

NOSQL WITH CACHING, SEARCH AND REAL-TIME ANALYTICS

James Gorlick
Basho

bit.ly/bdp-on-the-under-lin



@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

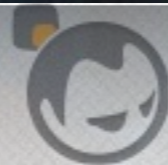


Word Count

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	Government Office Region (County or Unitary Authority District or Unitary Authority)	Local Government District or Unitary Authority	UK2 Spatial Unit	NUTS2 Spatial Unit	Station Facility Ownership	Station			
5131	ABW	Abbey Wood	London	Greater London	Greenwich	UKI2	Outer London	Southeastern			
3813	ABE	Aber	Wales - Cymru	Caerphilly - Caerffili	Caerphilly - Caerffili	UKL1	West Wales and the Valleys	Arriva Trains Wales			
3801	ACY	Abercynon		Rhondda Cynon Taff -	Rhondda Cynon Taff - Rhondda		West Wales and the				
3982	ABA	Aberdare									
8976	ABD	Aberdeen									
9090	AUR	Aberdour									
4435	AVY	Aberdovey									
4440	ABH	Abererch									
3602	AGV	Abergavenny									
2448	AGL	Abergele & Pensarn									
4303	AYW	Aberystwyth									
2552	ACR	Accrington									
8600	AAT	Achanalt									
8601	ACN	Achnasheen									
8602	ACH	Achnashellach									
7662	ACK	Acklington									
7286	ACL	Acle									
4529	ACG	Acocks Green									
2287	ACB	Acton Bridge									
Period and Financial year	Reporting Period	Days in period	Period beginning	Period ending	Bus journeys (m)	Underground journeys (m)	DLR Journeys (m)	Tram Journeys (m)	Overground Journeys (m)	Emirates Airline Journeys (m)	
01_10/11	1	31	1-Apr-10	1-May-10	189.1	90.5	6.3	2.3			
02_10/11	2	28	2-May-10	29-May-10	181.6	84.5	5.8	2.2			
03_10/11	3	28	30-May-10	26-Jun-10	175.9	84.3	5.8	2.1			
04_10/11	4	28	27-Jun-10	24-Jul-10	183.4	86.5	6.1	2.1			
05_10/11	5	28	25-Jul-10	21-Aug-10	160.4	82.9	5.8	2.0			
06_10/11	6	28	22-Aug-10	18-Sep-10	175.8	80.9	5.5	2.0			
07_10/11	7	28	19-Sep-10	16-Oct-10	189.8	88.7	6.3	2.3			
08_10/11	8	28	17-Oct-10	13-Nov-10	179.9	90.3	6.7	2.2	5.6		
09_10/11	9	28	14-Nov-10	11-Dec-10	178.8	90.6	6.4	2.3	5.4		
10_10/11	10	28	12-Dec-10	8-Jan-11	140.1	72.5	4.8	1.8	3.5		
11_10/11	11	28	9-Jan-11	5-Feb-11	183.0	84.4	6.3	2.1	5.2		
12_10/11	12	28	6-Feb-								
13_10/11	13	26	6-Mar-								
Period and Financial year	Reporting Period	Days in period	Period beginning	Period ending	Month	Number of Lost Customer Hours	Operated Kms (Peak and Off Peak)	% of Scheduled Operated	Excess Journey (mins)		
01_11/12	1	30	1-Apr-								
02_11/12	2	28	1-May-								
03_11/12	3	28	29-May-								
04_11/12	4	28	26-Jun-	02_11/12	2	28	28-May-11	May-11	2,322,089	5,444,365	97.3%
05_11/12	5	28	24-Jul-	03_11/12	3	28	25-Jun-11	Jun-11	1,929,445	5,463,580	97.6%
06_11/12	6	28	21-Aug-	04_11/12	4	28	23-Jul-11	Jul-11	2,300,769	5,465,180	96.8%
07_11/12	7	28	18-Sep-	05_11/12	5	28	20-Aug-11	Aug-11	2,059,198	5,595,332	97.1%
08_11/12	8	28	16-Oct-	06_11/12	6	28	17-Sep-11	Sep-11	2,462,085	5,516,502	96.8%
09_11/12	9	28	13-Nov-	07_11/12	7	28	15-Oct-11	Oct-11	2,155,102	5,512,264	97.5%
10_11/12	10	28	11-Dec-	08_11/12	8	28	12-Nov-11	Nov-11	2,054,499	5,657,701	97.5%
11_11/12	11	28	8-Jan-	09_11/12	9	28	10-Dec-11	Dec-11	1,989,066	5,694,703	97.4%
12_11/12	12	28	5-Feb-	10_11/12	10	28	7-Jan-12	Jan-12	2,043,901	5,275,158	96.4%
13_11/12	13	28	4-Mar-	11_11/12	11	28	4-Feb-12	Feb-12	2,073,884	5,730,948	97.1%
01_12/13	1	28	1-Apr-	12_11/12	12	28	3-Mar-12	Mar-12	2,604,341	5,559,547	95.9%
02_12/13	2	28	29-Apr-	13_11/12	13	28	31-Mar-12	Mar-12	1,997,431	5,682,740	97.4%
03_12/13	3	28	27-May-								
04_12/13	4	28	24-Jun-								
05_12/13	5	28	22-Jul-								

E_TOO MANY OPTIONS

ØMQ



Spark



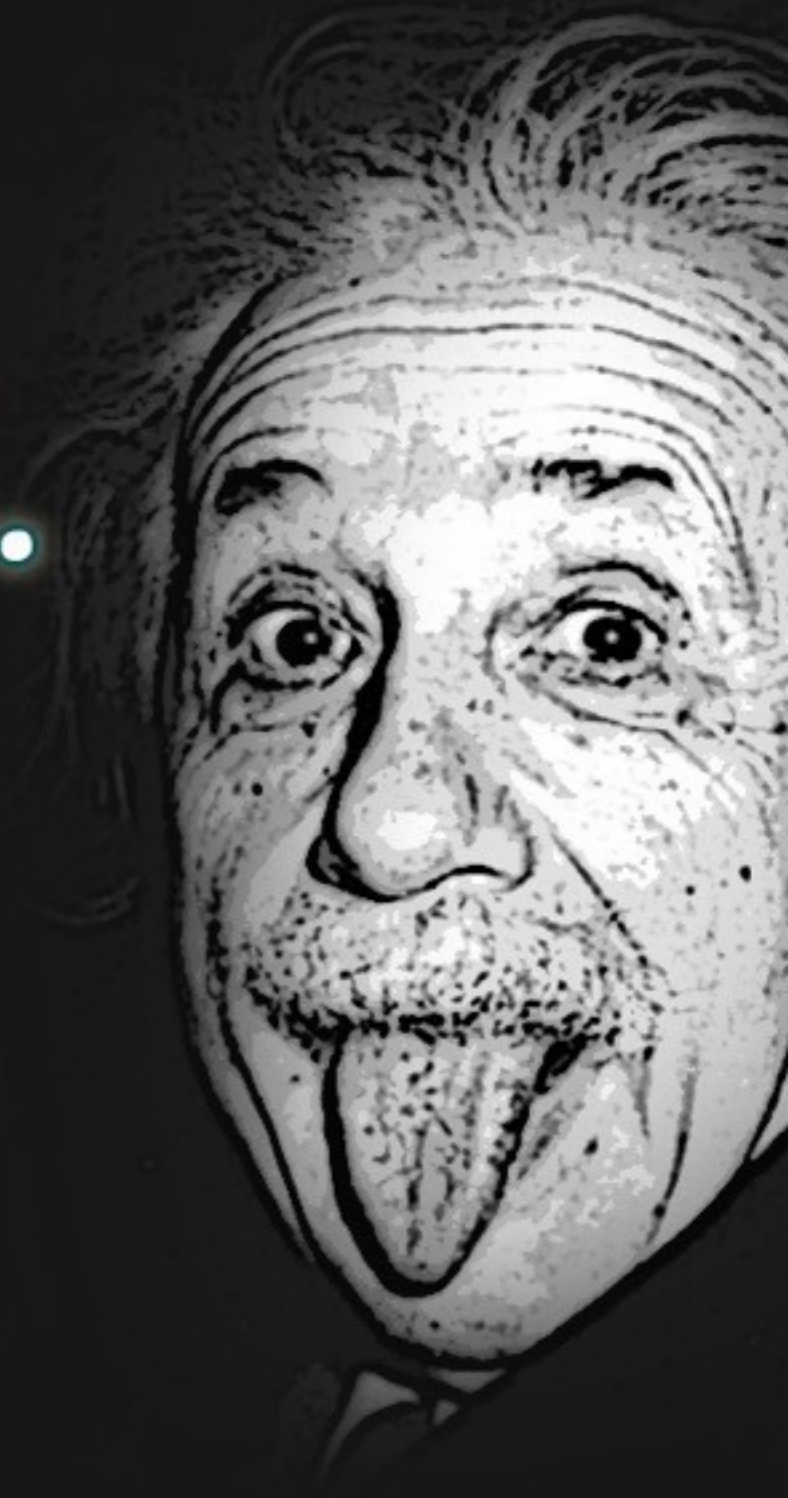
hadoop



Prescriptive, Open



Everyone **is**
a genius.

