



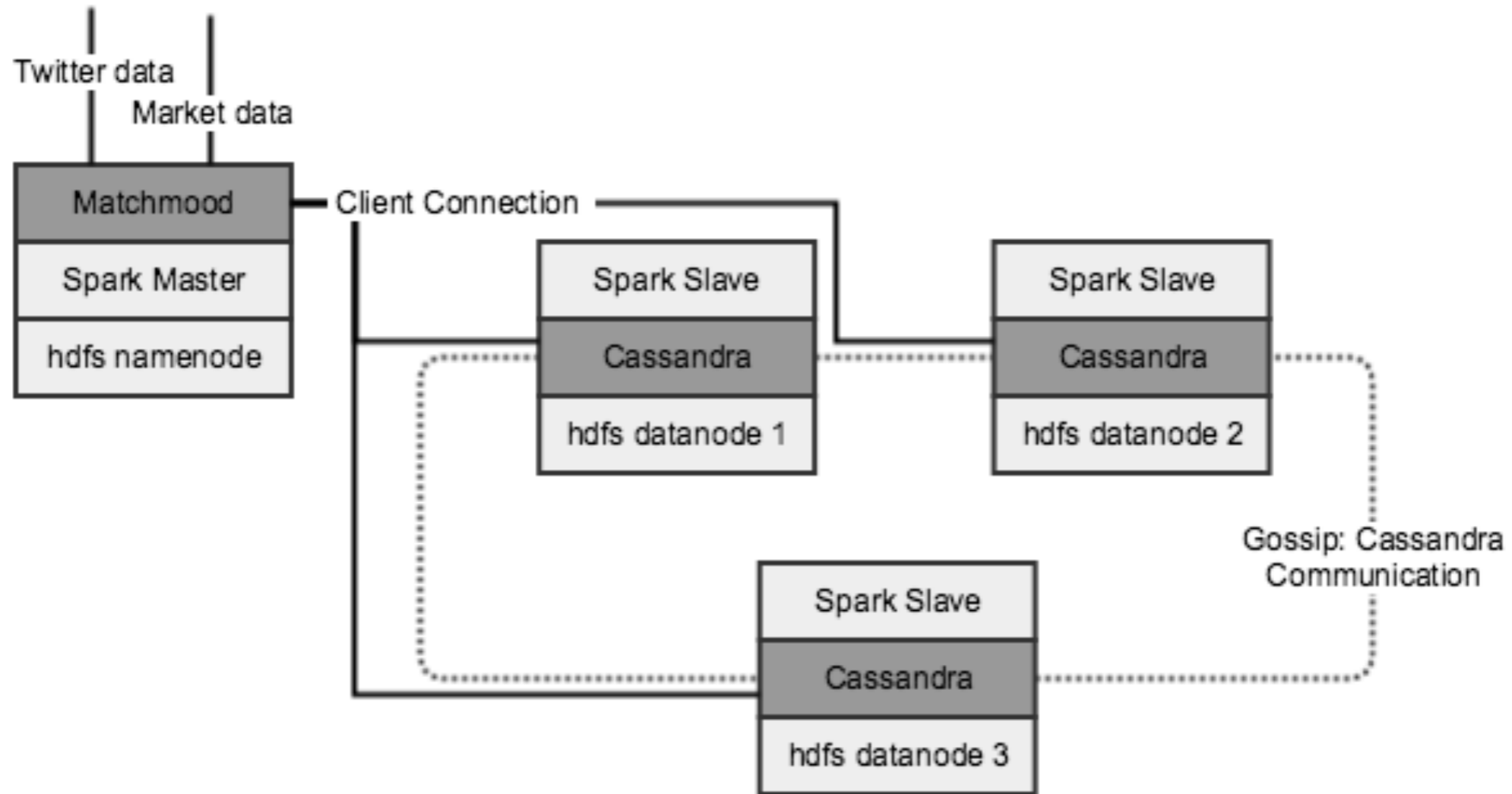
Matchmood

Sean Bugeja

sean.bugeja.12@um.edu.mt

Matthew Zammit

matthew.zammit.09@um.edu.mt



Architecture

MatchMood architecture

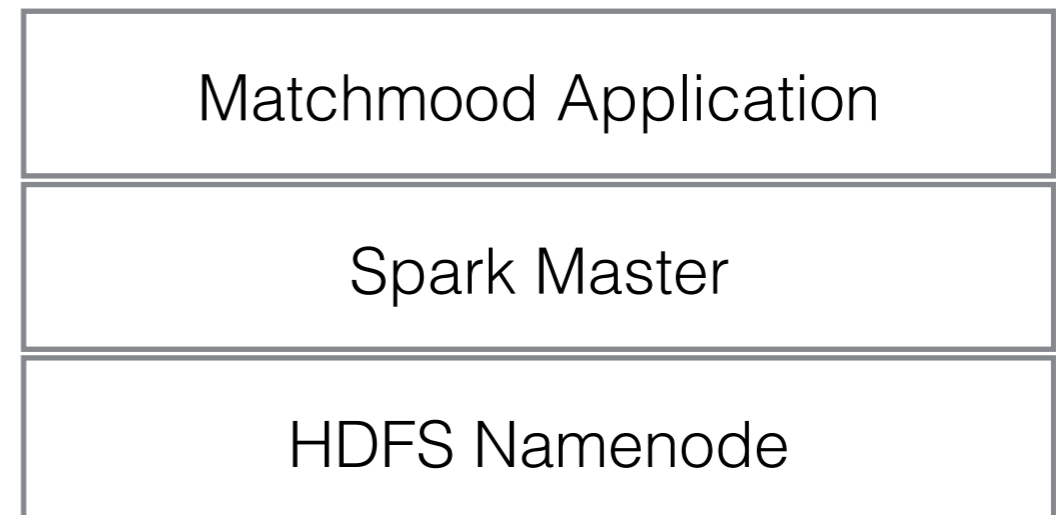
Data Node

- X3 data nodes
- Cassandra Database
