QUERYING SQL, NOSQL, AND NEWSQL DATABASES TOGETHER AND AT SCALE

Bapi Chatterjee

IBM, India Research Lab, New Delhi, India

DISCLAIMER

- The statements/views expressed in the presentation slides are those of the presenter and should not be attributed to IBM in any manner whatsoever.
- The definitions, facts, numbers, etc. are true to the best of my knowledge at the time when I retrieved them from their respective original sources.
- The presentation does contain contents from external sources and they have been duly acknowledged.

ACKNOWLEDGEMENTS

Srikanta Bedathur Jagannath

(IBM IRL, New Delhi)

"BIG DATA" IN THE REAL WORLD

Consider Patient data in Real world

Arrays ->	EKG Traces
Time series ->	Blood Oxygen
Time series ->	Blood Pressure
Arrays ->	EEG Traces

- Demographic <- Tables
- Caregiver Notes <-Documents
- Medical Charts <- Tables
- Lab test results <- Tables
- X-ray , MRI, ETC <- Images

Source: Tim Mattson, 2015

POLY DB ENGINES

- Tables, Time series (RDBMS) - MySQL, PostgreSQL, and Oracle
- Documents (Document Store) - Google BigTable, Apache Accumulo, MongoDB
- Arrays, Images (Array DBMS) - C-Store, HStore, SciDB, VoltDB, Graphulo

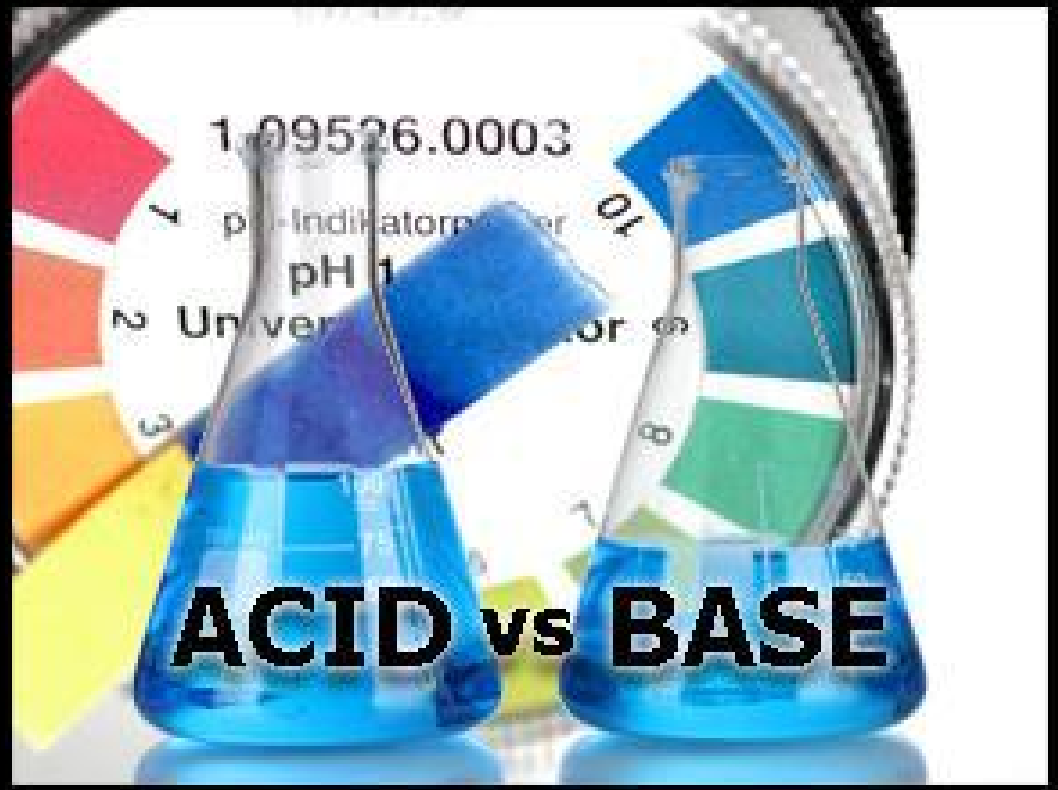
"POLY"QUERIES

- Complex analytics: Compute the FFT over all heartrate waveforms, grouped by patient and day
- Real-time decision making in SQL with streaming semantics: Raise an alarm if the heart rate over this window exceeds some threshold

