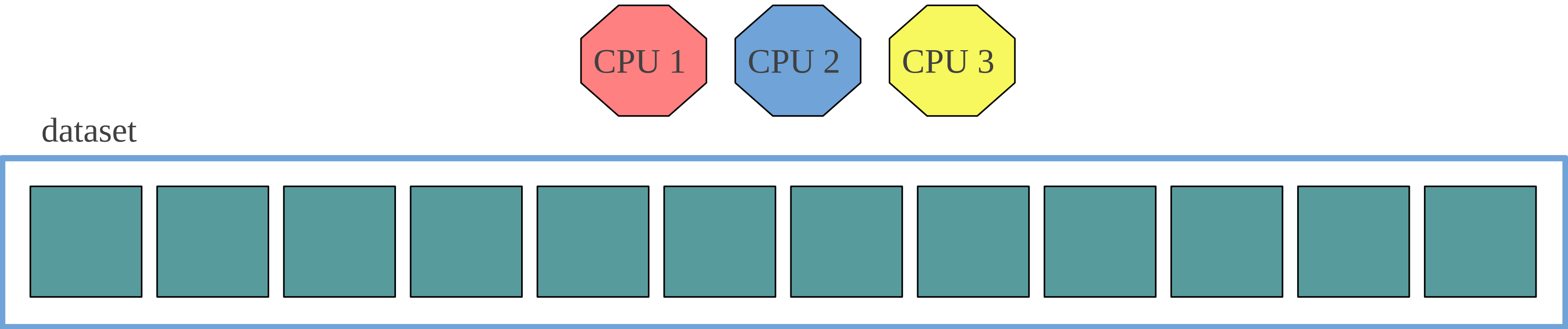




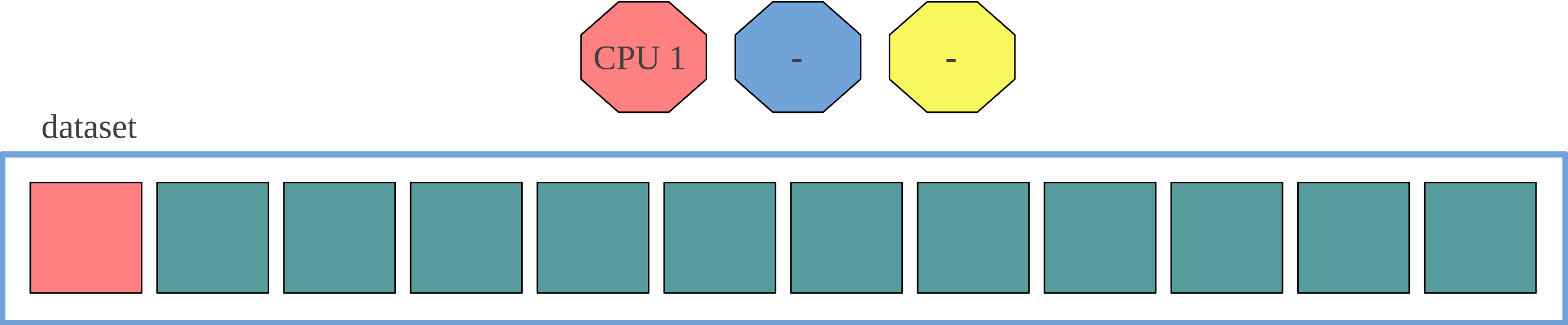
ANALYSIS OF GLOBAL TEMPERATURE

PART 1

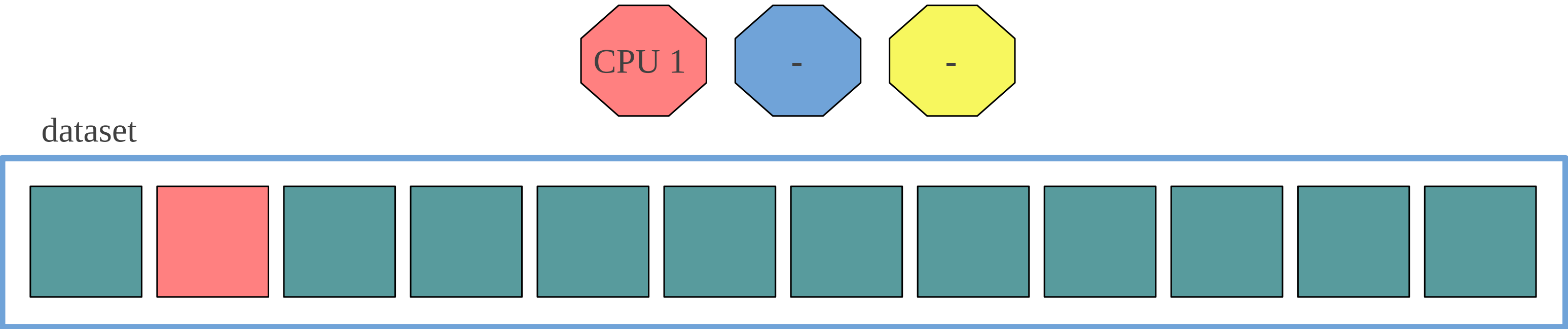
Data processing



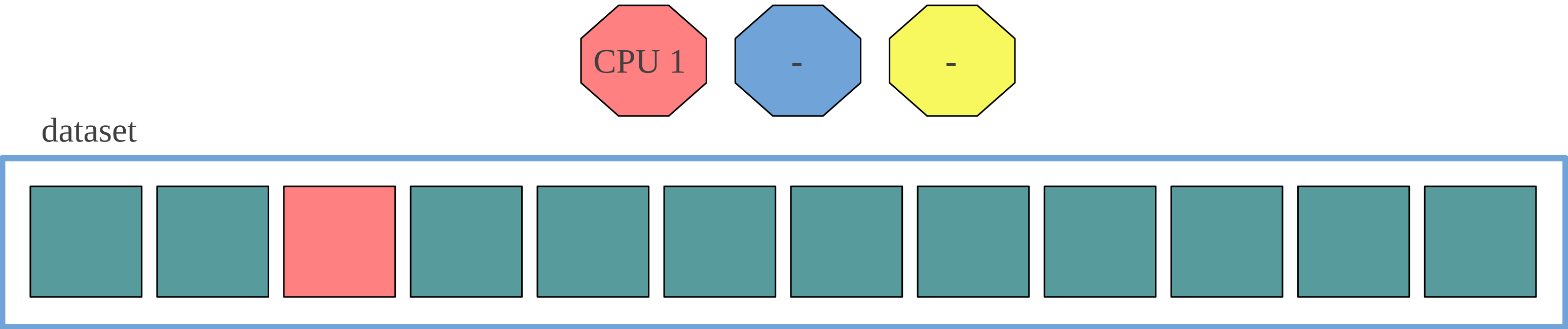
Sequential iteration



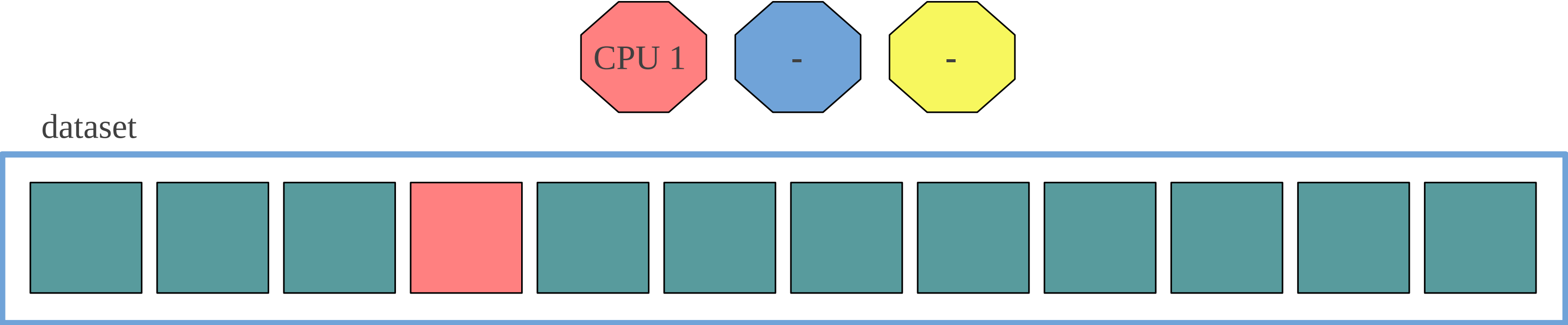
Sequential iteration



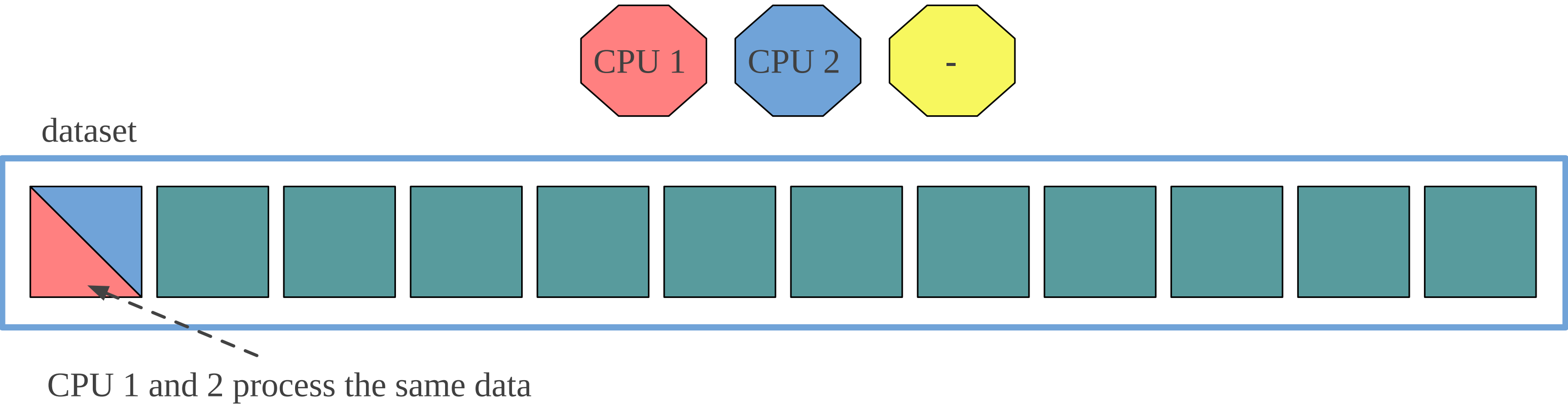
Sequential iteration



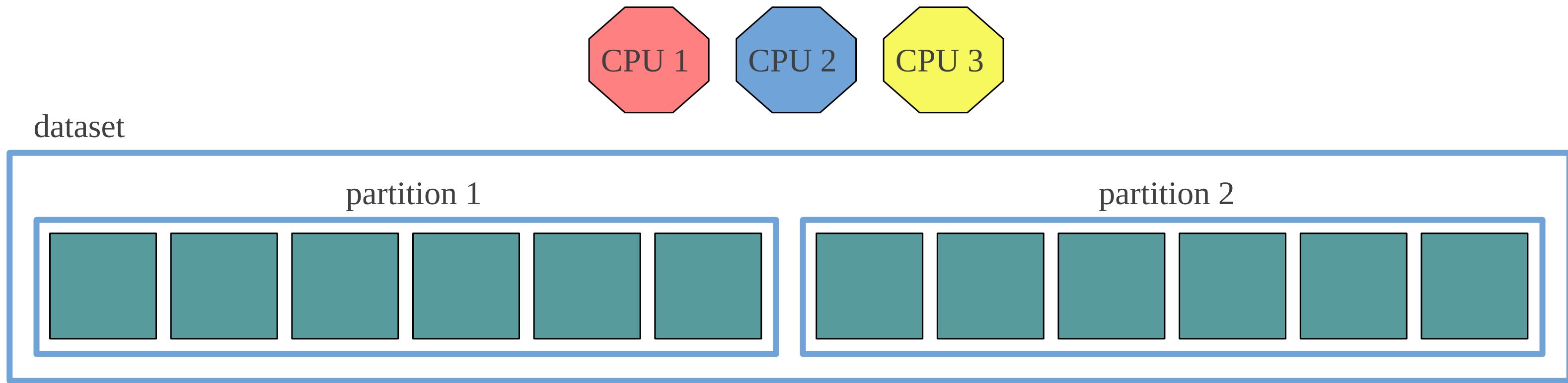
Sequential iteration



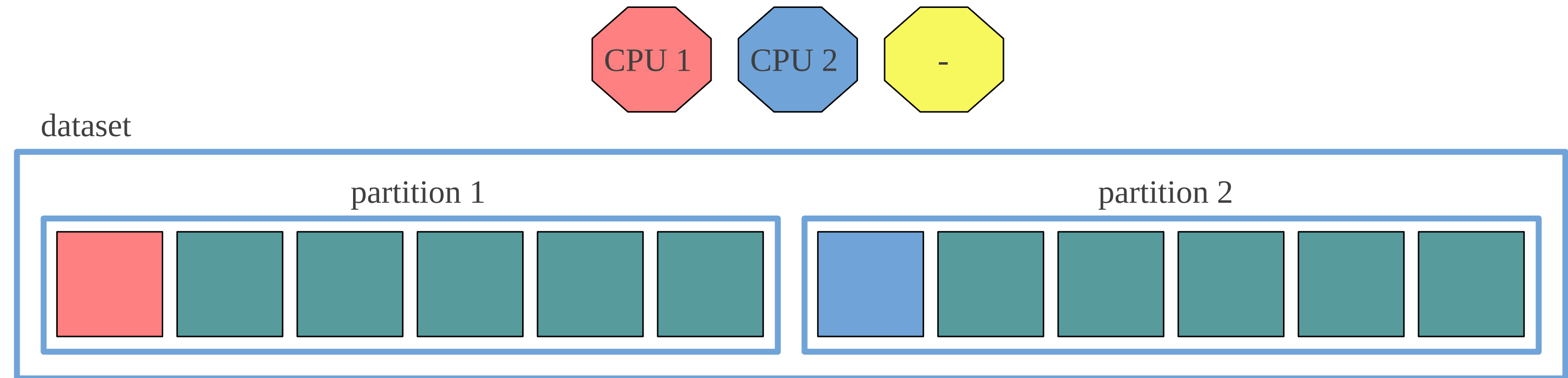
Useful work



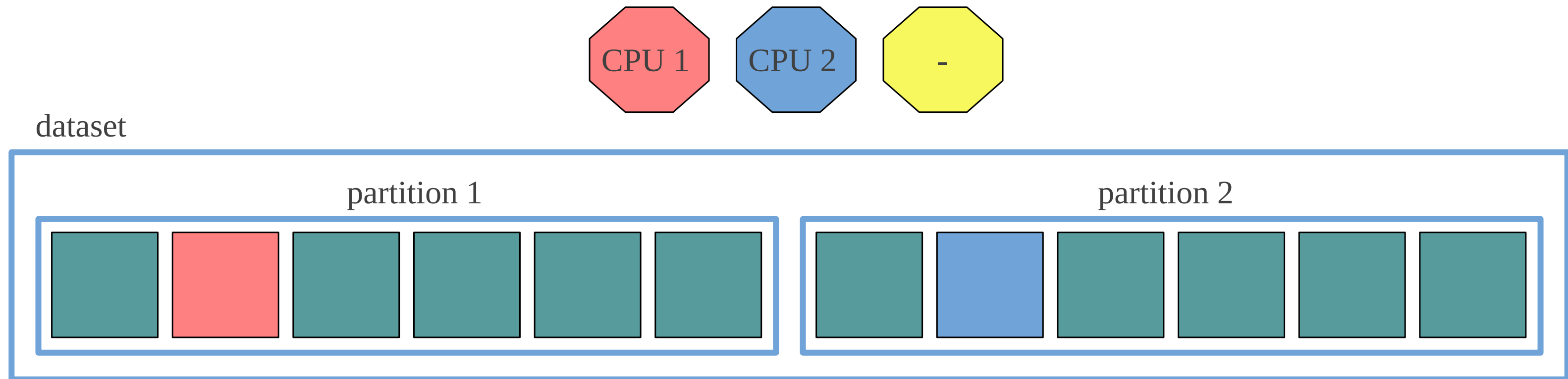
Partition the dataset



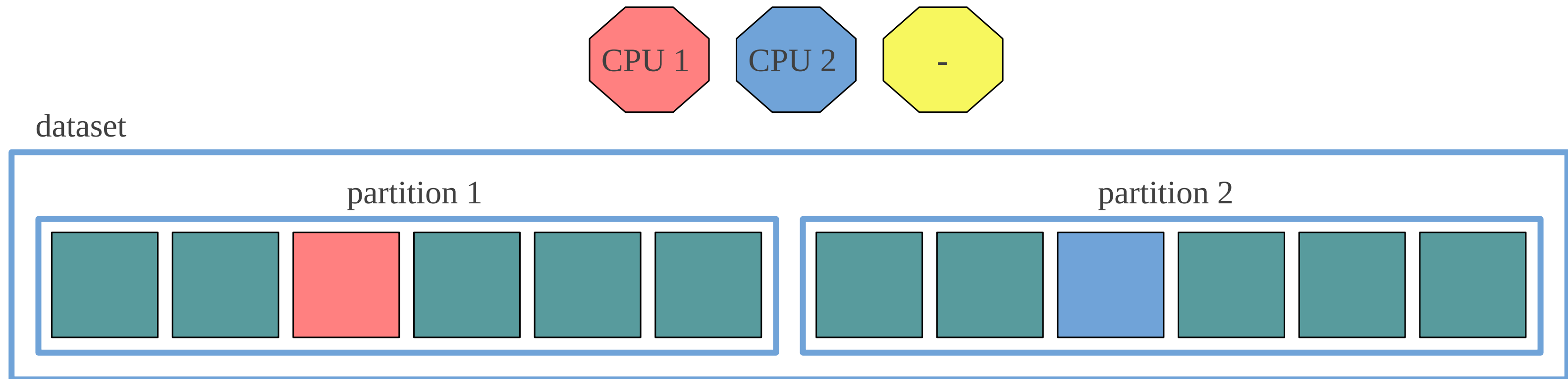
