A Short Intro to Go

CS 240

What’s this Weird Language I’ve Never Heard of?

“Go is a
compiled,
concurrent,
statically typed,
garbage-collected
language developed at Google”

- Rob Pike, 2012

Rob Pike is the Jeff Dean of distributed systems. Here’s the article the quote is from: https://talks.golang.org/2012/splash.article

What’s this Weird Language I’ve Never Heard of?

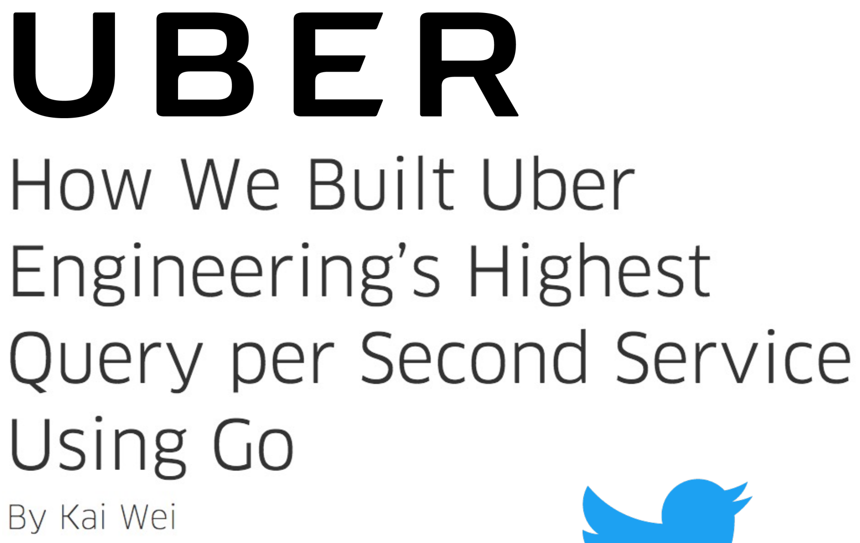
compiled	Like C, C++
concurrent	Like Erlang
statically typed	Like C, C++, Java
garbage-collected	Like Java and Python

Why Not Use Python, Java, C++, etc?

Built for Systems.
Go preserves efficiency but has good abstractions.
Easy multi threading and IO communication.

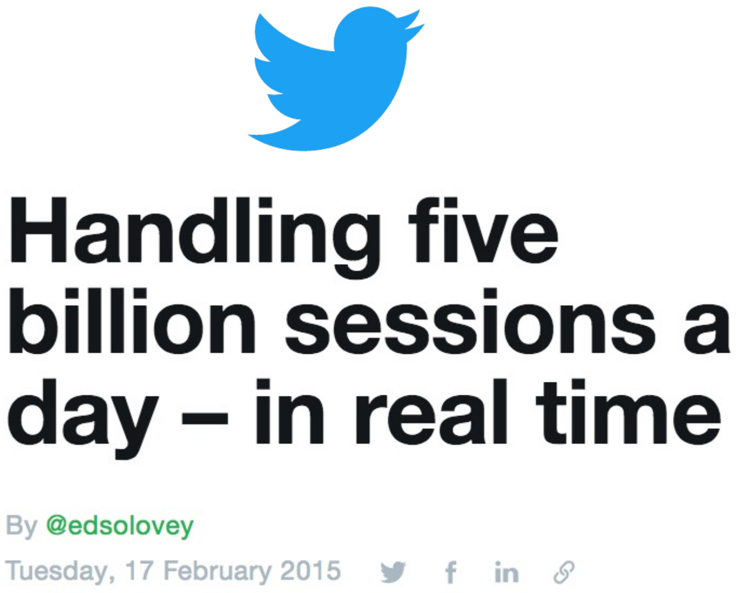
Develop quickly
Do many things efficiently *and at the same time*

Seems Google Specific. Who Else Actually Uses it?


UBER

How We Built Uber Engineering's Highest Query per Second Service Using Go

By Kai Wei


Handling five billion sessions a day – in real time

By @edsolovey

Tuesday, 17 February 2015

twitch

slack

SOUNDCLOUD

Dropbox

Official Self Reported List of Companies that use Go: https://github.com/golang/go/wiki/GoUsers

Why did *they* Choose Go?

“We built everything in Python because it was easy, but now it’s **slow**. So we switched to Go.”

- Most companies using Go

But How do I Use Go?

Start here:

https://tour.golang.org/list

Didn't install Go? Use the web IDE:

https://play.golang.org/

Other Resources:

Go for Pythonists

https://talks.golang.org/2013/go4python.slide#1

Go for Distributed Systems

https://talks.golang.org/2013/distsys.slide#1

Official Go Talks

https://github.com/golang/go/wiki/GoTalks

But How do I Use Go?

DEMO: go tour

Build Software for Any System

```
go build file.go
```

Compile an executable for your machine

```
env GOOS=windows GOARCH=amd64 go build file.go
```

Compile an executable for Windows with 64 bit processor

Format your Code

`COMMAND`	WHAT IT DOES
`gofmt file.go`	Format the file.go properly

DEMO: gofmt

Wait, I Have Questions!