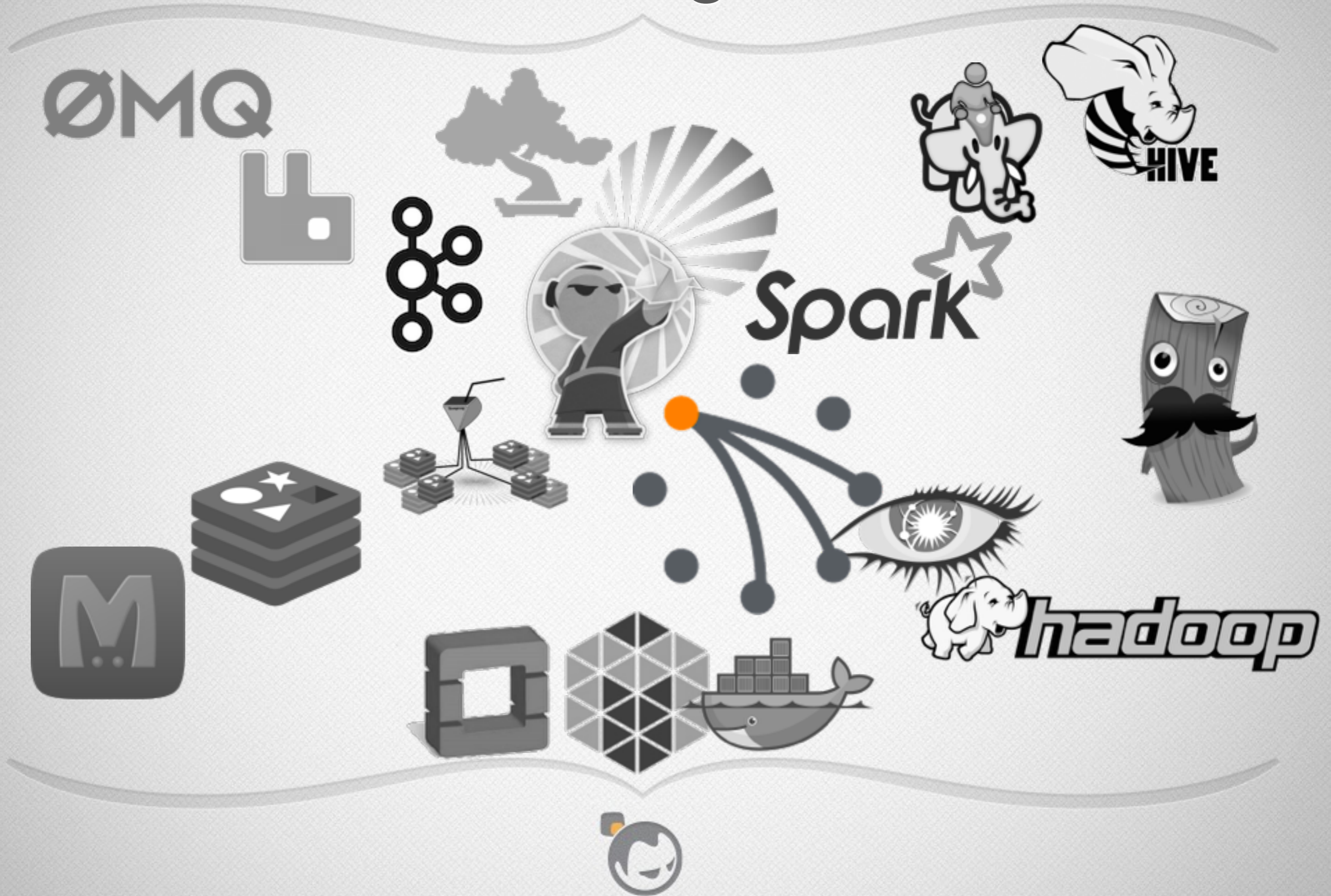


“ But if you judge a fish on its ability to climb a tree,
it will live its whole life believing that it is stupid ”

