

# NOSQL WITH CACHING, SEARCH AND REAL-TIME ANALYTICS

James Gorlick  
Basho

[bit.ly/bdp-on-the-underln](http://bit.ly/bdp-on-the-underln)



@paegun

# About Me

Application Developer with Lead and Project Management experience in a wide variety of business applications. Particularly interested in straight-through processing, removing data duplication, and increasing information availability throughout the organization in a secure manner. Previously with Dimensional Fund Advisors performed data science with much Fortran and many data sources while helping DFA "buy the market". Developed a number of in-house solutions for Myspace to serve purposes similar to what BDP does and will provide.

@paegun



# Word Count

Just as developer activities extend beyond writing and testing software into what were typically operational concerns, data solutions are extending beyond storing and retrieving data from a single database management system into the deployment and management of heterogeneous services.

A walk through the development and operational support of a reference application utilizing the Basho Data Platform, gains within the development and operations cycles will be demonstrated.









# Excellent Data

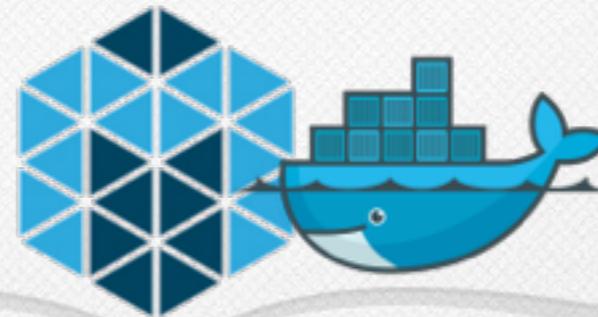
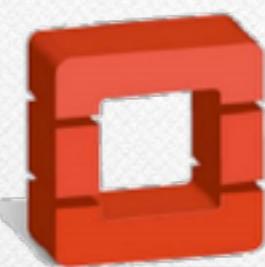
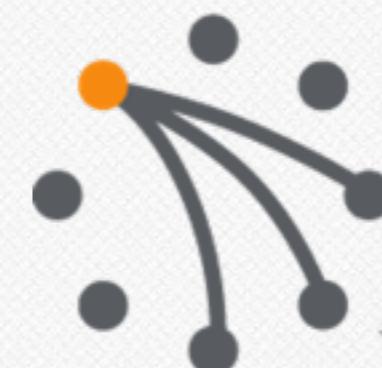
NLC	Origin TLC	Station Name	rnment Office Region (County or Unitary Authority)	ict or Unitary Auth	2 Spatial_Unit	NUTS2 Spatial Unit	Location Facility	Ownation
			London	Greater London	Greenwich	UKI2	Outer London	Southeastern
			Wales - Cymru	Caerphilly - Caerffili	Caerphilly - Caerffili	UKL1	West Wales and the Valleys	Arriva Trains Wales
			Rhondda Cynon Taff -	Rhondda Cynon Taff -	Rhondda Cynon Taff -		West Wales and the	
5131	ABW	Abbey Wood						
3813	ABE	Aber						
3801	ACY	Abercynon						
3982	ABA	Aberdare						
8976	ABD	Aberdeen						
9090	AUR	Aberdour	01_10/11	1	31	1-Apr-10	1-May-10	189.1
			02_10/11	2	28	2-May-10	29-May-10	181.6
4435	AVY	Aberdovey	03_10/11	3	28	30-May-10	26-Jun-10	175.9
4440	ABH	Abererch	04_10/11	4	28	27-Jun-10	24-Jul-10	183.4
3602	AGV	Abergavenny	05_10/11	5	28	25-Jul-10	21-Aug-10	160.4
			06_10/11	6	28	22-Aug-10	18-Sep-10	175.8
			07_10/11	7	28	19-Sep-10	16-Oct-10	189.8
2448	AGL	Abergele & Pensarn	08_10/11	8	28	17-Oct-10	13-Nov-10	179.9
			09_10/11	9	28	14-Nov-10	11-Dec-10	178.8
4303	AYW	Aberystwyth	10_10/11	10	28	12-Dec-10	8-Jan-11	140.1
2552	ACR	Accrington	11_10/11	11	28	9-Jan-11	5-Feb-11	183.0
8600	AAT	Achanalt	12_10/11	12	28	6-Feb-11		
8601	ACN	Achnasheen	13_10/11	13	26	6-Mar-11		
8602	ACH	Achnashellach	01_11/12	1	30	1-Apr-11		
			02_11/12	2	28	1-May-11		
7662	ACK	Acklington	03_11/12	3	28	29-May-11		
			04_11/12	4	28	26-Jun-11		
7286	ACL	Acle	05_11/12	5	28	24-Jul-11		
4529	ACG	Acocks Green	06_11/12	6	28	21-Aug-11		
			07_11/12	7	28	18-Sep-11		
2287	ACB	Acton Bridge	08_11/12	8	28	16-Oct-11		
			09_11/12	9	28	13-Nov-11		
			10_11/12	10	28	11-Dec-11		
			11_11/12	11	28	8-Jan-12		
			12_11/12	12	28	5-Feb-12		
			13_11/12	13	28	4-Mar-12		
			01_12/13	1	28	1-Apr-12		
			02_12/13	2	28	29-Apr-12		
			03_12/13	3	28	27-May-12		
			04_12/13	4	28	24-Jun-12		
			05_12/13	5	28	22-Jul-12		
			06_12/13	6	28	13-Nov-12		
			07_12/13	7	28	10-Dec-12		
			08_12/13	8	28	7-Jan-13		
			09_12/13	9	28	4-Feb-13		
			10_12/13	10	28	3-Mar-13		
			11_12/13	11	28	31-Mar-13		
			12_12/13	12	28	19-Apr-13		
			13_12/13	13	28	16-May-13		

# E\_TOOMANYOPTIONS

ØMQ



Spark



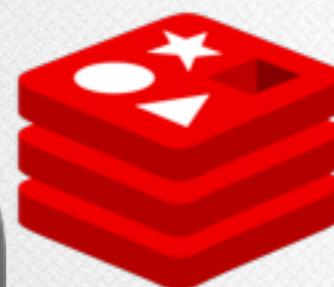
hadoop

# Prescriptive, Open

ØMQ



Spark



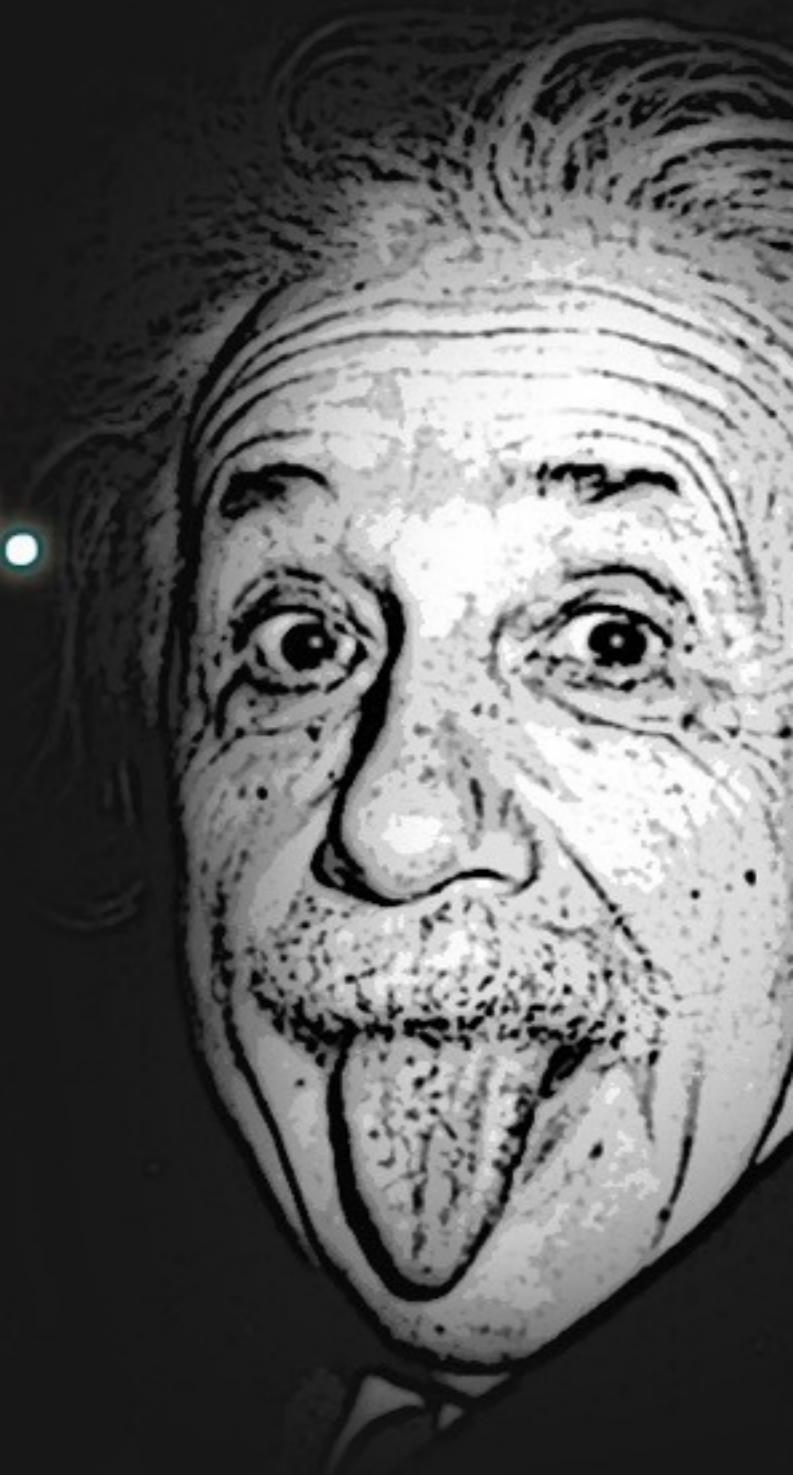
hadoop



# Everyone IS a genius.

“ But if you judge a fish on its ability to climb a tree,  
it will live its whole life believing that it is stupid ”

