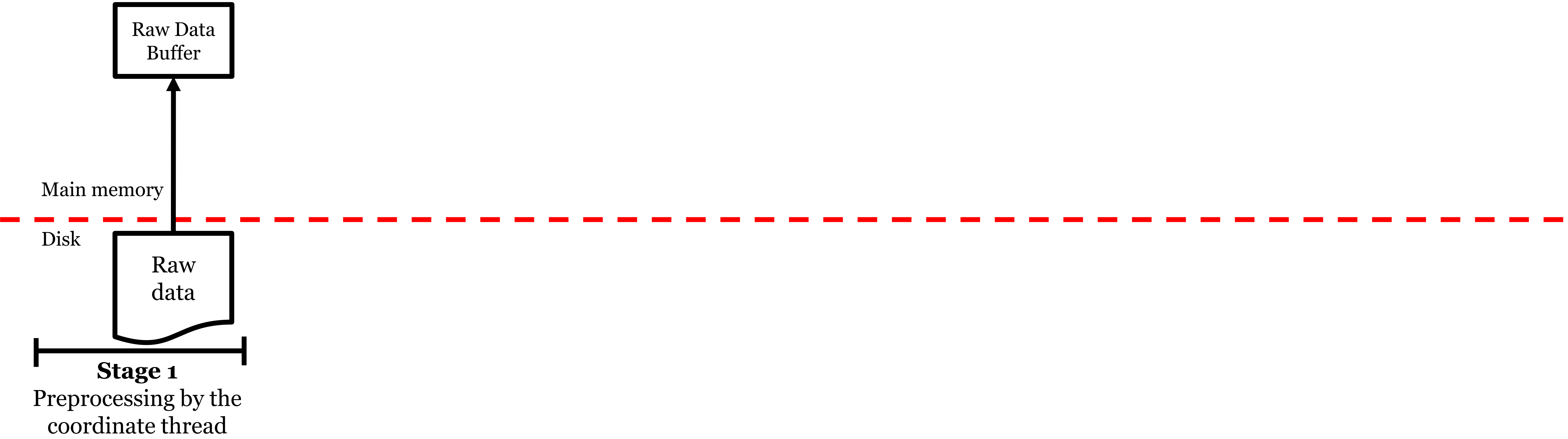
ParIS+: Parallel Indexing of Sequence

- Completely mask the CPU latency during index creation

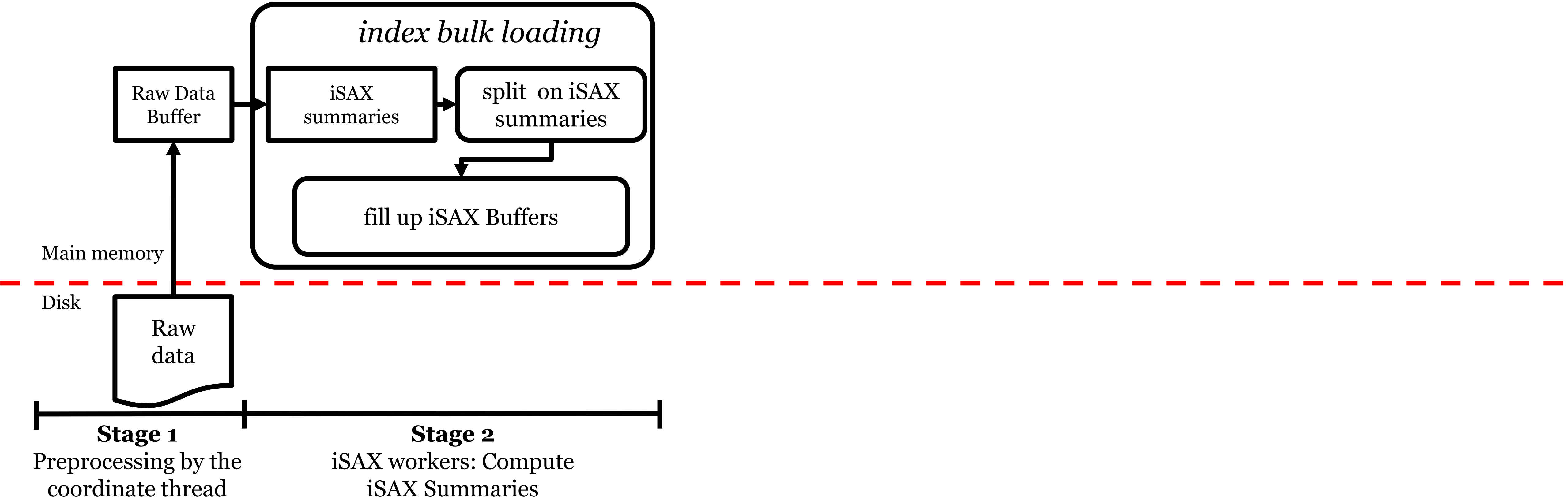
- Query answering 1-3 orders of magnitude faster than previous approaches

Peng, Fatourou, Palpanas
IEEE Trans. on Knowledge & Data Engineering 2021
IEEE Big Data 2018

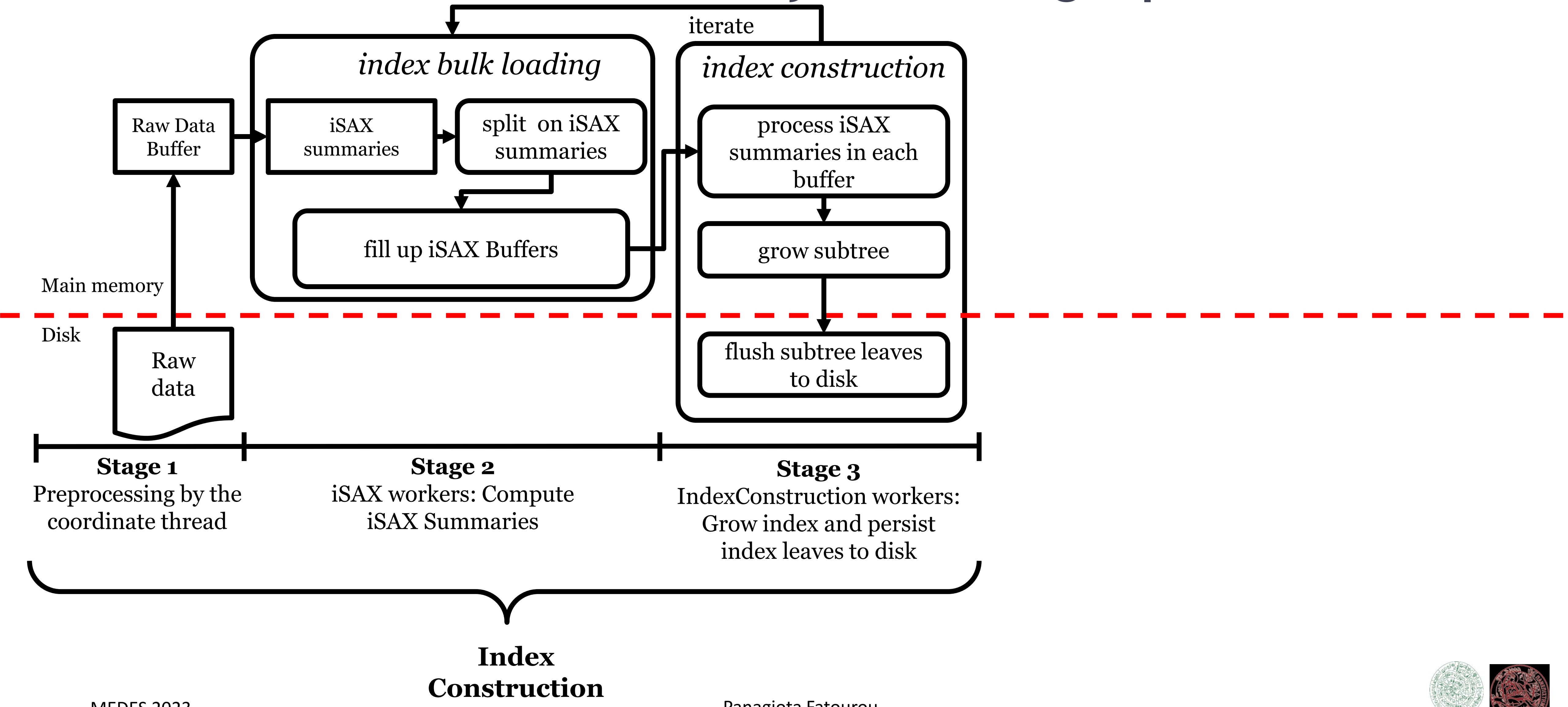
Index Creation and Query Answering Pipeline



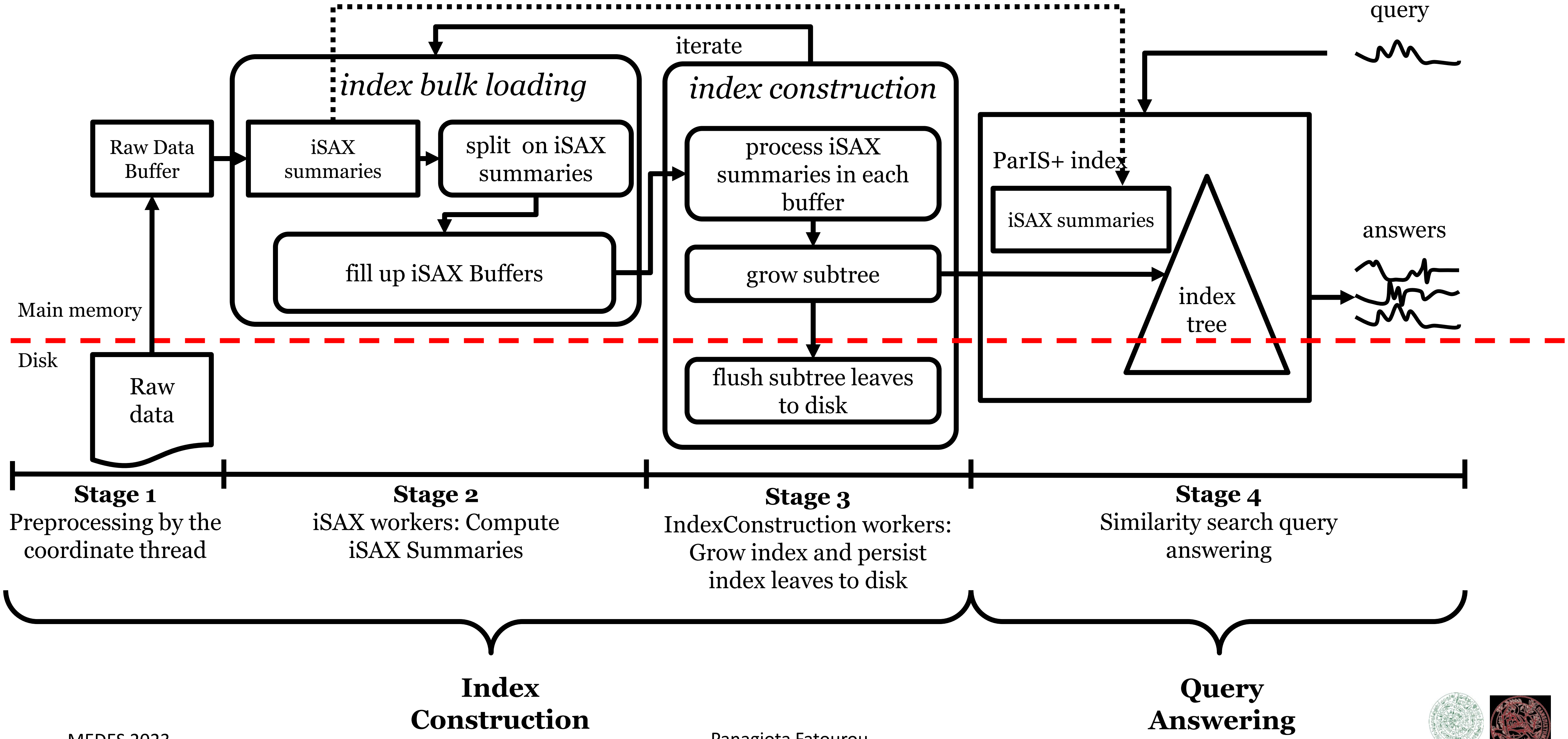
Index Creation and Query Answering Pipeline



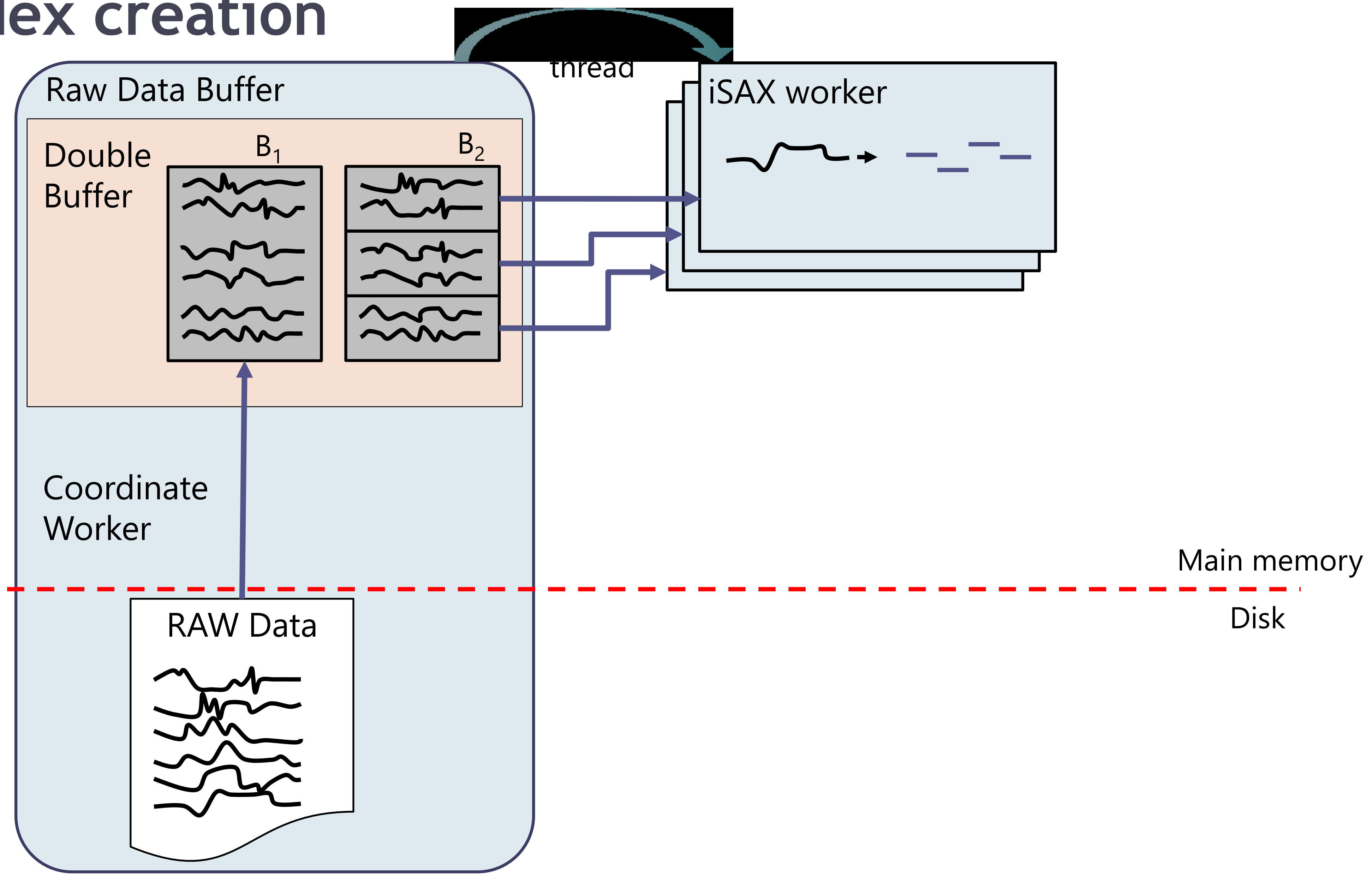
Index Creation and Query Answering Pipeline



