



Teradata

tutorialspoint

SIMPLY EASY LEARNING

www.tutorialspoint.com



<https://www.facebook.com/tutorialspointindia>



<https://twitter.com/tutorialspoint>

About the Tutorial

Teradata is a popular Relational Database Management System (RDBMS) suitable for large data warehousing applications. It is capable of handling large volumes of data and is highly scalable. This tutorial provides a good understanding of Teradata Architecture, various SQL commands, Indexing concepts and Utilities to import/export data.

Audience

This tutorial is designed for software professionals who are willing to learn Teradata concepts and become a Teradata developer. By the end of this tutorial, you will have gained intermediate level of expertise in Teradata.

Prerequisites

You should have a basic understanding of Relational concepts and basic SQL. It will be good if you have worked with any other RDBMS product.

Copyright & Disclaimer

© Copyright 2018 by Tutorials Point (I) Pvt. Ltd.

All the content and graphics published in this e-book are the property of Tutorials Point (I) Pvt. Ltd. The user of this e-book is prohibited to reuse, retain, copy, distribute or republish any contents or a part of contents of this e-book in any manner without written consent of the publisher.

We strive to update the contents of our website and tutorials as timely and as precisely as possible, however, the contents may contain inaccuracies or errors. Tutorials Point (I) Pvt. Ltd. provides no guarantee regarding the accuracy, timeliness or completeness of our website or its contents including this tutorial. If you discover any errors on our website or in this tutorial, please notify us at contact@tutorialspoint.com

Table of Contents

About the Tutorial	i
Audience.....	i
Prerequisites.....	i
Copyright & Disclaimer	i
Table of Contents	ii
PART 1: TERADATA BASICS.....	1
1. Teradata - Introduction.....	2
What is Teradata?	2
History of Teradata	2
Features of Teradata	2
2. Teradata – Installation	4
Installation Steps for Windows.....	4
Starting BTEQ.....	8
3. Teradata – Architecture	9
Components of Teradata.....	9
Storage Architecture	10
Retrieval Architecture	11
4. Teradata – Relational Concepts.....	12
5. Teradata – Data Types	14
6. Teradata – Tables.....	16
Table Types.....	16
Create Table	16
Alter Table	18
Drop Table	19
7. Teradata – Data Manipulation	20
Insert Records.....	20
Insert from Another Table.....	21
Update Records	22
Delete Records	23
8. Teradata – SELECT Statement	24
WHERE Clause	25
ORDER BY.....	25
GROUP BY	26
9. Teradata – Logical & Conditional Operators.....	27
BETWEEN.....	28
IN	28
NOT IN	29
10. Teradata – SET Operators	30
UNION.....	30
UNION ALL.....	31

INTERSECT	32
MINUS/EXCEPT	33
11. Teradata – String Manipulation	35
12. Teradata – Date/Time Functions.....	36
Date Storage.....	36
EXTRACT	36
INTERVAL.....	37
13. Teradata – Built-in Functions	39
14. Teradata – Aggregate Functions.....	40
15. Teradata – CASE & COALESCE	42
CASE Expression	42
COALESCE	43
NULLIF.....	44
16. Teradata – Primary Index.....	45
Unique Primary Index (UPI)	45
Non Unique Primary Index (NUPI).....	46
17. Teradata – Joins	47
INNER JOIN	47
OUTER JOIN	48
CROSS JOIN	50
18. Teradata – SubQueries.....	51
PART 2: TERADATA ADVANCED	53
19. Teradata – Table Types	54
Derived Table.....	54
Volatile Table.....	55
Global Temporary Table	55
20. Teradata – Space Concepts	57
Permanent Space	57
Spool Space	57
Temp Space	57
21. Teradata – Secondary Index.....	58
Unique Secondary Index (USI)	58
Non Unique Secondary Index (NUSI).....	58
22. Teradata – Statistics.....	59
Collecting Statistics.....	59
Viewing Statistics.....	60
23. Teradata – Compression	61
24. Teradata – EXPLAIN	62

Examples of EXPLAIN	62
Full Table Scan (FTS)	62
Unique Primary Index	63
Unique Secondary Index.....	63
Additional Terms	64
25. Teradata – Hashing Algorithm.....	65
26. Teradata – JOIN INDEX.....	67
Single Table Join Index.....	67
Multi Table Join Index	69
Aggregate Join Index	69
27. Teradata – Views	71
Create a View	71
Using Views	72
Modifying Views	72
Drop View	73
28. Teradata – Macros	74
Create Macros	74
Executing Macros	75
Parameterized Macros	76
Executing Parameterized Macros.....	76
29. Teradata – Stored Procedure	77
Creating Procedure.....	77
Executing Procedures	78
30. Teradata – JOIN Strategies	80
Join Methods	80
Merge Join	80
Nested Join	82
Product Join.....	82
31. Teradata – Partitioned Primary Index	83
32. Teradata – OLAP Functions	86
33. Teradata – Data Protection	89
Transient Journal	89
Fallback.....	89
Down AMP Recovery Journal	90
Cliques	90
Hot Standby Node	90
RAID	91
34. Teradata – User Management.....	92
Users.....	92
Accounts	93
Grant Privileges	93
Revoke Privileges.....	94
35. Teradata – Performance Tuning.....	95

36. Teradata – FastLoad97
 How FastLoad Works 97
 Executing a FastLoad Script 98
 FastLoad Terms..... 99

37. Teradata – MultiLoad.....100
 Limitation..... 100
 How MultiLoad Works..... 100
 Executing a MultiLoad Script 102

38. Teradata – FastExport103
 Executing a FastExport Script 104
 FastExport Terms..... 104

39. Teradata – BTEQ 105

40. Teradata – Questions & Answers 108

Part 1: Teradata Basics

1. Teradata - Introduction

What is Teradata?