Partition the dataset



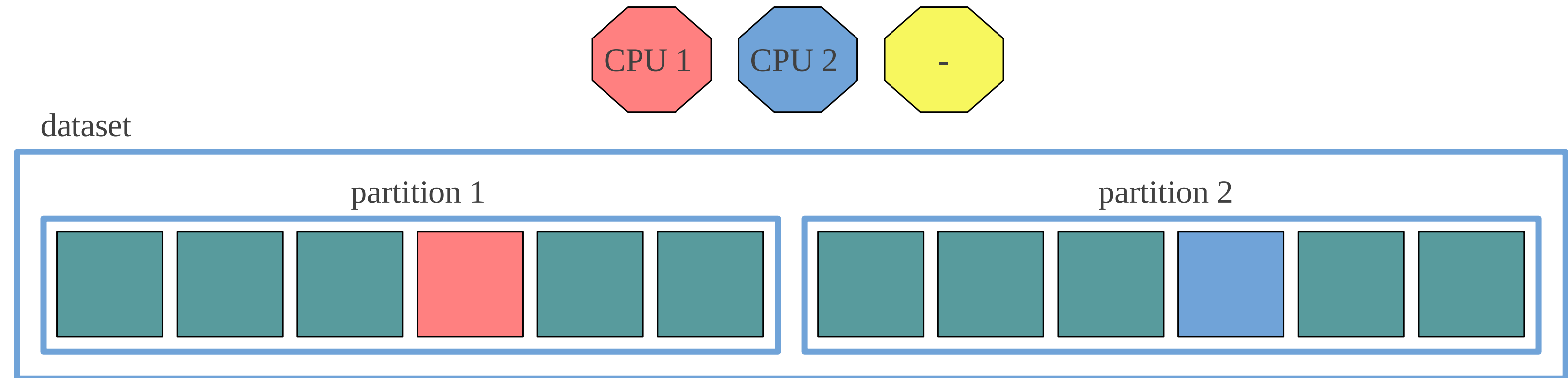
Partition the dataset



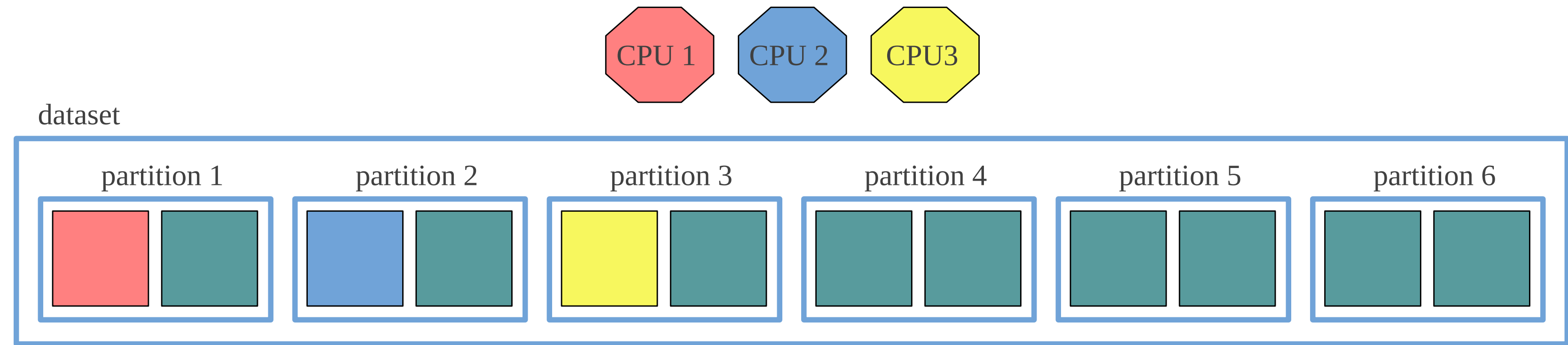
Partition the dataset



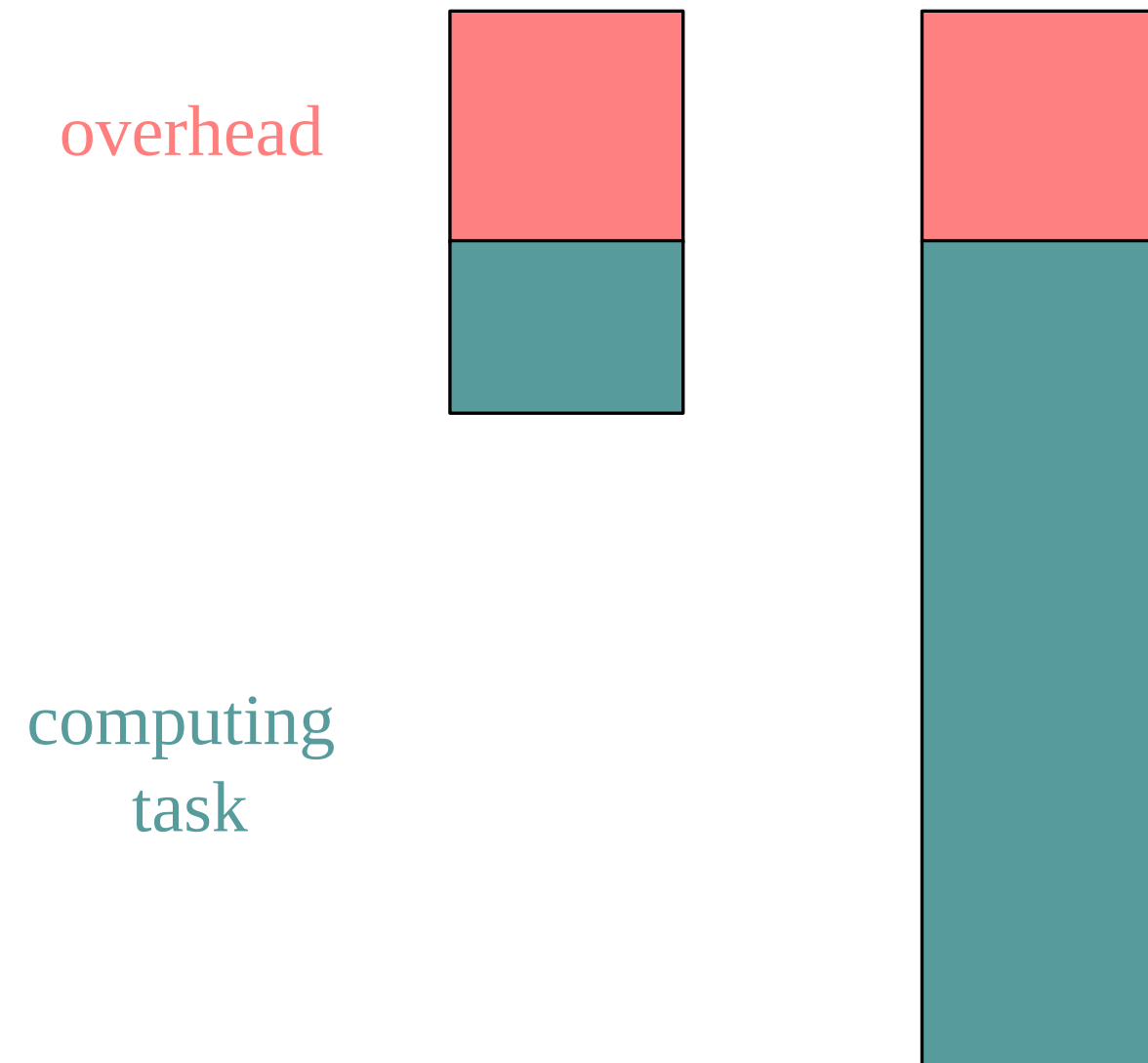
Partition the dataset



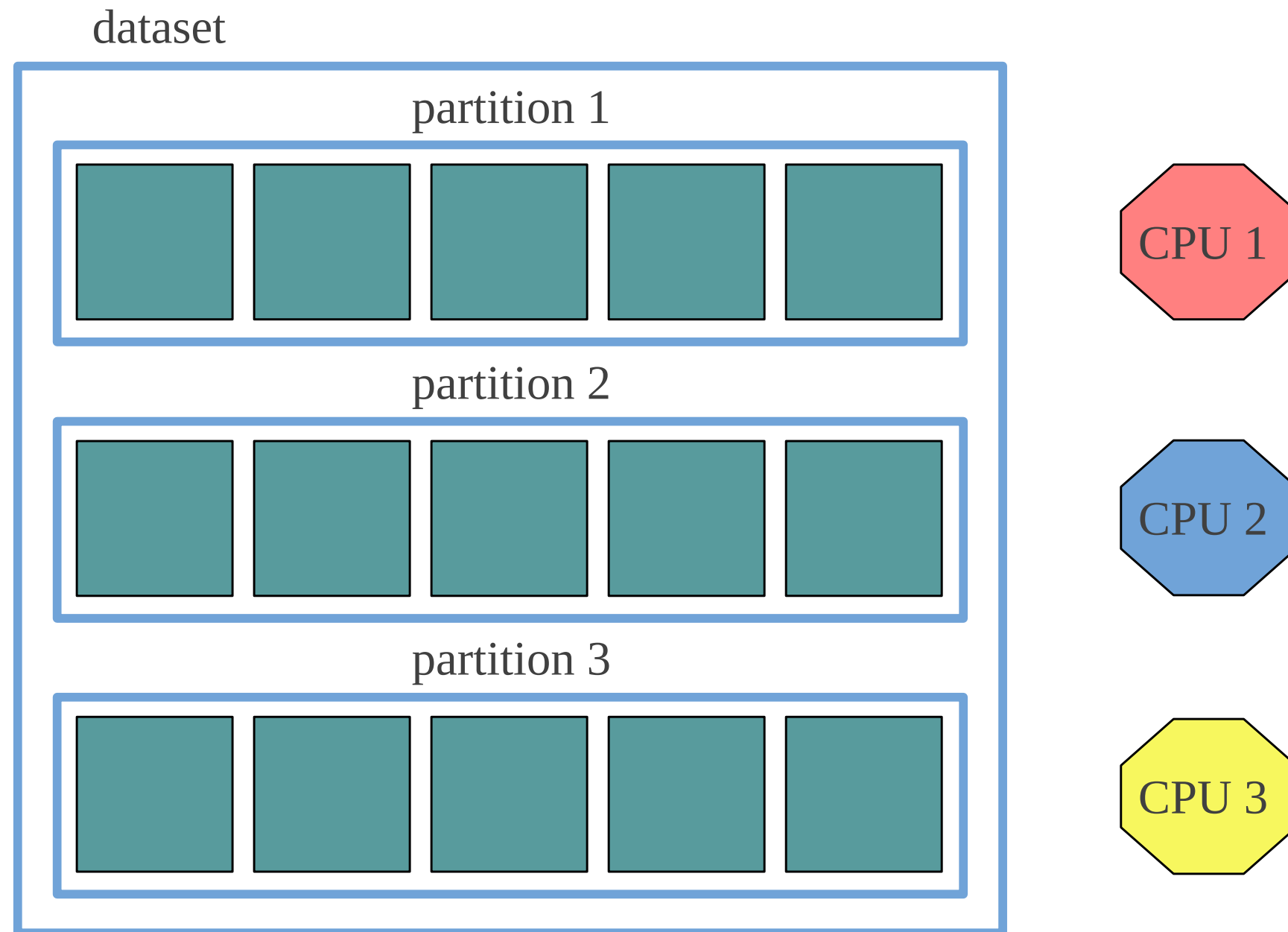
Partition the dataset



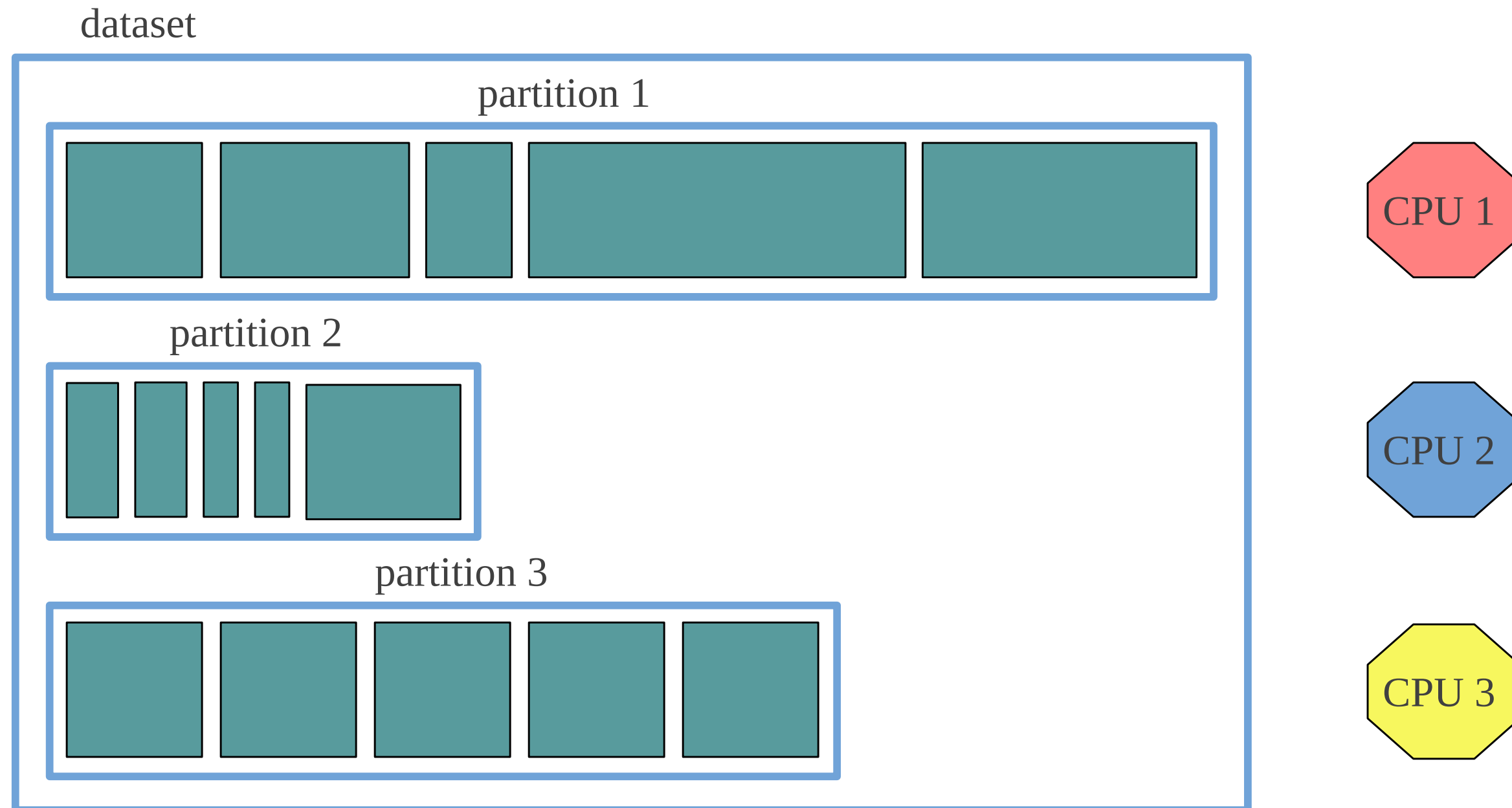
Size the partitions appropriately



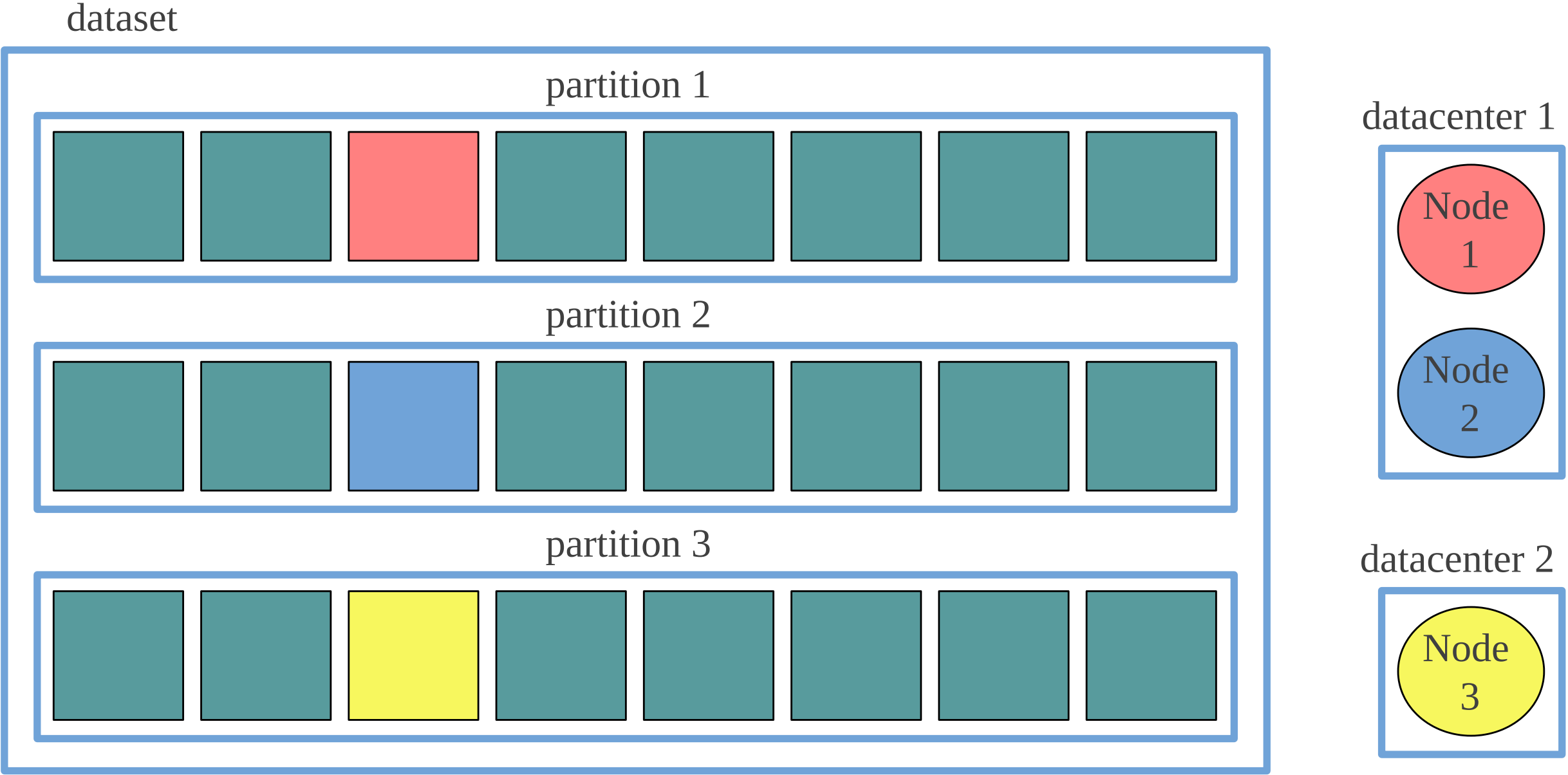
Tasks may require different computing power



Tasks may require different computing power



Scales to more than one computer



Important points

1. Size the partitions appropriately
2. Benchmark and tweak configuration for the task at hand