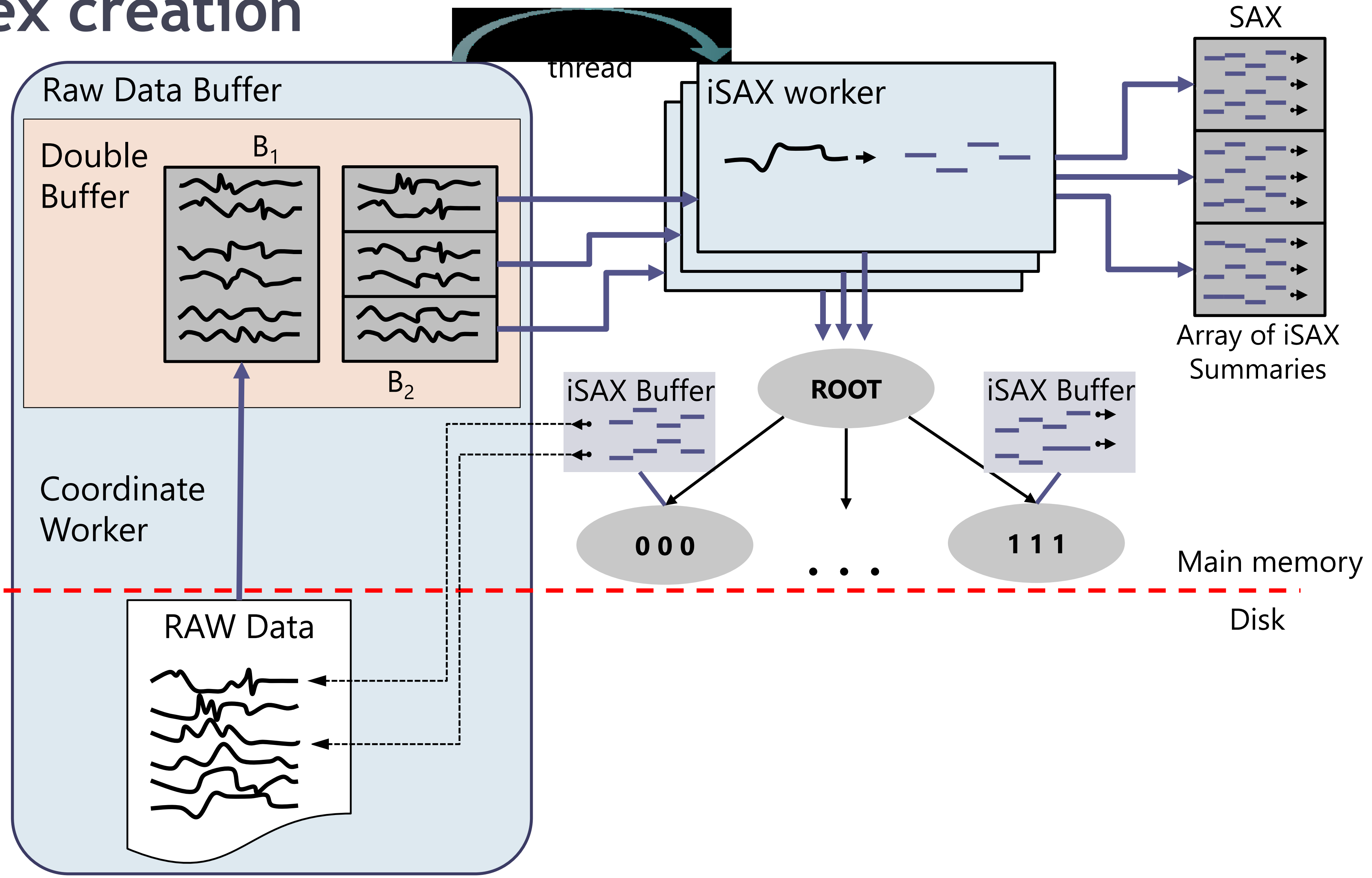
Index Creation and Query Answering Pipeline



