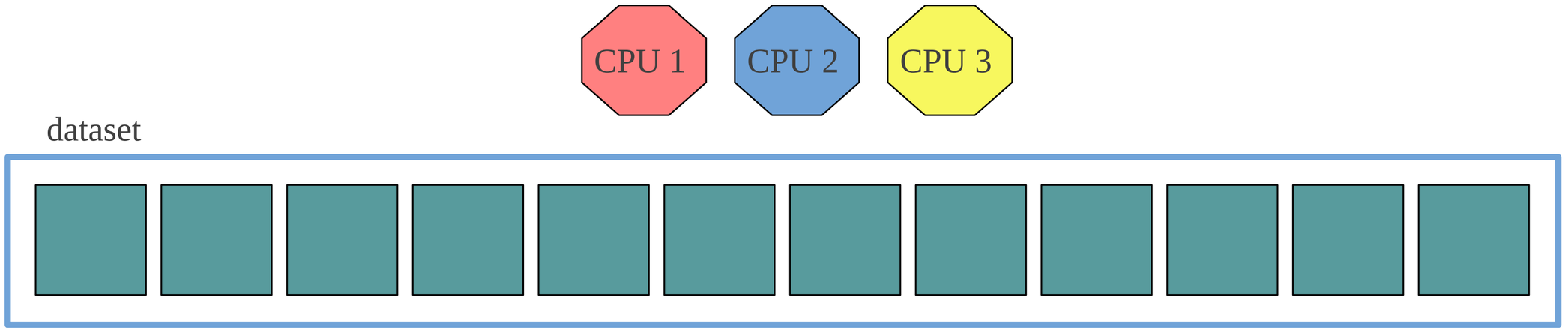
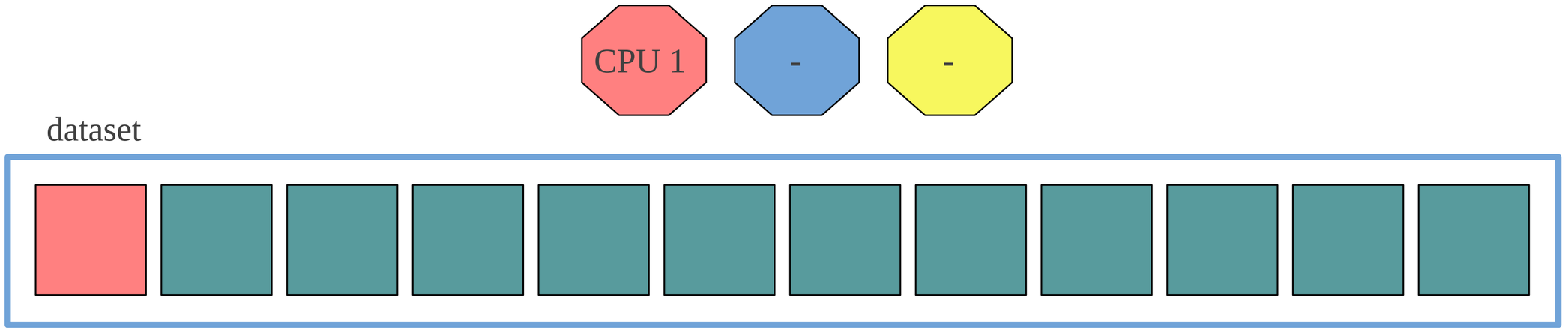
ANALYSIS OF GLOBAL TEMPERATURE
PART 1

