

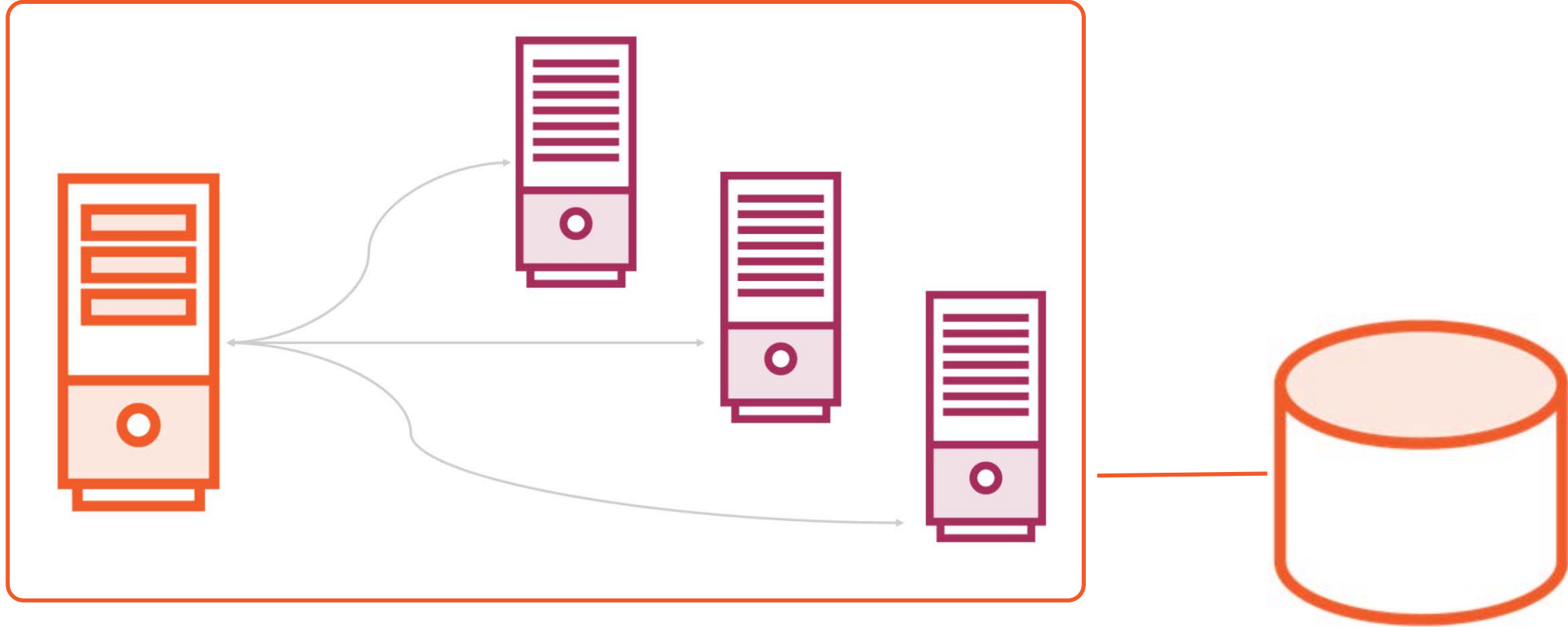
Running Hadoop on Windows



Elton Stoneman

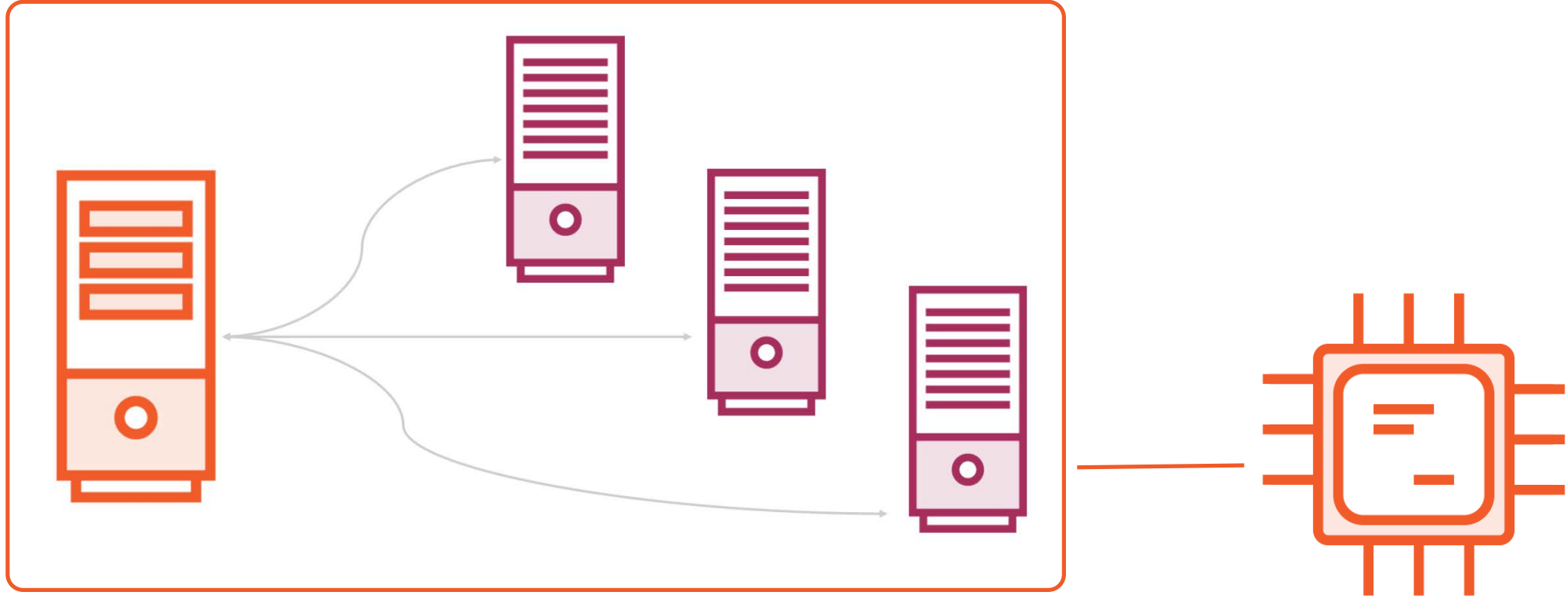
SOFTWARE ARCHITECT

@EltonStoneman blog.sixeyed.com

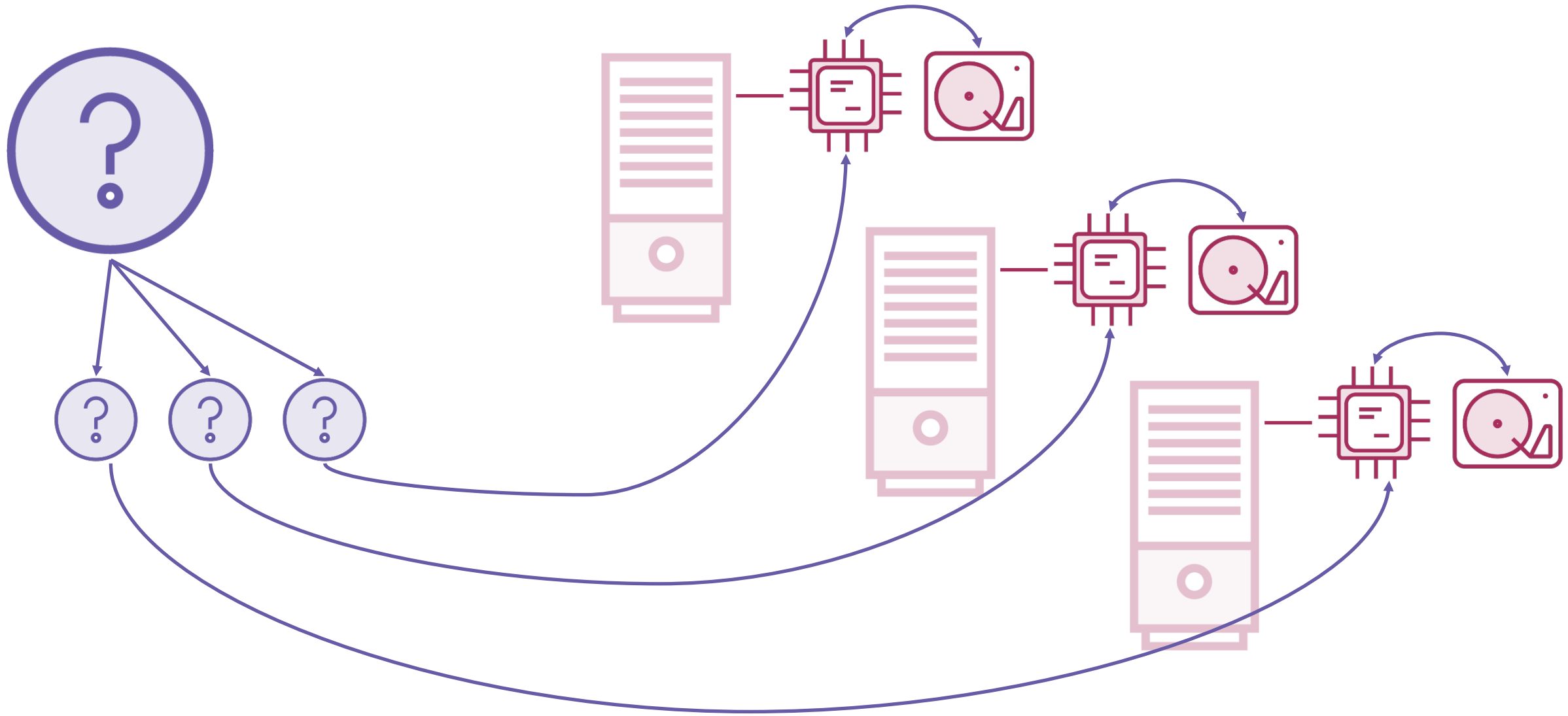


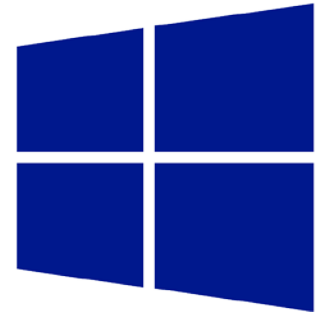
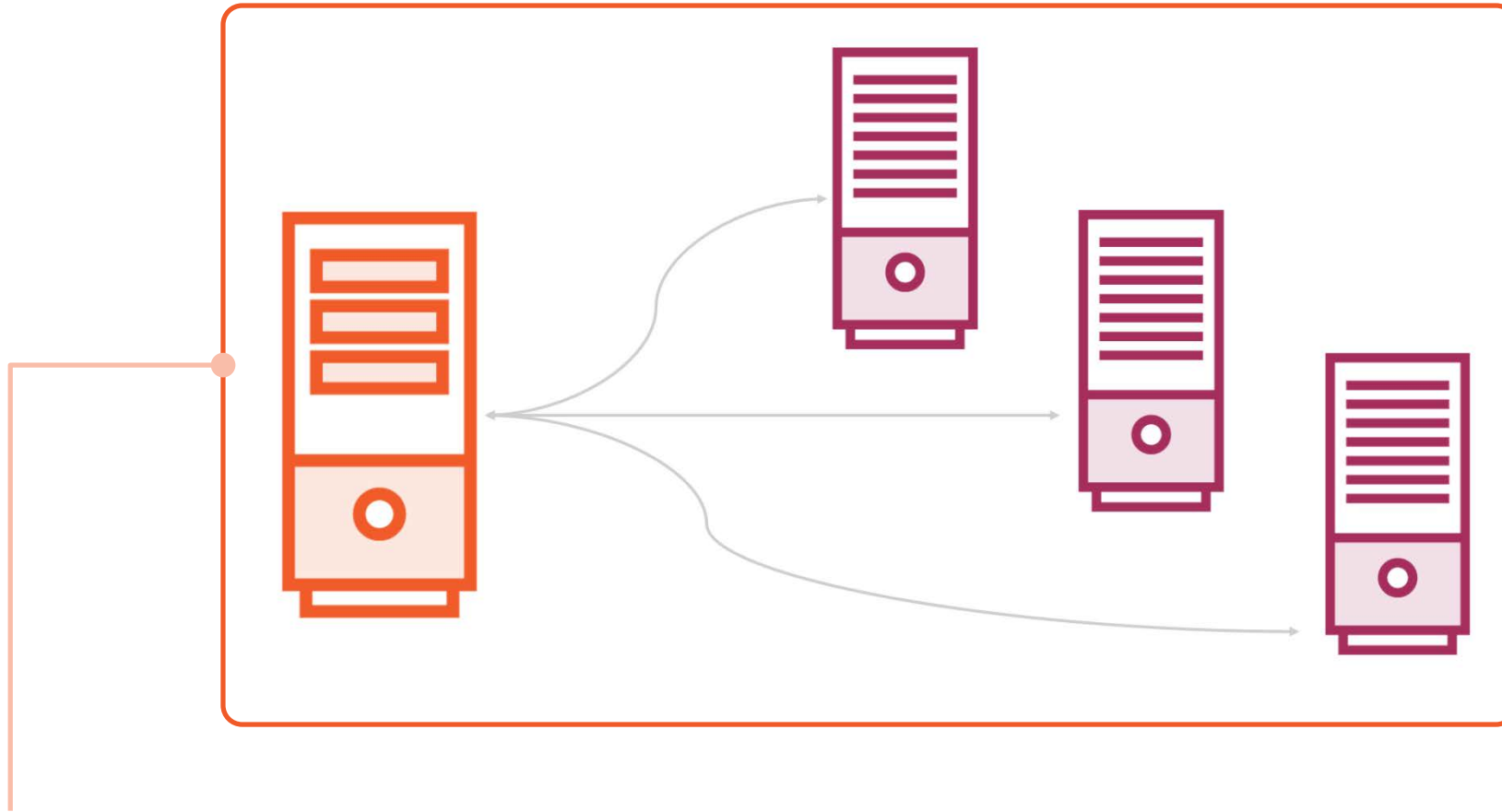
Hadoop Distributed File System