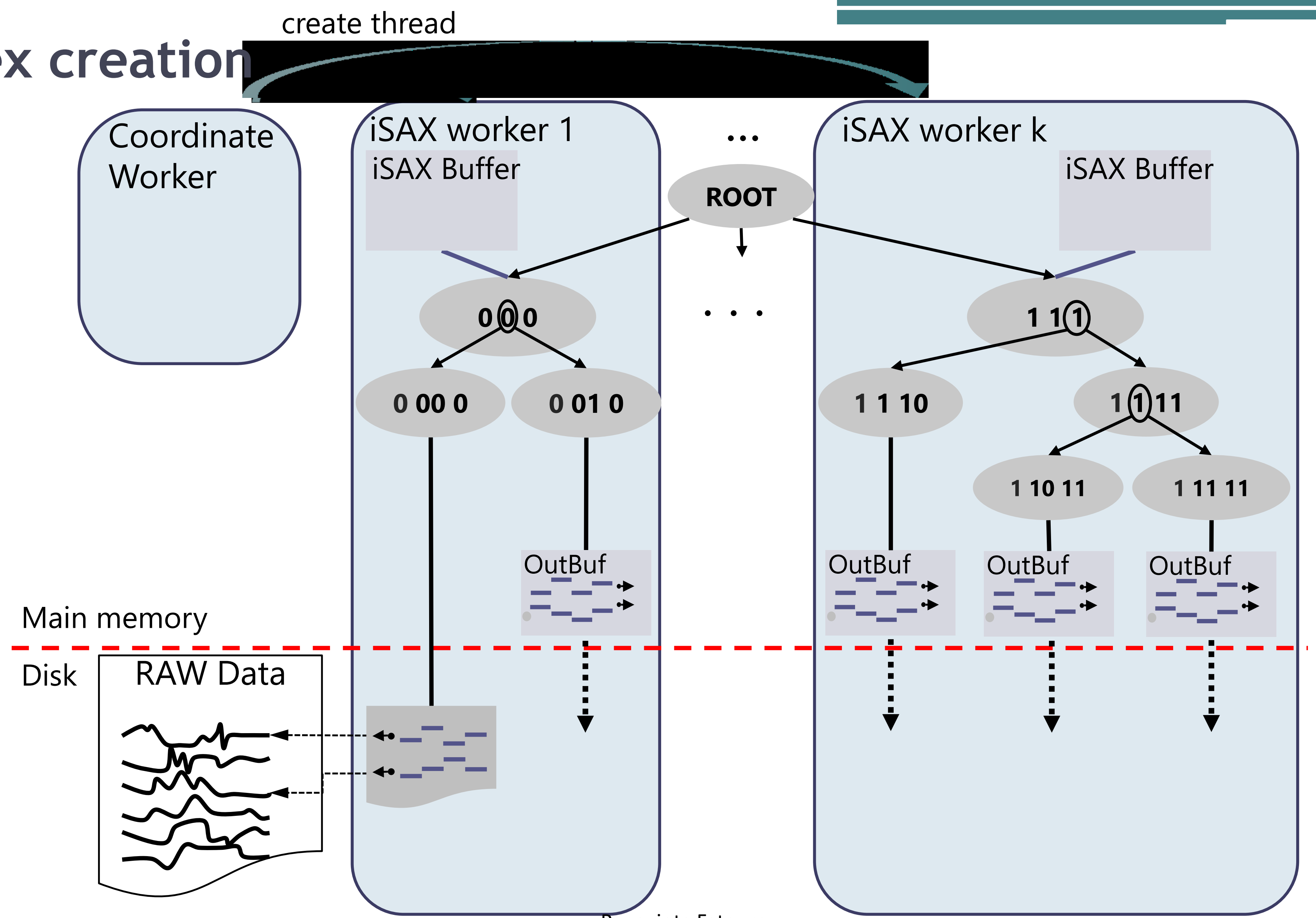
Index creation



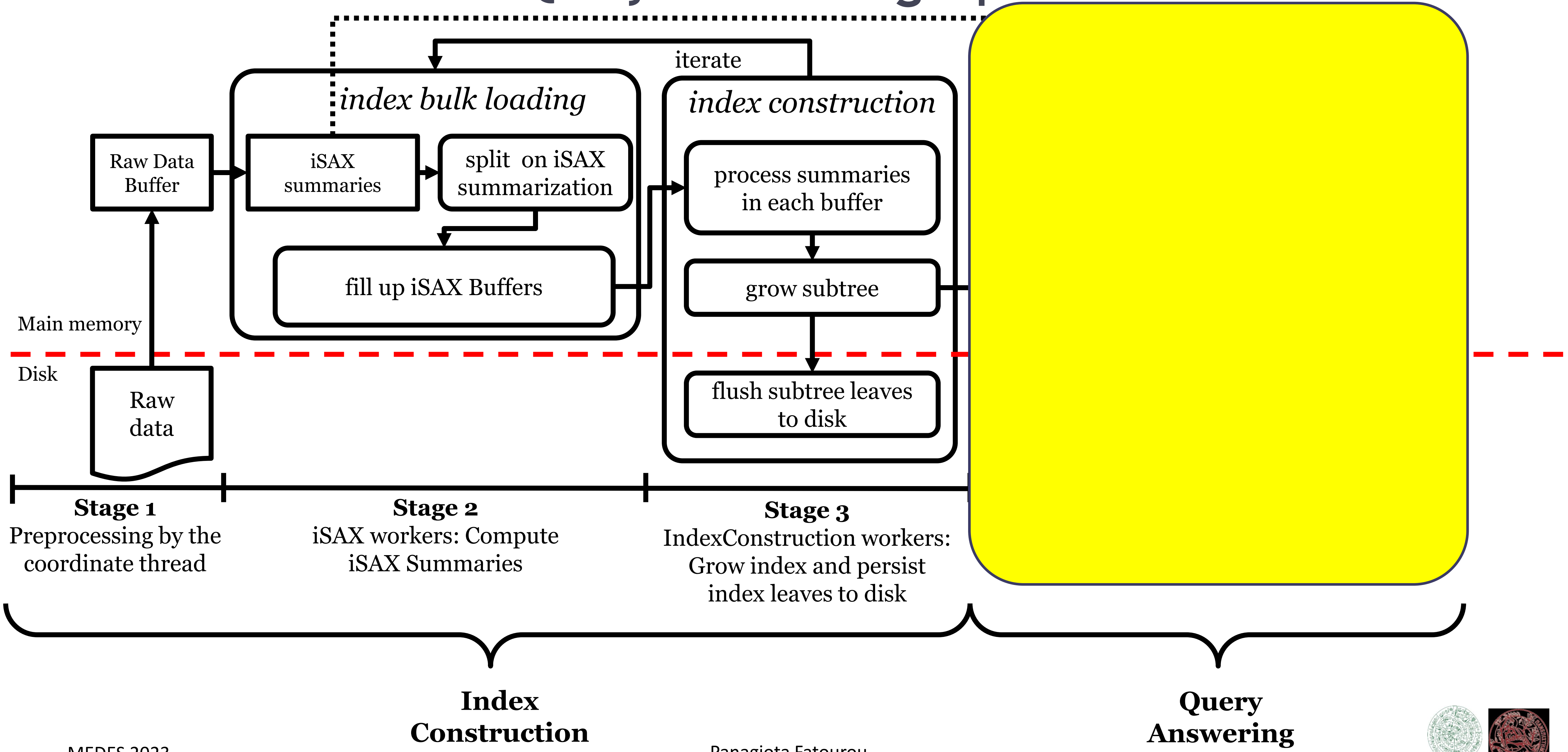
Index creation



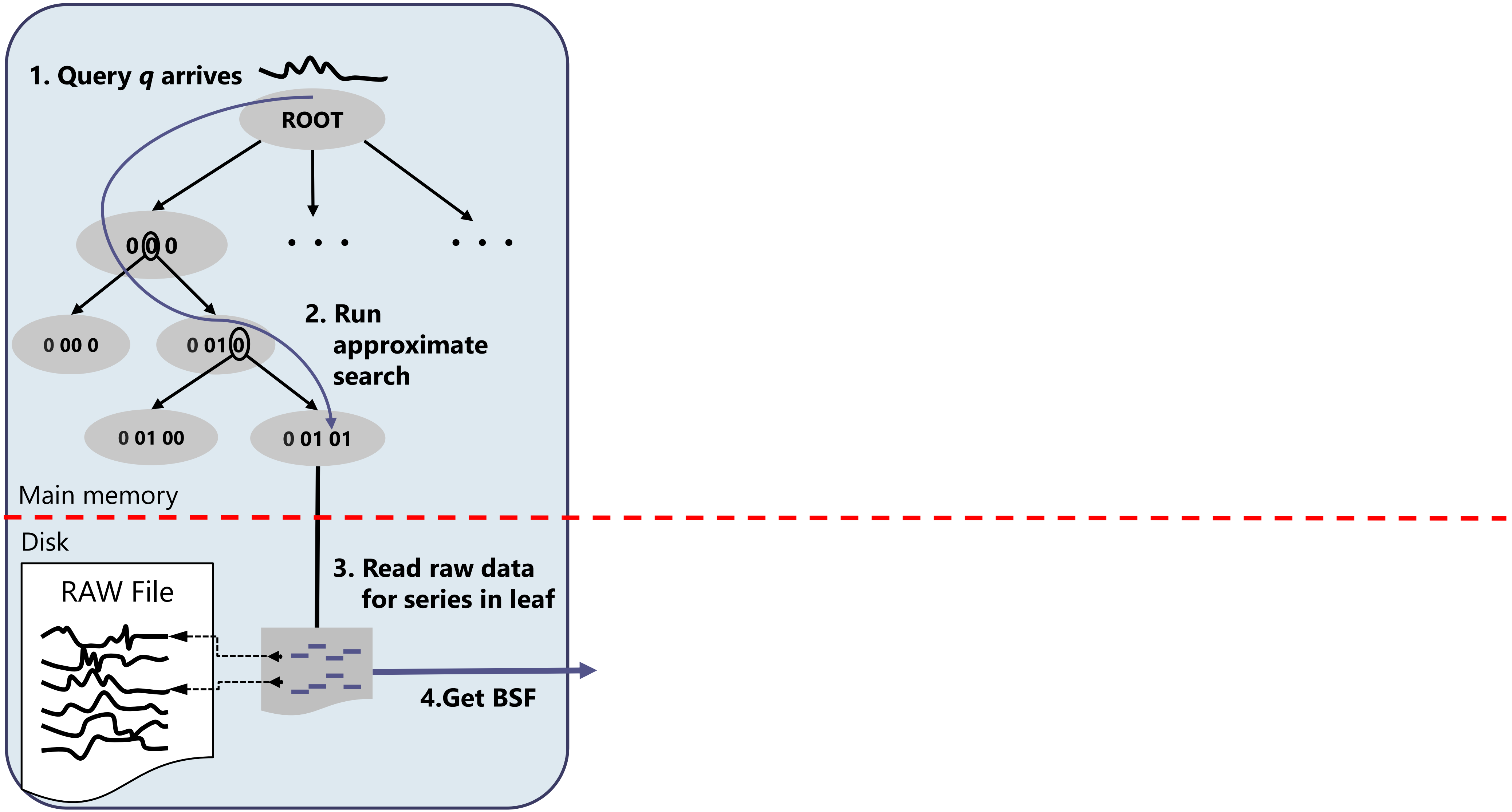
Index creation



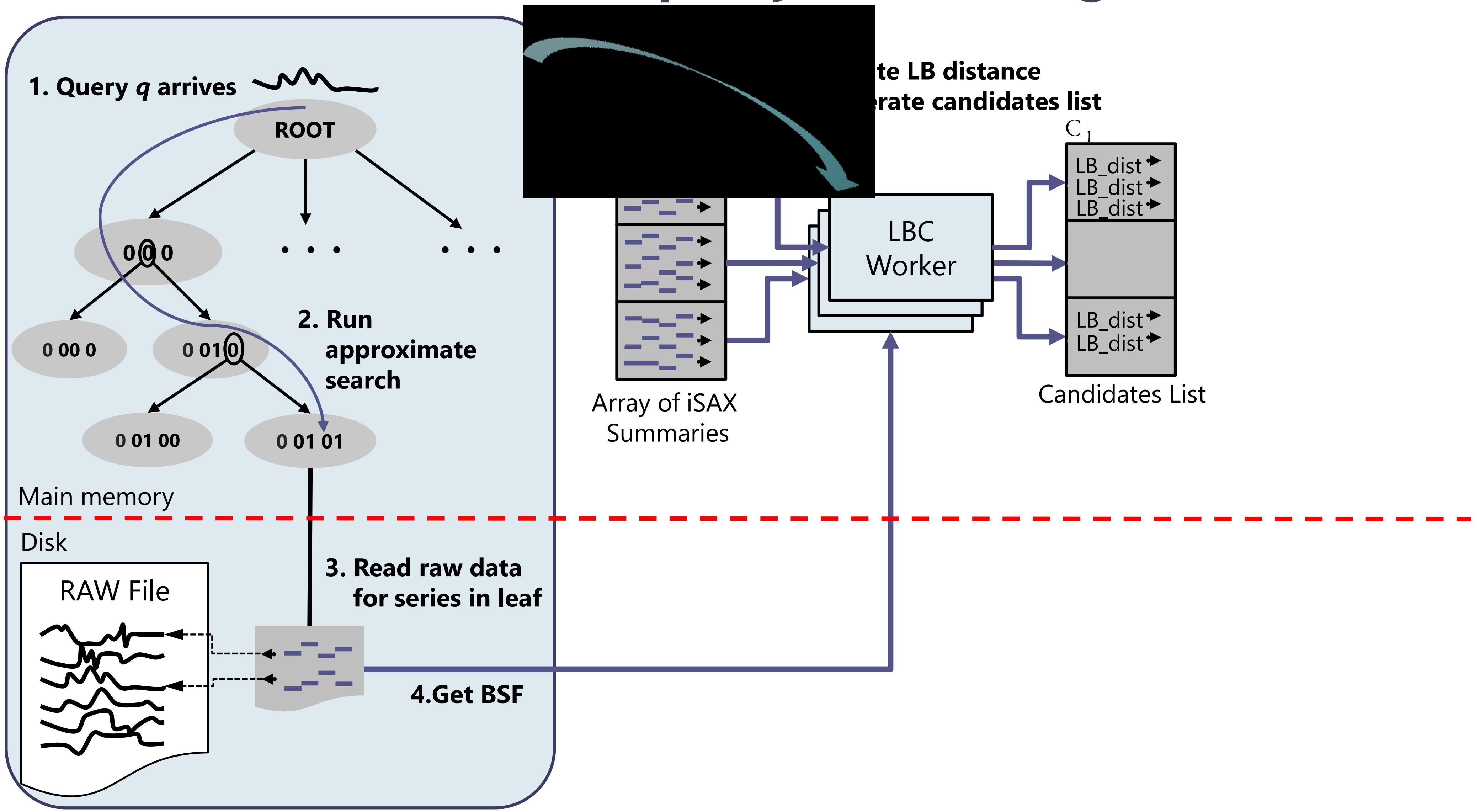
Index Creation and Query Answering Pipeline



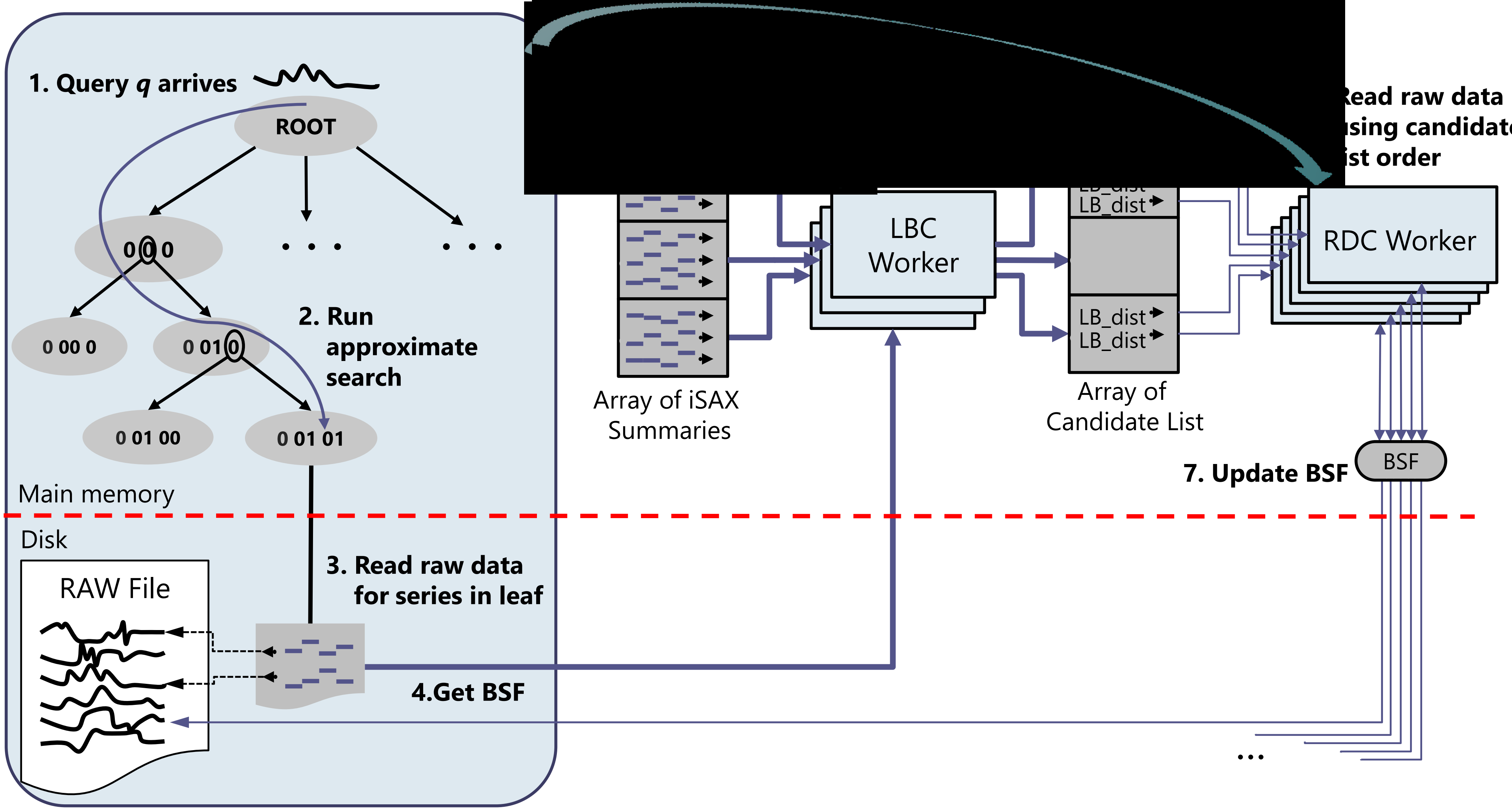
ParIS+ exact query answering



ParIS+ exact query answering



ParIS+ ex[redacted]



Experimental Setup

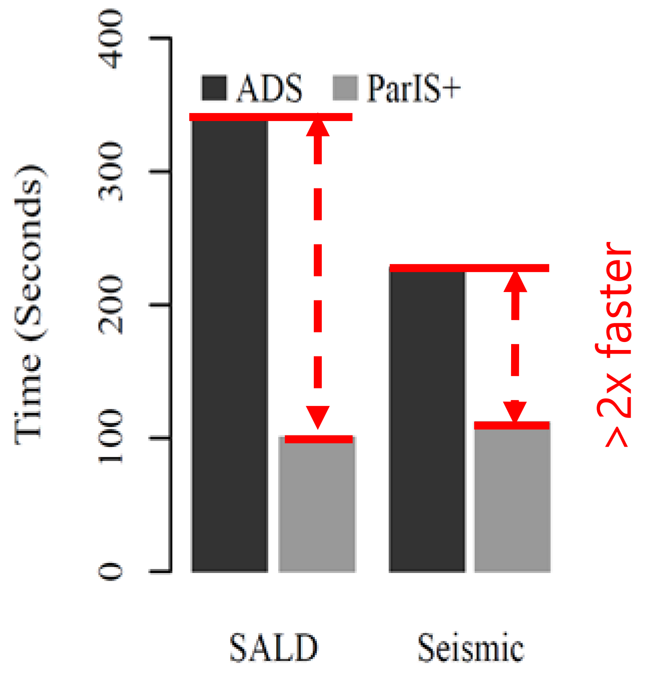
Machine: 2x Intel Xeon E5-2650 v4, 12 cores (24 hyper-threads)

Datasets: Random Dataset (50GB-200GB; series length: 256)
SALD Dataset (100GB; series length: 128): neuroscience MRI
Seismic Dataset (100GB; series length: 256): seismic signals

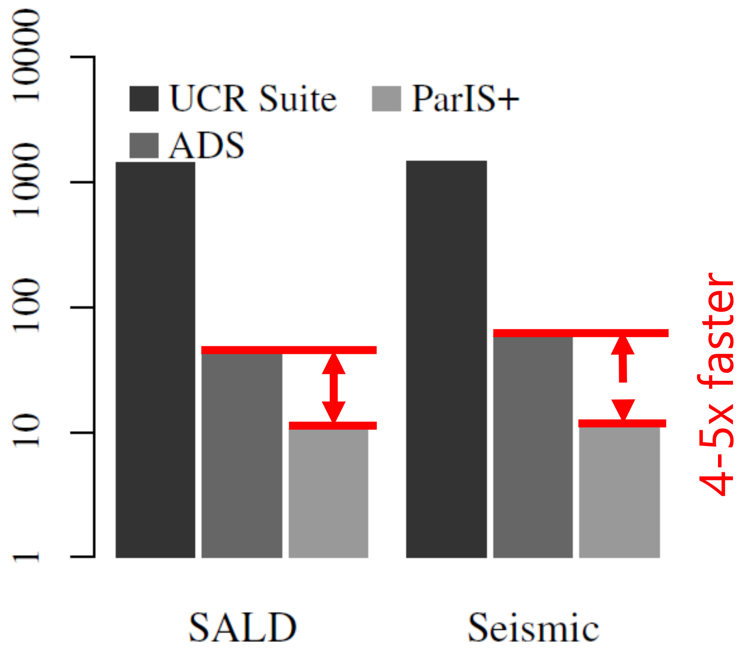
Algorithms: UCR Suite: serial scan method
ADS: disk-based index
ParIS+

Experiments

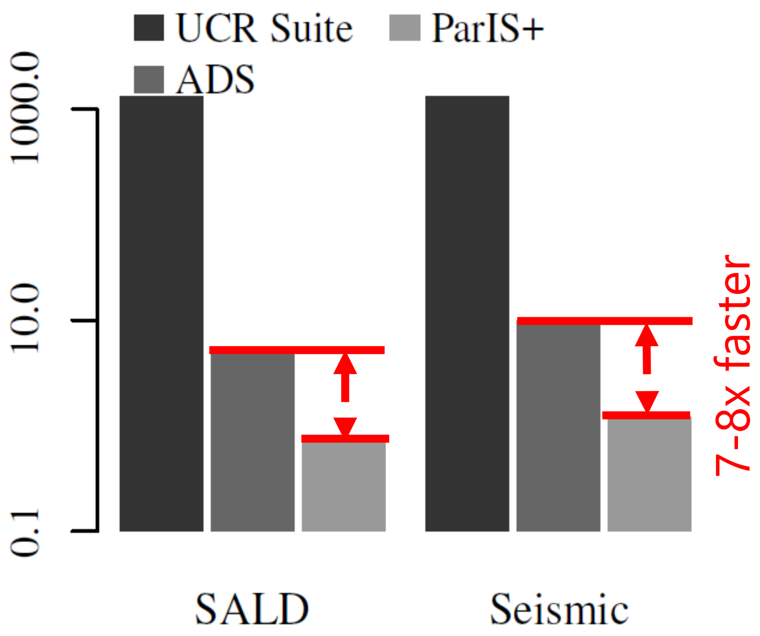
Time performance of indexing and exact query answering on real datasets



HDD



HDD



SSD

index creation: ParIS+ is >2x times faster than ADS

query answering HDD: ParIS+ is 4-5x faster than ADS

query answering SSD: ParIS+ is 7-8x faster than ADS

In-Memory Indexes

Motivation

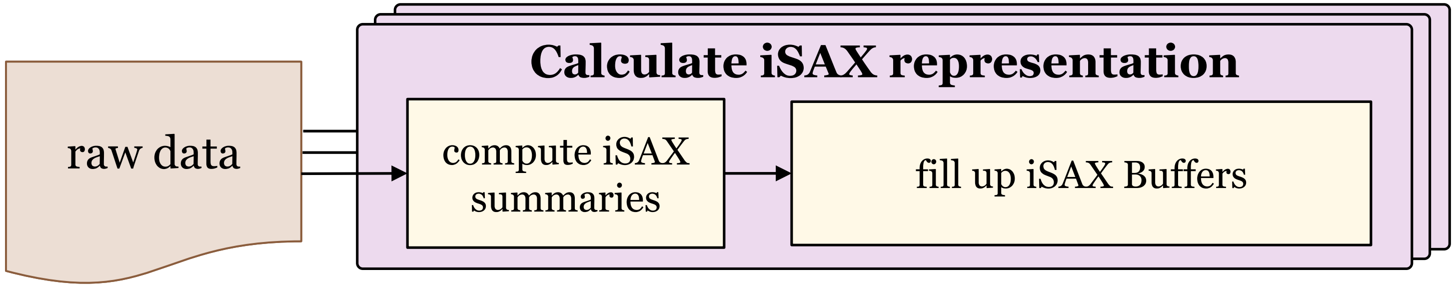
Airbus stores PB of data series, describing the behavior over time of various aircraft components
- vibrations of bearings in engines
- way pilots maneuver the plane through fly-by-wire system.