- Spark slaves
- hdfs YARN data node
- Storing market/twitter/
aggregated data
- Spark jobs run close to data



Name Node

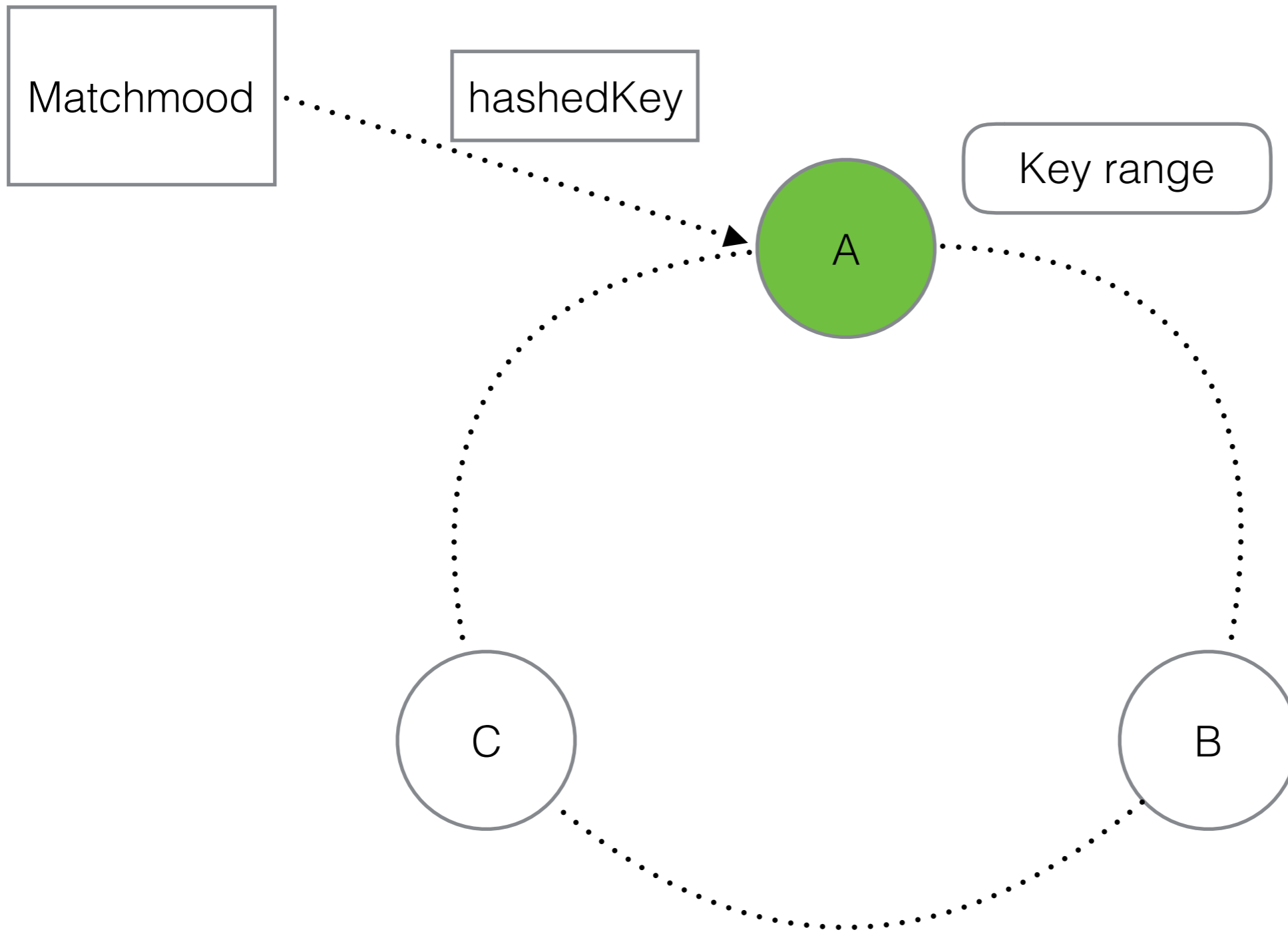
- 1 name node
- Spark Master
- hdfs YARN namenode
- Hosts Matchmood Application
- Hosts Moodbuilder Application