Go’s official
“Frequently Asked Questions (FAQ)”
https://golang.org/doc/faq

MapReduce

CS 240

Map Reduce

Wikipedia:
*“MapReduce is a **programming model** and an associated implementation for processing and generating **big data** sets with a **parallel, distributed** algorithm on a **cluster**.”*

In other words, a general and scalable solution to deal with big data computation on multiple machines.

Abstract Map Reduce

```
map(key, value) -> list(<k’, v’>)
```

- Apply function to (key, value) pair
- Outputs set of intermediate pairs

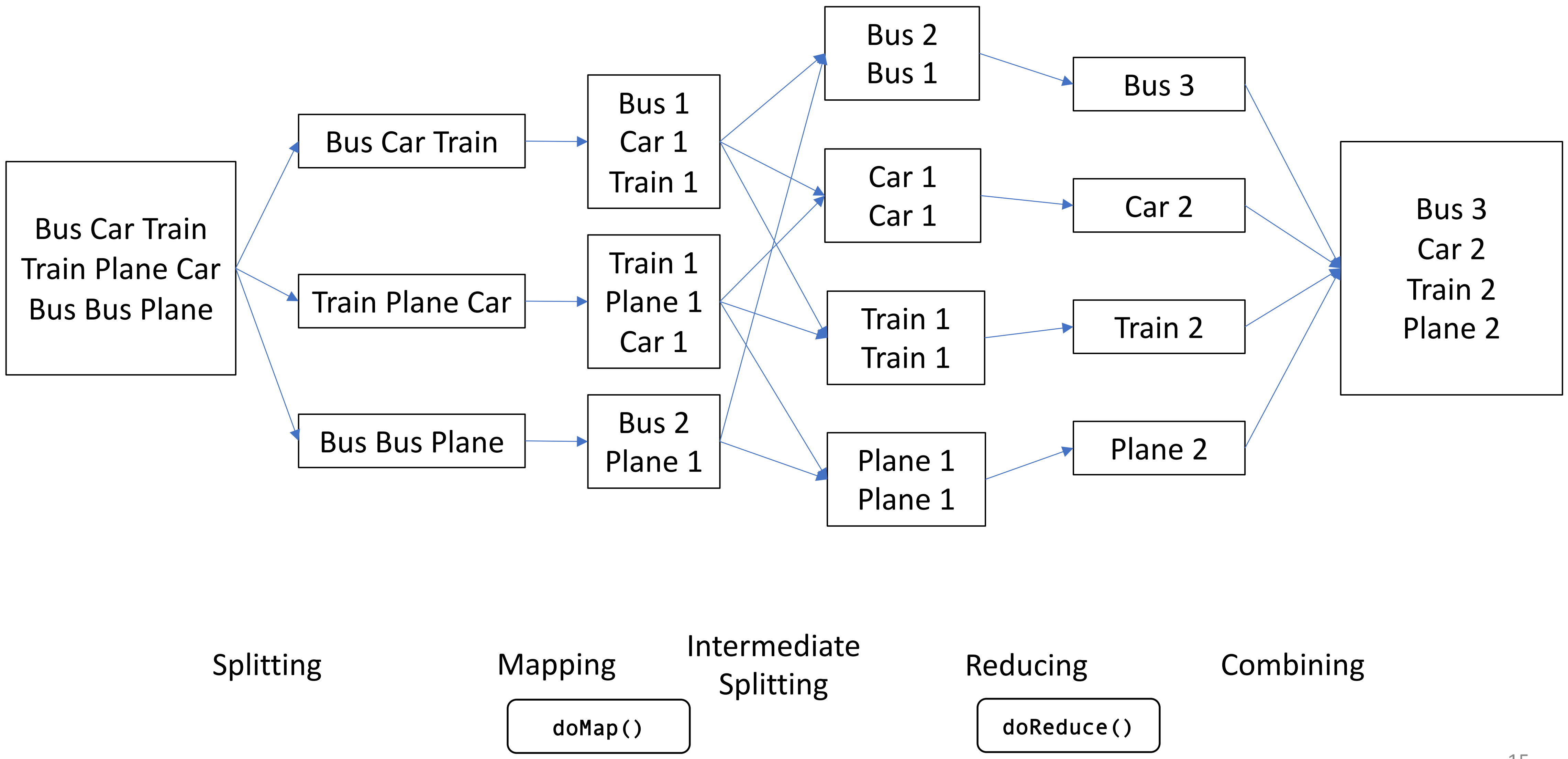
Split and distribute data

```
reduce(key, list<value>) -> <k’, v’>
```