~ Albert Einstein ~

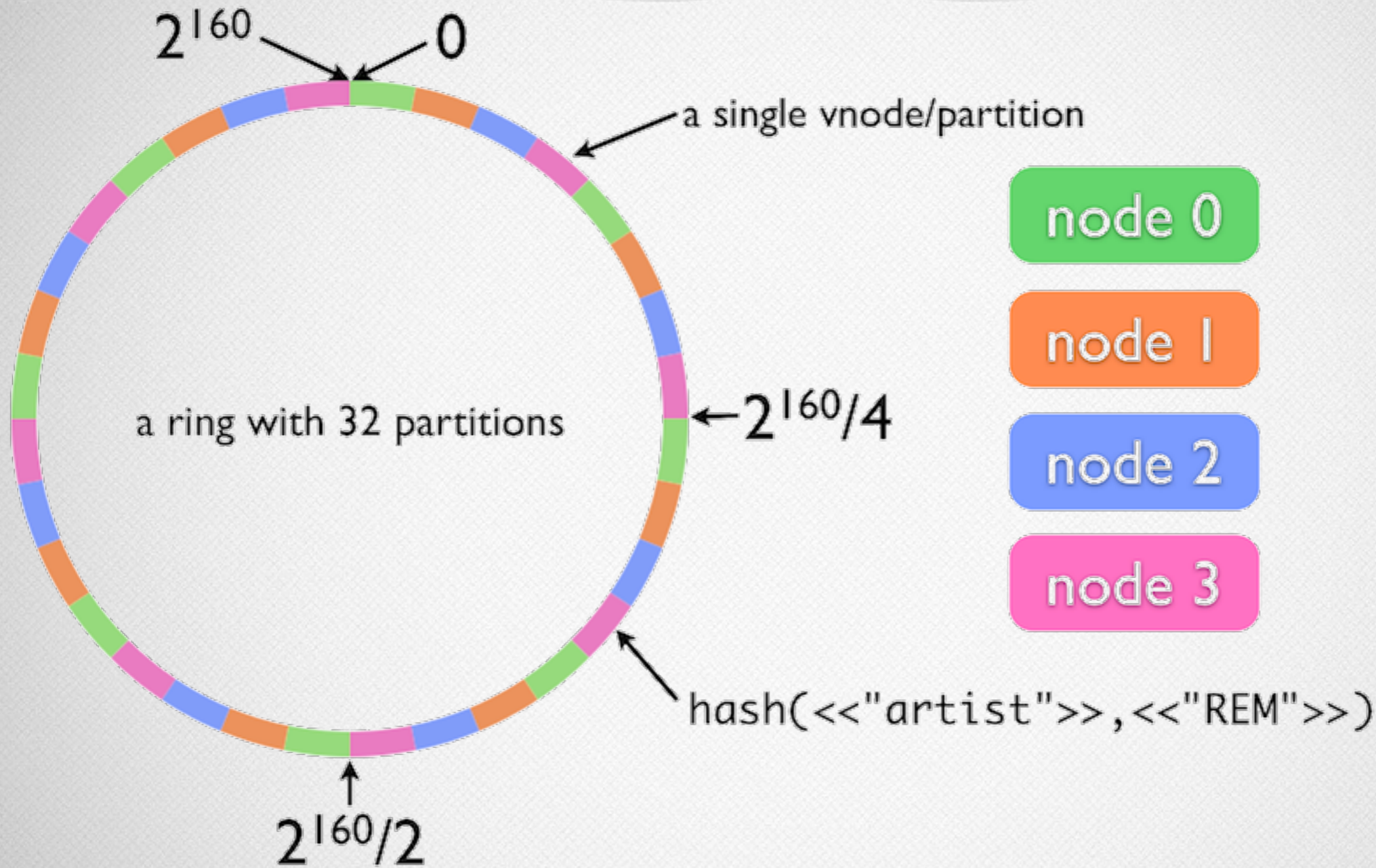


Storage

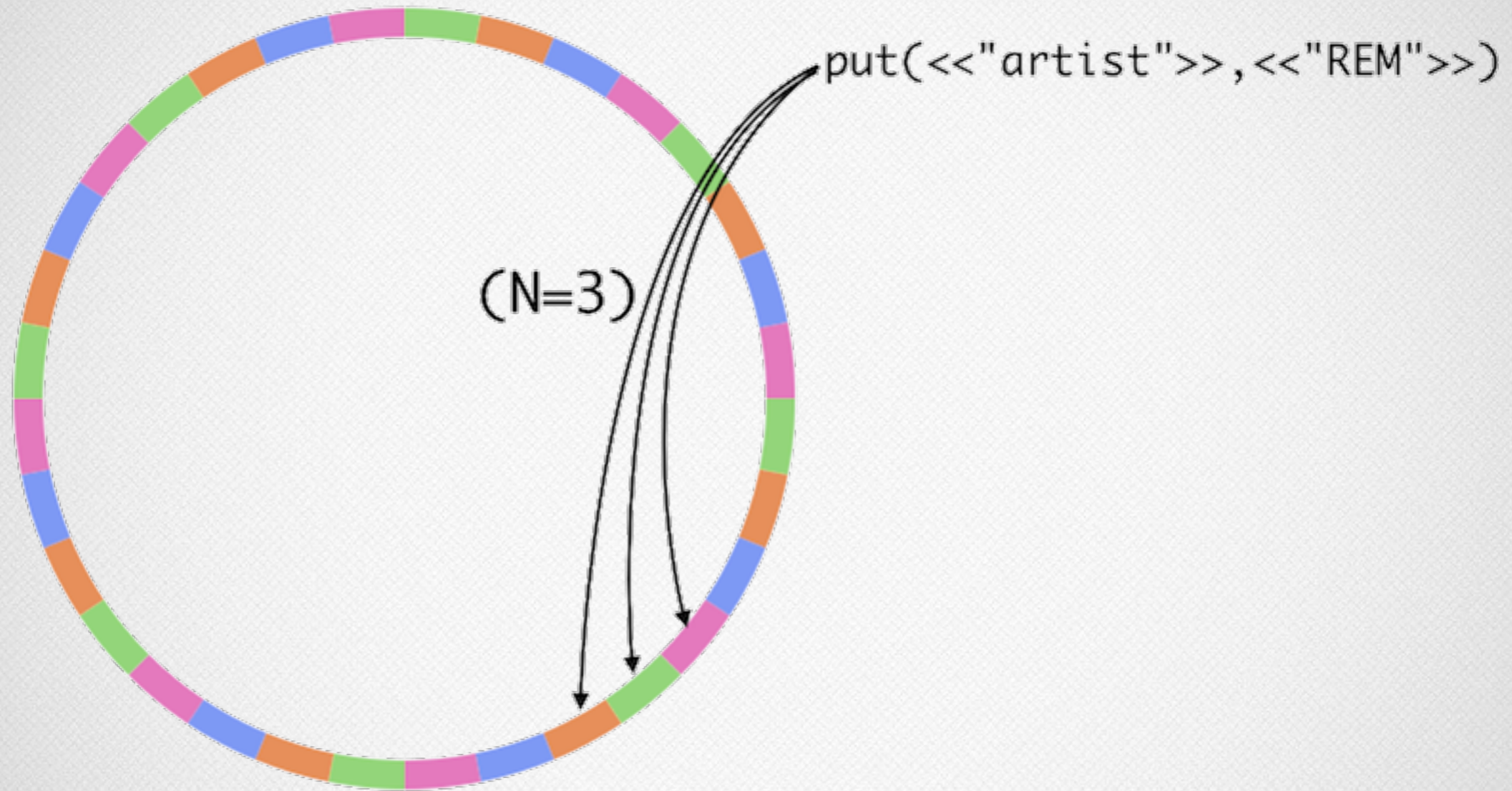




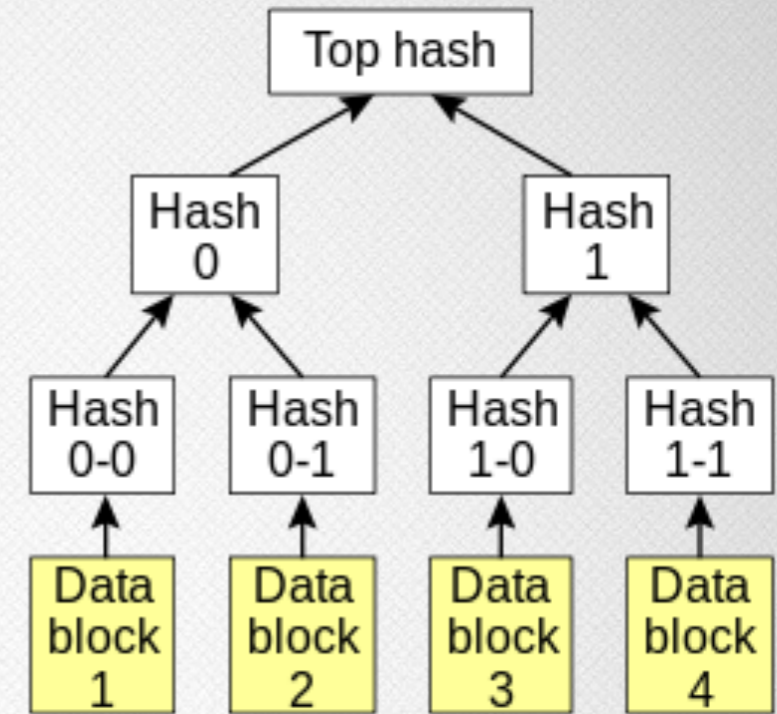
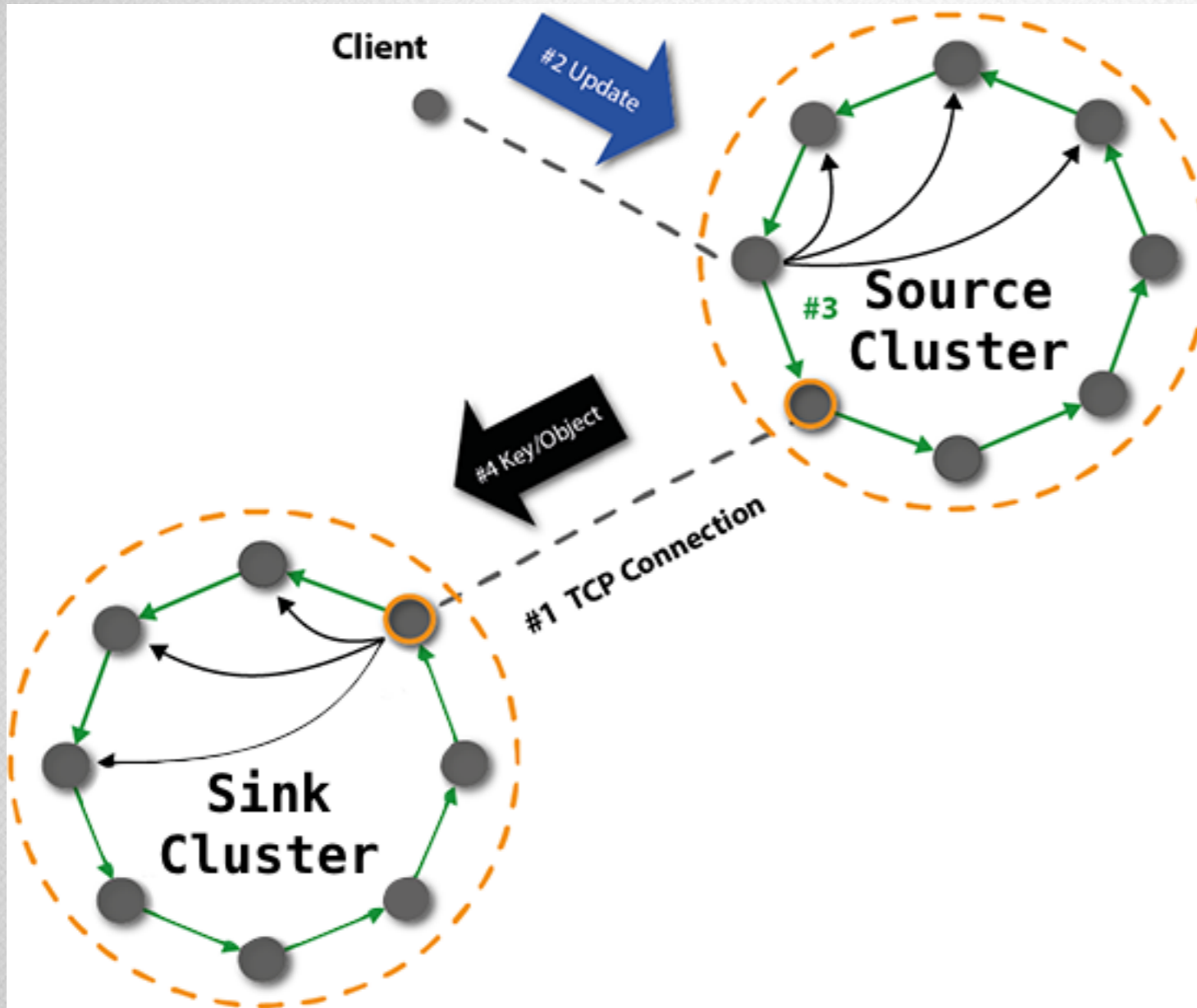
Consistent Hash (Route)



