

ORACLE®
Certified Professional

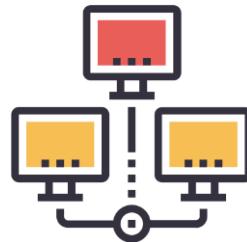


Microsoft Certified
Professional

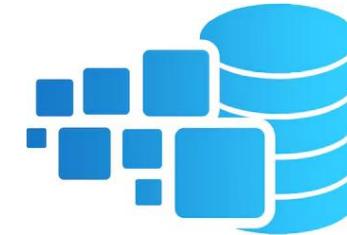
سہیل عمران



Computer
SCIENCE



BigData Analytics



Dr. SOHAIL IMRAN



Installation

Apache Drill is an innovative **schema-free SQL query engine** for Hadoop, NoSQL, and cloud storage. This open-source software enables users to analyze large-scale datasets from numerous sources directly, without needing to shift data across systems. Apache Drill provides the **flexibility** of on-the-fly data discovery, making it significantly adaptable to changing data formats and structures.

Drill supports a variety of NoSQL databases and file systems, including HBase, MongoDB, MapR-DB, HDFS, MapR-FS, Amazon S3, Azure Blob Storage, Google Cloud Storage, Swift, NAS, and local files. A single query can join data from multiple datastores. For example, you can join a user profile collection in MongoDB with a directory of event logs in Hadoop.

Drill's datastore-aware optimizer automatically restructures a query plan to leverage the datastore's internal processing capabilities. In addition, Drill supports data locality, so it's a good idea to co-locate Drill and the datastore on the same nodes.



Setup JAVA_HOME

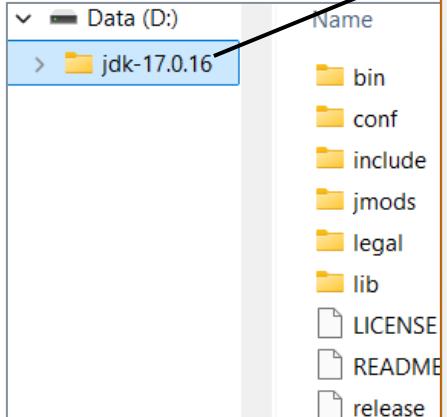


Recent download history

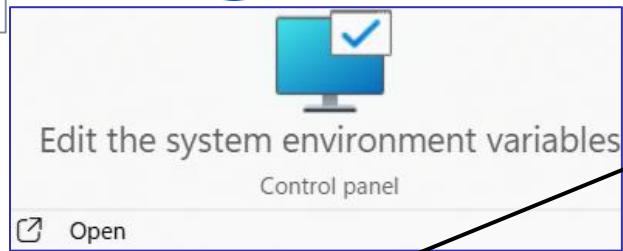
jdk-17.0.16_windows-x64_bin.zip
173 MB • 6 minutes ago