Yet Another Resource Negotiator

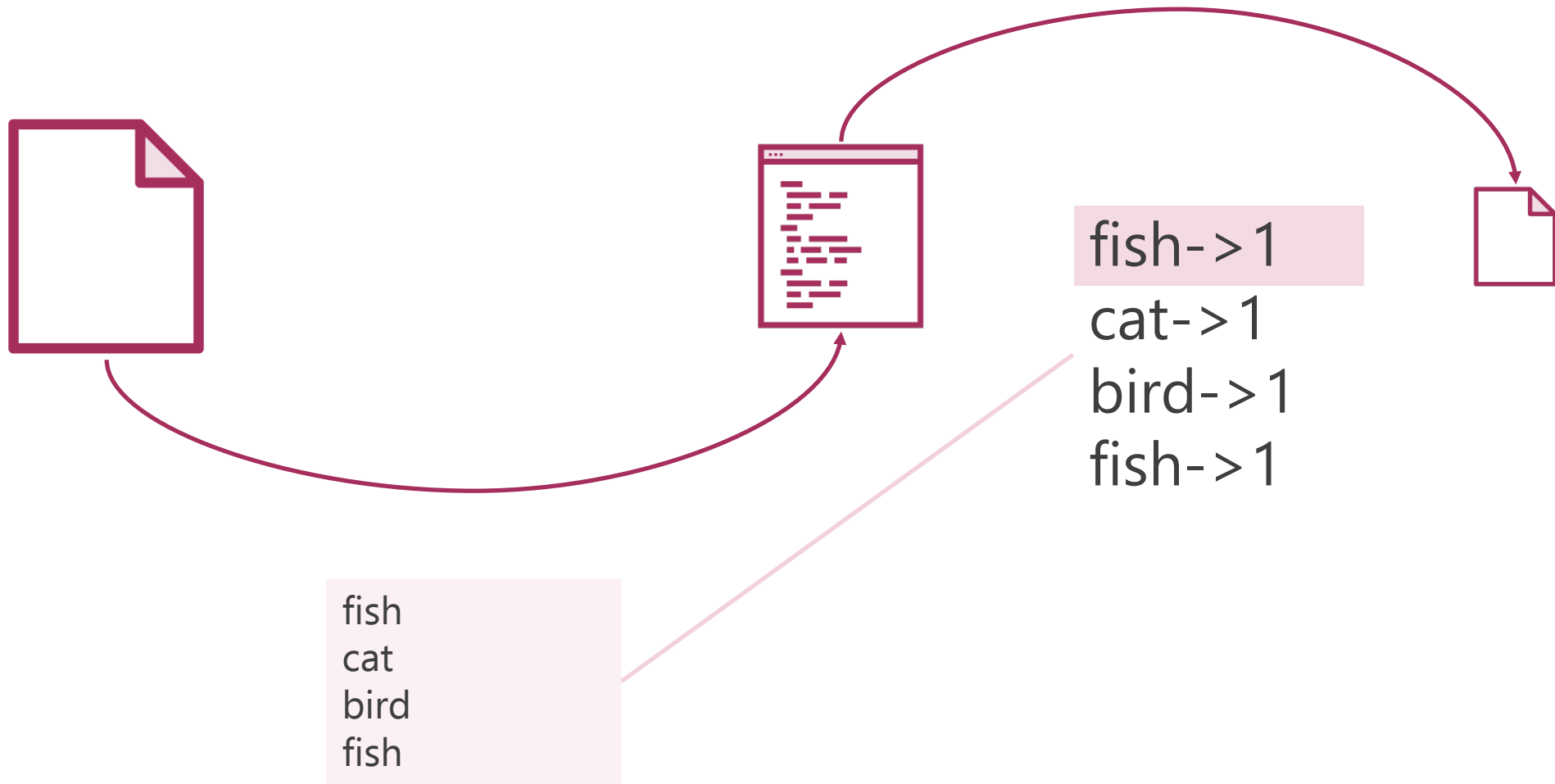


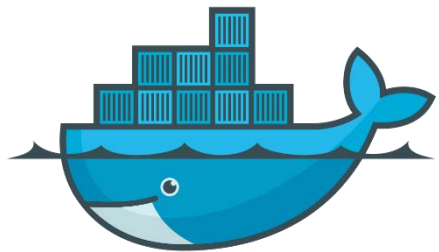
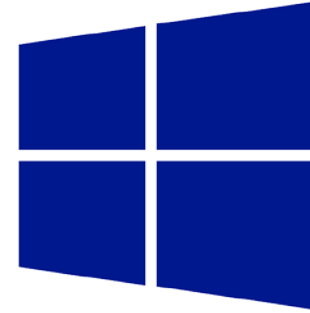




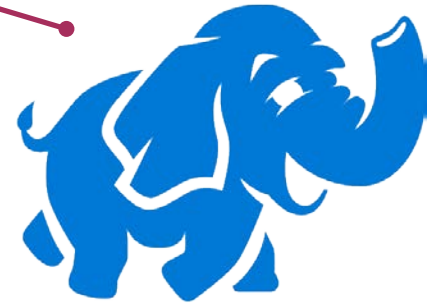
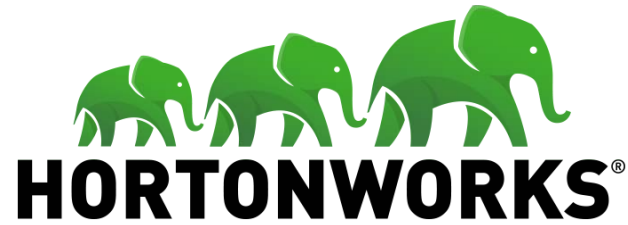
HADOOP WITH .NET POC

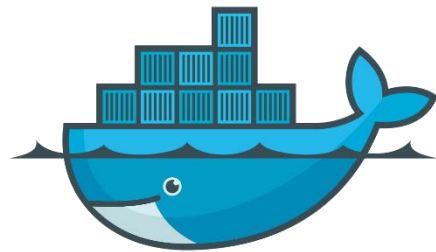
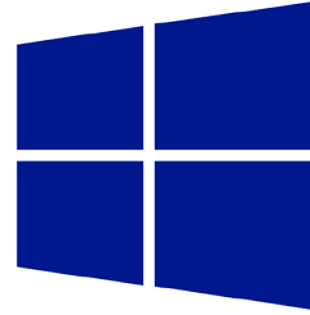
- Does Hadoop run well on Windows?
- Can you write queries in .NET?
- What about **fast queries** in .NET?
- Does the Hadoop ecosystem  Microsoft?