Teradata is one of the popular Relational Database Management System. It is mainly suitable for building large scale data warehousing applications. Teradata achieves this by the concept of parallelism. It is developed by the company called Teradata.

History of Teradata

Following is a quick summary of the history of Teradata, listing major milestones.

- 1979** – Teradata was incorporated
- 1984** – Release of first database computer DBC/1012
- 1986** – *Fortune* magazine names Teradata as 'Product of the Year'
- 1999** – Largest database in the world using Teradata with 130 Terabytes
- 2002** – Teradata V2R5 released with Partition Primary Index and compression
- 2006** – Launch of Teradata Master Data Management solution
- 2008** – Teradata 13.0 released with Active Data Warehousing
- 2011** – Acquires Teradata Aster and enters into Advanced Analytics Space
- 2012** – Teradata 14.0 introduced
- 2014** – Teradata 15.0 introduced

Features of Teradata

Following are some of the features of Teradata:

- **Unlimited Parallelism:** Teradata database system is based on Massively Parallel Processing (MPP) Architecture. MPP architecture divides the workload evenly across the entire system. Teradata system splits the task among its processes and runs them in parallel to ensure that the task is completed quickly.
- **Shared Nothing Architecture:** Teradata's architecture is called as Shared Nothing Architecture. Teradata Nodes, its Access Module Processors (AMPs) and the disks associated with AMPs work independently. They are not shared with others.
- **Linear Scalability:** Teradata systems are highly scalable. They can scale up to 2048 Nodes. For example, you can double the capacity of the system by doubling the number of AMPs.
- **Connectivity:** Teradata can connect to Channel-attached systems such as Mainframe or Network-attached systems.

- **Mature Optimizer:** Teradata optimizer is one of the matured optimizer in the market. It has been designed to be parallel since its beginning. It has been refined for each release.
- **SQL:** Teradata supports industry standard SQL to interact with the data stored in tables. In addition to this, it provides its own extension.
- **Robust Utilities:** Teradata provides robust utilities to import/export data from/to Teradata system such as FastLoad, MultiLoad, FastExport and TPT.
- **Automatic Distribution:** Teradata automatically distributes the data evenly to the disks without any manual intervention.

2. Teradata – Installation

Teradata provides Teradata express for VMWARE which is a fully operational Teradata virtual machine. It provides up to 1 terabyte of storage. Teradata provides both 40GB and 1TB version of VMware.

Prerequisites

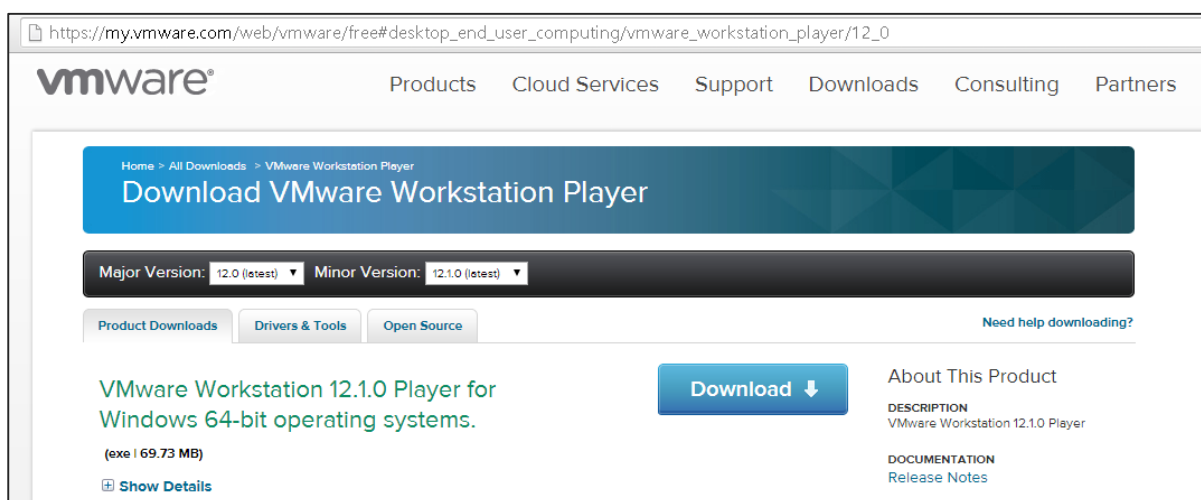
Since the VM is 64 bit, your CPU must support 64-bit.

Installation Steps for Windows

Step 1: Download the required VM version from the link, <http://downloads.teradata.com/download/database/teradata-express-for-vmware-player>

Step 2: Extract the file and specify the target folder.

