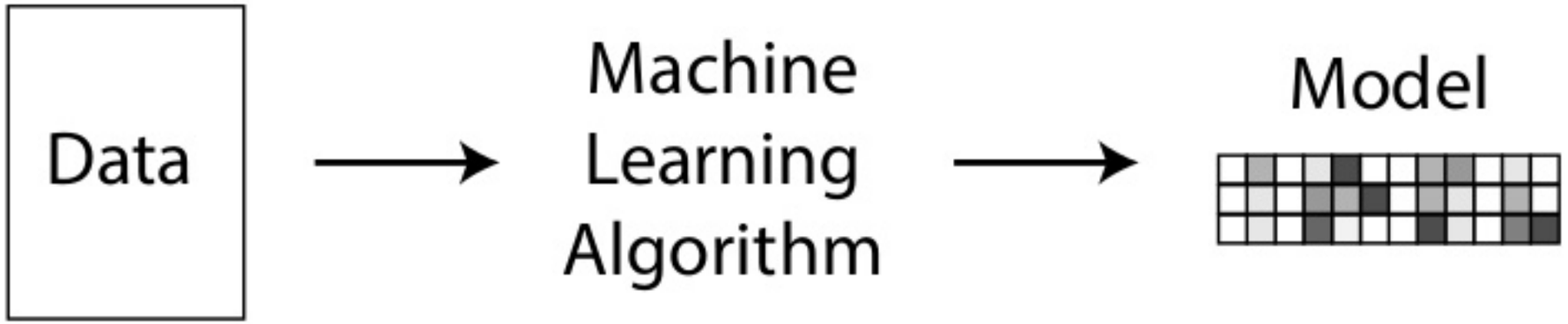
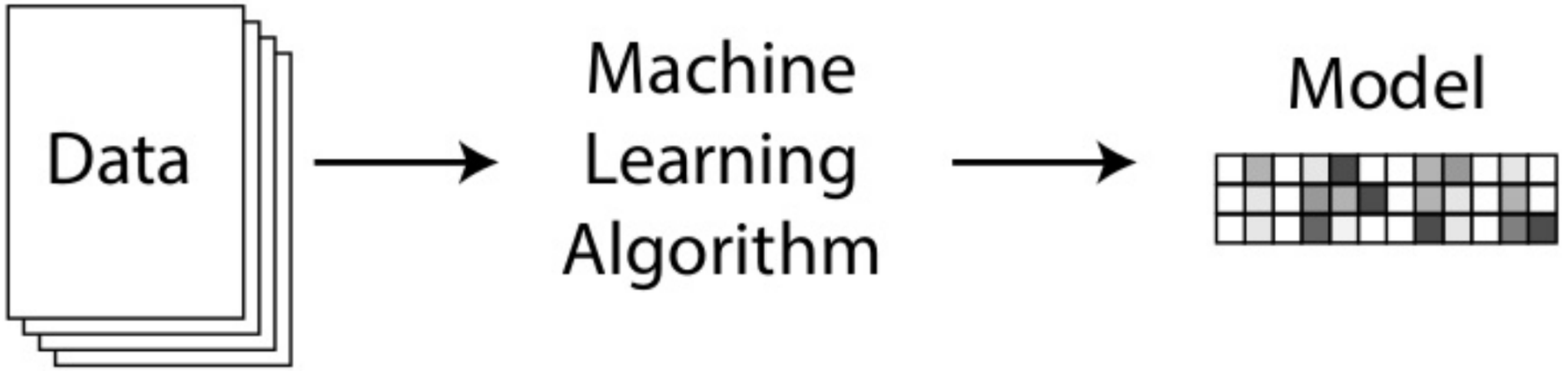
An Asynchronous Parameter Server for Spark

