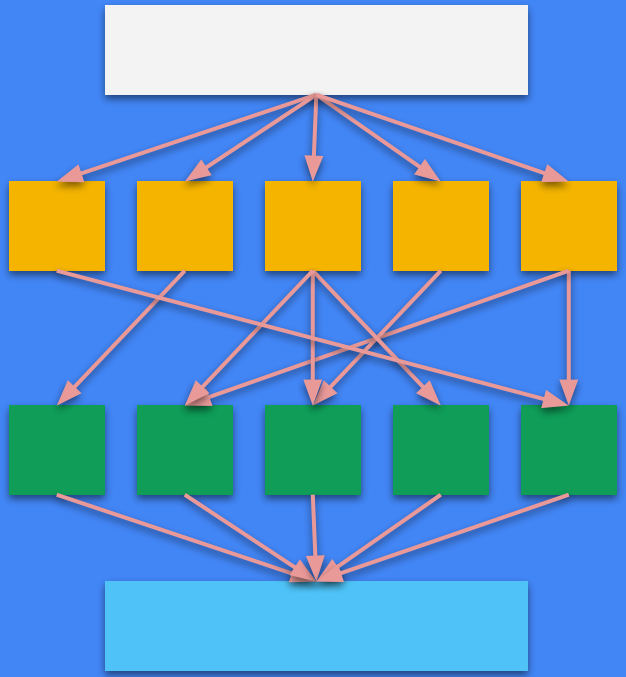
NYU Center for Data Science

Week 03.1: Map-Reduce

DS-GA 1004: Big Data

This week



1. **Introduction to Map-Reduce**
   (Dean & Ghemawat, 2008)

2. Criticisms of Map-Reduce
   (DeWitt & Stonebreaker, 2008)

Motivation: text indexing

- Say you have *N* documents (with *N* very large, e.g. the web), and you want to construct an index: **words → documents**
- On a single machine, this process takes **Ω(*N*)** time
- **Observation**: this problem is (almost) embarrassingly **parallel**
  - Whether any word appears in a document is **independent** of other documents
  - We should be able to process documents independently and **combine the results**

Indexing continued

- You could have multiple computers write to a shared database
  - With *M* machines, can we lower the time to $\Omega(N/M)$?
- You'll need some way to **distribute** work (and data) and **collect** results
- **Map-Reduce** (Dean & Ghemawat, 2004) provides a framework for this
- **Hadoop** (2008-) provides an open source implementation of Map-Reduce
  - … and supporting infrastructure for distributed computing

Power through restrictions

- **RDBMS/SQL** empowers us by **restricting** how we store and query **data**
- **Map-Reduce** empowers us by **restricting** how we implement **algorithms**

Why “map” and “reduce”?

- These are common operations in **functional programming**
  - E.g.: LISP, ML, Haskell, Scala…
- **map**(function *f*, values $[x_1, x_2, \ldots, x_n]$) → $[f(x_1), f(x_2), \ldots f(x_n)]$
  - **map** : function, list → list
- **reduce**(function *g*, values $[x_1, x_2, \ldots, x_n]$) → $g(x_1, \textbf{reduce}(g, [x_2, \ldots, x_n]))$
  - **reduce** : function, list → item

Example: sum of squares

- Define functions “**sum**” and “**square**”
  - **sum** : x, y → x + y **sum** : [ ] → 0
  - **square** : x → x * x

- **reduce**(**sum**, **map**(**square**, [x_1, x_2, …, x_n]))

Example: sum of squares

- Define functions “**sum**” and “**square**”
  - **sum** : x, y → x + y **sum** : [ ] → 0
  - **square** : x → x * x

- **reduce**(**sum**, **map**(**square**, [x_1, x_2, …, x_n])) → **sum**(**square**(x_1), **sum**(**square**(x_2), **sum**(⋯)))

Working with Map-Reduce

Conceptual framework

- You (the programmer) provide two functions: **mapper** and **reducer**
  - Can be arbitrarily complex, but simpler is better!
- The **mapper** consumes inputs, produces outputs of the form:
  (*key*, *value*)
- The **reducer** consumes a single *key* and list of *values*, and produces *values*