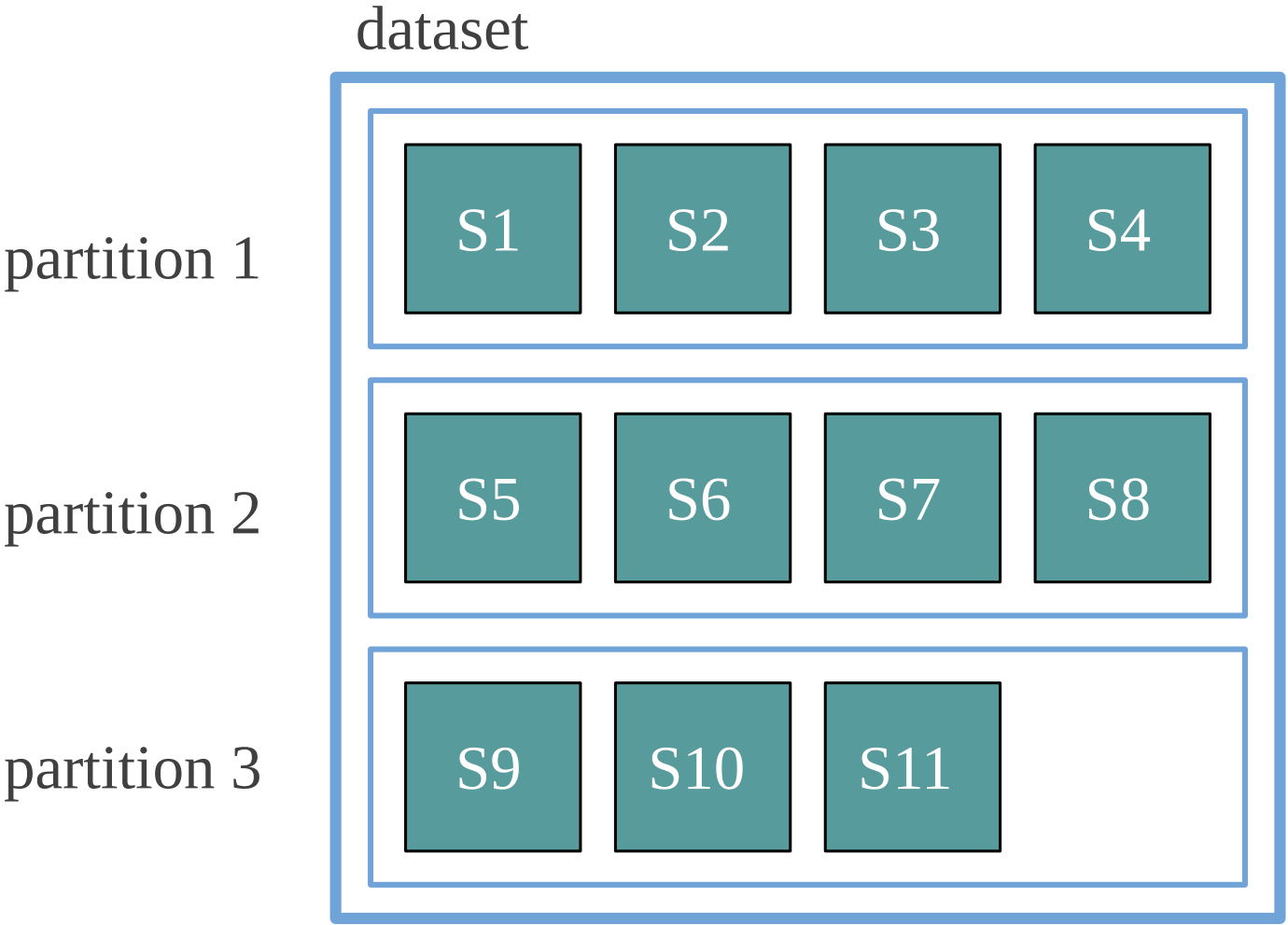
Important points

1. Size the partitions appropriately
2. Benchmark and tweak configuration for the task at hand
3. Parallel process MUST produce the SAME result as the sequential one

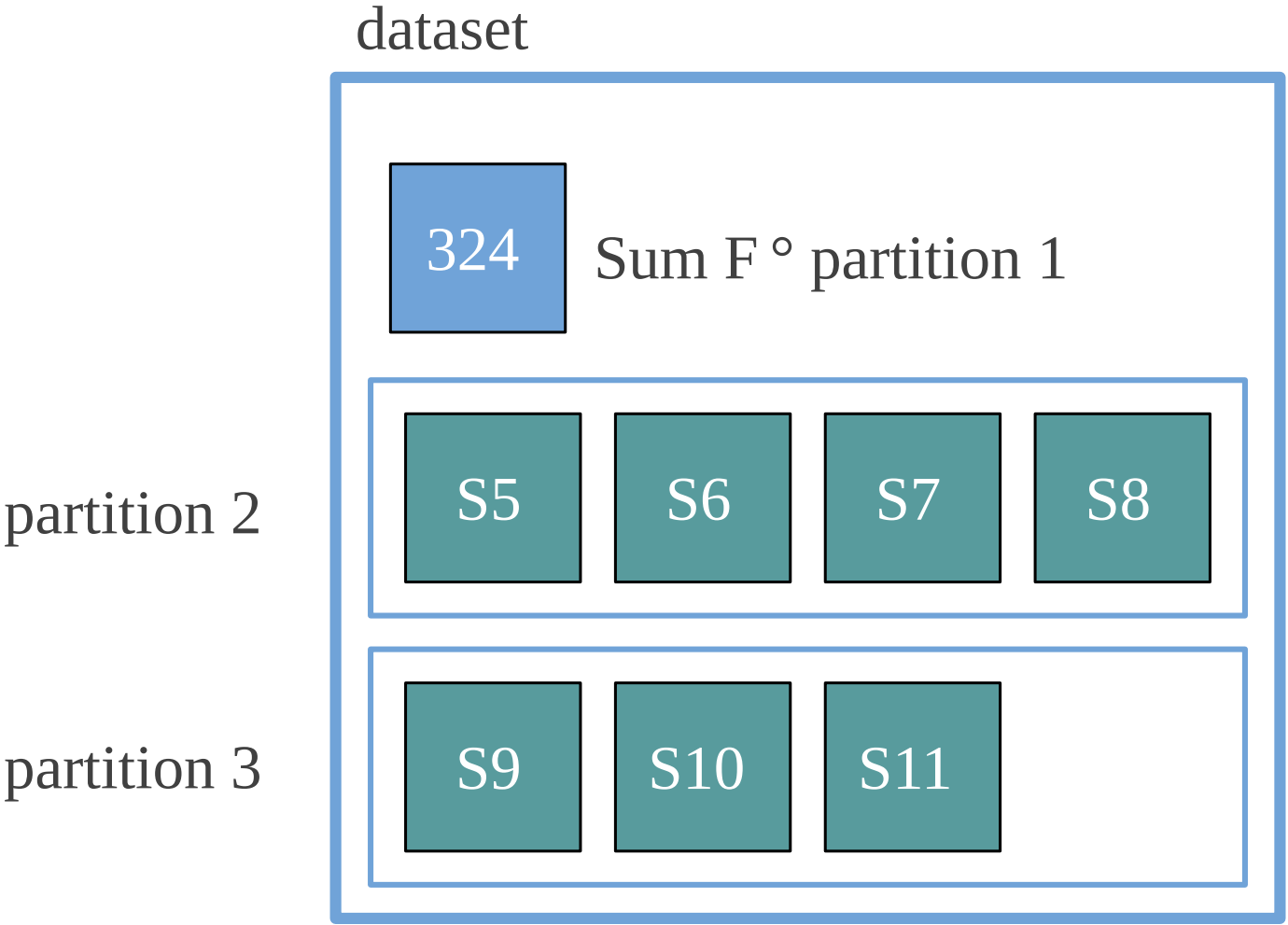
The background features a complex network of blue dots of varying sizes connected by thin, light blue lines. The dots are scattered across the frame, with some forming dense clusters and others standing alone. The lines create a web-like pattern that fills the entire background, giving it a technical or digital feel.