~ Albert Einstein ~

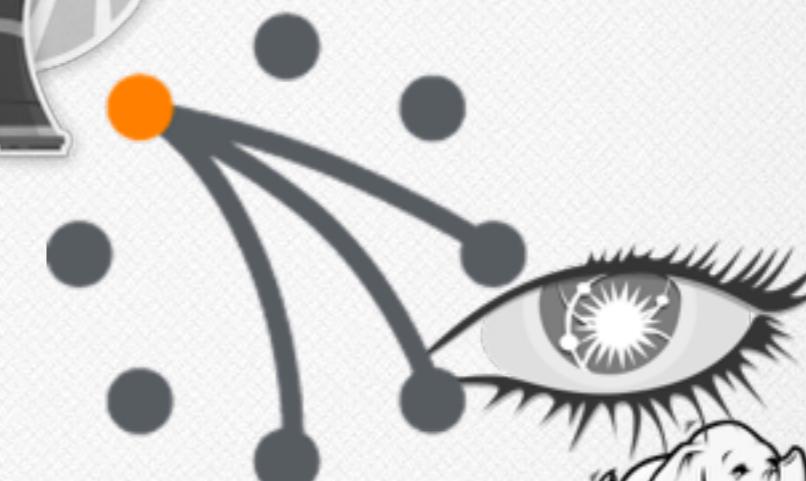


# Storage

ØMQ



Spark



ROUP

020 8805 4000  
www.underpin.com

M25 SCAFFOLDING

BUSH HILL CONSTRUCTION

U&M BASEMENTS

UNDERPIN & MAKEGOOD

U&M GROUP

020 8805 4000

M25 SCAFFOLDING

BUSH HILL CONSTRUCTION

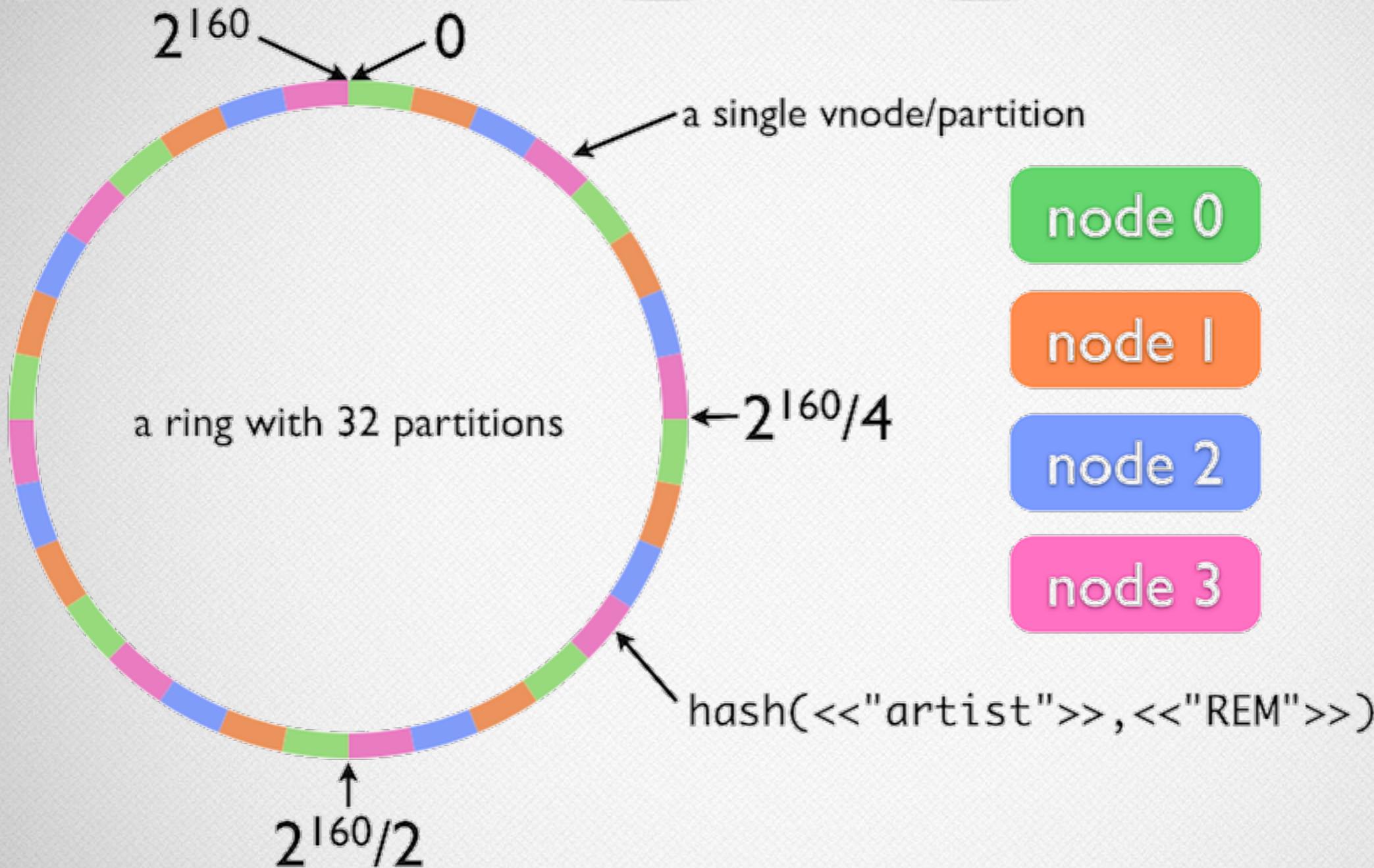
U&M BASEMENTS

UNDERPIN & MAKEGOOD

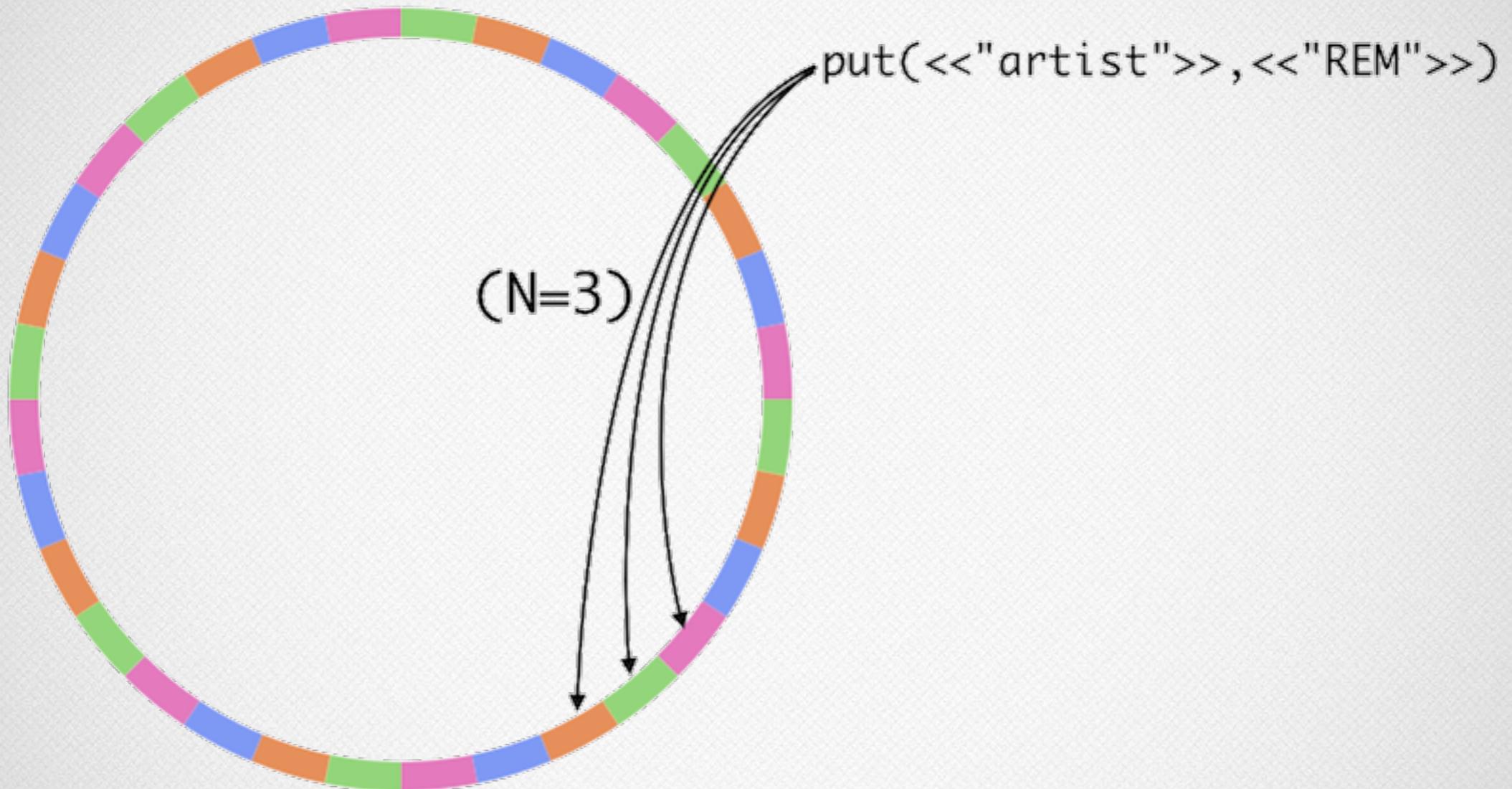
UDY



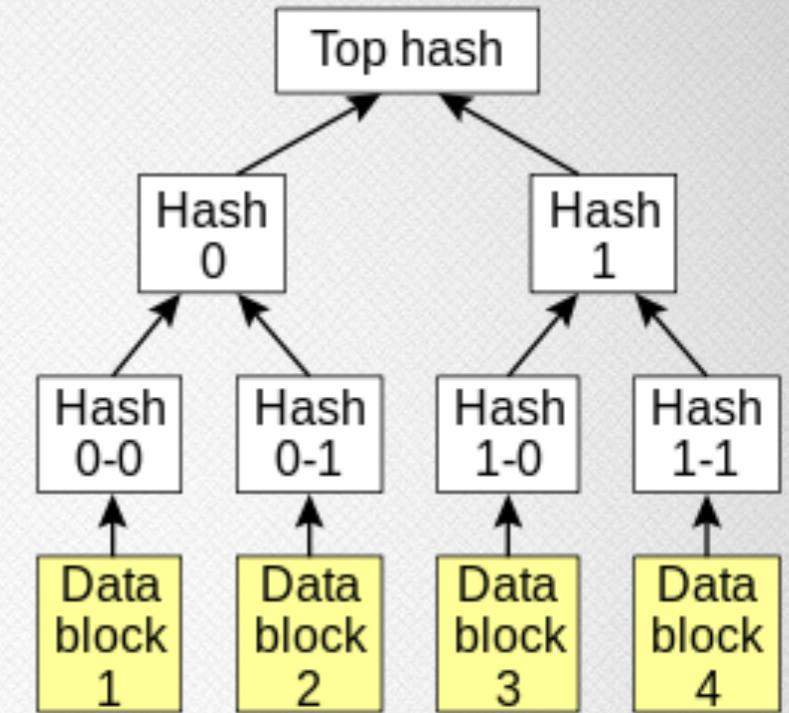
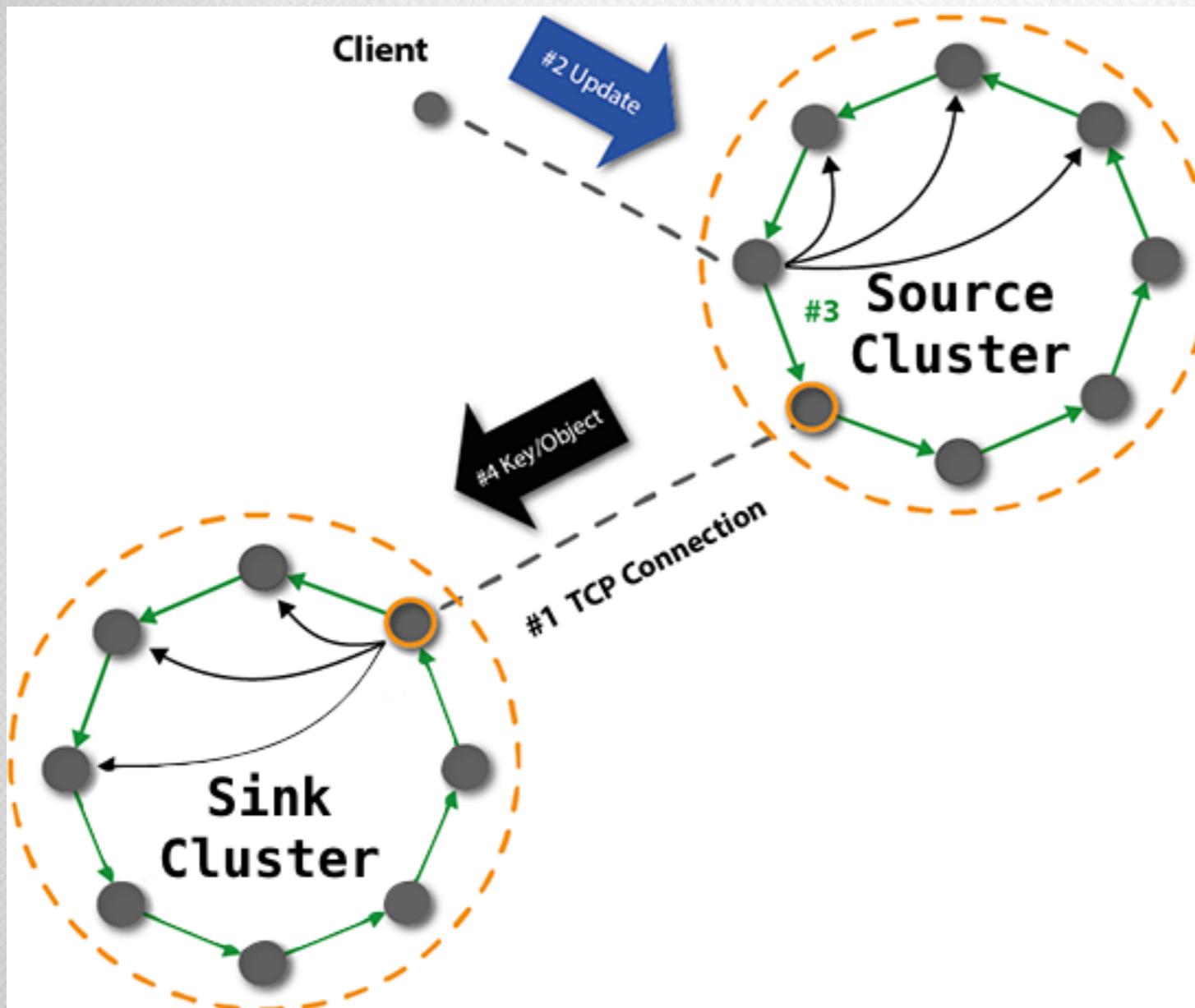
# Consistent Hash (Route)