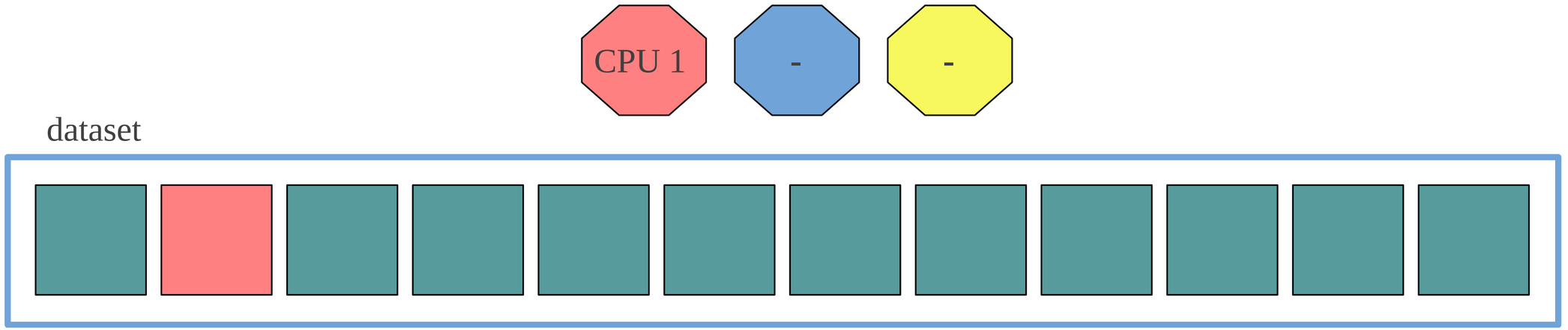
Data processing



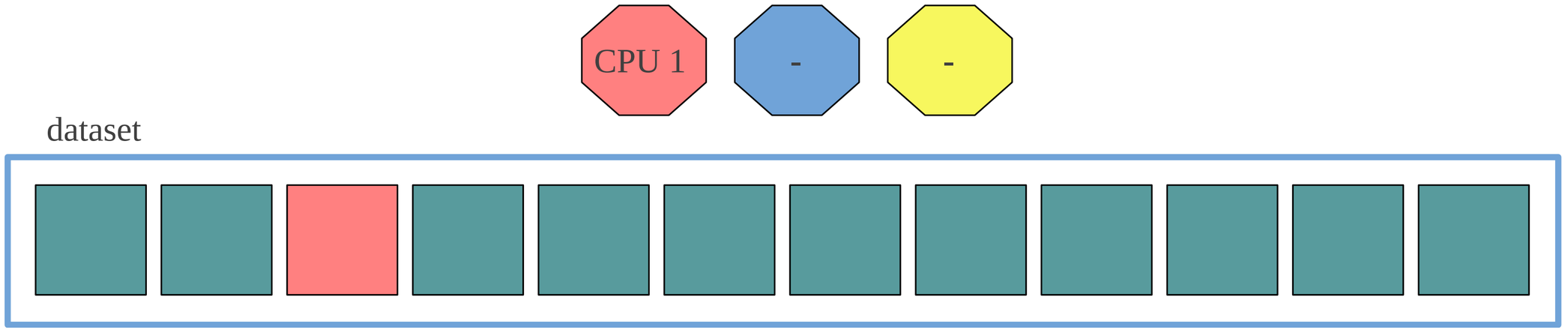
Sequential iteration



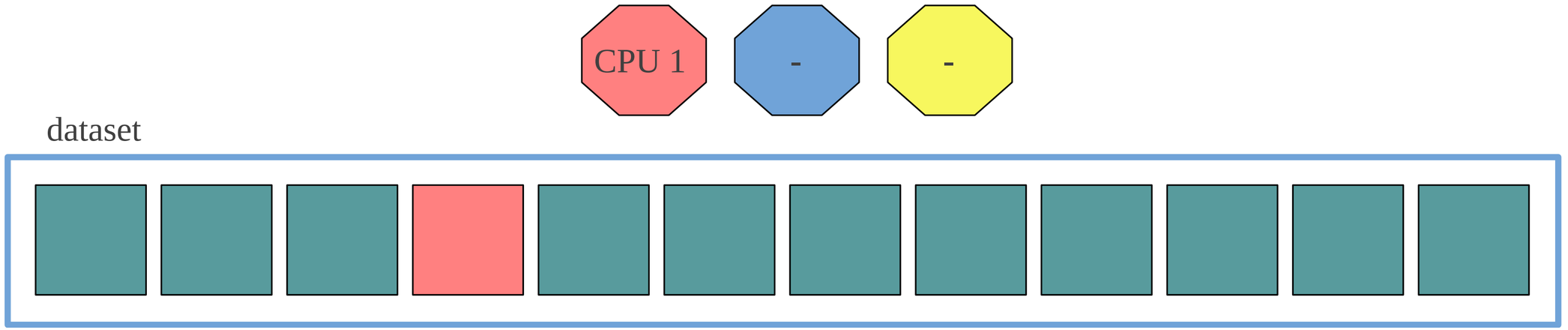
Sequential iteration



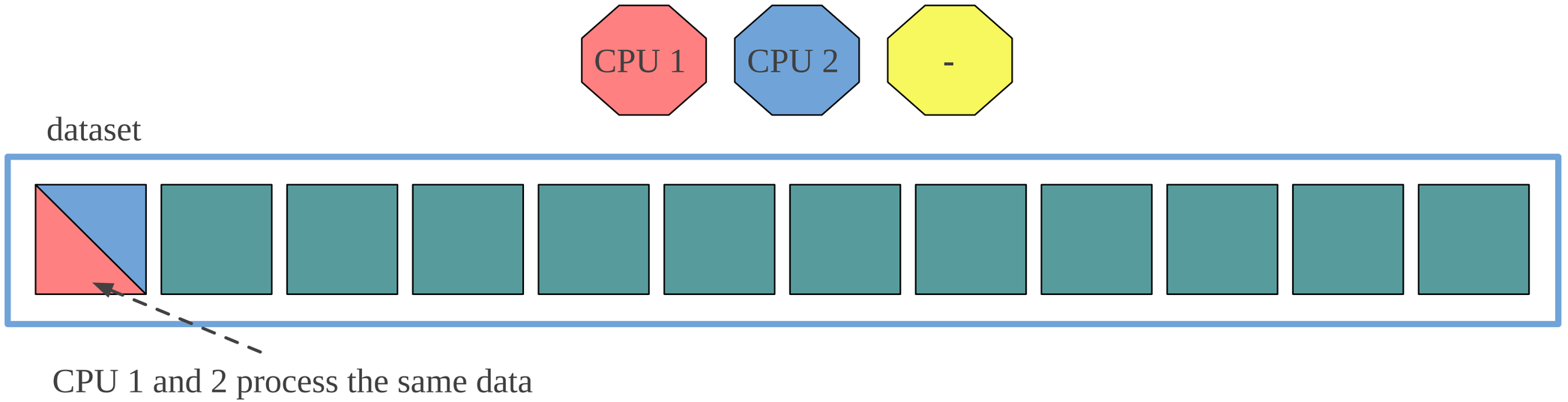
Sequential iteration



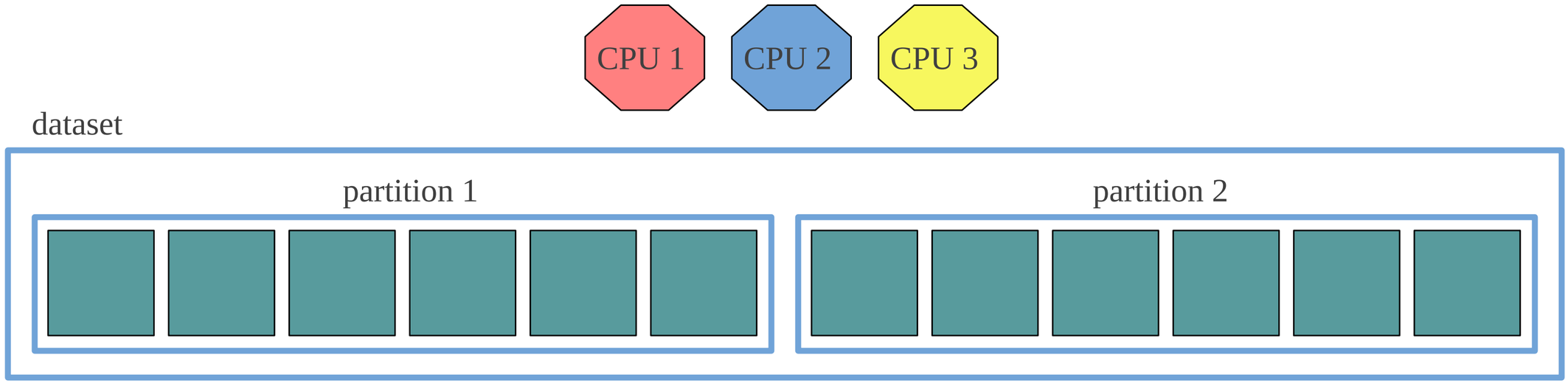
Sequential iteration



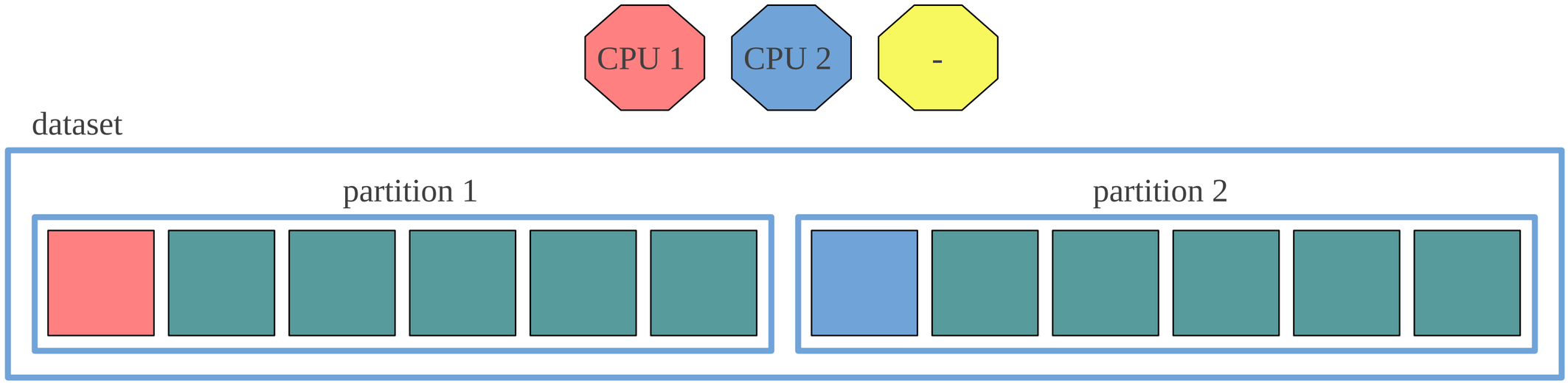
Useful work



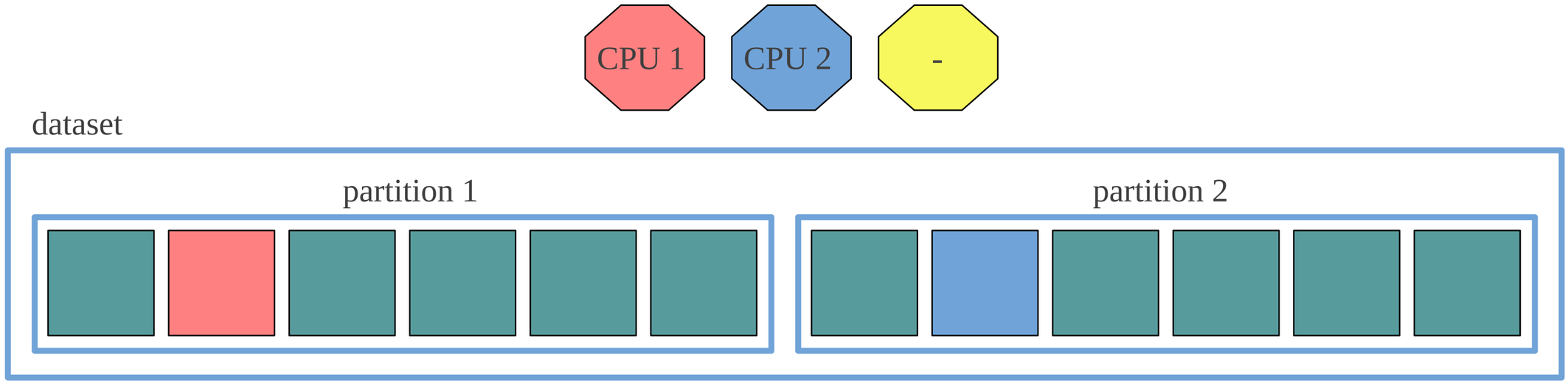
Partition the dataset



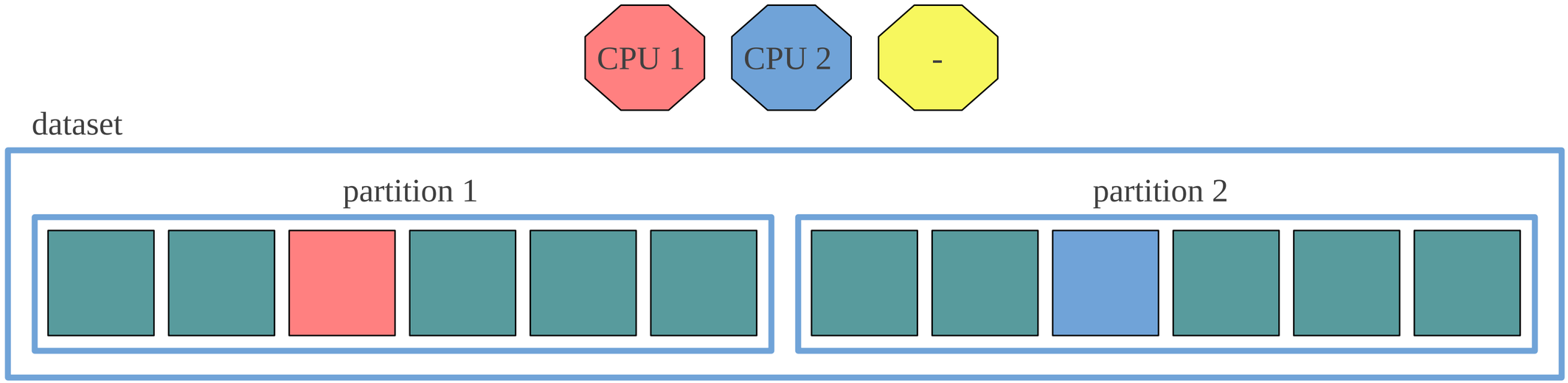
Partition the dataset



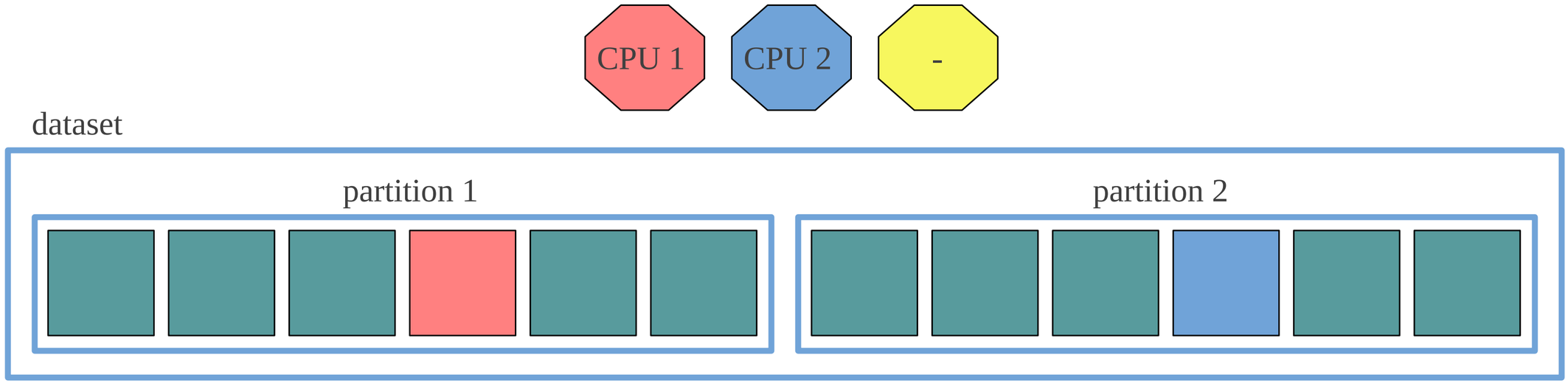
Partition the dataset



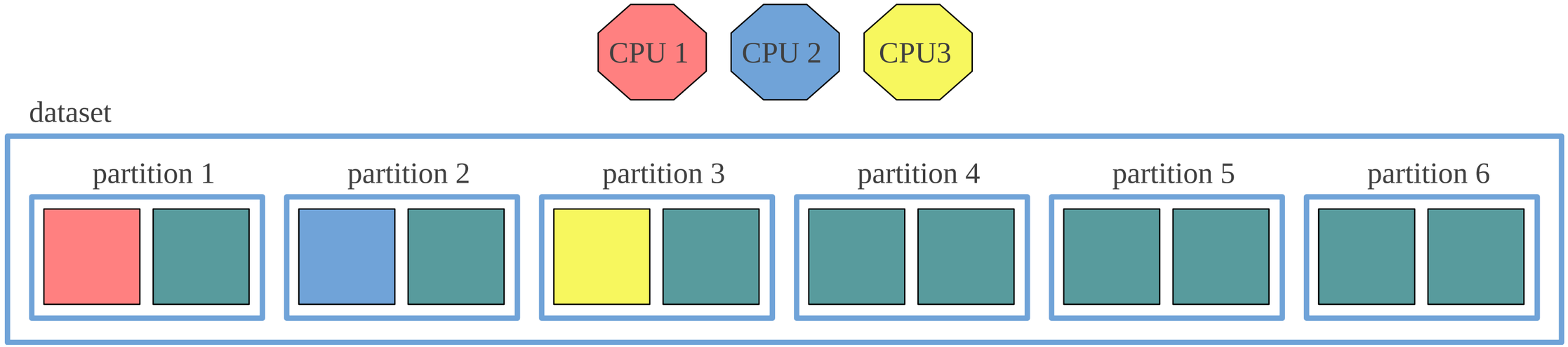
Partition the dataset



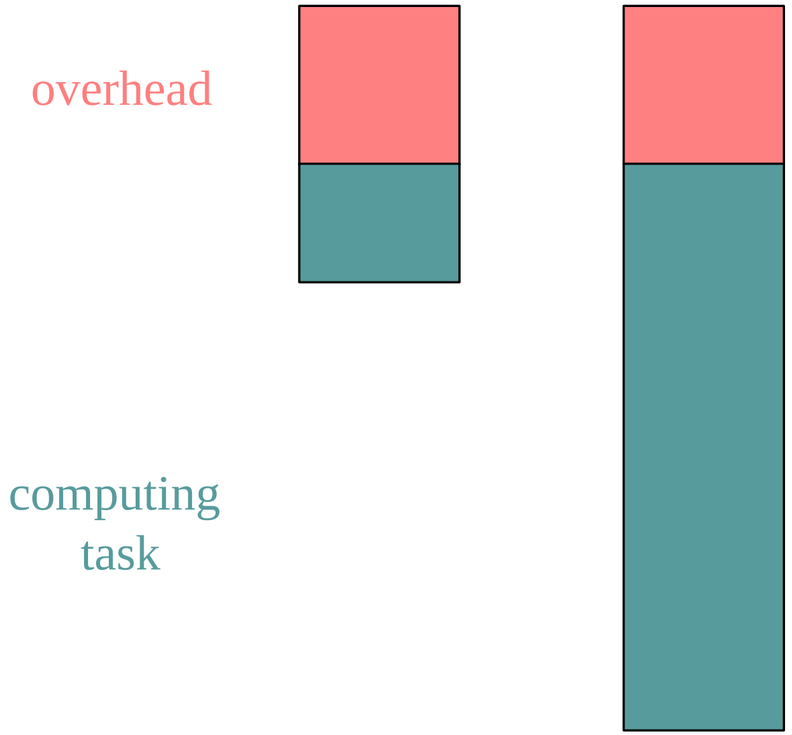
Partition the dataset



