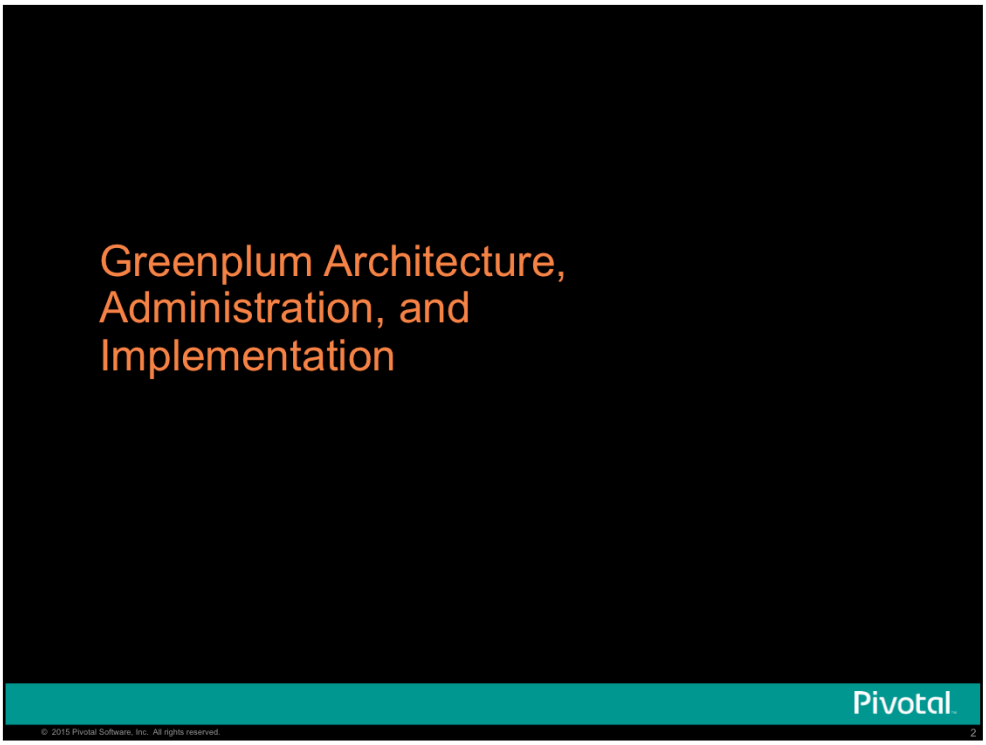
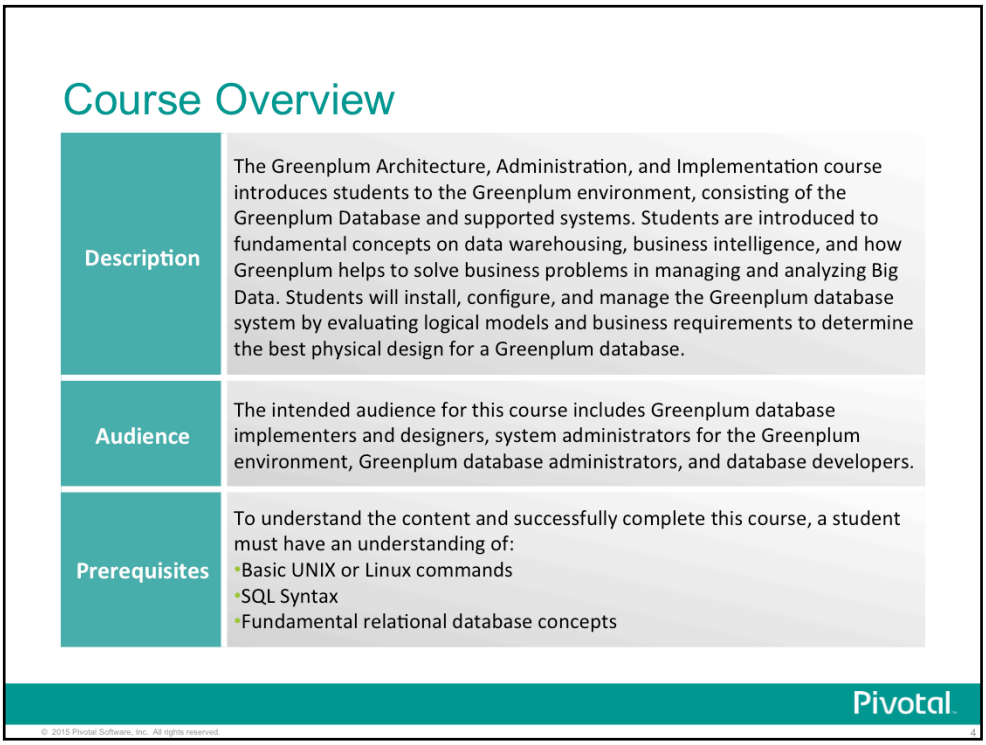

Pivotal
A NEW PLATFORM FOR A NEW ERA



Welcome to the Greenplum Architecture, Administration, and Implementation.