- Applies aggregation function to values
- Outputs result

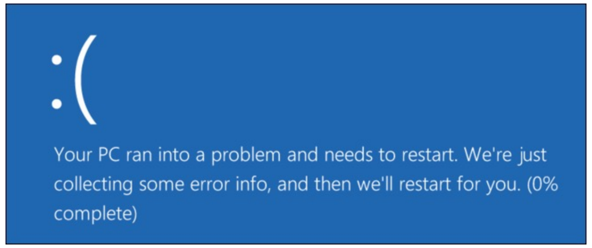
Aggregate and compute results

Word Count – The *Hello World* of Map Reduce



A Motivating Problem for Map Reduce

“Find me the closest Starbucks to KAUST.
Actually, I’ll give you a place and something to look for,
and you find me the closest one.
Here’s a 1 TB text file … good luck”


:(

Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you. (0% complete)

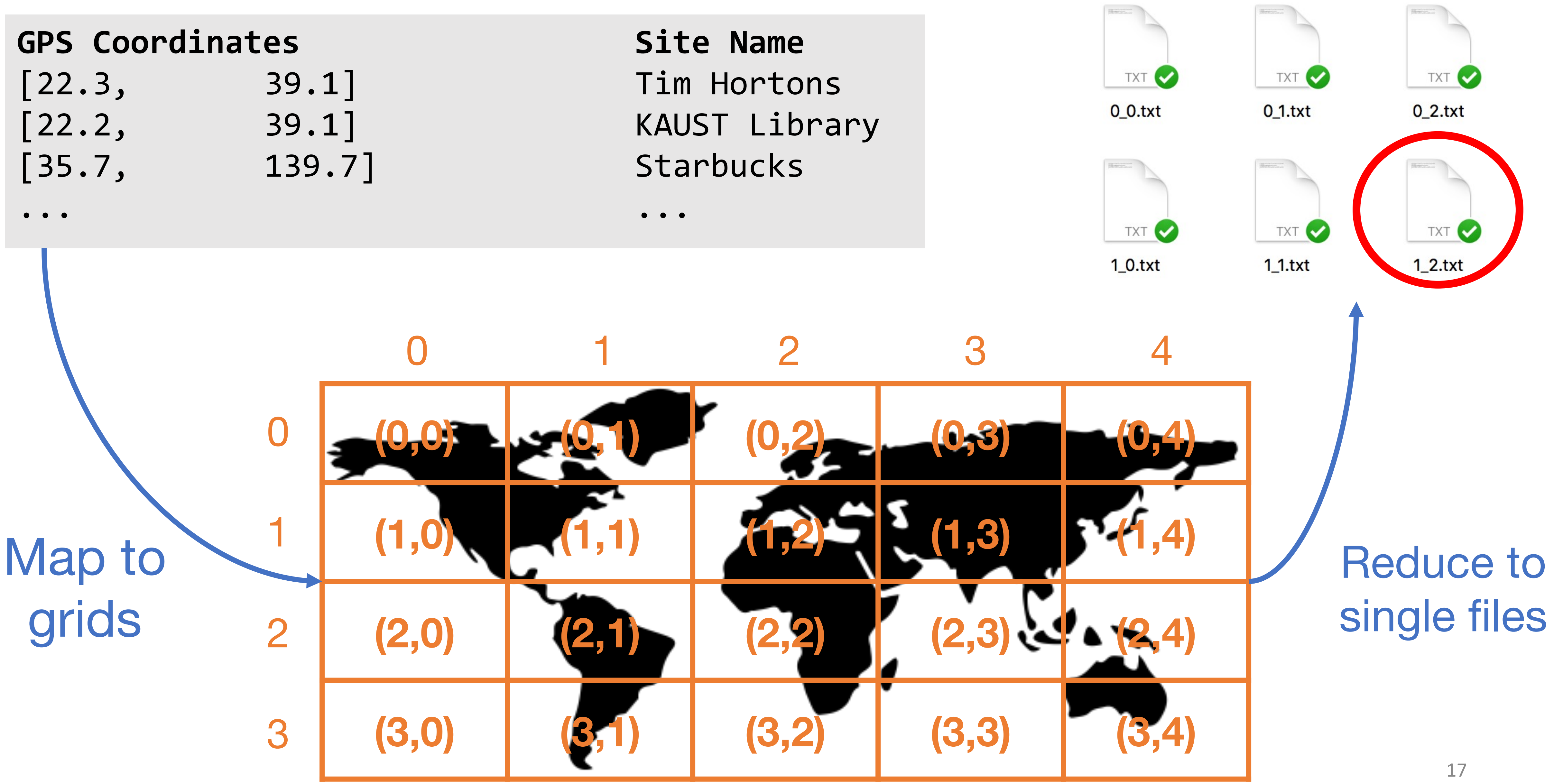
```
GPS Coordinates              Site Name
[22.3,     39.1]             Tim Hortons
[22.2,     39.1]             KAUST Library
[35.7,     139.7]            Starbucks
...                          ...
```