Map-Reduce flow

1. Map phase
   a. Distribute data to mappers
   b. Generate intermediate results (*key*, *value*)
2. Sort / shuffle phase
   a. Assign intermediate results to reducers (by *key*)
   b. Move data from mappers to reducers
3. Reduce phase
   a. Execute reducers and collect output

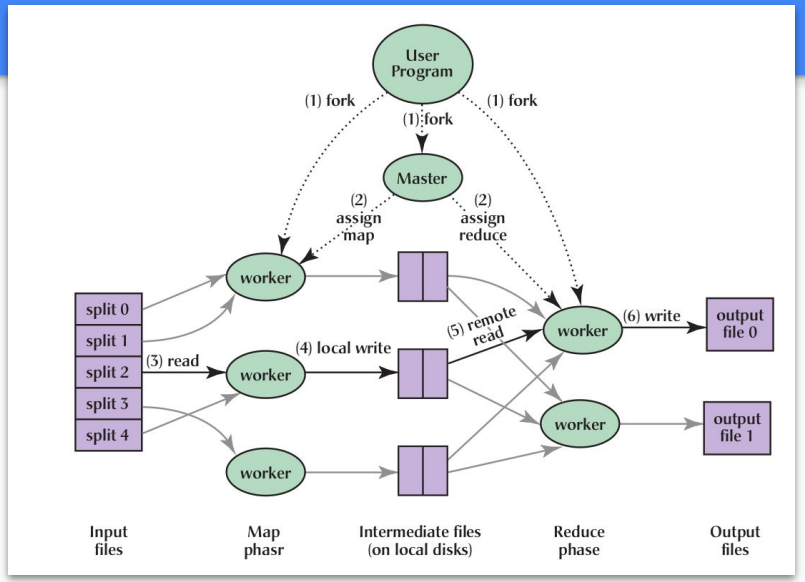


Figure adapted from Dean & Ghemawat, 2008

Ex.: word counting in a document collection

```
mapper(doc_id, doc_contents):
  for word in doc_contents:
    emit word, 1
```

```
reducer(word, counts):
  total_count = 0
  for count in counts:
    total_count += count
  emit total_count
```

Key idea:

- Make the mapper as simple as possible
- Let the MR framework route the intermediate results
- Reduce can be simple as well

INPUTS

doc_id: 1

Spinosaurus (meaning "spine lizard") is a genus of theropod dinosaur that lived in what now is North Africa, during the upper Albian to upper Turonian stages of the Cretaceous period, about 112 to 93.5 million ...