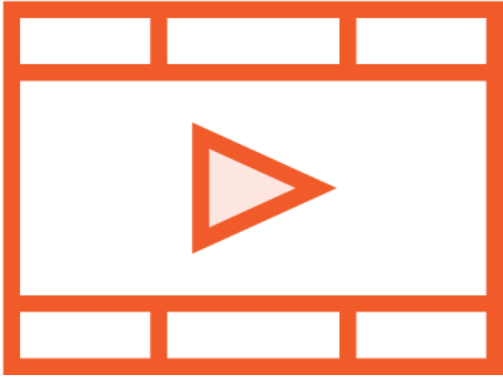


docker





docker

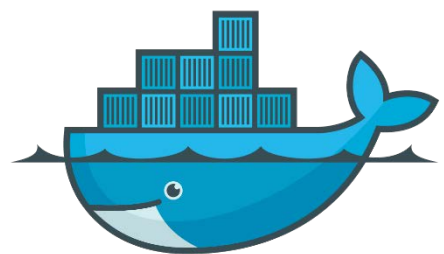


Docker and Containers: The Big Picture

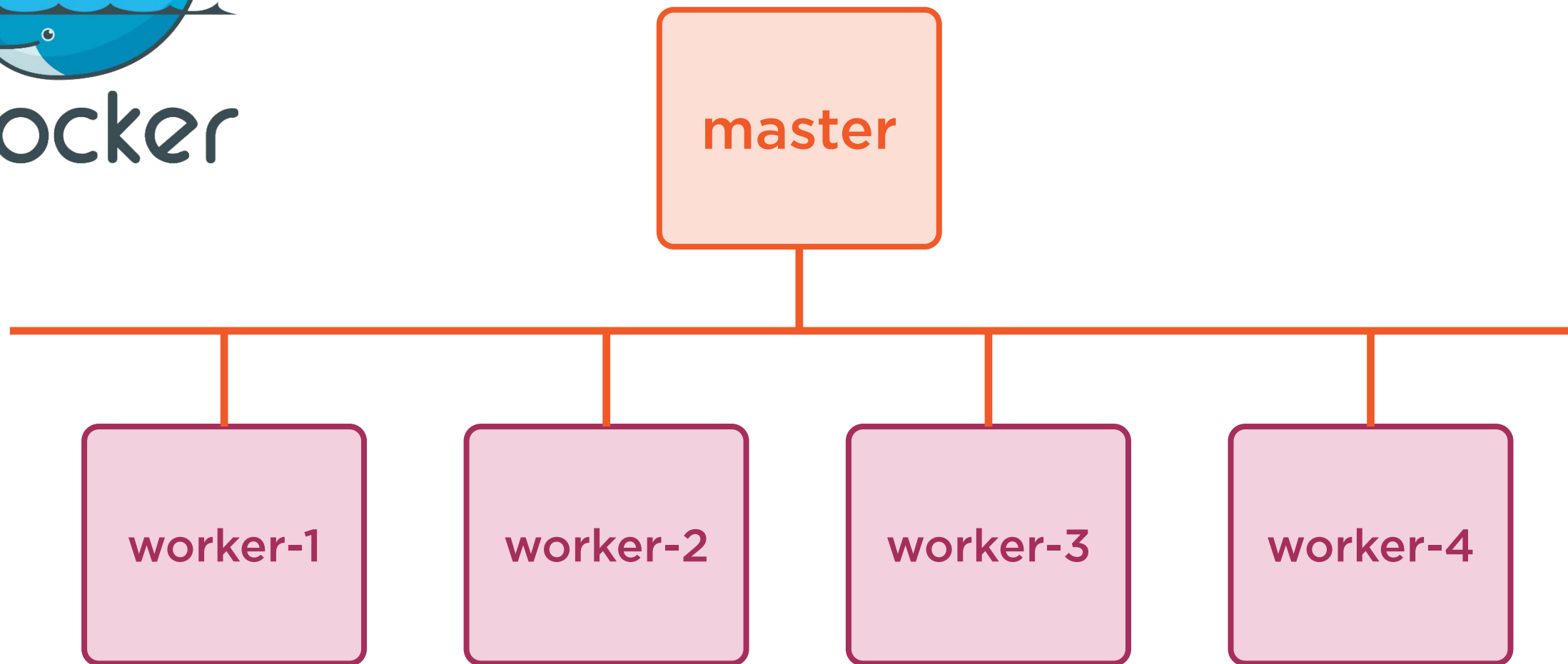


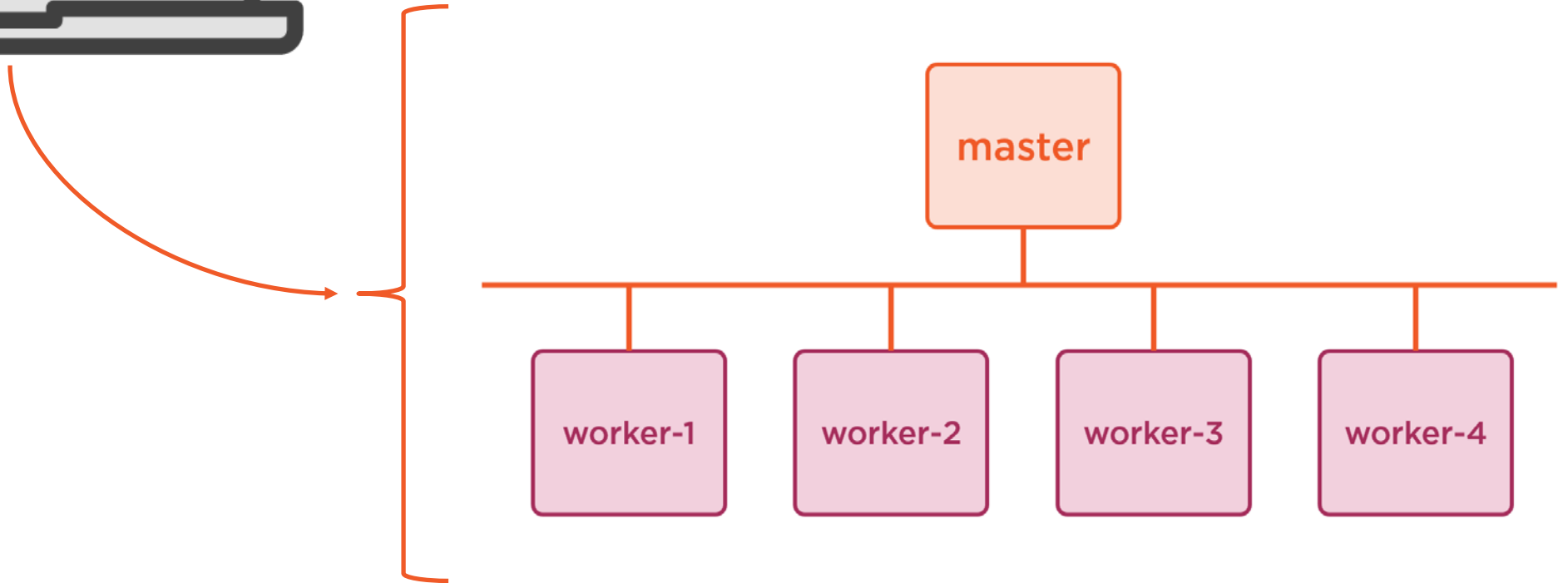
Nigel Poulton

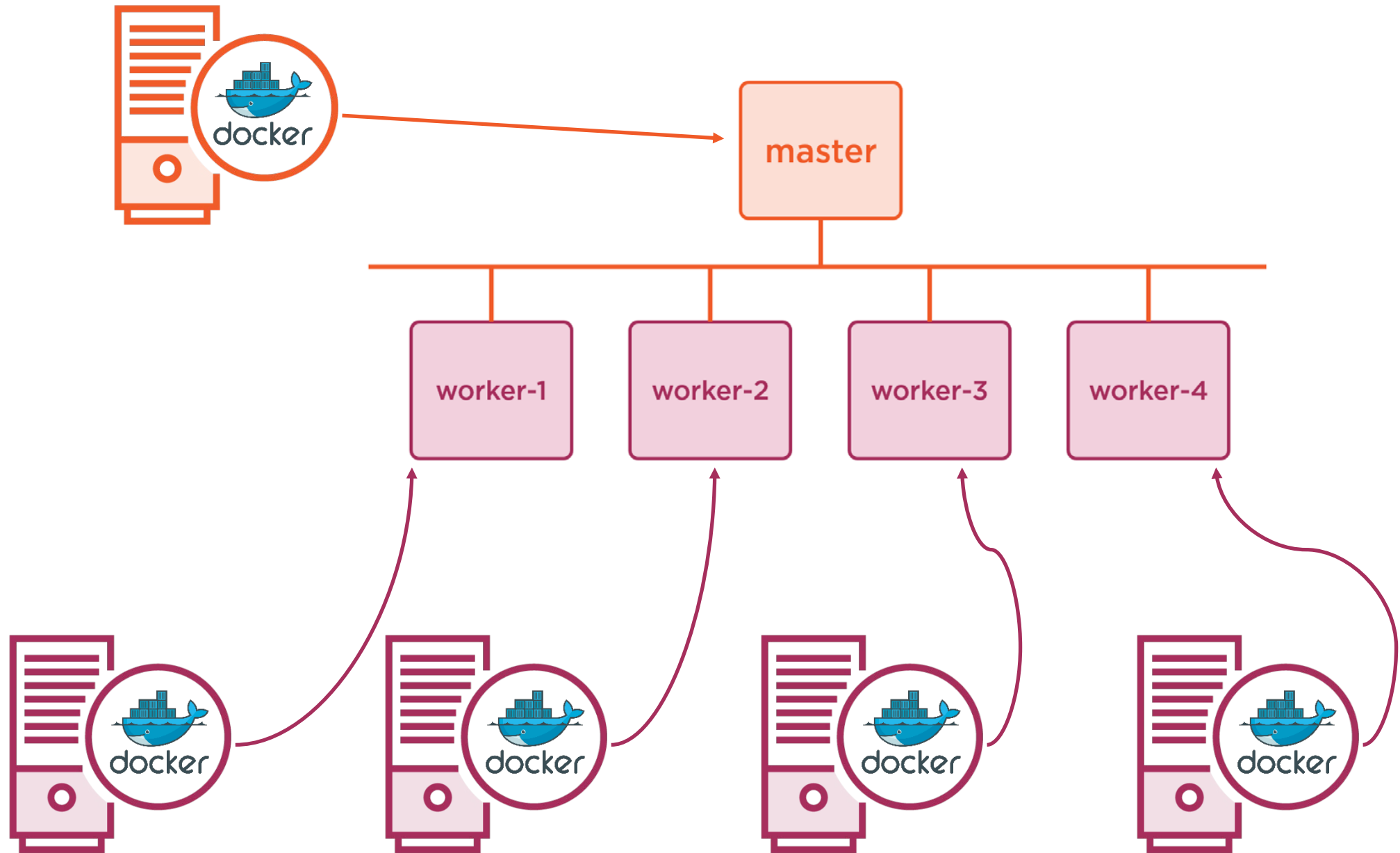
@nigelpoulton nigelpoulton.com



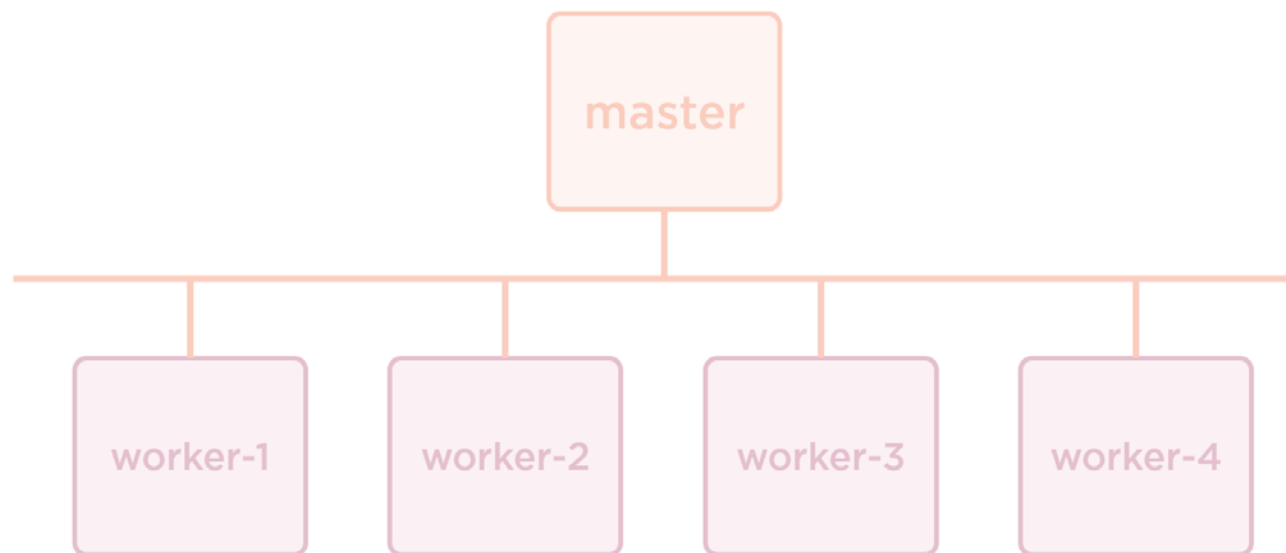
docker



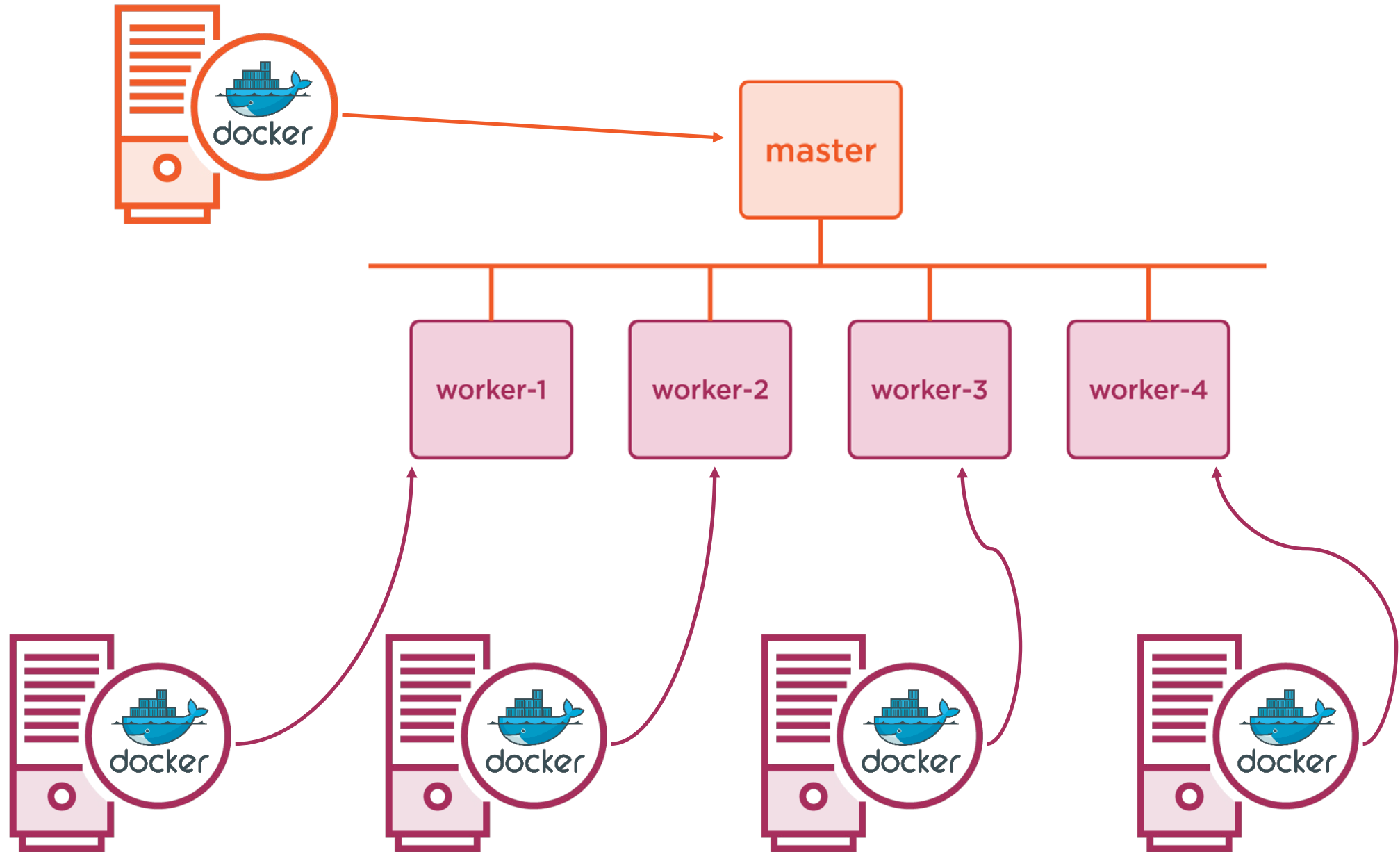


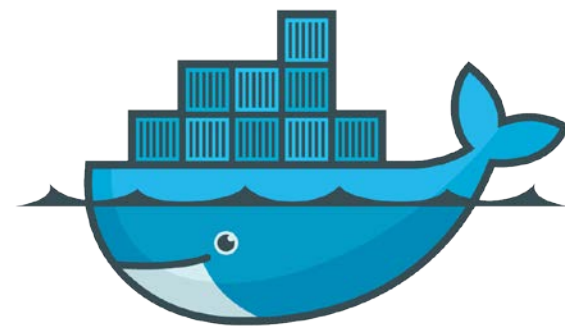
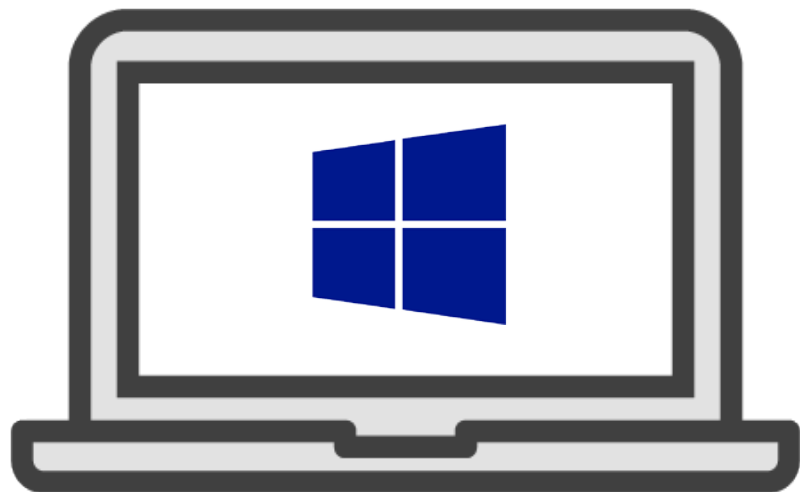


docker-compose **stop**

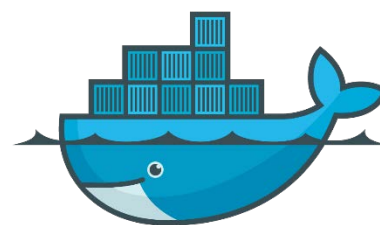
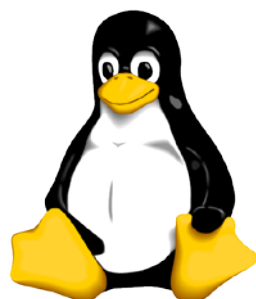
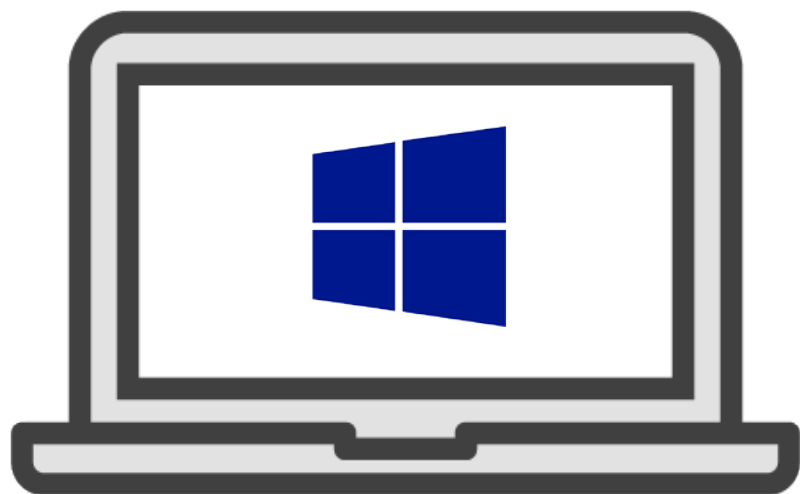


docker-compose **start**

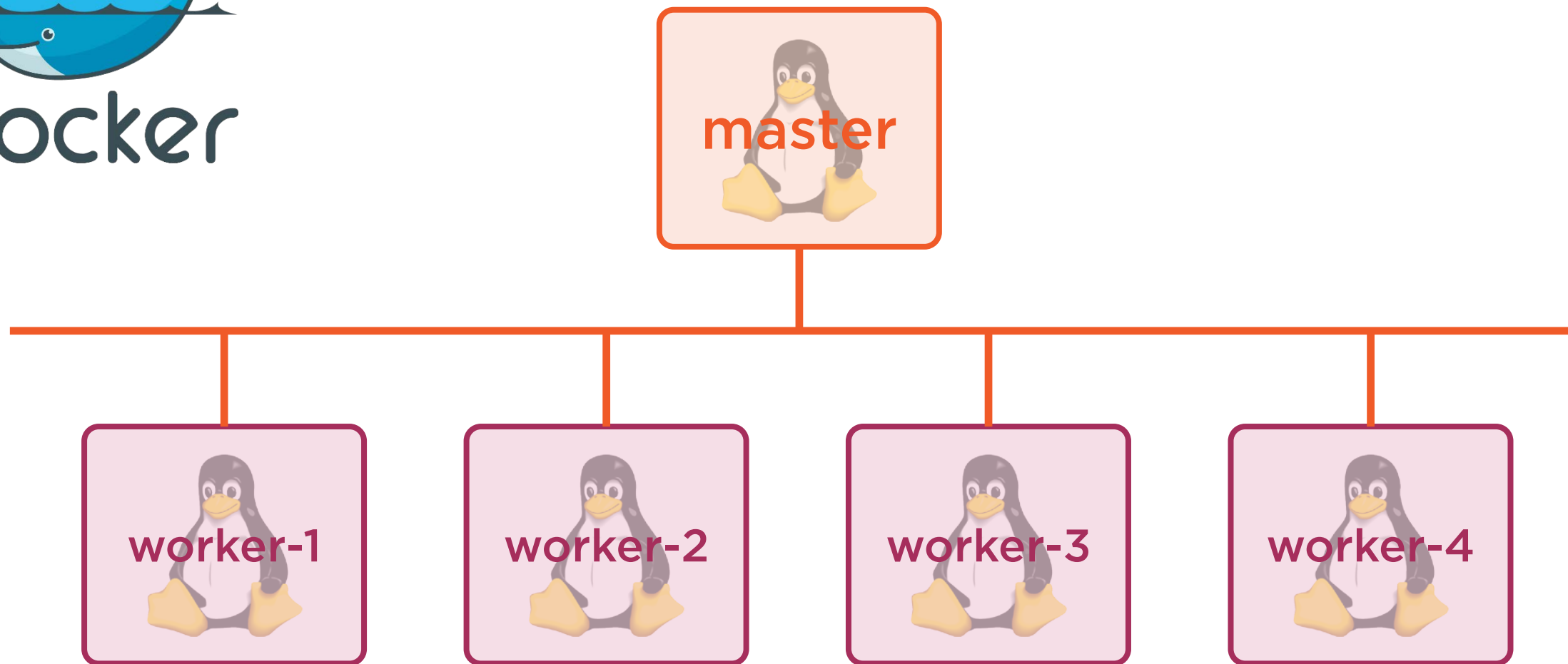
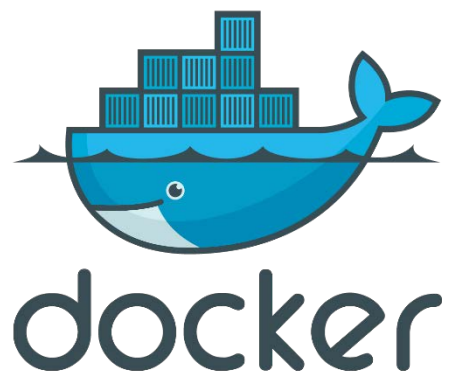