Analytics algorithms often operate on data subsets, which fit in memory:
- data relevant to landings from pilots

Astrophysics and neuroscience
- different, adhoc subsets of data need to be analyzed

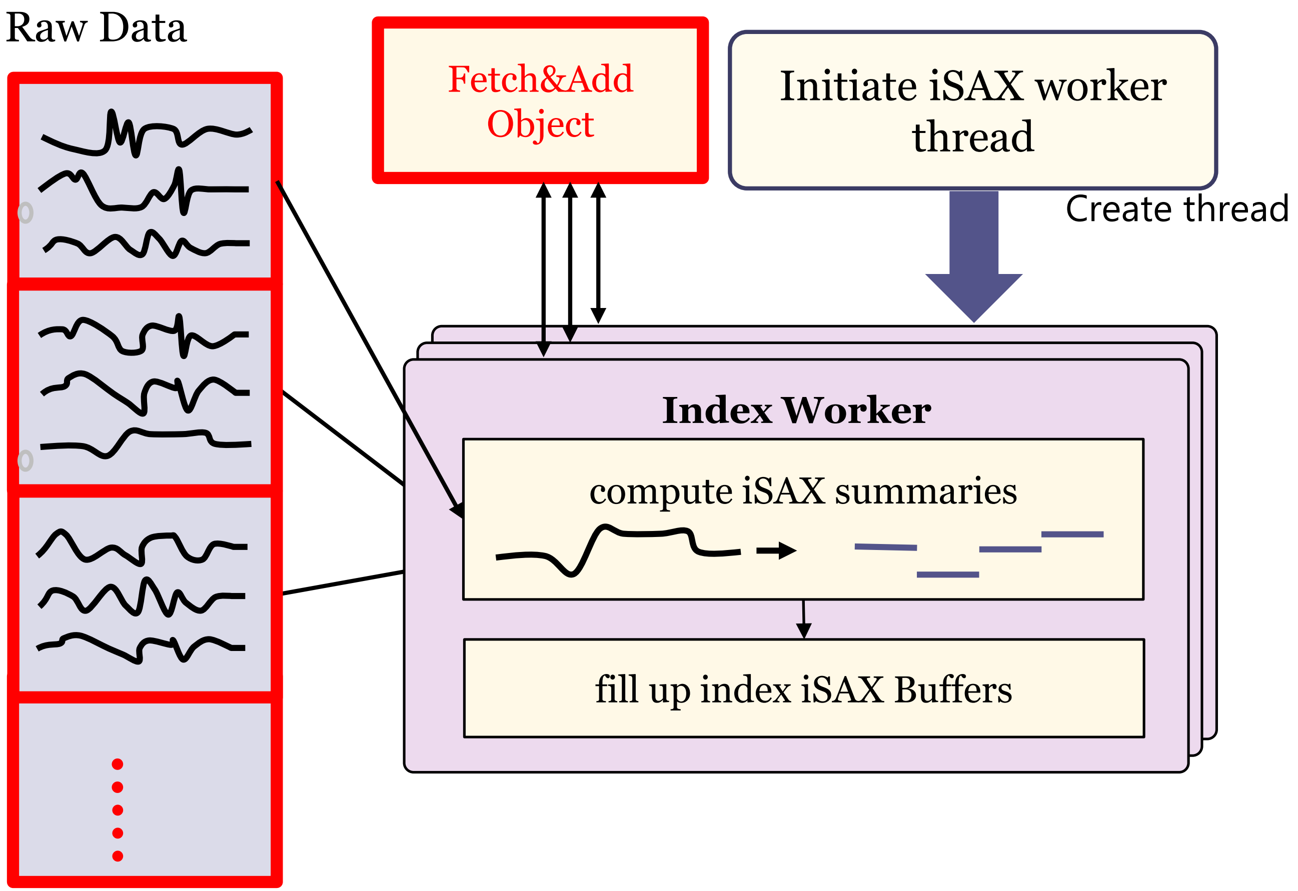
MESSI: In-memory data series index

Stage 1:
Load data to index



B. Peng, P. Fatourou and T. Palpanas, VLDB Journal 2021
IEEE International Conference on Data Engineering (ICDE) 2020

MESSI Index creation - Stage 1: Load data to Index

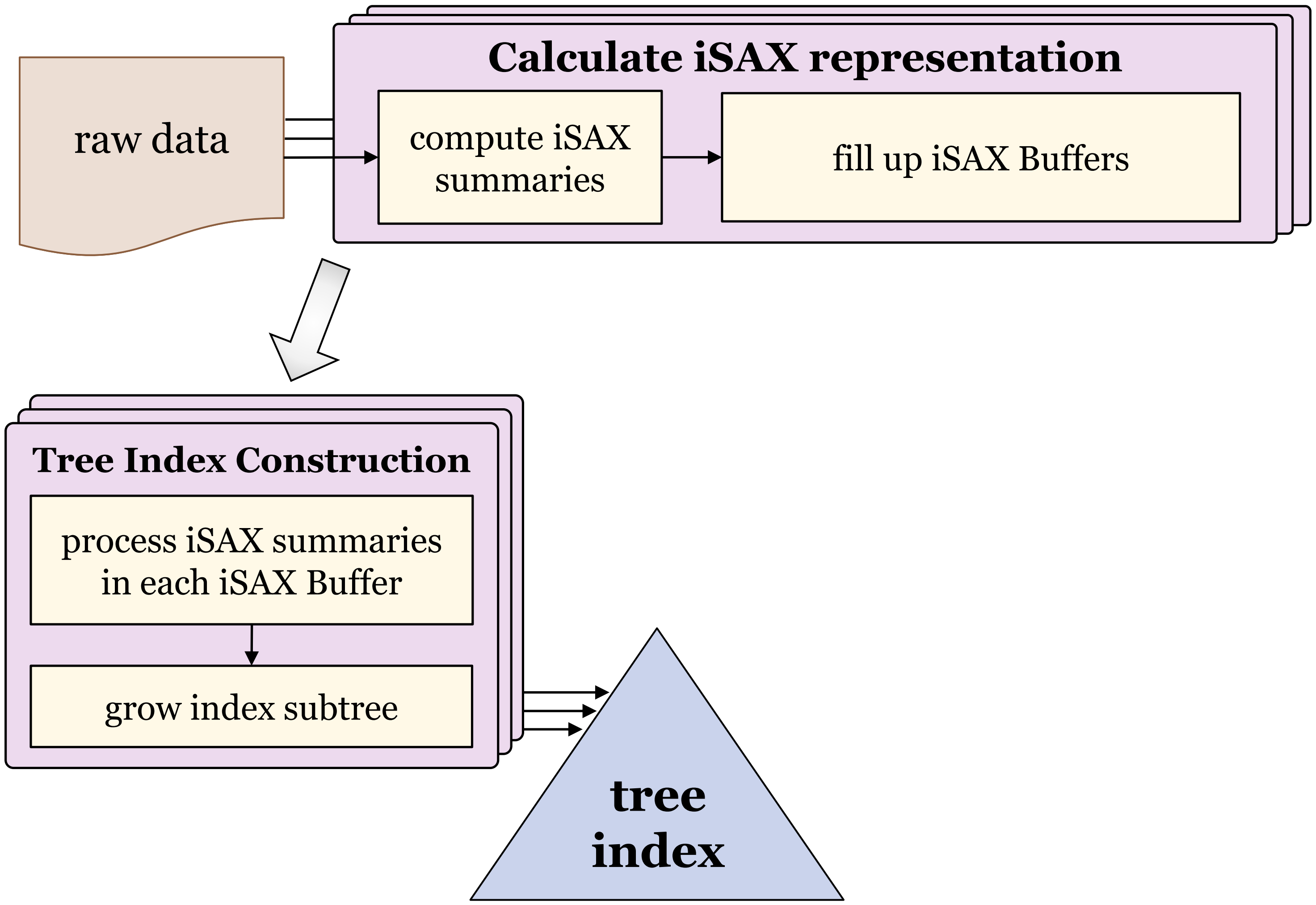


B. Peng, P. Fatourou and T. Palpanas, VLDB Journal 2021
IEEE International Conference on Data Engineering (ICDE) 2020

MESSI: In-memory data series index

Stage 1:
Load data to index

Stage 2:
Grow index

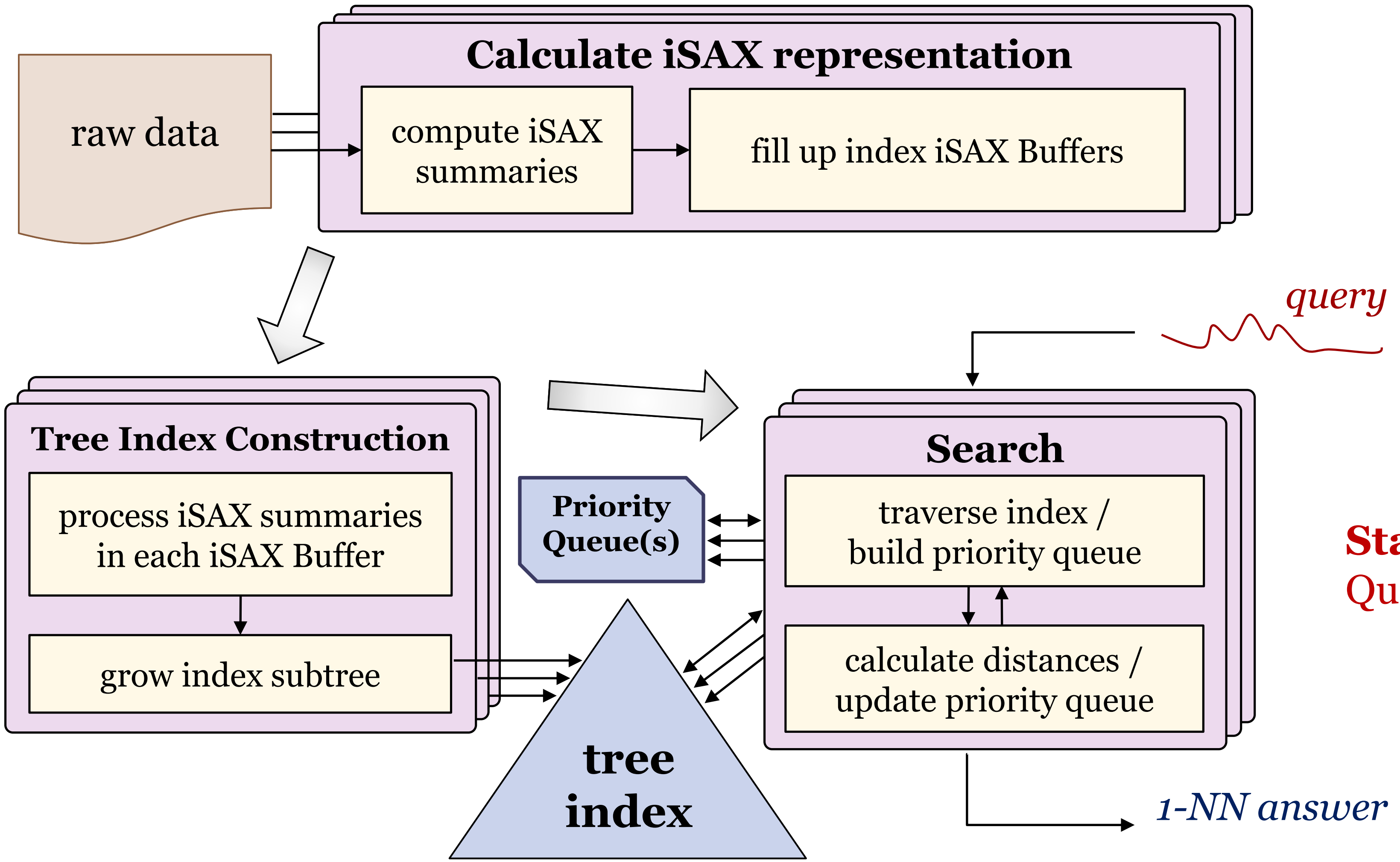


MESSI: In-memory data series index

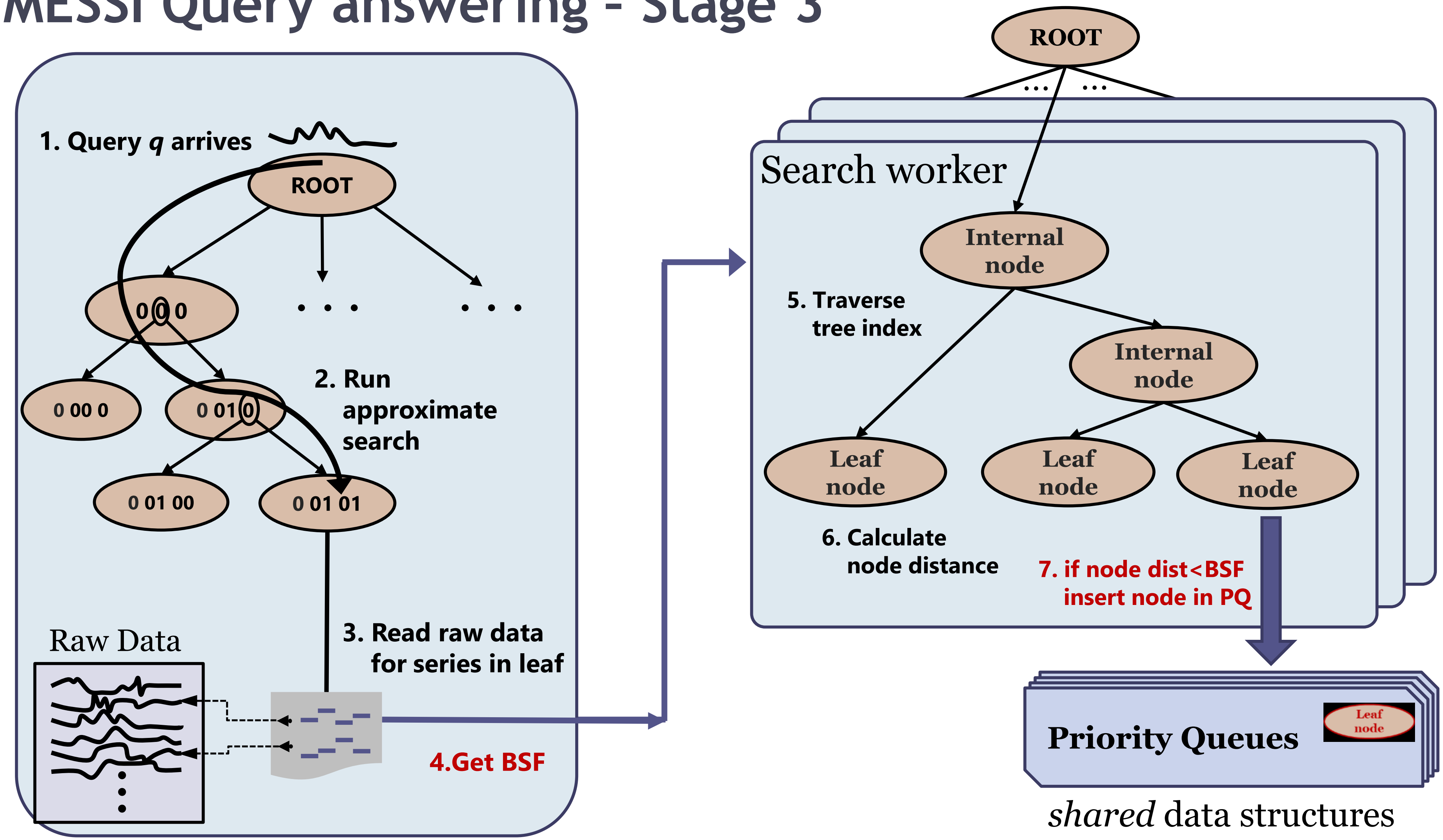
Stage 1:
Load data to index

Stage 2:
Grow index

Stage 3:
Query answering

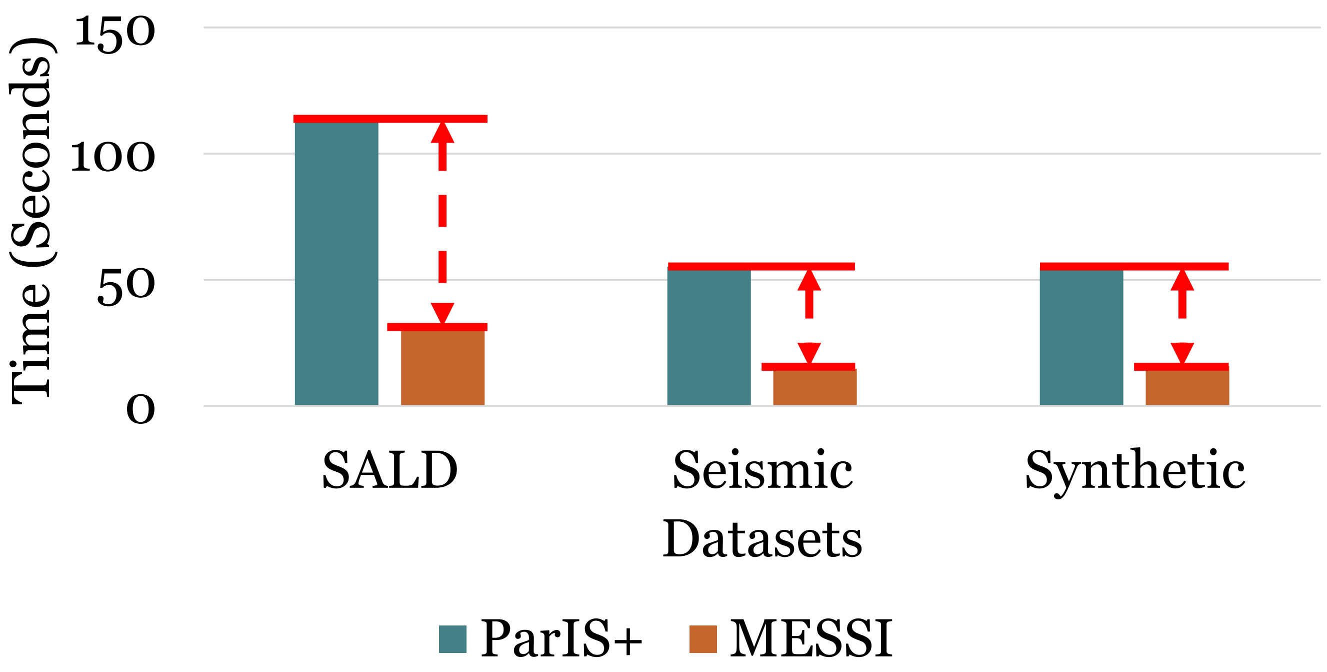


MESSI Query answering - Stage 3



Experiments

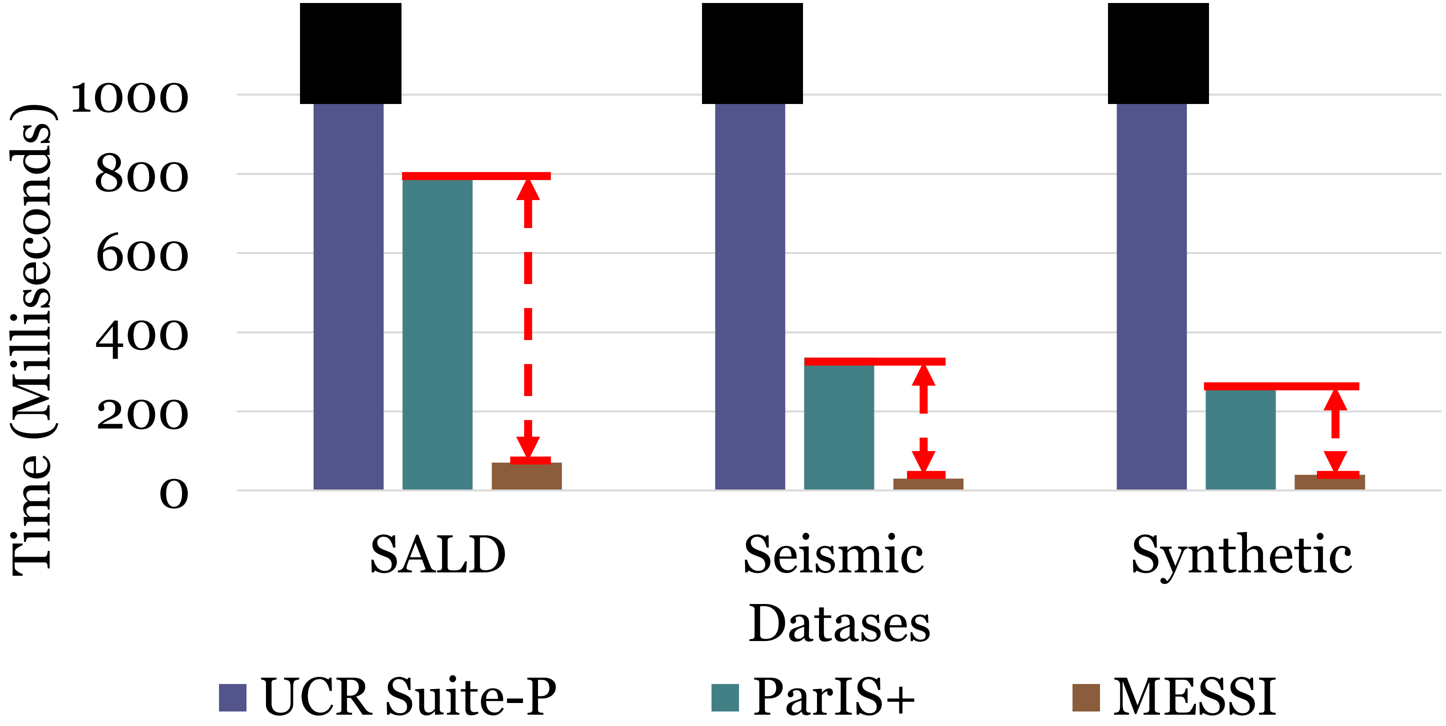
Time performance of index creation (100GB datasets)



MESSI up to 3.9x faster than **ParIS+ (in-memory)** for index creation

Experiments

Time performance of exact query answering



MESSI up to 11x faster than **ParIS+ (in-memory)** for query answering