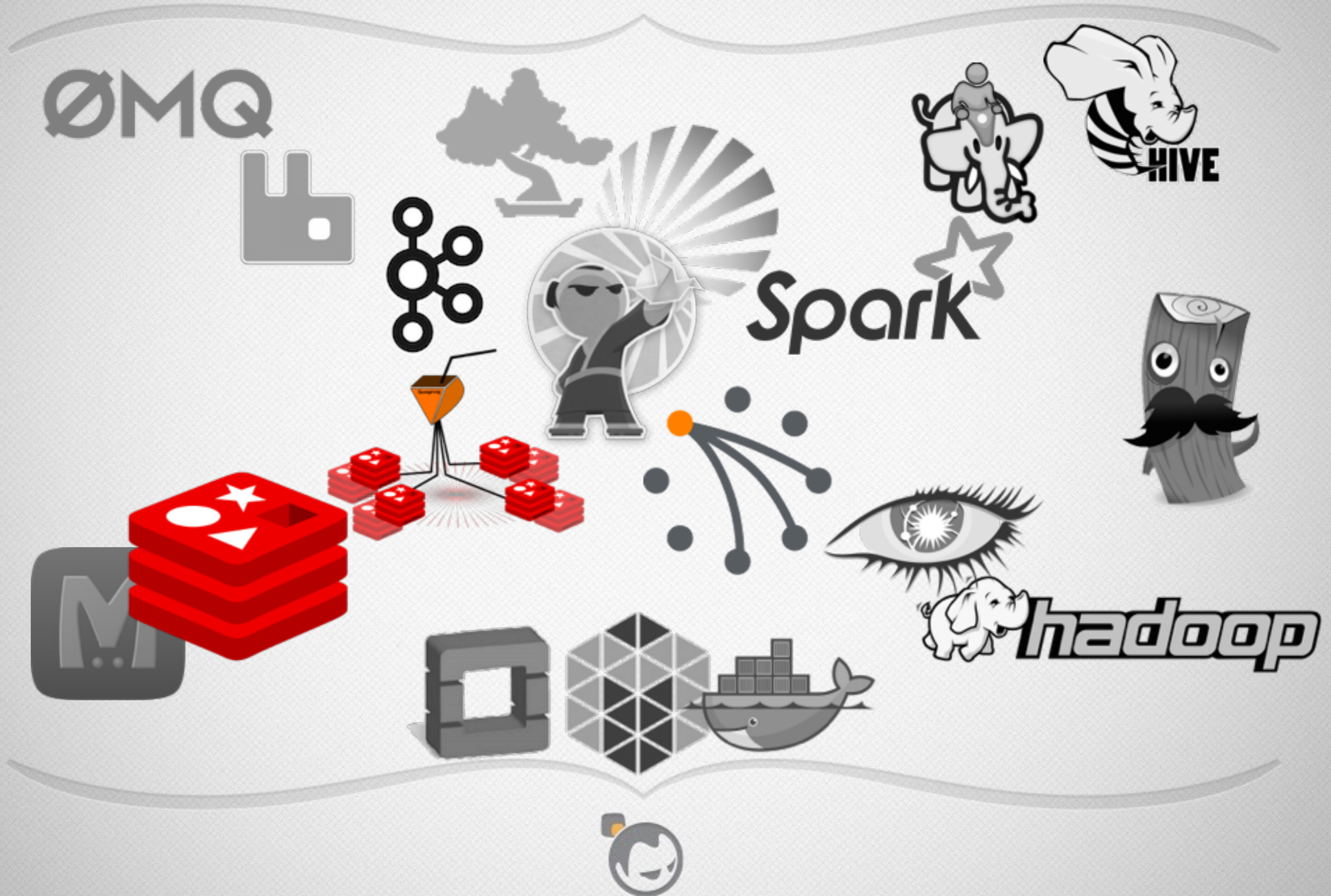
Ring Around The Rosie



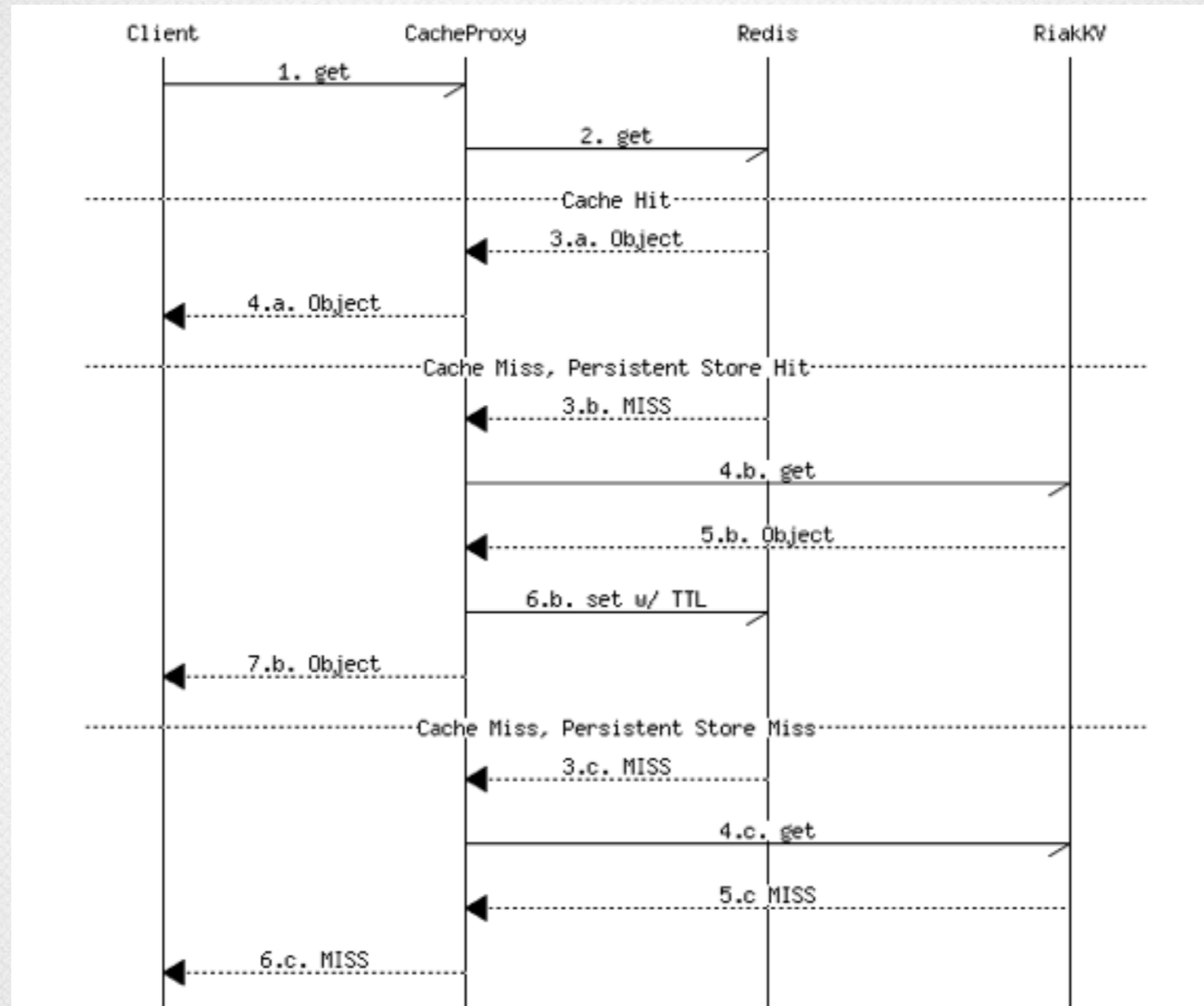
Multi Data Center



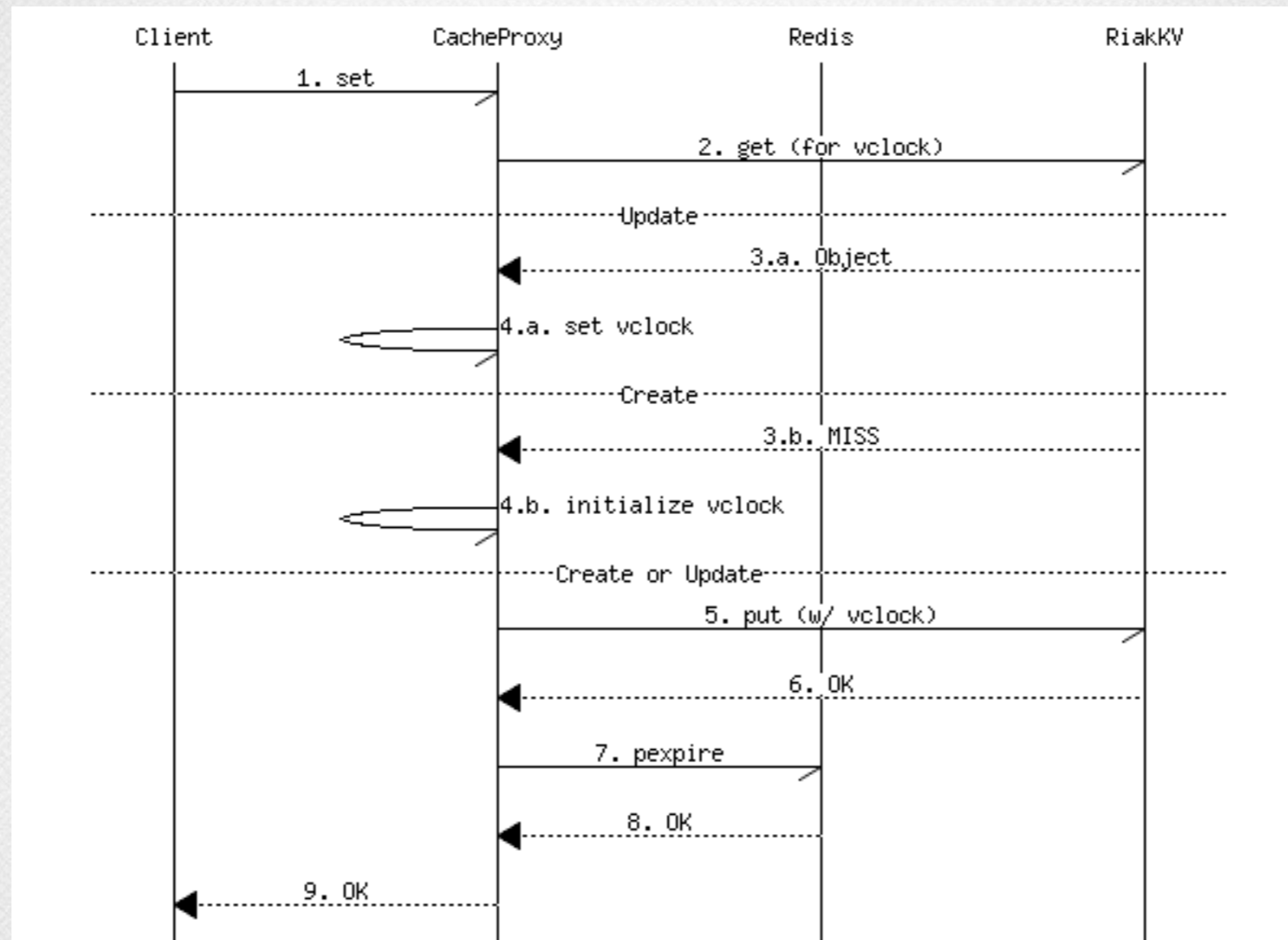