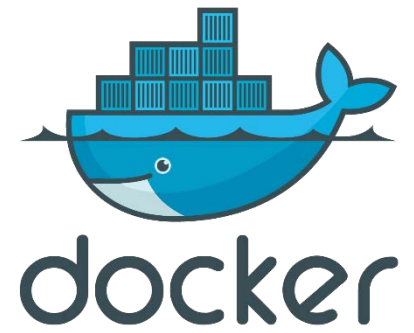
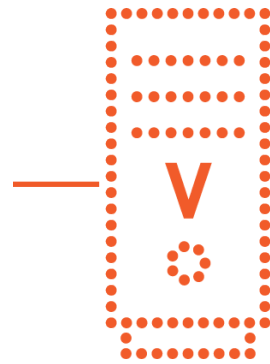
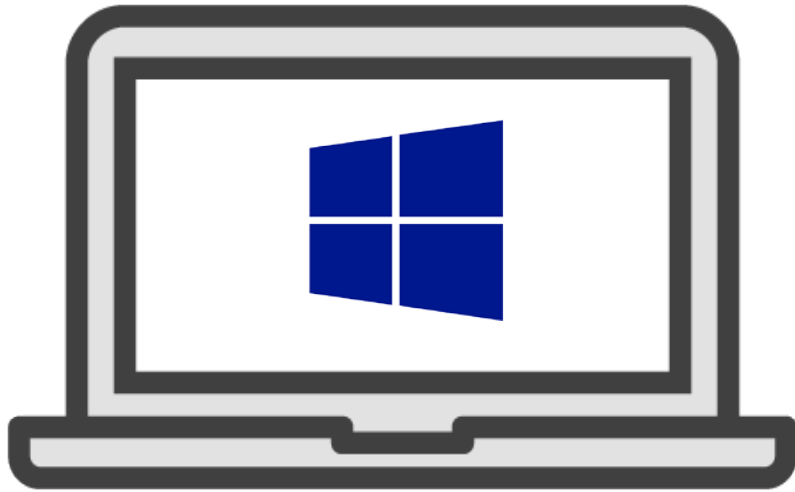


docker

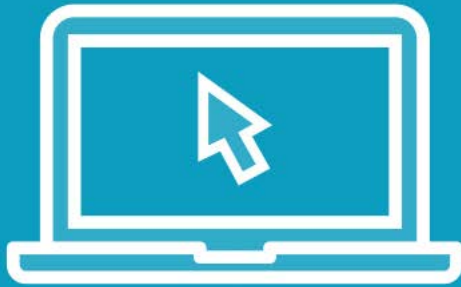


docker



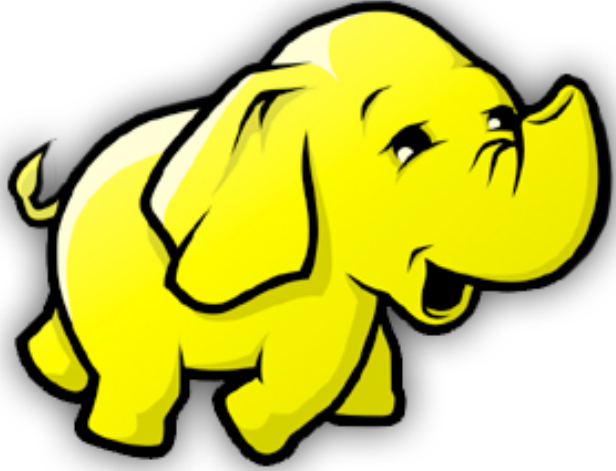


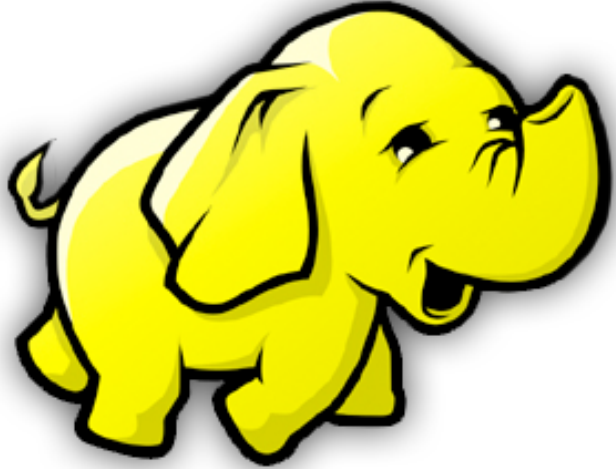
Demo



Hadoop in Docker

- Starting the cluster
- Connecting to the Master
- Running .NET Core jobs



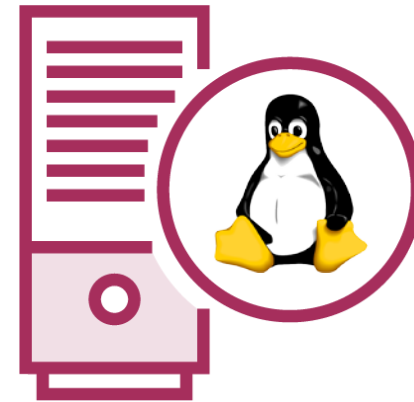
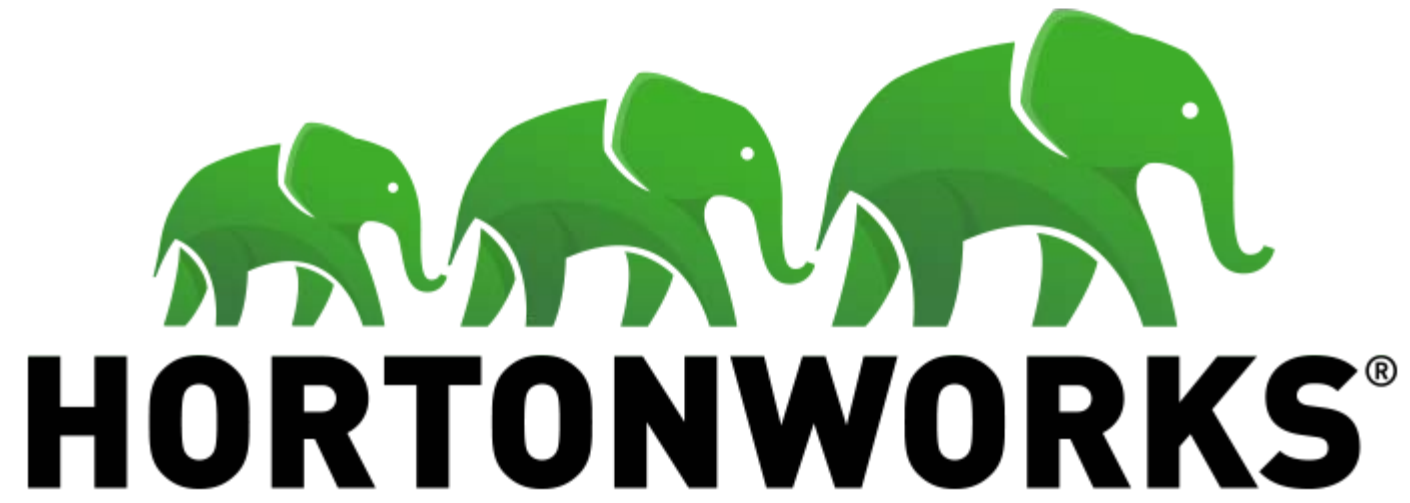


Distribution

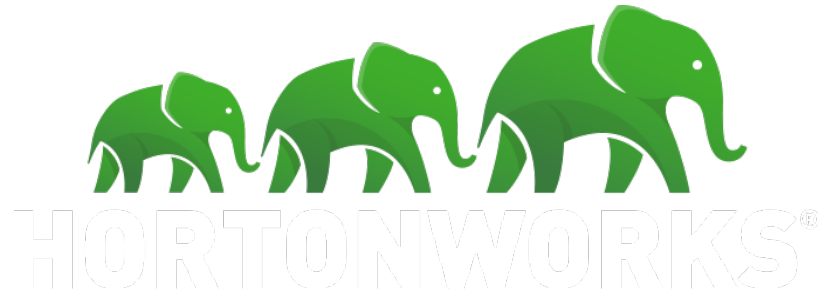
Packaged Components

Management UI

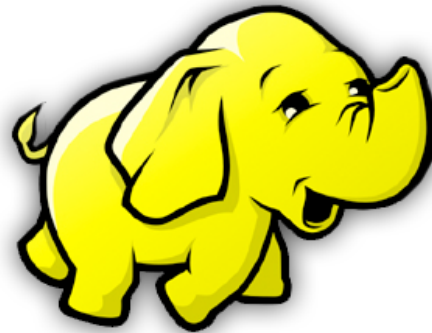
Support



HDP



APACHE
HBASE



Apache Ranger

HDP 2.4

Apache



2.7.1

2.7.2



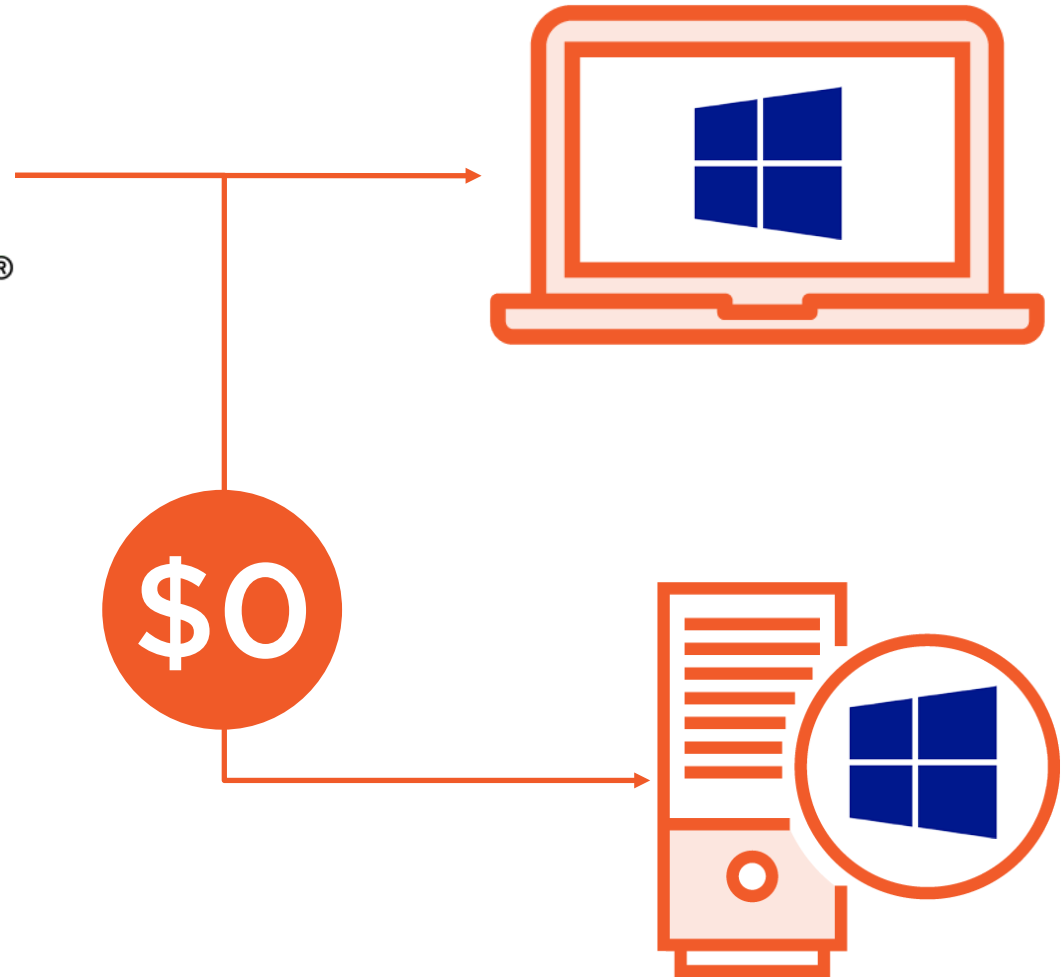
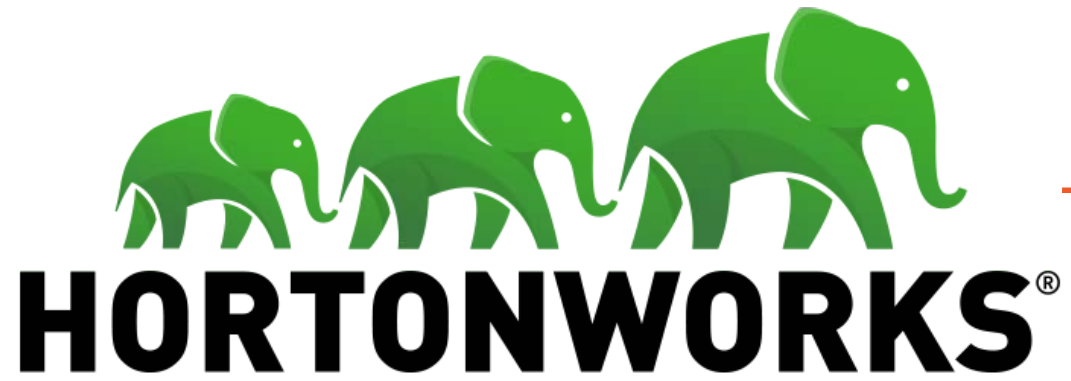
1.2.1

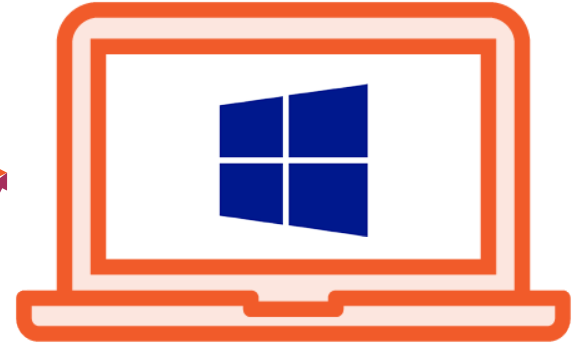
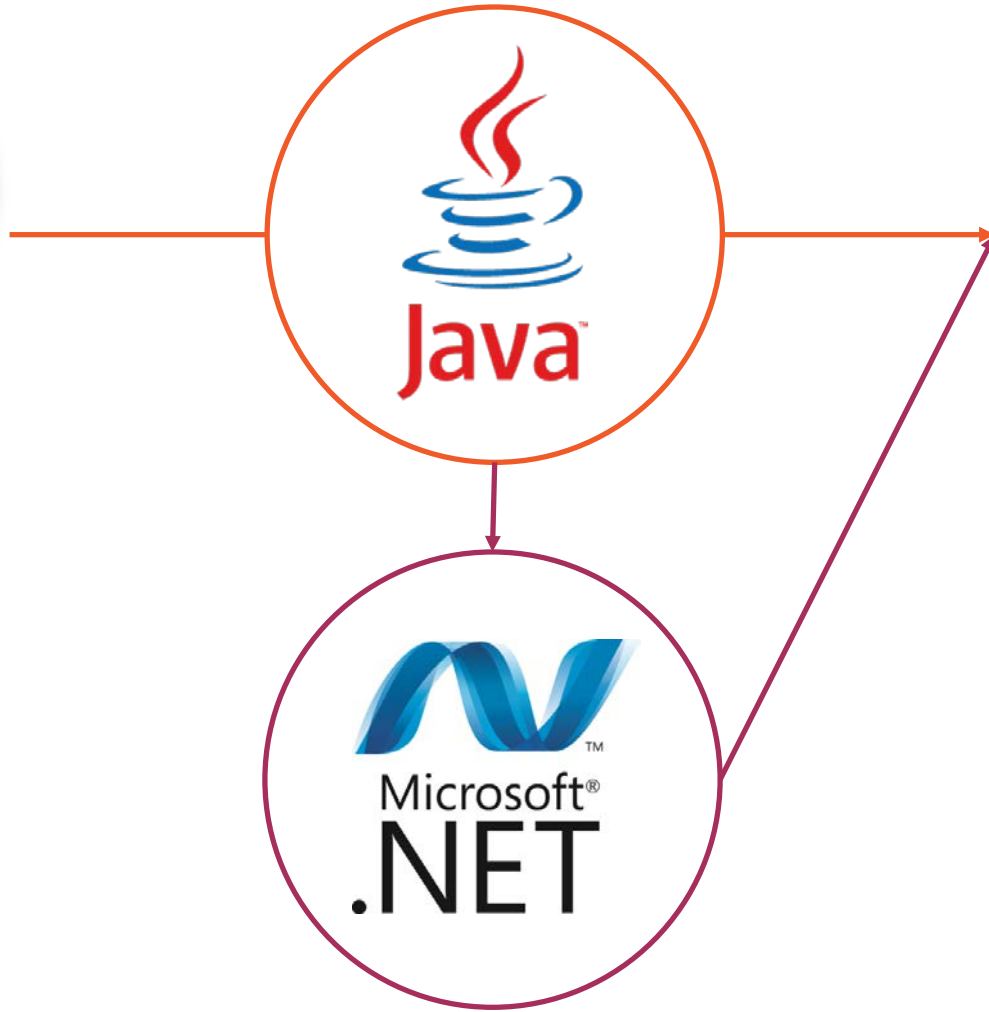
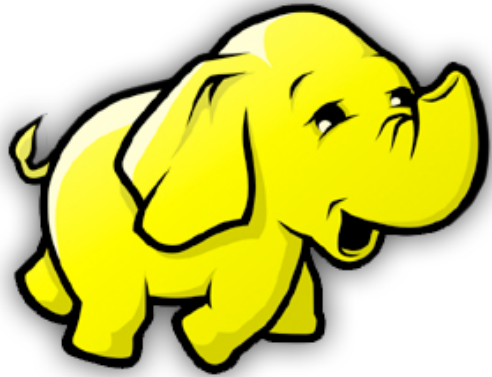
2.0.1