Revision Date: May 2015
Revision Number: MR-1CN-GRNADM.4.3.4.



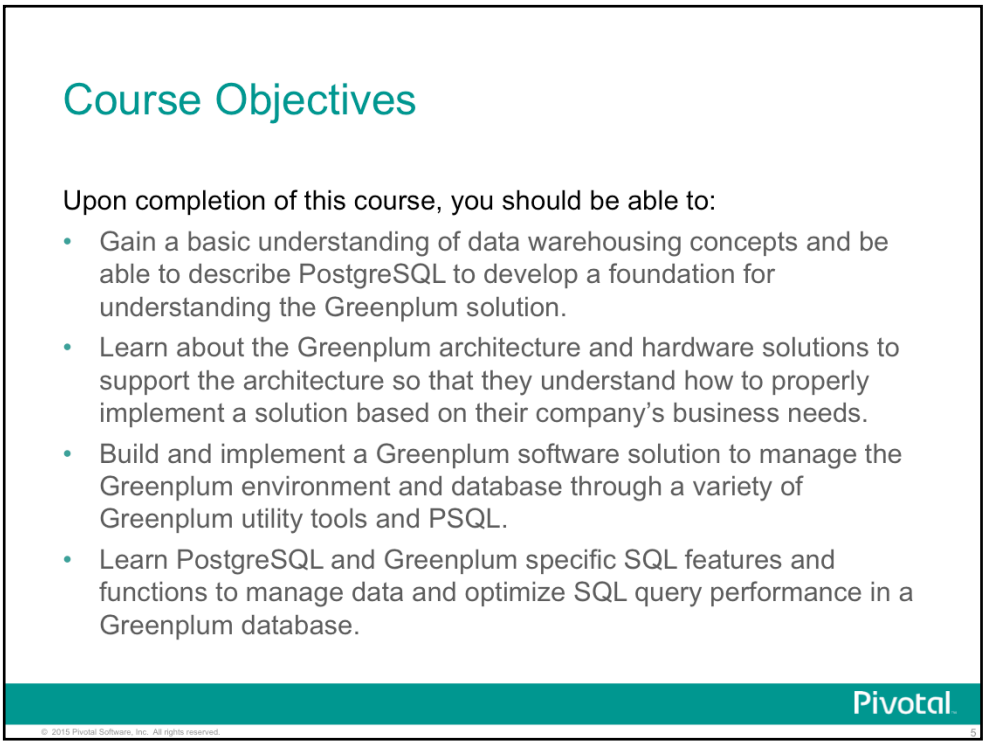
Course Overview

The Greenplum Architecture, Administration, and Implementation course provides you with the foundation you need to interact with, manage, and develop for your Greenplum environment, as an administrator, implementer, or developer. You will gain insight into the benefits that Greenplum offers to your organization in managing and analyzing your data.

Once you install and configure Greenplum, you will manage the data by evaluating logical models and business requirements to determine the best physical design for your database. You will consider specific performance impacts to retrieving your data based on how your data is stored within the database.

To successfully complete this course, you should have:

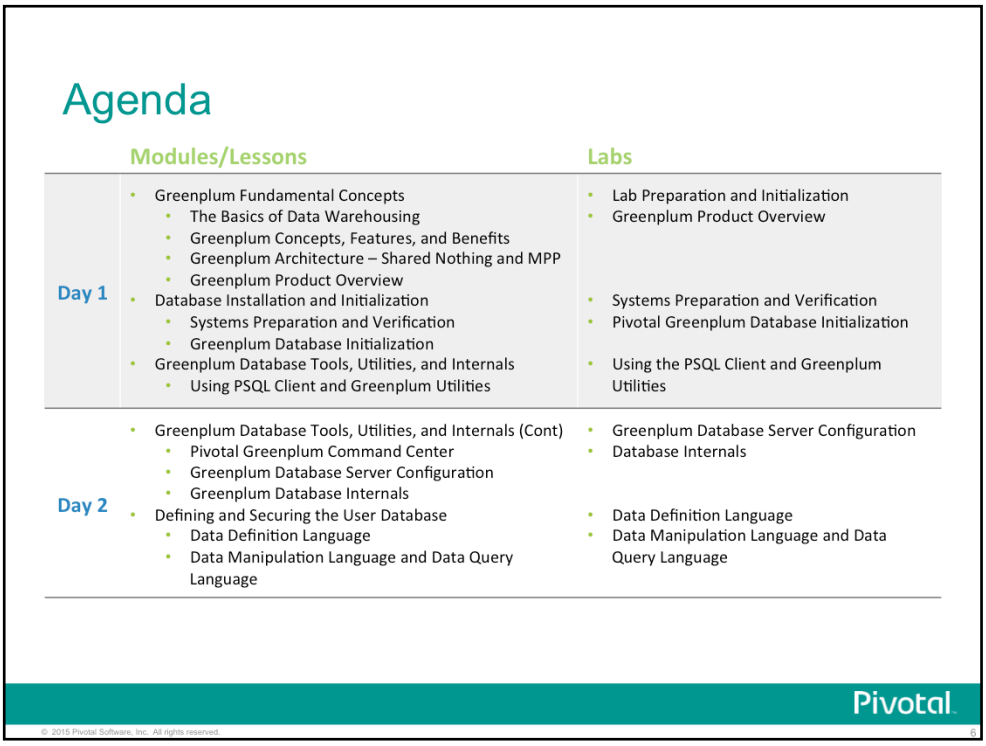
- A working knowledge of basic UNIX or Linux commands to help you navigate the Linux environment
- Some basic understanding of SQL and SQL syntax
- Fundamental relational database concepts



Course Objectives

On completion of the course, you should be able to:

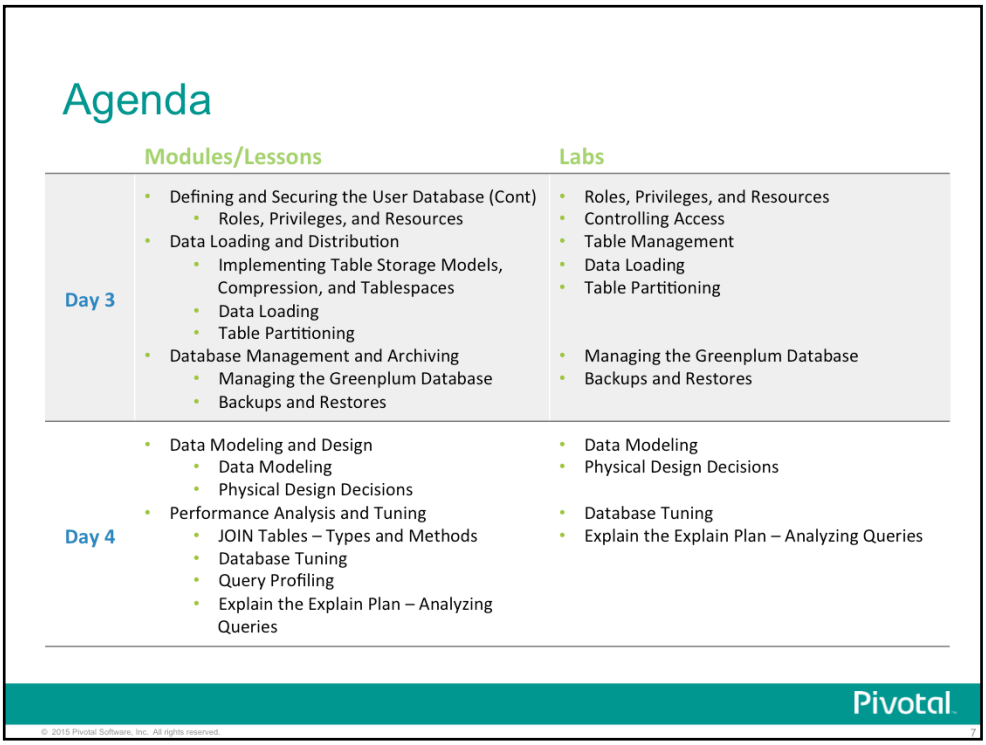
- Identify and explain data warehousing concepts and define what PostgreSQL is as well as its relationship to Greenplum
- Identify the major components within the Greenplum architecture and highlight the importance of certain hardware solutions that support the architecture
- Build and implement a Greenplum software solution in to manage the Greenplum database, database instances, and data using a variety of Greenplum utility tools and the PSQL client
- Use both PostgreSQL and Greenplum-specific SQL commands and functions to manage data and optimize SQL query performance in a Greenplum database