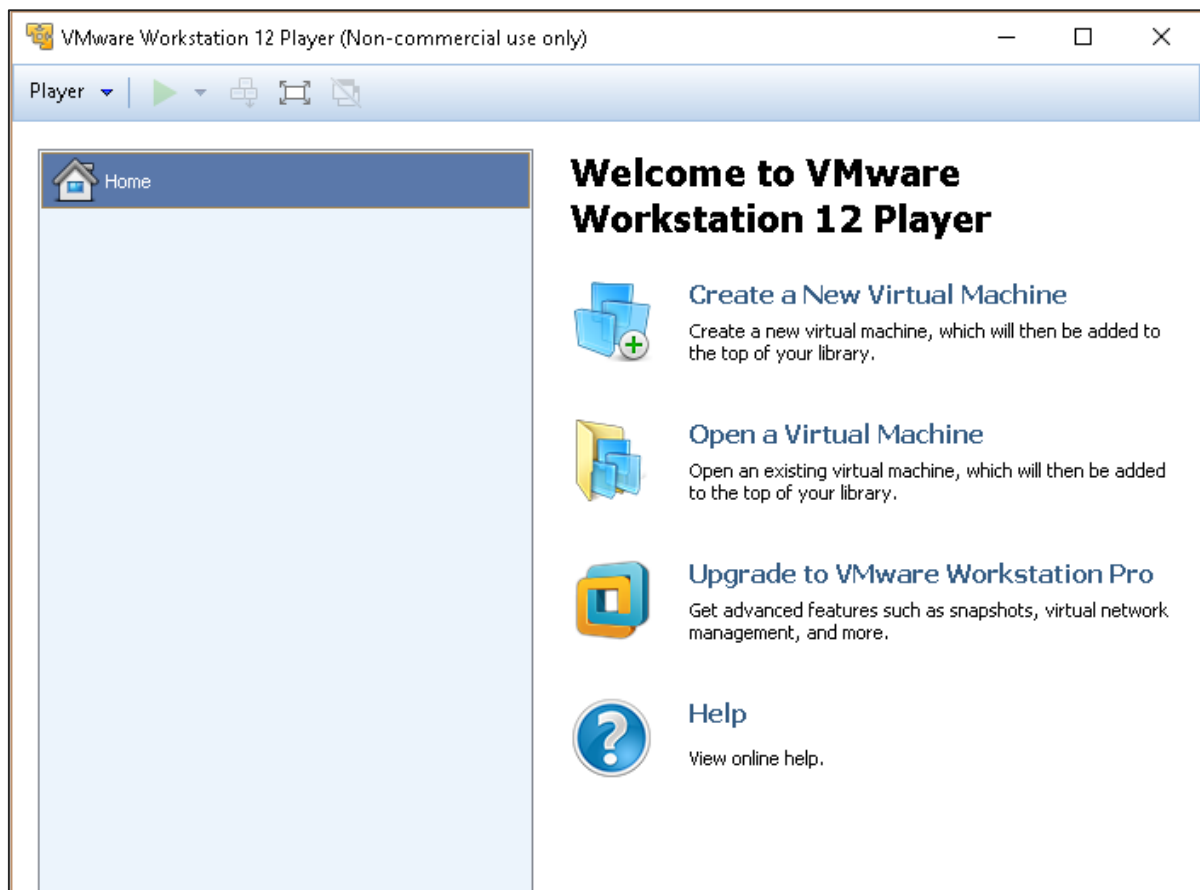
Step 3: Download the VMWare Workstation player from the link, <https://my.vmware.com/web/vmware/downloads>. It is available for both Windows and Linux. Download the VMWARE workstation player for Windows.



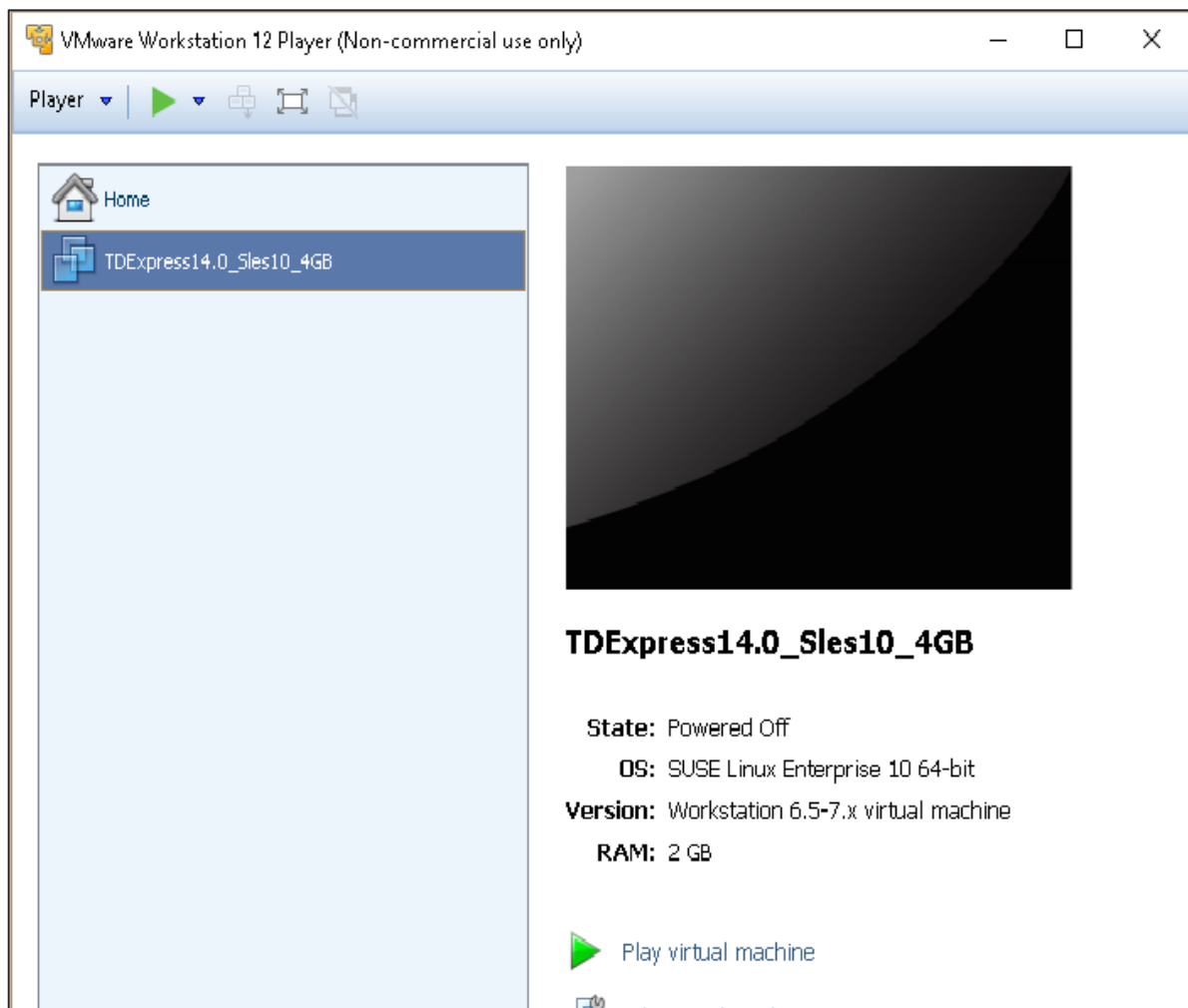
Step 4: Once the download is complete, install the software.