Rolf Jagerman
University of Amsterdam

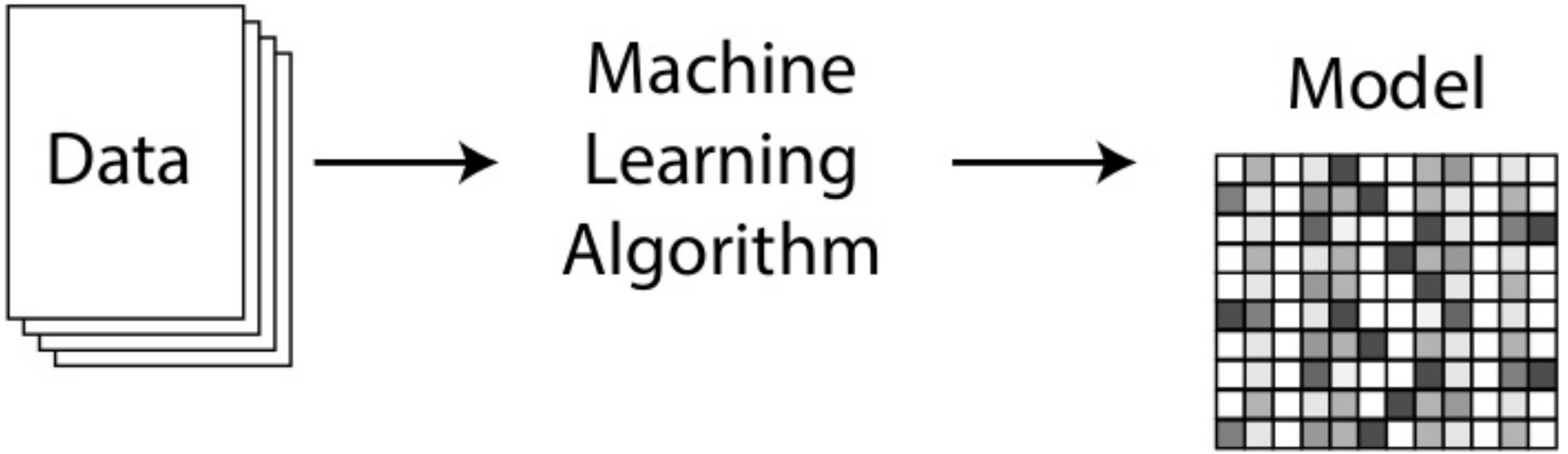
Motivation



Motivation

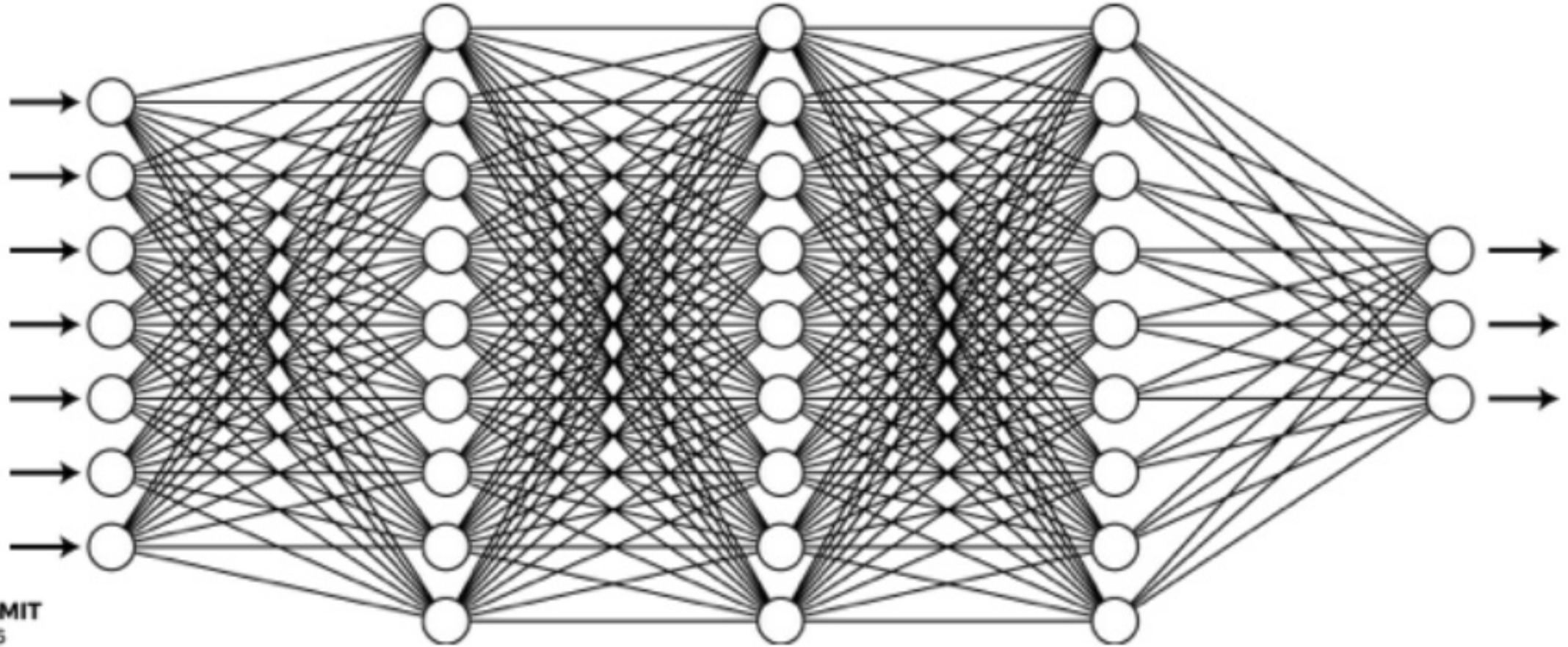


Motivation



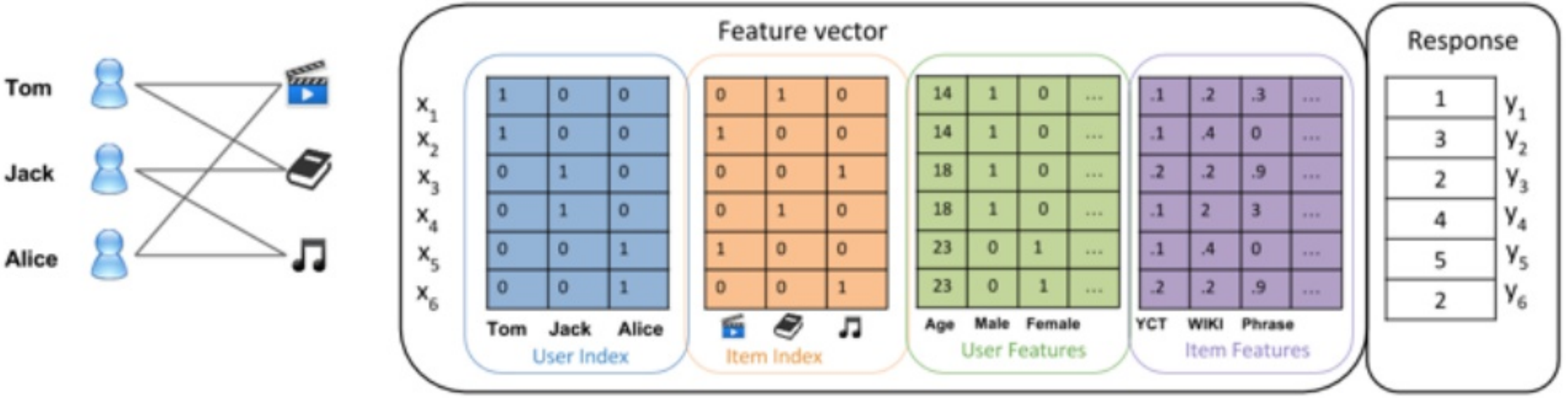
Motivation

Deep Learning



Motivation

Factorization Machines



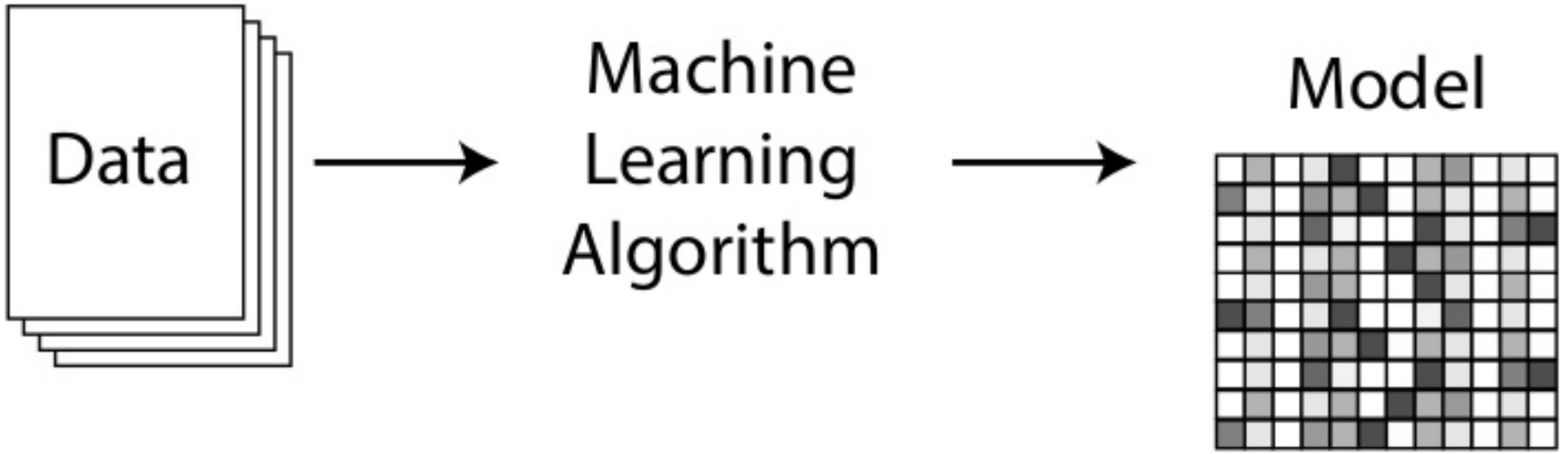
https://github.com/MLnick/glint-fm

Motivation

Topic Modeling (LDA)