SQL::NOSQL::NEWSQL

- SQL: Structured Query Language
- NoSQL: Not (only) SQL
- NewSQL: NoSQL but still SQL

SQL::NOSQL



ACID

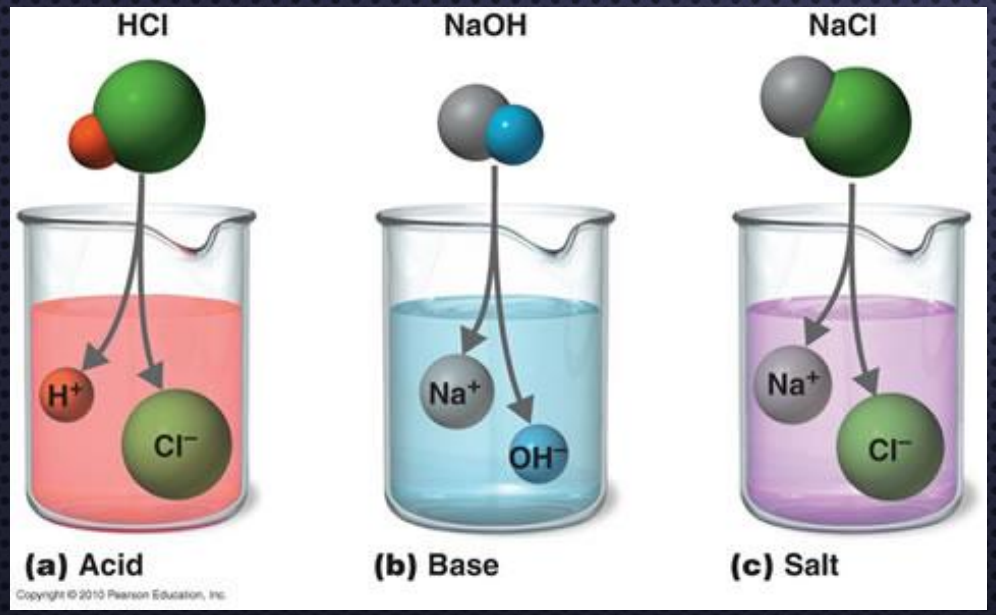
- **Atomicity** – Either the entire transaction complete or none
- **Consistency** – Any transaction will take the database from one consistent state to another with no broken constraints
- **Isolation** – Changes do not affect other users until committed
- **Durability** – Committed transactions can be recovered in case of system failure

BASE

- **Basic Availability** – Availability first even with partial consistency
- **Soft State** - Do away with consistency
- **Eventual Consistency** – Eventually, converge at a consistent state

(All about liveness, safety is ok to have but not an immediate requirement)