Extract

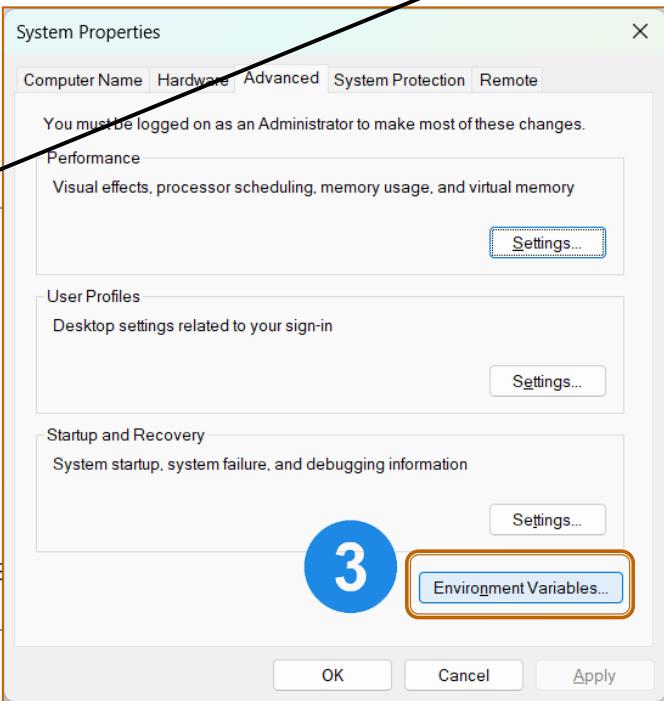
1



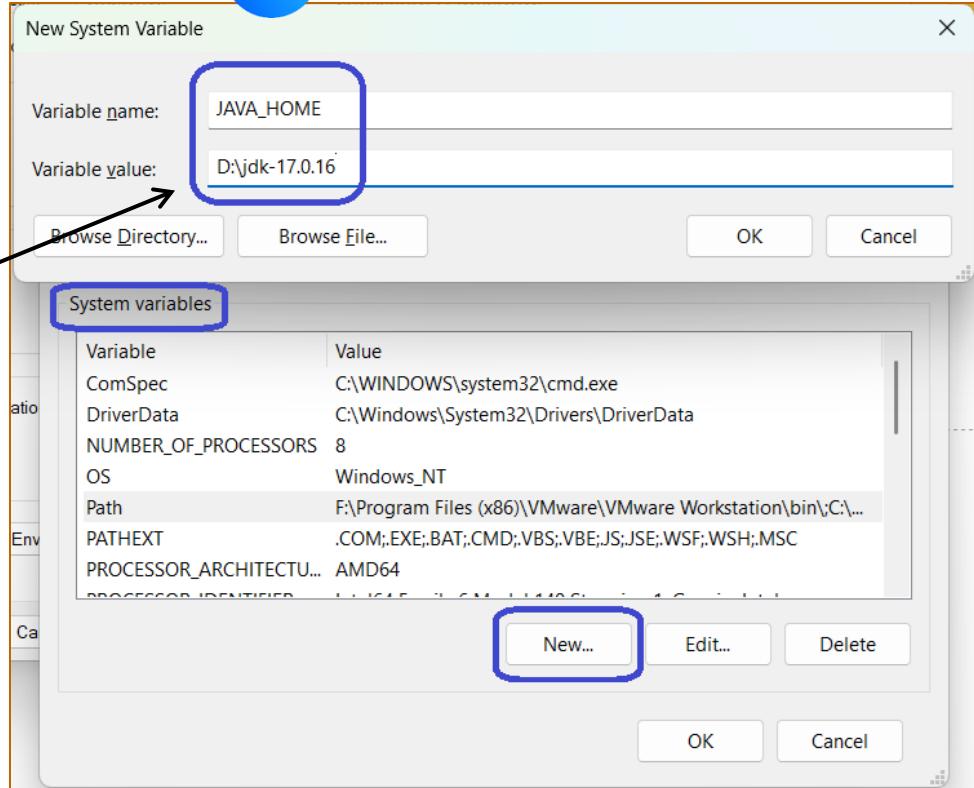
2



3



4



Setup - PATH



1

System variables

Variable	Value
ComSpec	C:\WINDOWS\system32\cmd.exe
DriverData	C:\Windows\System32\Drivers\DriverData
JAVA_HOME	D:\jdk-17.0.16
NUMBER_OF_PROCESSORS	8
OS	Windows_NT
Path	F:\Program Files (x86)\VMware\VMware Workstation\bin;C:\...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64