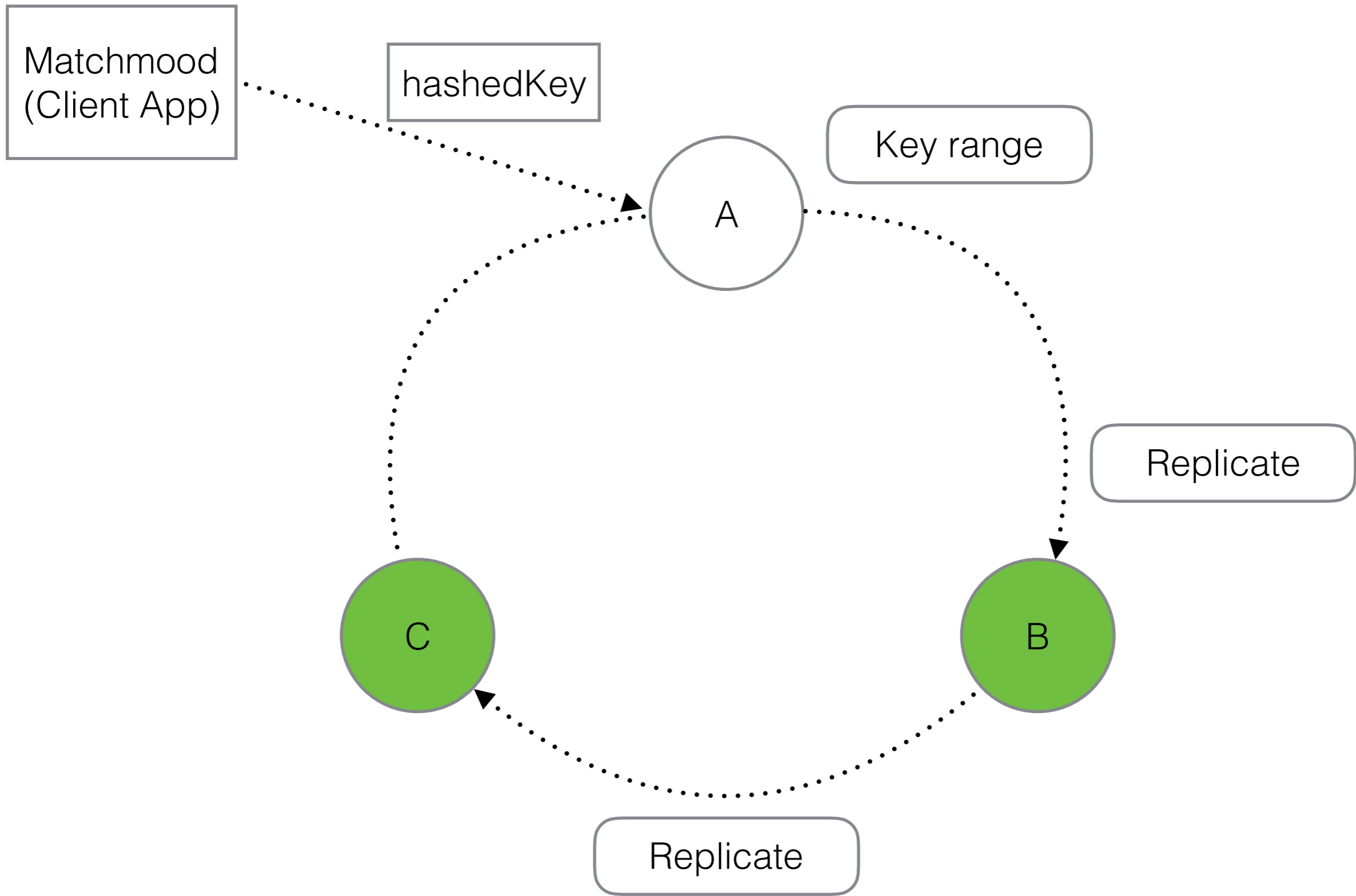


Cassandra

- Masterless cluster, Gossip communication
- Good for time series
- Fast Read/Writes
- CAP - Manageable
Consistency, Availability,
Partition