} In KAUST

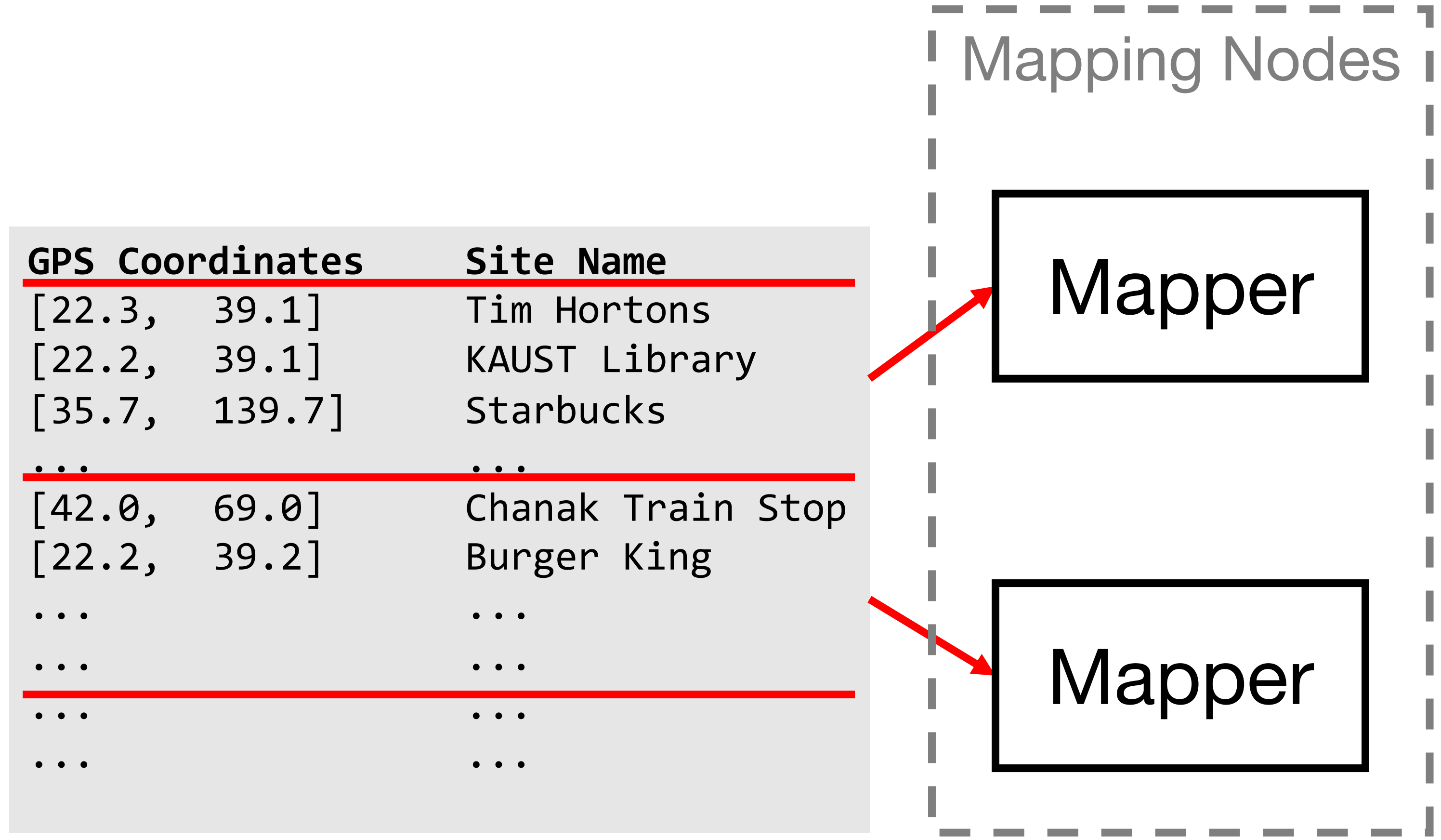
} In Tokyo, Japan

A Motivating Problem for Map Reduce

```
GPS Coordinates                Site Name
[22.3,       39.1]             Tim Hortons
[22.2,       39.1]             KAUST Library
[35.7,       139.7]            Starbucks
...                            ...
```