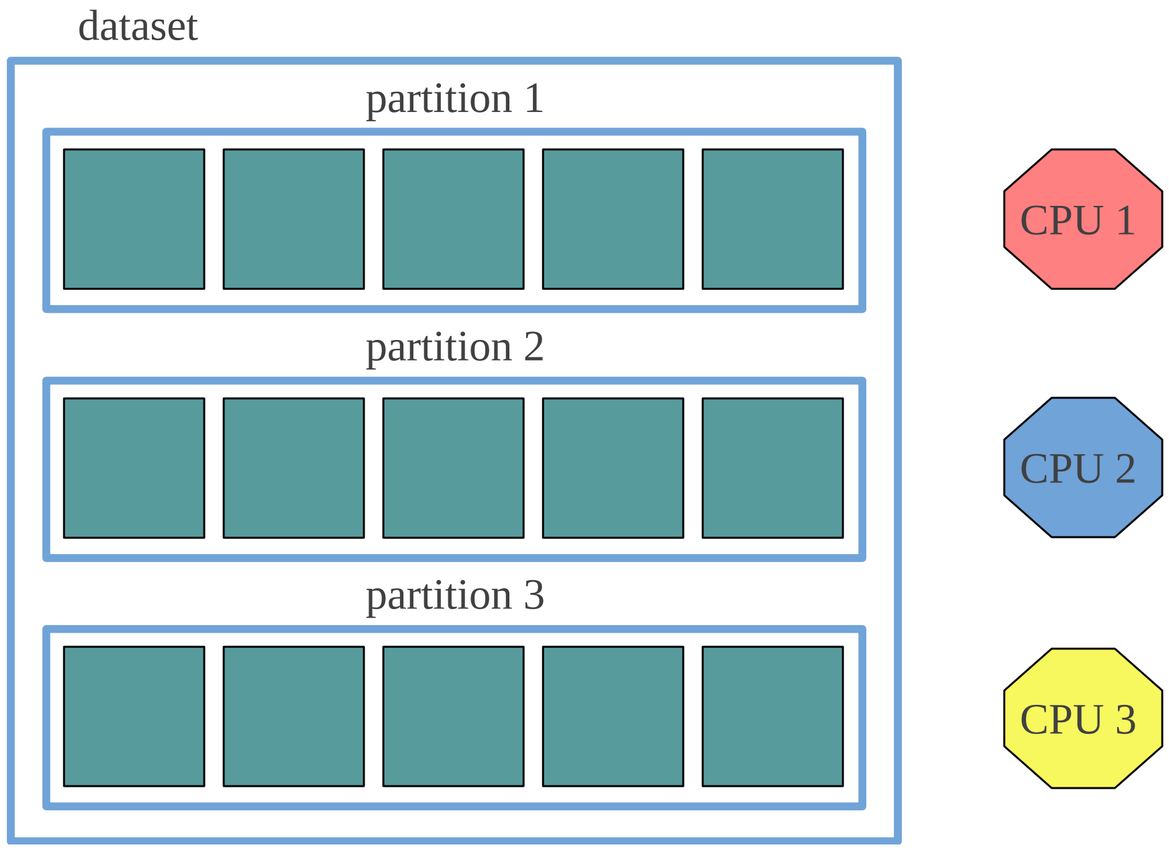
Partition the dataset



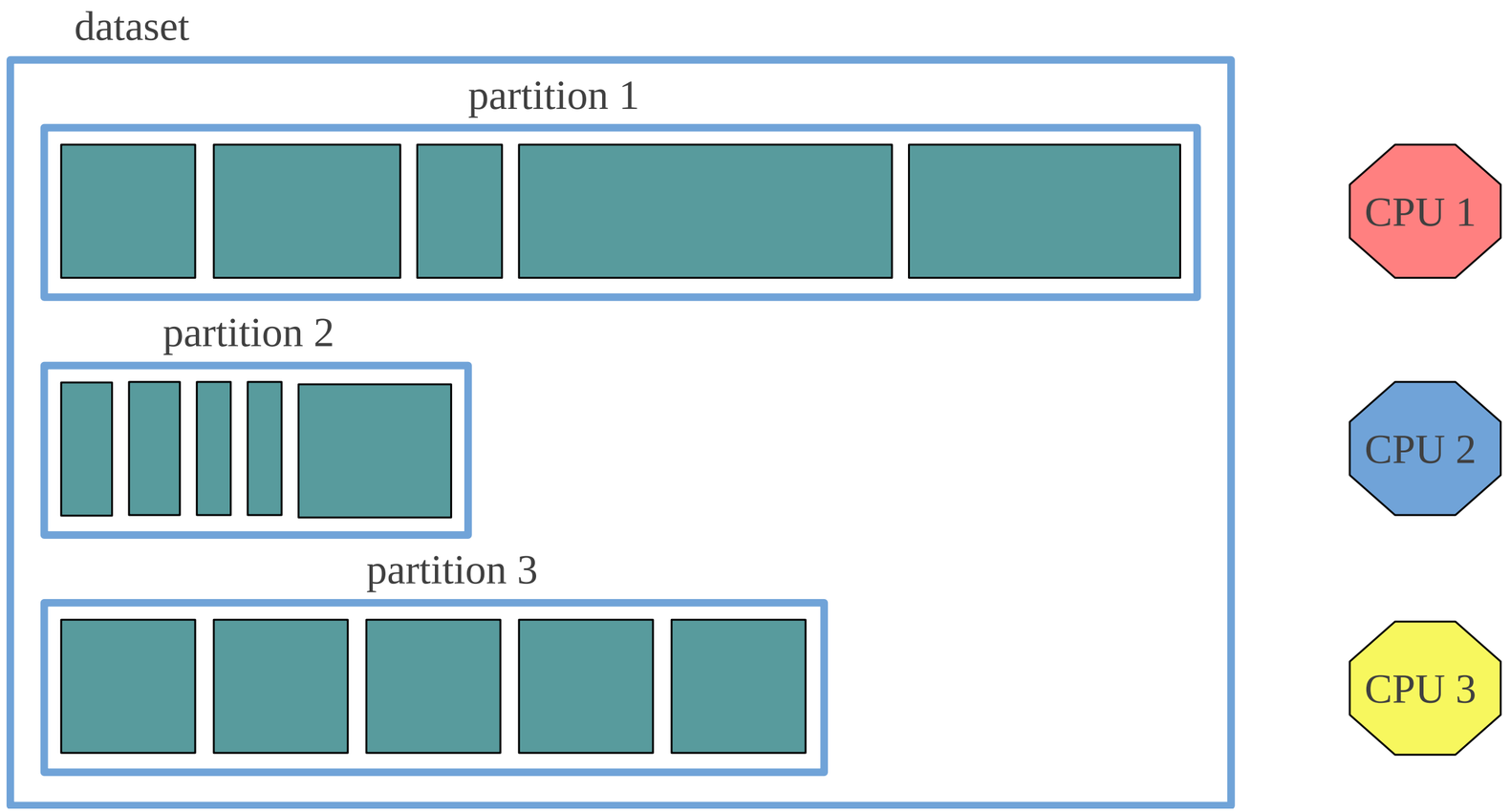
Size the partitions appropriately



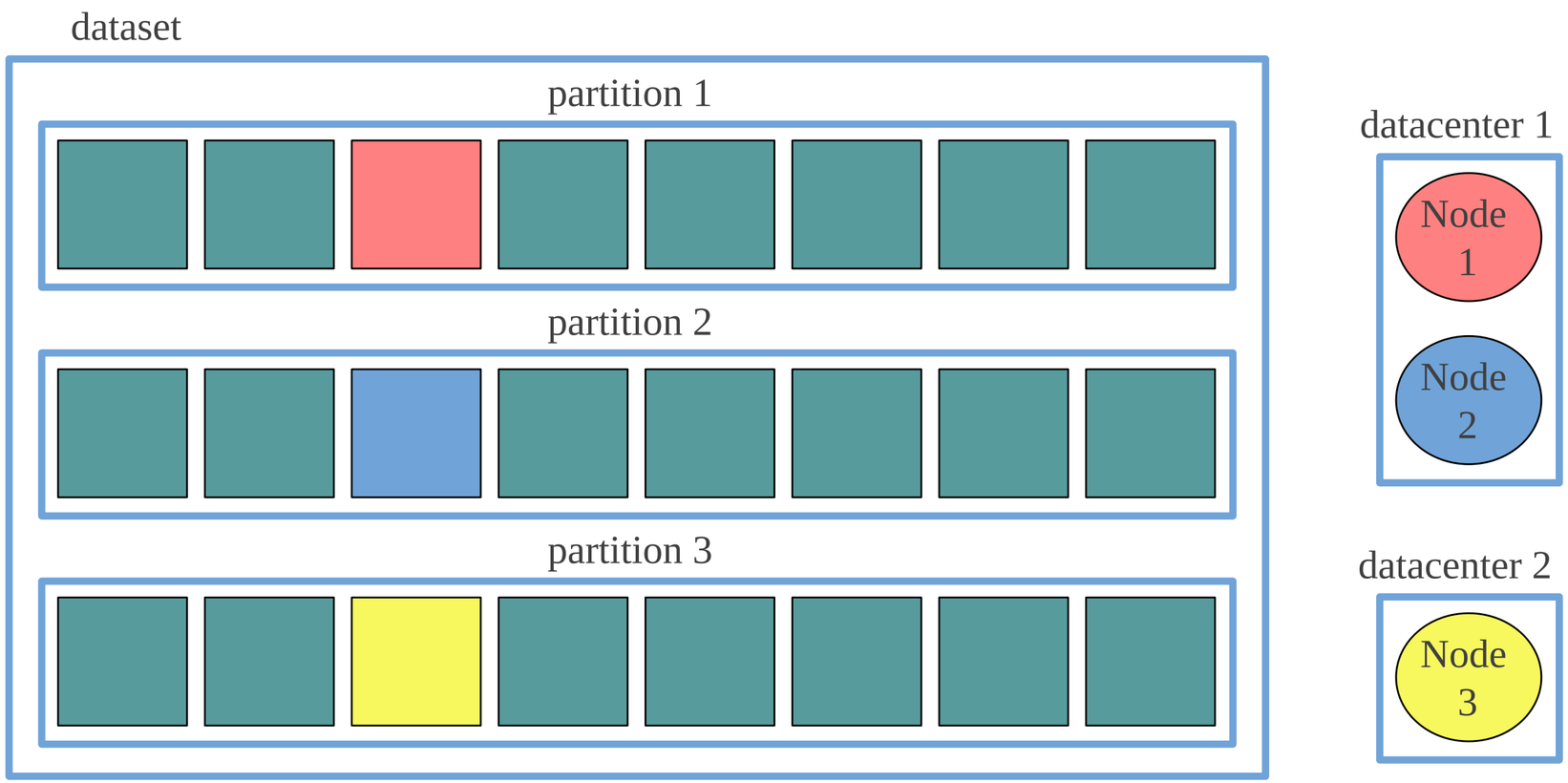
Tasks may require different computing power



Tasks may require different computing power



Scales to more than one computer



Important points

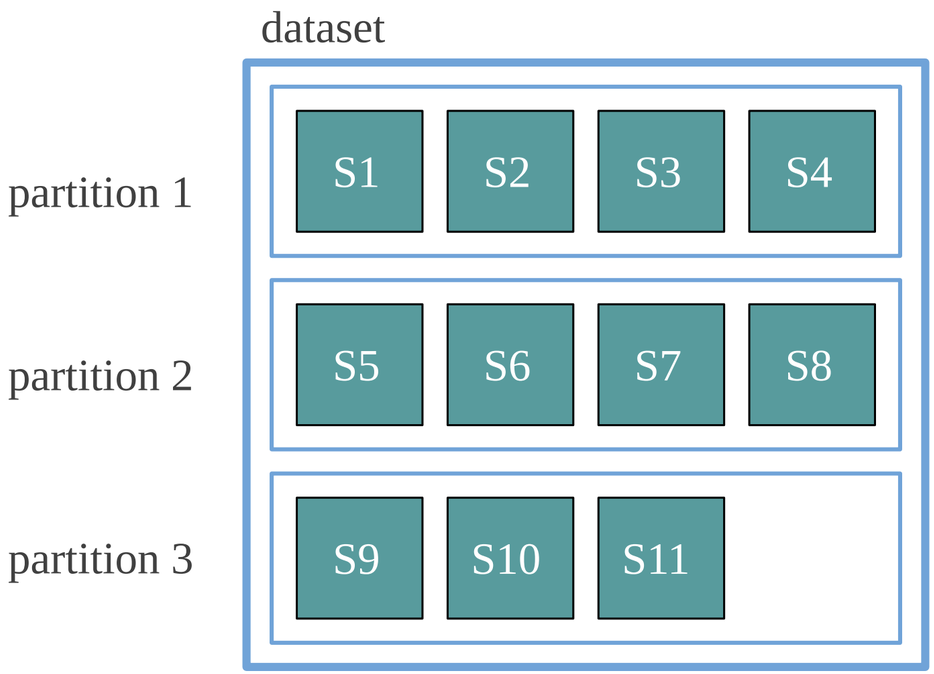
1. Size the partitions appropriately
2. Benchmark and tweak configuration for the task at hand

Important points

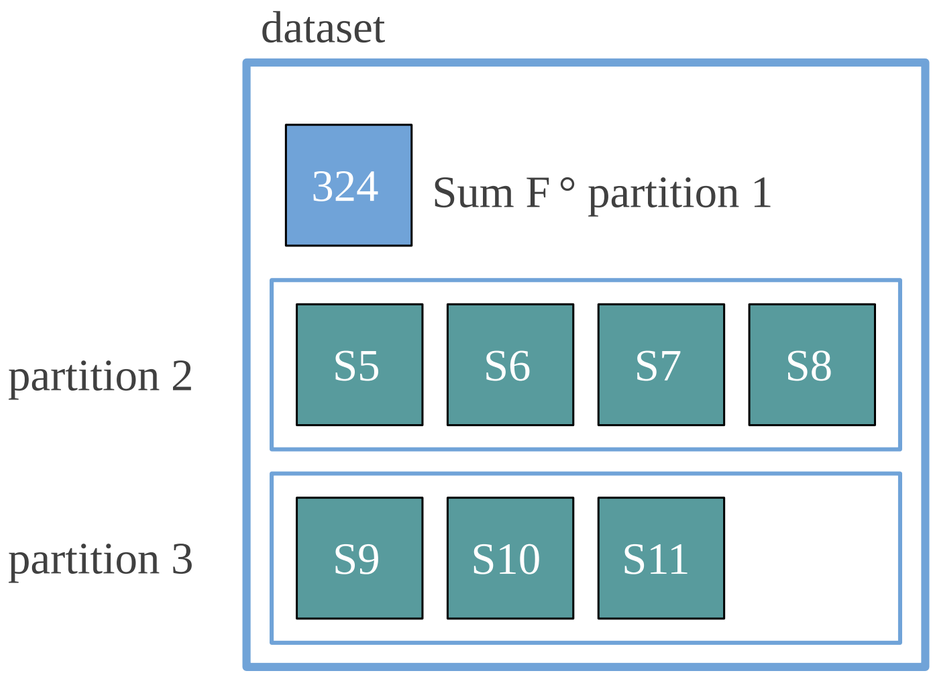
1. Size the partitions appropriately
2. Benchmark and tweak configuration for the task at hand
3. Parallel process MUST produce the SAME result as the sequential one