SQL::NOSQL::NEWSQL



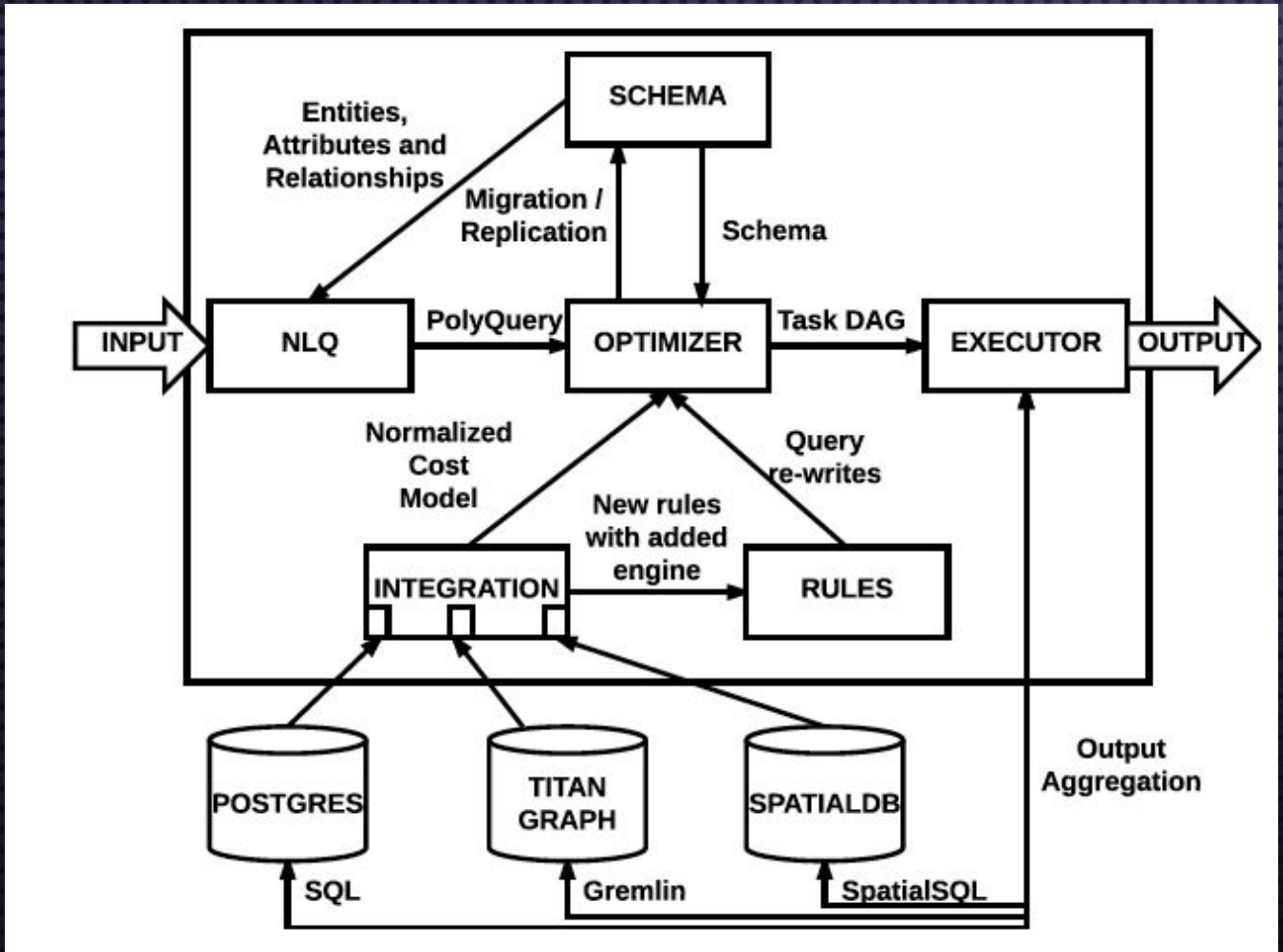
	SQL	NoSQL	NewSQL
Relational	Y	N	Y
Schema-less	N	Y	N
ACID Transactions	Y	N	Y
Horizontal Scalability	N	Y	Y
Performance Big Volume	N	Y	Y

SQL::NOSQL::NEWSQL

	SQL	NoSQL	NewSQL
Example	PostgreSQL	Accumulo	SciDB
Application	Transactions	Search	Analysis
Data Model	Relational Tables	Key-Value Pairs	Sparse Matrices
Math	Set Theory	Graph Theory	Linear Algebra
Consistency			
Volume			
Velocity			
Variety			
Analytics			
Usability			