Agenda

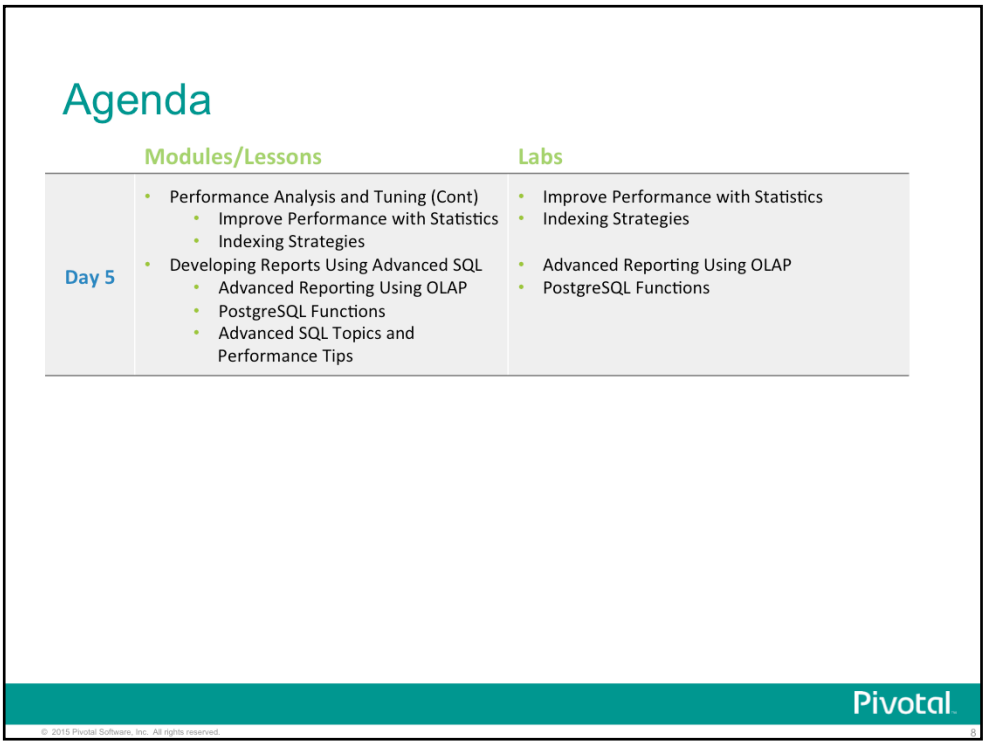
The course is a mixture of lecture and lab over a period of five days. The course is divided into eight modules which define specific high-level related concepts and activities for implementing and managing Greenplum. These modules are:

- **Greenplum Fundamental Concepts** – You will explore the concepts of data warehousing and how these concepts apply to Greenplum. You will then examine the features and benefits of implementing Greenplum in your environment. Finally, you will examine the concepts of shared nothing and MPP to understand how they help Greenplum to scale as well as it does.
- **Database Installation and Initialization** – A solid hardware solution greatly improves your Greenplum queries. You will take an in-depth look at the Greenplum product itself and examine these hardware solutions. You will prepare a system for Greenplum installation, perform the installation, and verify that all requirements are met before proceeding with initializing your Greenplum database. A variety of tools will allow you to connect to and manage the Greenplum environment. Once you learn how to use these tools, you will manage your Greenplum database instance.
- **Greenplum Database Tools, Utilities, and Internals** – Commonly used tools and utilities are examined in this module. In this module, you will examine and use utilities that provide access to the database and configure, monitor, and manage various aspects of the database.
- **Defining and Securing the User Database** – On installing the database, you must have a solid understanding of the database objects you will manage and the syntax available for managing these objects and your data. You will manage workload resources to protect not only data, but also the resources within the database environment. To do that, you will secure the environment by creating appropriate privileges and roles.



Agenda (continued)

- **Data Loading and Distribution** – Once the database has been created, you will load the data. You must make decisions on how the data will be distributed. You will therefore need to understand what table partitioning is and how it can aide you.
- **Database Management and Archive** – You will learn commands that will aid you in managing the database instance as well as creating data backups and restoring the data as necessary.
- **Data Modeling and Design** – While earlier modules examined data loading and some data distribution techniques, you will more closely examine how to model your data and make key physical design decisions which will impact how the data is distributed in your environment.
- **Performance Analysis and Tuning** – Another in-depth subject, this module takes a closer look at specific commands and methods that can impact how your data is retrieved. You look at methods of examining your queries so that you develop appropriate queries to maximize performance gained based on how your data is distributed. You examine how to improve statistical analysis of your data to improve the overall query performance. You also examine whether or not to index your tables, implement OLAP methods in your queries, create PostgreSQL functions, and apply specific performance tuning tips for your data and your queries.



Agenda (continued)

- **Developing Reports Using Advanced SQL** – Analysts use a combination of custom ad-hoc queries and Business intelligence tools to generate reports. The module highlights the various OLAP methods available to generate reports as well as how to build functions to access commonly used methods.

This slide is intentionally left blank.

Pivotal