TemperatureNotebook.scala

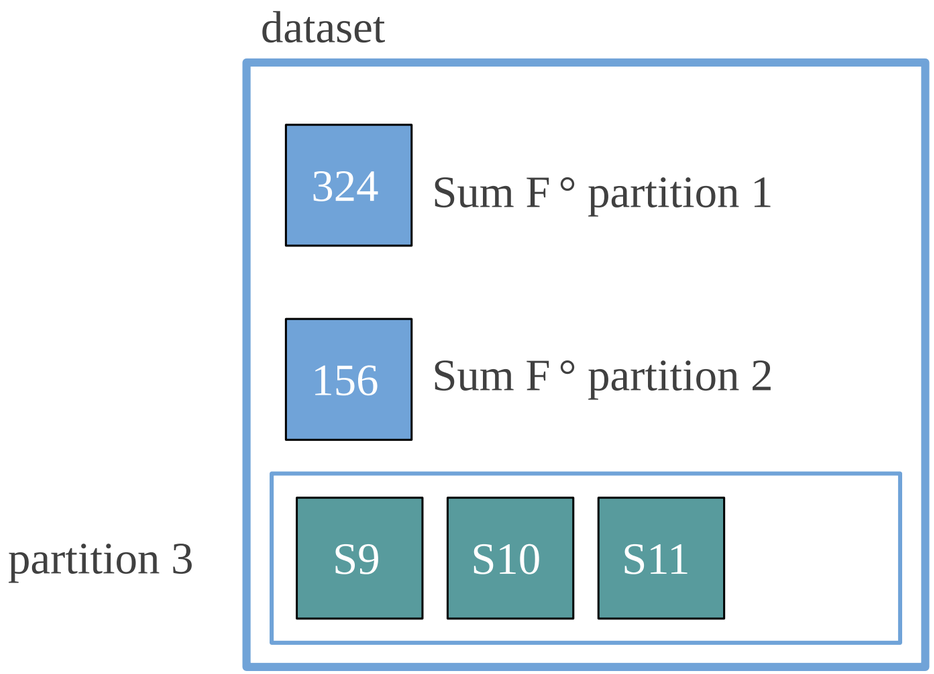
Average temperature



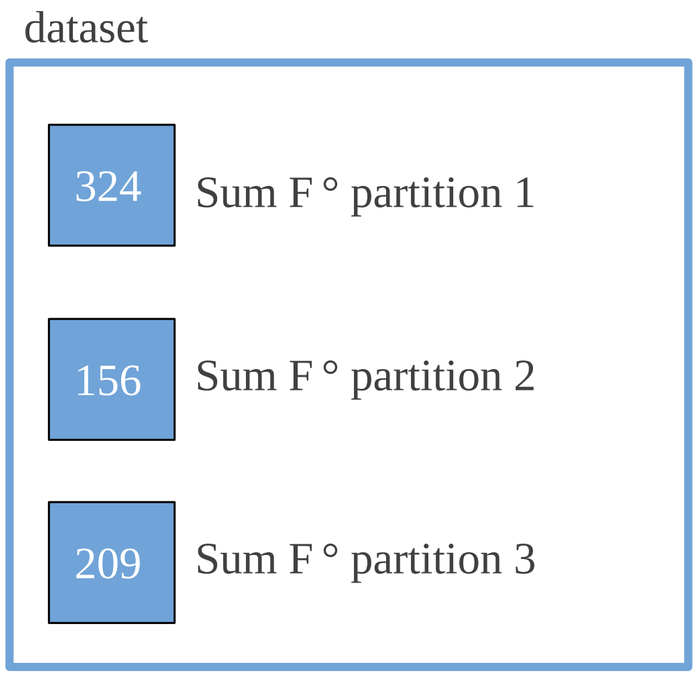
Sum temperatures per partition



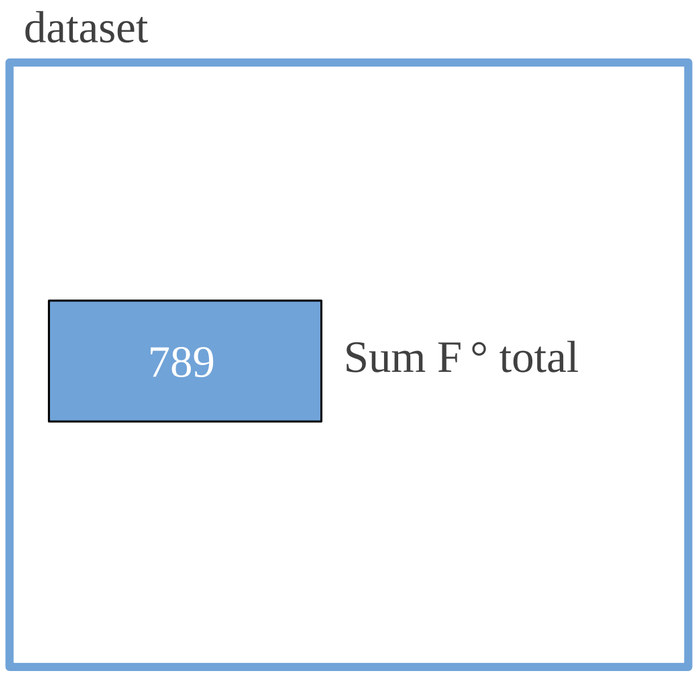
Sum temperatures per partition



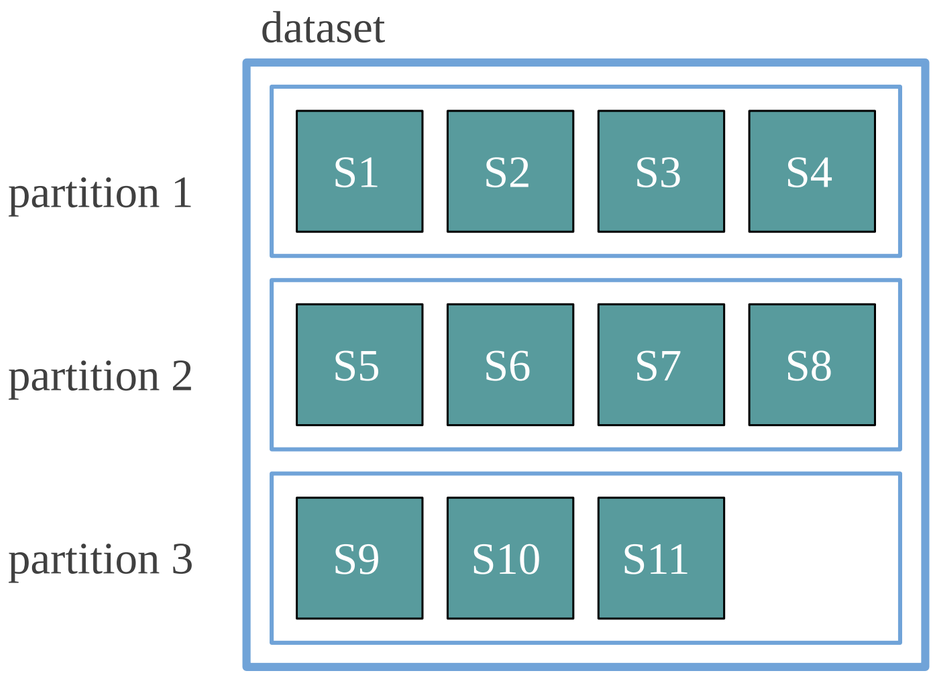
Sum temperatures per partition



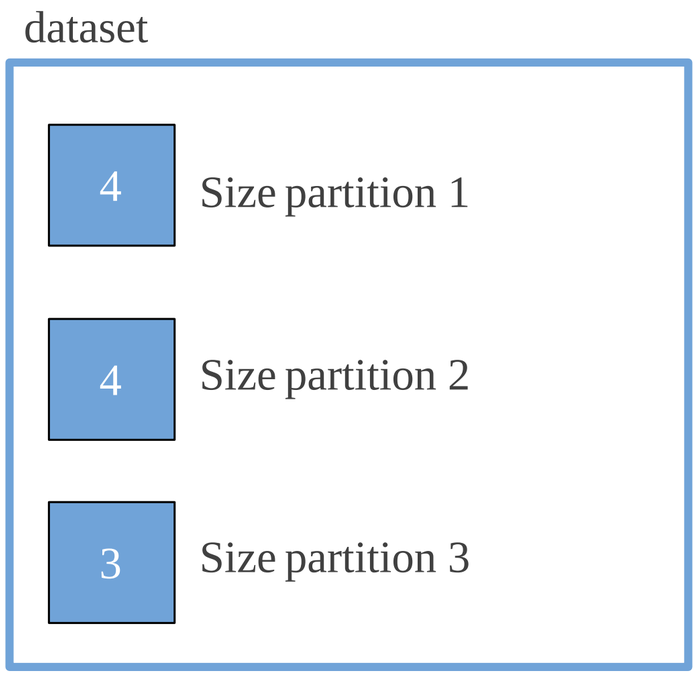
Sum-up all partitions



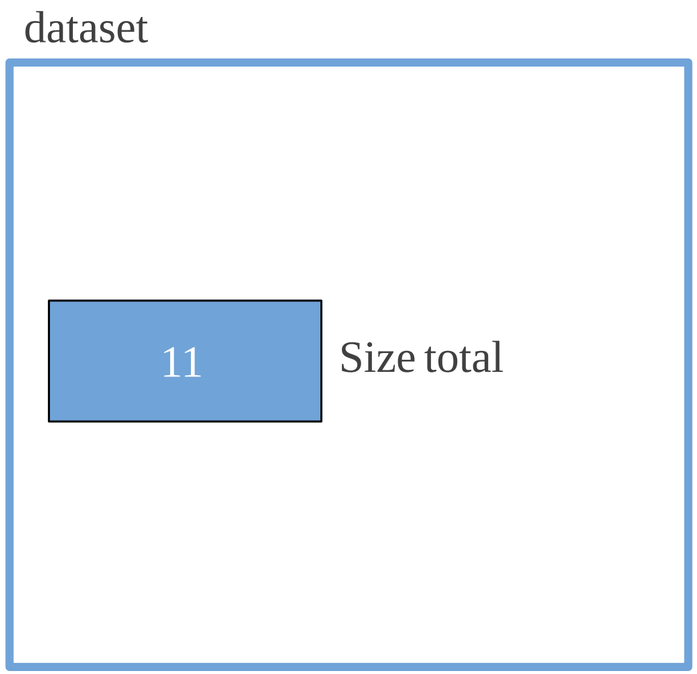
Size


dataset
partition 1
partition 2
partition 3
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11