Source: Hayden Jananthan et al., 2015

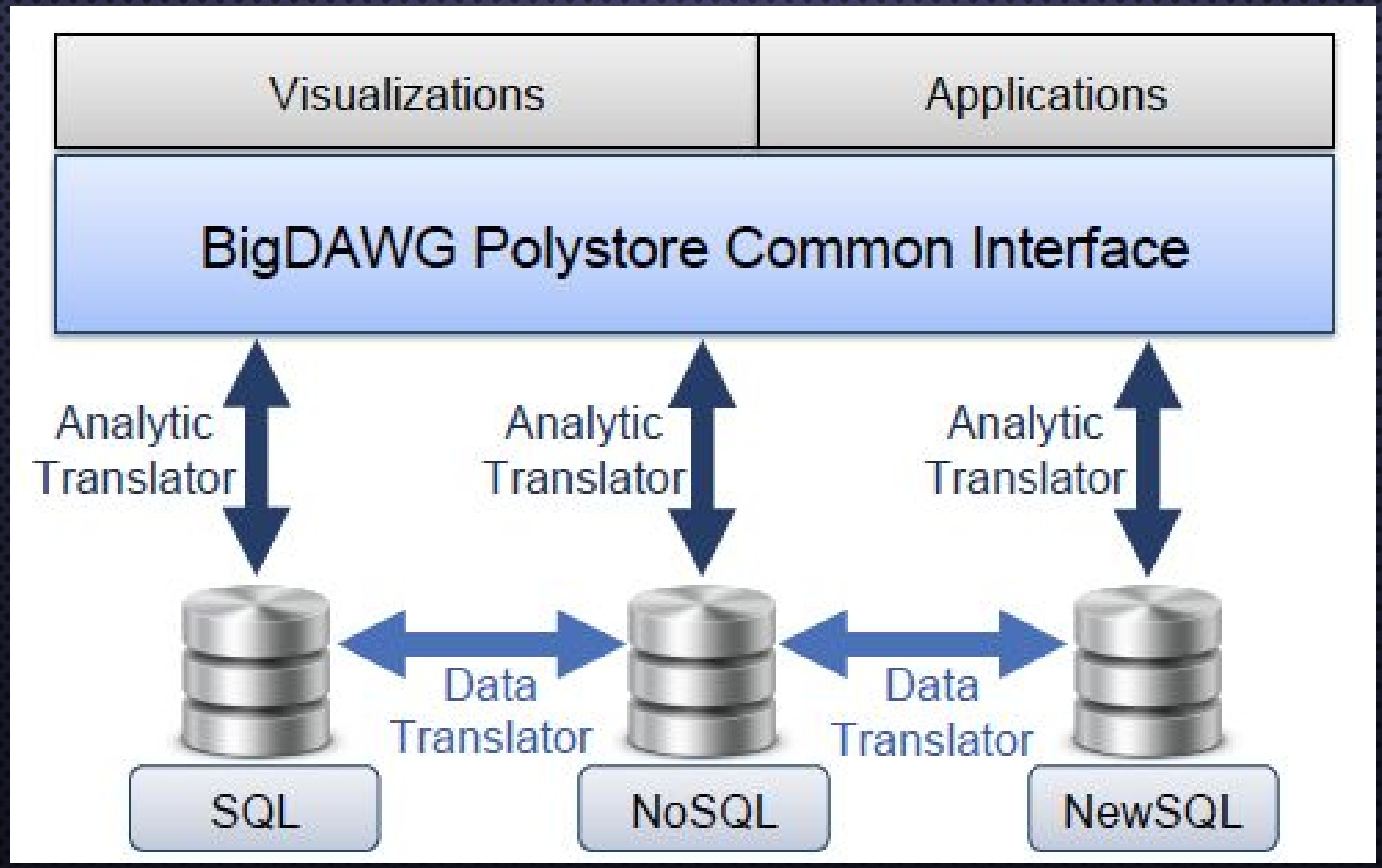
POLYSTORE



- Describe queries in a common language
- Break down the query execution into individual components
- Know where datasets are and what they contain
- Understand the query execution strength of each engine
- Support data transformation if required, but minimize its overheads
- Re-write queries into corresponding language
- And…deliver performance for complex queries