Tourism	Video games	Javascript	Biology
hotel	allianc	write	tag
local	hord	docum	protein
car	euro	var	hypothet
holidai	warcraft	return	gene
area	wow	prop	cytoplasm
golf	gold	function	prk
wed	warhamm	subset	locu

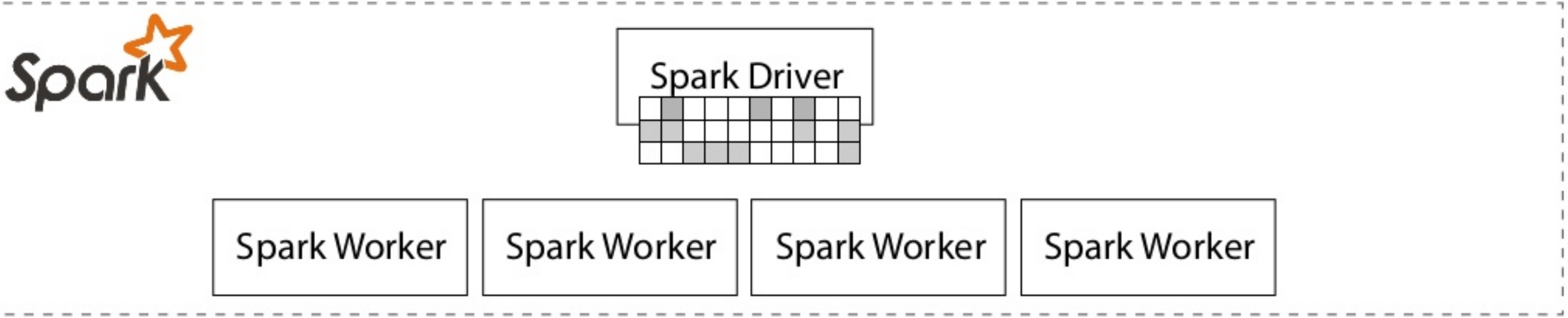
Problem

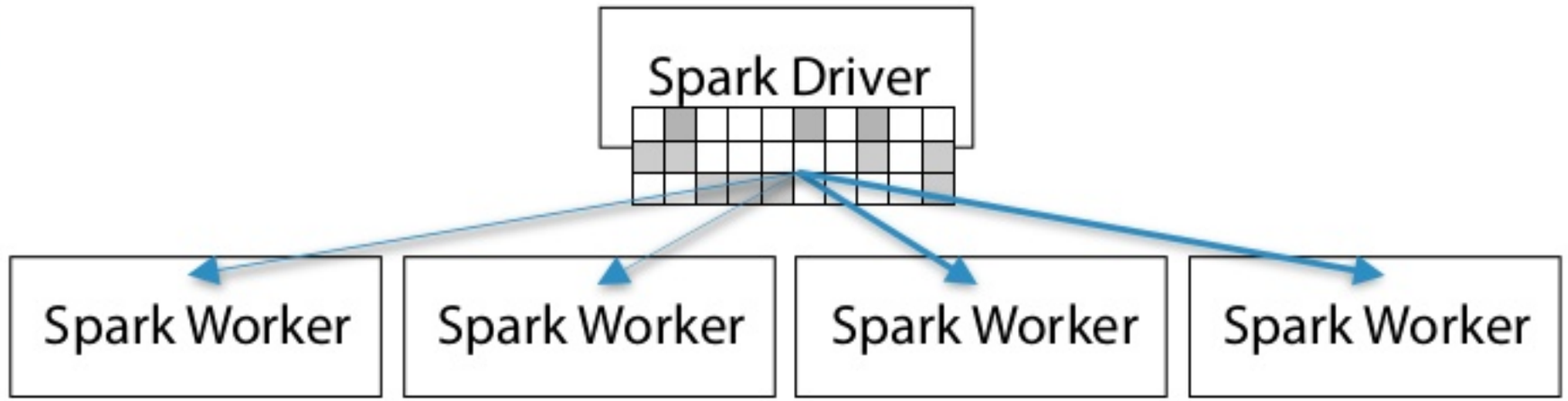


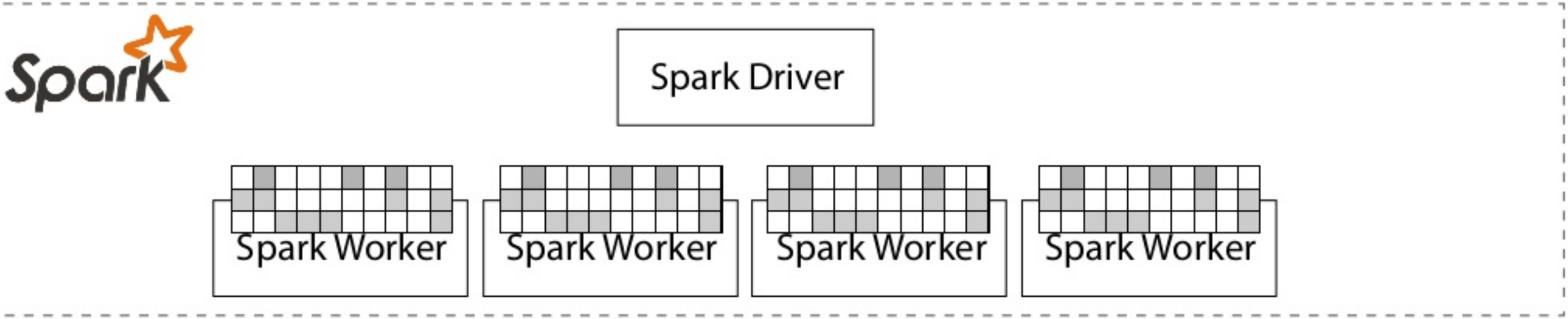
Model size exceeds memory of a single machine!