TemperatureNotebook.scala

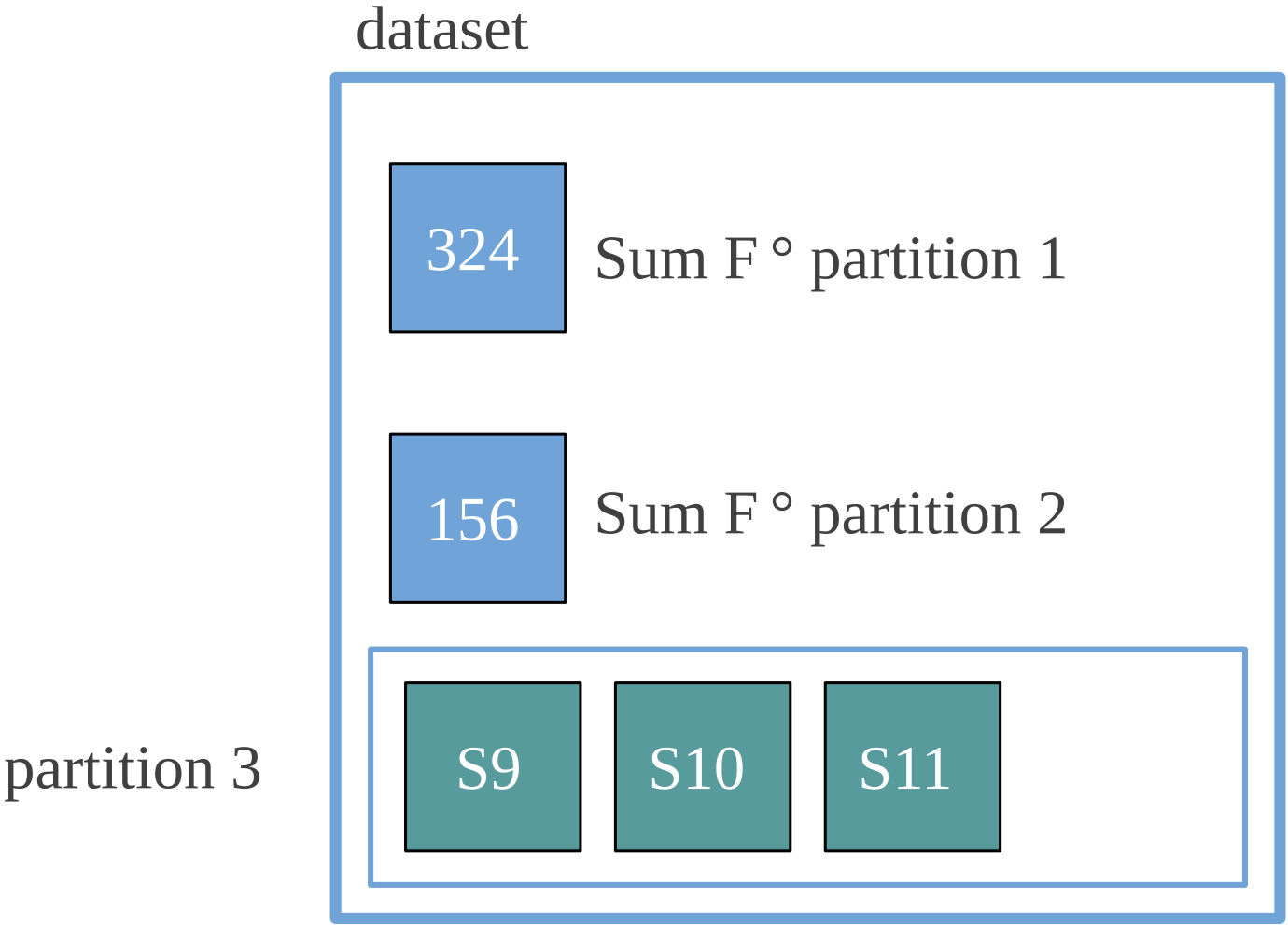
Average temperature



Sum temperatures per partition



Sum temperatures per partition



Sum temperatures per partition

dataset

324

Sum F ° partition 1

156

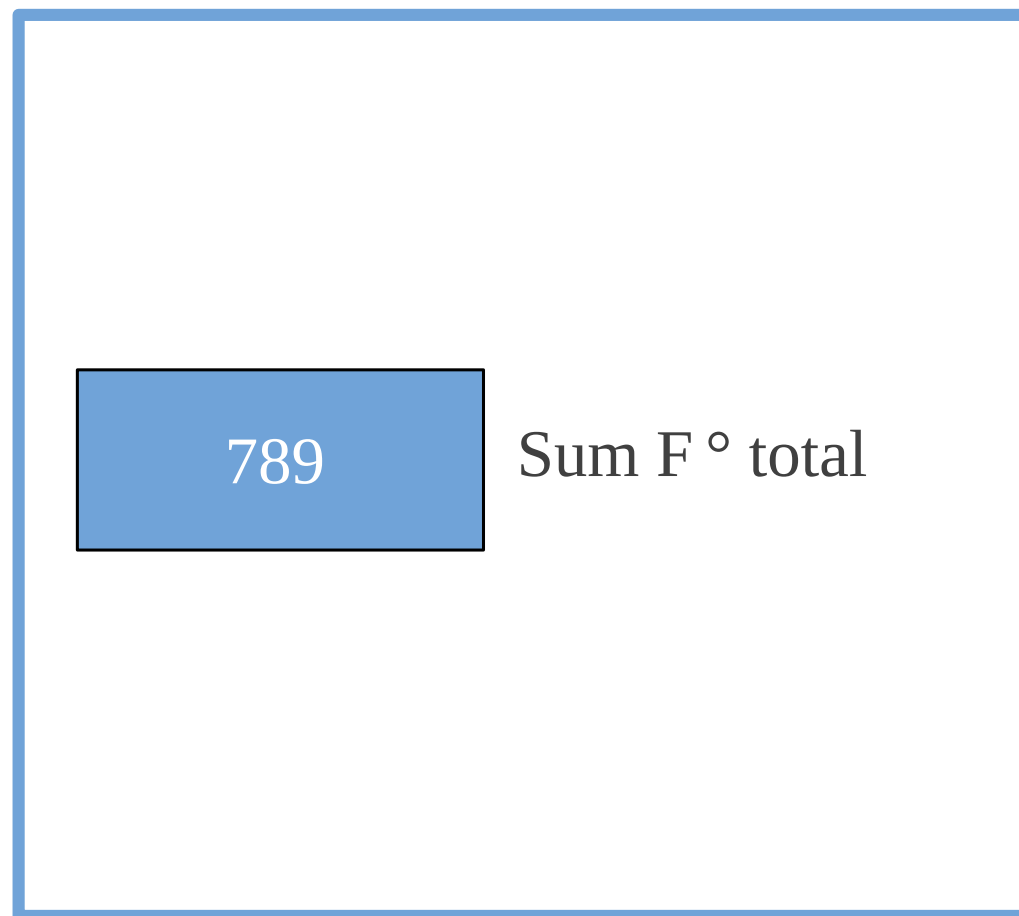
Sum F ° partition 2

209

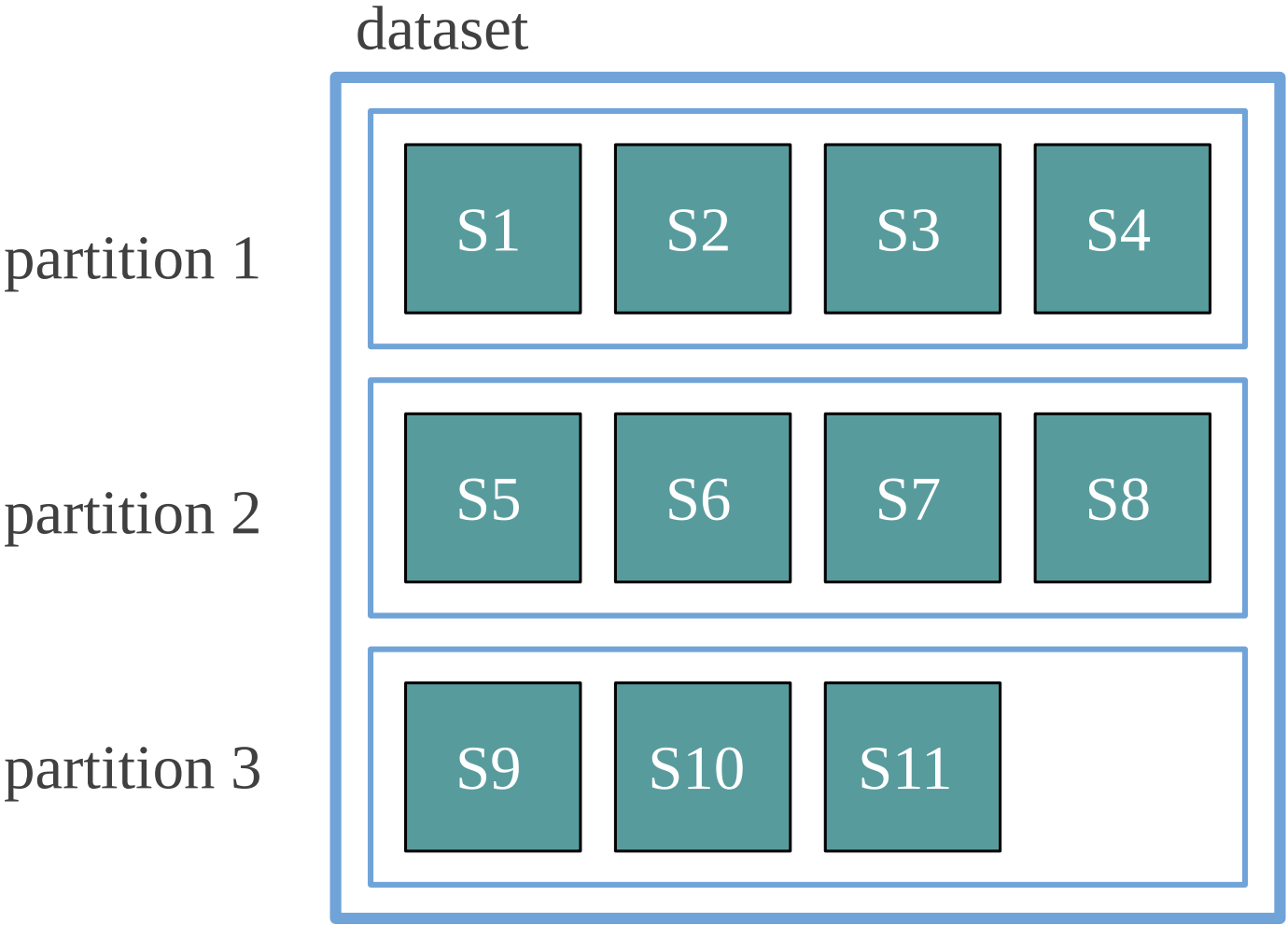
Sum F ° partition 3

Sum-up all partitions

dataset



Size



Calculate size per partition

dataset

4

Size partition 1

4

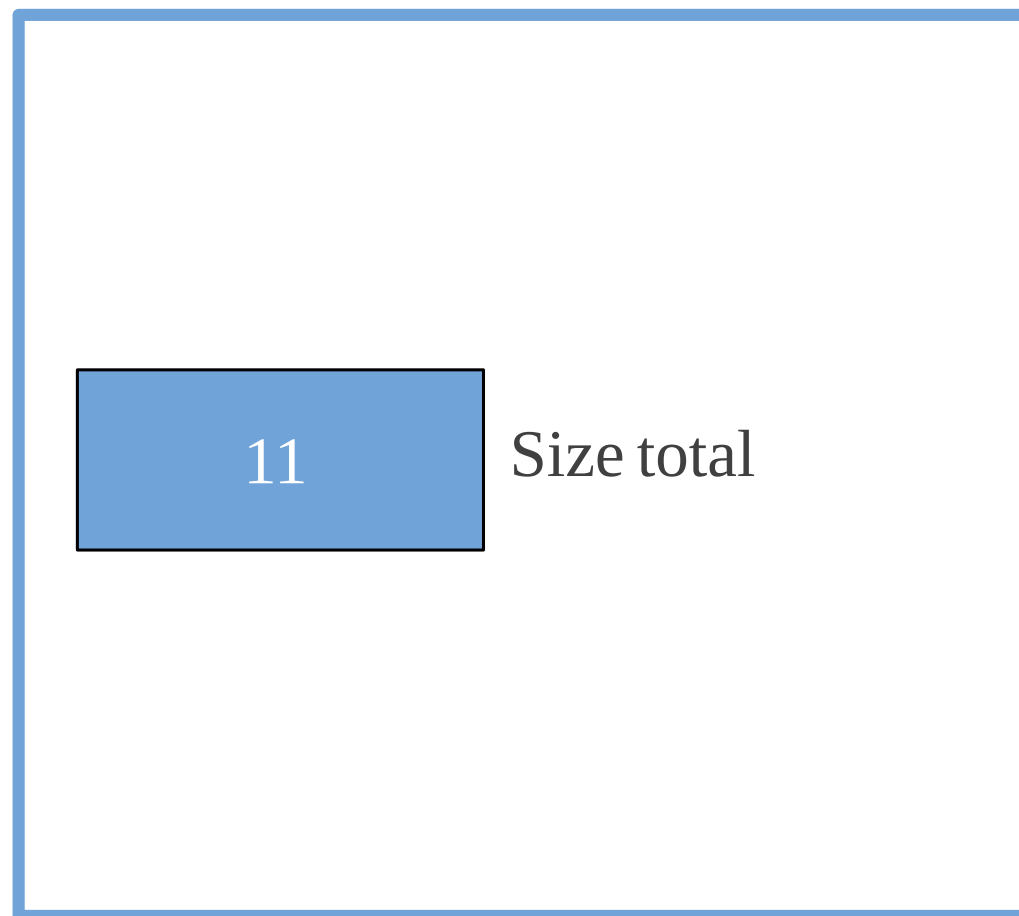
Size partition 2