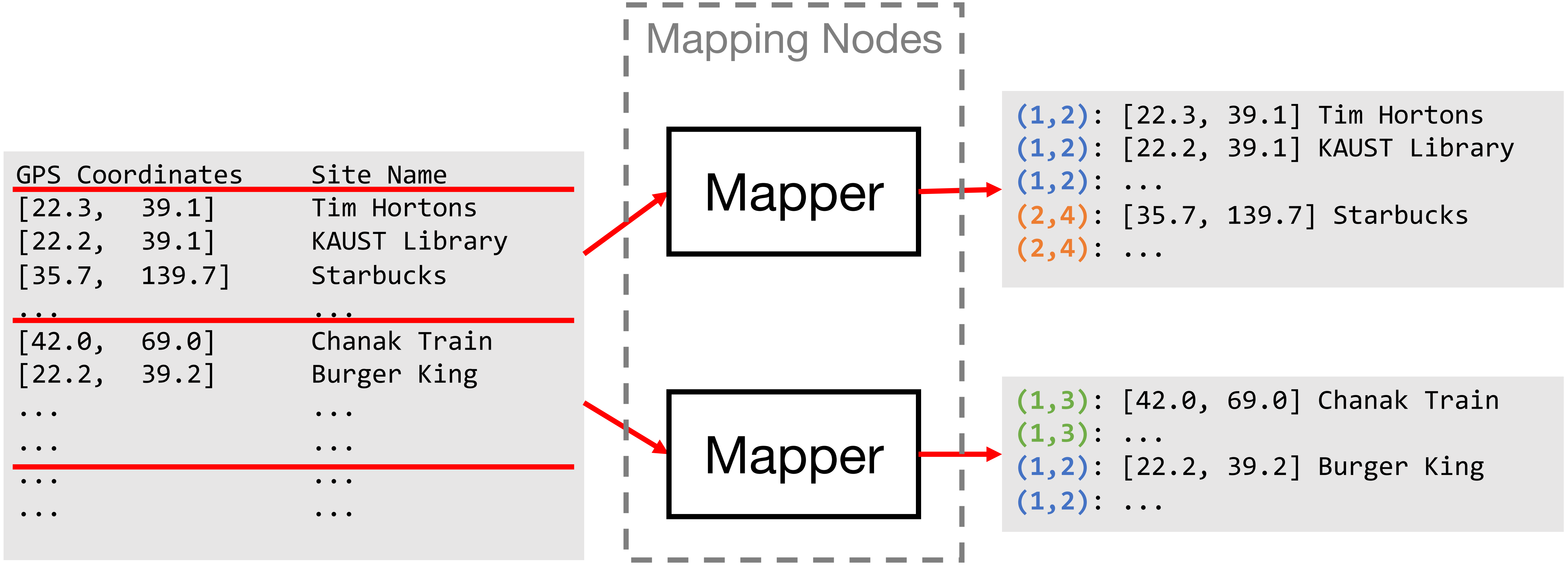
Split the File and Map Each Chunk Independently (1/2)



```
GPS Coordinates       Site Name
[22.3,   39.1]        Tim Hortons
[22.2,   39.1]        KAUST Library
[35.7,   139.7]       Starbucks
...                   ...
[42.0,   69.0]        Chanak Train Stop
[22.2,   39.2]        Burger King
...                   ...
...                   ...
...                   ...
...                   ...
```

Split the File and Map Each Chunk Independently (2/2)

```
KEY <grid>: VALUE <locations and name>
...
```



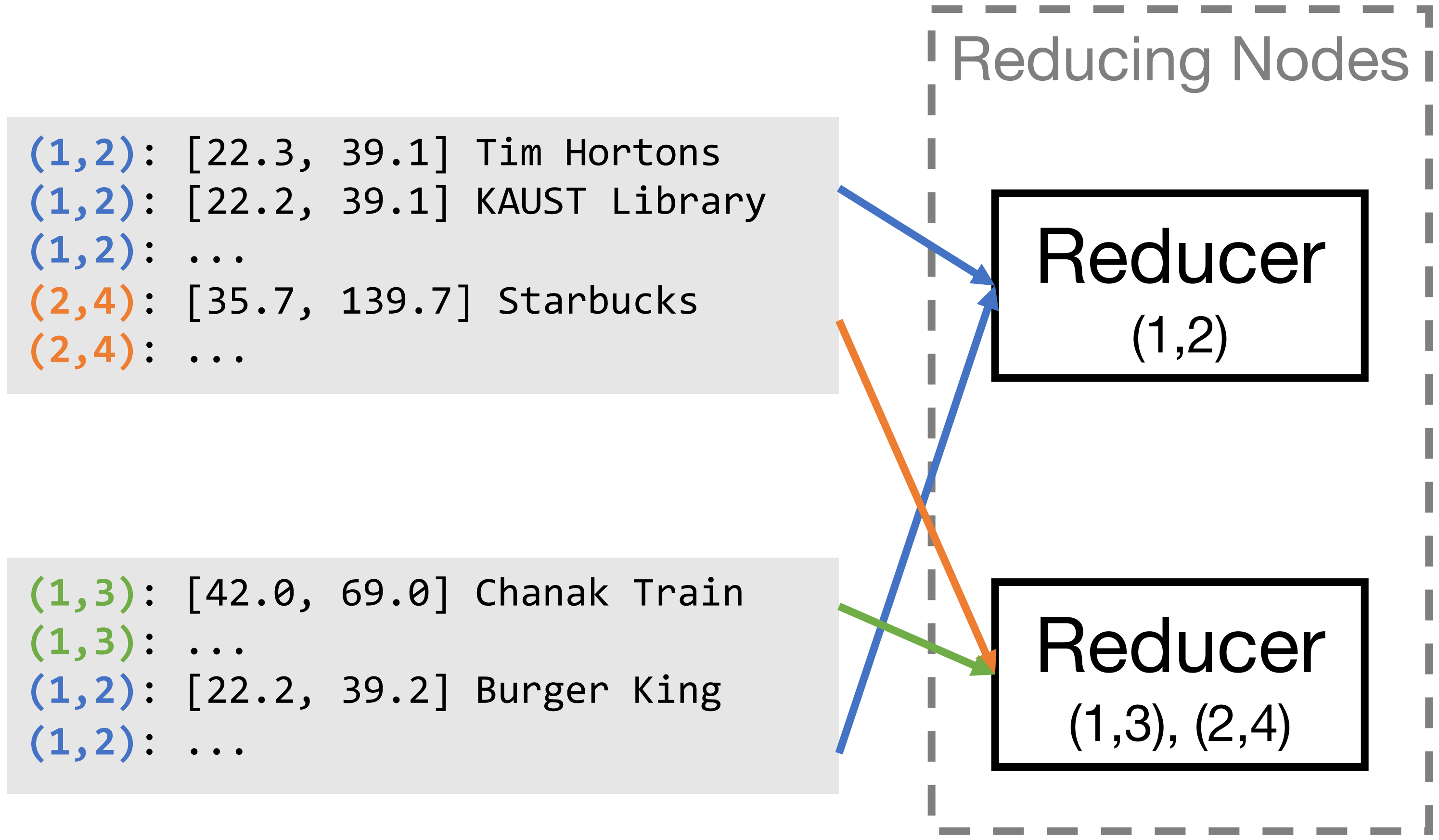
```
GPS Coordinates        Site Name
[22.3,   39.1]         Tim Hortons
[22.2,   39.1]         KAUST Library
[35.7,   139.7]        Starbucks
...                    ...
[42.0,   69.0]         Chanak Train
[22.2,   39.2]         Burger King
...                    ...
...                    ...
...                    ...
...                    ...
```

```
(1,2): [22.3, 39.1] Tim Hortons
(1,2): [22.2, 39.1] KAUST Library
(1,2): ...
(2,4): [35.7, 139.7] Starbucks
(2,4): ...
```

```
(1,3): [42.0, 69.0] Chanak Train
(1,3): ...
(1,2): [22.2, 39.2] Burger King
(1,2): ...
```