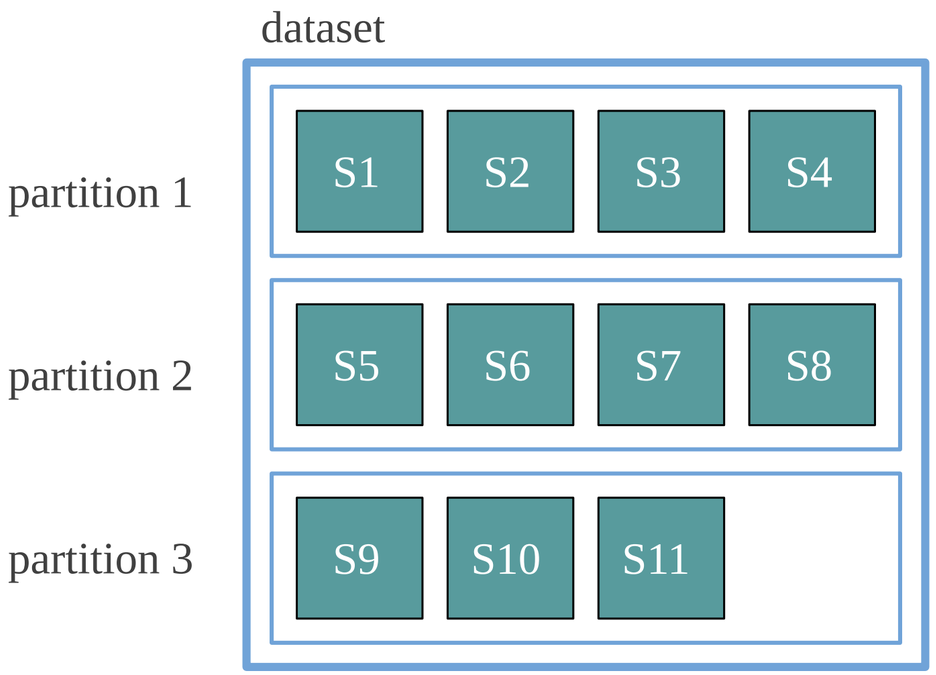
Calculate size per partition



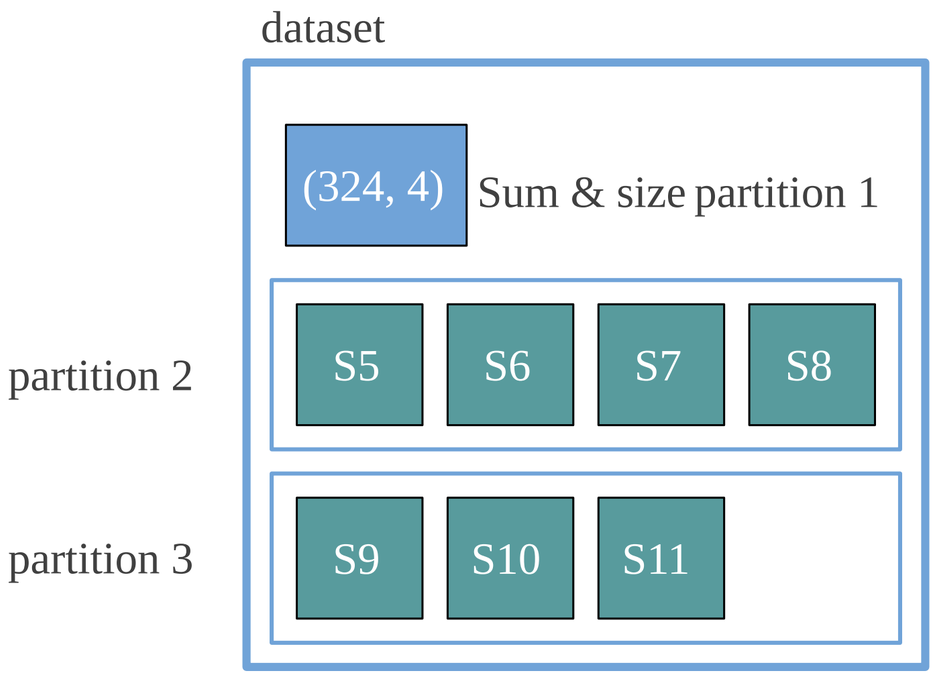
Sum-up all partitions



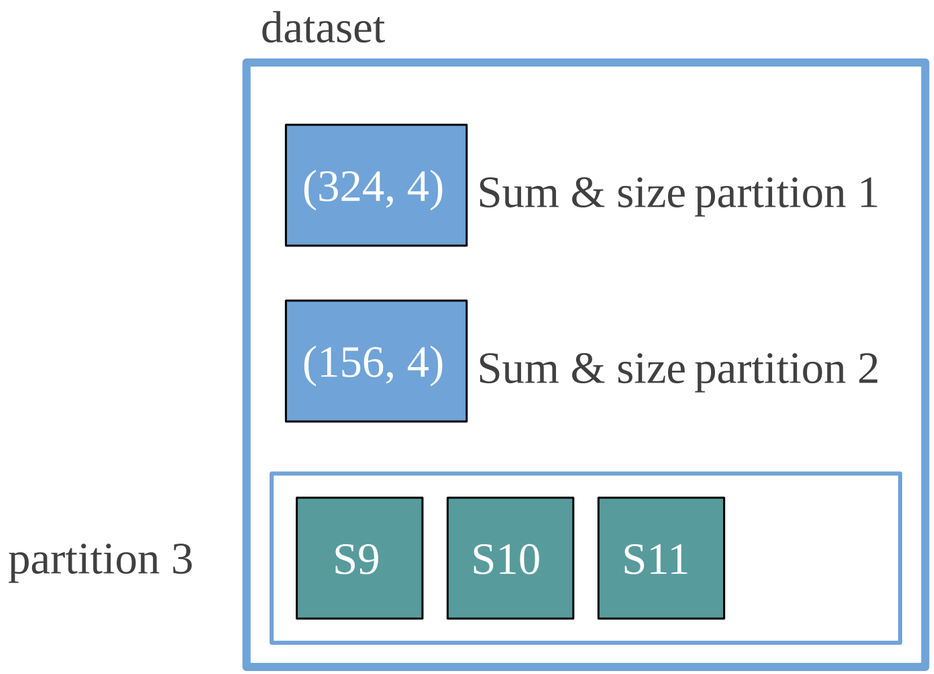
Average temperature in one pass



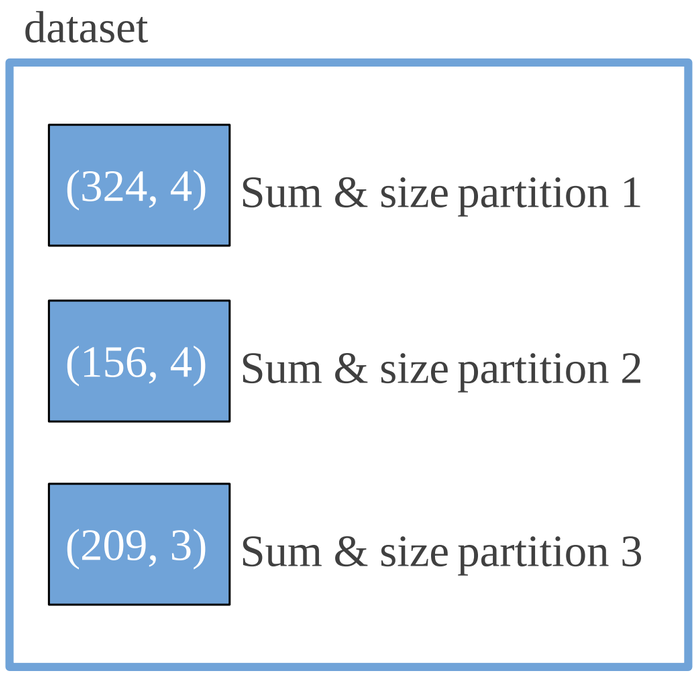
Calculate sum and size per partition



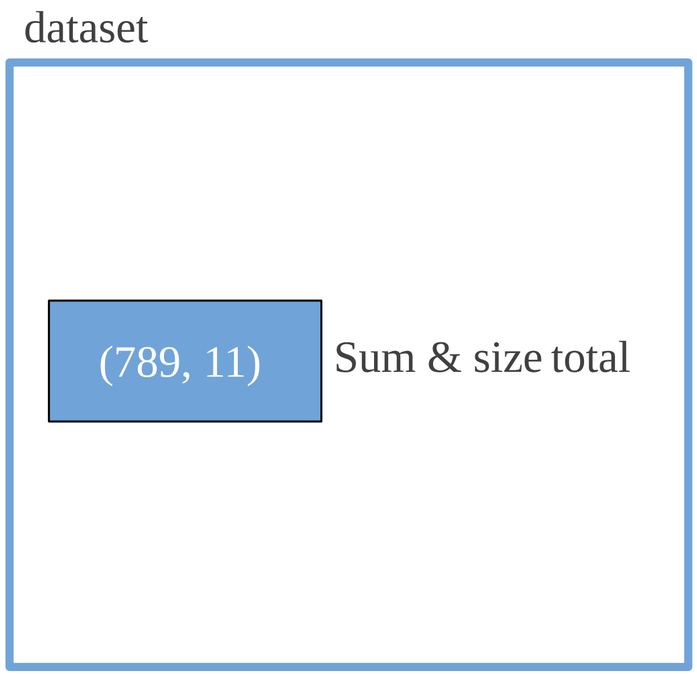
Calculate sum and size per partition