Would using GPUs help?

Challenges

- Limited GPU memory size (12GB of RAM)
  - much smaller than raw data

- Slow interconnect speeds (PCI-Express 3.0 x16 delivers 10GB/sec)
  - Moving raw data needed by individual queries prohibitively expensive
  - Even moving small ad hoc subsets of data required by queries incurs a prohibitively high cost (0.4% of a 100 GB dataset requires > 40msec).
  - Raw data cannot be processed in the GPU.

- Non-sophisticated Streaming Processors (GPU cores)
  - not suited for supporting complex data structures/branching required in tree-like indexing

SING Main Ideas

Peng, Fatourou, Palpanas,
IEEE Int. Conference on Data Engineering (ICDE) 2020

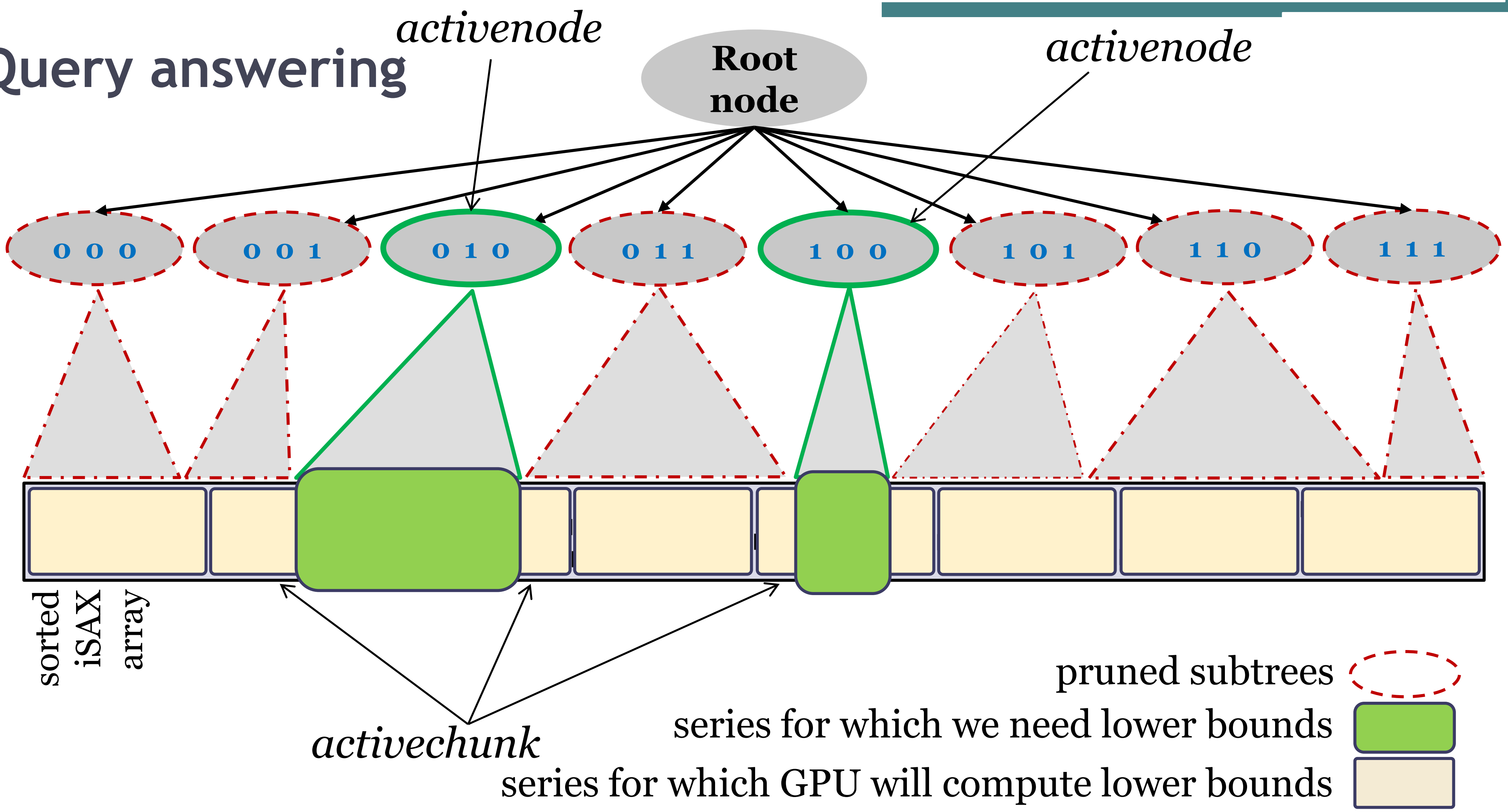
- Execute an in-order traversal of the root subtrees and store i-SAX summaries in an array in order
- Store this array in GPU memory
  - ✓ Serves large data series collections using limited GPU memory
  - ✓ Performs mostly consecutive in-memory accesses

- New pruning strategy that ignores entire root subtrees
  - ✓ Reduces the number of lower distance computations that the GPU should execute.

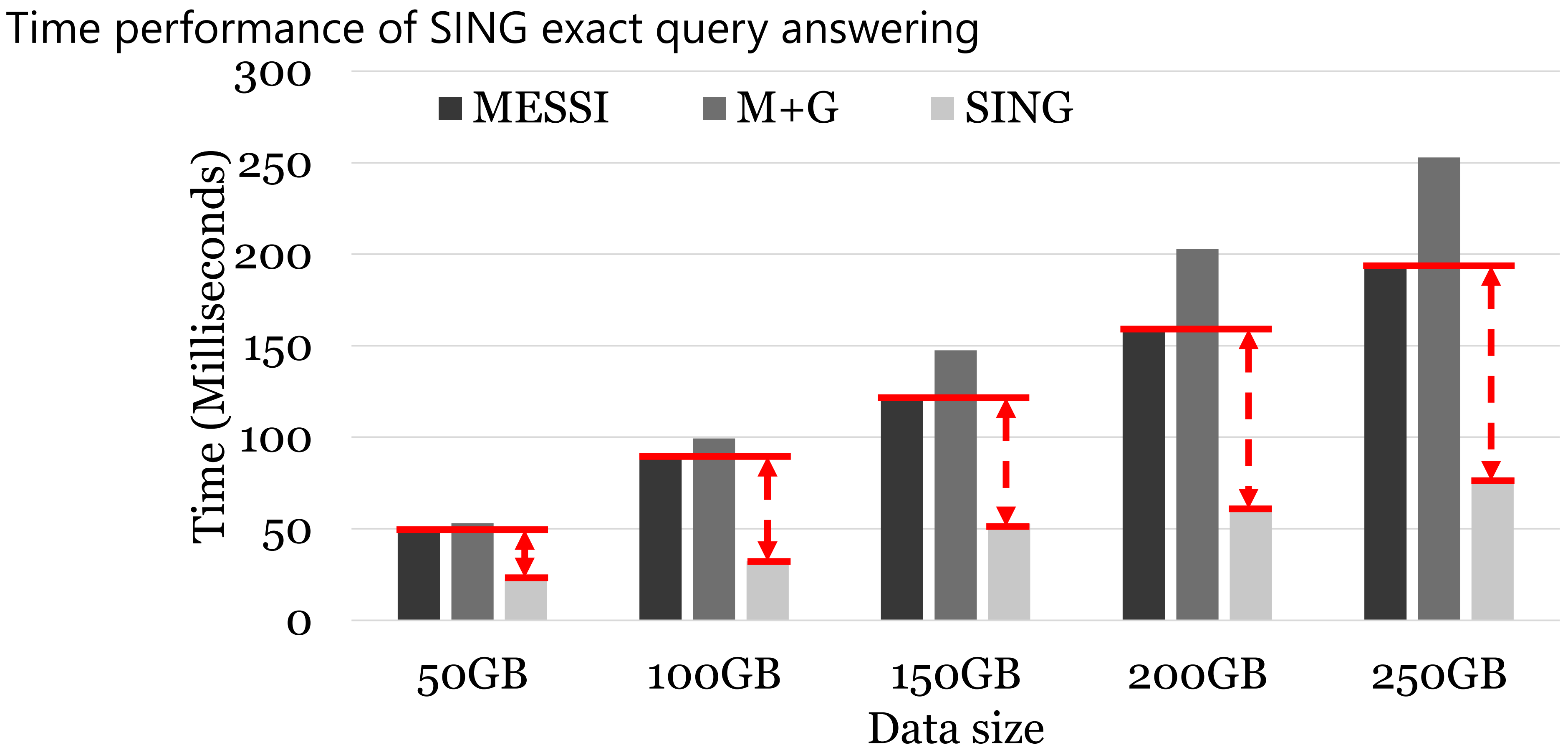
- Novel way to compute lower bound distances
  - ✓ Avoids the use of a dictionary
  - ✓ Employ a simple polynomial function