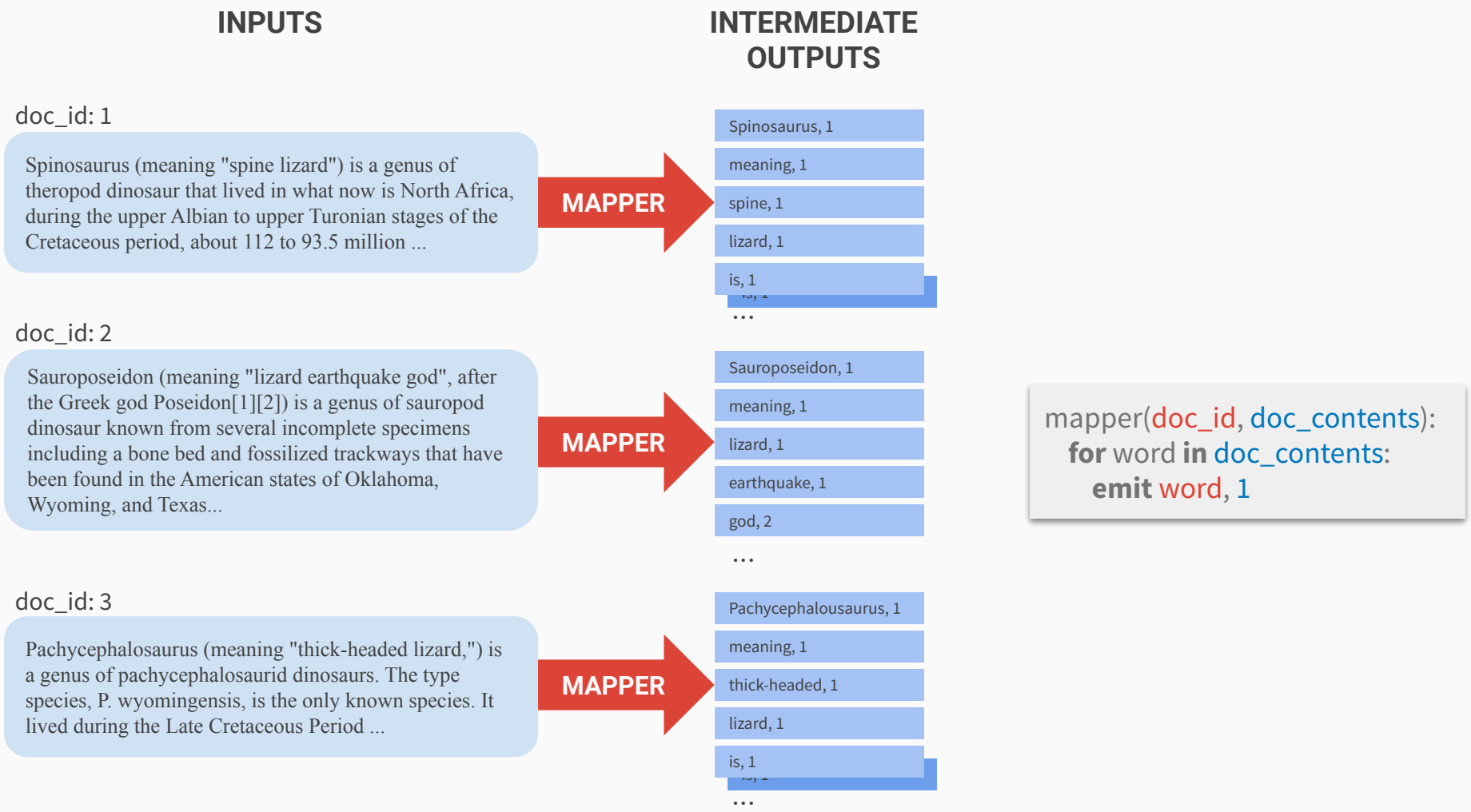
doc_id: 2

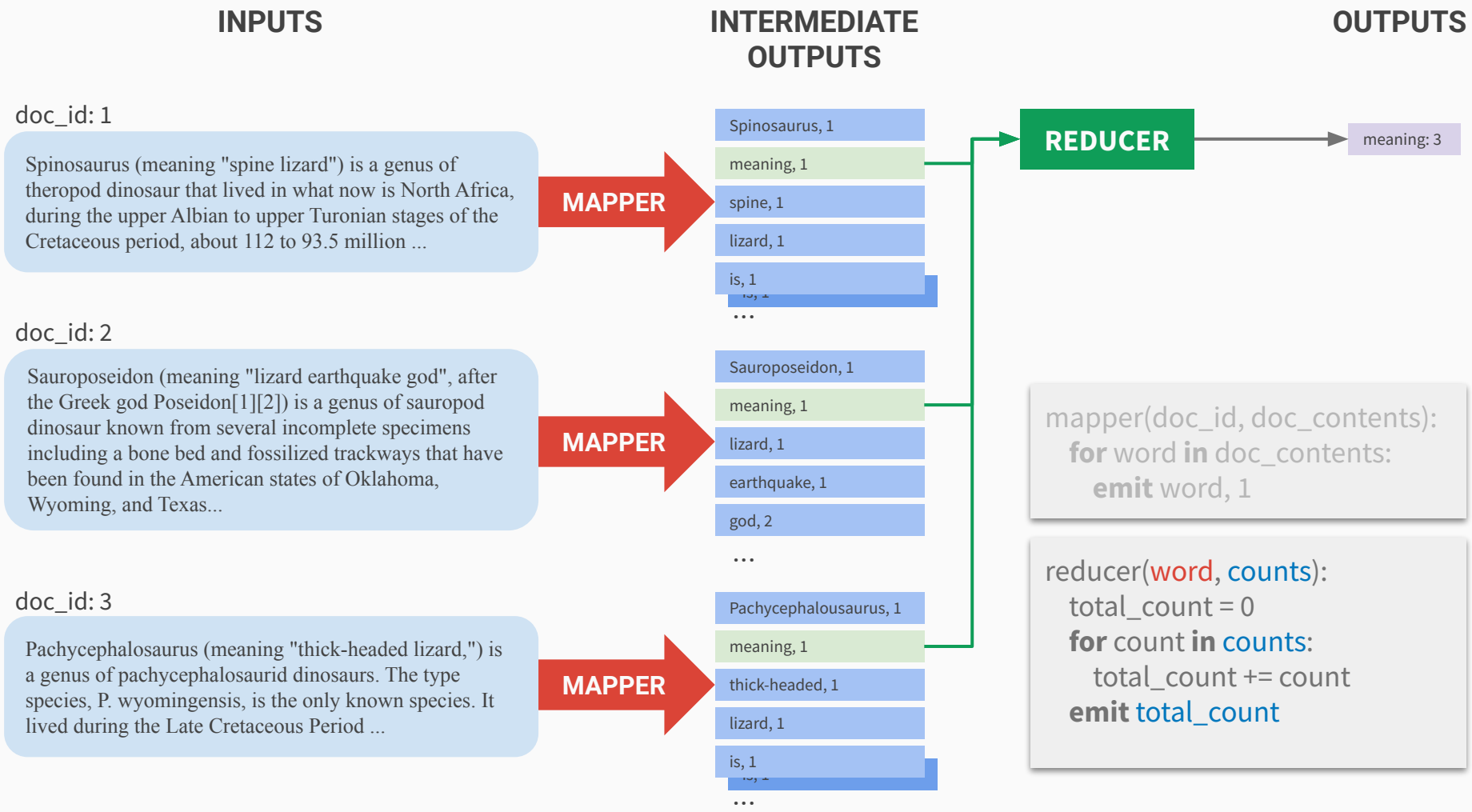
Sauroposeidon (meaning "lizard earthquake god", after the Greek god Poseidon[1][2]) is a genus of sauropod dinosaur known from several incomplete specimens including a bone bed and fossilized trackways that have been found in the American states of Oklahoma, Wyoming, and Texas...

