
Guru99 Provides [FREE ONLINE TUTORIAL](#) on Various courses like

[Java](#) | [MIS](#) | [MongoDB](#) | [BigData](#) | [Cassandra](#) | [Web Services](#)

[SQLite](#) | [JSP](#) | [Informatica](#) | [Accounting](#) | [SAP Training](#) | [Python](#)

[Excel](#) | [ASP Net](#) | [HBase](#) | [Testing](#) | [Selenium](#) | [CCNA](#) | [NodeJS](#)

[TensorFlow](#) | [Data Warehouse](#) | [R Programming](#) | [Live Projects](#) | [DevOps](#)

Top 19 Ab initio Interview Questions & Answers

1) Mention what is Abinitio?

“Abinitio” is a latin word meaning “from the beginning.” Abinitio is a tool used to extract, transform and load data. It is also used for data analysis, data manipulation, batch processing, and graphical user interface based parallel processing.

2) Explain what is the architecture of Abinitio?

Architecture of Abinitio includes

- GDE (Graphical Development Environment)
- Co-operating System
- Enterprise meta-environment (EME)
- Conduct-IT

3) Mention what is the role of Co-operating system in Abinitio?

The Abinitio co-operating system provide features like

- Manage and run Abinitio graph and control the ETL processes
- Provide Ab initio extensions to the operating system
- ETL processes monitoring and debugging
- Meta-data management and interaction with the EME

4) Explain what does dependency analysis mean in Abinitio?

In Ab initio, dependency analysis is a process through which the EME examines a project entirely and traces how data is transferred and transformed- from component-to-component, field-by-field, within and between graphs.

5) Explain how Abinitio EME is segregated?

Abinitio is logically divided into two segments

- Data Integration Portion
- User Interface (Access to the meta-data information)



6) Mention how can you connect EME to Abinitio Server?

To connect with Ab initio Server, there are several ways like

- Set AB_AIR_ROOT
- Login to EME web interface- [http://serverhost:\[serverport\]/abinitio](http://serverhost:[serverport]/abinitio)
- Through GDE, you can connect to EME data-store
- Through air-command

7) List out the file extensions used in Abinitio?

The file extensions used in Abinitio are

- **.mp**: It stores Ab initio graph or graph component
- **.mpc**: Custom component or program
- **.mdc**: Dataset or custom data-set component
- **.dml**: Data manipulation language file or record type definition
- **.xfr**: Transform function file
- **.dat**: Data file (multifile or serial file)

8) Mention what information does a .dbc file extension provides to connect to the

database?

The .dbc extension provides the GDE with the information to connect with the database are

- Name and version number of the data-base to which you want to connect
- Name of the computer on which the data-base instance or server to which you want to connect runs, or on which the database remote access software is installed
- Name of the server, database instance or provider to which you want to link

9) Explain how you can run a graph infinitely in Ab initio?

To execute graph infinitely, the graph end script should call the .ksh file of the graph. Therefore, if the graph name is abc.mp then in the end script of the graph it should call to abc.ksh. This will run the graph for infinitely.

10) Mention what the difference between “Look-up” file and “Look is up” in Abinitio?

Lookup file defines one or more serial file (Flat Files); it is a physical file where the data for the Look-up is stored. While Look-up is the component of abinitio graph, where we can save data and retrieve it by using a key parameter.

11) Mention what are the different types of parallelism used in Abinitio?

Different types of parallelism used in Abinitio includes

- **Component parallelism:** A graph with multiple processes executing simultaneously on separate data uses parallelism
- **Data parallelism:** A graph that works with data divided into segments and operates on each segments respectively, uses data parallelism.
- **Pipeline parallelism:** A graph that deals with multiple components executing simultaneously on the same data uses pipeline parallelism. Each component in the pipeline read continuously from the upstream components, processes data and writes to downstream components. Both components can operate in parallel.

12) Explain what is Sort Component in Abinitio?

The Sort Component in Abinitio re-orders the data. It comprises of two parameters “Key” and “Max-core”.

- **Key:** It is one of the parameters for sort component which determines the collation order
- **Max-core:** This parameter controls how often the sort component dumps data from memory to disk

13) Mention what dedup-component and replicate component does?

- **Dedup component:** It is used to remove duplicate records

- **Replicate component:** It combines the data records from the inputs into one flow and writes a copy of that flow to each of its output ports

14) Mention what is a partition and what are the different types of partition components in Abinitio?

In Abinitio, partition is the process of dividing data sets into multiple sets for further processing. Different types of partition component includes

- **Partition by Round-Robin:** Distributing data evenly, in block size chunks, across the output partitions
- **Partition by Range:** You can divide data evenly among nodes, based on a set of partitioning ranges and key
- **Partition by Percentage:** Distribution data, so the output is proportional to fractions of 100
- **Partition by Load balance:** Dynamic load balancing
- **Partition by Expression:** Data dividing according to a DML expression
- **Partition by Key:** Data grouping by a key

15) Explain what is SANDBOX?

A SANDBOX is referred for the collection of graphs and related files that are saved in a single directory tree and behaves as a group for the purposes of navigation, version control, and migration.

16) Explain what is de-partition in Abinitio?

De-partition is done in order to read data from multiple flow or operations and are used to re-join data records from different flows. There are several de-partition components available which includes **Gather, Merge, Interleave, and Concatenation.**

17) List out some of the air commands used in Abinitio?

Air command used in Abinitio includes

- **air object ls :** It is used to see the listings of objects in a directory inside the project
- **air object rm :** It is used to remove an object from the repository
- **air object versions-verbose :** It gives the version history of the object.

Other air command for Abinitio include air object cat, air object modify, air lock show user, etc.

18) Mention what is Rollup Component?

Roll-up component enables the users to group the records on certain field values. It is a multiple stage function and consists initialize 2 and Rollup 3.

19) Mention what is the syntax for m_dump in Abinitio?

The syntax for m_dump in Abinitio is used to view the data in multfile from unix prompt. The command for m_dump includes

- **m_dump a.dml a.dat:** This command will print the data as it manifested from GDE when we view data in formatted text
- **m_dump a.dml a.dat>b.dat:** The output is re-directed in b.dat and will act as a serial file.b.dat that can be referred when it is required.