CACHE

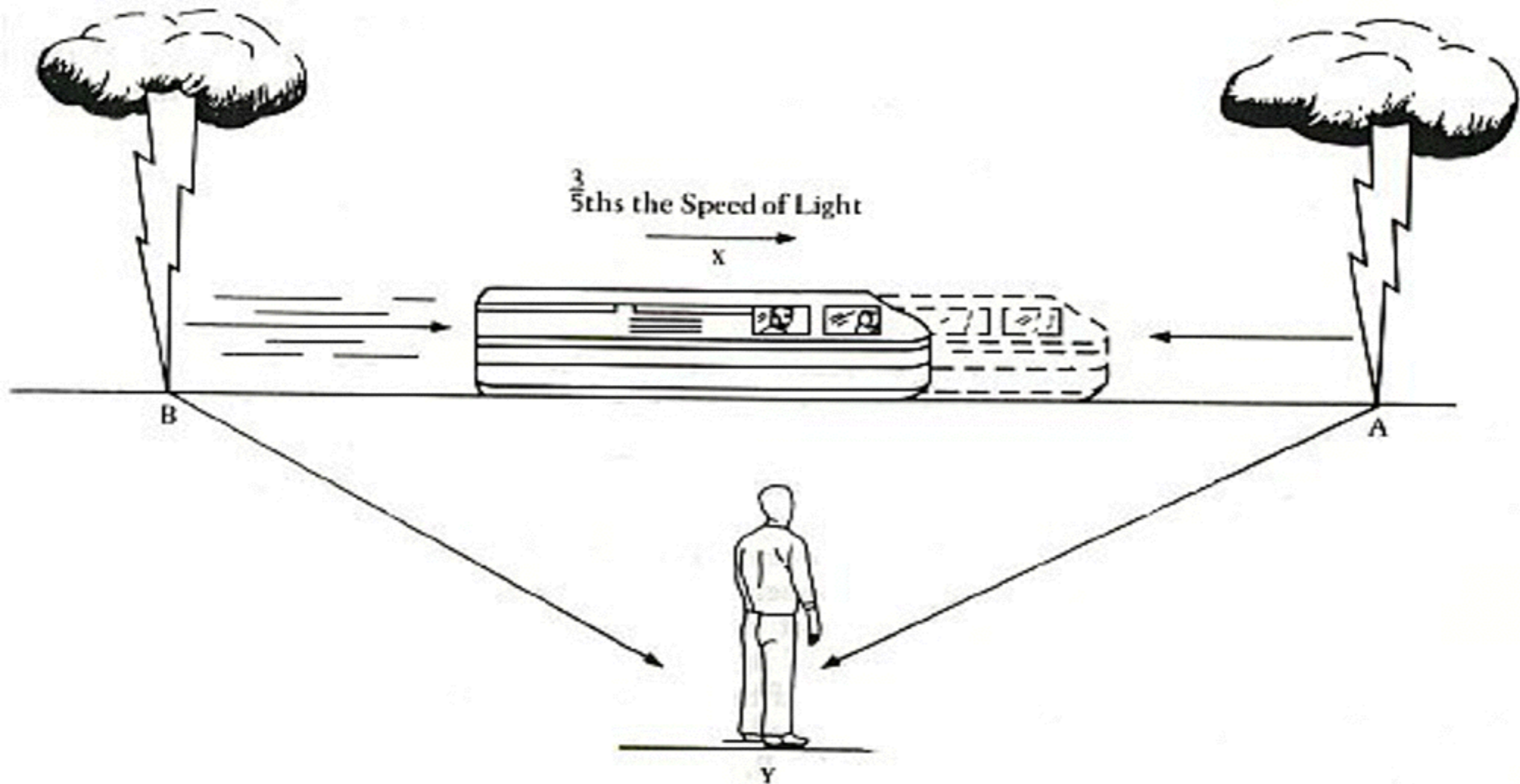


Read-Through



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



INDEX



CREATE

C



READ

R



UPDATE

U