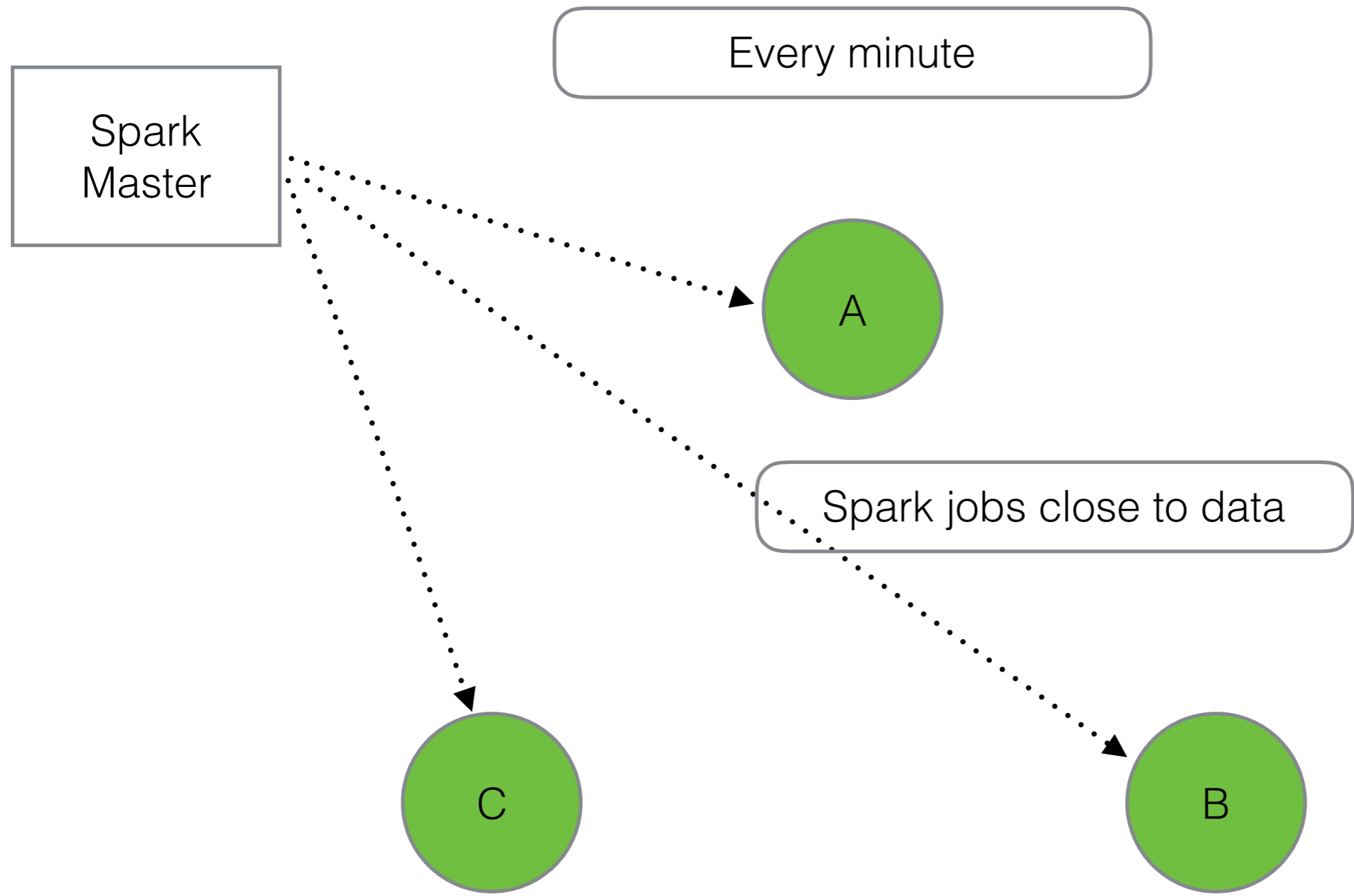




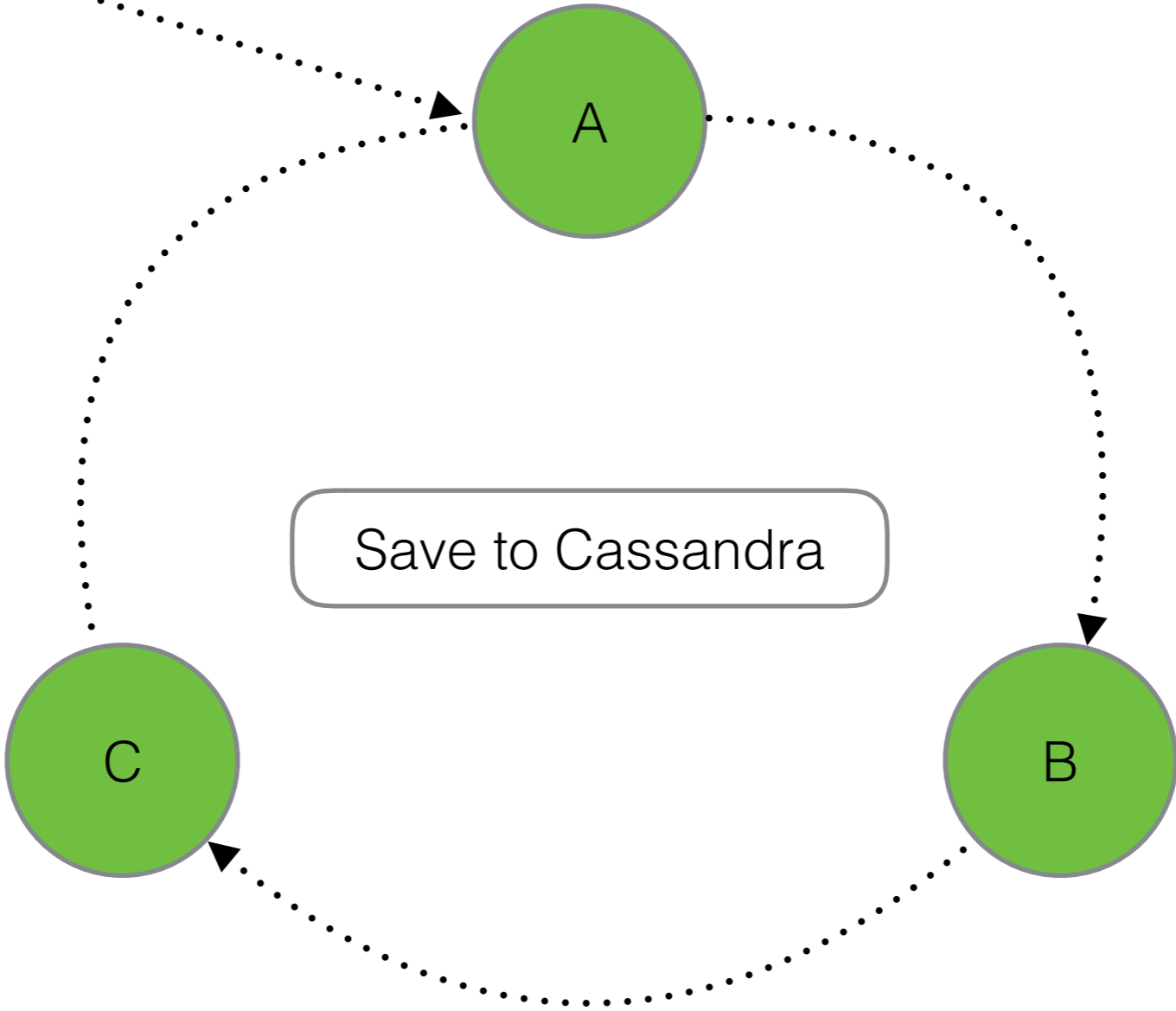
Spark

- Batch/Real time large-scale data processing
- MapReduce in memory
- Make use of RDD(Resilient Distributed Datasets)
- Uniform API(s); Spark SQL