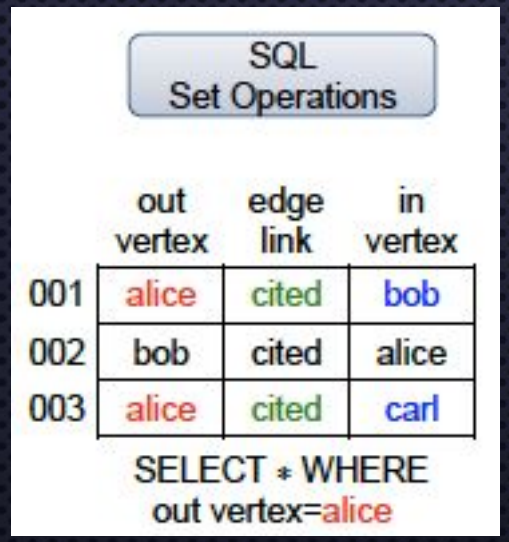
Source: Srikanta Bedathur et al., 2016

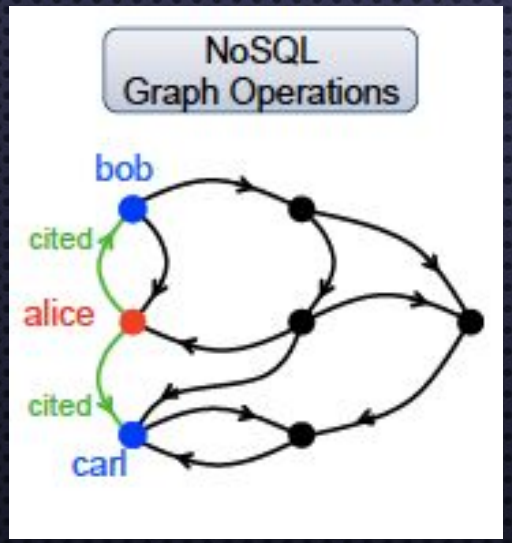
POLYSTORE: BIGDAWG

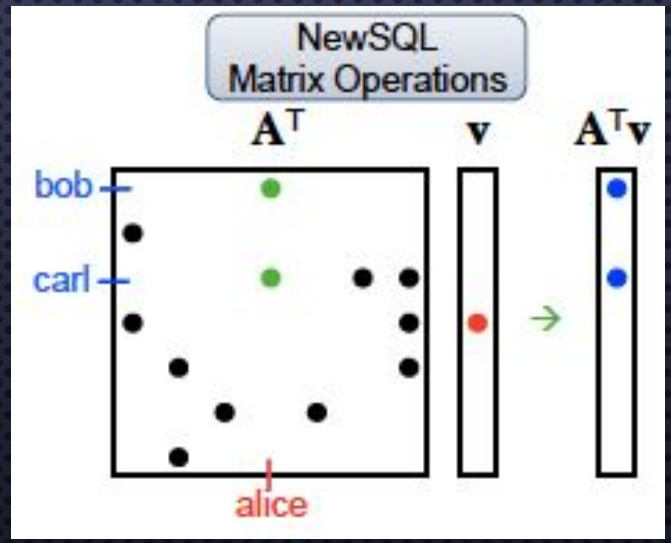


Source: Jeremi Kepner et al., 2016

POLYSTORE: MATHEMATICS







PolyAlgebra

- Mathematical underpinning for queries in a PolyStore.
- To encompass relational, graph, document, spatial, spatio-temporal, etc.
- Problem: Discovering a PolyAlgebra.
- Problem: Optimizing a Query Language based on a PolyAlgebra.

Source: Jeremi Kepner et al., 2016

POLYSTORE: MATHEMATICS

Integrating Data Model: D4M

- D4M: Dynamic Distributed Dimensional Data Model.
- Foundation of D4M: Associative array.
- Provide a generalization of sparse matrices.
- Constitute a function between a set of tuples and a value space.
- As a data structure, return a value given some number of keys.
- In practice, associative arrays support linear algebraic operations such as summation, union, intersection, multiplication and element-wise operations.
- Associative arrays have one-to-one relationship with key-value store databases, sparse matrices and adjacency matrices of graphs.

	SQL	NoSQL	NewSQL	Polystore
Example	PostgreSQL	Accumulo	SciDB	BigDAWG
Application	Transactions	Search	Analysis	All
Data Model	Relational Tables	Key-Value Pairs	Sparse Matrices	Associative Arrays
Math	Set Theory	Graph Theory	Linear Algebra	Associative Algebra
Consistency	■			■
Volume		■	■	■
Velocity		■	■	■
Variety		■		■
Analytics			■	■
Usability	■			■

Source: Jeremi Kepner et al., 2016

ASSOCIATIVE ARRAY: INTUITION

A	Artist	Date	Duration	Genre
053013ktnA1	Bandayde	2013-05-30	5:14	Electronic
053013ktnA2	Kastle	2013-05-30	3:07	Electronic
063012ktnA1	Kitten	2010-06-30	4:38	Rock
082812ktnA1	Kitten	2012-08-28	3:25	Pop

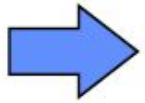
- Associative arrays are generalization of sparse matrices.
- Intuitively, an array is an Associative array if each row and column has a unique label.

Source: Jeremi Kepner et al., 2016

ASSOCIATIVE ARRAY: CONSTRUCTION

Input Data

Time	Col1	Col2	Col3
2001-01-01	a		a
2001-01-02	b	b	
2001-01-03		c	c



	01-01-2001	02-01-2001	03-01-2001
Col1|a	1		
Col1|b		1	
Col2|b		1	
Col2|c			1
Col3|a	1		
Col3|c			1

	Col1|a	Col1|b	Col2|b	Col2|c	Col3|a	Col3|c
01-01-2001	1				1	
02-01-2001		1	1			
03-01-2001				1		1