(KEY) can appear in multiple mappers

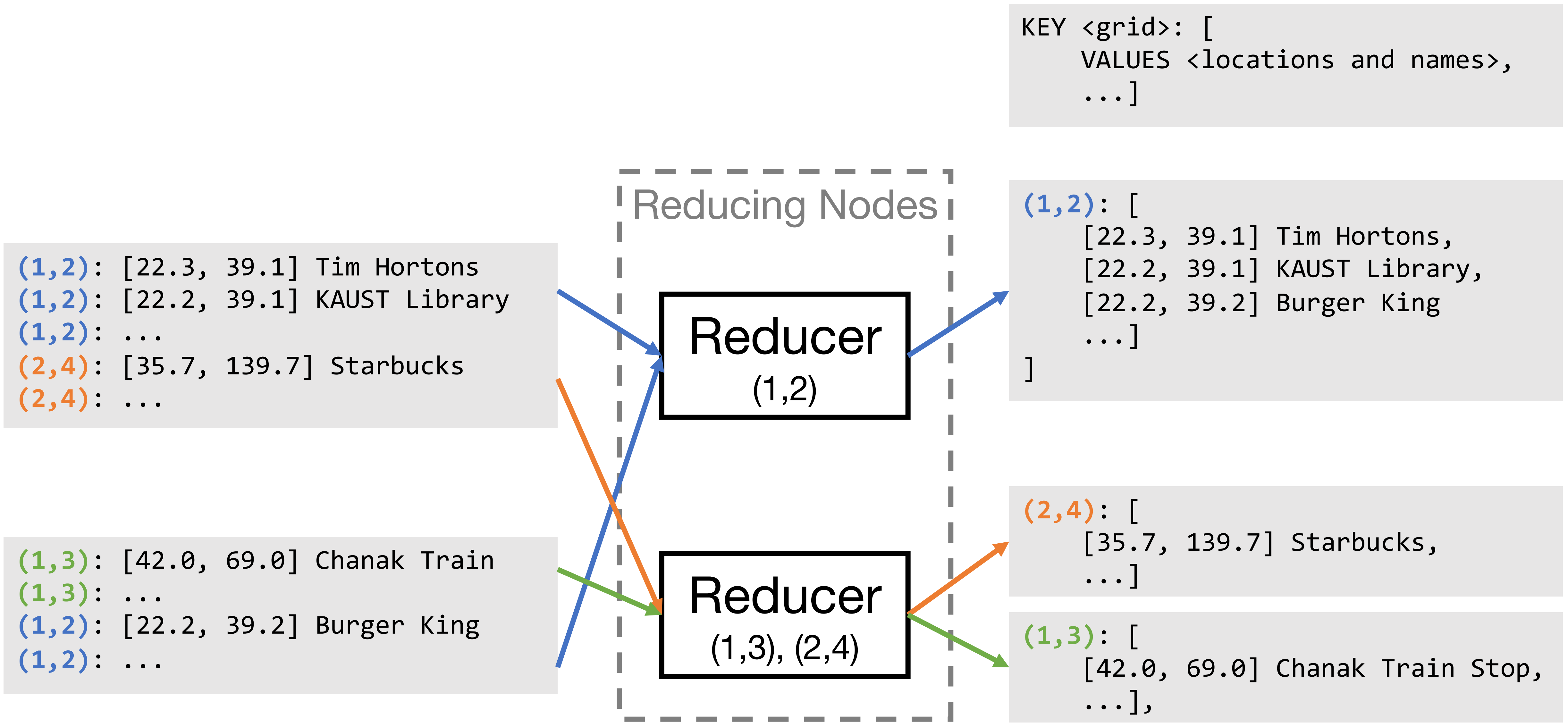
Collect the Mapper Results and Reduce to Single Files (1/2)



```
(1,2): [22.3, 39.1] Tim Hortons
(1,2): [22.2, 39.1] KAUST Library
(1,2): ...
(2,4): [35.7, 139.7] Starbucks
(2,4): ...
```

```
(1,3): [42.0, 69.0] Chanak Train
(1,3): ...
(1,2): [22.2, 39.2] Burger King
(1,2): ...
```