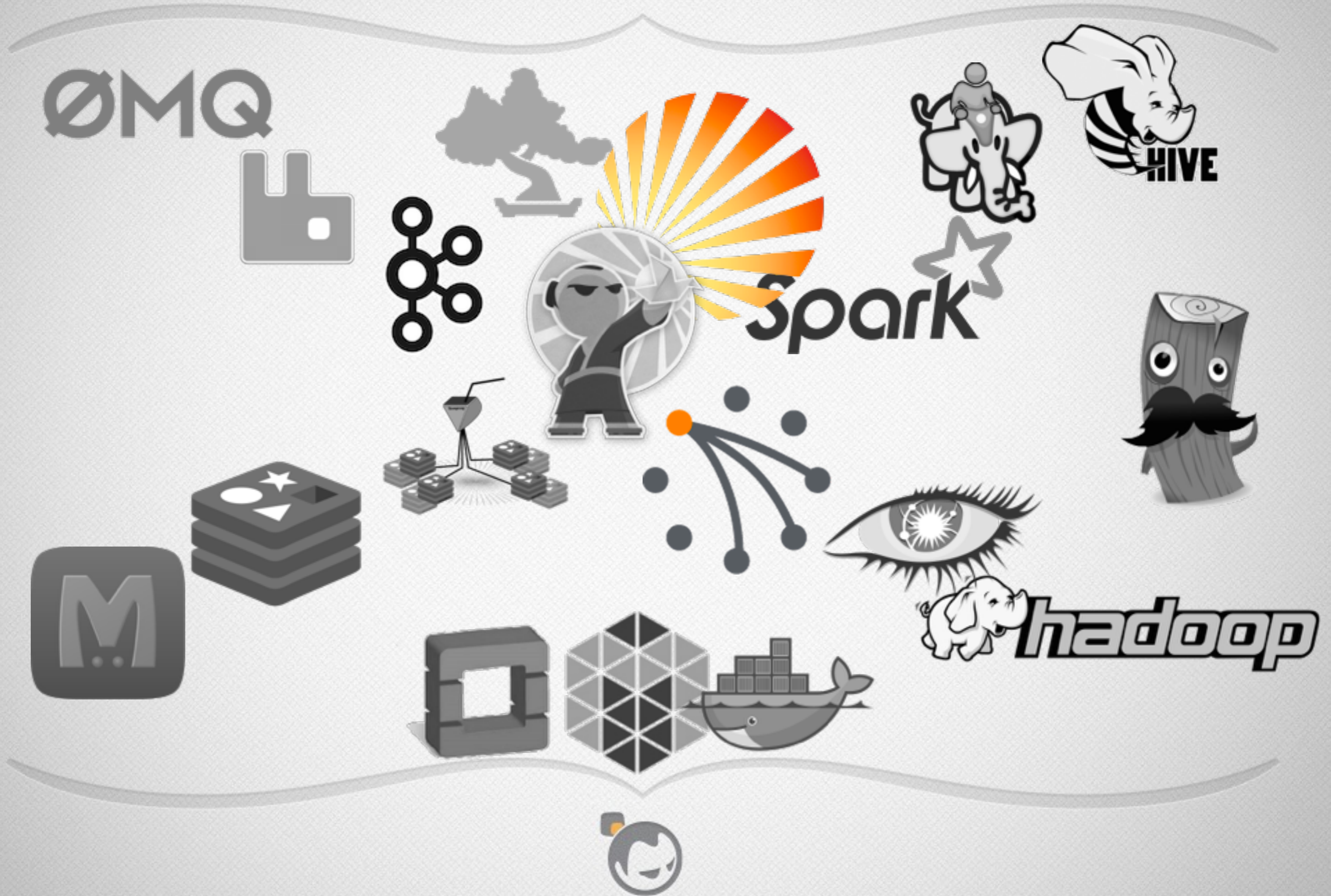


DELETE

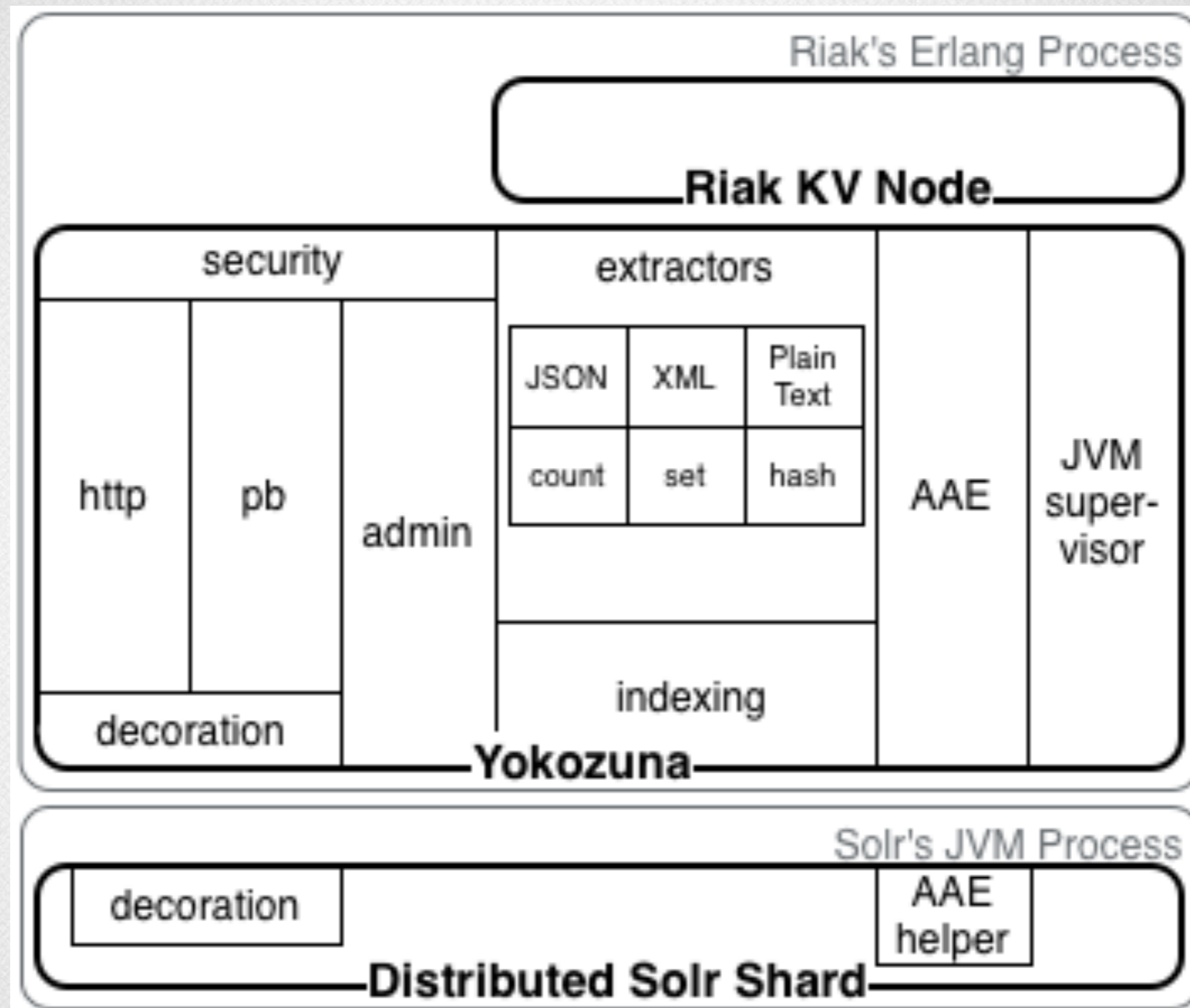
D



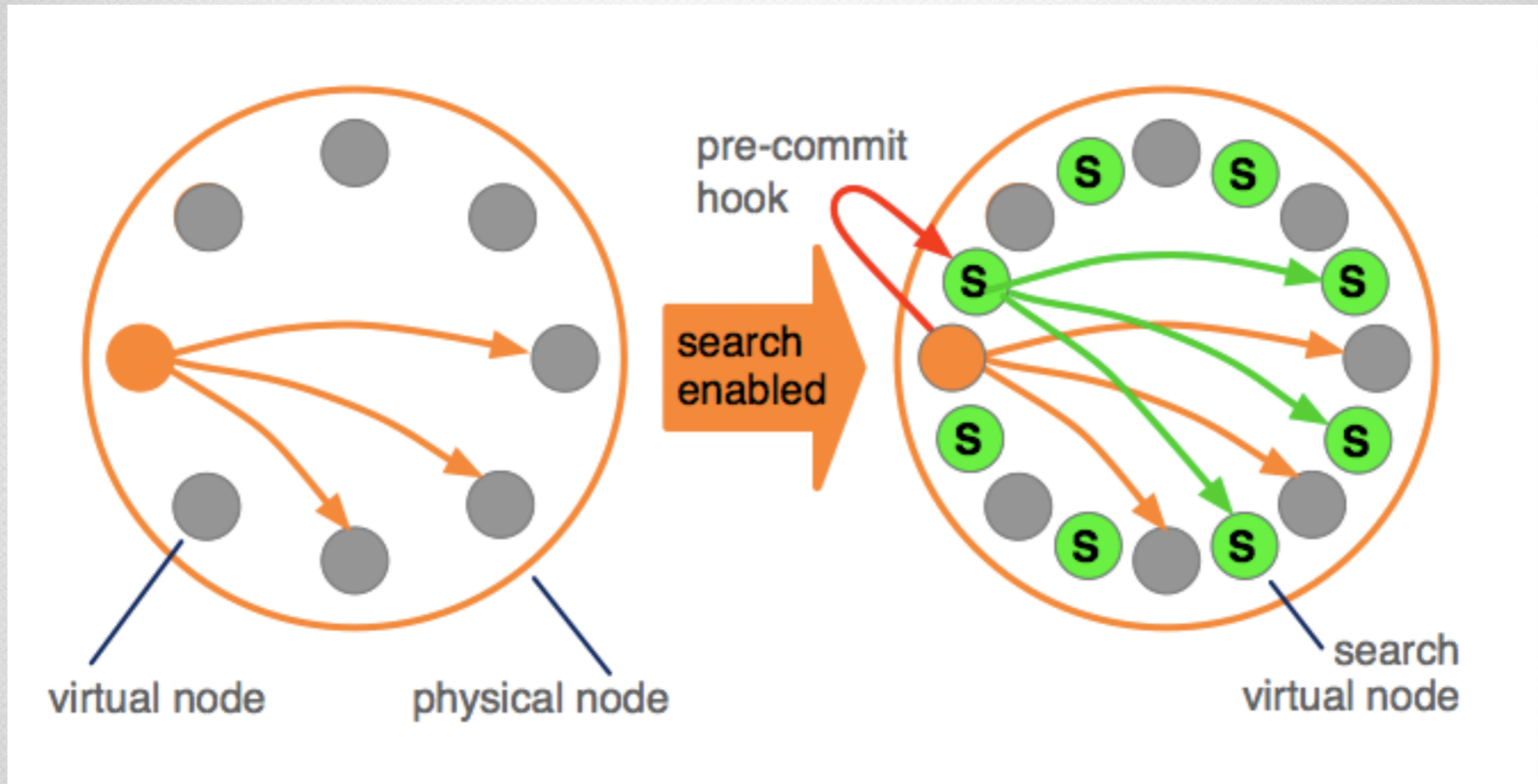
Search



Ridealong