Graphs
Adjacency Matrix

Matrices
- Straightforward if Boolean
- Same as tables, else

Source: Jeremi Kepner et al., 2016

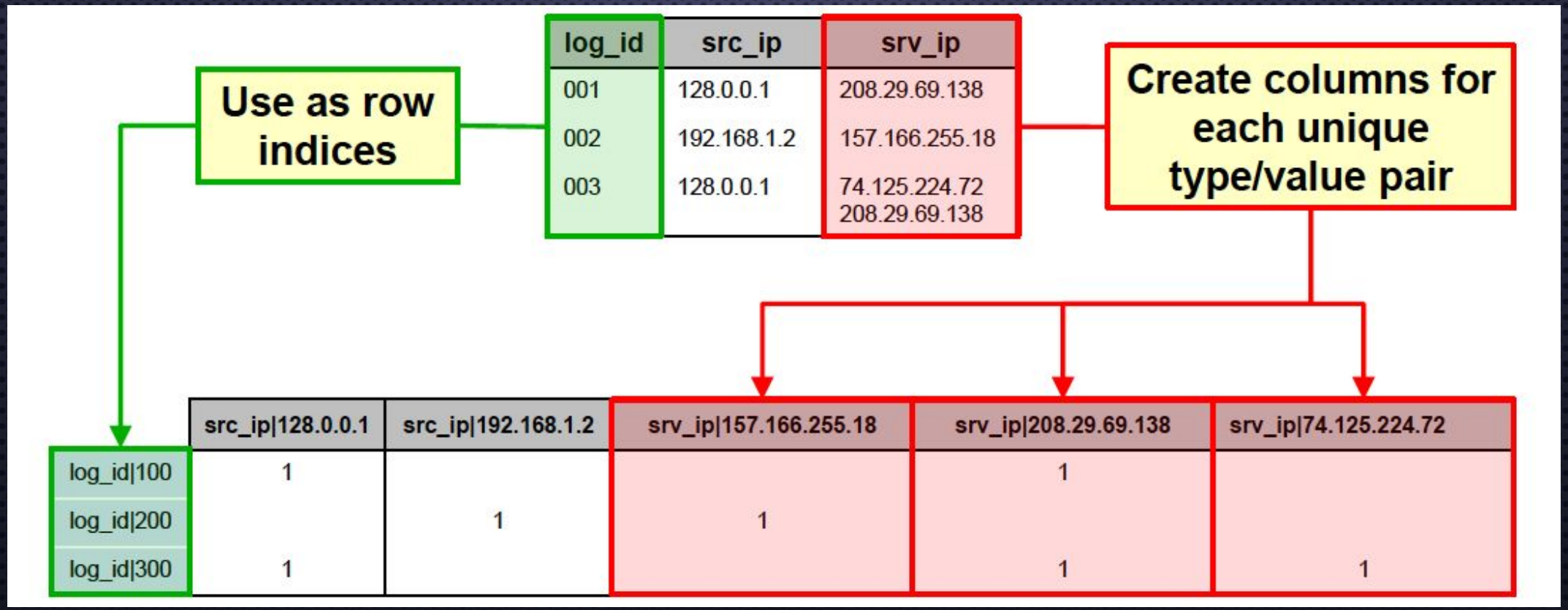
ASSOCIATIVE ARRAY: MATHEMATICS

- Closure: All mathematical operations on two (or more) associative arrays return associative array.
- Given associative arrays A, B, and C, associative array addition is denoted by A + B = C.
  - This is equivalent to inserting new rows in a table.
- Given associative arrays A, B, and C, associative array element-wise multiplication is denoted A * B = C
  - This is equivalent to selecting rows from a table.
- In practice, all computations are restricted to the nonzero rows and nonzero columns of the associative array representation of relations.
- In many computations, the only operations that need to be specified are the identities, the additive inverse and multiplicative annihilator.
  - $v + 0 = v$; $v * 1 = v$; $v + -v = 0$; $v * 0 = 0$; $v \in V$
  - $(V, +, *, 0, 1)$ form a semiring.