Collect the Mapper Results and Reduce to Single Files (2/2)



```
KEY <grid>: [
    VALUES <locations and names>,
    ...]
```

```
(1,2): [22.3, 39.1] Tim Hortons
(1,2): [22.2, 39.1] KAUST Library
(1,2): ...
(2,4): [35.7, 139.7] Starbucks
(2,4): ...
```

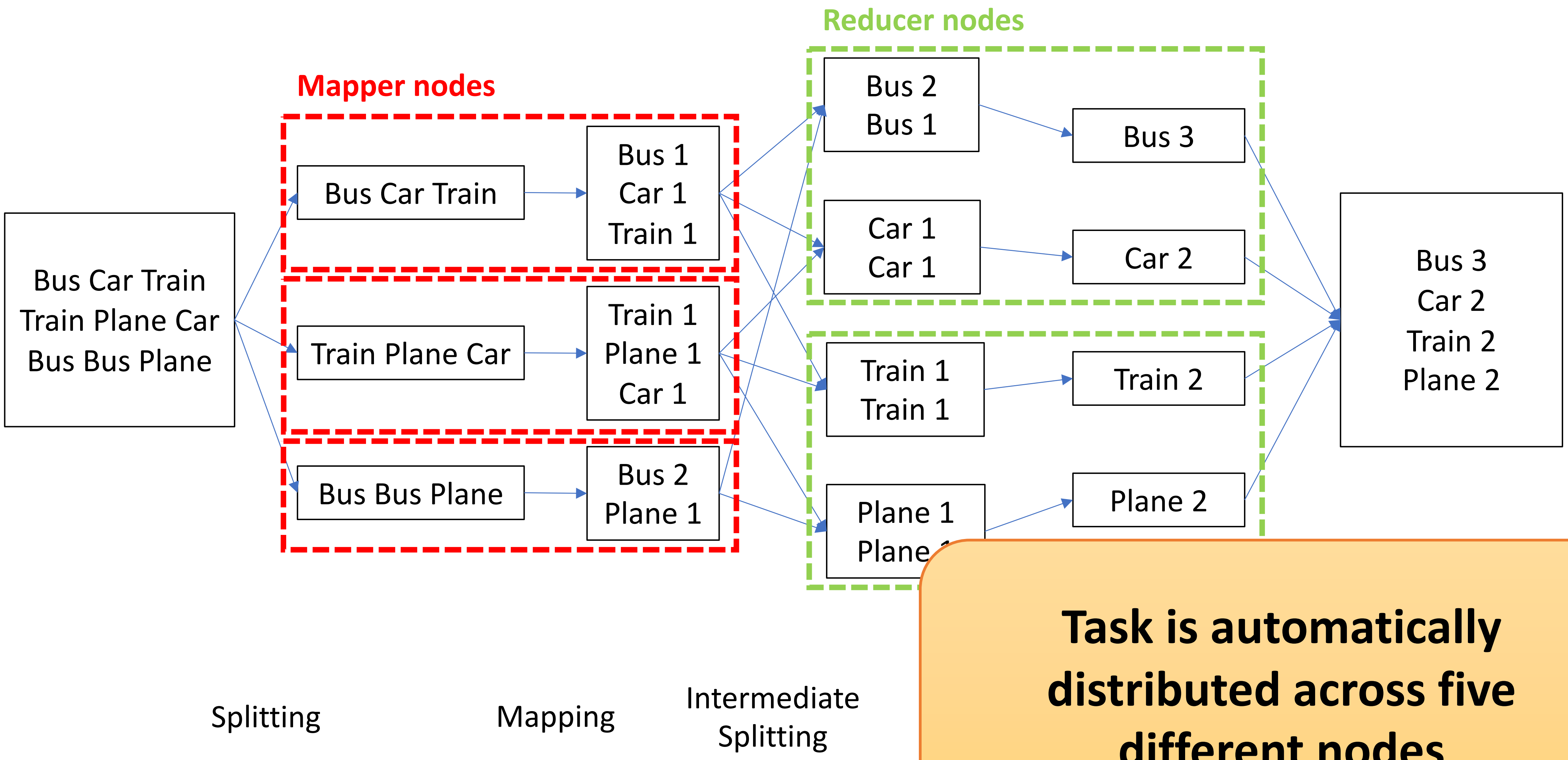
```
(1,3): [42.0, 69.0] Chanak Train
(1,3): ...
(1,2): [22.2, 39.2] Burger King
(1,2): ...
```

```
(1,2): [
    [22.3, 39.1] Tim Hortons,
    [22.2, 39.1] KAUST Library,
    [22.2, 39.2] Burger King
    ...]
]
```

```
(2,4): [
    [35.7, 139.7] Starbucks,
    ...]
```

```
(1,3): [
    [42.0, 69.0] Chanak Train Stop,
    ...],
```