- Effectively divide the workload among the GPU and CPU cores, and orchestrates their parallel execution.
  - ✓ CPU workers start processing candidate answers without waiting for the GPU computation to complete.
  - ✓ Split the work into chunks
  - ✓ GPU streams the result of the work on each chunk to the CPU threads
  - ✓ CPU completes the computation for corresponding data series

SING Query answering



Experiments

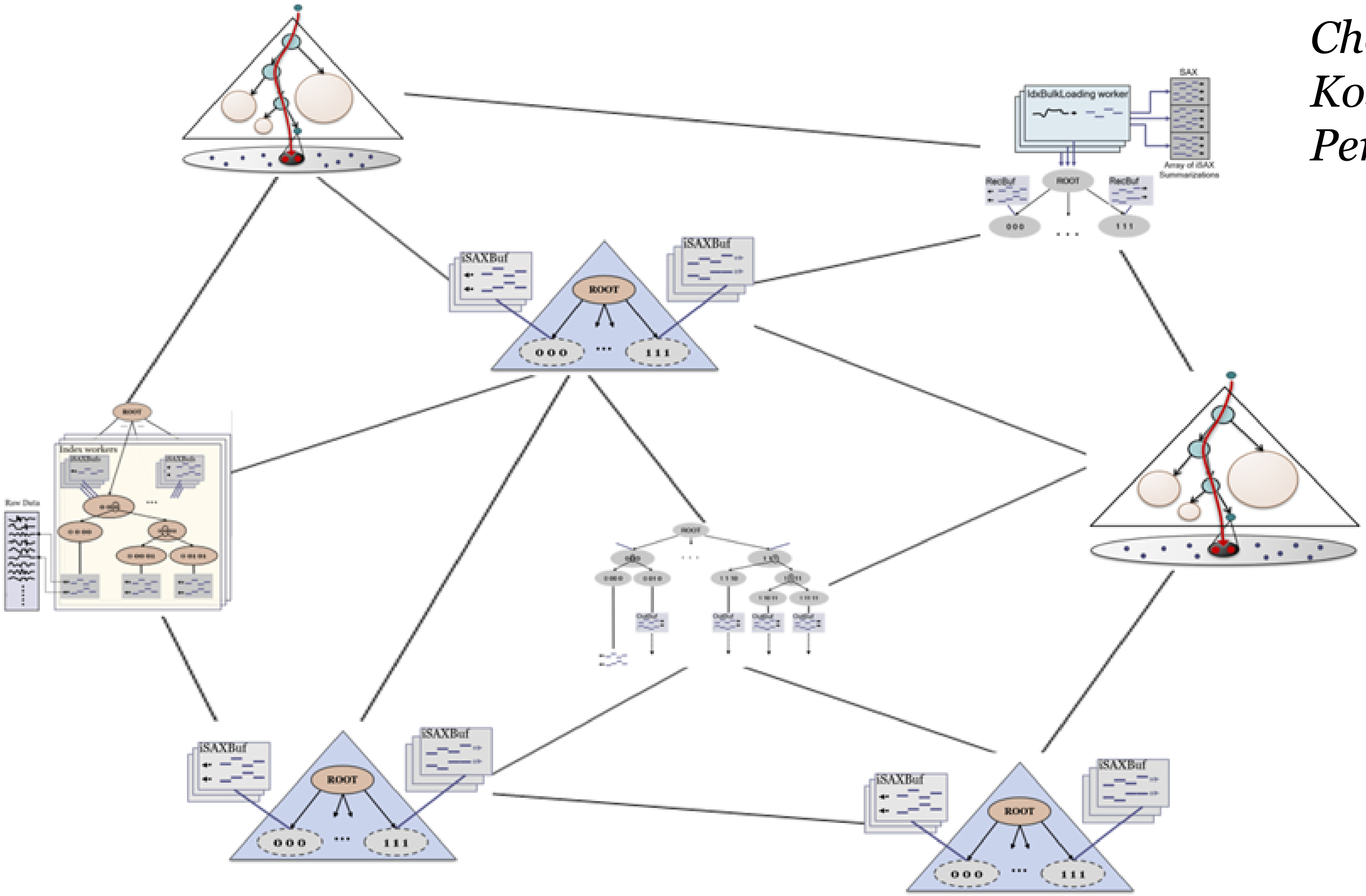


SING becomes increasingly faster than **MESSI** as dataset size grows

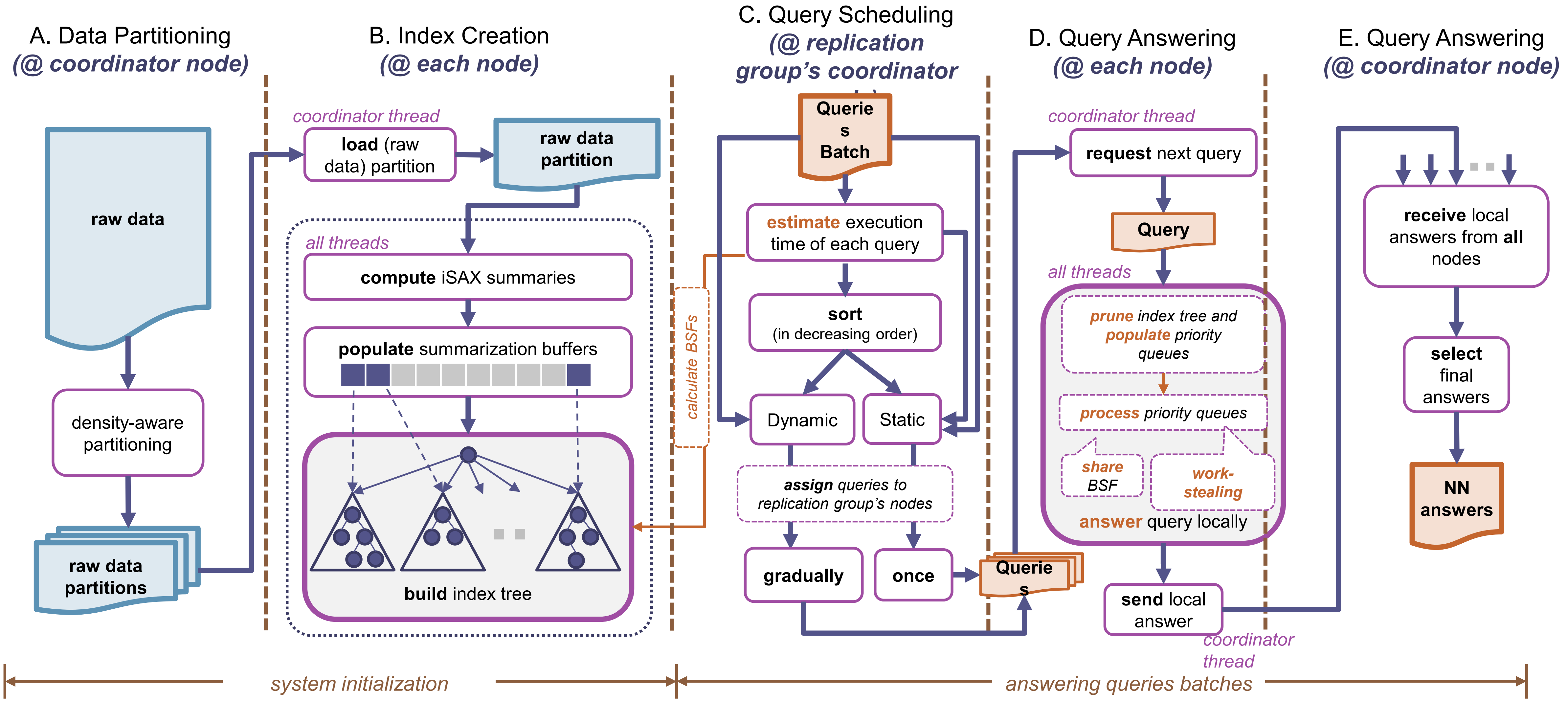
A Journey in the Land of Distributed Data Series Processing

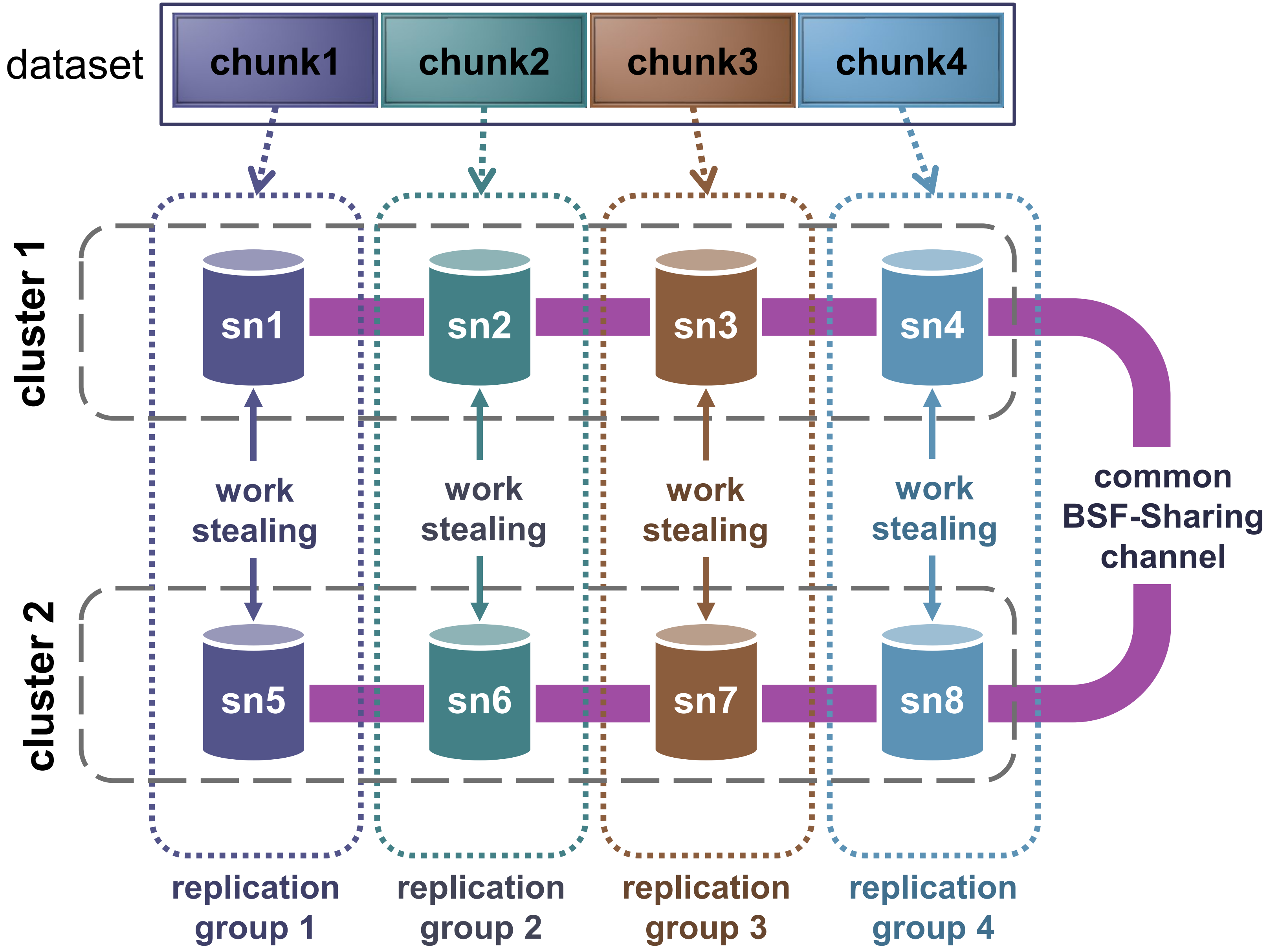
Odyssey: Techniques and Methodology

Chatzakis, Fatourou, Kosmas, Palpanas, Peng, VLDB End. 2023

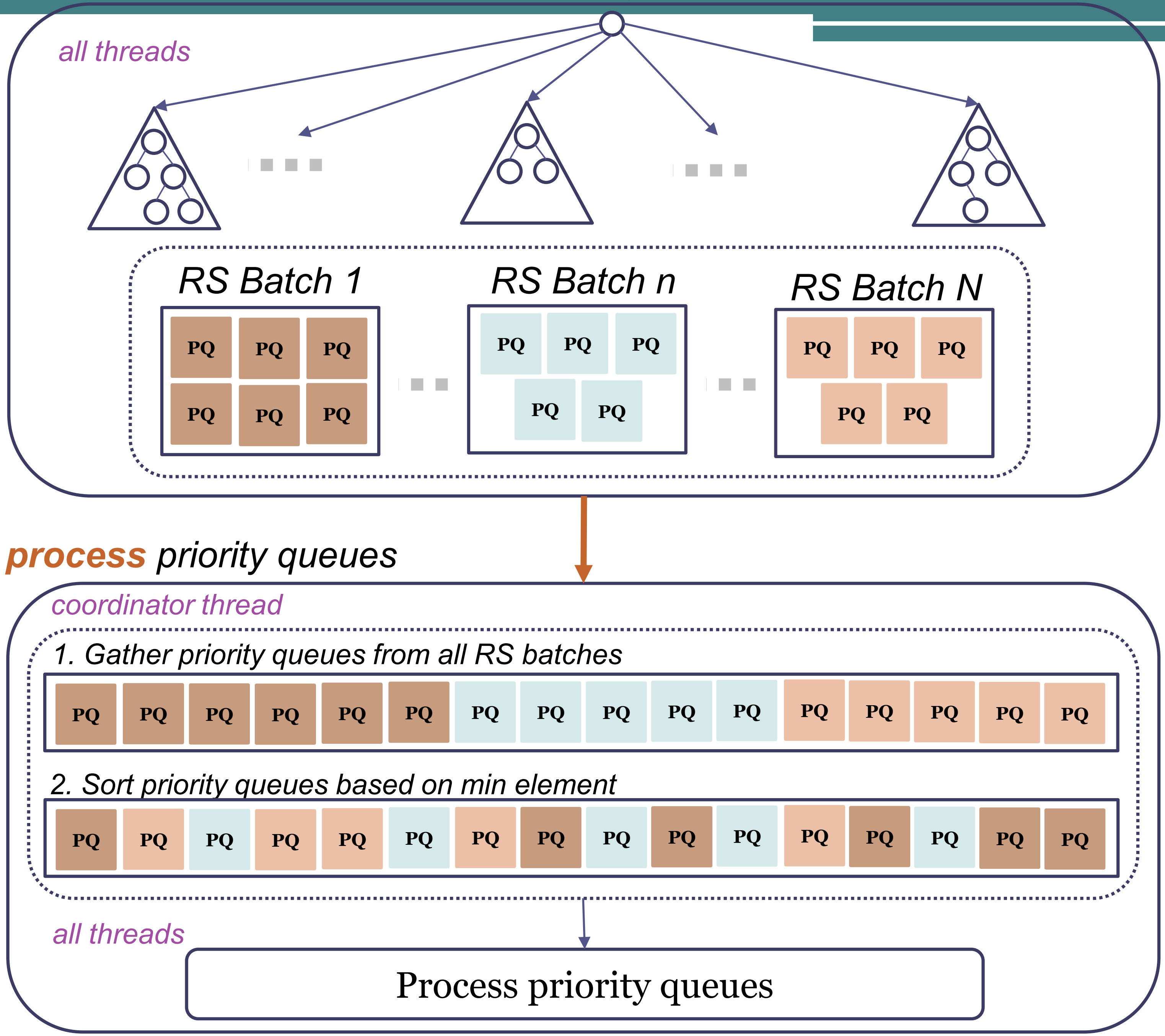


Techniques and Methodology




dataset
chunk1
chunk2
chunk3
chunk4
cluster 1
sn1
sn2
sn3
sn4
work stealing
work stealing
work stealing
work stealing
common BSF-Sharing channel
cluster 2
sn5
sn6
sn7
sn8
replication group 1
replication group 2
replication group 3
replication group 4