A semiring is a set together with two binary operators S(+,*) and additive and multiplicative identity elements 1 and 0, respectively, satisfying the conditions: Additive associativity, Additive commutativity, Multiplicative associativity, Left and right distributivity

Source: Hayden Jananthan et al., 2015

EXAMPLE



Source: Jeremi Kepner et al., 2016

EXAMPLE

Query Operation	SQL	D4M
Select all	SELECT * FROM T	E(:,:)
Select column	SELECT src_ip FROM T	E(:,StartsWith('src_ip| '))
Select sub-column	SELECT src_ip FROM T WHERE src_ip=128.0.0.1	E(:,'src_ip|128.0.0.1 ')
Select sub-matrix	SELECT * FROM T WHERE src_ip=128.0.0.1	E(Row(E(:,'src_ip|128.0.0.1 '))),:)

Source: Jeremi Kepner et al., 2016

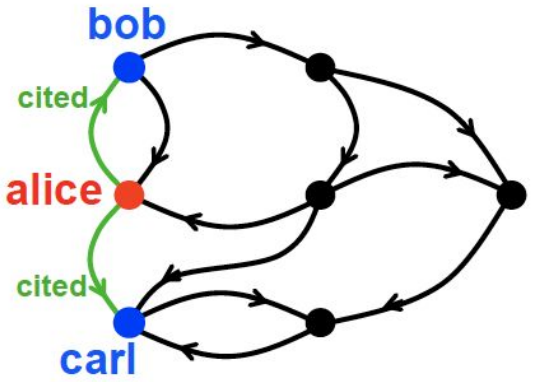
POLYSTORE: MATRIX ALGEBRA

SQL
Set Operations

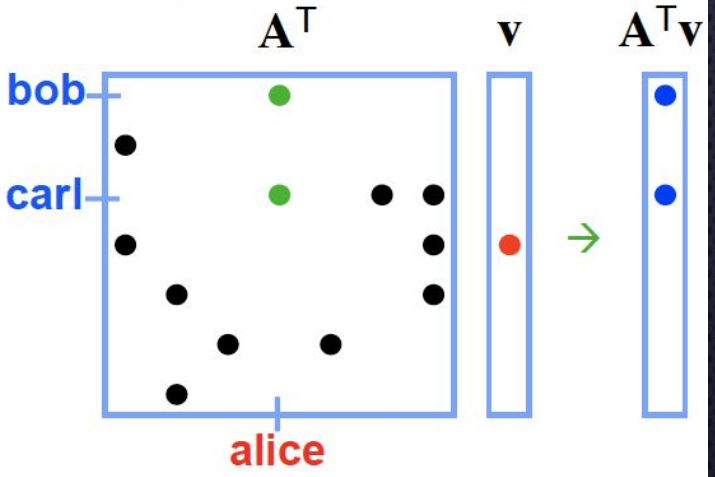
	from	link	to
001	alice	cited	bob
002	bob	cited	alice
003	alice	cited	carl

SELECT
WHERE from=alice

NoSQL
Graph Operations



NewSQL
Linear Algebra



Associative Array Algebra Provides a Unified Mathematics for SQL, NoSQL, NewSQL

$$A = \mathbb{S}^{N \times M}(k_1, k_2, v, \oplus) \qquad (k_1, k_2, v) = A \qquad C = A^T \qquad C = A \oplus B \qquad C = A \otimes C \qquad C = A\,B = A \oplus . \otimes B$$

Operations in All Representations are Equivalent

Source: Jeremi Kepner et al., 2016

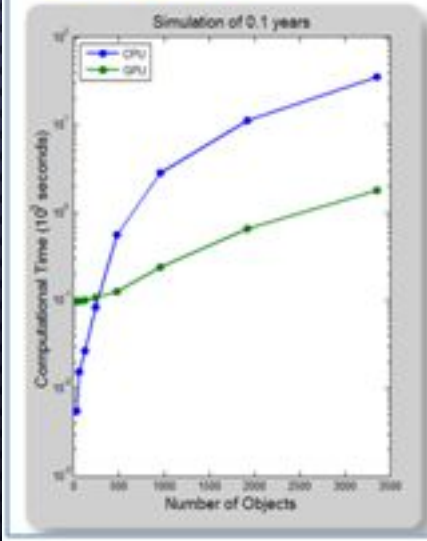
HPC FOR POLYSTORE QUERIES

- BLAS: Basic Linear Algebra Subprograms
- pMatlab
- Matlab-GPU/CUDA

MathWorks

Benchmark: N-Body Simulation

CPU vs GPU



Objects	CPU (10^3s)	GPU (10^3 s)	Speed up
60	0.015	0.099	0.15
120	0.027	0.099	0.27
240	0.083	0.108	0.76
480	0.559	0.126	4.42
960	2.83	0.241	11.77
1920	11.3	0.655	17.17
3360	35.3	1.822	19.38

Thank you!