Spark
Master

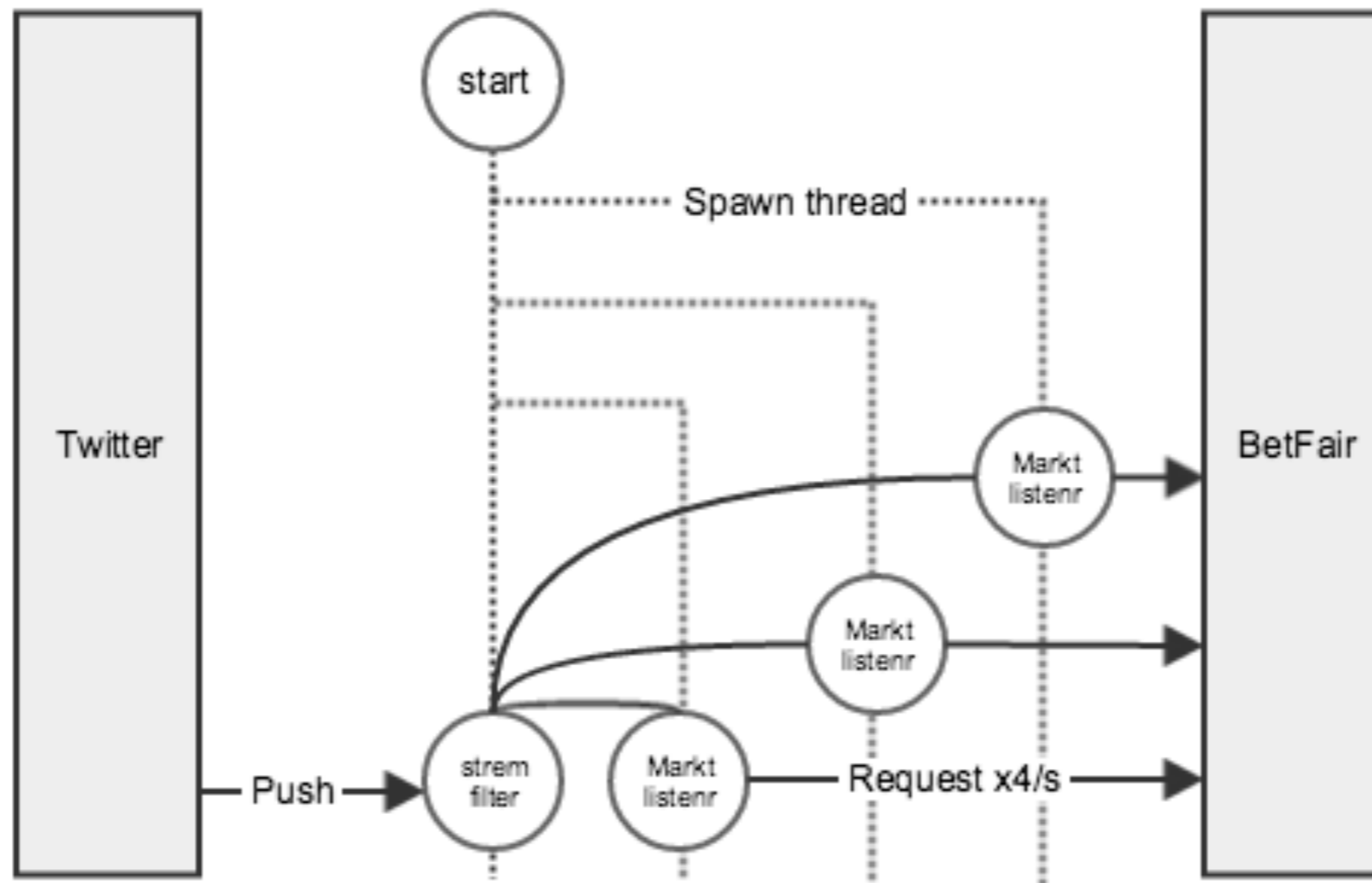


A

Save to Cassandra

C

B



Application

MatchMood

Market Data

- Betfair API
- JSON-RPC
- Request every 250 ms for a given market
- Match Odds market; Odds for the selections 1X2