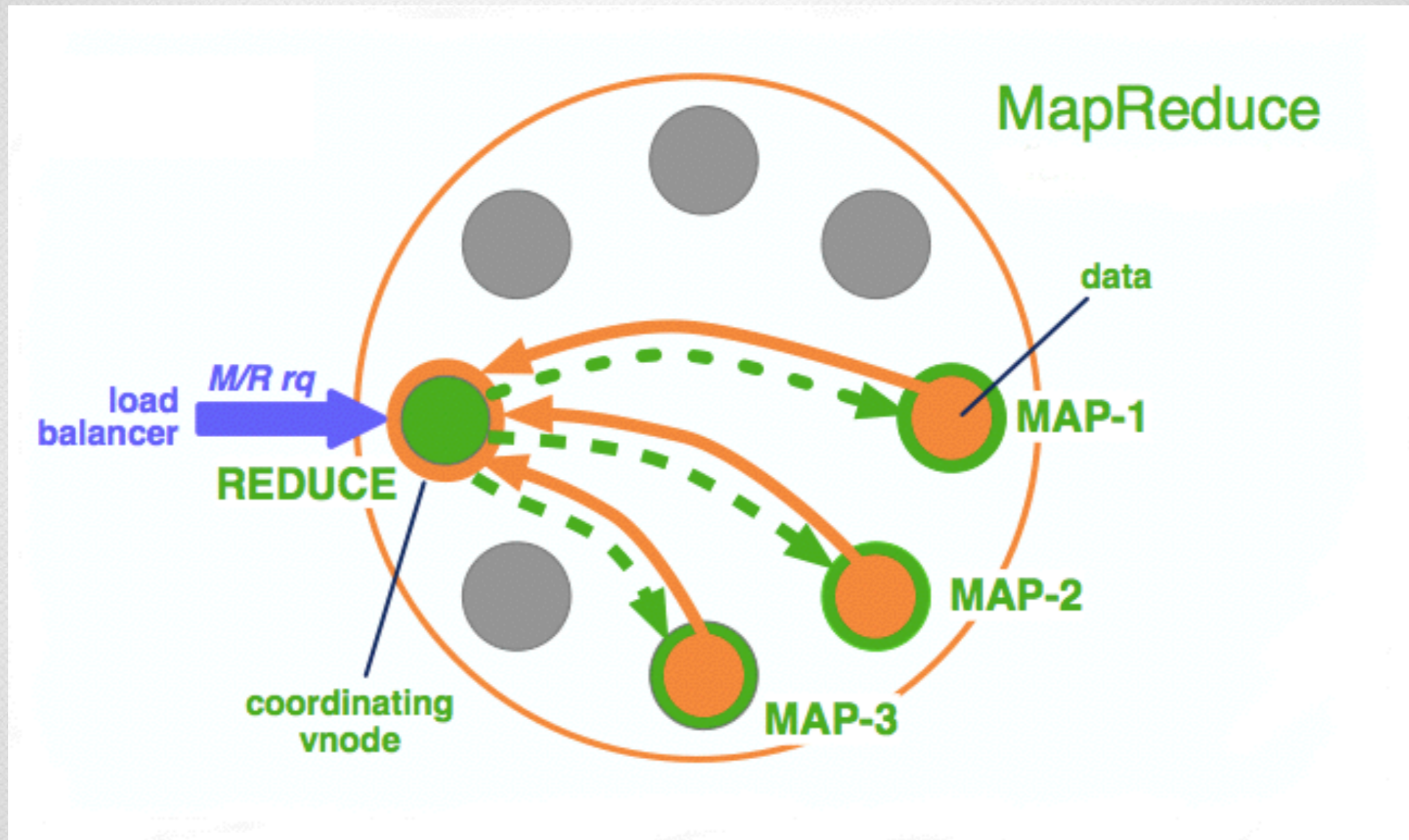
Multi Service



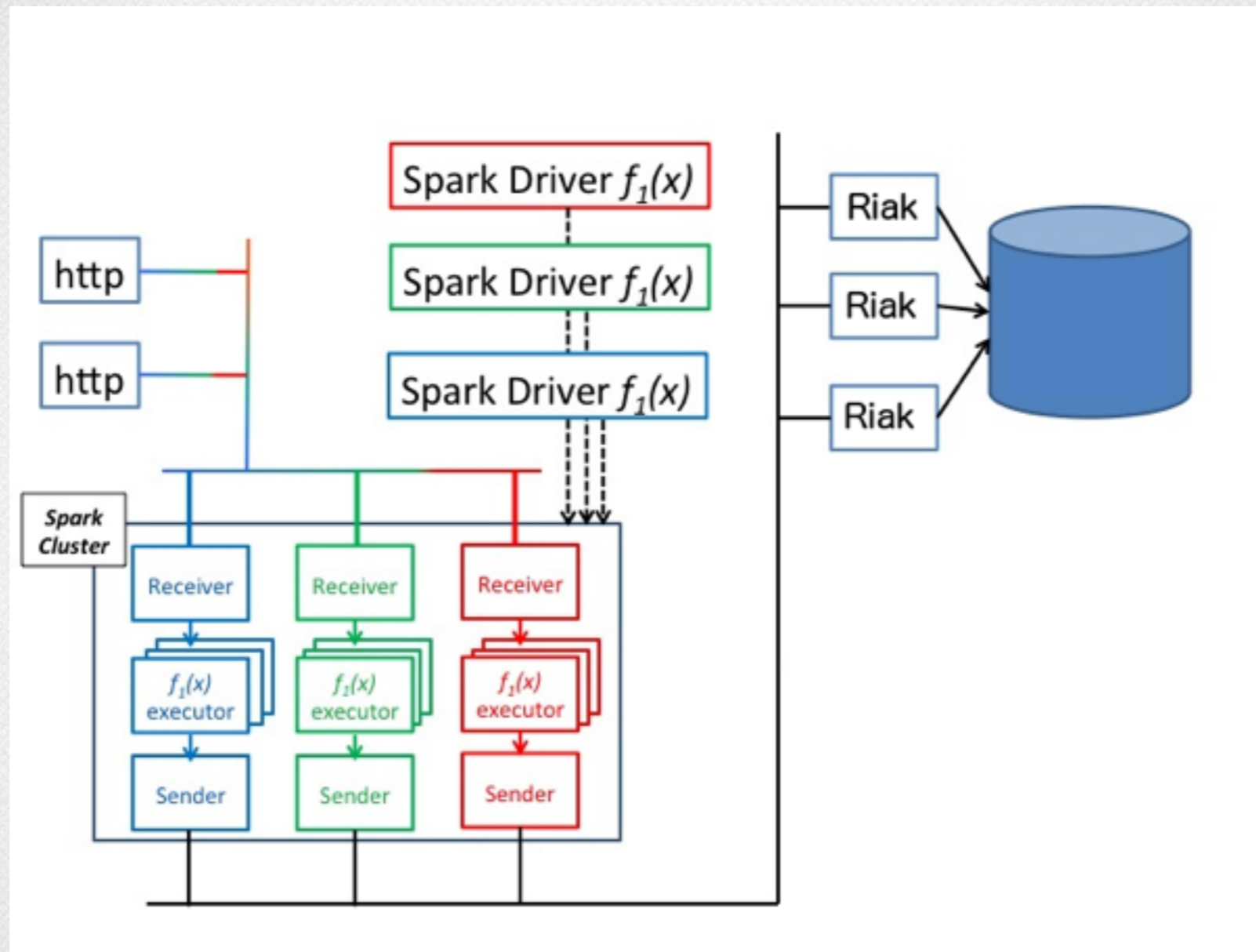
Analysis



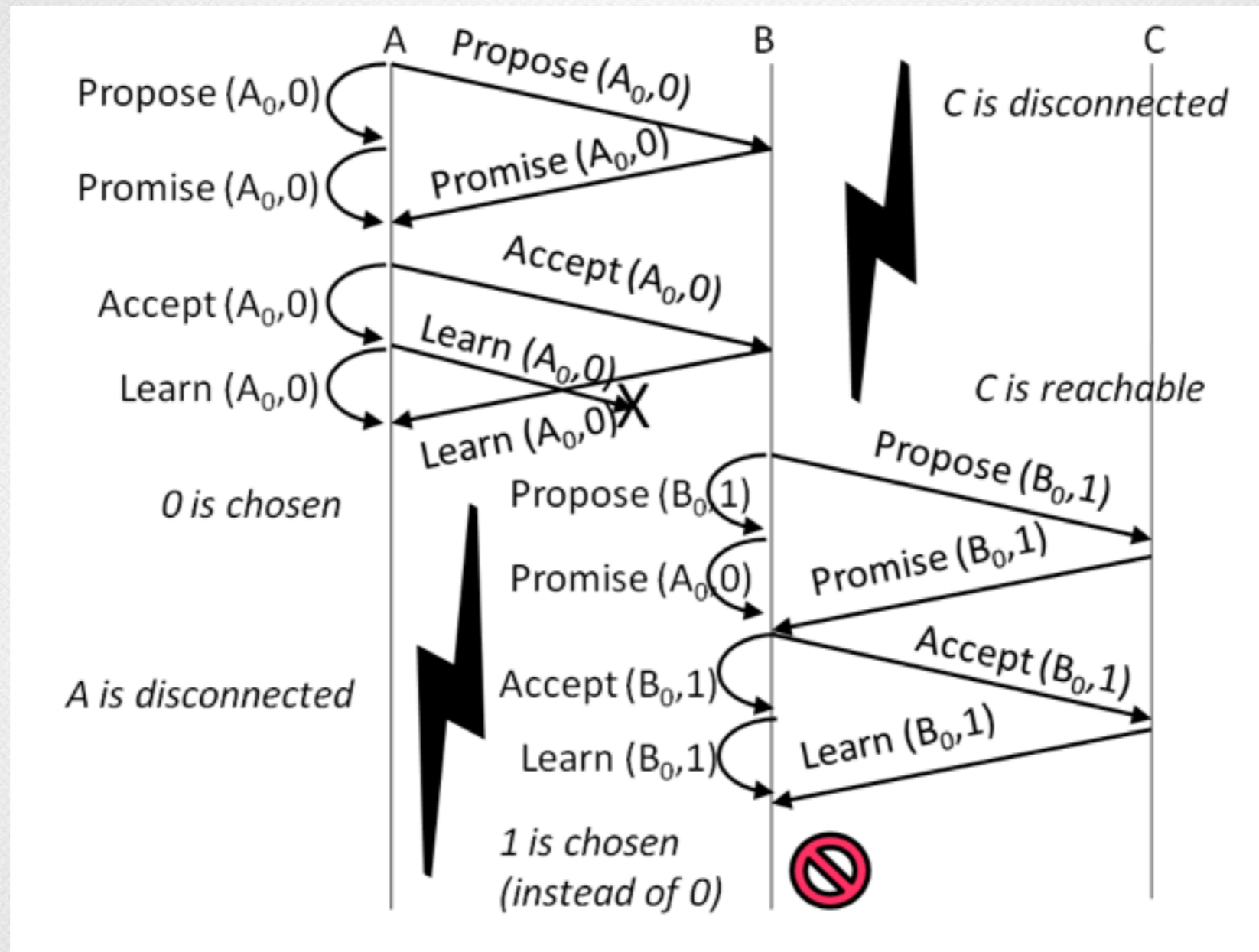
Ship The Compute?