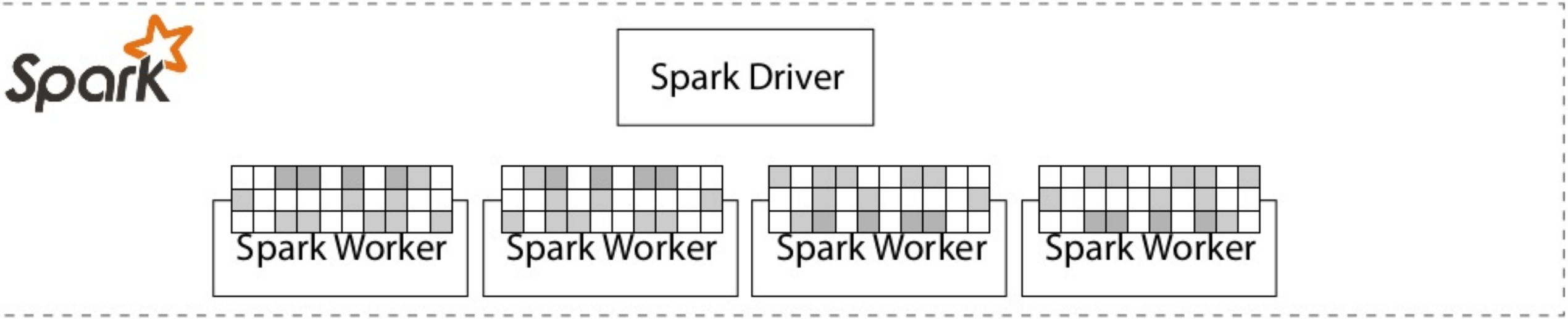
Spark Driver

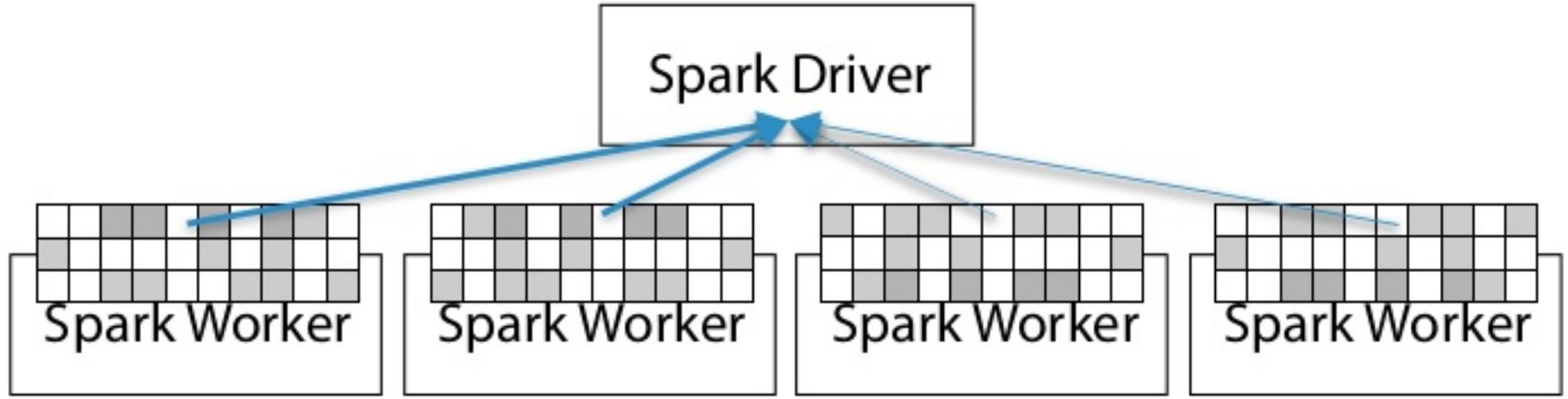
Spark Worker | Spark Worker | Spark Worker | Spark Worker

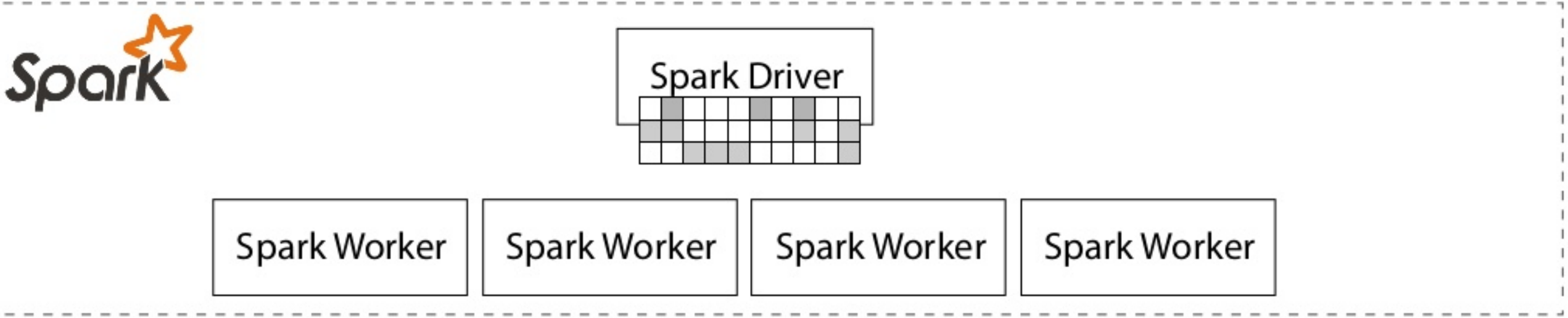

Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker


Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker


Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker


Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker


Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker

Parameter Server

- A machine learning framework
- Distributes a model over multiple machines
- Offers two operations:
  - **Pull**: query parts of the model
  - **Push**: update parts of the model

Parameter Server

- Machine learning update equation:

$$w_i \leftarrow w_i + \Delta$$

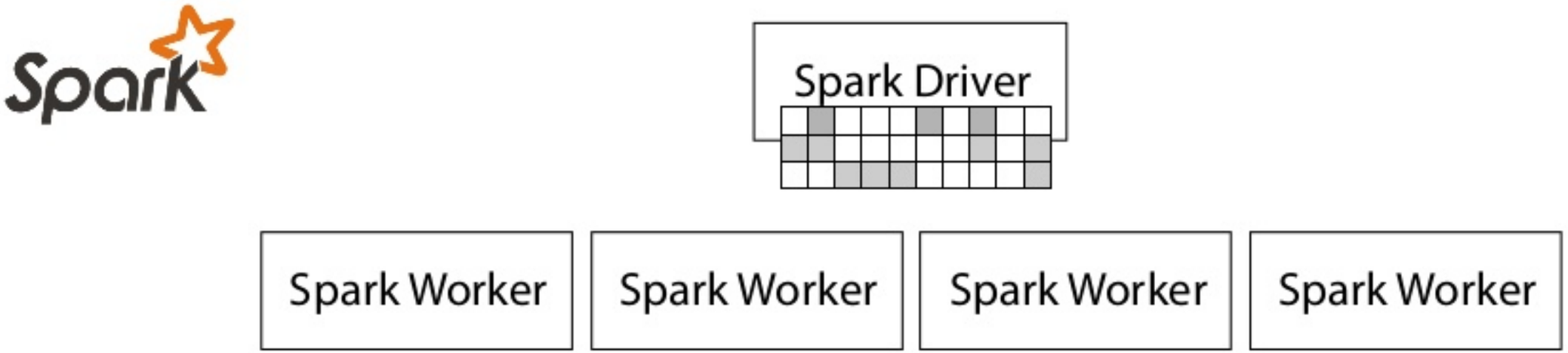
- (Stochastic) gradient descent
- Collapsed Gibbs sampling for topic modeling