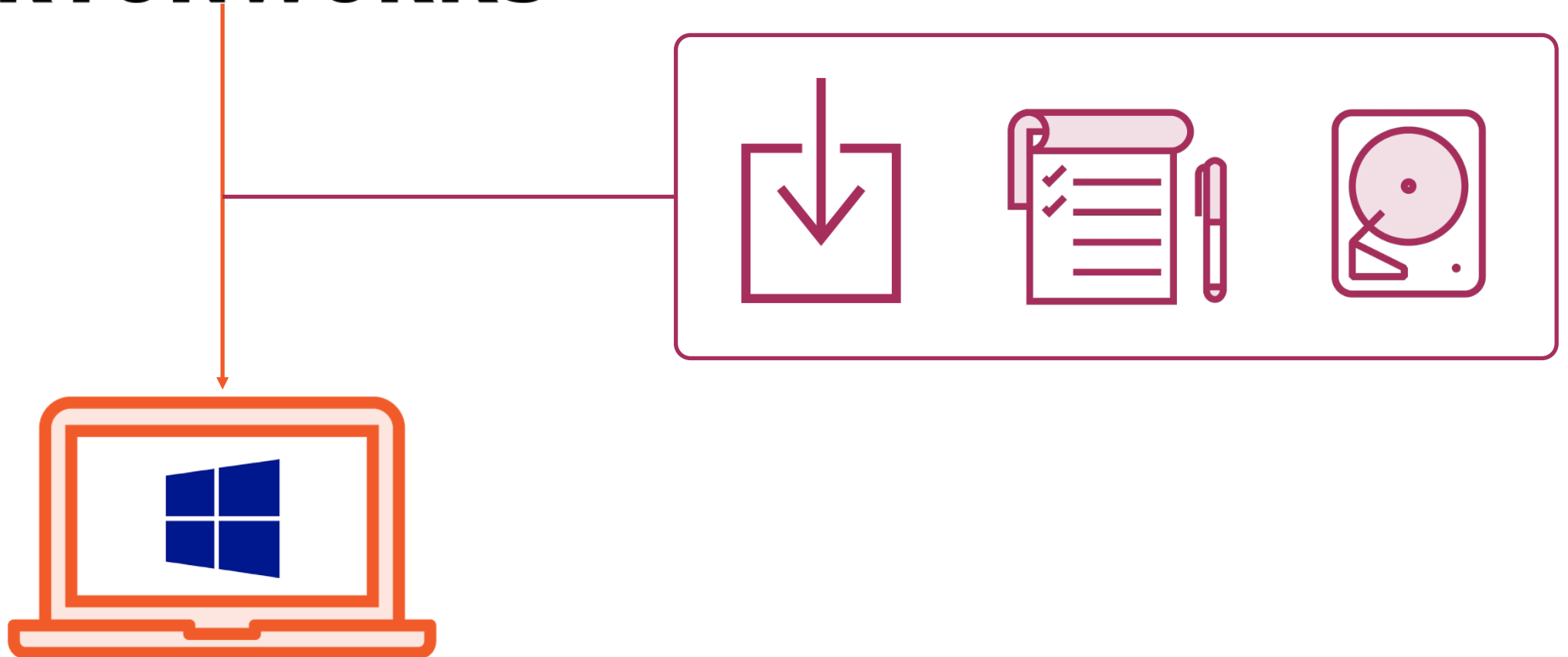


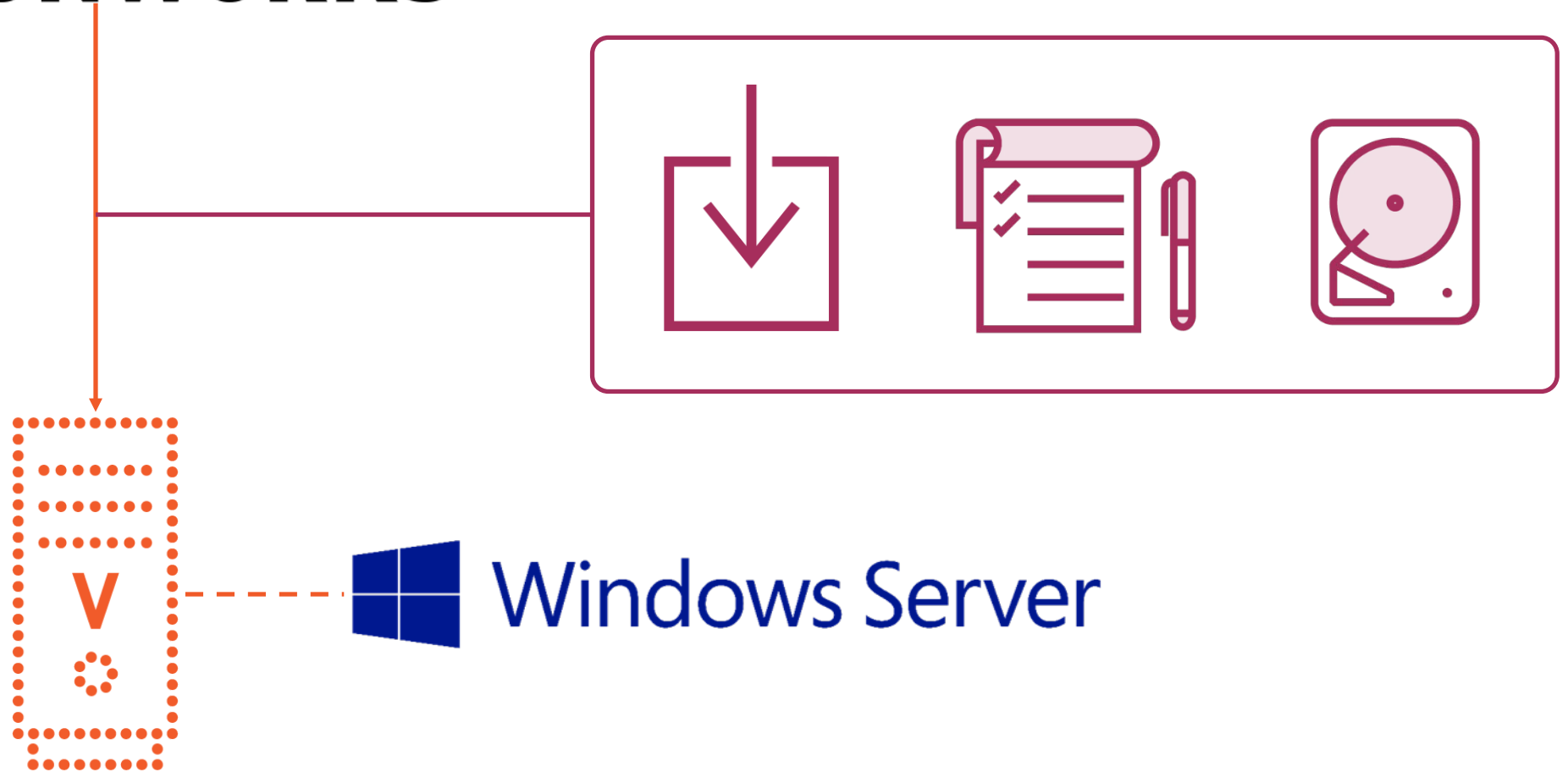
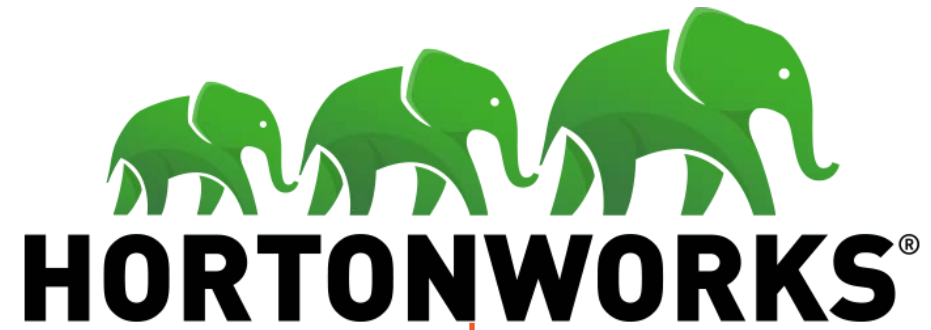
1.6.0

1.6.1

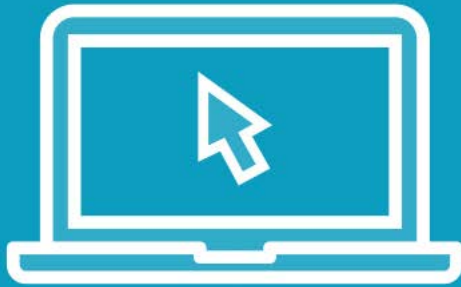






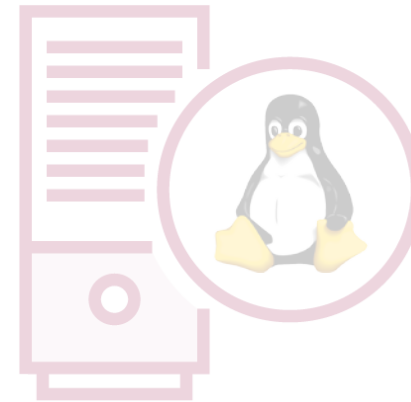


Demo



Hortonworks HDP

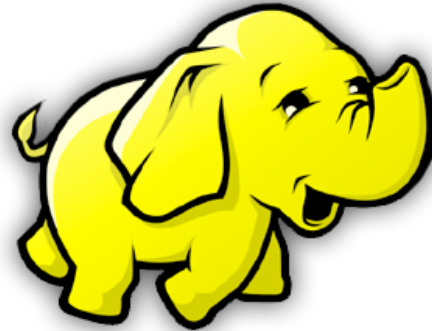
- Installing on Windows Server
- Starting Hadoop services
- Running .NET jobs

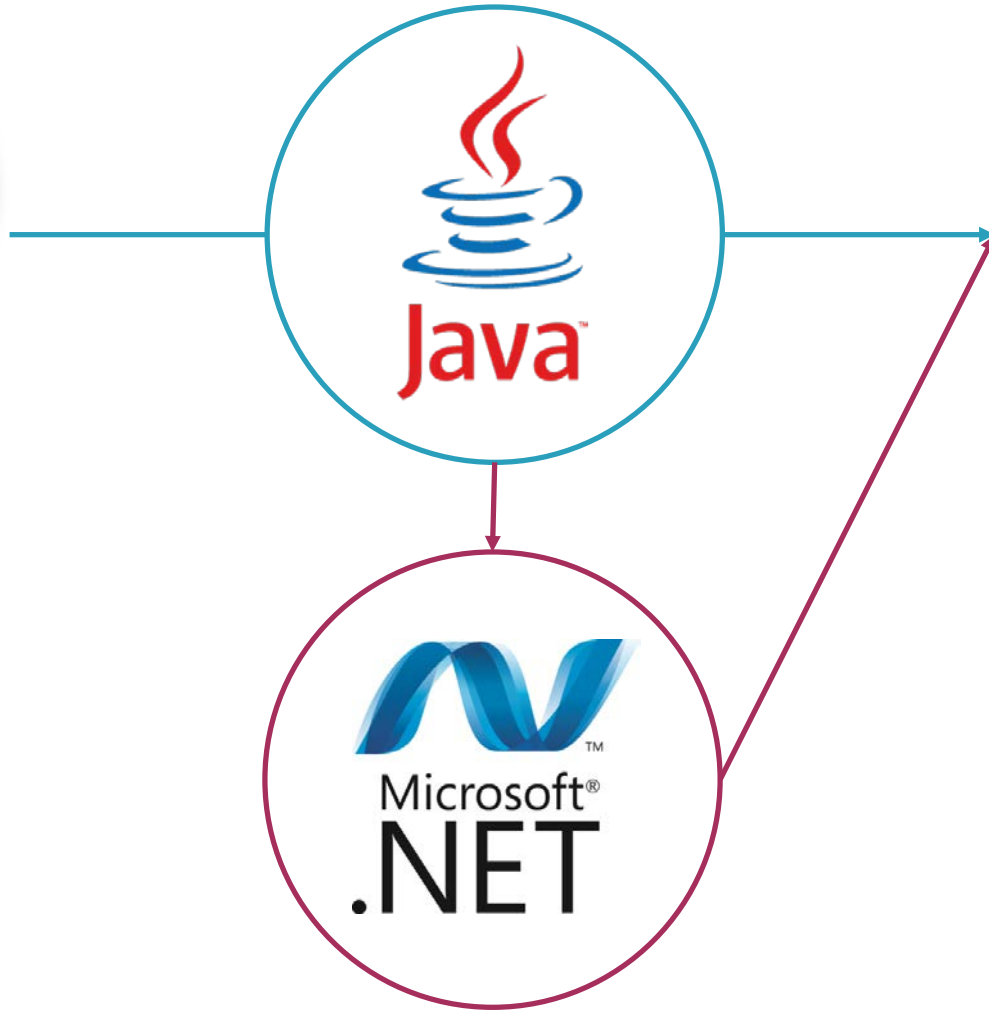
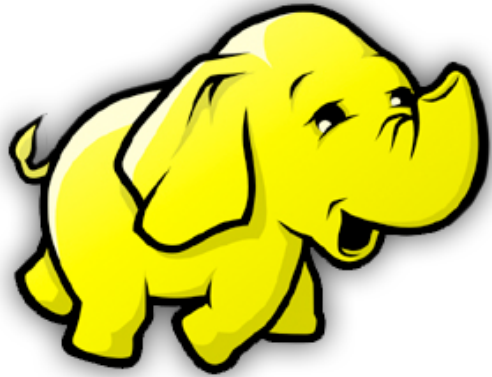