prune index tree and *populate* priority queues

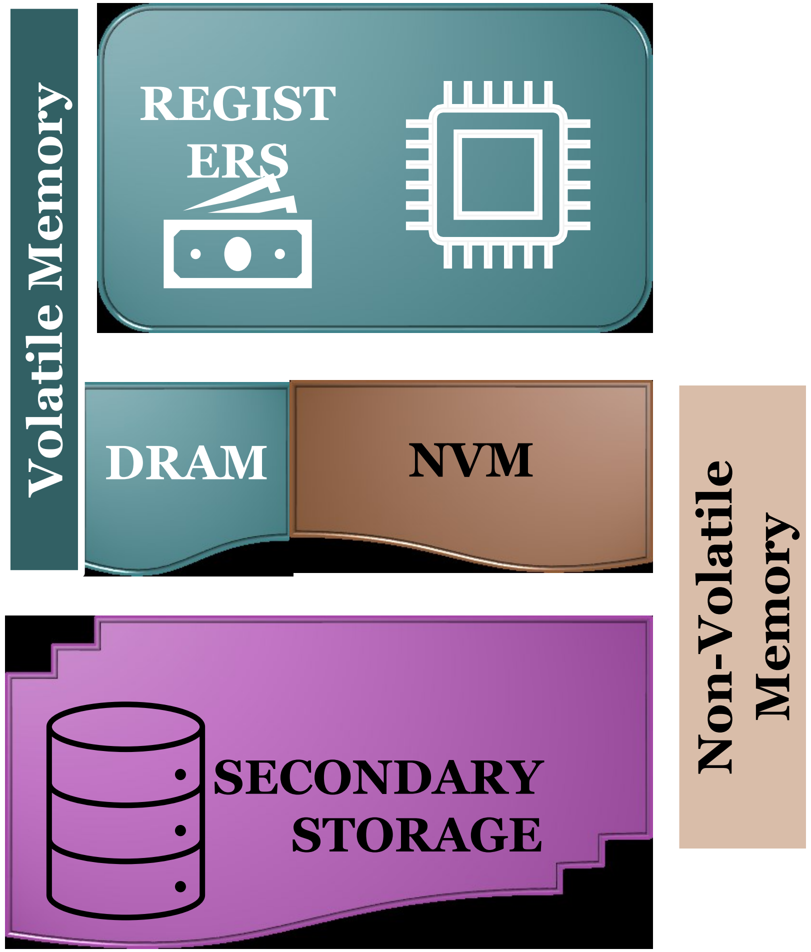


Persist: How to Compute with Non-Volatile Main Memory?

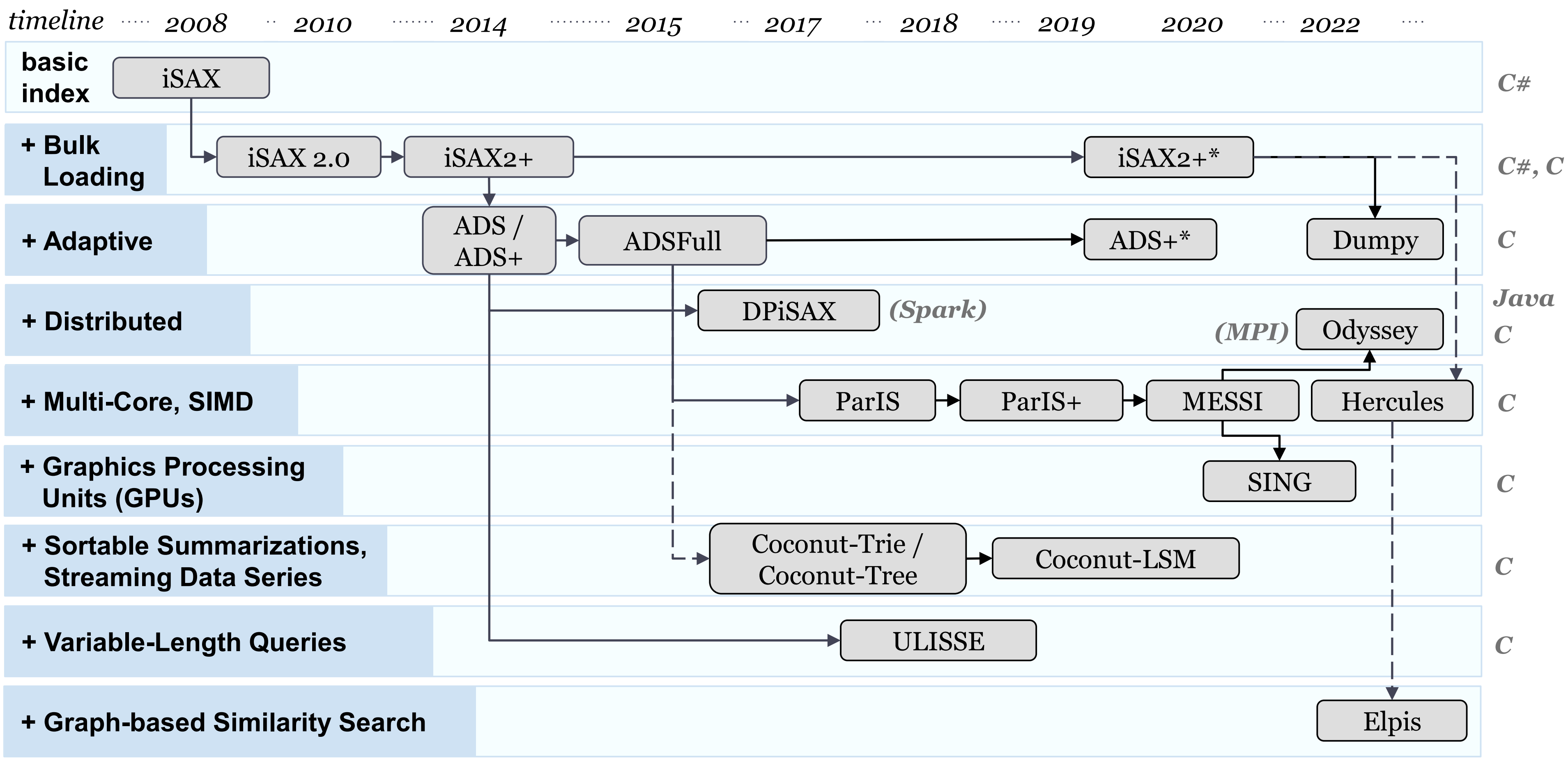
Funded under the 2nd HFRI call for faculty members of Greek Universities

- Non Volatile Main Memory:
  - ✓ Large size, cheap
  - ✓ Recoverability of computation after failures

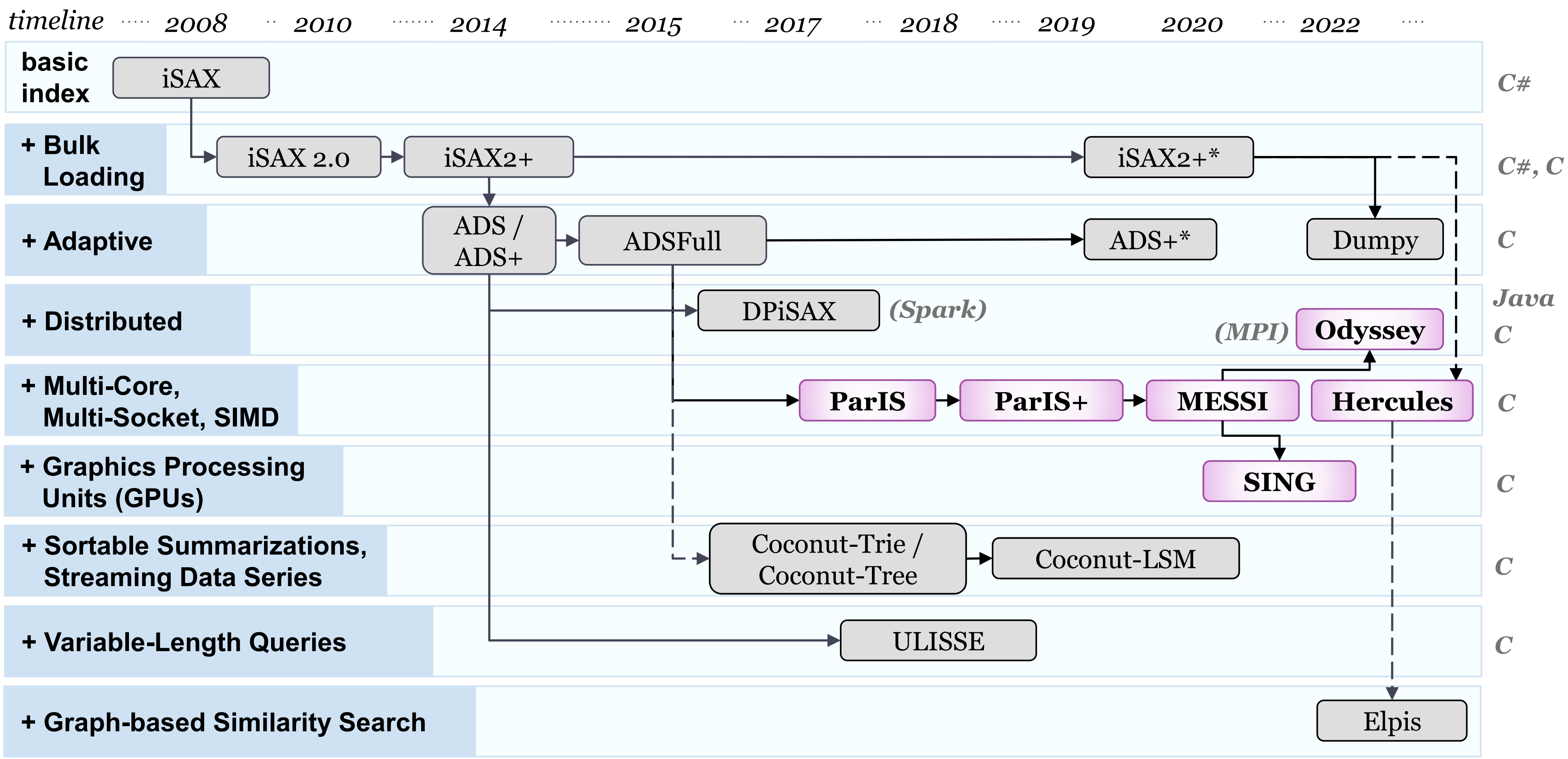
How does using NVM affect data series processing?