Spread The Load



No Pets

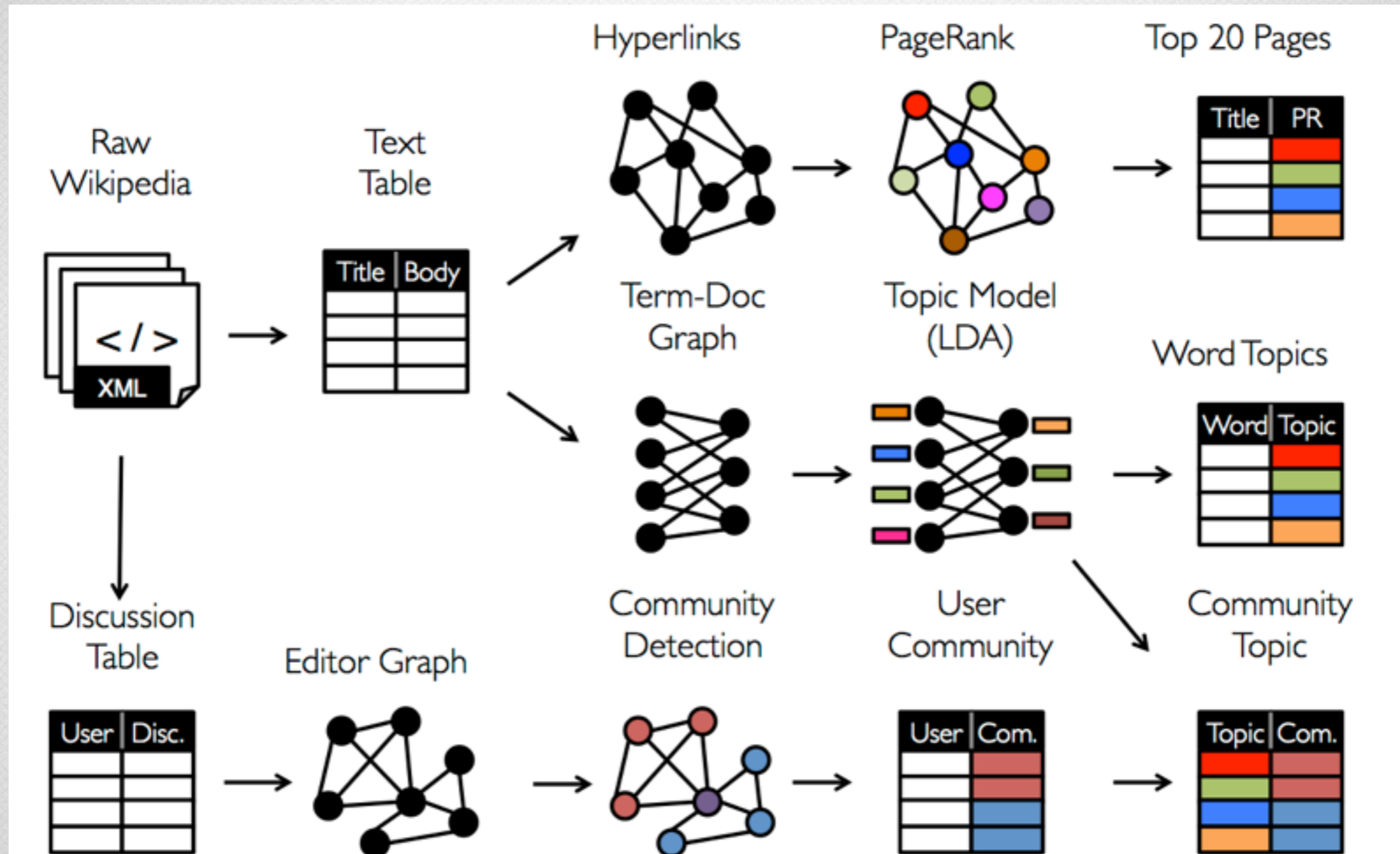




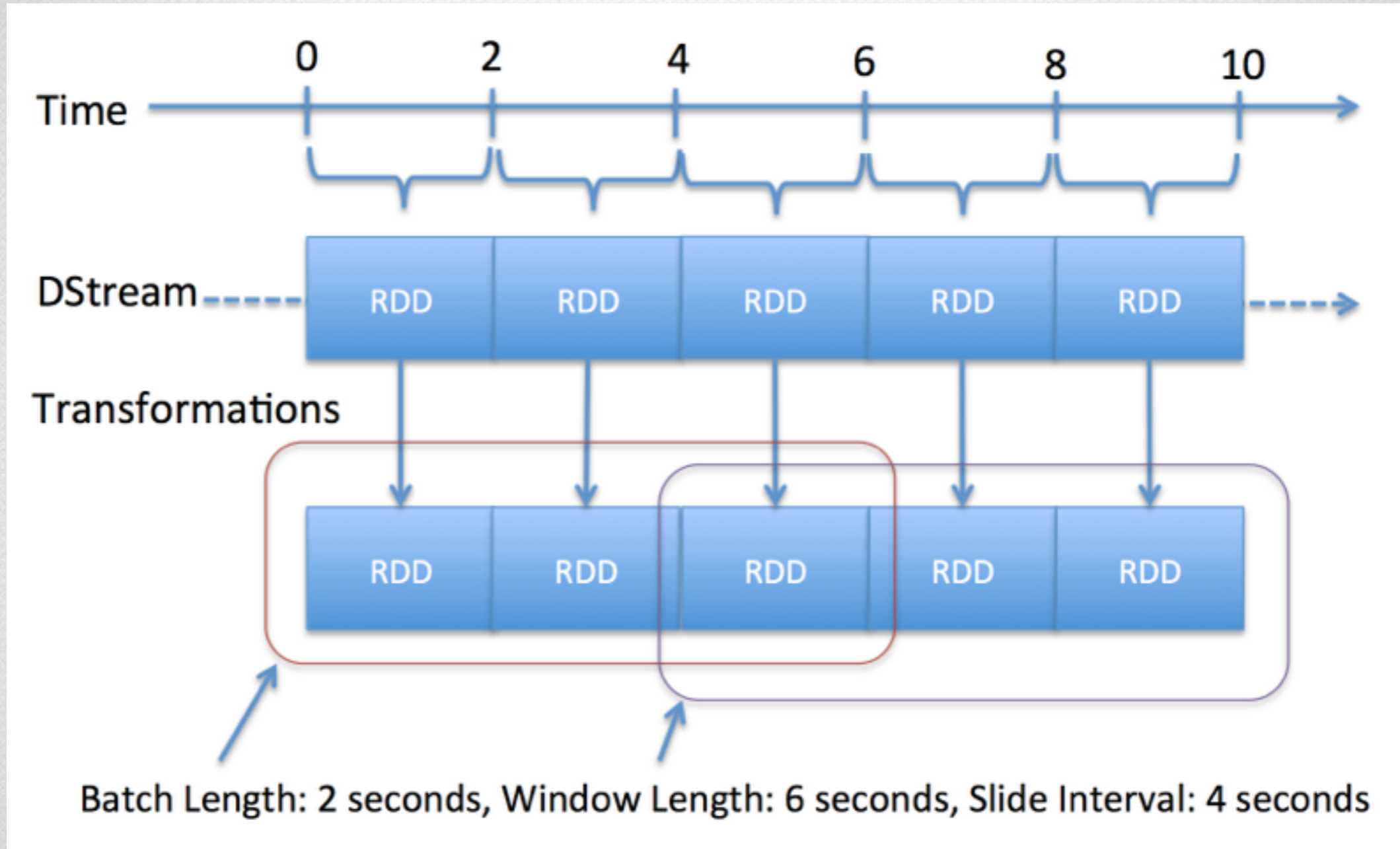
JULIA'S
BOAT
BOAT
JULIA'S



Graph Reality



Window Seat





QUEUE

ØMQ



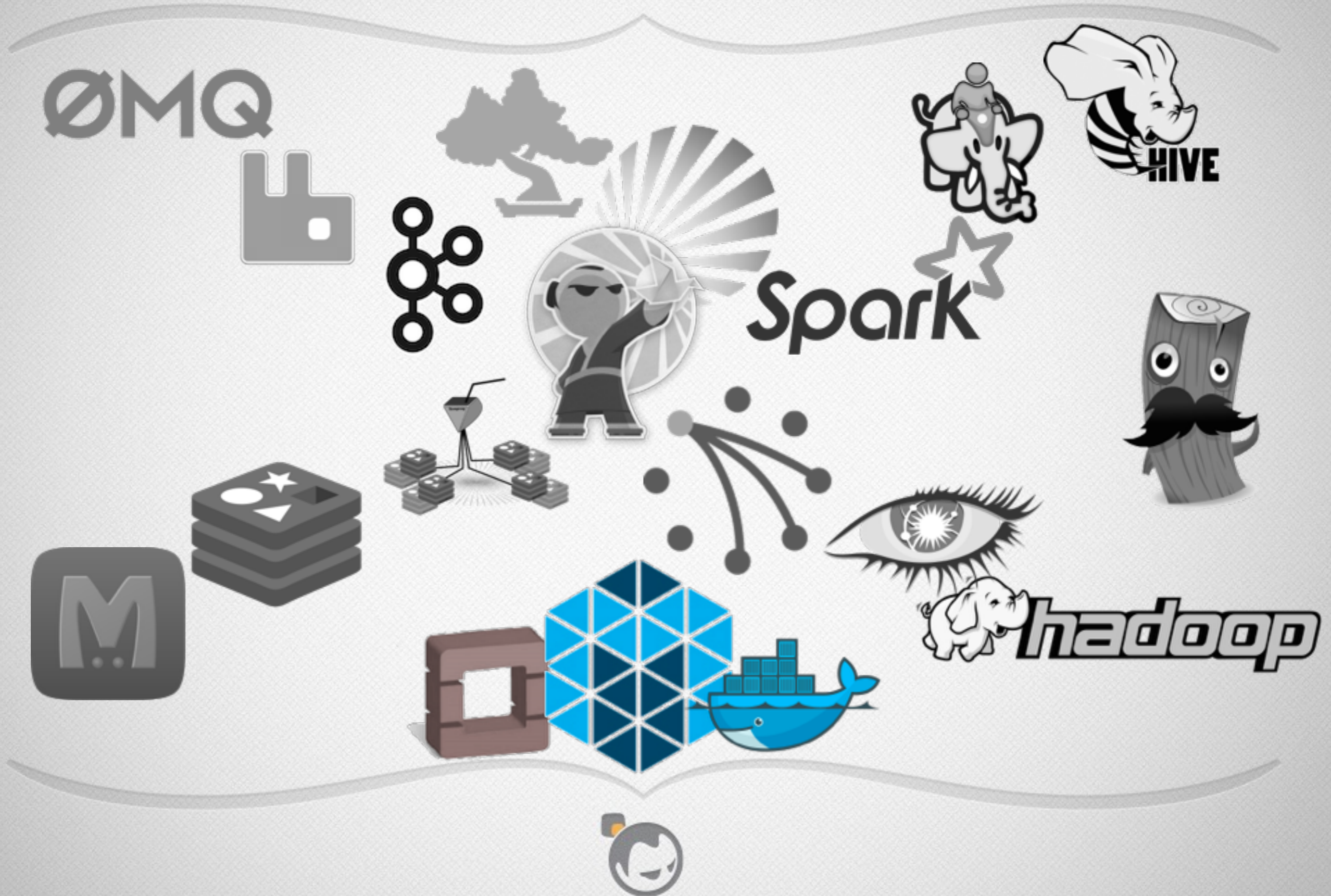
Spark



hadoop



HOST



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?