3

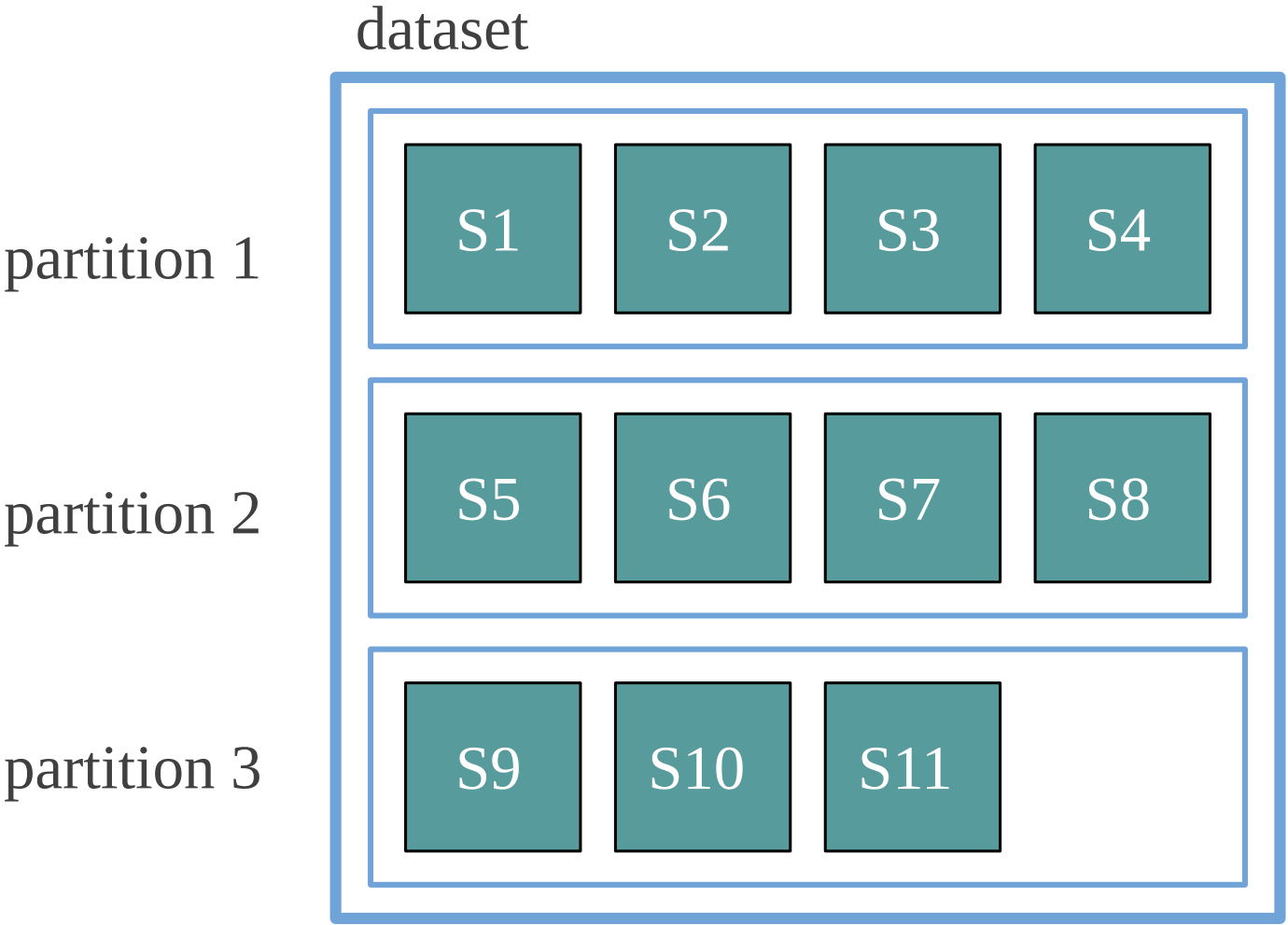
Size partition 3

Sum-up all partitions

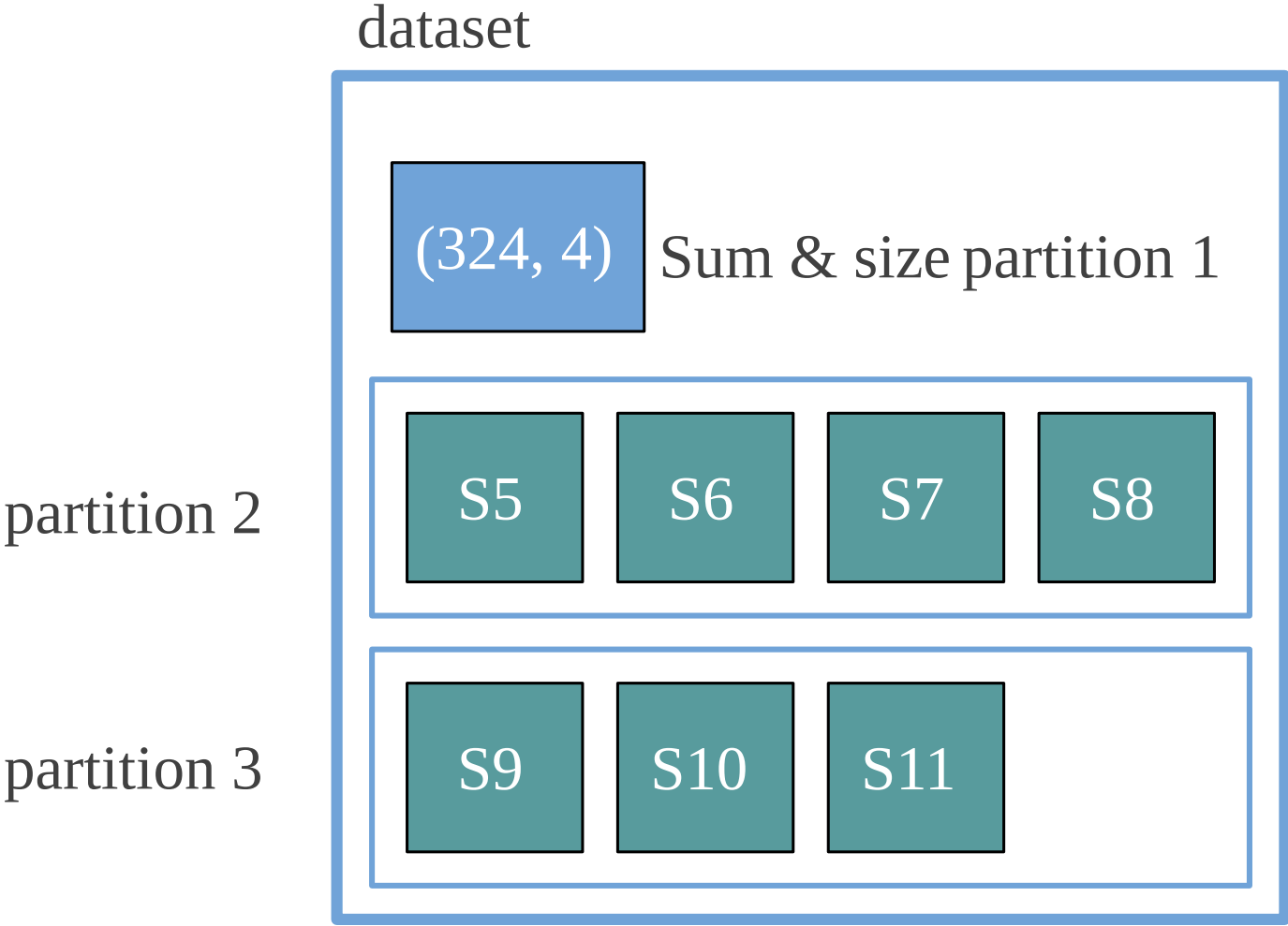
dataset



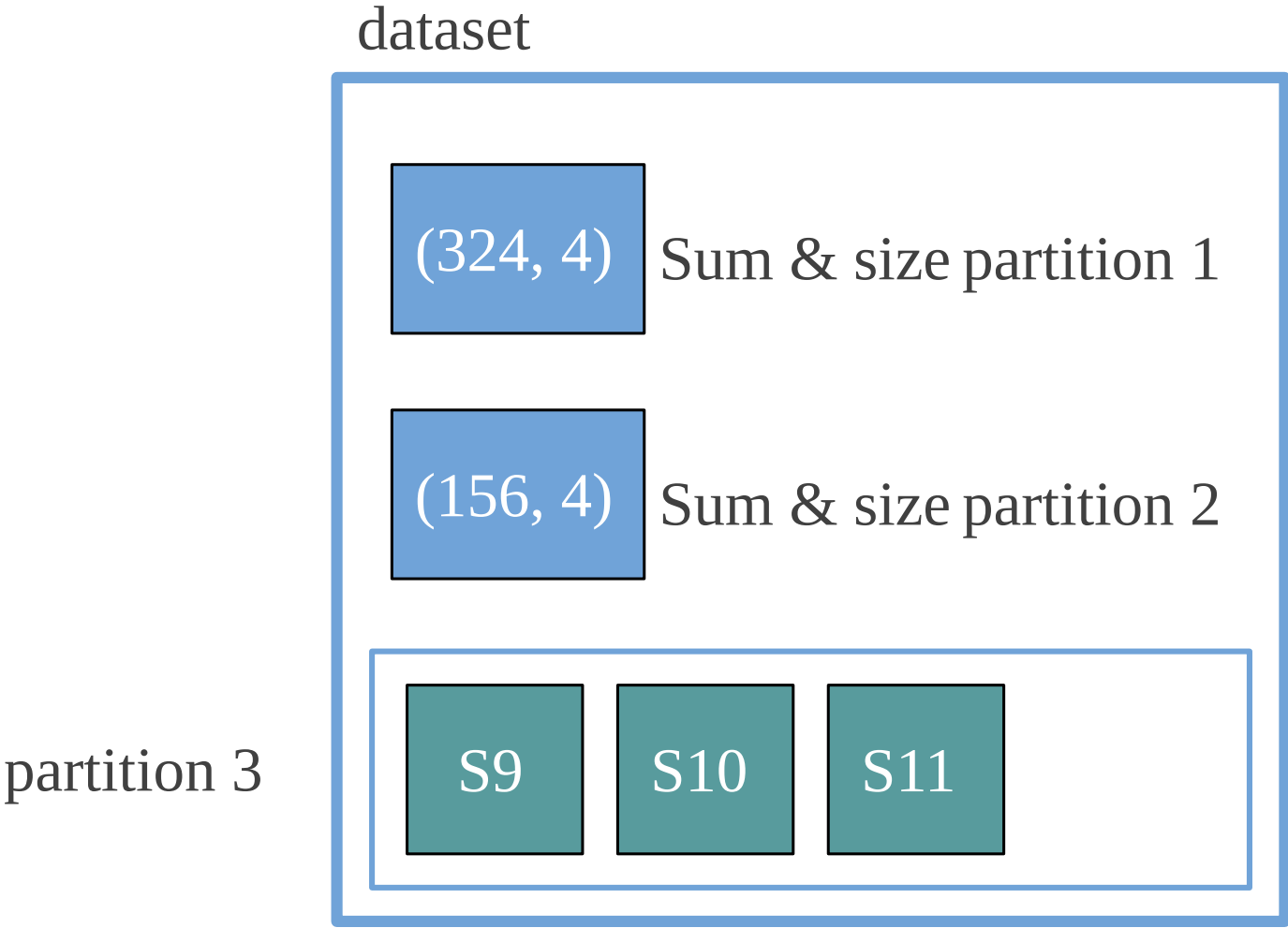
Average temperature in one pass



Calculate sum and size per partition



Calculate sum and size per partition



Calculate sum and size per partition

dataset

(324, 4) Sum & size partition 1

(156, 4) Sum & size partition 2

(209, 3) Sum & size partition 3

Sum-up all partitions

dataset

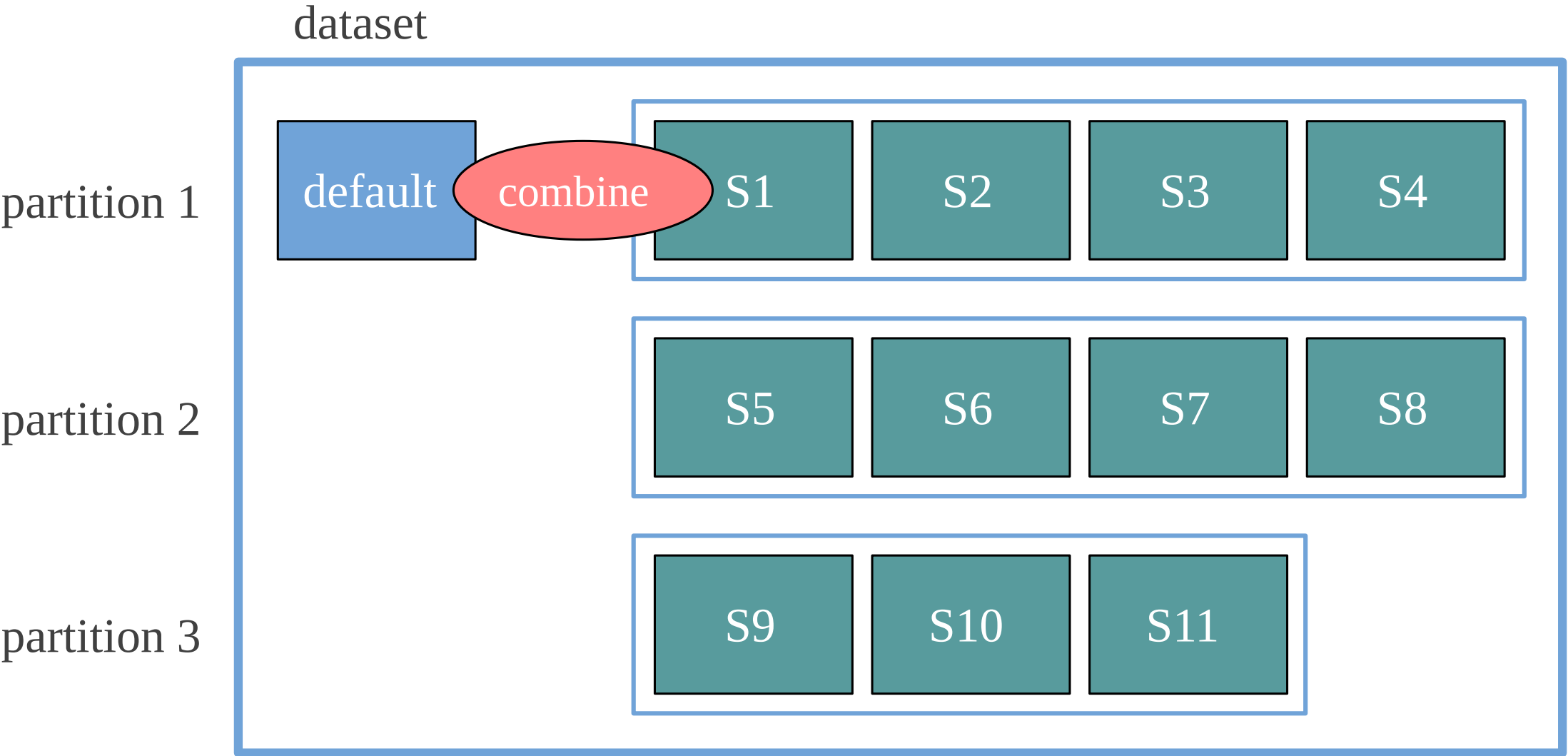
(789, 11)