Primary Key						
Partition	Ordering					
tag	match_time	time_uuid	home	draw	away	volume
	match_time	time_uuid	home	draw	away	volume
	match_time	time_uuid	home	draw	away	volume
	match_time	time_uuid	home	draw	away	volume

Market Data Example

tag	match_time	time_uuid	away	draw	home	time_stamp	volume
CRYMNU	80	93f29cb8-1f7c-11e6-9298-0e3a819fea4c	262	212	680	2016-05-21 17:50:59+0000	9.4112e+06
CRYMNU	80	93c90fba-1f7c-11e6-b064-be1c0c70fd60	262	214	680	2016-05-21 17:50:59+0000	9.4033e+06
CRYMNU	80	939f9428-1f7c-11e6-bf14-4d22df91e81b	258	214	680	2016-05-21 17:50:59+0000	9.4252e+06
CRYMNU	80	9375de94-1f7c-11e6-a565-93e35715720a	262	214	680	2016-05-21 17:50:59+0000	9.4046e+06
CRYMNU	80	934ca74a-1f7c-11e6-85a2-7b8c5059dae8	260	214	660	2016-05-21 17:50:58+0000	9.4182e+06
CRYMNU	80	93233c48-1f7c-11e6-845b-47796a8479cc	256	214	680	2016-05-21 17:50:58+0000	9.3911e+06
CRYMNU	80	92f9b24c-1f7c-11e6-83fc-9df7d955785b	264	212	680	2016-05-21 17:50:58+0000	9.4117e+06
CRYMNU	80	92d04a56-1f7c-11e6-a7dc-d4c7709ca6d8	260	214	680	2016-05-21 17:50:58+0000	9.4027e+06
CRYMNU	80	92a69468-1f7c-11e6-9381-bdb4e1a22726	262	214	680	2016-05-21 17:50:57+0000	9.4195e+06
CRYMNU	80	927cbc24-1f7c-11e6-92ea-e97f30ee84be	264	214	680	2016-05-21 17:50:57+0000	9.4045e+06
CRYMNU	80	92532238-1f7c-11e6-8ede-2c883285127c	262	212	660	2016-05-21 17:50:57+0000	9.4186e+06
CRYMNU	80	9229bd1c-1f7c-11e6-ab19-b14e6304370b	264	212	660	2016-05-21 17:50:56+0000	9.4064e+06
CRYMNU	80	9200364a-1f7c-11e6-b254-b0002fd690dc	258	214	680	2016-05-21 17:50:56+0000	9.3909e+06
CRYMNU	80	91d6785a-1f7c-11e6-836c-175df91b464b	258	214	680	2016-05-21 17:50:56+0000	9.3909e+06
CRYMNU	80	91ad06be-1f7c-11e6-bb46-3267deb49f52	258	214	680	2016-05-21 17:50:56+0000	9.3909e+06
CRYMNU	80	9183810e-1f7c-11e6-b856-9d897d02275b	258	216	680	2016-05-21 17:50:55+0000	9.3922e+06
CRYMNU	80	9159f186-1f7c-11e6-bac3-b489eec61701	262	212	660	2016-05-21 17:50:55+0000	9.4186e+06
CRYMNU	80	91306c30-1f7c-11e6-80c1-152fc0ed35c8	258	216	680	2016-05-21 17:50:55+0000	9.3922e+06
CRYMNU	80	91070ee4-1f7c-11e6-86d9-73781b98afc5	262	210	660	2016-05-21 17:50:55+0000	9.4094e+06
CRYMNU	80	90dd8c04-1f7c-11e6-8f75-1fd25eed8299	262	214	660	2016-05-21 17:50:54+0000	9.4048e+06

Twitter Data



- Twitter API
- JSON
- Push service through filtered stream
- Listen to keywords, hashtags, handles

Primary Key								
Partition	Ordering							
tag	match_time	time_uuid	entity	text	anger	positive	disgust
	match_time	time_uuid	entity	text	anger	positive	disgust
	match_time	time_uuid	entity	text	anger	positive	disgust
	match_time	time_uuid	entity	text	anger	positive	disgust