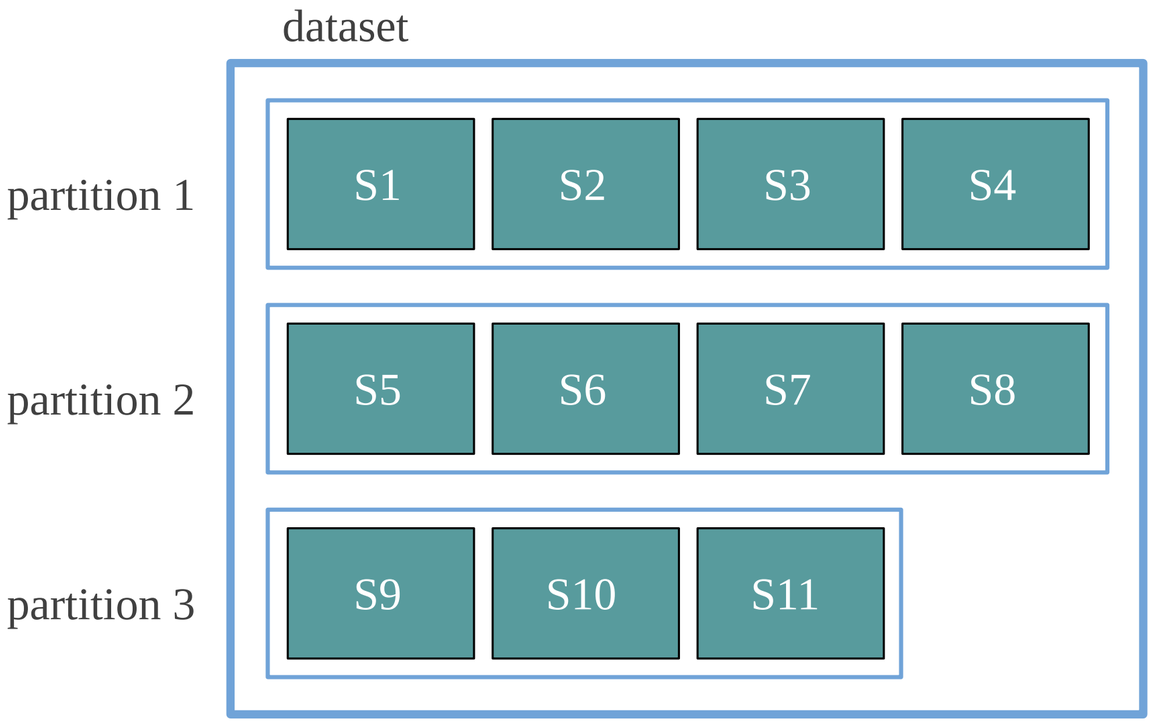
Calculate sum and size per partition



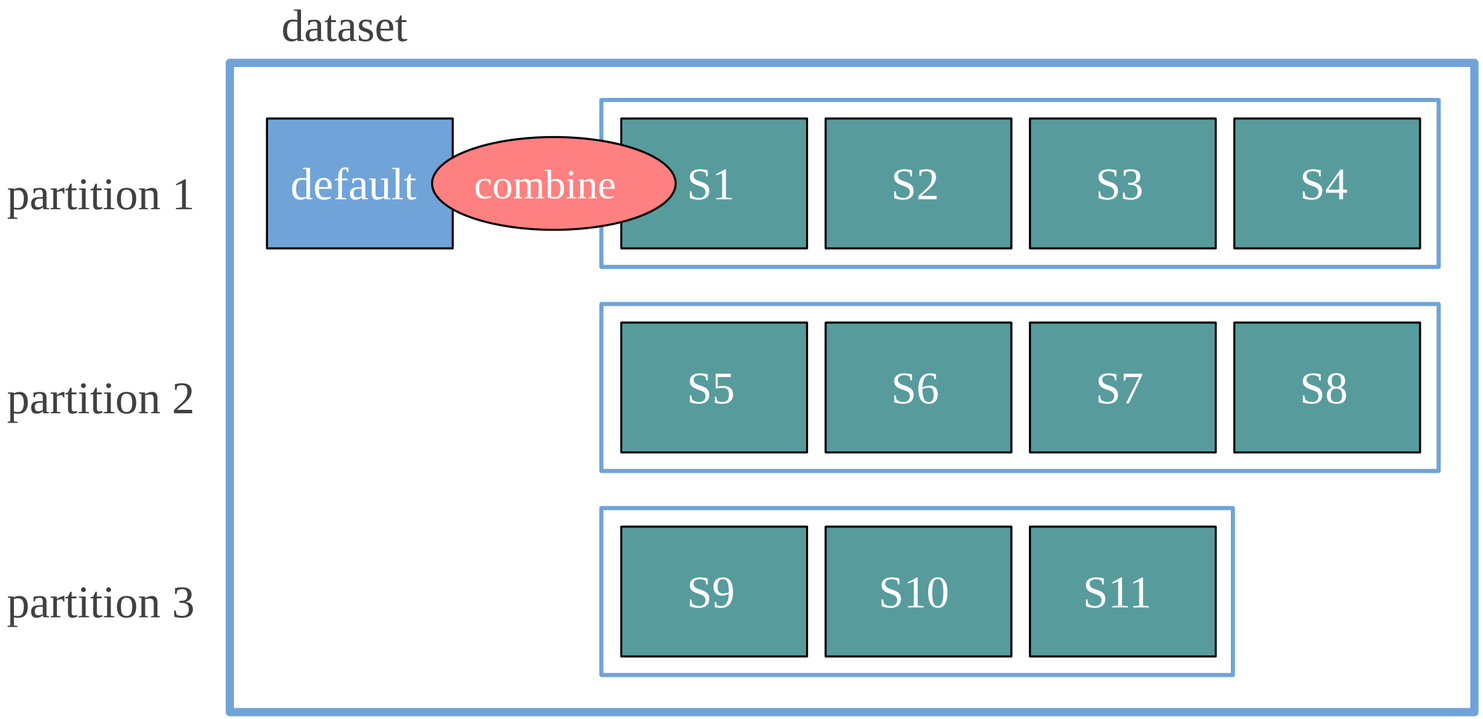
Sum-up all partitions



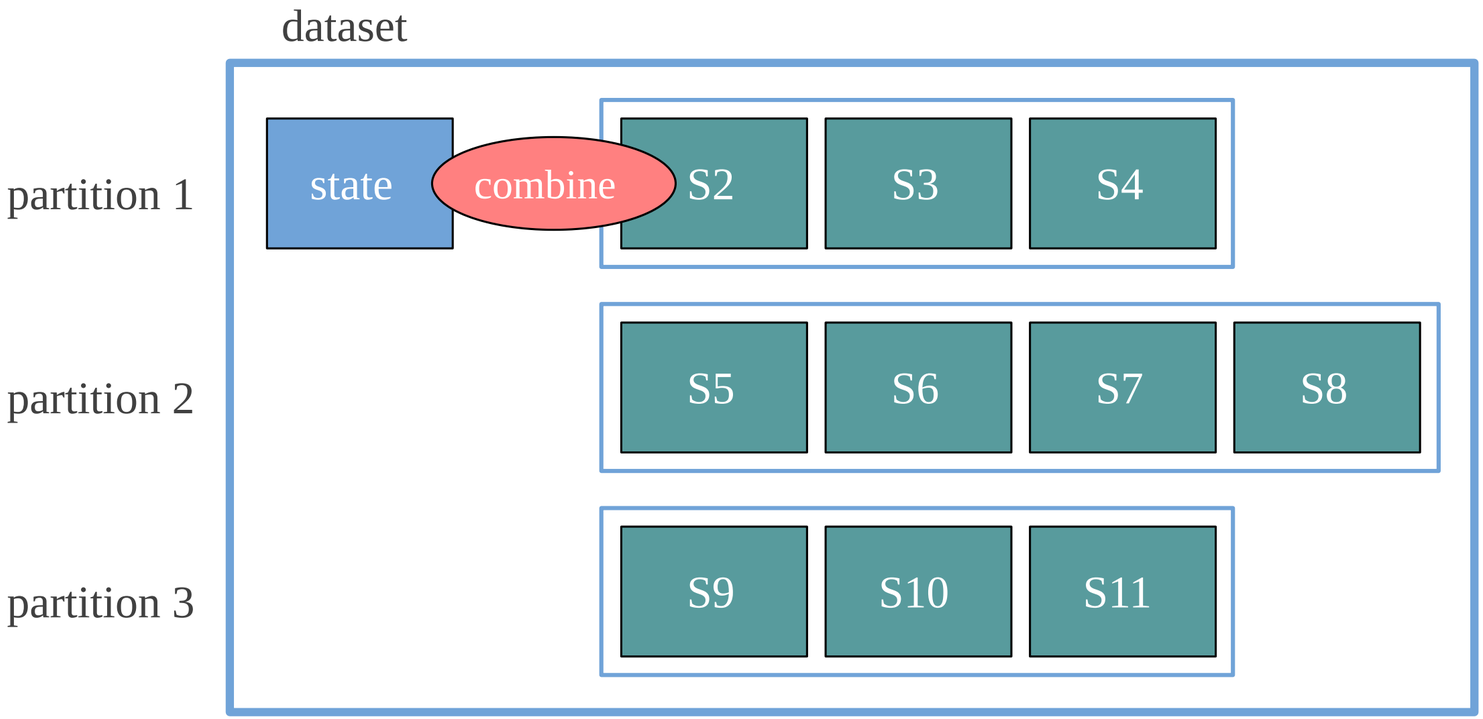
Fold partitions



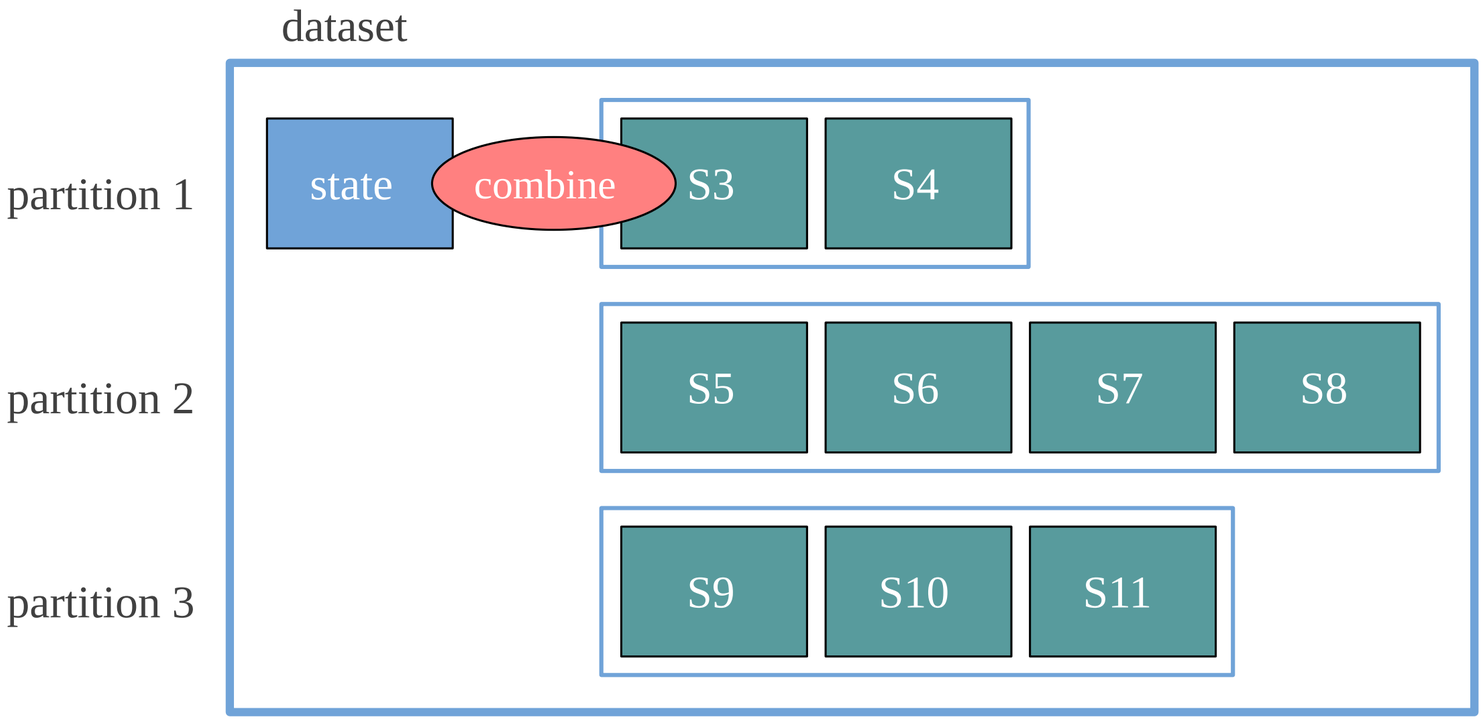
Fold partitions



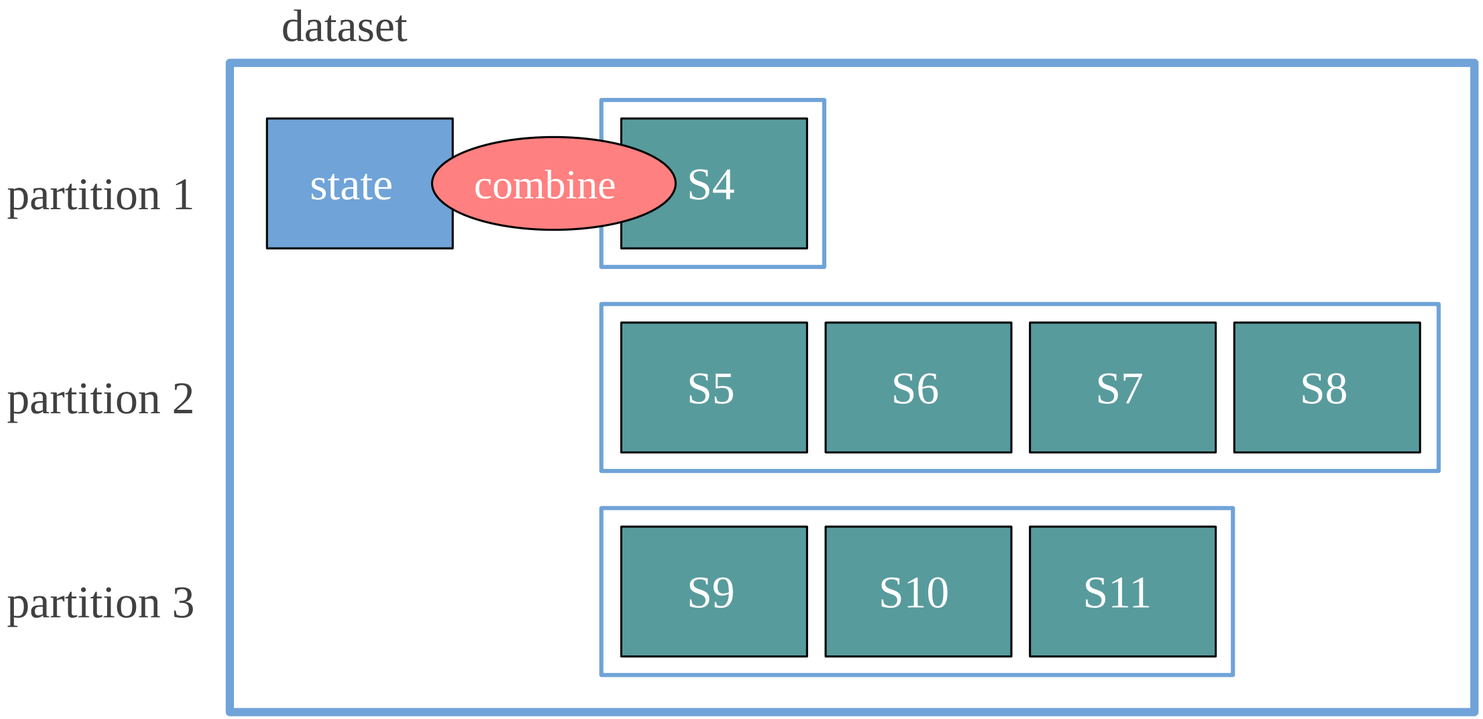
Fold partitions



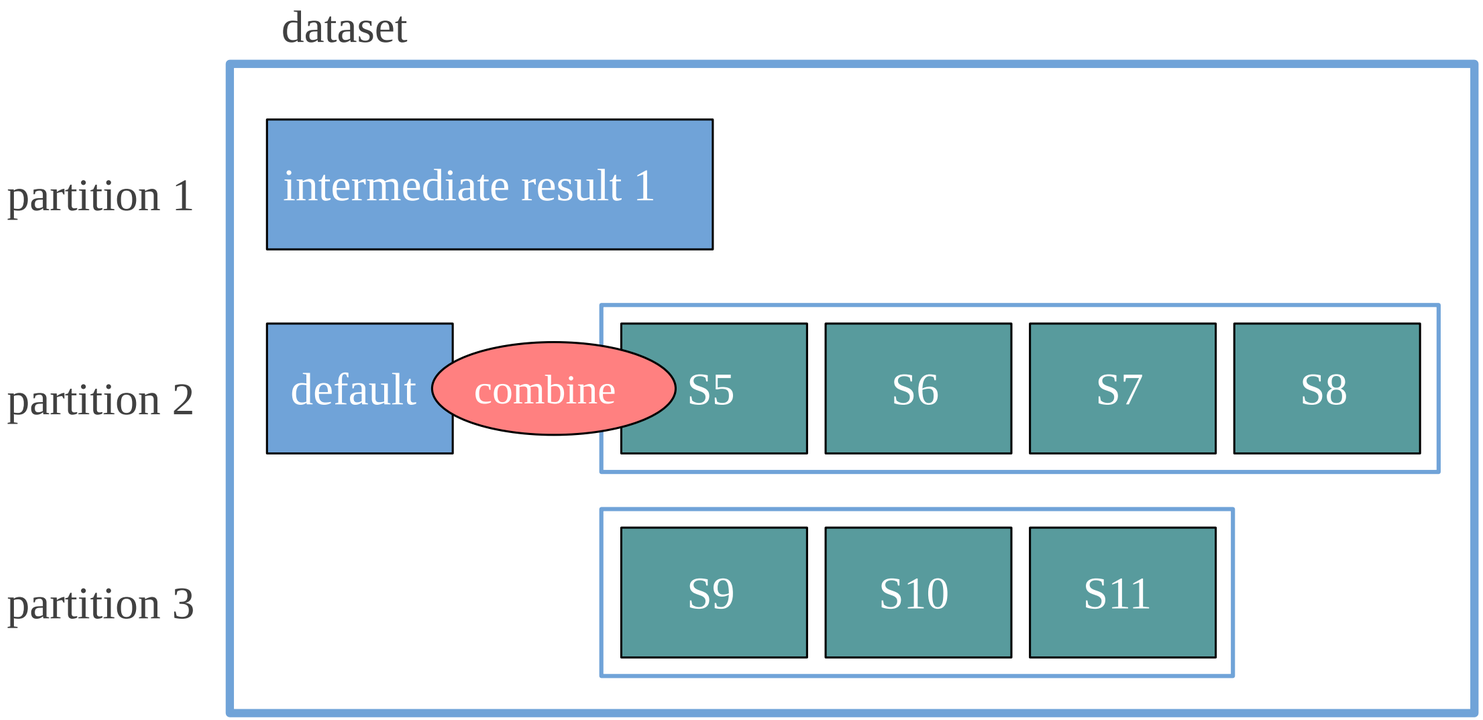
Fold partitions