Step 5: After the installation is complete, run the VMWARE client.

Step 6: Select 'Open a Virtual Machine'. Navigate through the extracted Teradata VMWare folder and select the file with extension .vmdk.

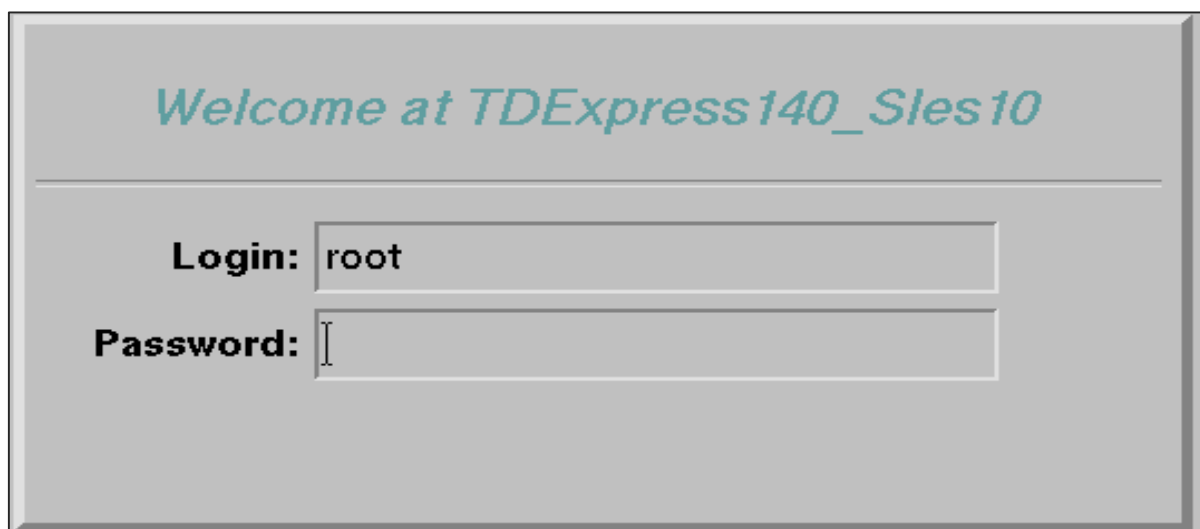


Step 7: Teradata VMWare is added to the VMWare client. Select the added Teradata VMware and click 'Play Virtual Machine'.