Sum & size total

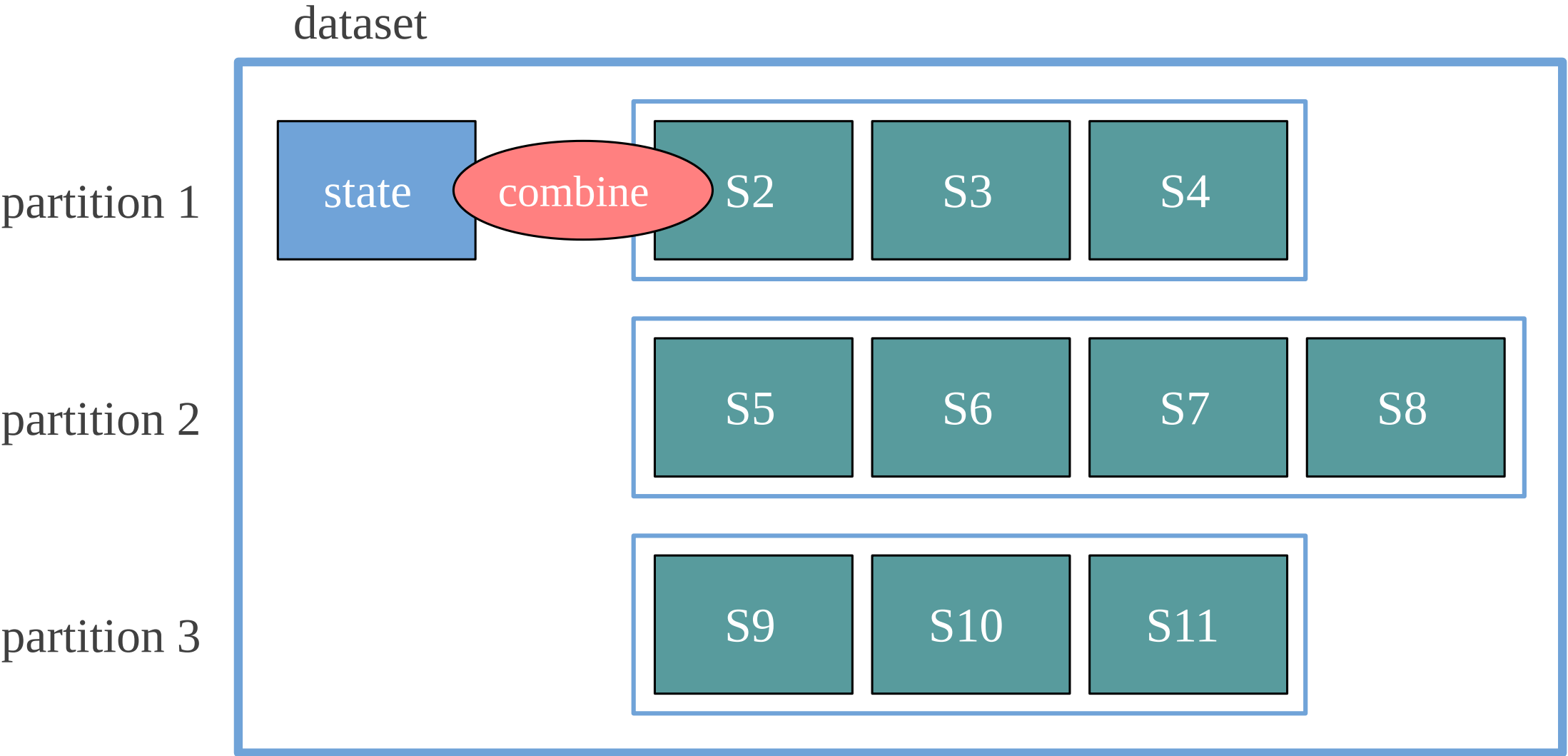
Fold partitions



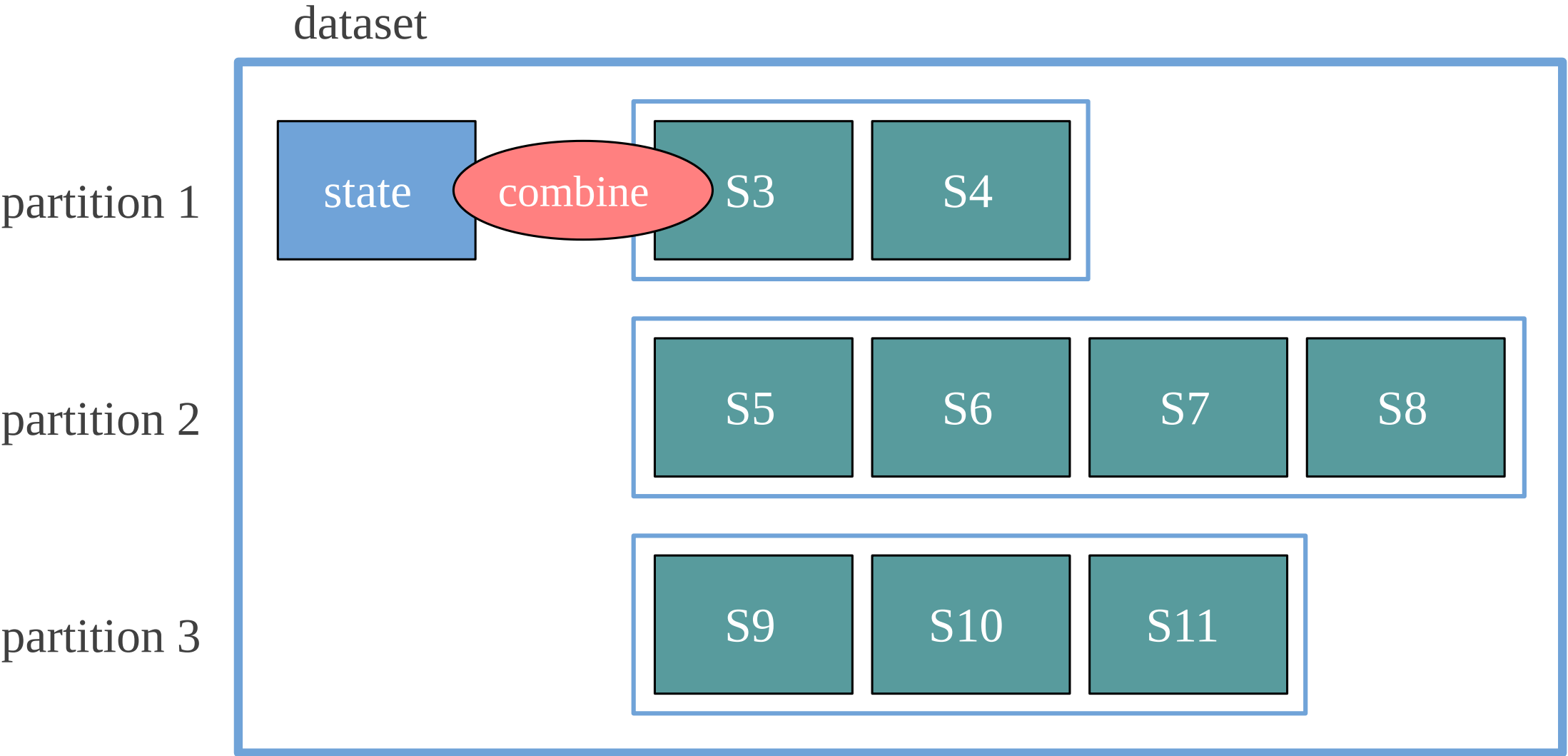
Fold partitions



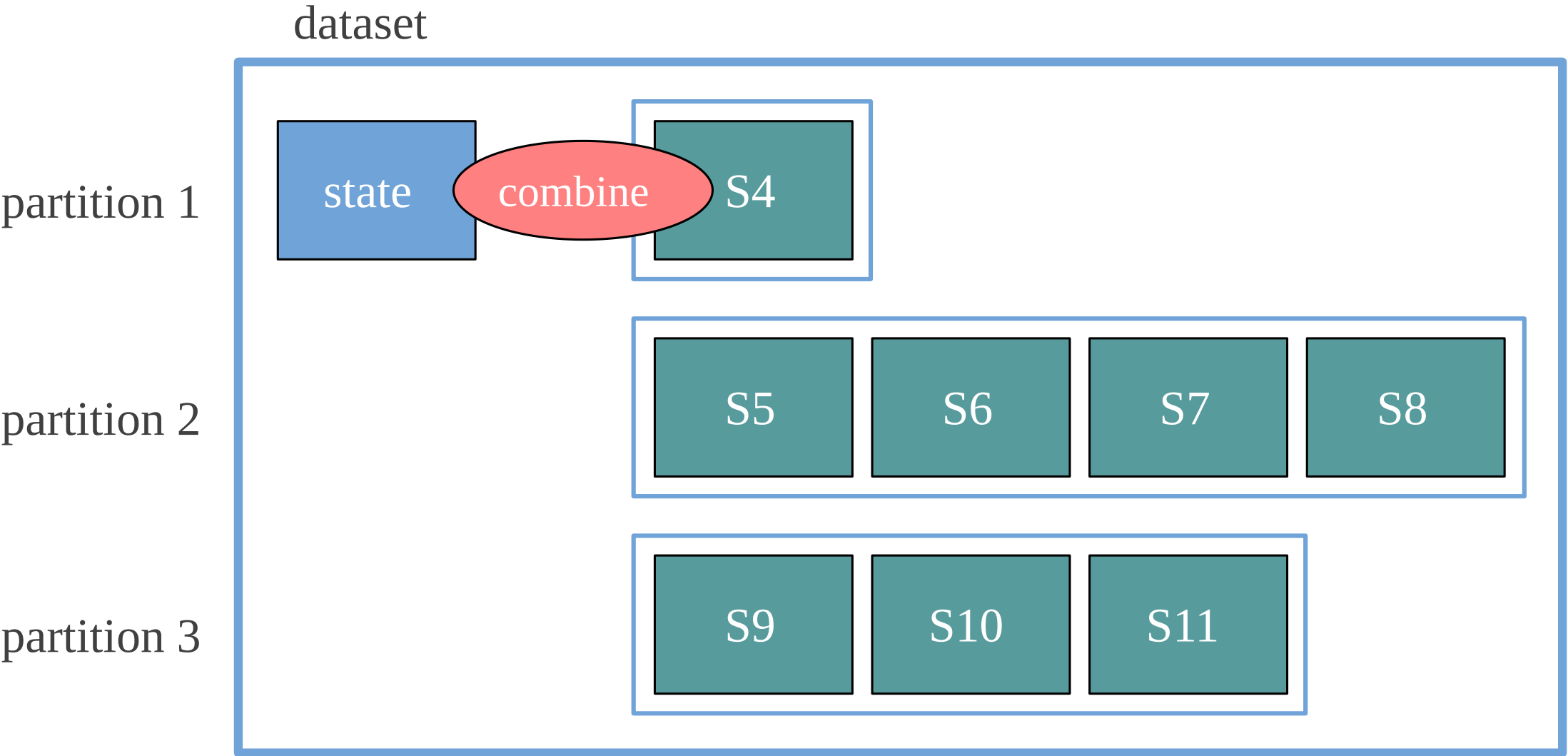
Fold partitions



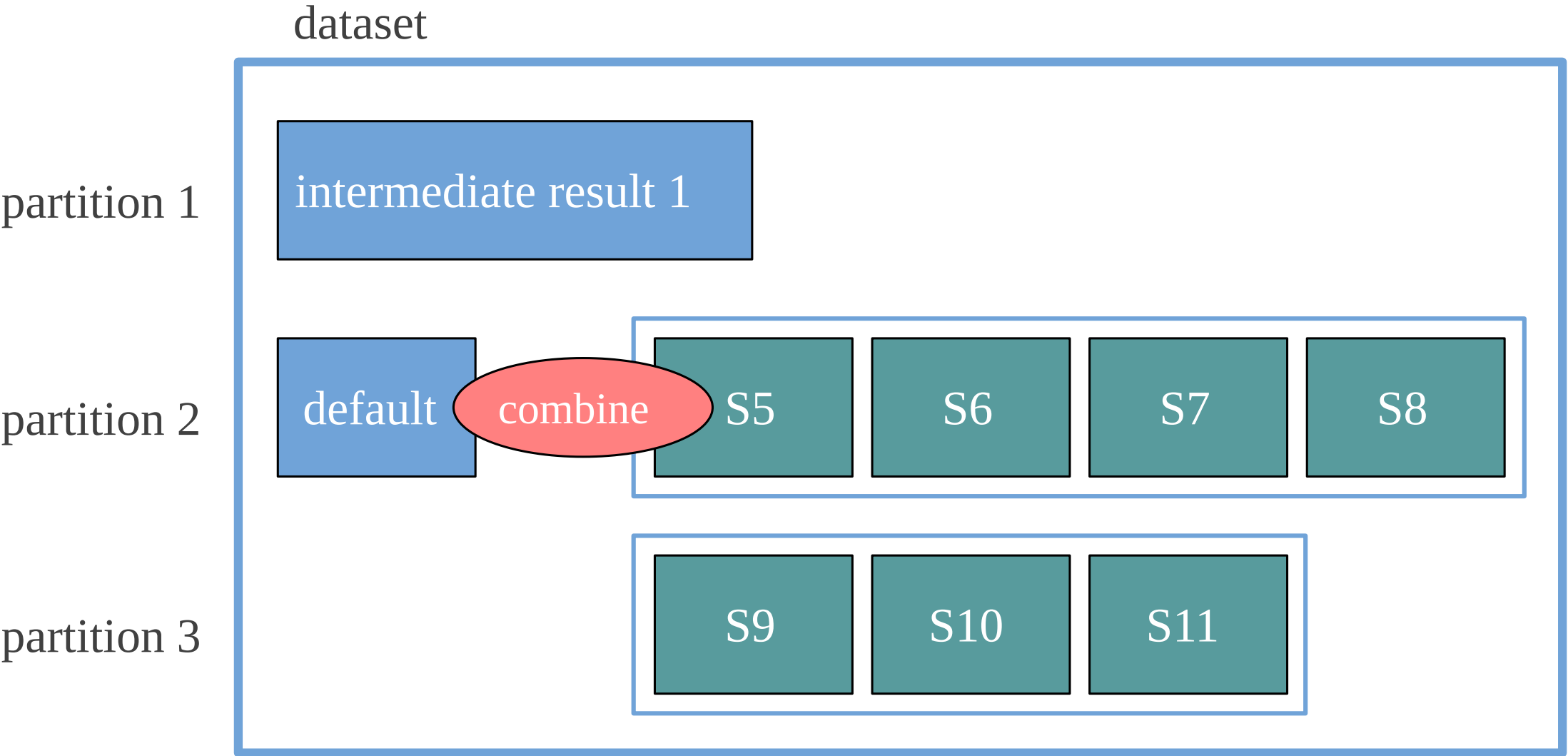
Fold partitions



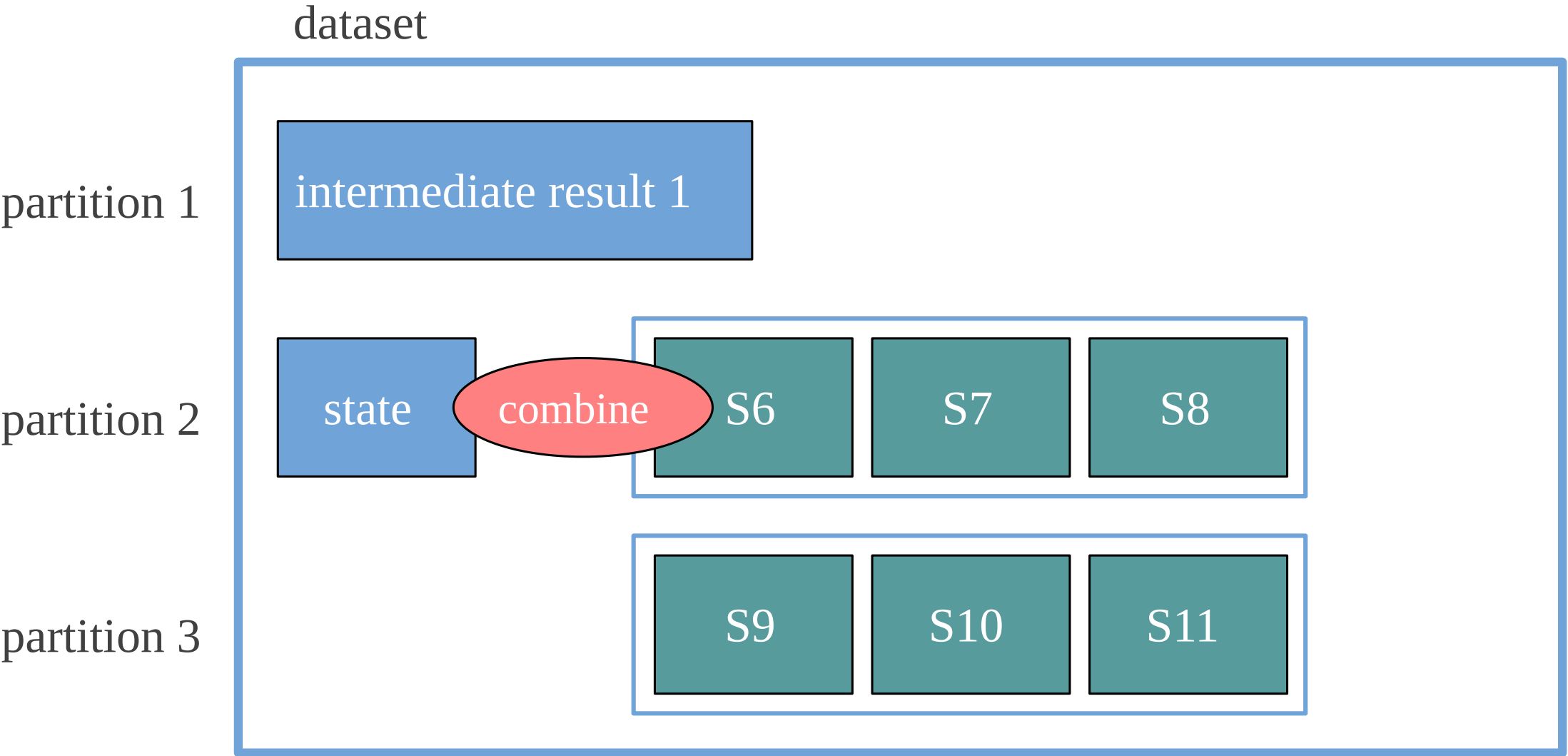
Fold partitions



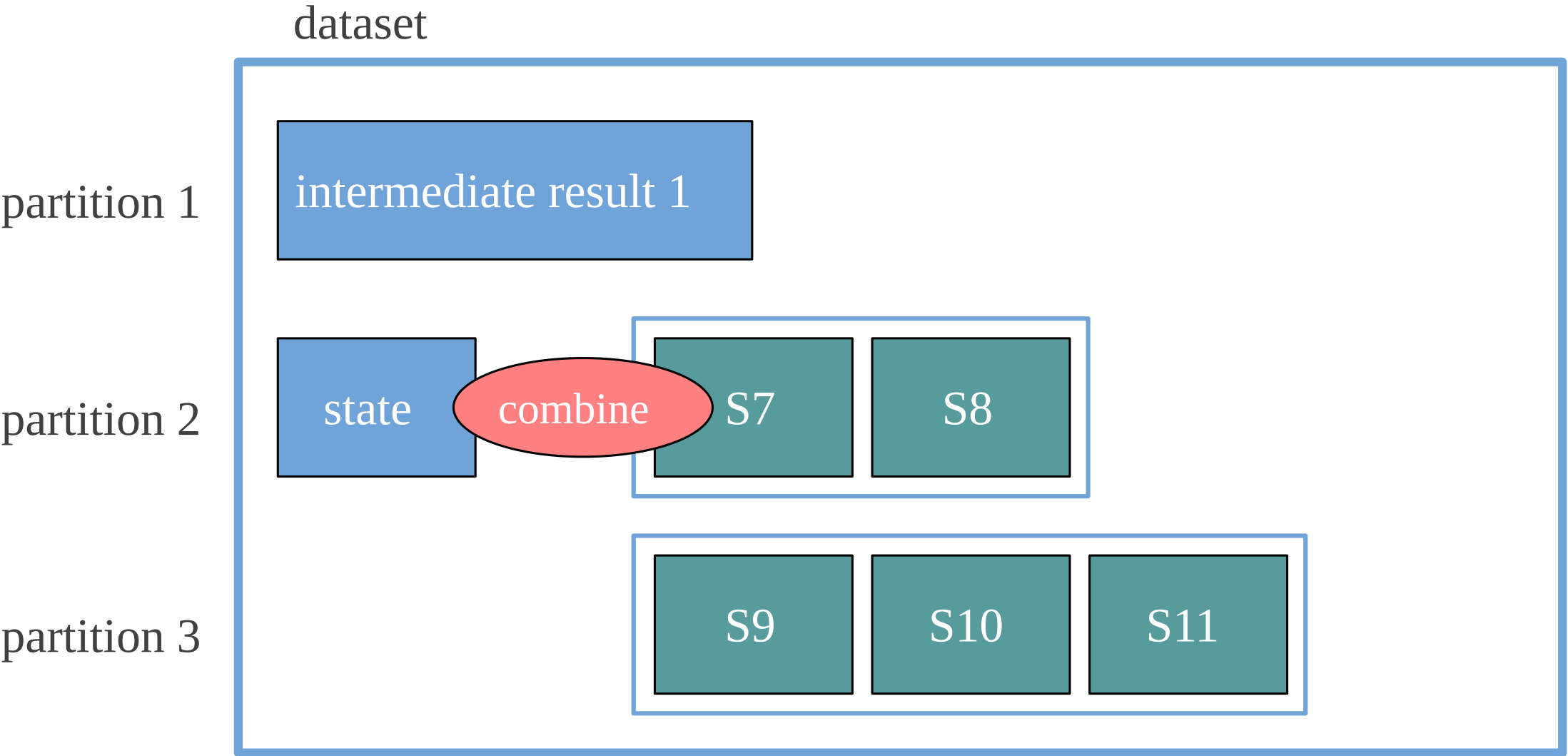
Fold partitions



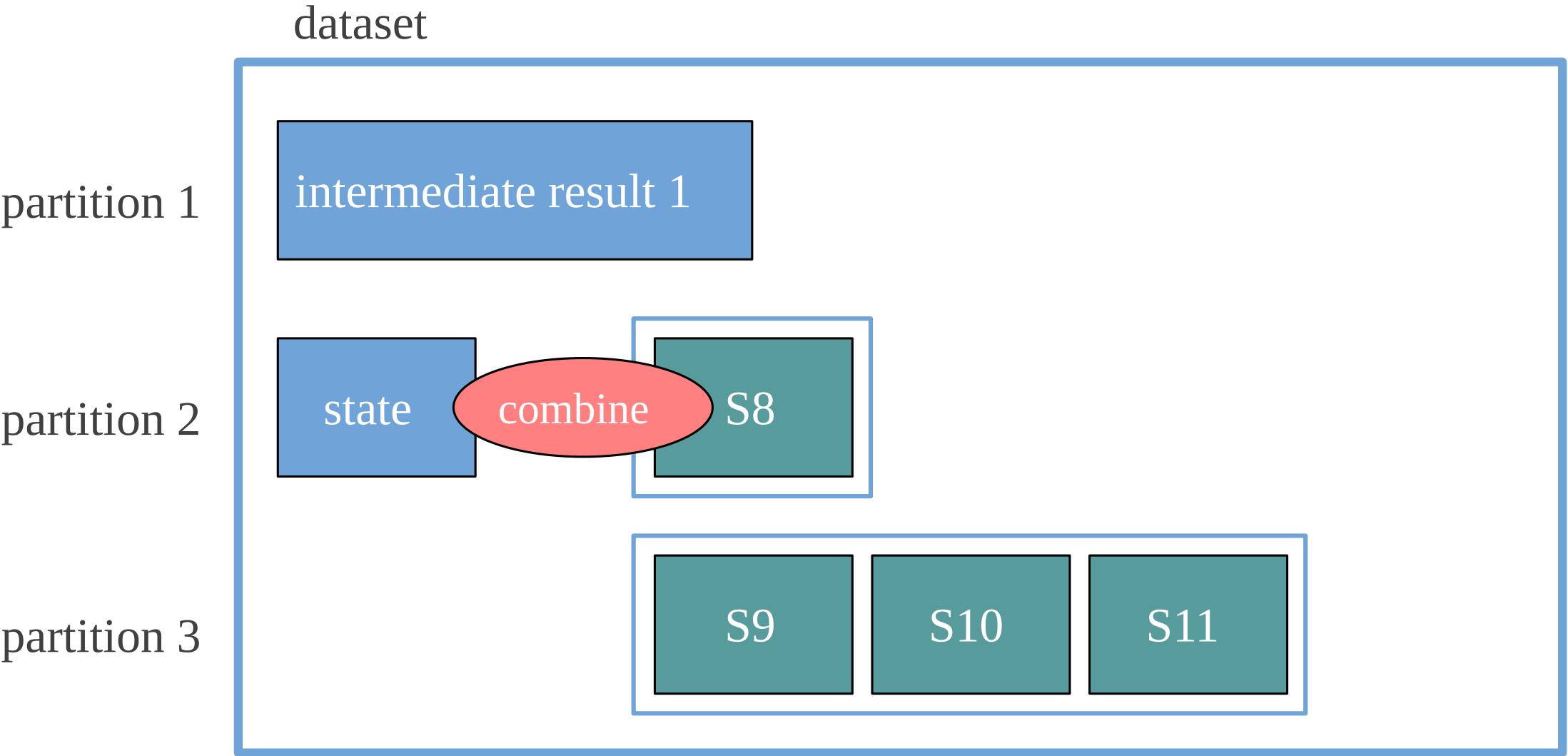
Fold partitions