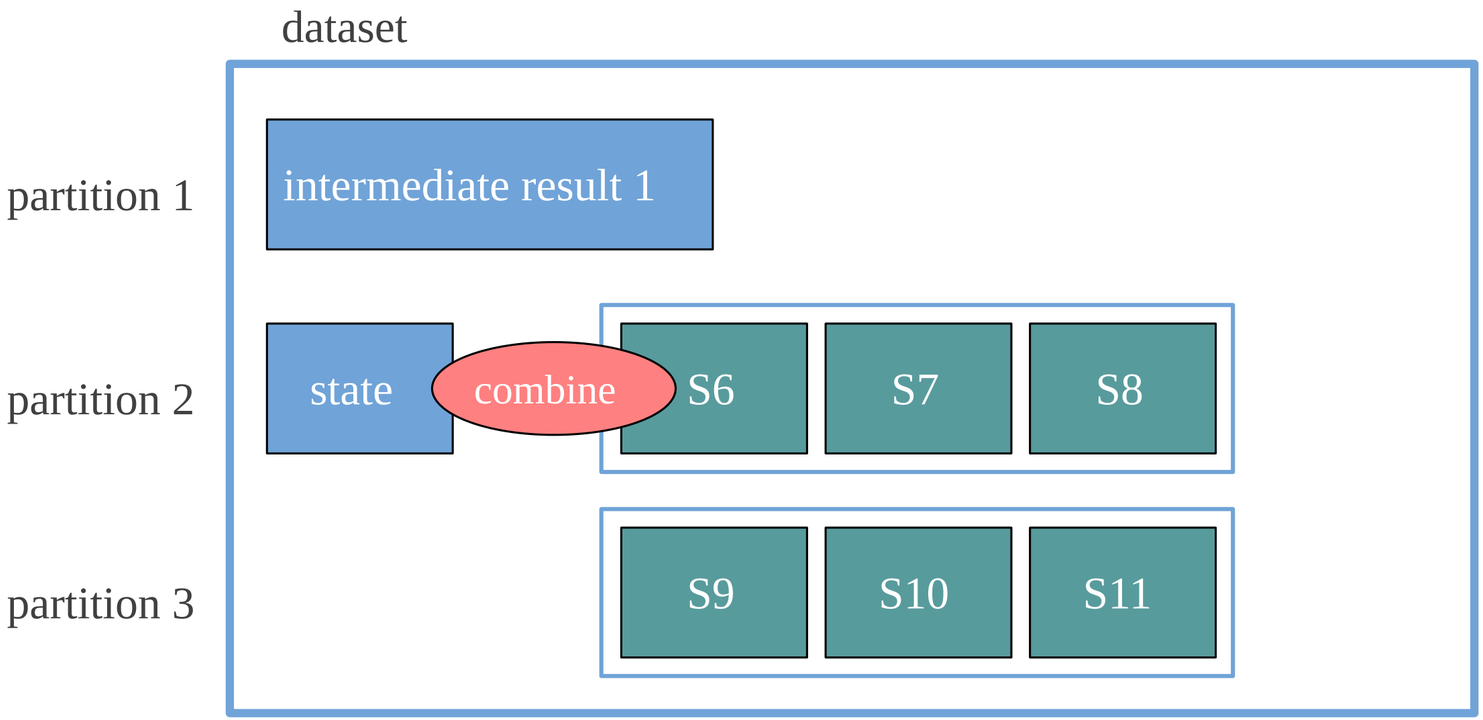
Fold partitions



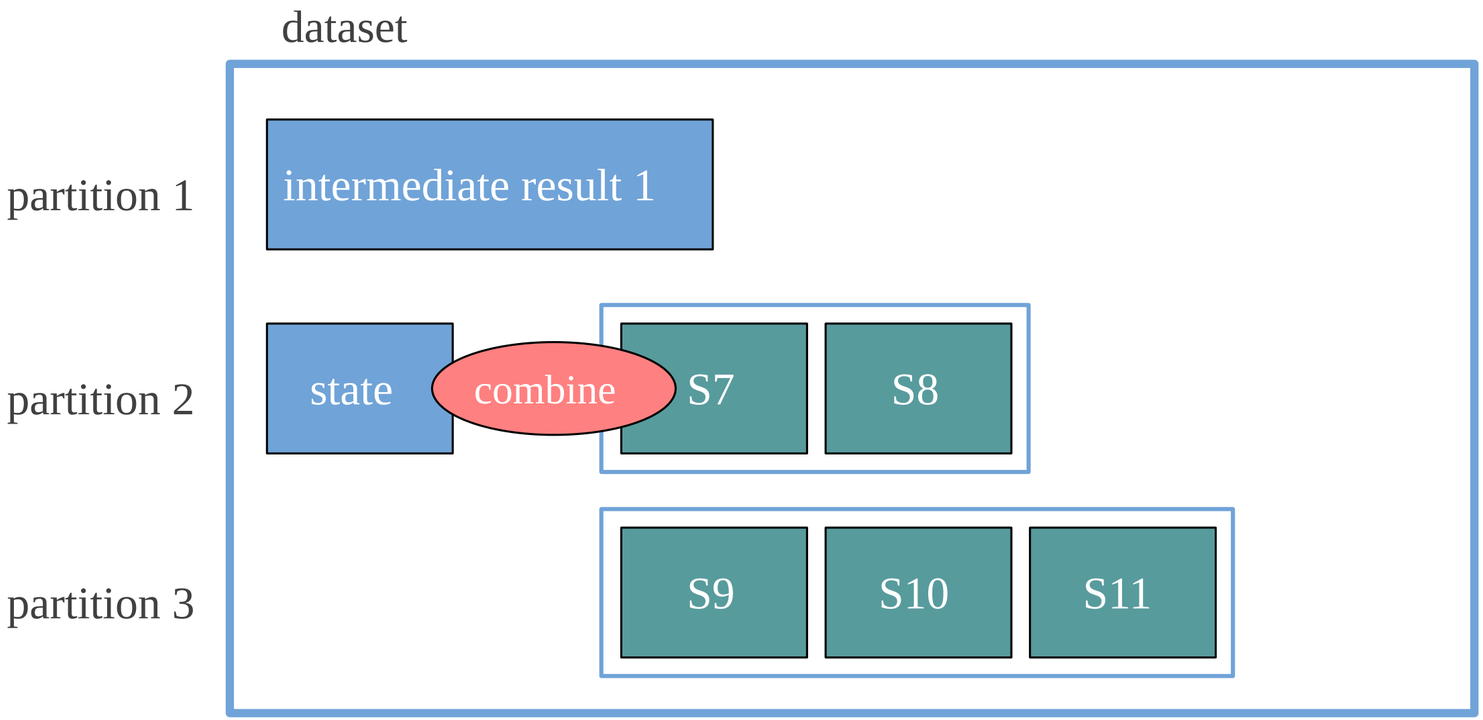
Fold partitions



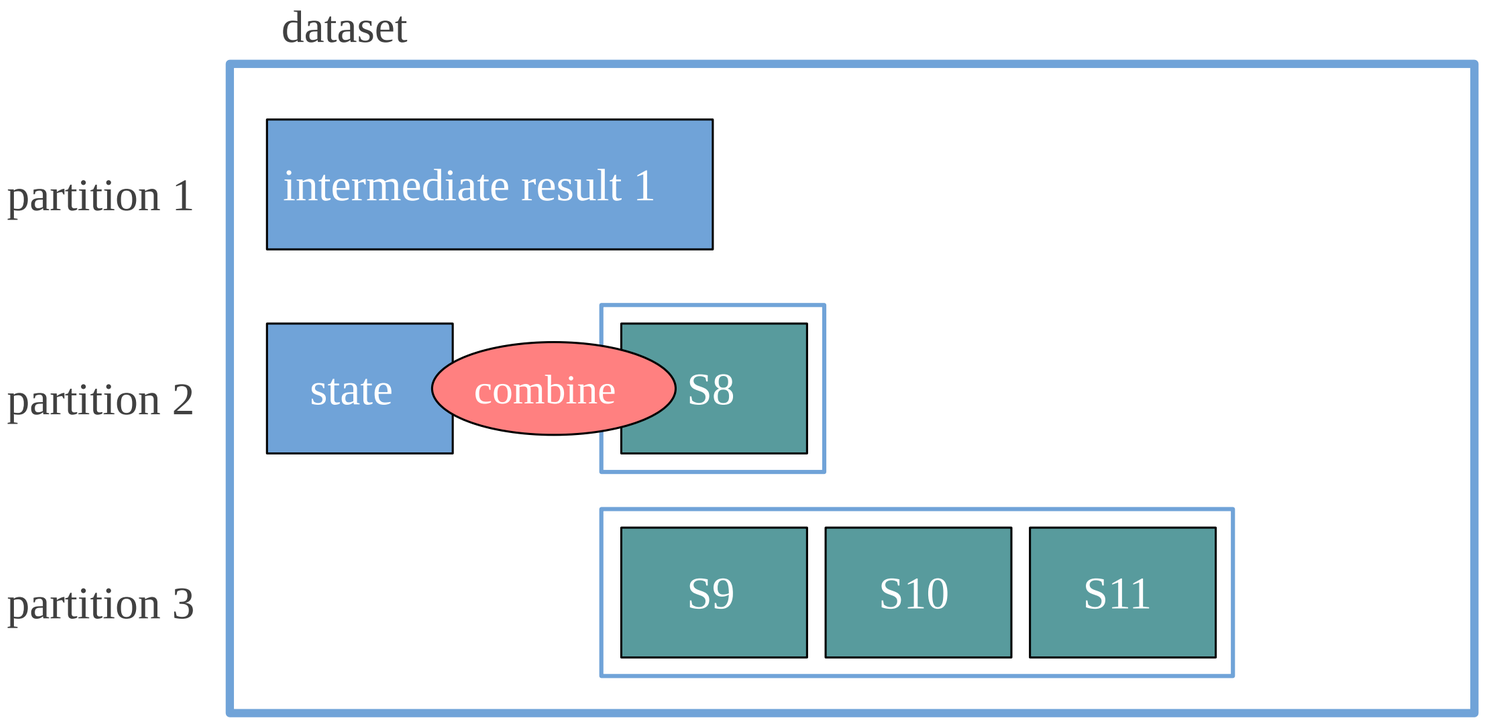
Fold partitions



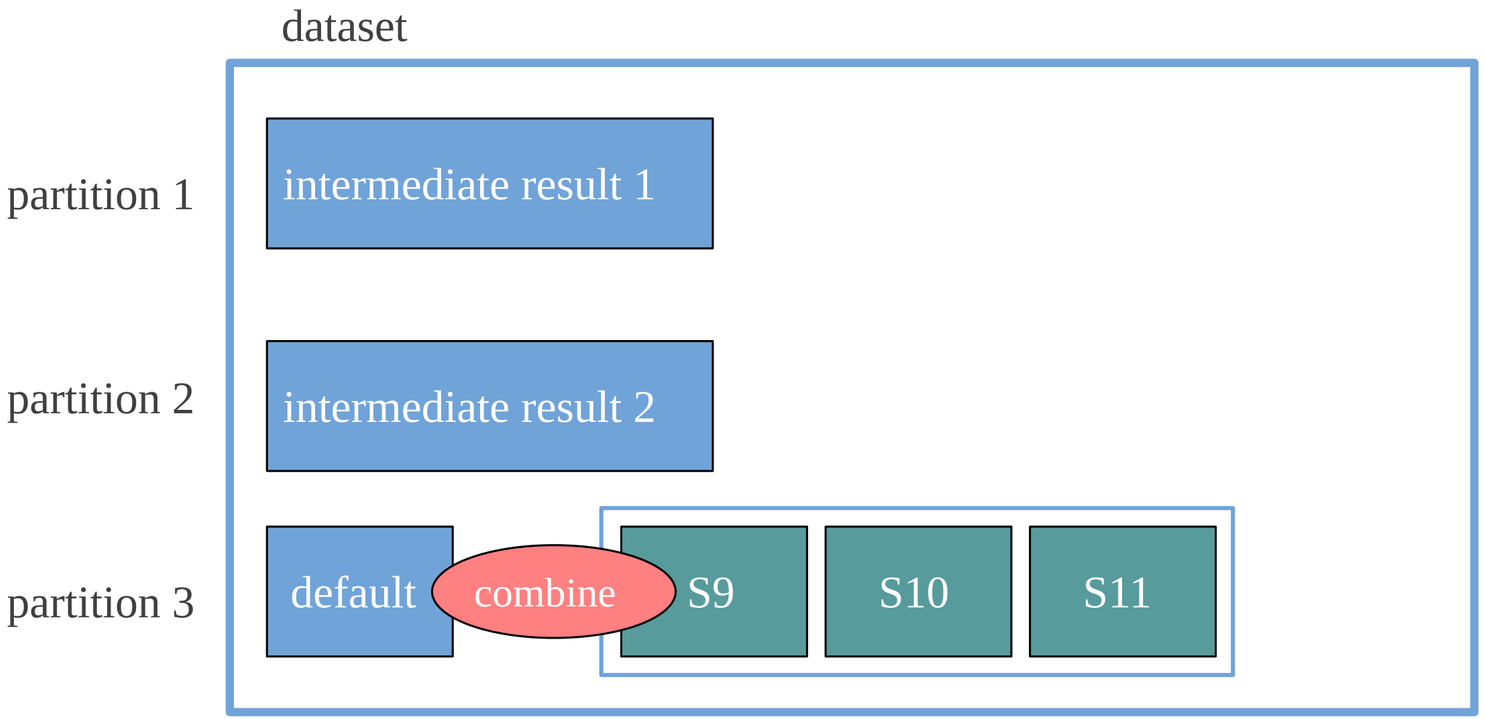
Fold partitions



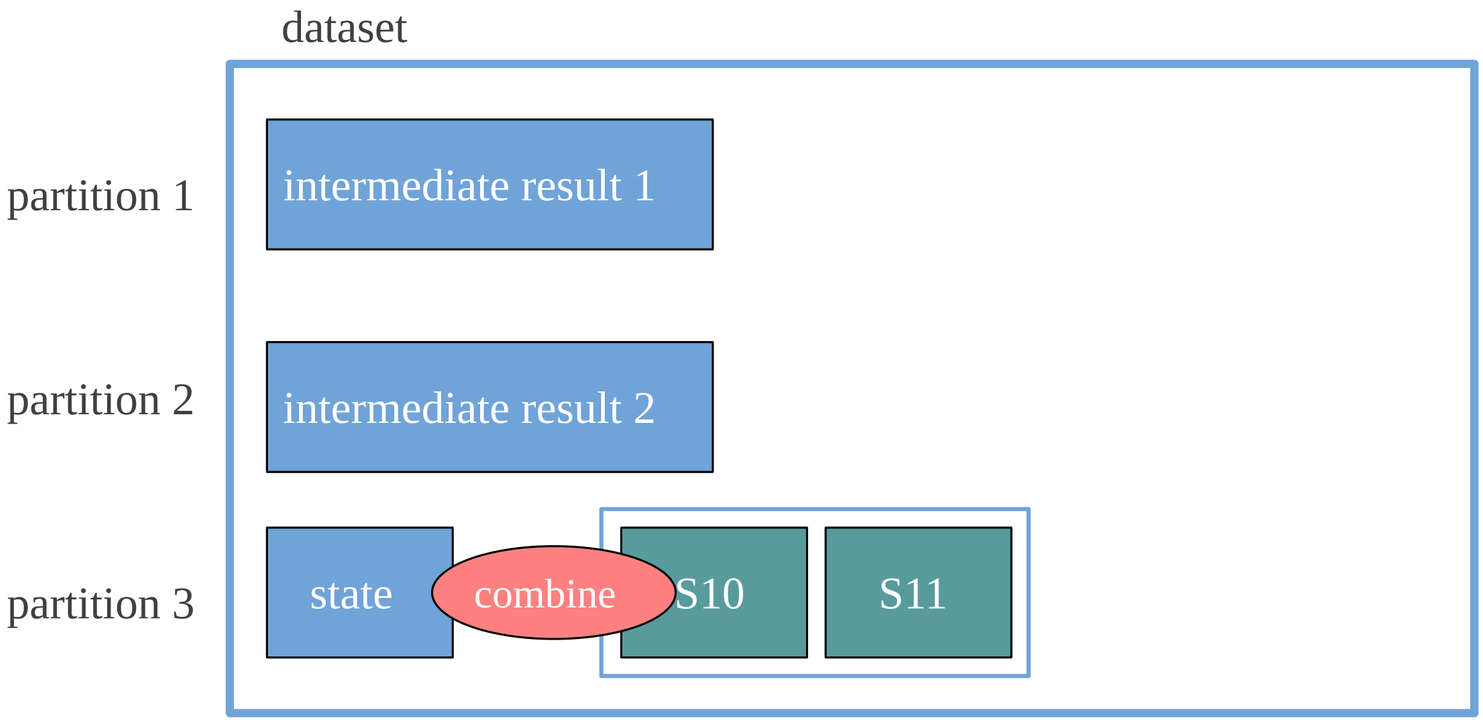
Fold partitions



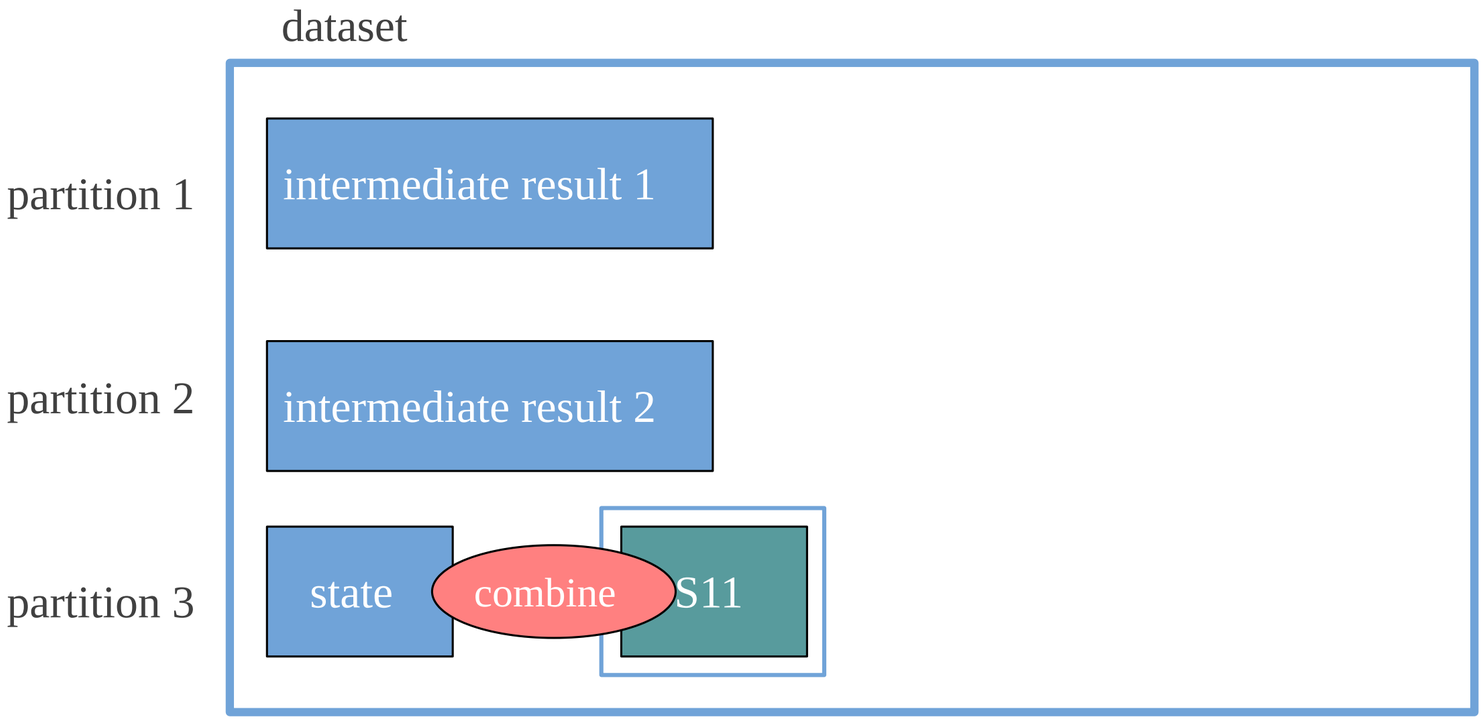
Fold partitions



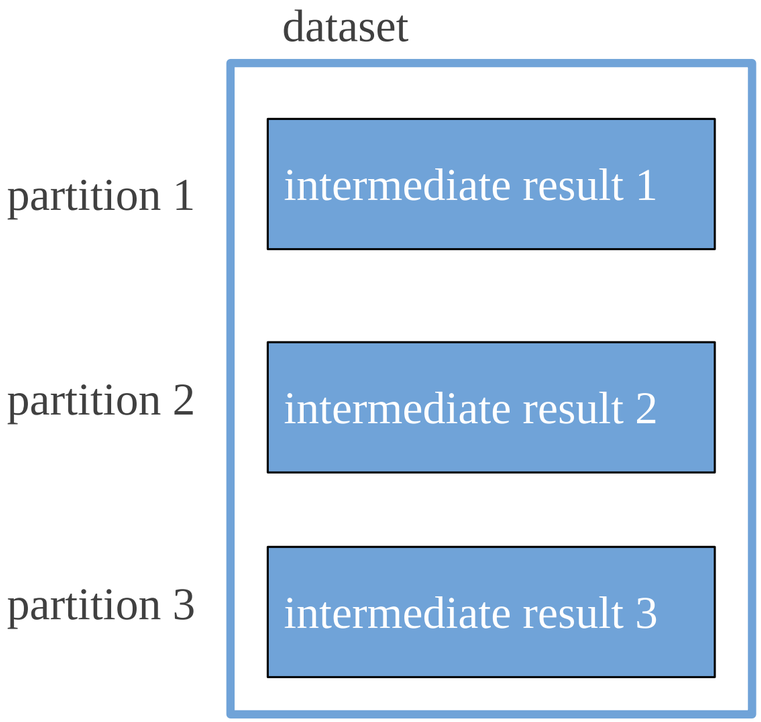