New... Edit... Delete

2

Edit environment variable

F:\Program Files (x86)\VMware\VMware Workstation\bin\
%SystemRoot%\system32
%SystemRoot%
%SystemRoot%\System32\Wbem
%SYSTEMROOT%\System32\WindowsPowerShell\v1.0\
%SYSTEMROOT%\System32\OpenSSH\
%USERPROFILE%\dnx\bin
C:\Program Files\Microsoft DNX\Dnvm\
C:\Program Files (x86)\Windows Kits\8.1\Windows Performance ...
C:\Program Files\Microsoft SQL Server\120\Tools\Binn\
C:\Program Files (x86)\Microsoft SQL Server\150\DT\Binn\
C:\Program Files\Azure Data Studio\bin
C:\Program Files (x86)\nodejs\
D:\jdk-17.0.16\bin

New Edit... Browse... Delete Move Up Move Down Edit text... OK Cancel

3
`java -version`

Windows PowerShell

```
PS C:\Users\DELL> java -version
java version "1.7.0_80"
Java(TM) SE Runtime Environment (build 1.7.0_80-b15)
Java HotSpot(TM) 64-Bit Server VM (build 24.80-b11, mixed mode)
PS C:\Users\DELL>
```

download and extract



<https://drill.apache.org/download/>

Download

Drill 1.22.0 was released on June 28, 2025.

Drill for Hadoop 3 and non-Hadoop environments

Find an Apache Mirror

Direct File Download

Client Drivers (ODBC/JDBC)



Drill for Hadoop 2 environments

Official support for Hadoop 2 was dropped along with support for Java 8 with the release of Drill 1.22.0. Users who would like to create their own builds of Drill 1.22.0 or newer for Java 8 and Hadoop 2 are directed to [the developer documentation](#).

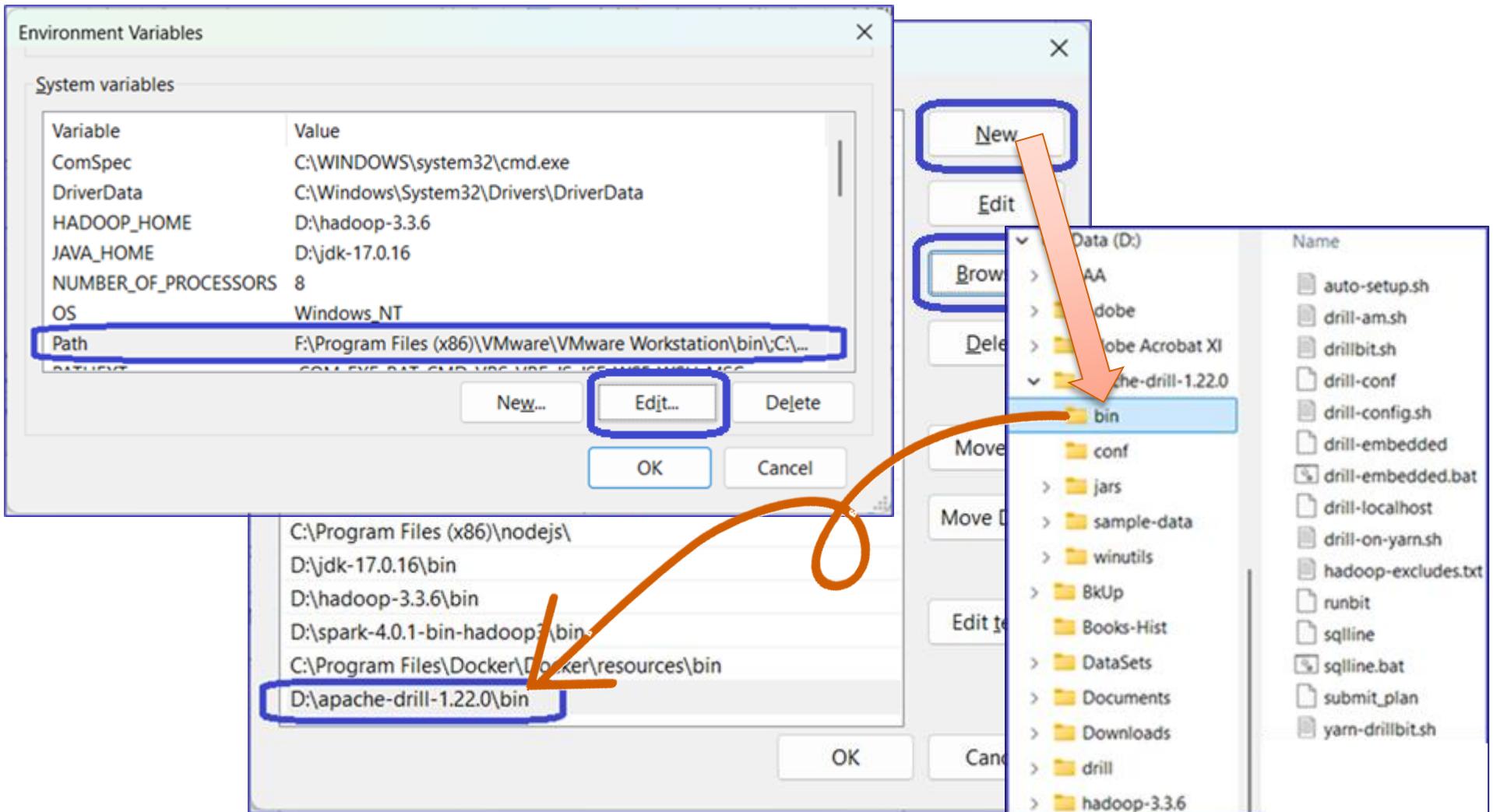
Additional resources for Drill 1.22.0:

- [Blog post](#)
- [Release notes \(JIRA\)](#)
- Source code: [tar.gz](#) [ZIP](#).
- Verification Links:
 - [SHA](#)
 - [ASC](#)
 - [KEYS](#)
 - [How to Verify Downloaded Files](#)

If you're looking for an older release, see the [release notes](#).

big
data
ANALYTICS

Environment Variables



Start



```
Windows PowerShell PS C:\Users\DELL> drill-embedded.bat
Apache Drill 1.22.0
"A little SQL for your NoSQL."
apache drill> |
```

Browser



<http://localhost:8047/storage>

localhost:8047/storage

Apache Drill Query Profiles Storage Metrics Threads Logs Options Documentation

Plugin Management

Create Export all

Enabled Storage Plugins			Disabled Storage Plugins				
cp	Update	Disable	Export	box	Update	Enable	Export
dfs	Update	Disable	Export	cassandra	Update	Enable	Export
				drill	Update	Enable	Export
				dropbox	Update	Enable	Export
				druid	Update	Enable	Export
				elastic	Update	Enable	Export
				googlesheets	Update	Enable	Export
				hbase	Update	Enable	Export

hive	Update	Enable	Export
http	Update	Enable	Export
kafka	Update	Enable	Export
kudu	Update	Enable	Export
mongo	Update	Enable	Export
oci	Update	Enable	Export
opentsdb	Update	Enable	Export
phoenix	Update	Enable	Export
rdbms	Update	Enable	Export
s3	Update	Enable	Export
splunk	Update	Enable	Export

Query - file



```
SELECT * FROM dfs.`D:\DataSets\emp.json` LIMIT 3;
```

Apache Drill **Query** Profiles Storage Metrics Threads Logs Options Documentation

Sample SQL query: `SELECT * FROM cp.`employee.json` LIMIT 20` ×

Query type: SQL Physical Logical

Query **Hint:** Use *Ctrl+Enter* to submit

```
1 SELECT * FROM dfs.`D:\DataSets\emp.json` LIMIT 3;
```

Submit Reset Limit results to rows Default schema: schema i i

Query Submitted

Waiting for results...