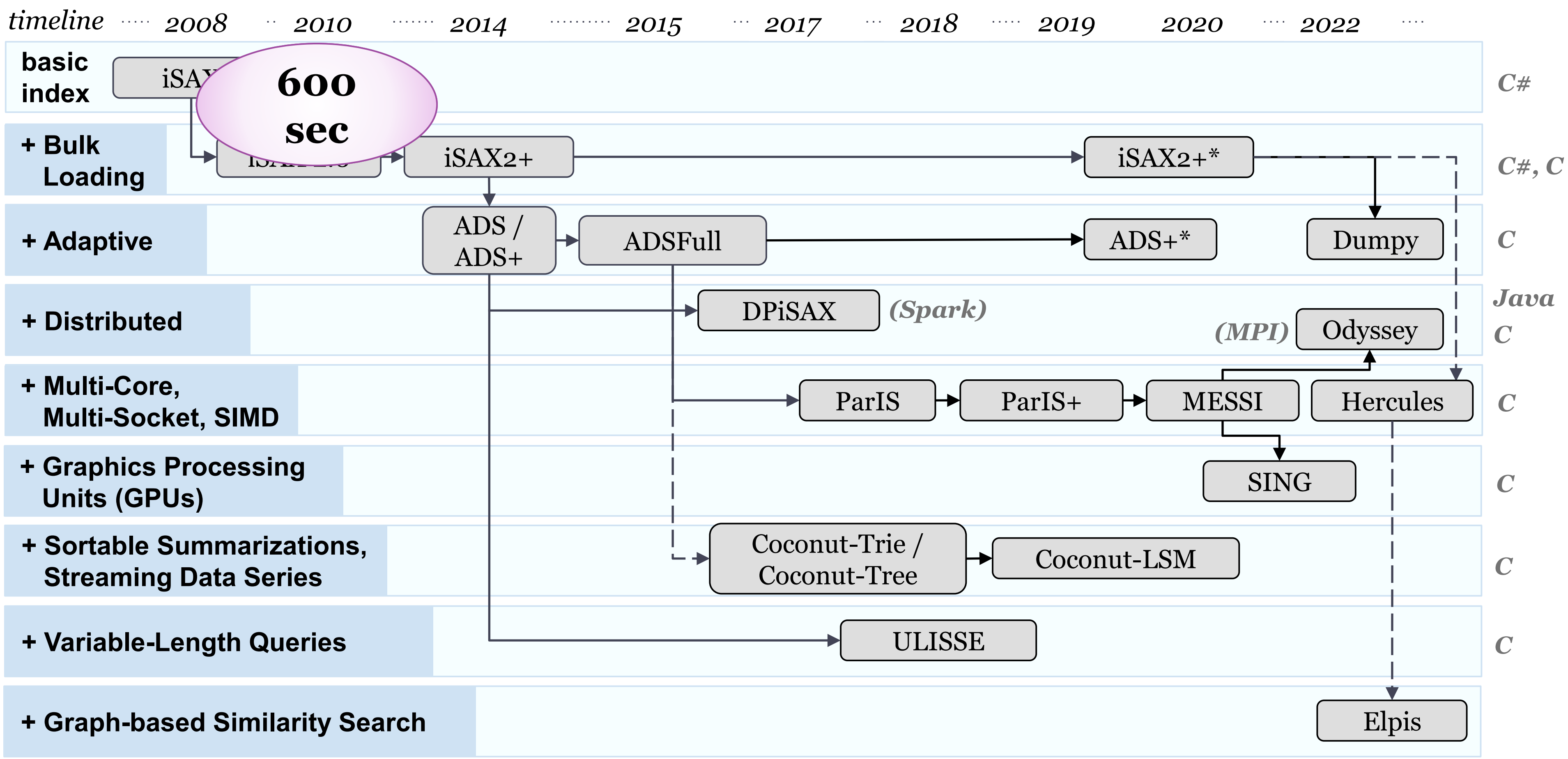
iSAX Index Family Lineage Tree



iSAX Index Family Lineage Tree

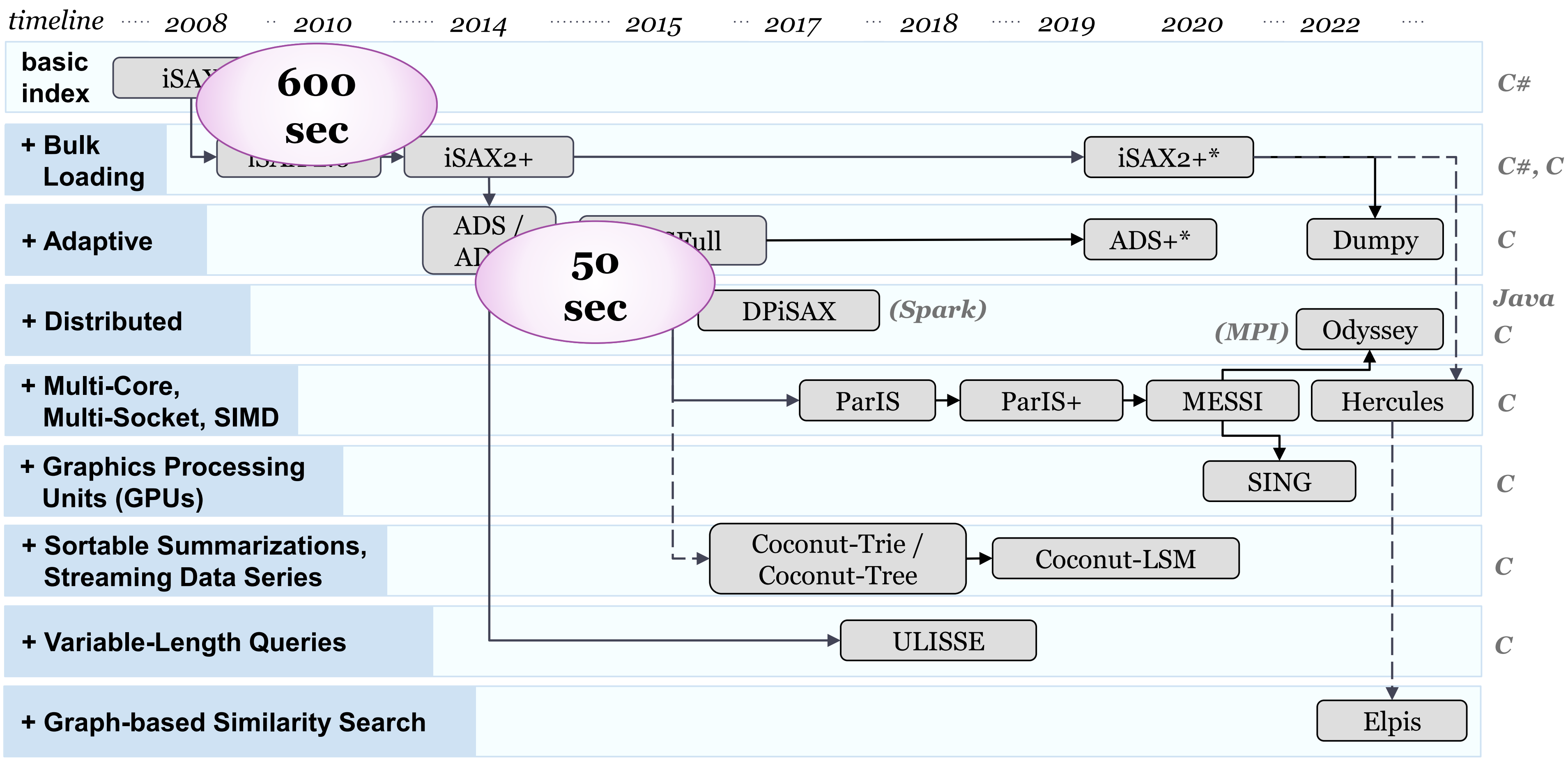


iSAX Index Family Lineage Tree



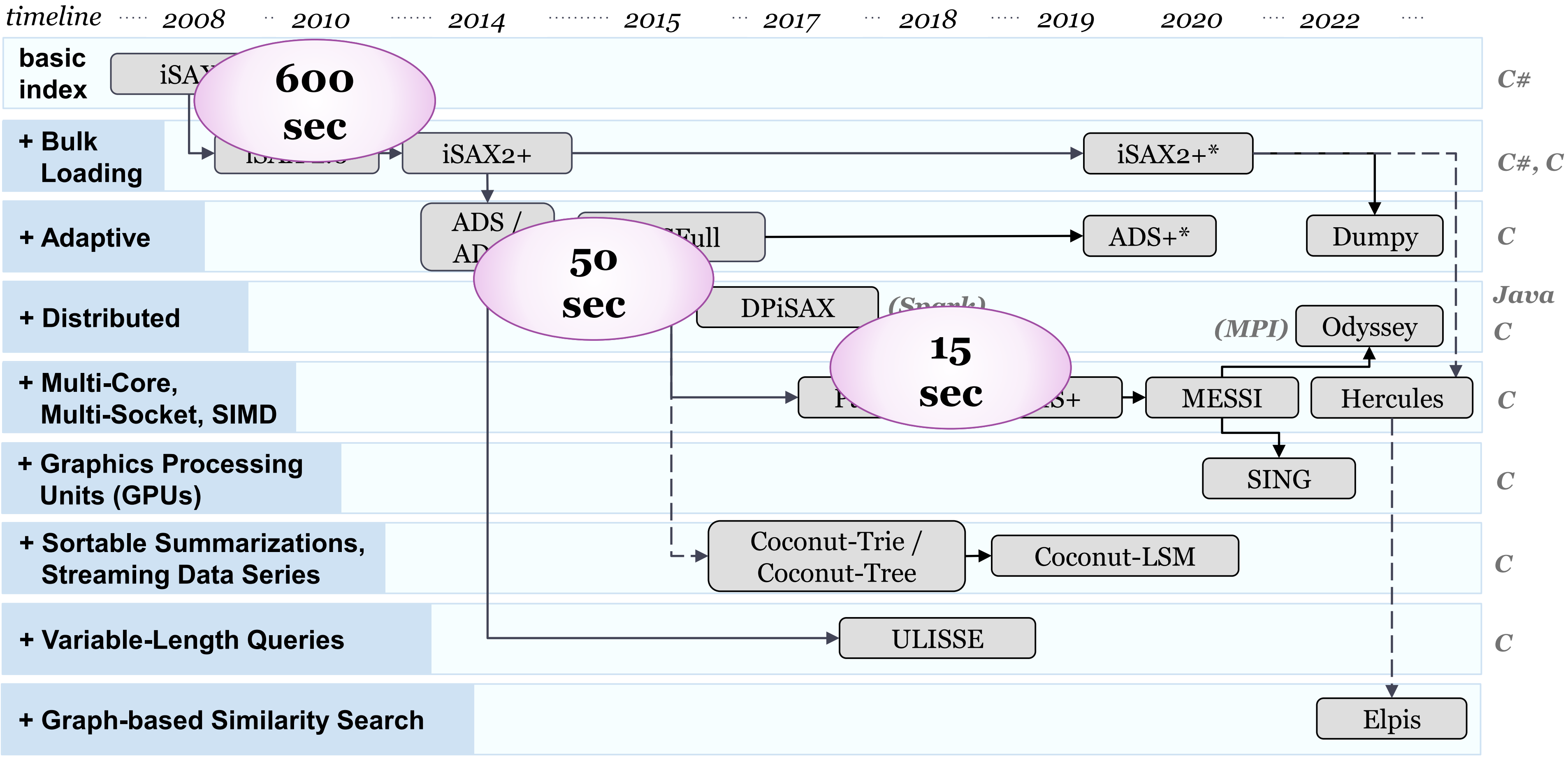
execution time for **1 similarity search query on a 100GB dataset *on disk***

iSAX Index Family Lineage Tree



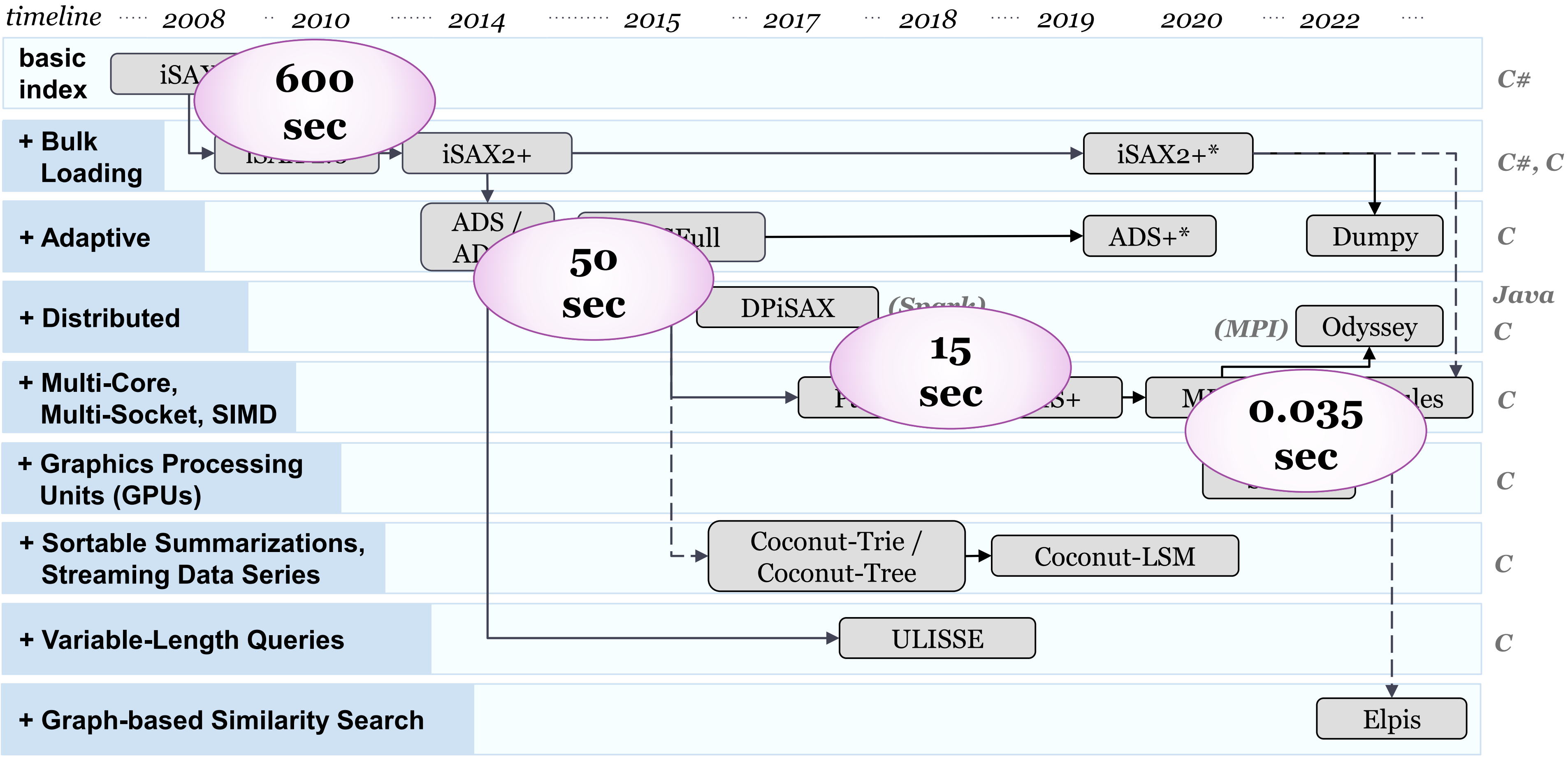
execution time for **1 similarity search query on a 100GB dataset *on disk***

iSAX Index Family Lineage Tree



execution time for **1 similarity search query on a 100GB dataset *on disk***

iSAX Index Family Lineage Tree



execution time for **1 similarity search query on a 100GB dataset *in memory***

Open Research Directions

- Extend existing work to other types of queries
  - **approximate** queries, **subsequence** similarity search
- Examine other modern hardware solutions
  - Field-Programmable Gate Arrays (**FPGA**s)
- Parallelization of other types of indexes that are not in the iSAX family.
  - DStree [Wang, Wang, Pei, Wang, Huang. A Data-adaptive and Dynamic Segmentation Index for Whole Matching on Time Series. PVLDB, 6(10):793-804, 2013.]
- Variable-length queries - Collections of variable length data series

Team and Current Collaborations

With Other Universities

- *Carnegie Mellon University, USA*
- *University of Toronto*, Canada
- *York University*, Canada
- *Technion*, Israel
- *Ben-Gurion University of the Negev*, Israel
- *École Polytechnique Fédérale de Lausanne* (EPFL), Switzerland
- *University of Paris (ex Paris Descartes University and ex Paris Diderot University),* France
- Telecom Paris, France
- Chinese Academy of Sciences
- University of Edinburgh
- University of Athens, University of Thessaloniki, ATHENA Research Center

Students and Postdocs

- Eleftherios Kosmas
- Nikos Kallimanis
- Botao Peng
- Karima Echichabi
- Ohad Ben-Baruch
- Naama Ben-David
- Yihan Shun
- Vassilis Gabrielatos
- Manolis Papadospiridakis
- Manos Chatzakis
- George Paterakis
- Konstantinos Chatzinikolaou
- Yuanhao Wei

Thank you!

Panagiota Fatourou

http://www.ics.forth.gr/~faturu/
faturu@csd.uoc.gr

Part of the presented work was done while P. Fatourou was working at LIPADE, Université Paris Cité

Université de Paris · data Intelligence Institute of Paris · UNIVERSITY OF CRETE · FORTH Foundation for Research & Technology - Hellas