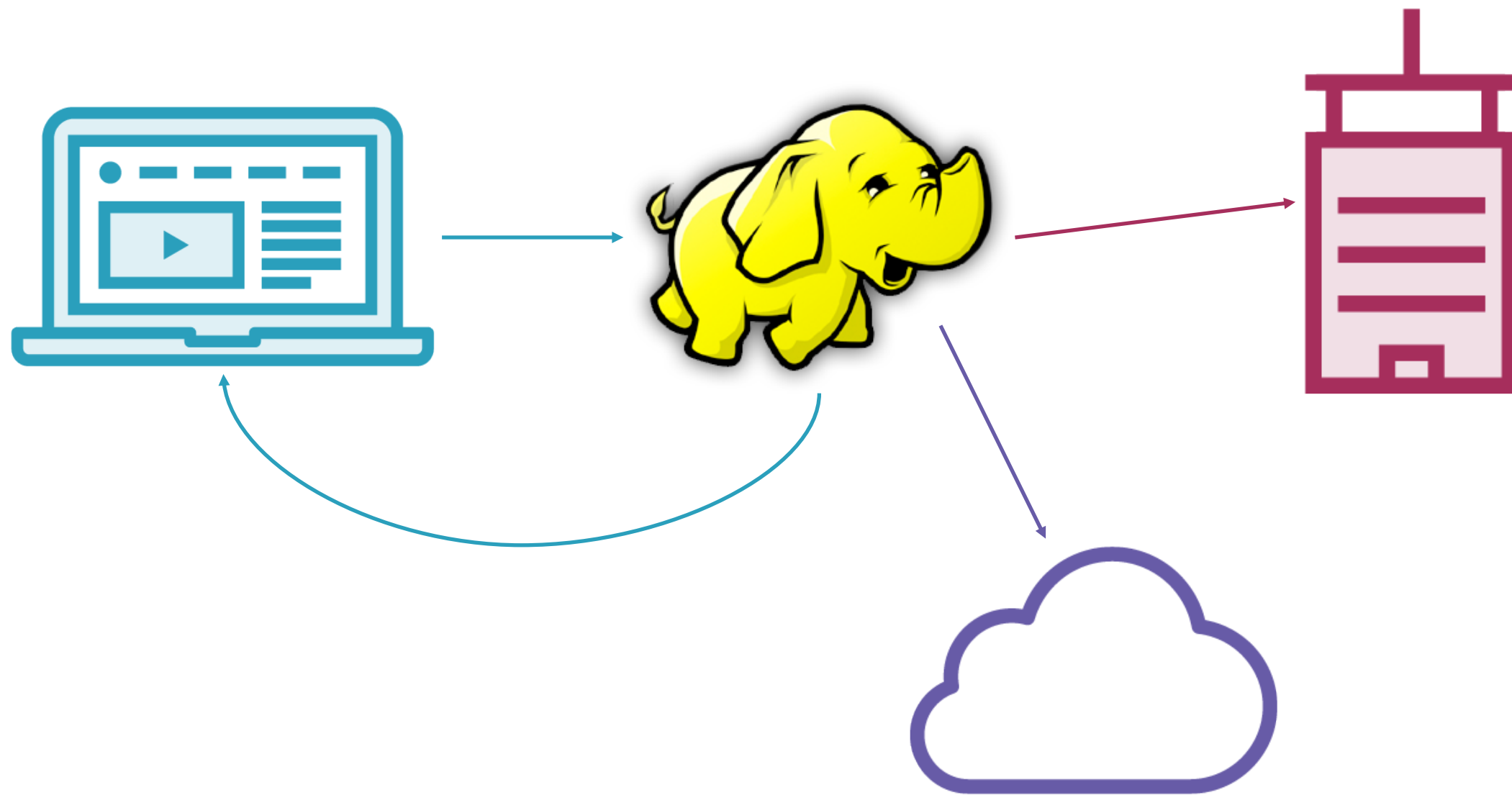
Big Data Platform

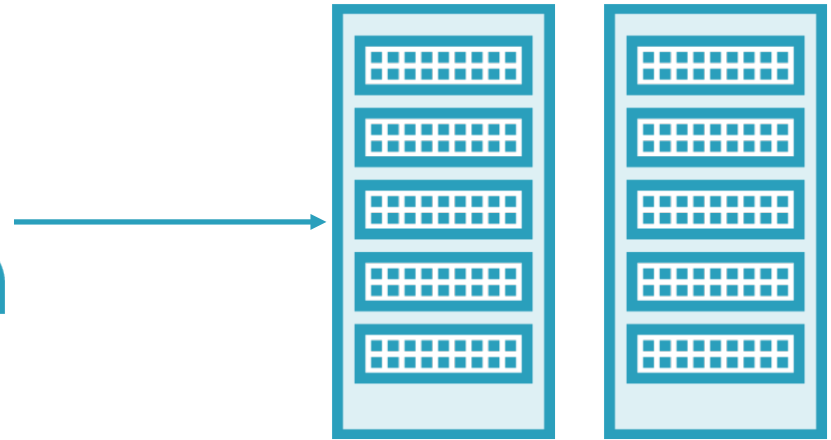
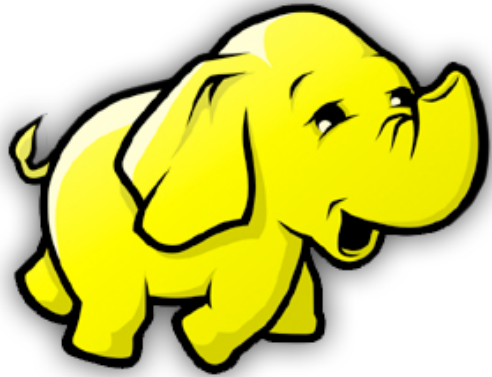