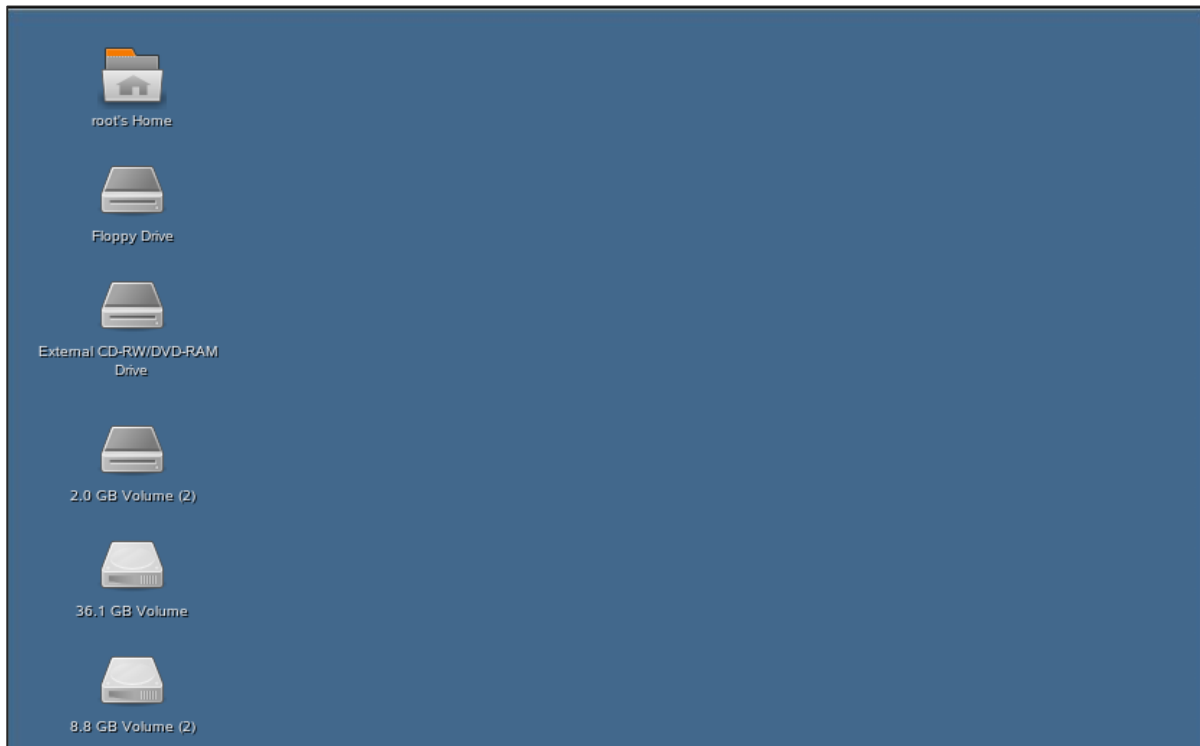


Step 8: If you get a popup on software updates, you can select 'Remind Me Later'.

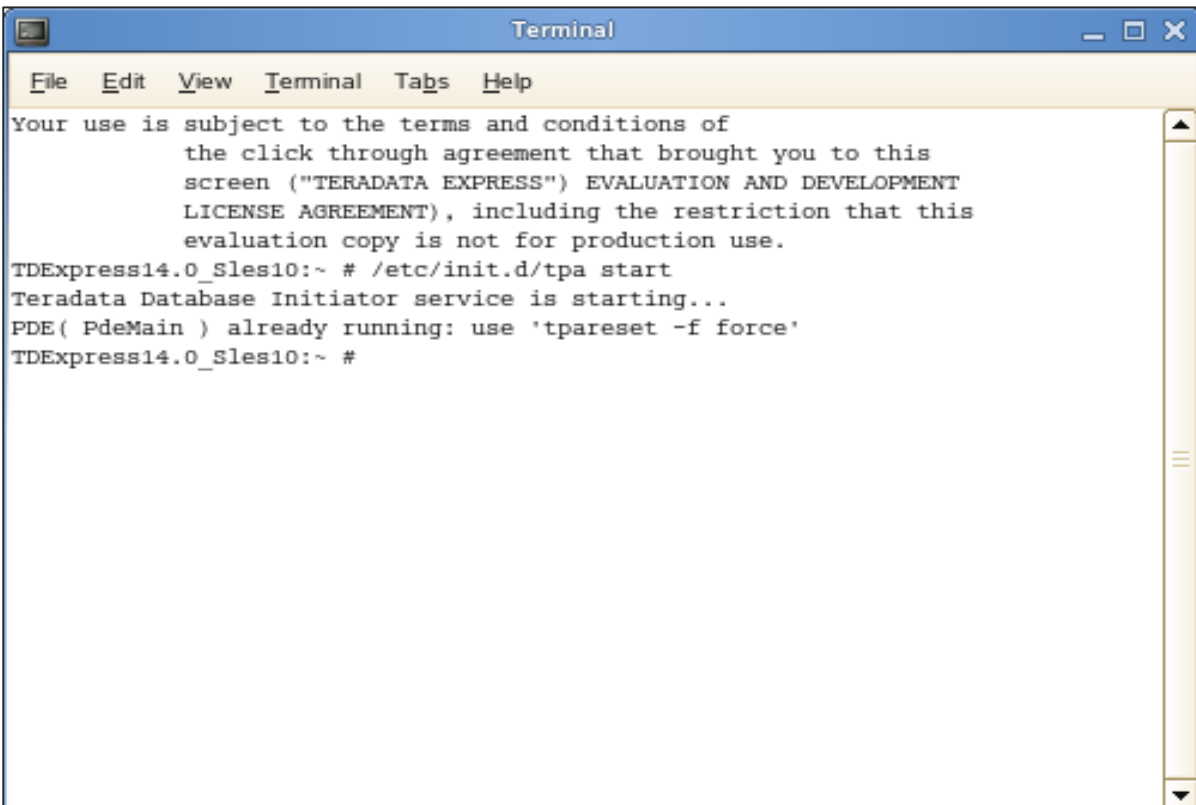
Step 9: Enter the user name as root, press tab and enter password as root and again press Enter.



Step 10: Once the following screen appears on the desktop, double-click on 'root's home'. Then double-click on 'Genome's Terminal'. This will open the Shell.



Step 11: From the following shell, enter the command `/etc/init.d/tpa start`. This will start the Teradata server.