# Ring Around The Rosie

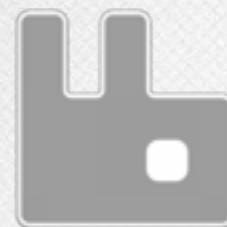


# Multi Data Center

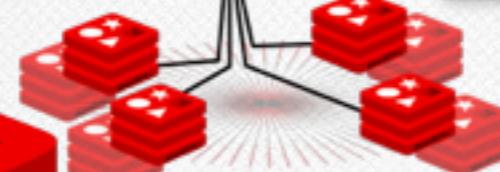


# CACHE

ØMQ



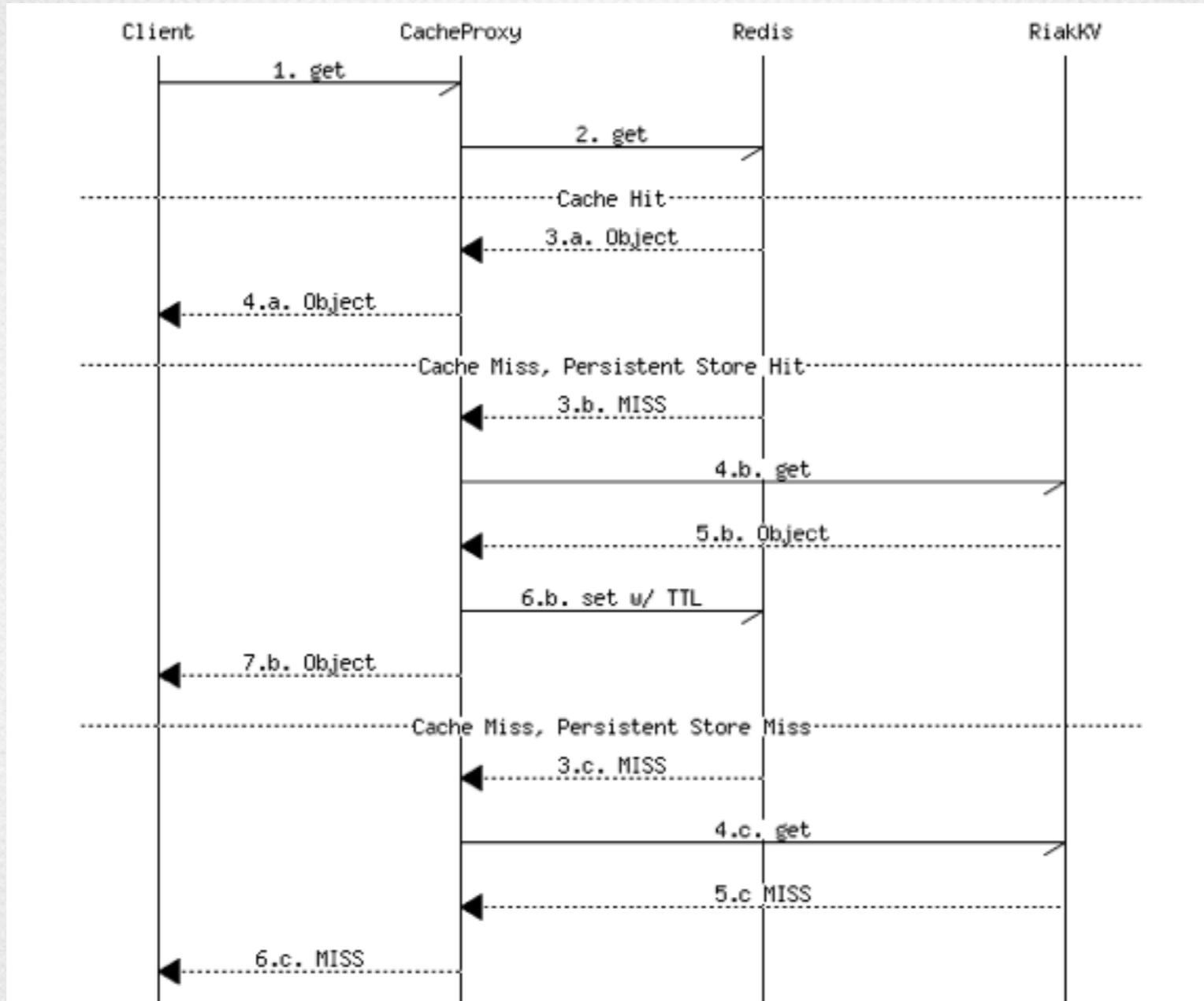
Spark



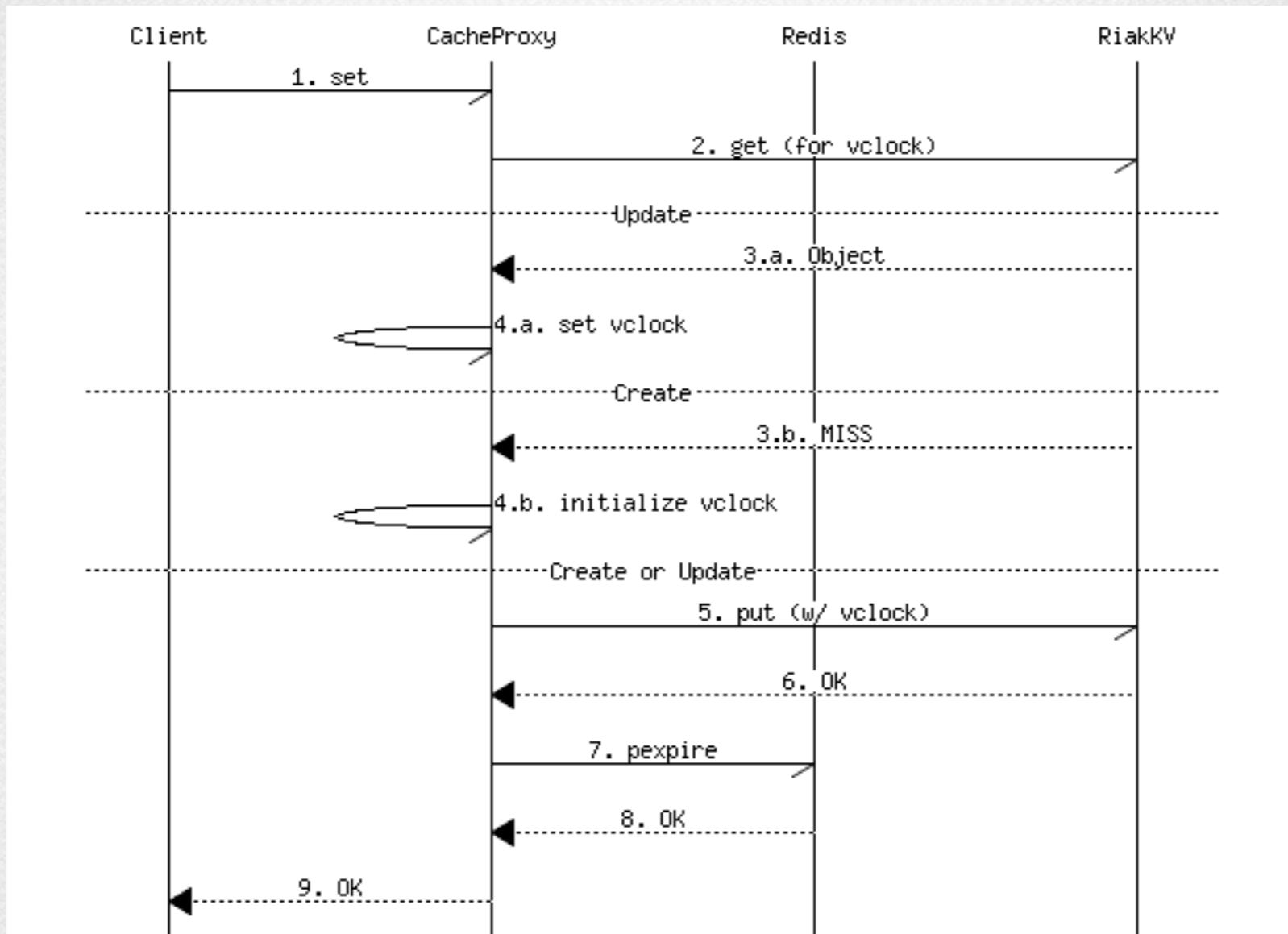
hadoop

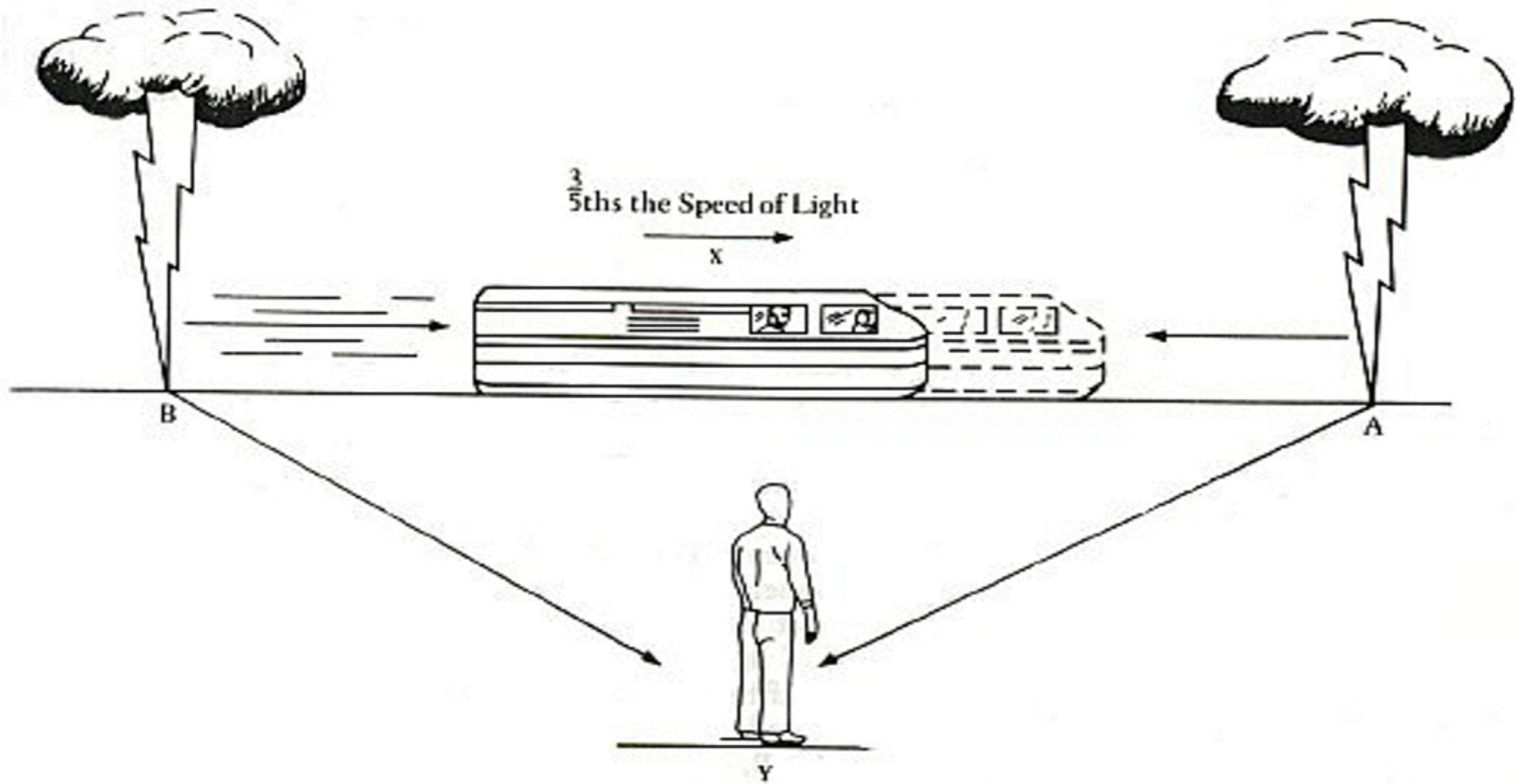


# Read-Through



# Write-Around

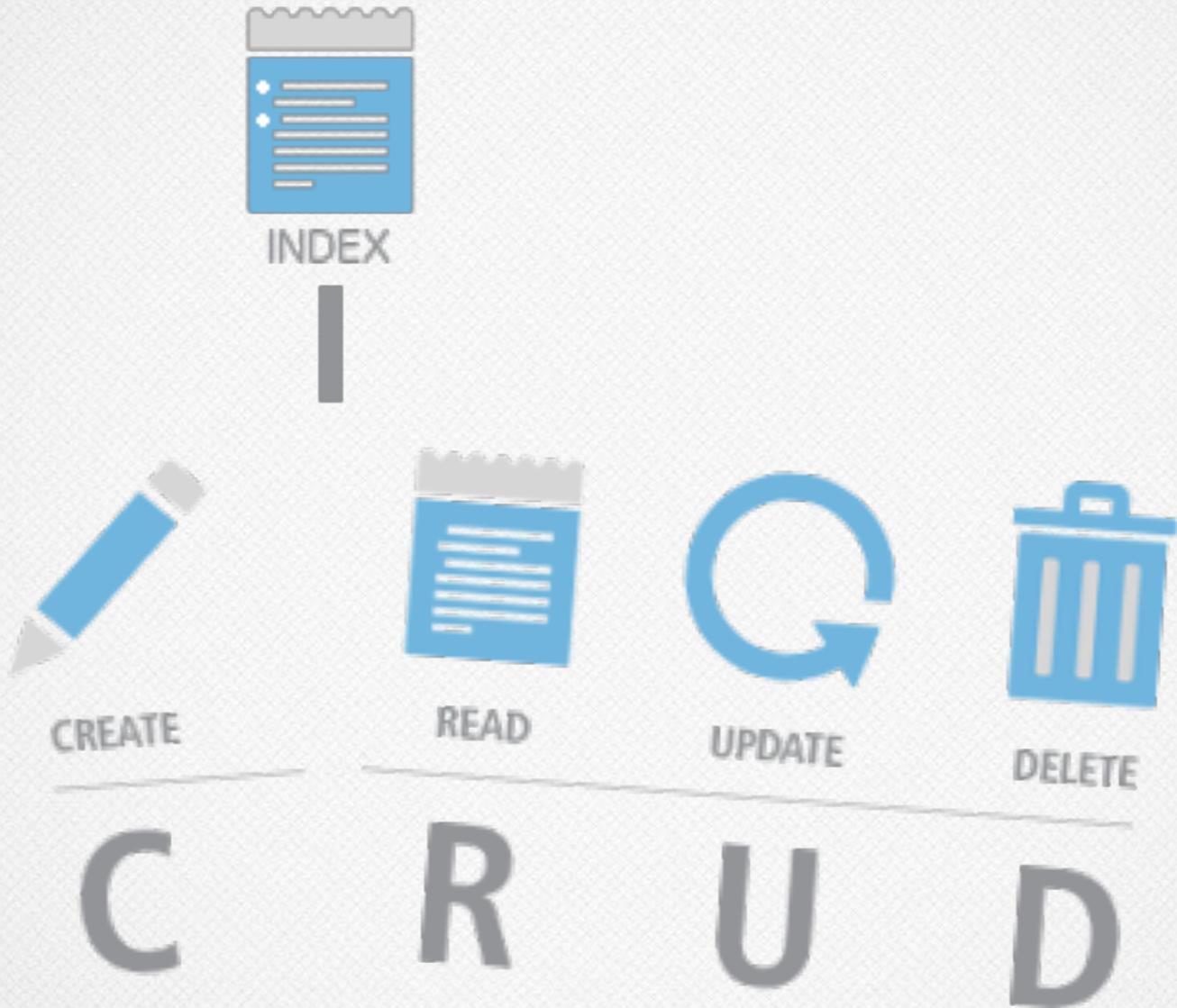




In Einstein's special train example, the light from A will arrive at X before that from B. Hence X will observe the lightning at A as happening before that at B. Y, however, will observe the bolts of lightning to be simultaneous. This is an example of how observations from reference frames moving at great speeds relative to each other reveal a different timing of events.

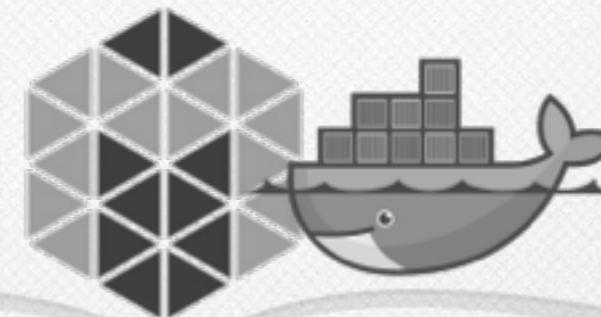
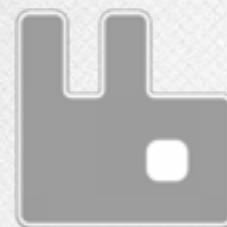