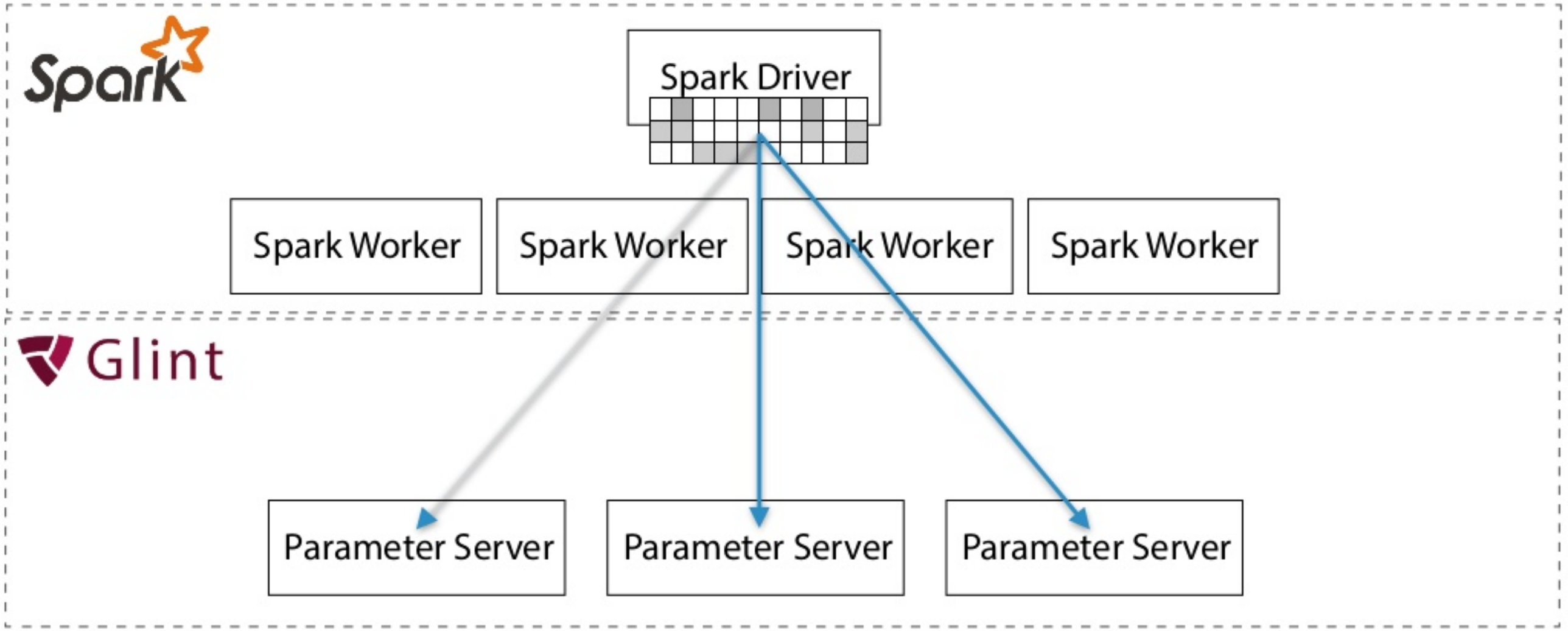
Parameter Server

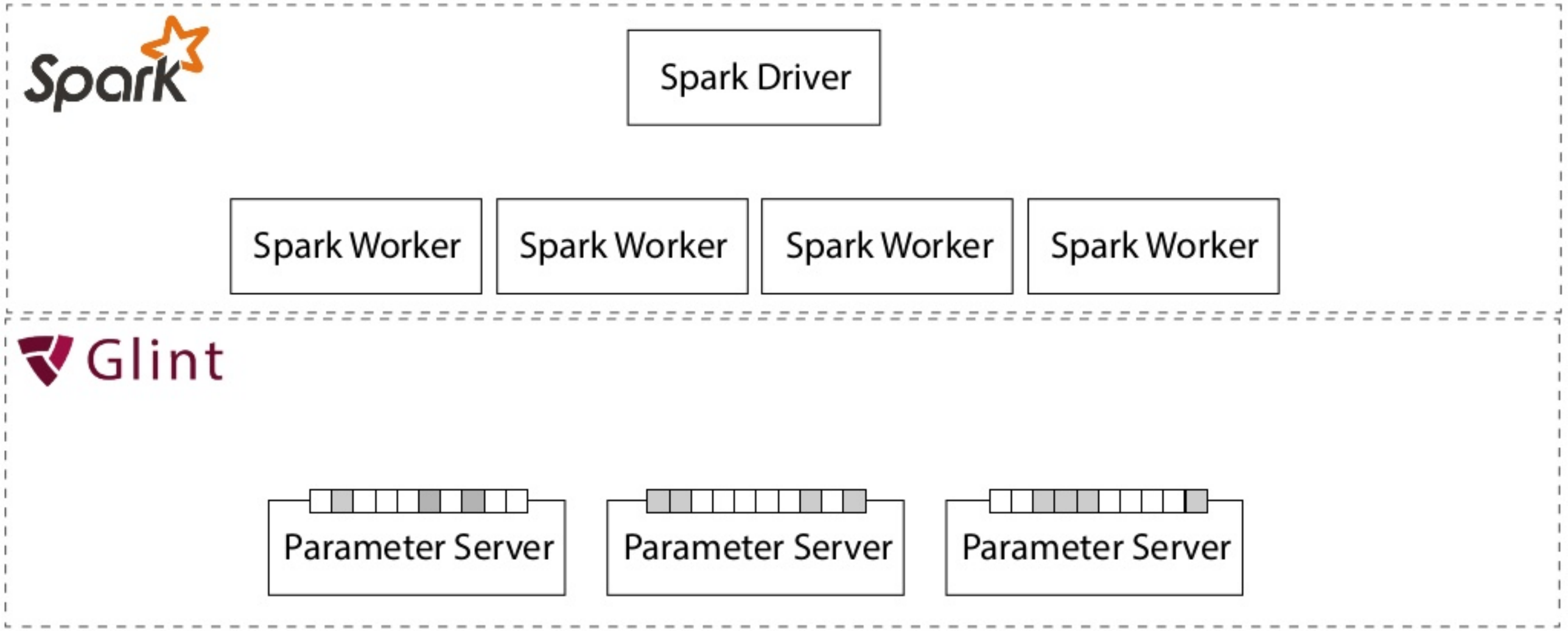
- Machine learning update equation:

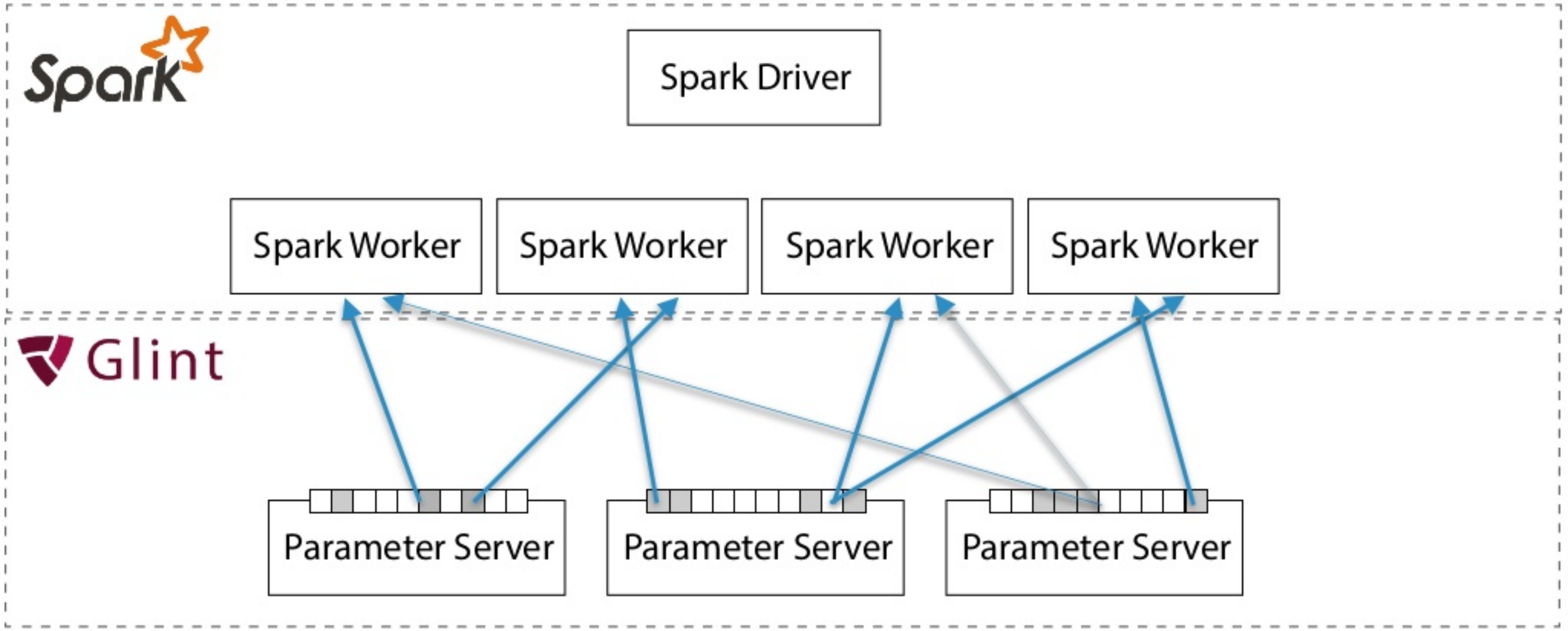
$$w_i \leftarrow w_i + \Delta$$

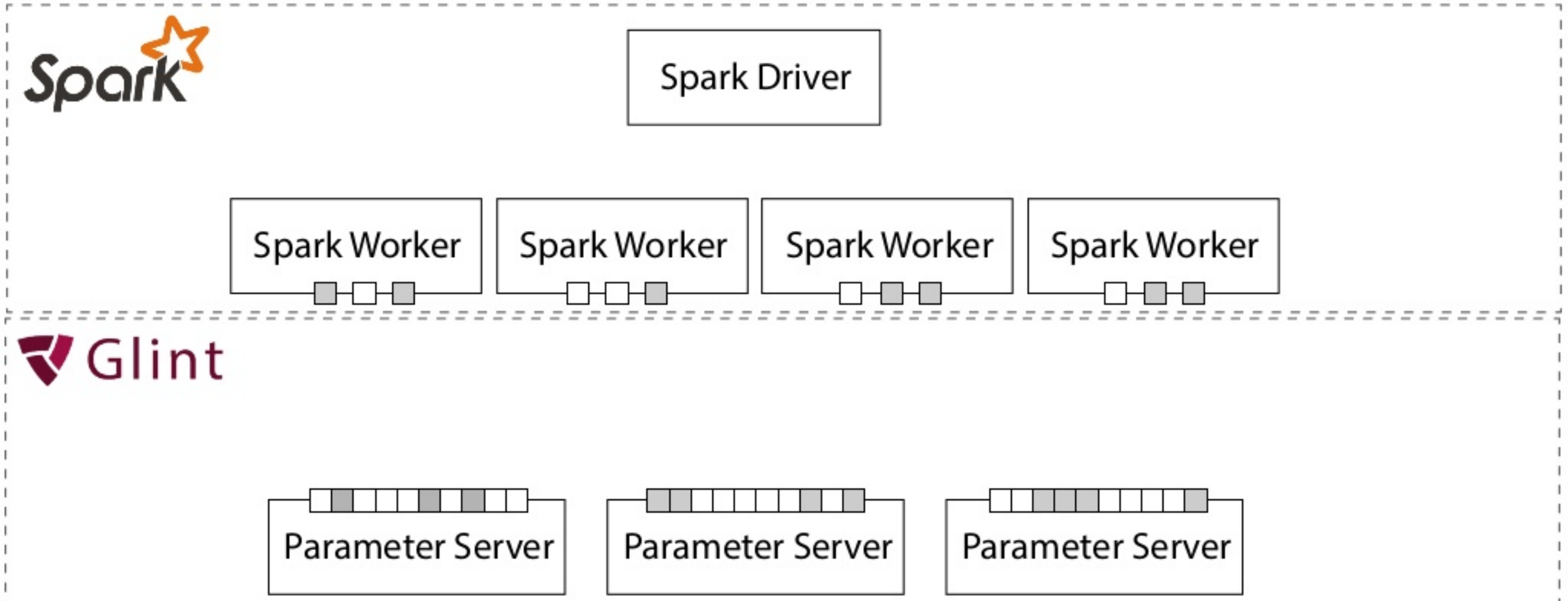
- (Stochastic) gradient descent
- Collapsed Gibbs sampling for topic modeling
- Aggregate **push** updates via addition (+)