Spark

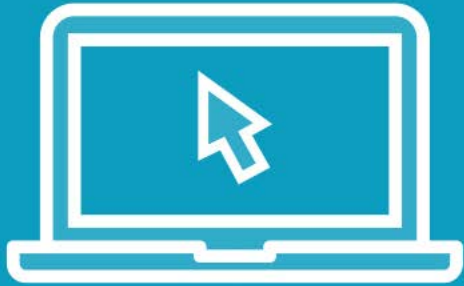






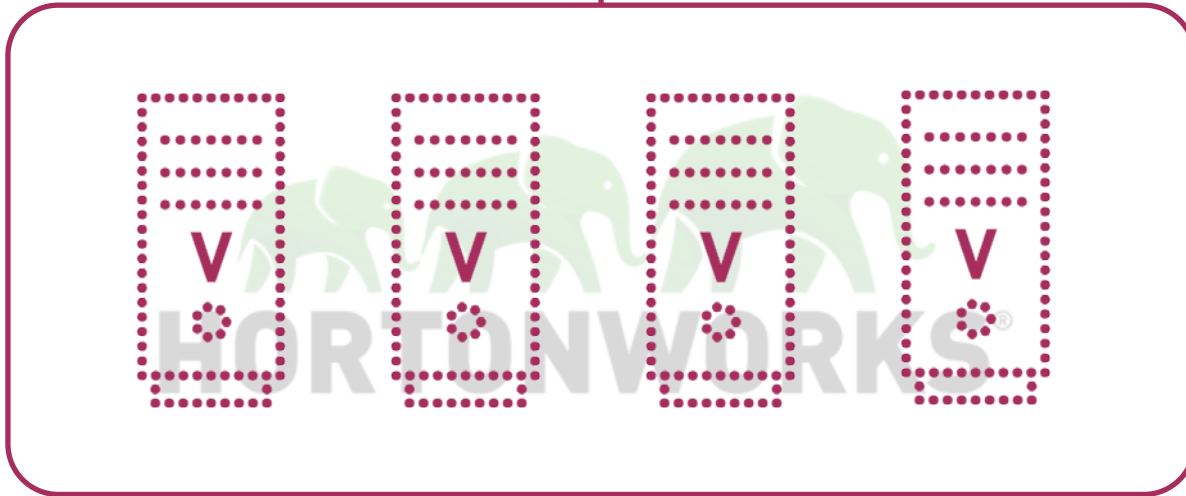


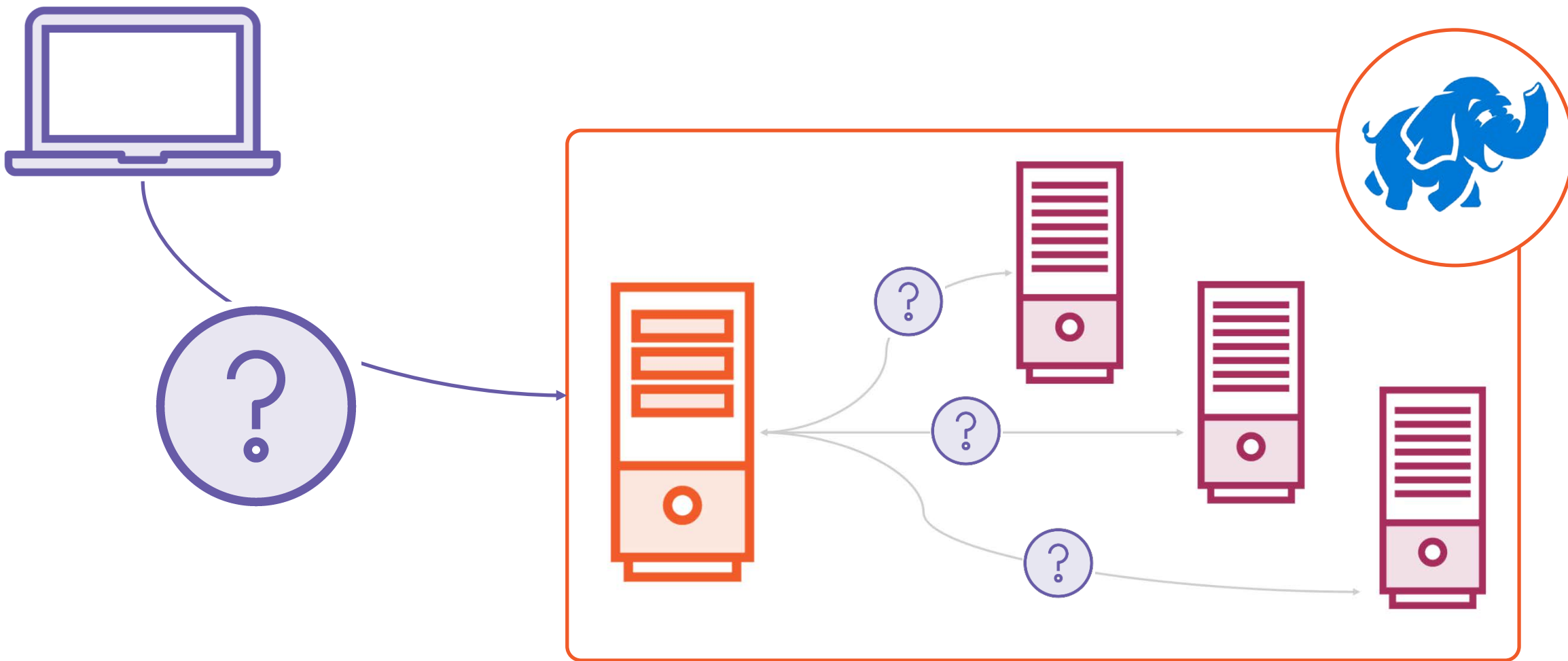
Demo

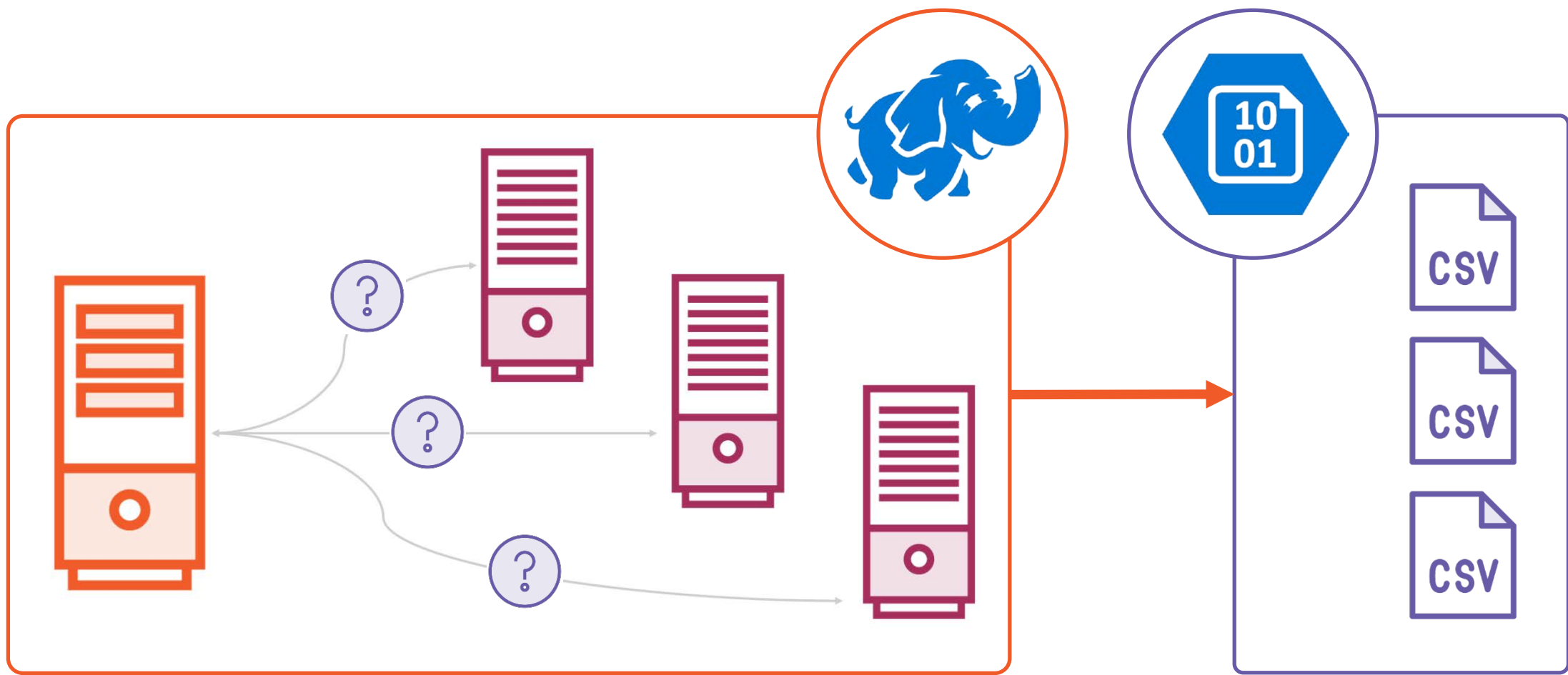


Syncfusion Big Data Platform

- Installing on Windows 10
- Using the Dashboard
- Running .NET jobs

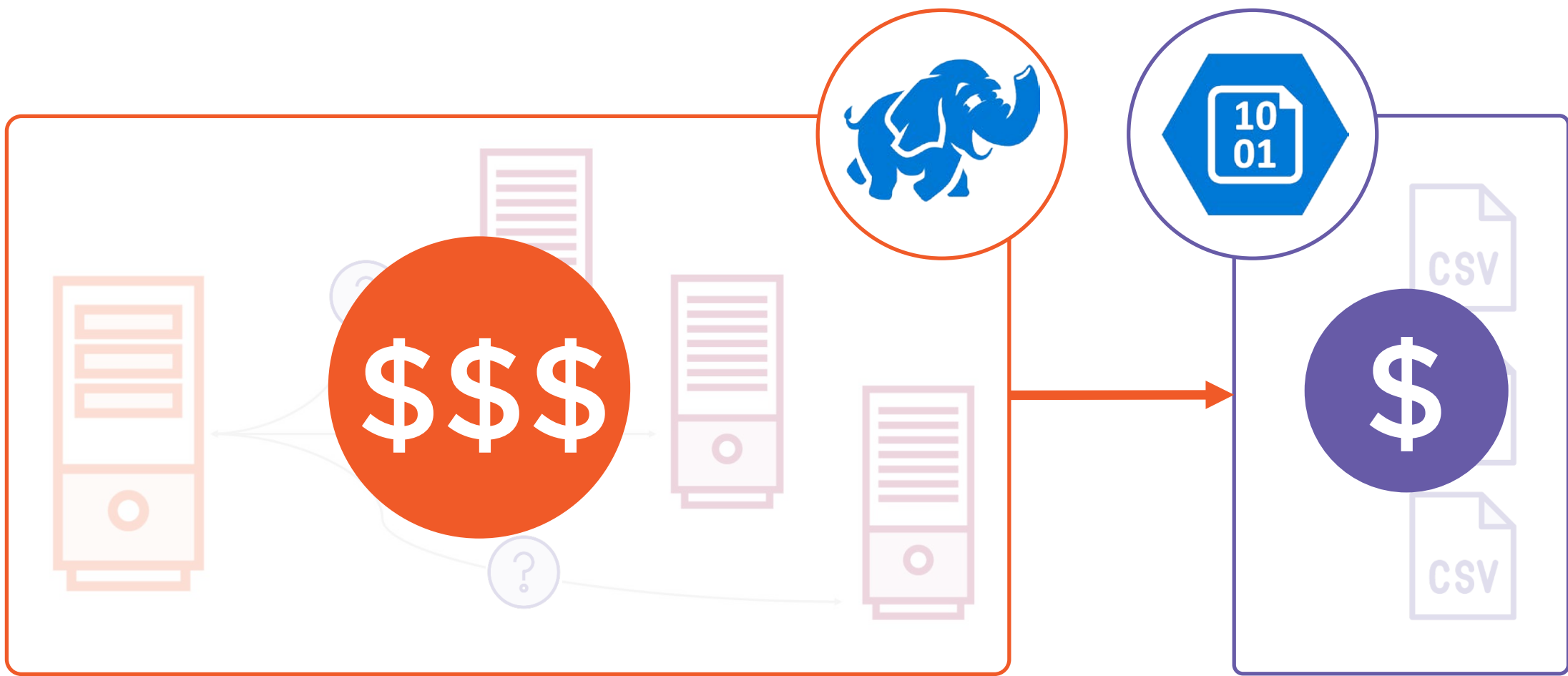






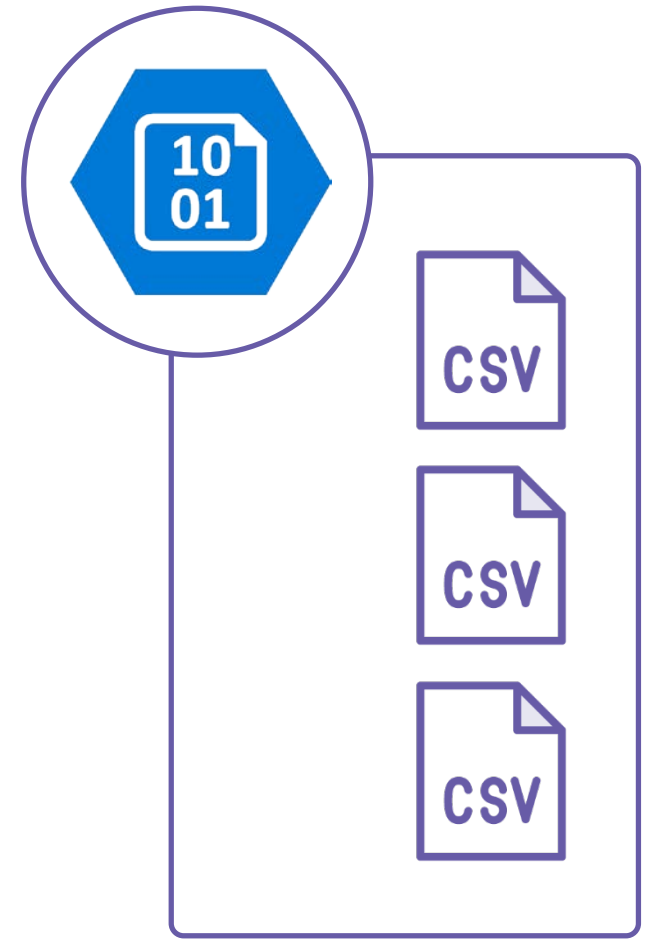
