



DEEP  
LEARNING  
INSTITUTE



DLI Accelerated Data Science Teaching Kit

# Lecture 12.2 - How HBase Scales Up Storage



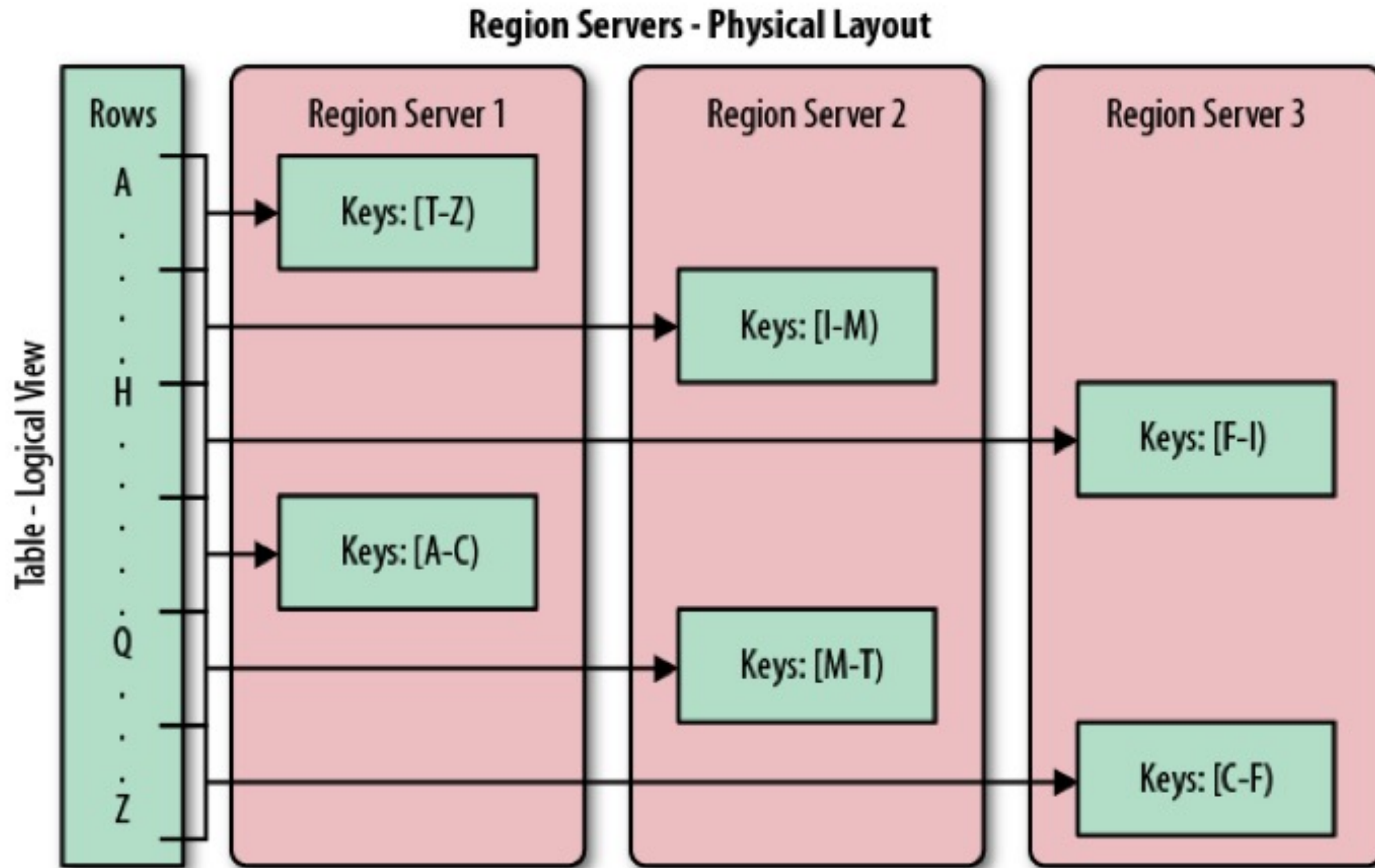
The Accelerated Data Science Teaching Kit is licensed by NVIDIA, Georgia Institute of Technology, and Prairie View A&M University under the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

# How HBase scales up storage?

Automatically **divide** contiguous ranges of rows into **regions**

Start with **one** region, split into two when getting too large, and so on.

# How HBase scales up storage?



Excellent Summary: <http://blog.cloudera.com/blog/2013/04/how-scaling-really-works-in-apache-hbase/>

# Concise way to describe all these?

## HBase data model

- Sparse, distributed, persistent, multidimensional map
- Indexed by **row** key + **column** key + **timestamp**

(Table, RowKey, Family, Column, Timestamp) → Value



DEEP  
LEARNING  
INSTITUTE



DLI Accelerated Data Science Teaching Kit

# Thank You