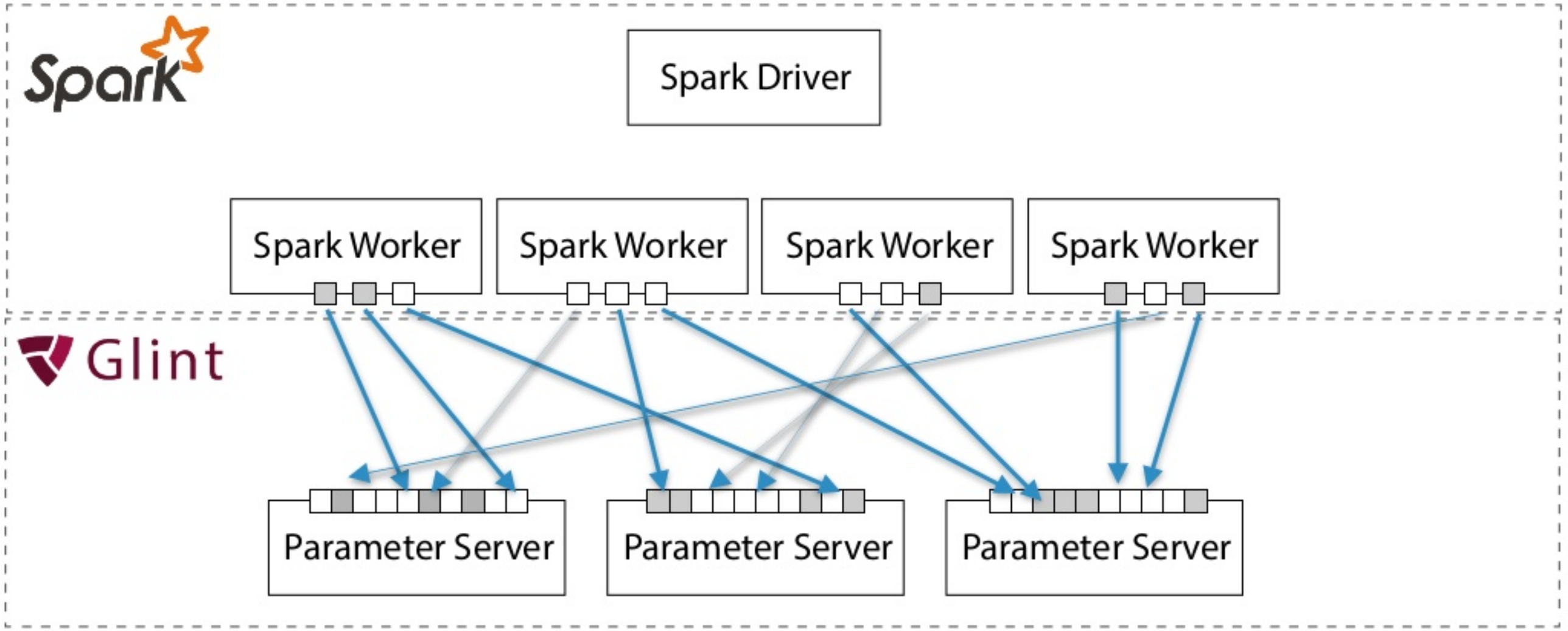

Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server

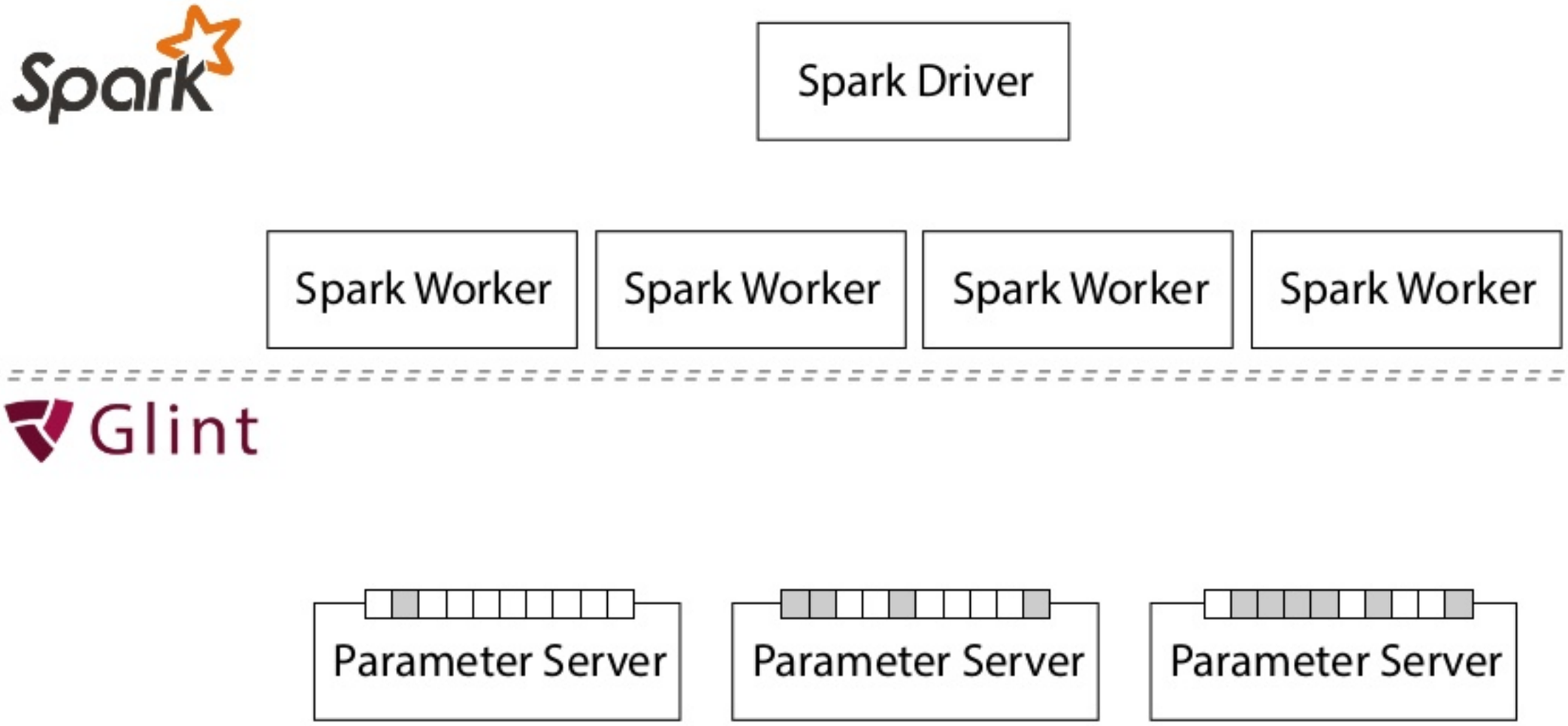

Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server


Spark
Spark Driver
Spark Worker
Spark Worker
Spark Worker
Spark Worker
Glint
Parameter Server
Parameter Server
Parameter Server