Word Count – The *Hello World* of Map Reduce



Task is automatically distributed across five different nodes

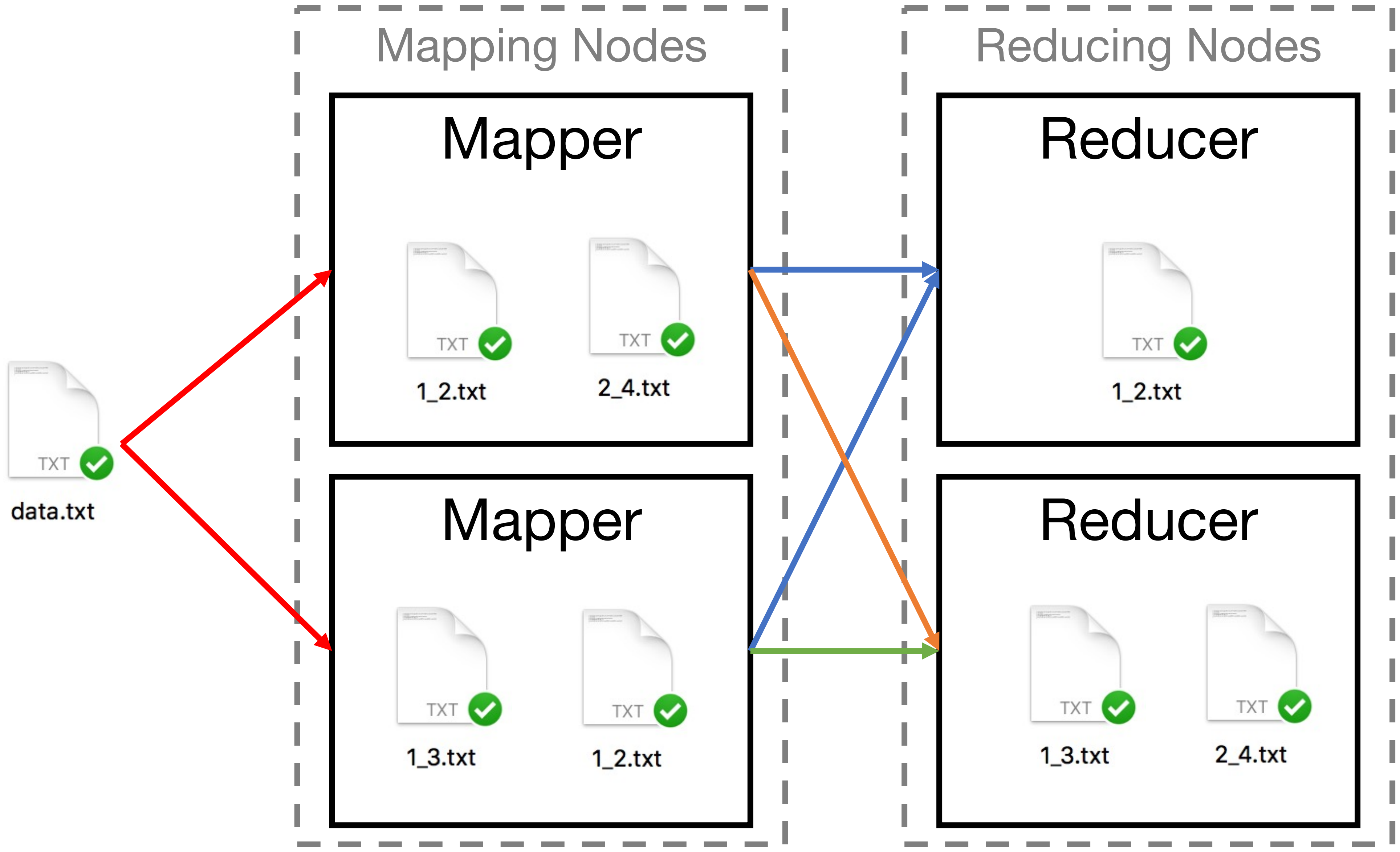
Hadoop: An open-source implementation

Apache Hadoop is the most popular open-source implementation of MapReduce

Runs on top of a distributed filesystem (HDFS)

Try their MapReduce tutorial:
https://hadoop.apache.org/docs/r3.3.1/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html

How Hadoop Does it



Some Advice for the Assignments

- Write modular code
- Use comments (even to yourself)
- Don’t forget `gofmt` (graded)
- The clearer your code is, the more we can help with bugs