# Excuse Me Miss



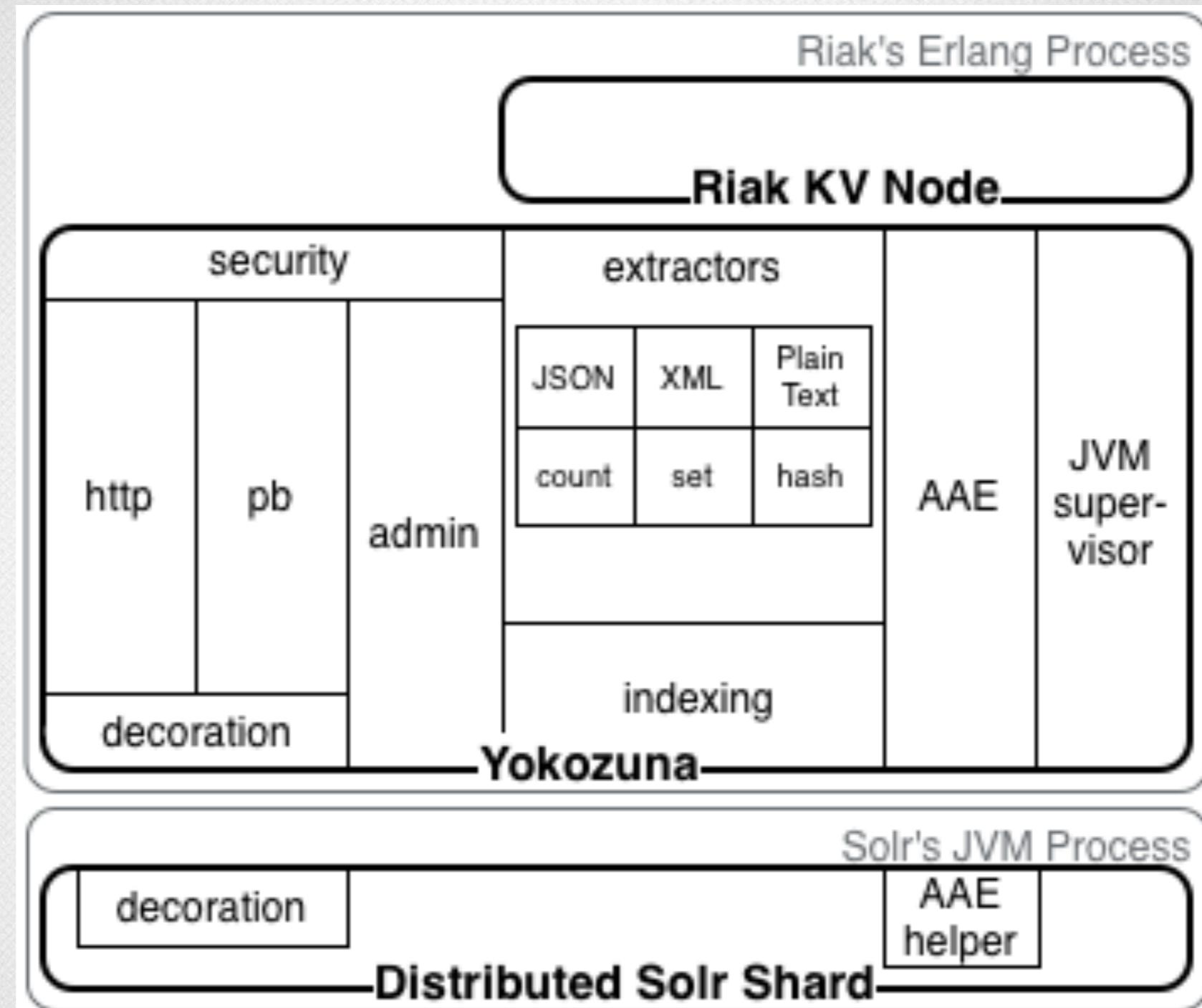
# Search

ØMQ

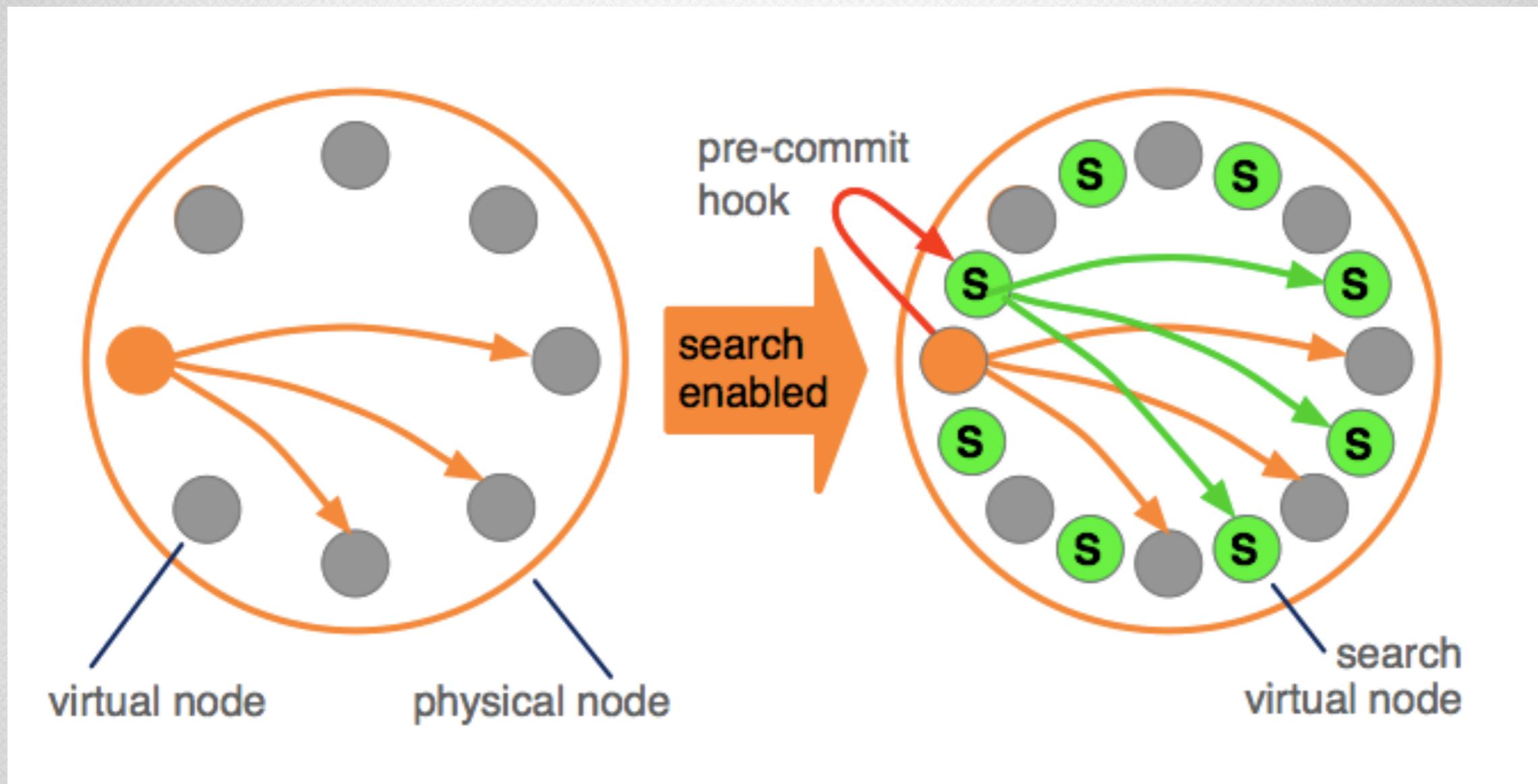


*hadoop*

# Ridealong



# Multi Service

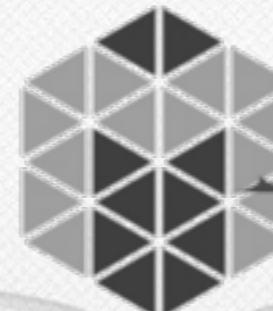
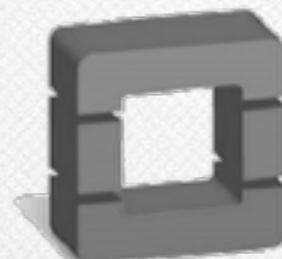