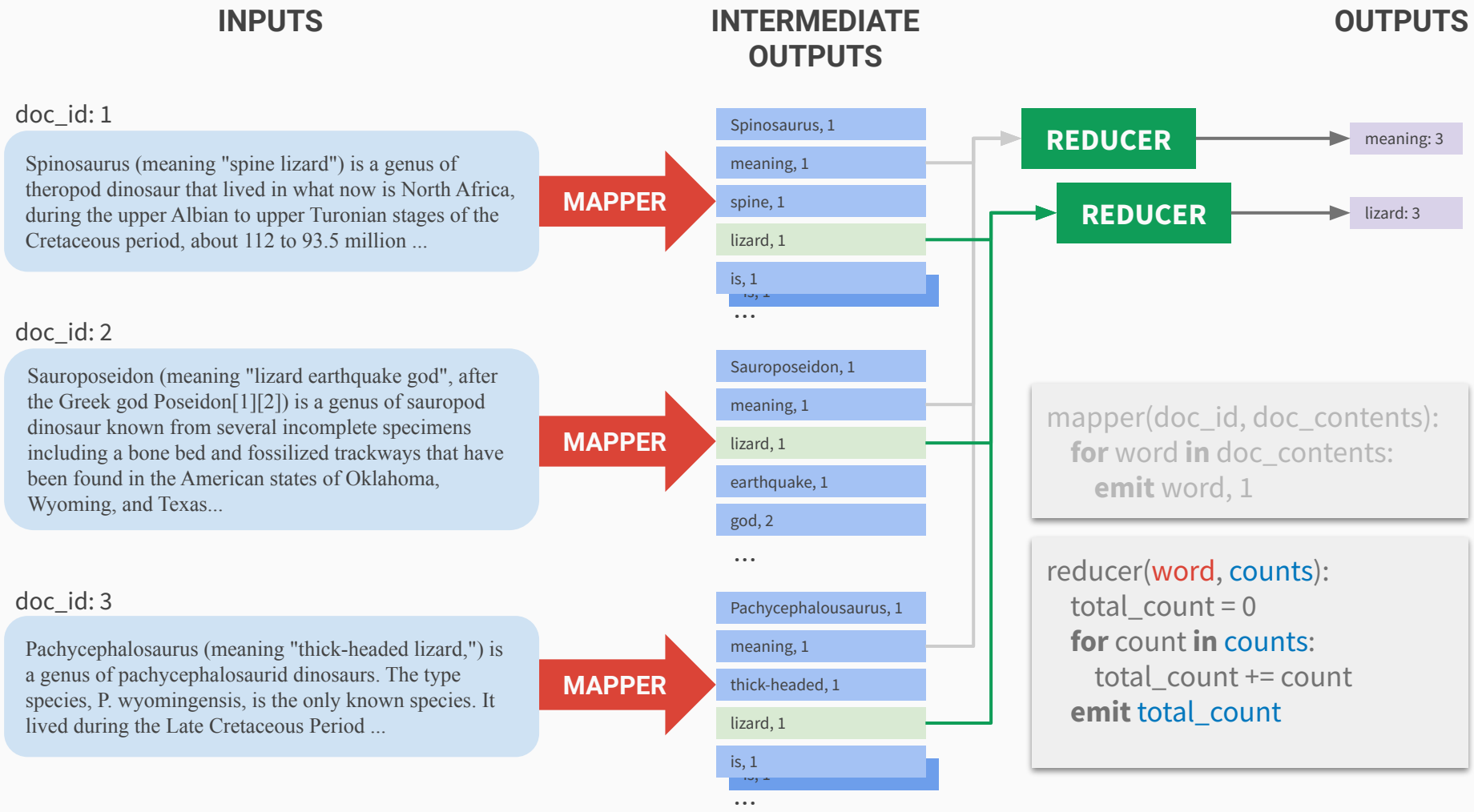
doc_id: 3

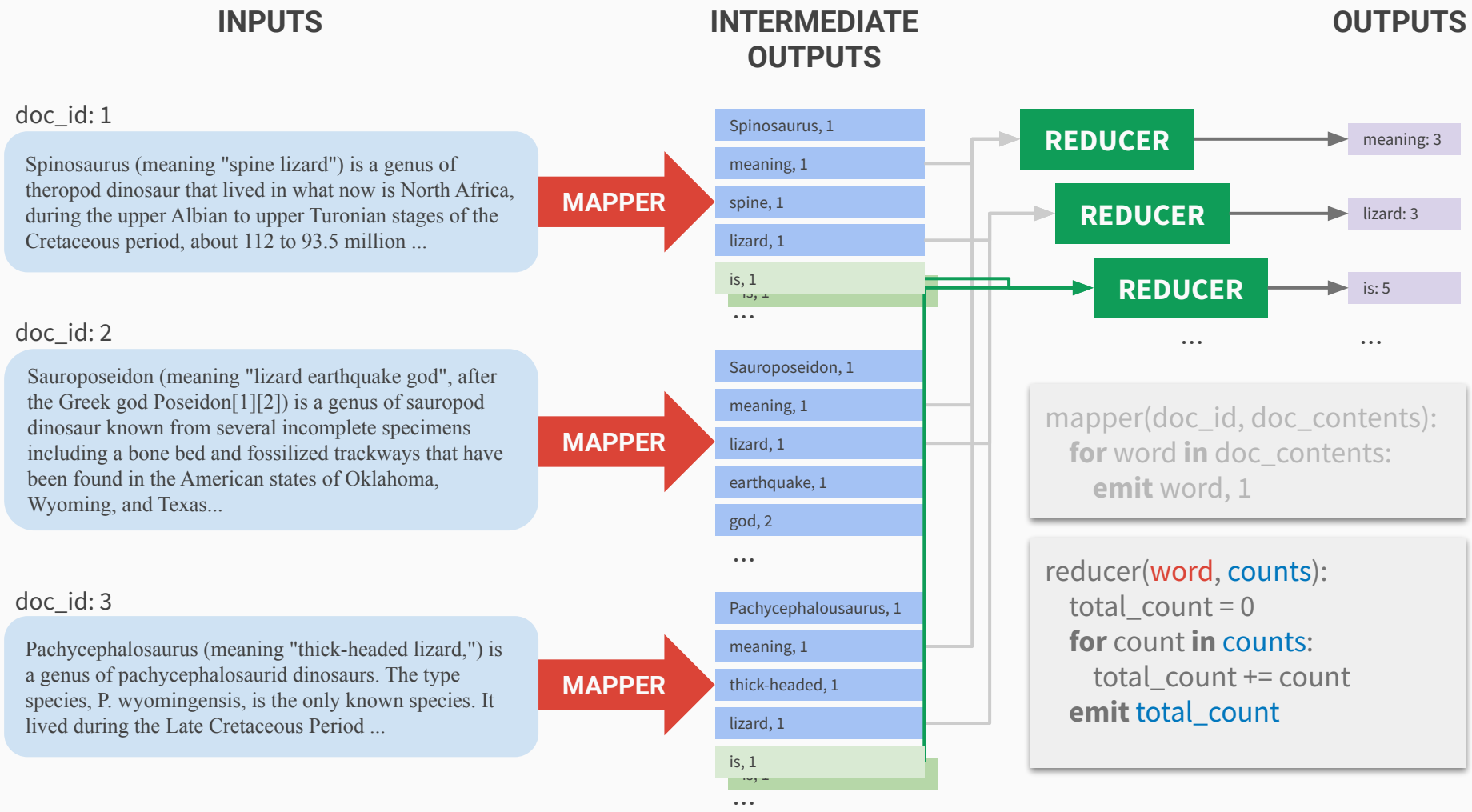
Pachycephalosaurus (meaning "thick-headed lizard,") is a genus of pachycephalosaurid dinosaurs. The type species, P. wyomingensis, is the only known species. It lived during the Late Cretaceous Period ...

```
mapper(doc_id, doc_contents):
  for word in doc_contents:
    emit word, 1
```


