Understanding Databricks/Spark Performance Tuning

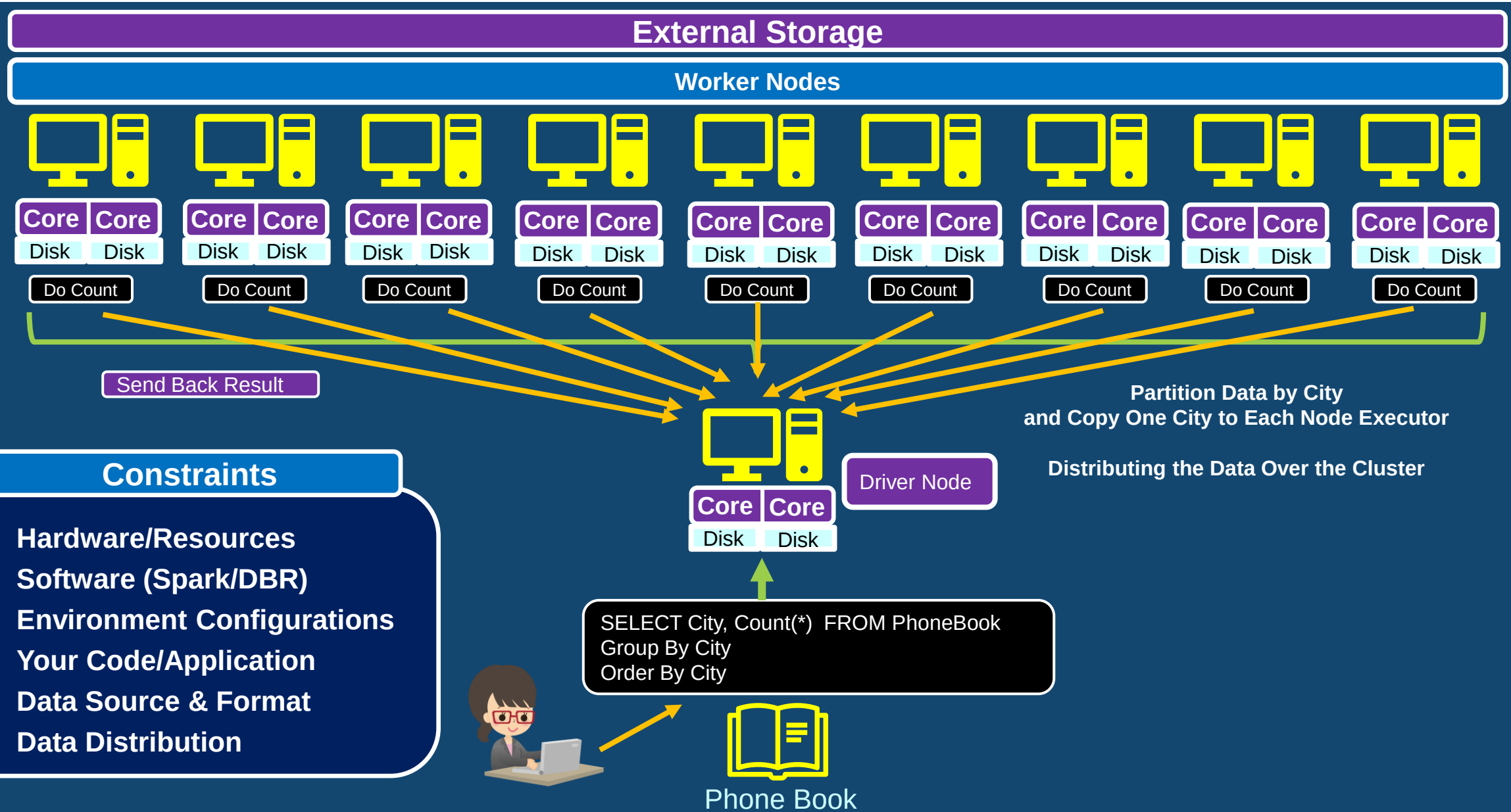
Lesson 01: The Spark Architecture & Bottlenecks

by Bryan Cafferky from my YouTube channel

Where We're Going?

- Databricks/Apache Spark Architecture
- Performance Bottlenecks
- Tuning Options Vary By Platform
- Wrap Up

Spark Cluster



Constraints

- Hardware/Resources
- Software (Spark/DBR)
- Environment Configurations
- Your Code/Application
- Data Source & Format
- Data Distribution

Options Vary By Platform

Delta Lake vs. Non Delta Lake

Open-Source Spark

Databricks

Wrapping Up

- **Databricks/Apache Spark Architecture**
- **Perfo**
- **Tuning Options Vary By Platform**

Thank You!