# Analysis

ØMQ



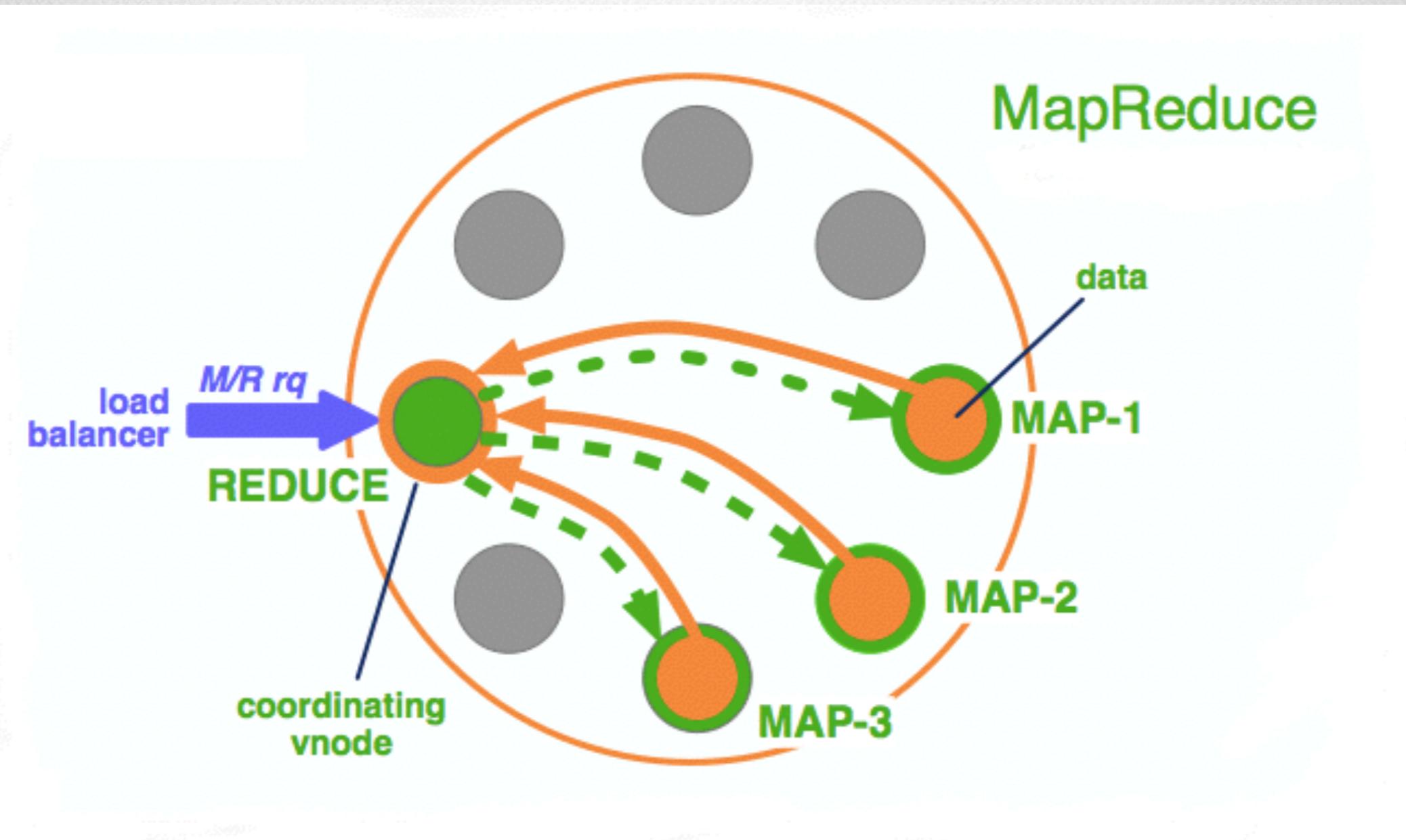
Spark



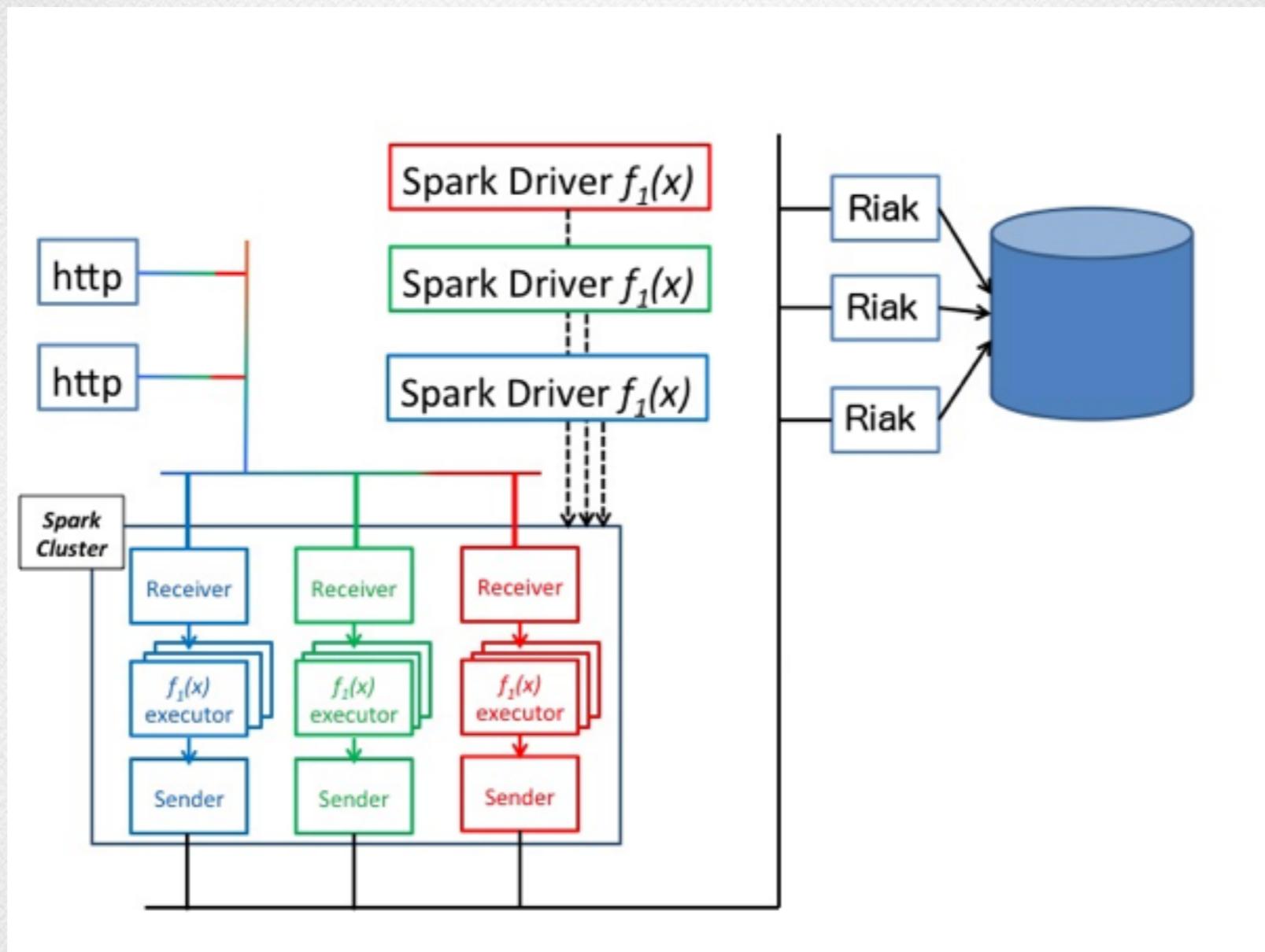
hadoop



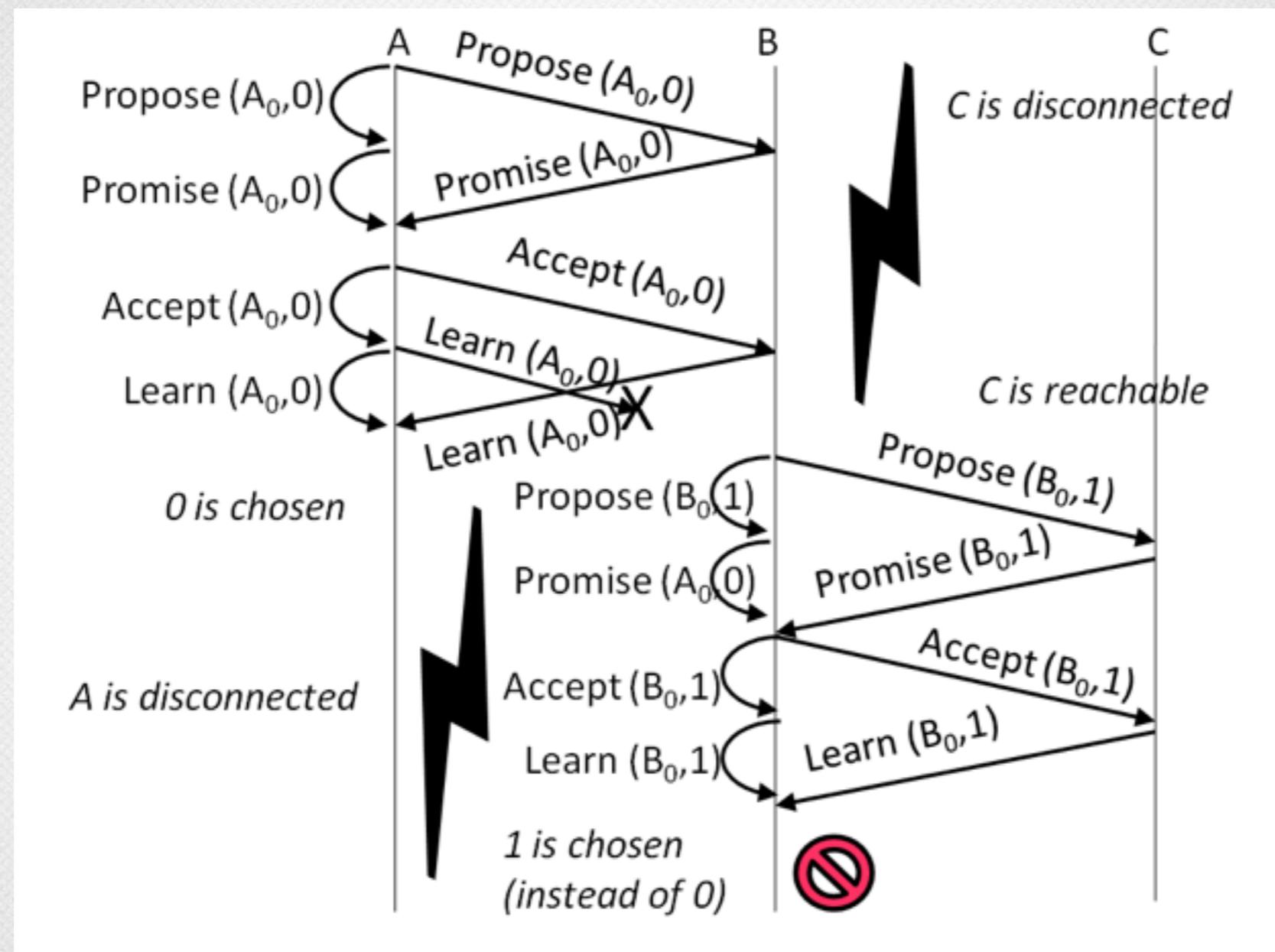
# Ship The Compute?