Args passed to Filter

- ['@Swansea', '@MCFC', '#SWACITY', '#SWAvCITY', '#SWA', '#swans', '#CITY', '#mcfc', '@WatfordFC', '@SunderlandAFC', '#watfordfcSUN', '#watfordfcvSUN', '#watfordfc', '#SUN', '#SAFC', '@SouthamptonFC', '@CPFC', '#SaintsFCCRY', '#SaintsFCvCRY', '#SaintsFC', '#CRY', '#CPFC', '@stokecity', '@whufc_official', '#STKWHU', '#STKvWHU', '#STK', '#SCFC', '#WHU', '#WestHam', '@ManUtd', '@afcbournemouth', '#MNUBOU', '#MNUvBOU', '#MNU', '#mufc', '#BOU', '#afcb', '@NUFC', '@SpursOfficial', '#NEWTOT', '#NEWvTOT', '#NEW', '#NUFC', '#TOT', '#COYS', '@ChelseaFC', '@LCFC', '#CHELEI', '#CHEvLEI', '#CHE', '#CFC', '#LEI', '#lcfc', '@Everton', '@NorwichCityFC', '#EFCNOR', '#EFCvNOR', '#EFC', '#COYB', '#NOR', '#OTBC', '#NCFC', '@Arsenal', '@AVFCOfficial', '#ARSAVL', '#ARSvAVL', '#ARS', '#Arsenal', '#AVL', '#avfc', '@WBAFCofficial', '@LFC', '#WBALIV', '#WBAvLIV', '#WBA', '#LIV', '#LFC']

Twitter Data Example

match_time	text
60	Manchester United Wazza has been great as From midfield to defence to attack ! Cmon United ! Bring the Cup home ! #WeAreUnited #FACupFinal
60	What we thinking I'm going 2 -0 Manchester United Martial & Rashford
60	Need goals and I believe in the boys Manchester United #mufclive
60	The least enjoyable 45 minutes of football I've ever experienced #cpfc #mufc
60	Youthful , Rashford & Martial doing it right surely #ComeOnManUtd #MJFC #FACupFinal
60	PEEP PEEEEEEEEEP! Crystal Palace Manchester United #cpfc #mufc
60	What we thinking I'm going 2 -0 Manchester United Martial & Rashford
60	Going Manchester United win in the #FACupFinal Rooney and rashford to score #weareunited #bbcfacup
60	I reckon a red card will get produced second half #mufc
60	Manchester United go go united. WE ARE THE UNITED
60	Come on Manchester United , let's bring home the trophy #MJFC
60	Come on the boys Manchester United
60	#CrystalPalace #ManUnited #Live On HD To Mobile Or PC #WeAreUnited #MUFCvCPFC #mufc #cpfc #Wembley #facupfinal 57
60	Manchester United: A quick hello to former #mufc team-mates for . #WeAreUnited
60	So far : Rooney pretty good Rashford has pace Martial needs more freedom We're getting a red card before FT #mufc
60	Manchester United: HT: Palace 0 #mufc 0. We've had plenty of possession & the better chances but it's level at the break. #WeAreUnited
60	So far : Rooney pretty good Rashford has pace Martial needs more freedom We're getting a red card before FT #mufc
60	I'm hoping and praying for Manchester United and #Marseille wins tonight and #sevilla win tomorrow. #finals #TIMCup #FACup
60	Manchester United: HT: Palace 0 #mufc 0. We've had plenty of possession & the better chances but it's level at the break. #WeAreUnited
60	What we thinking I'm going 2 -0 Manchester United Martial & Rashford

Twitter Data Example (cont)

entity	positive	negative	sentiment	anger	fear	trust	anticipation	disgust	joy	sadness	subjectivity	surprise
ManUtd	0.142857	0	1	0	0	0.285714	0.285714	0.142857	0	0	0.75	0.142857
ManUtd	0.333333	0	0	0	0	0.333333	0.333333	0	0	0	0	0
ManUtd	0	0	0	0	0	0.5	0.5	0	0	0	0	0
NoMention	0	0.2	0.333333	0.2	0	0.2	0.4	0	0	0	0.633333	0
ManUtd	1	0	0.261905	0	0	0	0	0	0	0	0.474868	0
MirrorFootball	0	0	0	0	0	0.333333	0.666667	0	0	0	0	0
ManUtd	0.333333	0	0	0	0	0.333333	0.333333	0	0	0	0	0
ManUtd	0	0.166667	0.8	0.166667	0	0.166667	0.333333	0	0	0.166667	0.4	0
NoMention	0	0	-0.055556	0	0	0	0	0	0	0	0.055556	0
ManUtd	0	0	0	0	0	0.5	0.5	0	0	0	0	0
ManUtd	0	0.142857	0	0.142857	0	0.285714	0.285714	0	0	0.142857	0	0
ManUtd	0	0	0	0	0	0.5	0.5	0	0	0	0	0
NoMention	0	0	0.136364	1	0	0	0	0	0	0	0.5	0
WayneRooney	0	0	0.166667	0	0	0	0	0	0	0	0.25	0
allforunited	0.076923	0.230769	0.258333	0.153846	0	0.230769	0.230769	0	0	0.076923	0.516667	0
ManUtd	0.166667	0	0.5	0	0	0.166667	0.166667	0.166667	0.166667	0	0.5	0.166667
allforunited	0.076923	0.230769	0.258333	0.153846	0	0.230769	0.230769	0	0	0.076923	0.516667	0
BorussiaDortmun	0	0	0.55	0	0	0.5	0.5	0	0	0	0.3	0
ManUtd	0.166667	0	0.5	0	0	0.166667	0.166667	0.166667	0.166667	0	0.5	0.166667
ManUtd	0.333333	0	0	0	0	0.333333	0.333333	0	0	0	0	0

MoodBuilder

- Application written in Scala
- 1 Minute intervals - aggregate data
- 2 main jobs
- Run as near real time or rebuilding jobs(for older data)