Fold partitions



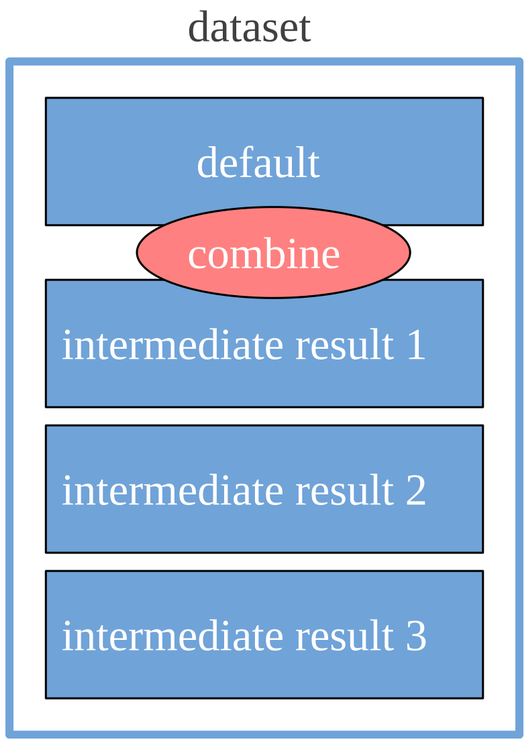
Fold partitions



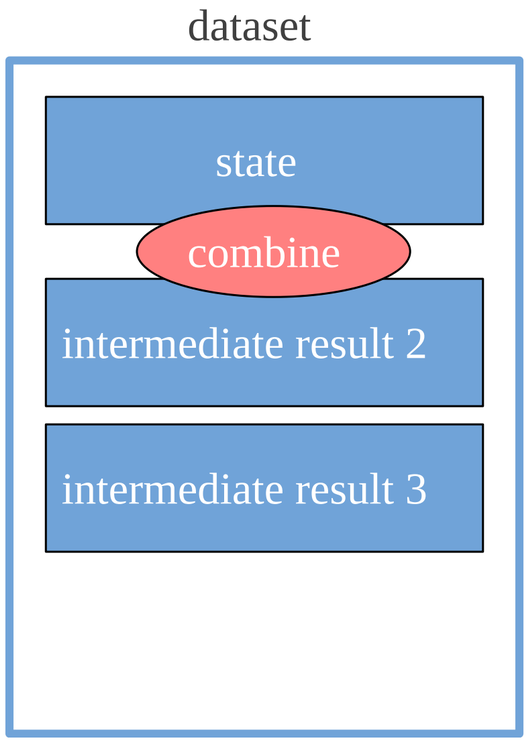
Fold partitions



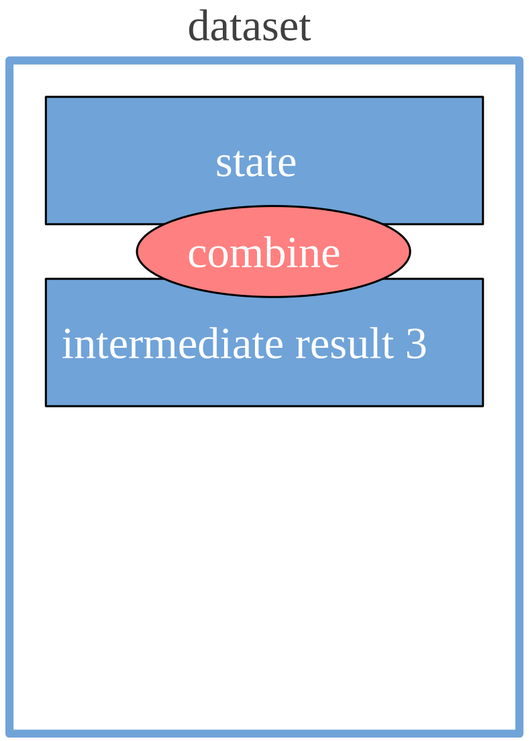
Fold intermediate results



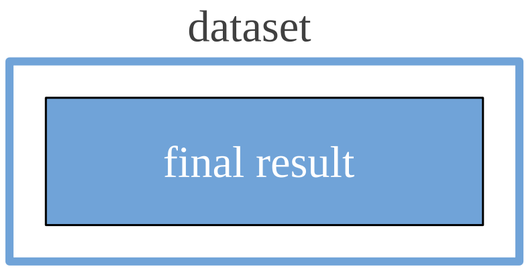
Fold intermediate results



Fold intermediate results



Fold intermediate results



TemperatureExercises.scala