# Spread The Load

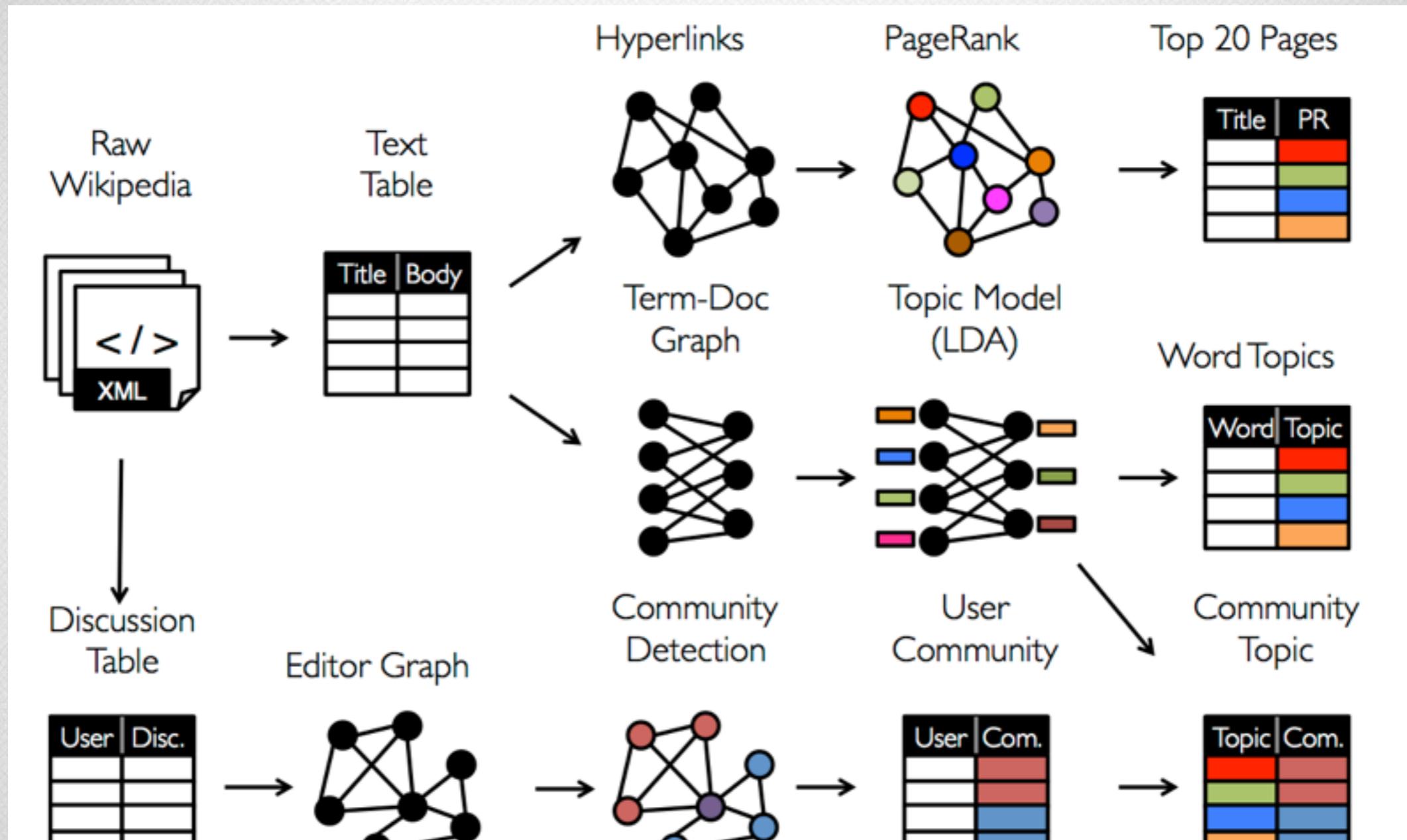


# No Pets

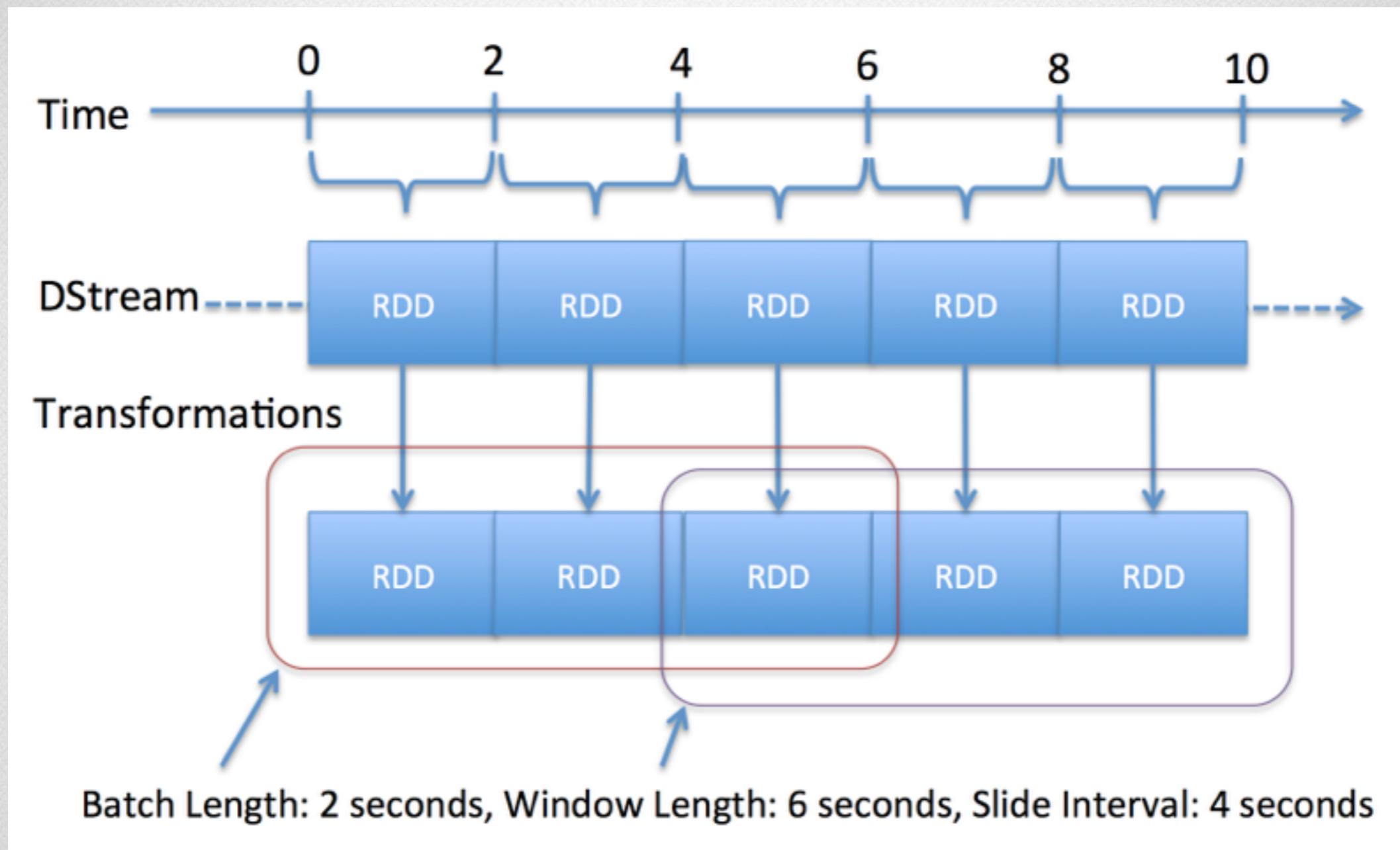


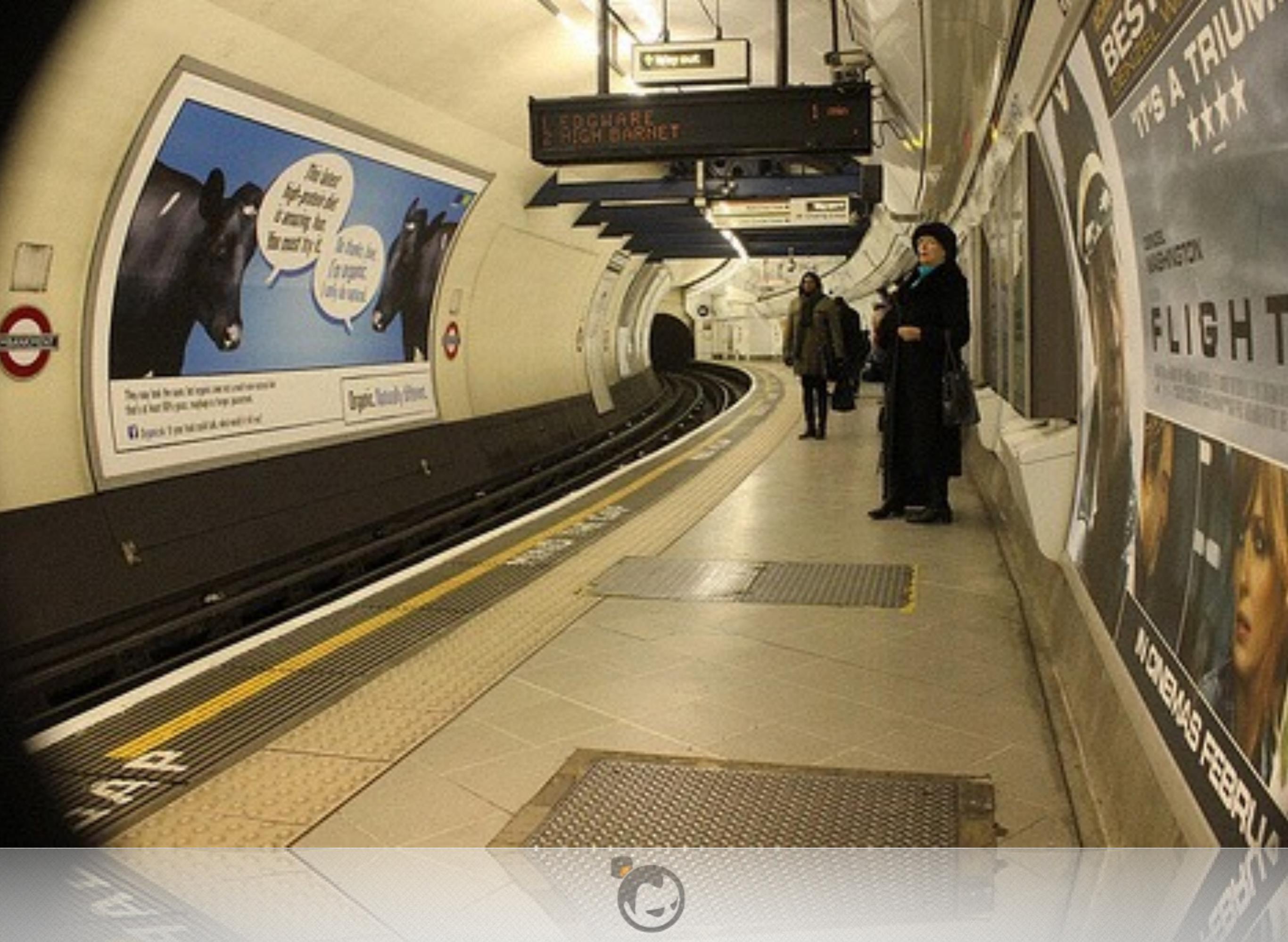


# Graph Reality



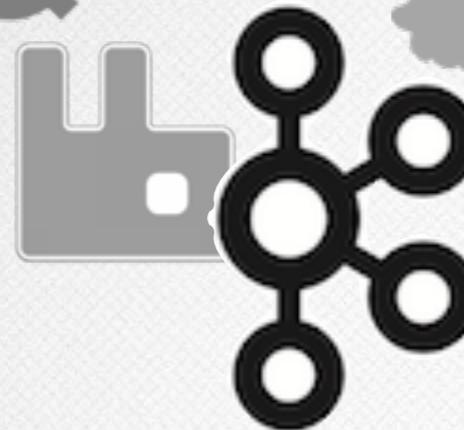
# Window Seat



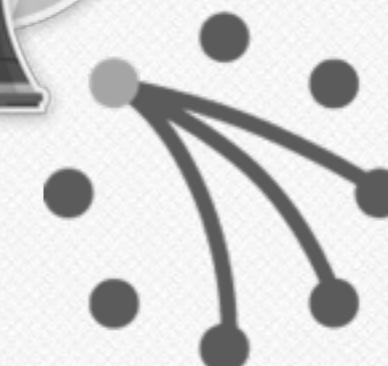


# QUEUE

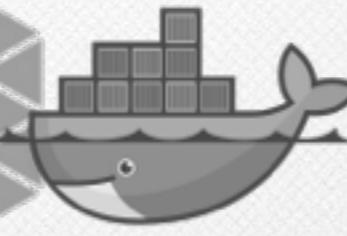
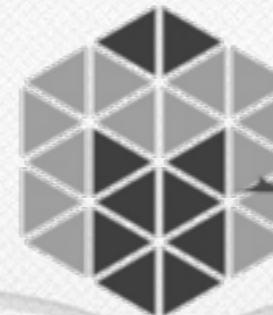
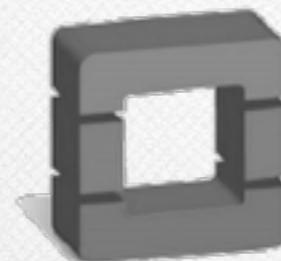
ØMQ



Spark



hadoop