Aggregate Twitter Data

- Select data from last_minute from Cassandra twitter_data table
- Map for each tuple
- (tag,entity) -> (fear,positive,negative,.....,joy)
- ReduceByKey (tag,entity) -> sum each tuple value ((fear,pos,neg,....,joy))
- Save result to Cassandra table

Primary Key								
Partition	Ordering							
tag	match_time	count	entity	fear	anger	positive	disgust
	match_time	count	entity	fear	anger	positive	disgust
	match_time	count	entity	fear	anger	positive	disgust
	match_time	count	entity	fear	anger	positive	disgust

Aggregate Twitter Data Example

tag	match_time	count	entity	anger	anticipation	disgust	fear	joy	negative	positive	sadness	sentiment	subjectivity	surprise
CRYMNU	69	135	ManUtd	1.44405	31.225	30.32619	0.859524	1.44286	2.9869	30.97619	1.22857	7.5265	26.80057	1.28571
CRYMNU	69	25	CPFC	0.806818	18.68106	0.790909	0.090909	0.2	0.806818	0.2	0.715909	3.13125	10.14167	0.2
CRYMNU	69	19	NoMention	0	2.97619	1.84286	0.342857	1.34286	0	1.34286	0.142857	0.443283	4.41111	0.842857
CRYMNU	69	8	UltimateFanLive	0	4	1	0	0	0	2	0	-0.25	0.333333	0
CRYMNU	69	5	EmiratesFACup	0.733333	1.2	0.5	0.5	0.166667	0	0	0	-0.887879	1.94167	0.2
CRYMNU	69	4	allforunited	0.507692	0.661538	0	0	0	0.661538	0.153846	0.353846	1.225	1.8	0
CRYMNU	69	3	BBCSport	0	0.333333	0.25	0	0.25	0.333333	0.25	0	0	0.2	0.25
CRYMNU	69	3	FullTimeDEVILS	0	0	0	0	0	0	0	0	0.359375	1	0
CRYMNU	69	2	Cantonaesque_	0.75	0.5	0	0	0	0.25	0	0	0.75	1.01667	0.25
CRYMNU	69	2	UnitedGazette	0	1	0	0	0	0	0	0	0	0	0
CRYMNU	69	2	WayneRooney	0	0	0	0	0	0	0	0	0.166667	0.25	0
CRYMNU	69	2	goldnewsuk	0	1	0	0	0	1	0	0	0.538095	0.869841	0
CRYMNU	69	1	BBCMTD	0	0	0.5	0	0	0	0.5	0	0.0625	1	0
CRYMNU	69	1	BenzahraHussem	0.5	0	0	0	0	0	0	0	0.068182	0.75	0.5
CRYMNU	69	1	DoverAthletic	0.2	0.4	0	0	0	0.2	0	0	0	0	0
CRYMNU	69	1	DuncanCastles	0.142857	0.142857	0.142857	0	0.142857	0.142857	0	0.142857	0.1	0.8	0
CRYMNU	69	1	ManUtd1_	0	0.666667	0	0	0	0	0	0	0	0	0
CRYMNU	69	1	PayetForward	0	0	0	1	0	0	0	0	0	0.033333	0
CRYMNU	69	1	UtdHQ	0.25	0.25	0	0	0	0.25	0	0	0.577778	0.488889	0
CRYMNU	69	1	footballtips	0	0	0.5	0	0	0	0	0	-0.3	0.2	0.5

Aggregate Market Data

- Select data from last minute from Cassandra table market_data
- Map tuple for each operation (Home, Away, Draw, Volume)
- ReduceByKey for each tuple (max, min, open, close, average, volume difference)
- Save result to Cassandra table

Primary Key								
Partition	Ordering							
tag	type	match_time	open	close	min	max	sma	volume
		match_time	open	close	min	max	sma	volume
		match_time	open	close	min	max	sma	volume
		match_time	open	close	min	max	sma	volume

Aggregate Market Data Example

tag	selection_type	matchtime	close	max	min	open	sma	volume
CRYMNU	HOME	69	550	570	550	560	560.46082	32267
CRYMNU	HOME	68	560	570	540	560	552.33643	22709
CRYMNU	HOME	67	550	560	540	550	549.38678	47603
CRYMNU	HOME	66	560	570	550	560	556.91241	48575
CRYMNU	HOME	65	560	570	550	550	558.99084	20595
CRYMNU	HOME	64	560	560	550	560	557.87036	35081
CRYMNU	HOME	63	560	560	550	560	557.77777	13503
CRYMNU	HOME	62	560	560	550	560	554.93085	16594
CRYMNU	HOME	61	550	560	540	560	550.32861	30869

RESTful API

- GET <http://api.angrylogic.com/matches/>
- GET <http://api.angrylogic.com/matches/<TAG>/market/>
- GET <http://api.angrylogic.com/matches/<TAG>/entities/>
- GET <http://api.angrylogic.com/matches/<TAG>/entities/<ENTITY>>
- GET <http://api.angrylogic.com/matches/<TAG>/entities/<ENTITY>/minutes/>

Demonstration

- RESTful API
- <https://youtu.be/AGyrniRent8>
- GUI Application