Demo

Experiments

Task

- LDA topic modeling
- 1,000 topics
- 27TB data set (ClueWeb12)

Experiments

Setup

- 30 Spark workers (16 CPU cores each)
- 3.7TB RAM total
- Interconnected over 10Gb/s ethernet

Experiments

Approach

- Glint for model storage (billions of parameters)
- LDA approximation algorithm for runtime improvements
- A small loss in model quality is acceptable

Experiments

Comparing three methods:

1. Glint LDA
2. MLLib EM LDA
3. MLLib Online LDA

Experiments

Perplexity (topic model quality)

Data (GB)	# Topics	Glint	MLLib EM	MLLib Online
50	20	6,108	**-1.3%**	**+0.4%**
100	20	5,731	**-5.4%**	**-4.6%**
150	20	5,427	**-10.7%**	**-8.2%**
200	20	6,021	**-5.1%**	**+4.0%**
200	**40**	5,813	**+1.2%**	**-1.3%**
200	**60**	5,520	**+4.8%**	**-5.5%**
200	**80**	5,861	**+6.8%**	**-0.5%**

Experiments

Runtime (minutes)

Data (GB)	# Topics
50	20
100	20
150	20
200	20
200	**40**
200	**60**
200	**80**

Glint	MLLib EM	MLLib Online
6.3	9.7	16.3
7.1	14.2	17.8
8.9	14.1	19.6
10.8	22.3	21.5
11.9	23.7	57.5
13.4	32.4	131.0
14.7	34.4	233.2

Experiments

Shuffle Write (GB)

Data (GB)	# Topics
50	20
100	20
150	20
200	20
200	**40**
200	**60**
200	**80**

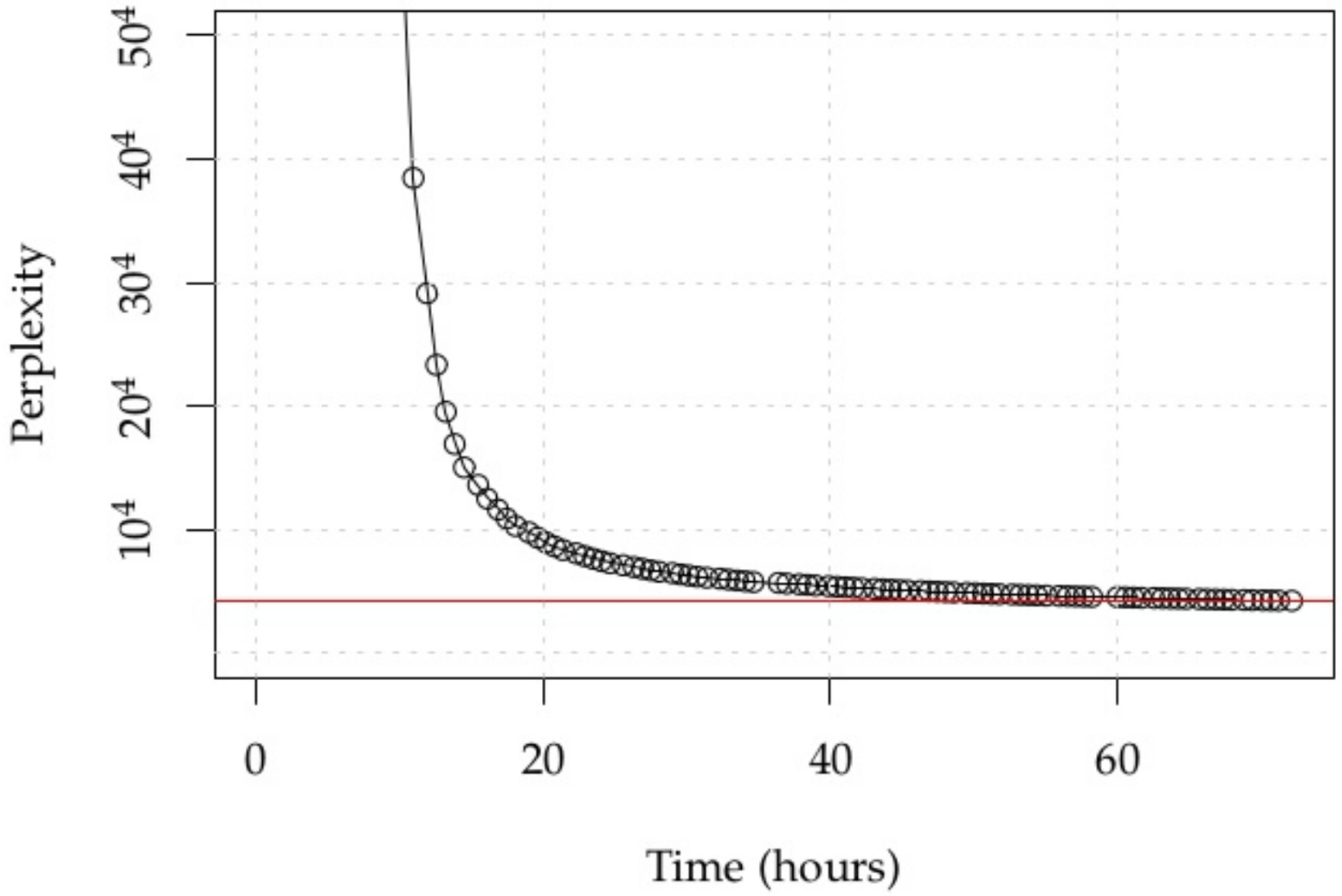
Glint	MLLib EM	MLLib Online
	3.3	
	4.6	
No Shuffle Write	5.5	No Shuffle Write
	6.2	
	12.1	
	18.0	
	23.9	

Experiments

- MLLib could not scale beyond 200GB or 100 topics due to task and job failures

- Glint can compute a topic model on the full **27TB** with **1,000** topics

Experiments



Conclusion

- Glint is a parameter server for Spark
- Machine learning for very large models
- Asynchrony enables highly flexible threading
- Extremely easy to use
- Outperforms MLLib on LDA topic modeling

Future work

- Better fault tolerance (using Chord/DHT)
- User defined functions for aggregation
- Support for sparse models
- Implementing other algorithms
  - Deep learning
  - Linear models

THANK YOU

Glint

github.com/rjagerman/glint