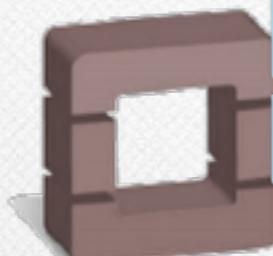
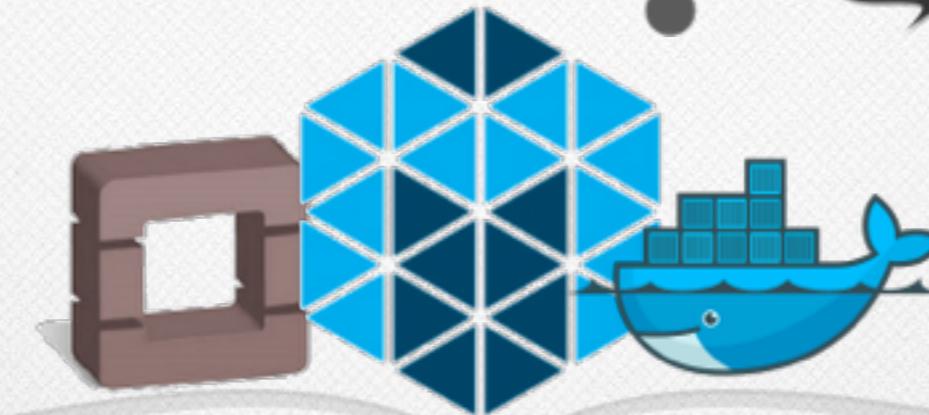


# HOST

ØMQ



Spark



hadoop



# For More Information

- London Datastore - <http://data.london.gov.uk/>
- Basho Data Platform - <http://docs.basho.com/dataplatform/latest/>
- Spark - <http://spark.apache.org/>
- Redis - <http://redis.io>
- Kafka - <http://kafka.apache.org/>
- Solr - <http://lucene.apache.org/solr/>
- Riak - <http://docs.basho.com/riak/latest/>



# REHASH?