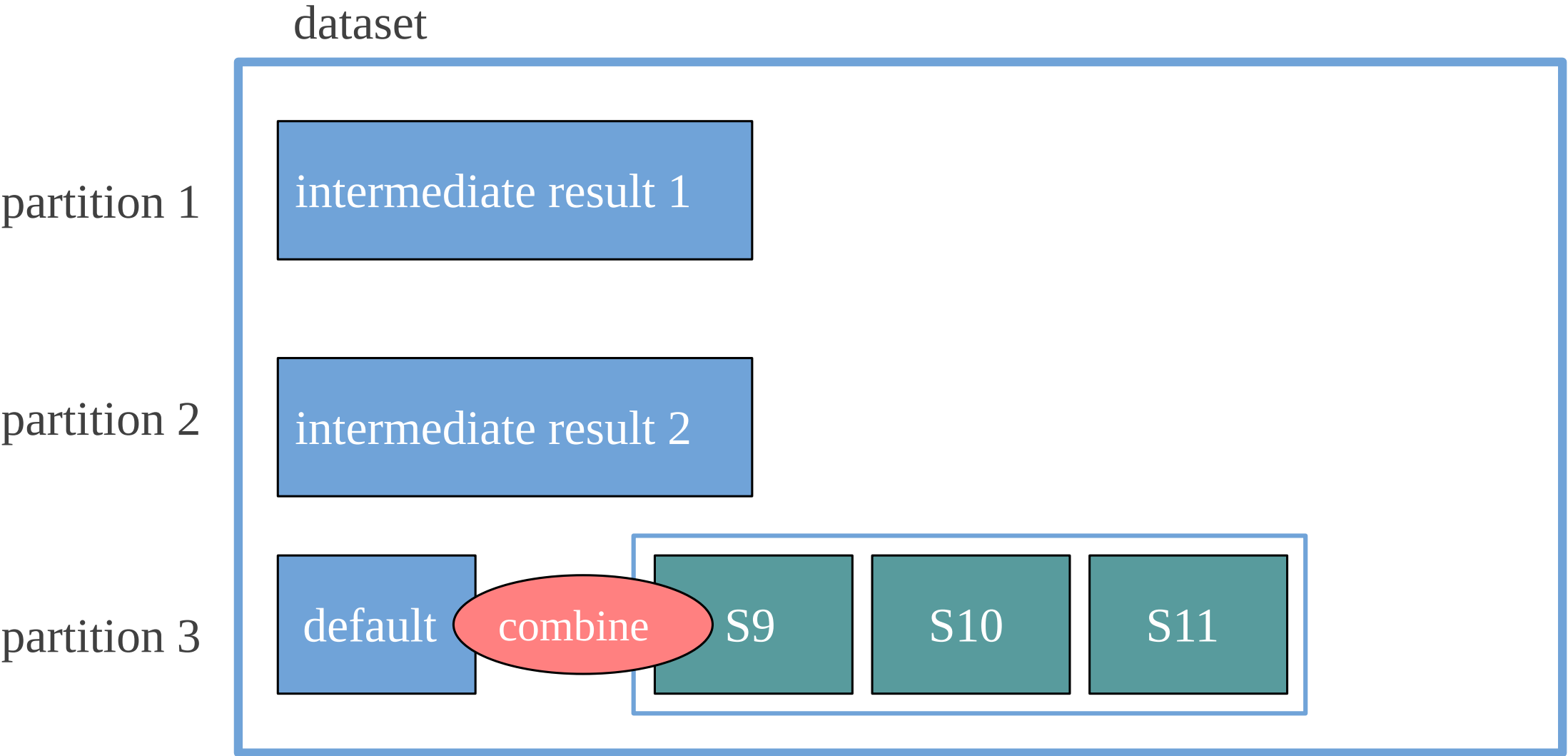
Fold partitions



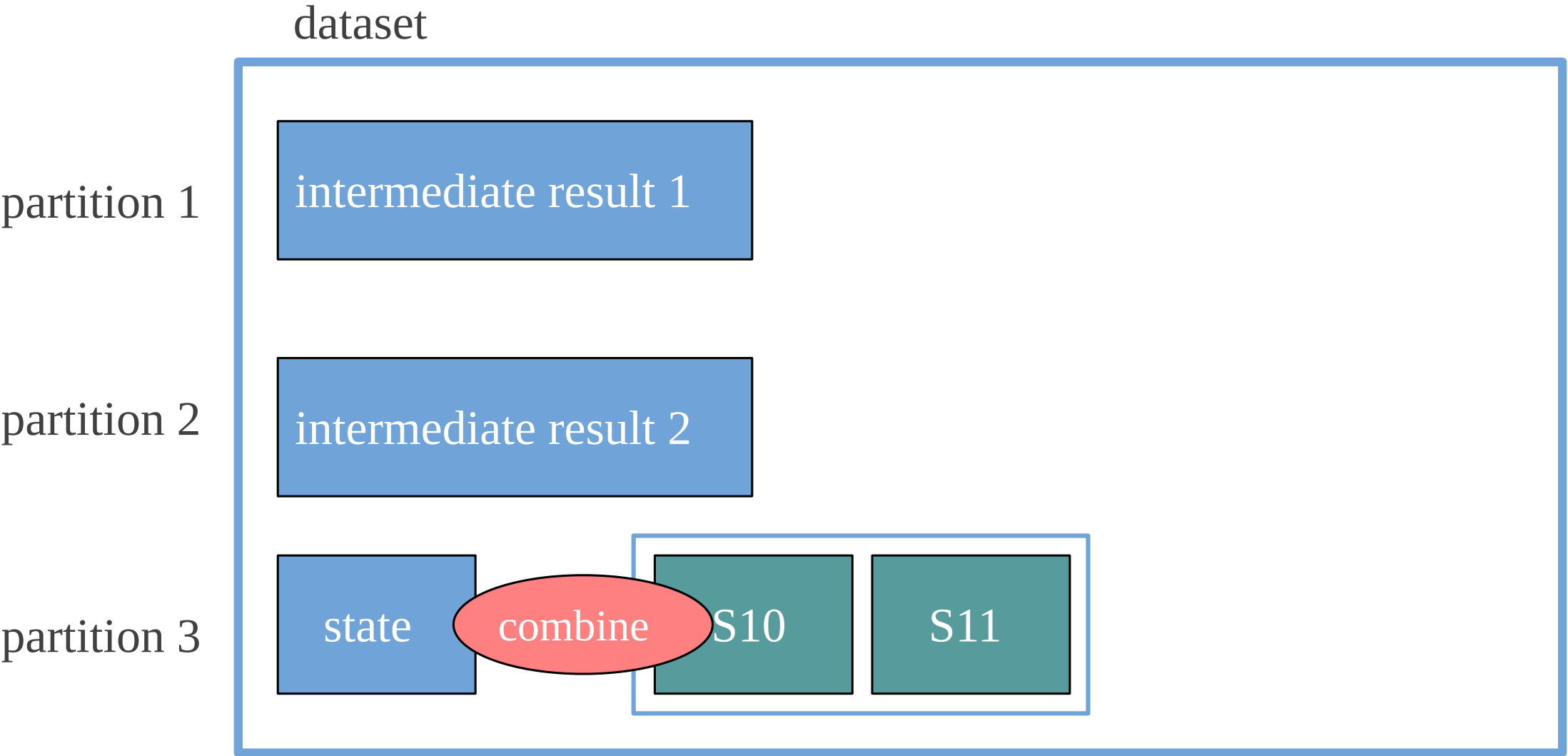
Fold partitions



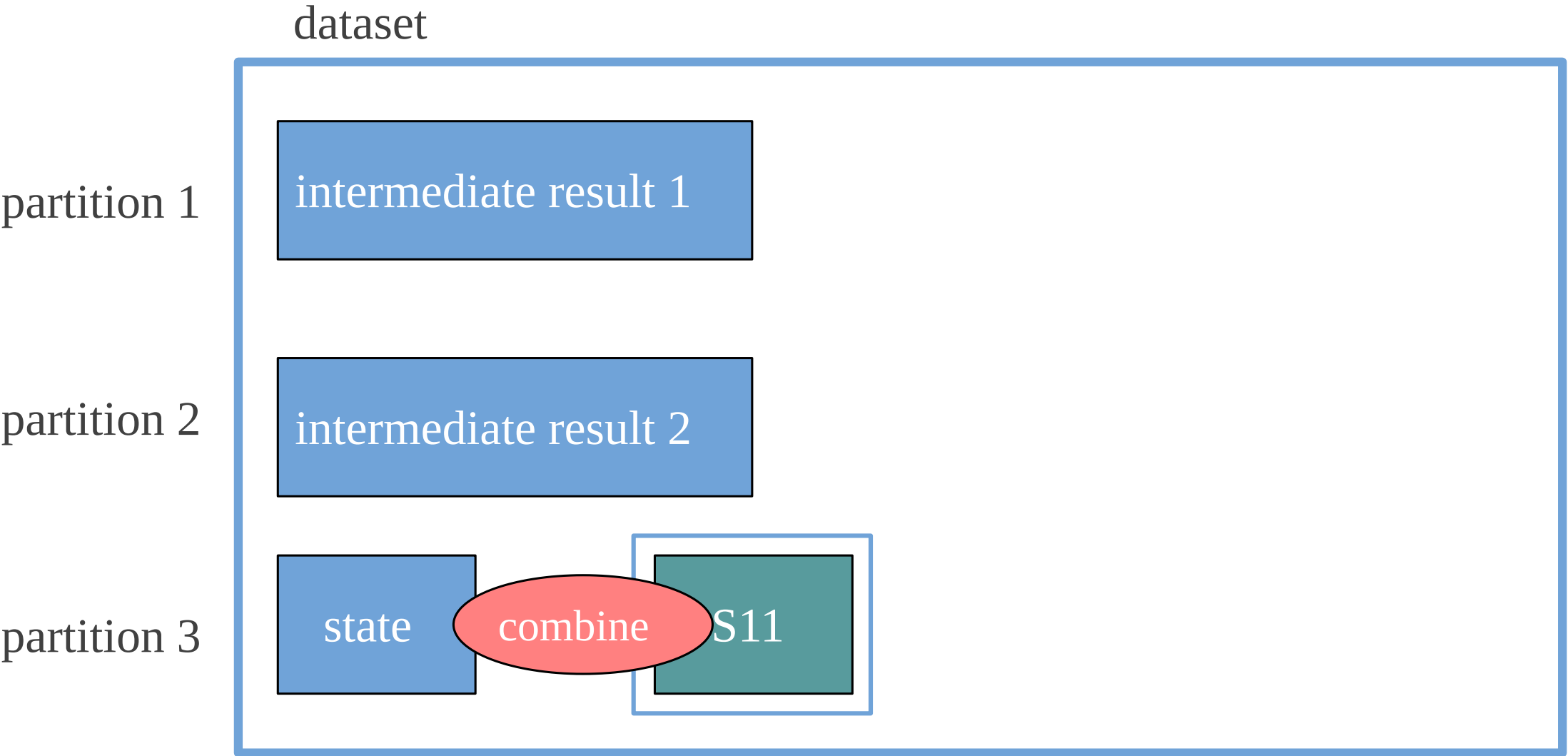
Fold partitions



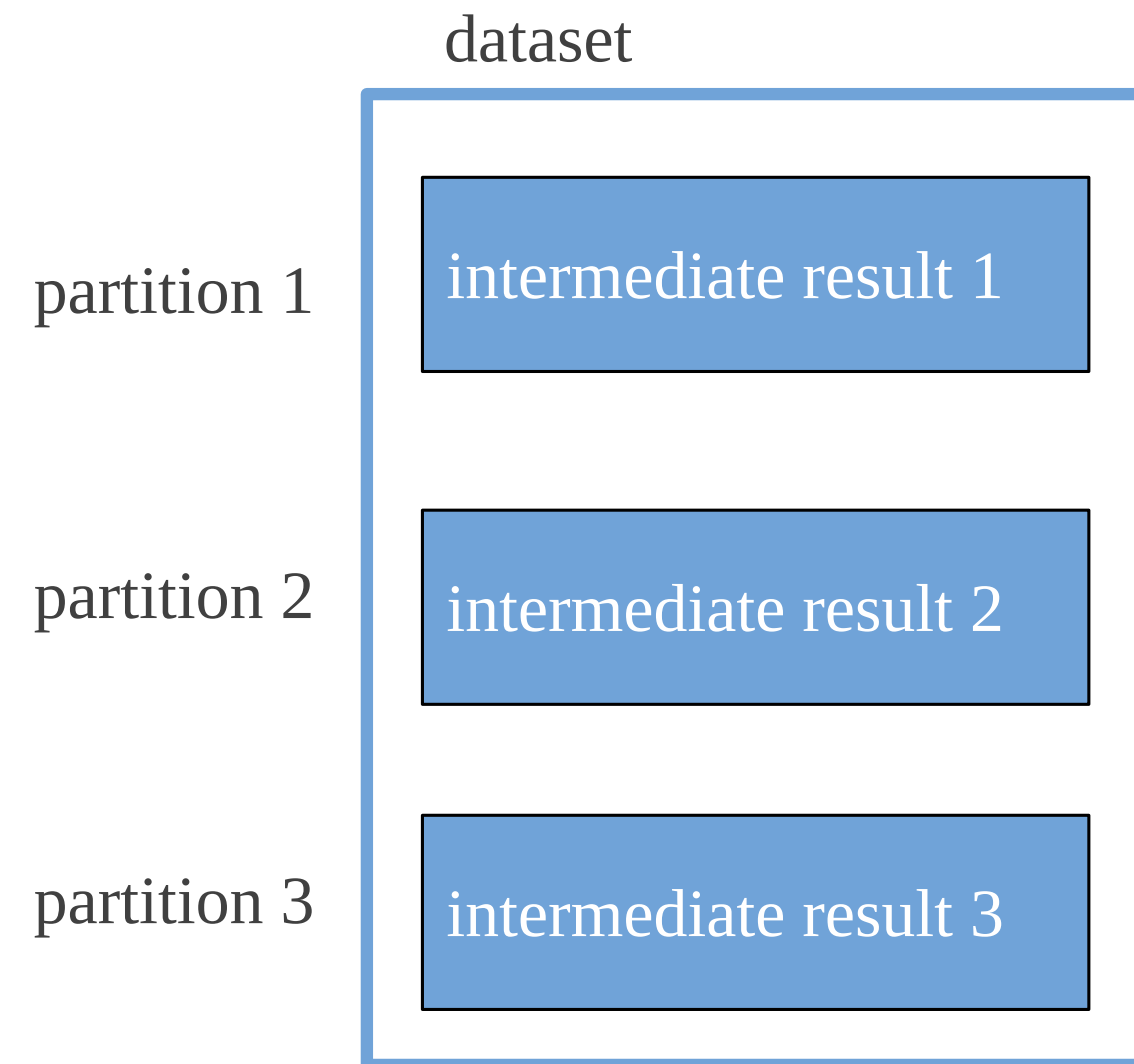
Fold partitions



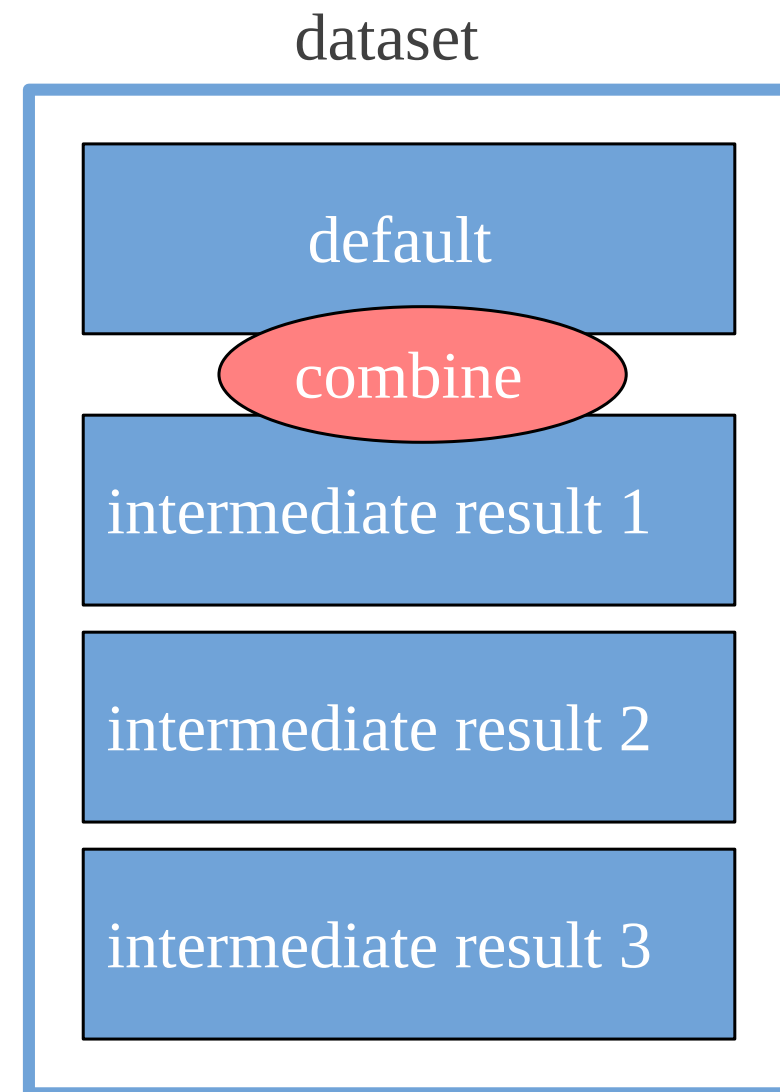
Fold partitions



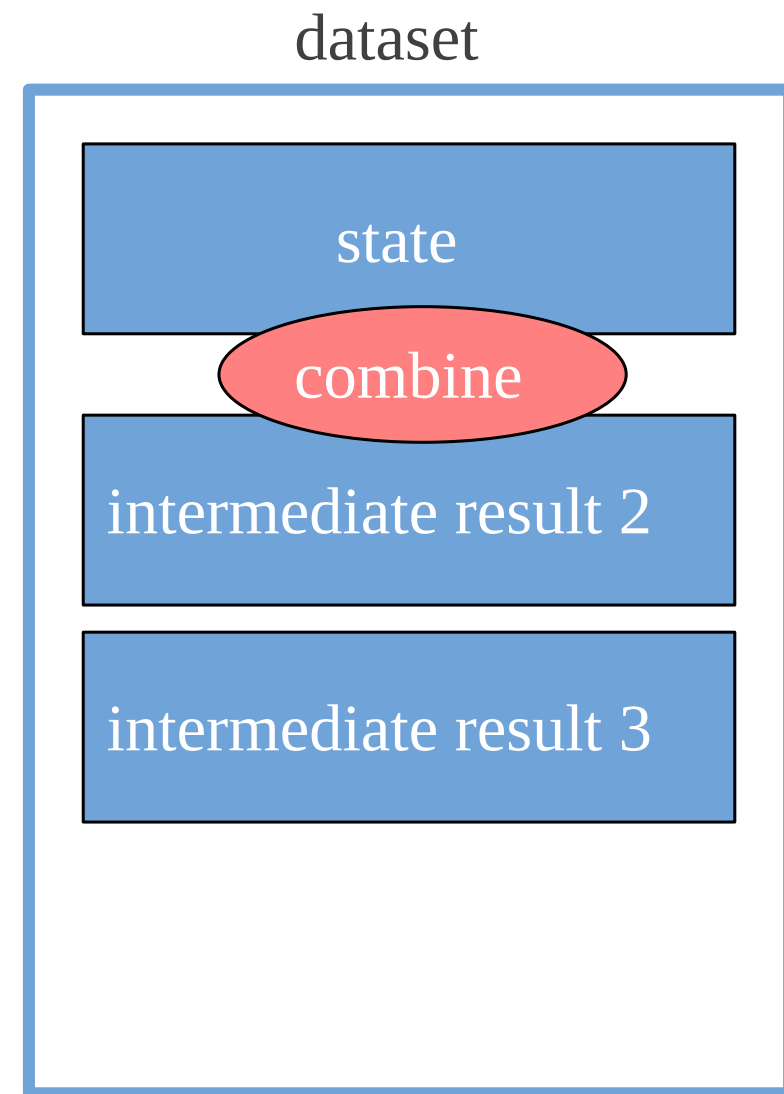
Fold partitions



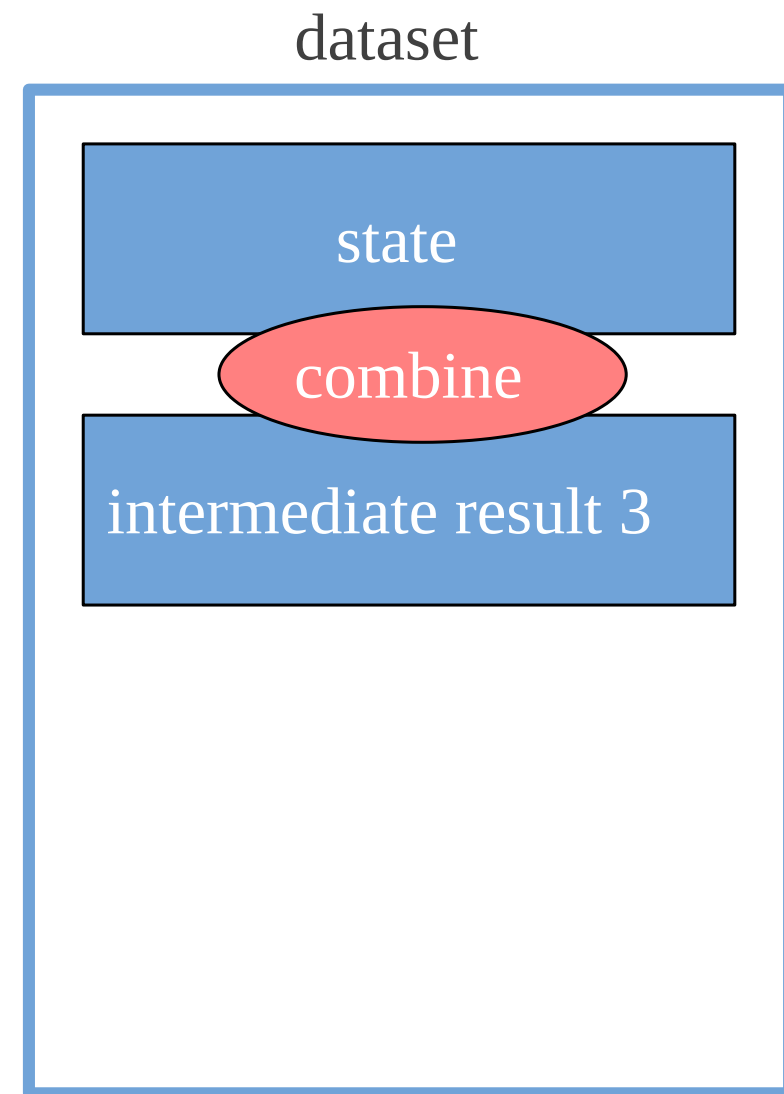
Fold intermediate results



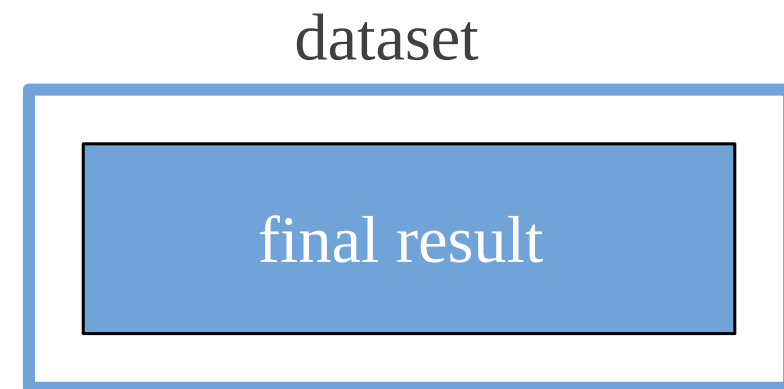
Fold intermediate results



Fold intermediate results



Fold intermediate results



The background of the slide features a complex, abstract network of light blue lines and dots. The dots, representing nodes, are of varying sizes and are interconnected by thin, light blue lines, creating a web-like structure that spans the entire frame. The overall aesthetic is clean and modern, with a light blue color palette.

TemperatureExercises.scala