INPUTS
INTERMEDIATE OUTPUTS
doc_id: 1
Spinosaurus (meaning "spine lizard") is a genus of theropod dinosaur that lived in what now is North Africa, during the upper Albian to upper Turonian stages of the Cretaceous period, about 112 to 93.5 million ...
MAPPER
Spinosaurus, 1
meaning, 1
spine, 1
lizard, 1
is, 1
...
doc_id: 2
Sauroposeidon (meaning "lizard earthquake god", after the Greek god Poseidon[1][2]) is a genus of sauropod dinosaur known from several incomplete specimens including a bone bed and fossilized trackways that have been found in the American states of Oklahoma, Wyoming, and Texas...
MAPPER
Sauroposeidon, 1
meaning, 1
lizard, 1
earthquake, 1
god, 2
...
doc_id: 3
Pachycephalosaurus (meaning "thick-headed lizard,") is a genus of pachycephalosaurid dinosaurs. The type species, P. wyomingensis, is the only known species. It lived during the Late Cretaceous Period ...
MAPPER
Pachycephalousaurus, 1
meaning, 1
thick-headed, 1
lizard, 1
is, 1
...
mapper(doc_id, doc_contents):
for word in doc_contents:
emit word, 1


INPUTS
INTERMEDIATE OUTPUTS
OUTPUTS
doc_id: 1
Spinosaurus (meaning "spine lizard") is a genus of theropod dinosaur that lived in what now is North Africa, during the upper Albian to upper Turonian stages of the Cretaceous period, about 112 to 93.5 million ...
doc_id: 2
Sauroposeidon (meaning "lizard earthquake god", after the Greek god Poseidon[1][2]) is a genus of sauropod dinosaur known from several incomplete specimens including a bone bed and fossilized trackways that have been found in the American states of Oklahoma, Wyoming, and Texas...
doc_id: 3
Pachycephalosaurus (meaning "thick-headed lizard,") is a genus of pachycephalosaurid dinosaurs. The type species, P. wyomingensis, is the only known species. It lived during the Late Cretaceous Period ...
MAPPER
MAPPER
MAPPER
Spinosaurus, 1
meaning, 1
spine, 1
lizard, 1
is, 1
...
Sauroposeidon, 1
meaning, 1
lizard, 1
earthquake, 1
god, 2
...
Pachycephalousaurus, 1
meaning, 1
thick-headed, 1
lizard, 1
is, 1
...
REDUCER
meaning: 3
mapper(doc_id, doc_contents):
for word in doc_contents:
emit word, 1
reducer(word, counts):
total_count = 0
for count in counts:
total_count += count
emit total_count