Please don't close this window

empno	ename	hiredate	sal	comm	deptno	dir0		
7839	KING	{"\$date": "2001-11-16T19:00:00.000Z"}	5000	0	10	DataSets		
7698	BLA	{"\$date": "2001-04-30T19:00:00.000Z"}	2850	0	30	DataSets		
7782	CLARK	MANAGER	7839	{"\$date": "2001-06-08T19:00:00.000Z"}	2450	0	10	DataSets

Showing 1 to 3 of 3 entries

Previous 1 Next

MongoDB service



Services

File Action View Help

Services (Local)

MongoDB Server (MongoDB)

[Start the service](#)

Description:
MongoDB Database Server
(MongoDB)

Start service MongoDB Server (MongoDB) on Local Computer

Service Control

Windows is attempting to start the following service on Local Computer...

MongoDB Server (MongoDB)

Name	Description	Status	Automatic	Network Se...
Microsoft Windows SMS Ro...	Routes mess...	Manual (Trigg...	Local System	
Mobile Connectivity Manag...	This service ...	Disabled	Local Service	
MongoDB Server (MongoDB)	MongoDB D...	Running	Manual (Trigg...	Local System
Start	Signal aggre...	Manual	Local System	
Stop	Provides abil...	Disabled	Local Service	
Pause	Maintains a ...	Manual	Local System	
Resume	Network Co...	Manual (Trigg...	Local Service	
Restart	Brokers con...	Manual (Trigg...	Local System	
All Tasks	Manages ob...	Manual	Local System	
Refresh	Provides Dir...	Manual (Trigg...	Local System	
Properties	Identifies th...	Manual	Network Se...	
Help	Collects and ...	Manual	Network Se...	
	The Network...	Manual (Trigg...	Local System	
	This service ...	Running	Automatic	Local Service
	Provides net...	Running	Manual	Local System

Extended / Standard /

Start service MongoDB Server (MongoDB) on Local Computer



Enable Mongo

Apache Drill Query Profiles Storage Metrics Threads Logs

hive Update Enable Export

http Update Enable Export

kafka Update Enable Export

kudu Update Enable Export

mongo Update Enable Export

oci Update Enable Export

opentsdb Update Enable Export

phoenix Update Enable Export

rdbms Update Enable Export

s3 Update Enable Export

splunk Update Enable Export

Plugin Management

Create Export all

Enabled Storage Plugins

cp Update Disable Export

dfs Update Disable Export

mongo Update Disable Export

Apache Drill Query Profiles Storage Metrics

Configuration

```
1 {  
2   "type": "mongo",  
3   "connection": "mongodb://localhost:27017/",  
4   "pluginOptimizations": {  
5     "supportsProjectPushdown": true,  
6     "supportsFilterPushdown": true,  
7     "supportsAggregatePushdown": true,  
8     "supportsSortPushdown": true,  
9     "supportsUnionPushdown": true,  
10    "supportsLimitPushdown": true  
11  },  
12  "batchSize": 100,  
13  "enabled": true,  
14  "authMode": "SHARED_USER"  
15 }
```

Back Update Disable Export Delete

```
graph LR; A[mongo] --> B[config JSON]; A --> C[mongo]
```

Query - MongoDB



```
SELECT * FROM mongo.mydb.emp LIMIT 3;
```

Query type: SQL Physical Logical

Query

```
1 SELECT * FROM mongo.mydb.emp LIMIT 3;
```

_id	empno	ename	job	hiredate	sal	deptno	mgr
h\xE1\x1BQh\x97\xF3\xAD\xAE\xDC\xDBz	7839	KING	PRESIDENT	2001-11-17T00:00	5000	10	null
h\xE1\x1BQh\x97\xF3\xAD\xAE\xDC\xDB{	7698	BLAKE	MANAGER	2001-05-01T00:00	2850	30	7839
h\xE1\x1BQh\x97\xF3\xAD\xAE\xDC\xDB	7782	CLARK	MANAGER	2001-06-09T00:00	2450	10	7839

Showing 1 to 3 of 3 entries

Previous 1 Next

big
data
ANALYTICS