A screenshot of a terminal window titled 'Terminal'. The window has a menu bar with 'File', 'Edit', 'View', 'Terminal', 'Tabs', and 'Help'. The terminal content shows a copyright notice for Teradata Express, followed by the command `/etc/init.d/tpa start` being executed. The output indicates that the Teradata Database Initiator service is starting, and that the PDE (PdeMain) is already running. The prompt returns to the user.

```
Terminal
File Edit View Terminal Tabs Help
Your use is subject to the terms and conditions of
the click through agreement that brought you to this
screen ("TERADATA EXPRESS") EVALUATION AND DEVELOPMENT
LICENSE AGREEMENT), including the restriction that this
evaluation copy is not for production use.
TDEExpress14.0_Sles10:~ # /etc/init.d/tpa start
Teradata Database Initiator service is starting...
PDE( PdeMain ) already running: use 'tpareset -f force'
TDEExpress14.0_Sles10:~ #
```

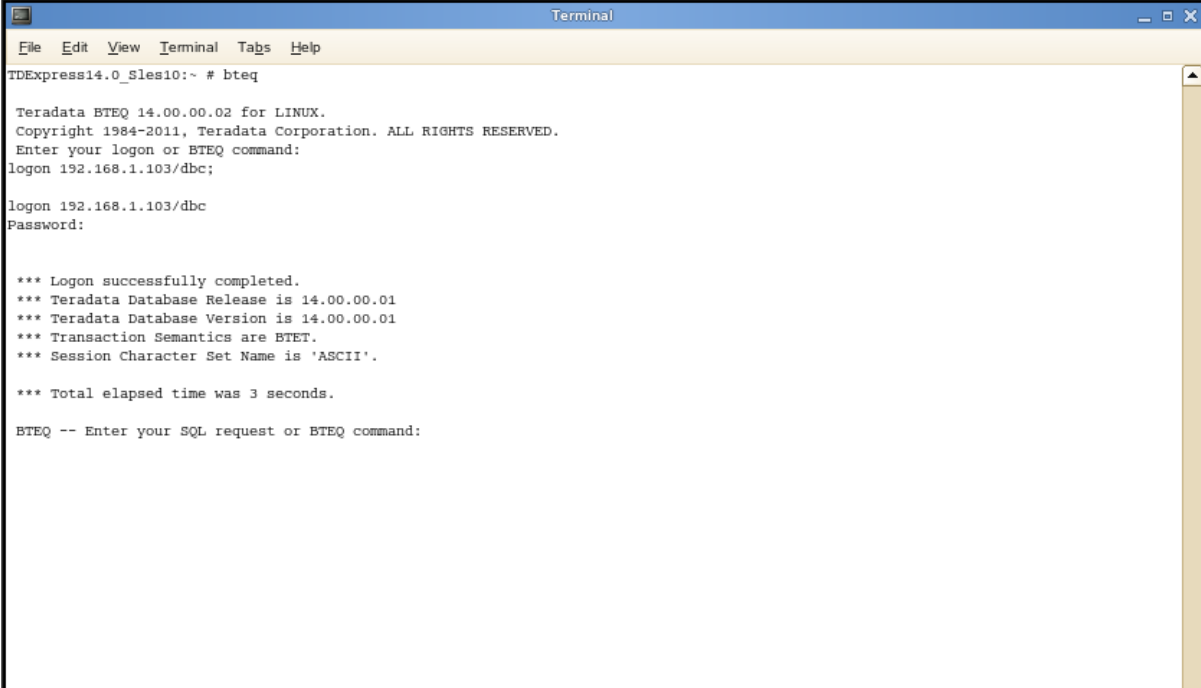
Starting BTEQ

BTEQ utility is used to submit SQL queries interactively. Following are the steps to start BTEQ utility.

Step 1: Enter the command `/sbin/ifconfig` and note down the IP address of the VMWare.

Step 2: Run the command `bteq`. At the logon prompt, enter the command.

Logon `<ipaddress>/dbc,dbc;` and enter At the password prompt, enter password as `dbc;`