INPUTS
INTERMEDIATE OUTPUTS
OUTPUTS
doc_id: 1
Spinosaurus (meaning "spine lizard") is a genus of theropod dinosaur that lived in what now is North Africa, during the upper Albian to upper Turonian stages of the Cretaceous period, about 112 to 93.5 million ...
doc_id: 2
Sauroposeidon (meaning "lizard earthquake god", after the Greek god Poseidon[1][2]) is a genus of sauropod dinosaur known from several incomplete specimens including a bone bed and fossilized trackways that have been found in the American states of Oklahoma, Wyoming, and Texas...
doc_id: 3
Pachycephalosaurus (meaning "thick-headed lizard,") is a genus of pachycephalosaurid dinosaurs. The type species, P. wyomingensis, is the only known species. It lived during the Late Cretaceous Period ...
MAPPER
MAPPER
MAPPER
Spinosaurus, 1
meaning, 1
spine, 1
lizard, 1
is, 1
...
Sauroposeidon, 1
meaning, 1
lizard, 1
earthquake, 1
god, 2
...
Pachycephalousaurus, 1
meaning, 1
thick-headed, 1
lizard, 1
is, 1
...
REDUCER
REDUCER
meaning: 3
lizard: 3
mapper(doc_id, doc_contents):
for word in doc_contents:
emit word, 1
reducer(word, counts):
total_count = 0
for count in counts:
total_count += count
emit total_count


INPUTS
INTERMEDIATE OUTPUTS
OUTPUTS
doc_id: 1
Spinosaurus (meaning "spine lizard") is a genus of theropod dinosaur that lived in what now is North Africa, during the upper Albian to upper Turonian stages of the Cretaceous period, about 112 to 93.5 million ...
doc_id: 2
Sauroposeidon (meaning "lizard earthquake god", after the Greek god Poseidon[1][2]) is a genus of sauropod dinosaur known from several incomplete specimens including a bone bed and fossilized trackways that have been found in the American states of Oklahoma, Wyoming, and Texas...
doc_id: 3
Pachycephalosaurus (meaning "thick-headed lizard,") is a genus of pachycephalosaurid dinosaurs. The type species, P. wyomingensis, is the only known species. It lived during the Late Cretaceous Period ...
MAPPER
MAPPER
MAPPER
Spinosaurus, 1
meaning, 1
spine, 1
lizard, 1
is, 1
Sauroposeidon, 1
meaning, 1
lizard, 1
earthquake, 1
god, 2
Pachycephalosaurus, 1
meaning, 1
thick-headed, 1
lizard, 1
is, 1
REDUCER
REDUCER
REDUCER
meaning: 3
lizard: 3
is: 5
mapper(doc_id, doc_contents):
for word in doc_contents:
emit word, 1
reducer(word, counts):
total_count = 0
for count in counts:
total_count += count
emit total_count

Map-Reduce in practice

Sorting and shuffling

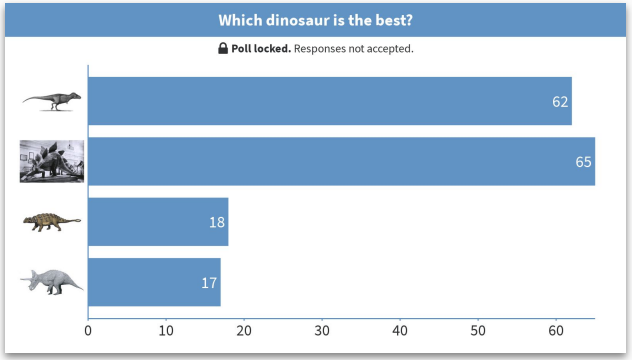
- Say the mapping stage produces some *K* total intermediate keys
- And we have some *R* reducer nodes
- Each key with one of the reducers by a **partition function**
- Often done by hashing the key: $k \rightarrow \textbf{hash}(k) \textbf{ mod } R$
- Between map and reduce, intermediate are **shuffled** over the network

Key → reducer assignment

- **All values** for a given key *k* go to **exactly one reducer**
- Conversely: a reducer acting on some intermediate key *k* needs to see all associated values
- This will have consequences!