A terminal window titled "Terminal" showing the execution of the BTEQ utility. The prompt is "TDExpress14.0_Sles10:~ # bteq". The output includes: "Teradata BTEQ 14.00.00.02 for LINUX. Copyright 1984-2011, Teradata Corporation. ALL RIGHTS RESERVED. Enter your logon or BTEQ command: logon 192.168.1.103/dbc; logon 192.168.1.103/dbc Password: *** Logon successfully completed. *** Teradata Database Release is 14.00.00.01 *** Teradata Database Version is 14.00.00.01 *** Transaction Semantics are BTET. *** Session Character Set Name is 'ASCII'. *** Total elapsed time was 3 seconds. BTEQ -- Enter your SQL request or BTEQ command:"/>

```
TDExpress14.0_Sles10:~ # bteq

Teradata BTEQ 14.00.00.02 for LINUX.
Copyright 1984-2011, Teradata Corporation. ALL RIGHTS RESERVED.
Enter your logon or BTEQ command:
logon 192.168.1.103/dbc;

logon 192.168.1.103/dbc
Password:

*** Logon successfully completed.
*** Teradata Database Release is 14.00.00.01
*** Teradata Database Version is 14.00.00.01
*** Transaction Semantics are BTET.
*** Session Character Set Name is 'ASCII'.

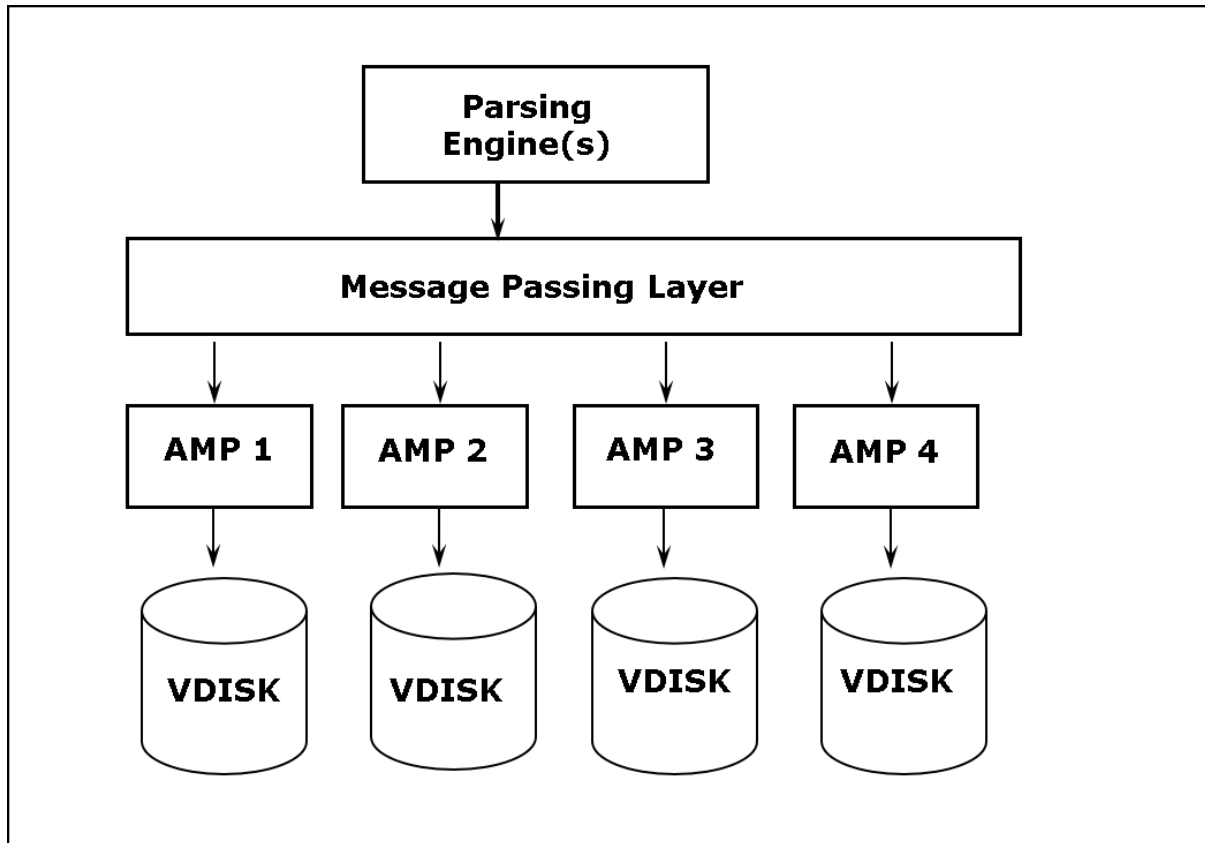
*** Total elapsed time was 3 seconds.

BTEQ -- Enter your SQL request or BTEQ command:
```

You can log into Teradata system using BTEQ and run any SQL queries.

3. Teradata – Architecture

Teradata architecture is based on Massively Parallel Processing (MPP) architecture. The major components of Teradata are Parsing Engine, BYNET and Access Module Processors (AMPs). The following diagram shows the high level architecture of a Teradata Node.



Components of Teradata

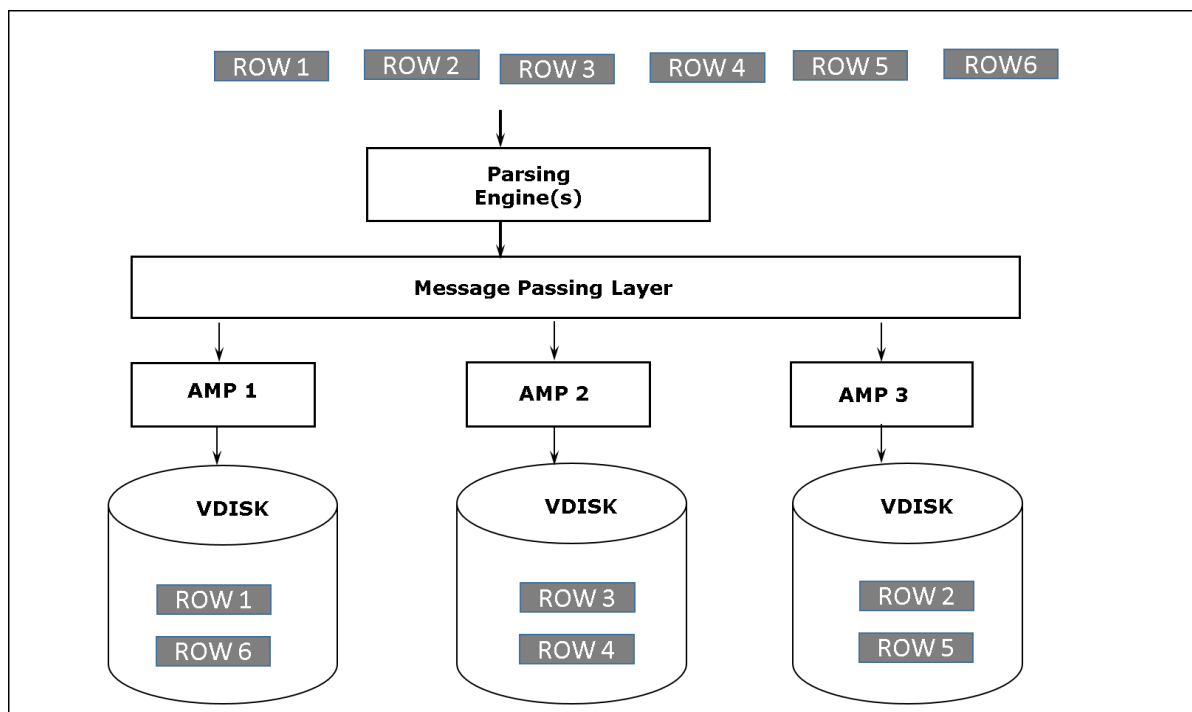
The key components of Teradata are as follows:

- **Node:** It is the basic unit in Teradata System. Each individual server in a Teradata system is referred as a Node. A node consists of its own operating system, CPU, memory, own copy of Teradata RDBMS software and disk space. A cabinet consists of one or more Nodes.
- **Parsing Engine:** Parsing Engine is responsible for receiving queries from the client and preparing an efficient execution plan. The responsibilities of parsing engine are:
 - Receive the SQL query from the client.
 - Parse the SQL query check for syntax errors.
 - Check if the user has required privilege against the objects used in the SQL query.

- Check if the objects used in the SQL actually exists.
 - Prepare the execution plan to execute the SQL query and pass it to BYNET.
 - Receives the results from the AMPs and send to the client.
- **Message Passing Layer:** Message Passing Layer called as BYNET, is the networking layer in Teradata system. It allows the communication between PE and AMP and also between the nodes. It receives the execution plan from Parsing Engine and sends to AMP. Similarly, it receives the results from the AMPs and sends to Parsing Engine.
 - **Access Module Processor (AMP):** AMPs, called as Virtual Processors (vprocs) are the one that actually stores and retrieves the data. AMPs receive the data and execution plan from Parsing Engine, performs any data type conversion, aggregation, filter, sorting and stores the data in the disks associated with them. Records from the tables are evenly distributed among the AMPs in the system. Each AMP is associated with a set of disks on which data is stored. Only that AMP can read/write data from the disks.

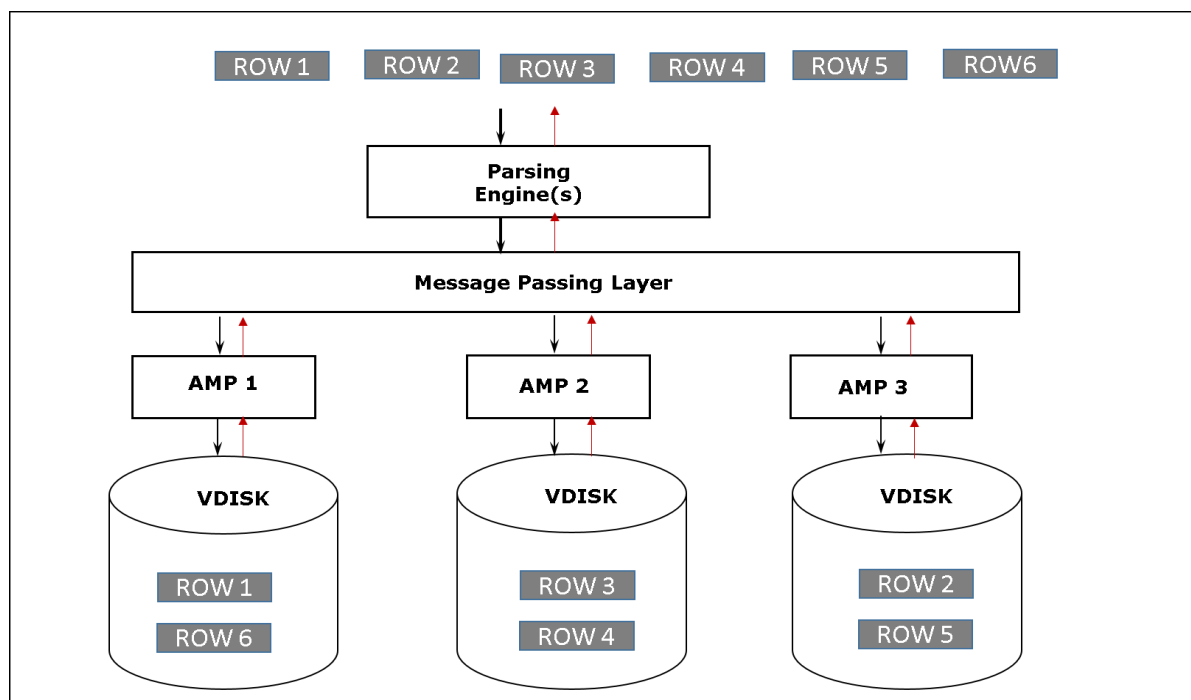
Storage Architecture

When the client runs queries to insert records, Parsing engine sends the records to BYNET. BYNET retrieves the records and sends the row to the target AMP. AMP stores these records on its disks. Following diagram shows the storage architecture of Teradata.



Retrieval Architecture

When the client runs queries to retrieve records, the Parsing engine sends a request to BYNET. BYNET sends the retrieval request to appropriate AMPs. Then AMPs search their disks in parallel and identify the required records and sends to BYNET. BYNET then sends the records to Parsing Engine which in turn will send to the client. Following is the retrieval architecture of Teradata.



End of ebook preview

If you liked what you saw...

Buy it from our store @ <https://store.tutorialspoint.com>