Key-skew

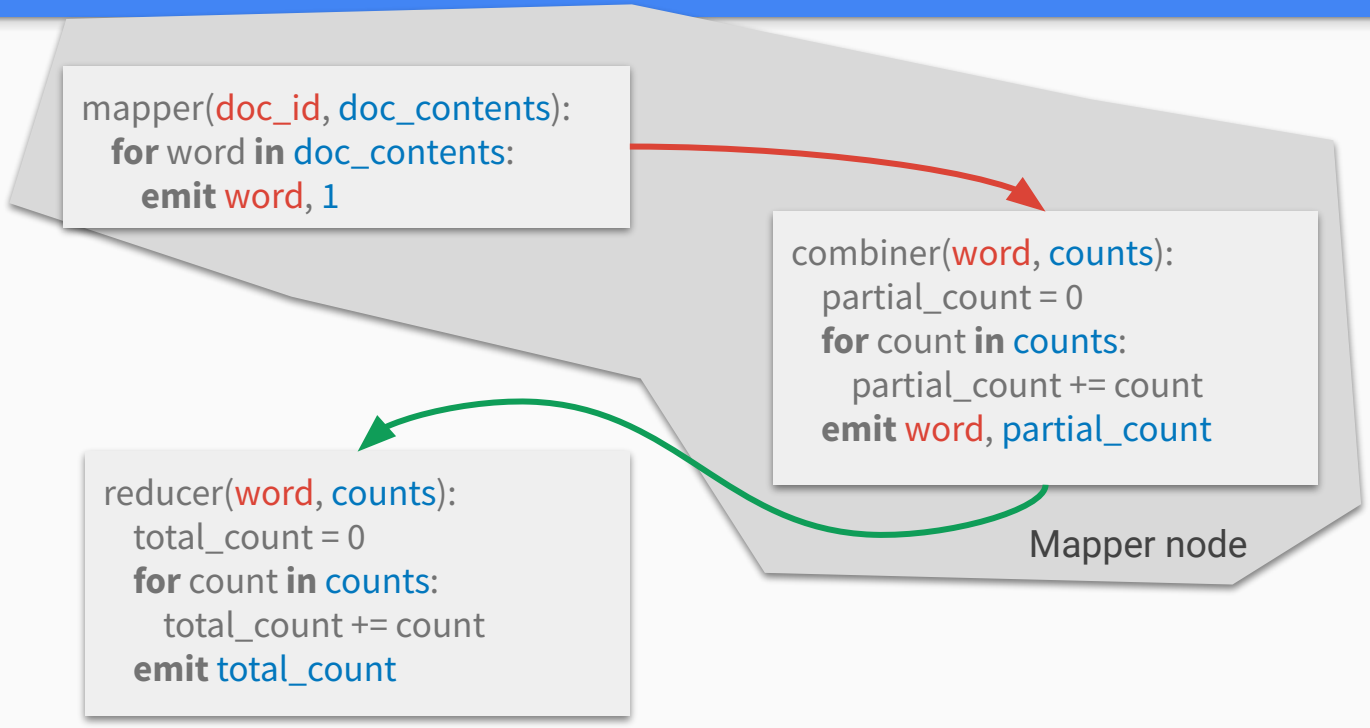
- What happens when the intermediate key distribution is **unbalanced**?
- All values for the **same key** must go to the **same reducer**
- **Different reducers** will have **different work loads**
- This is called **key skew** (or **data skew**)
  - It's a **bad thing!**



Combiners

- Key-skew leads to **high latency**
  - Reducer time scales (at least) like # values per key
- Lots of keys ⇒ lots of communication
  - Shuffling data is expensive!
- We can sometimes simplify the reducer's job by having mappers pre-reduce (**combine**) data before shuffling

Combiner example: word count



This works because summation is **commutative** and **associative:**

A + B = B + A

A + B + C = (A + B) + C

When that happens, you can re-use the **reducer** code as a **combiner**!

Heuristics for using MR well

- Have **fewer mappers than inputs**
- Have **fewer reducers than intermediate keys**
  - This keeps nodes busy
- **Combiners** can help, but sometimes a fancier map is better
- Sometimes you can be clever with sorting to reduce communication

Summary

- The Map-Reduce framework simplifies how we think about distributed computation
- MR was critical to the development of large-scale data analysis
- But it’s not without drawbacks…