Compute

Storage

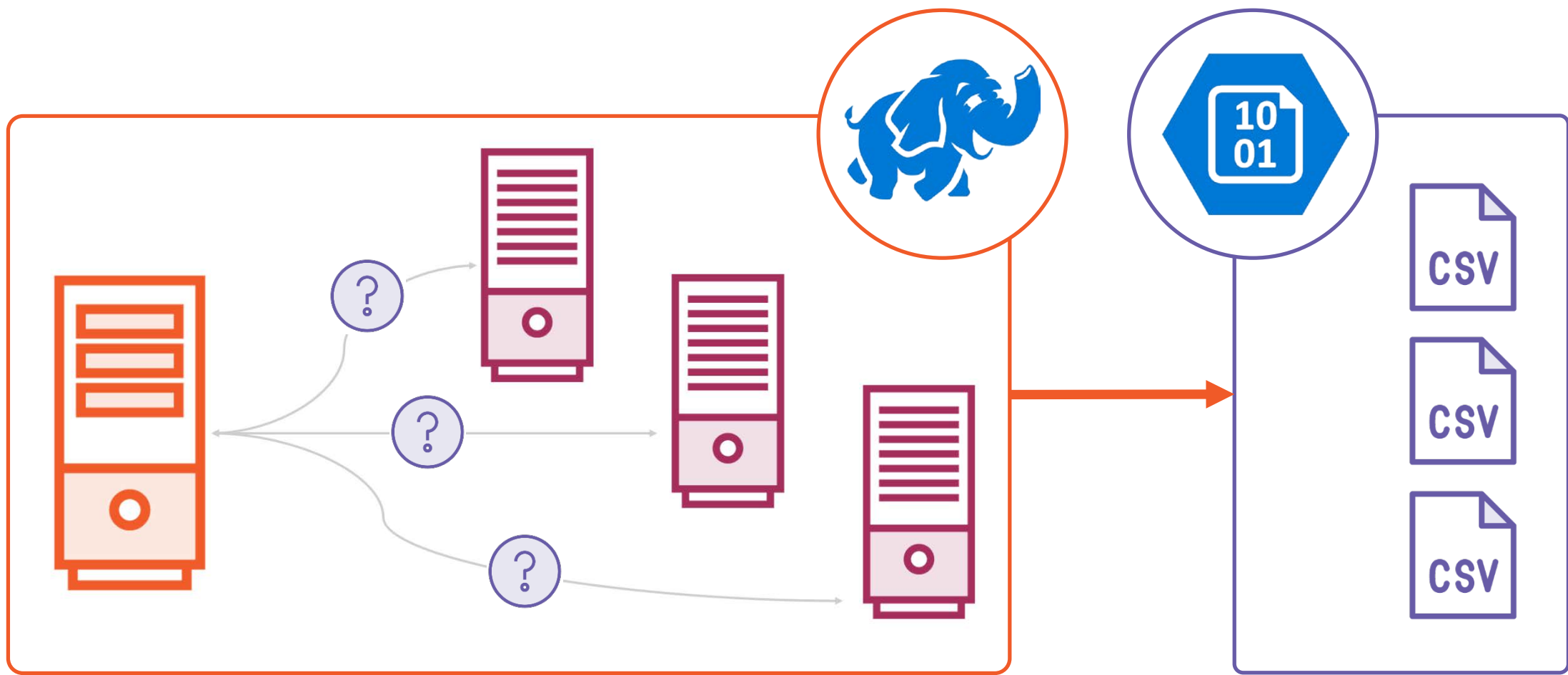


Compute

Storage

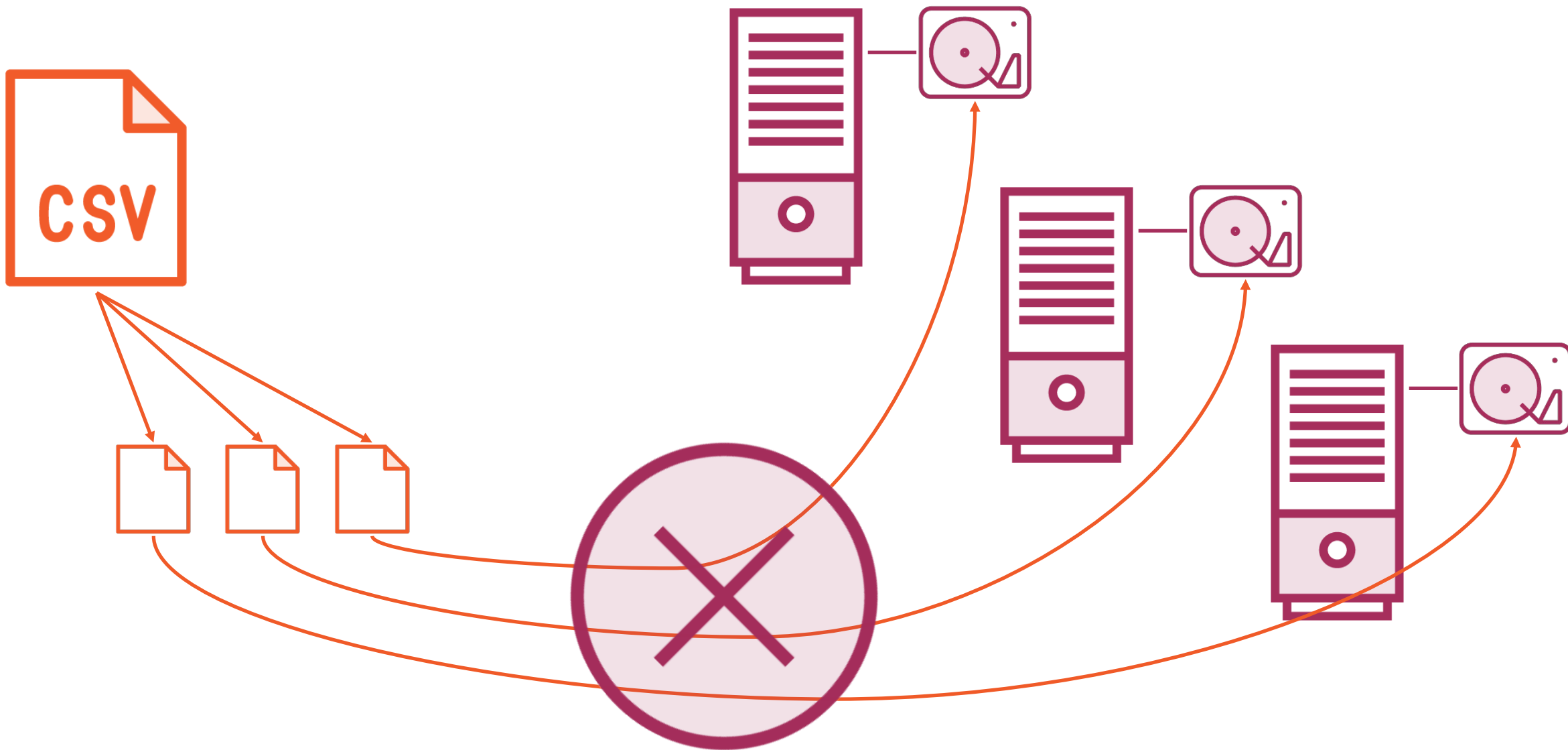


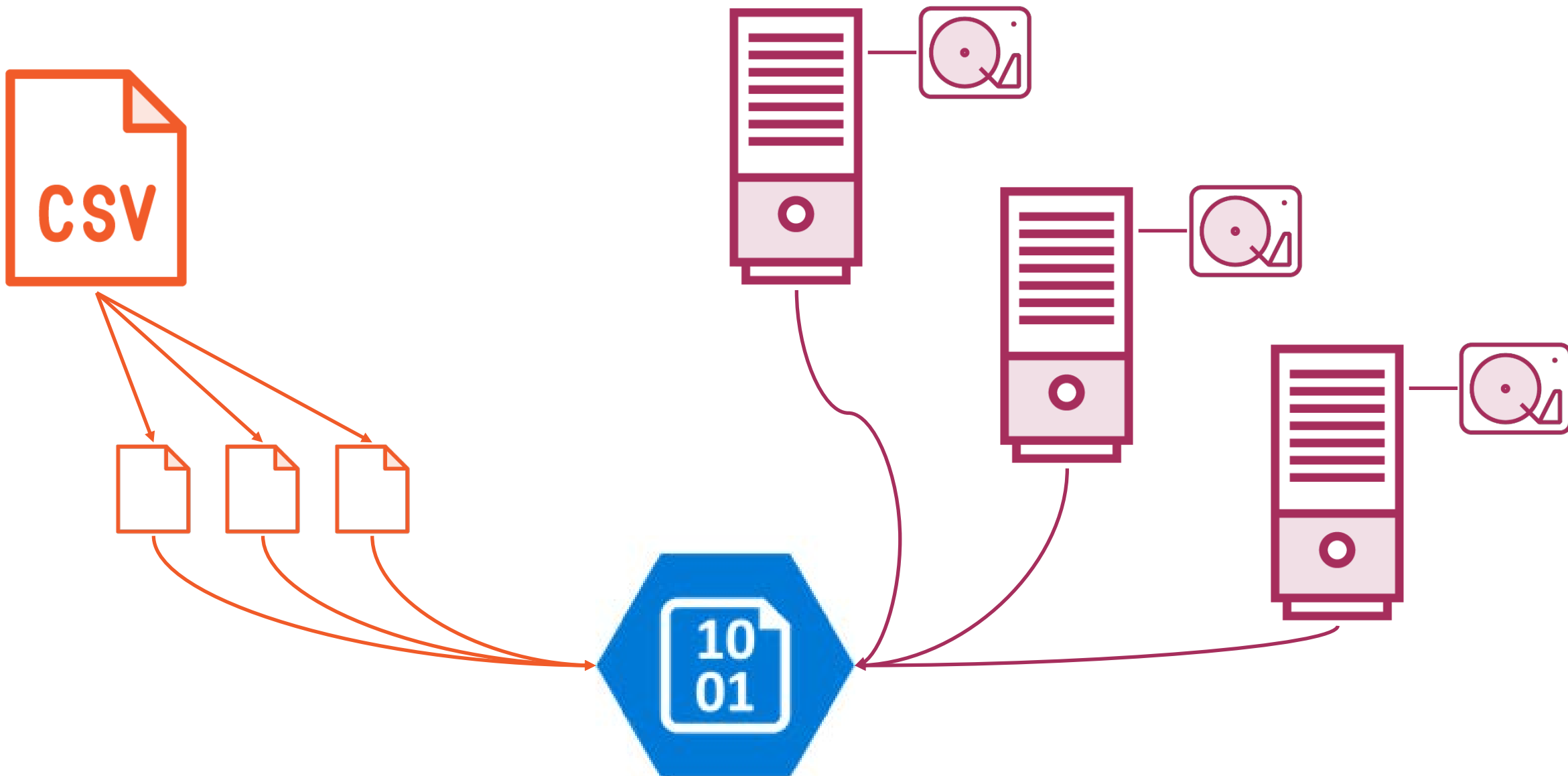
Storage

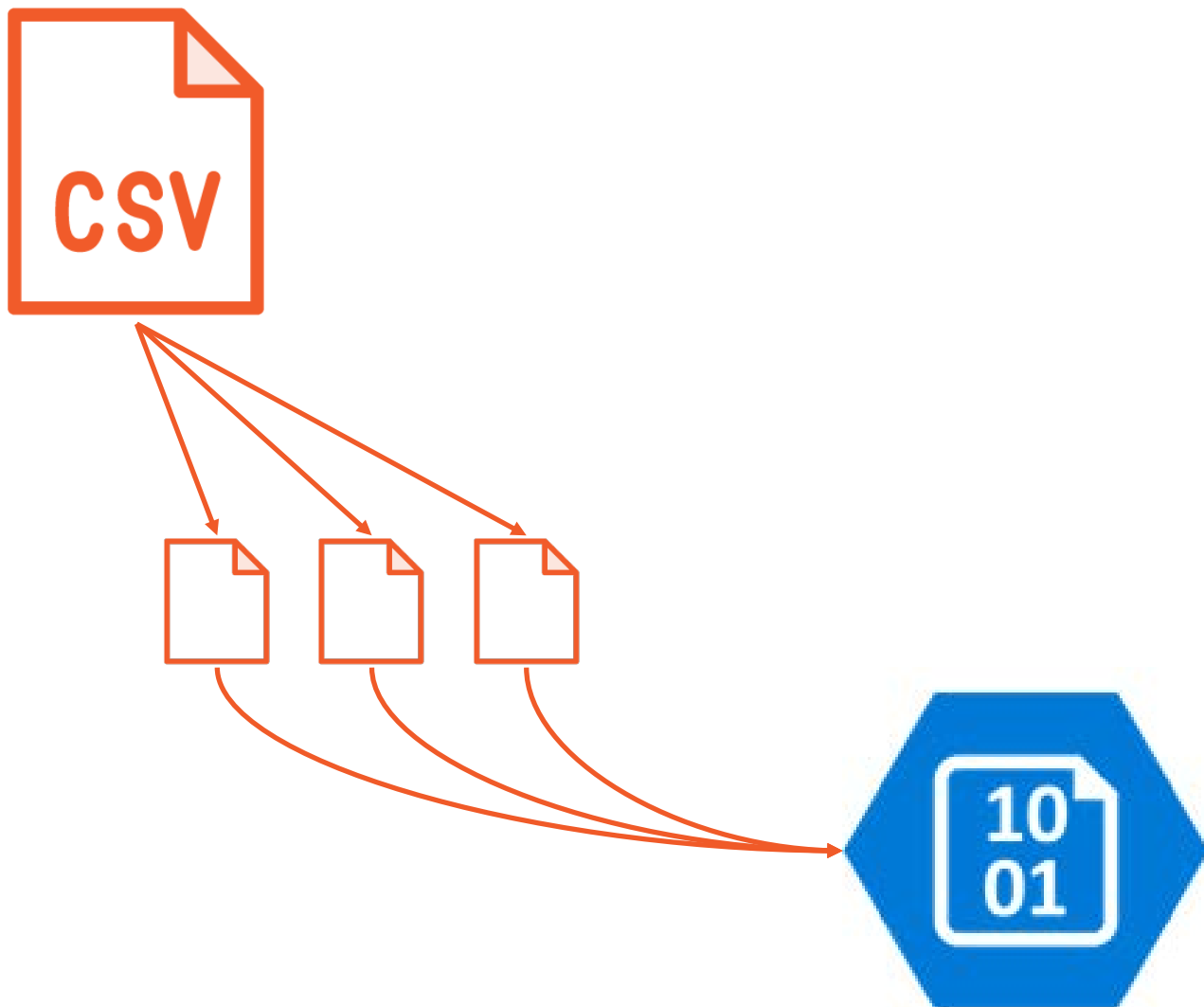


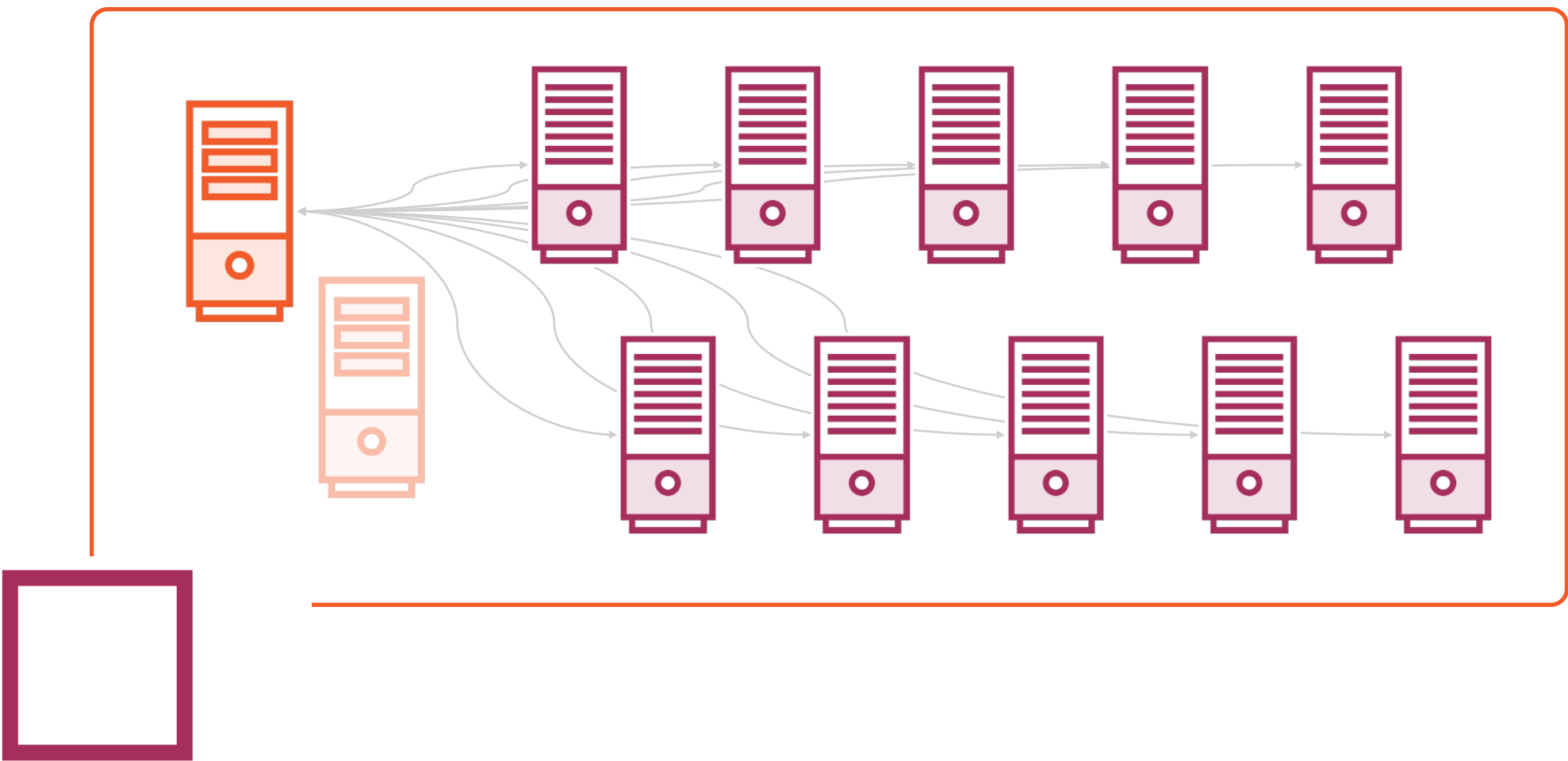
Compute-2

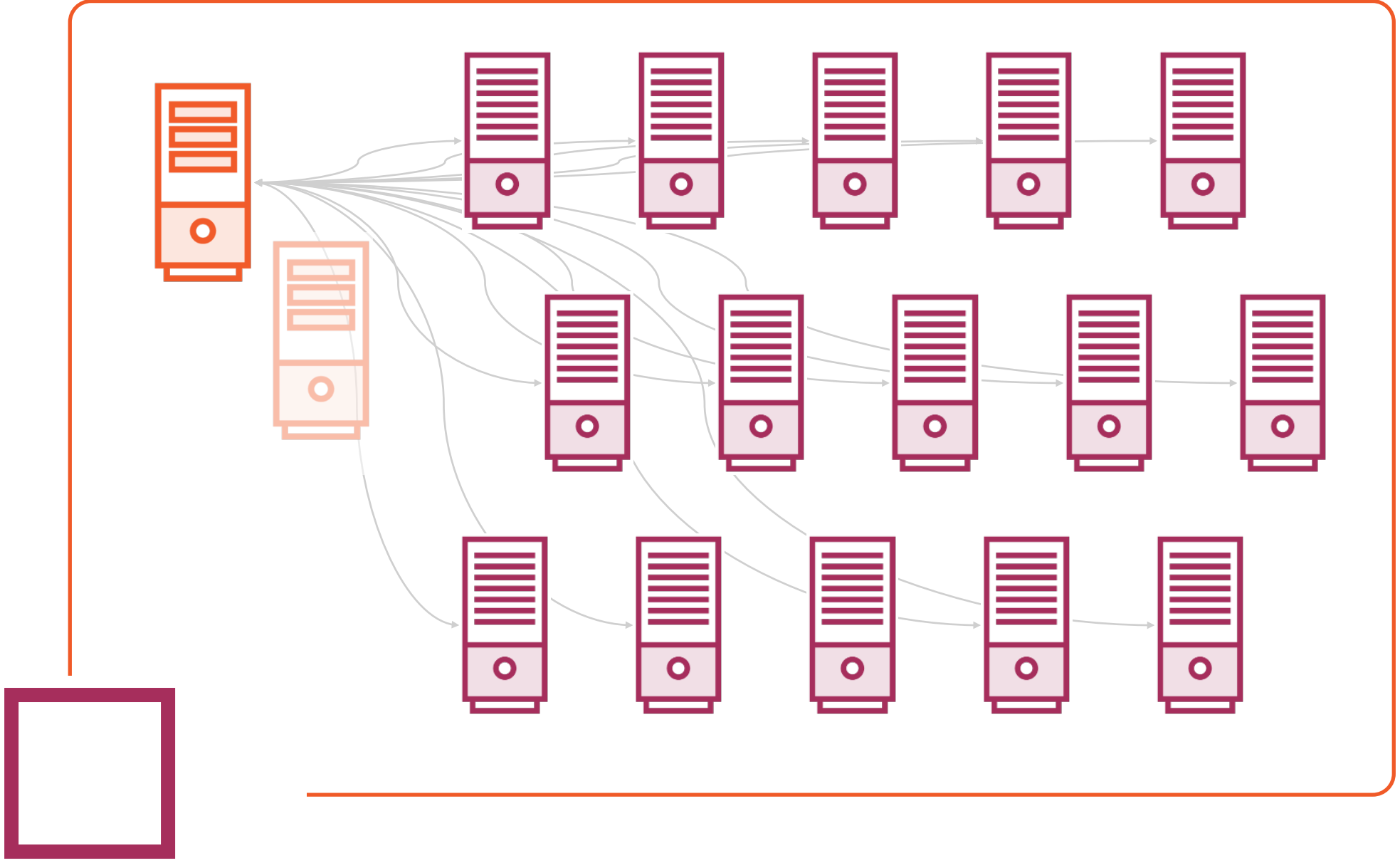
Storage

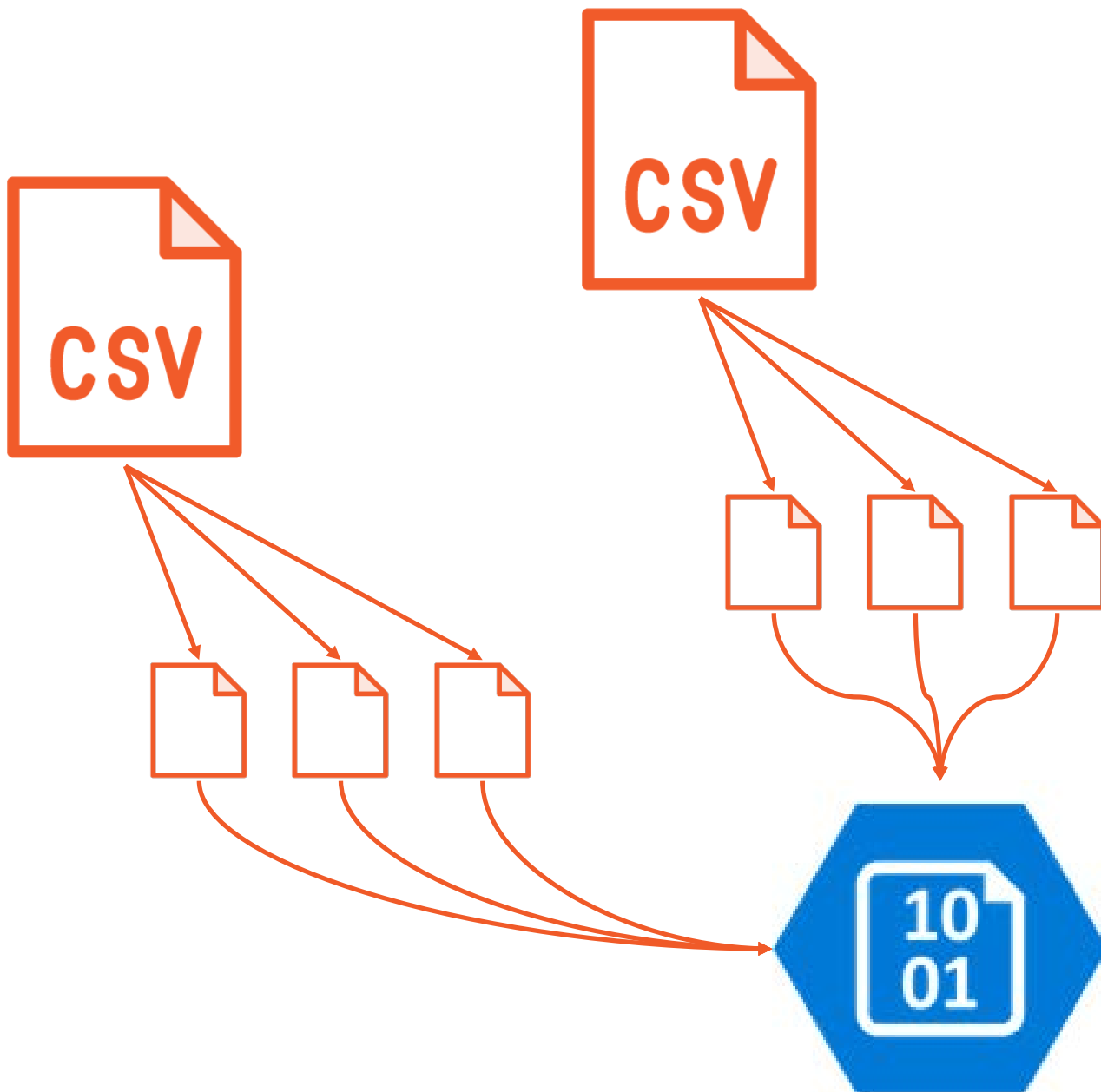




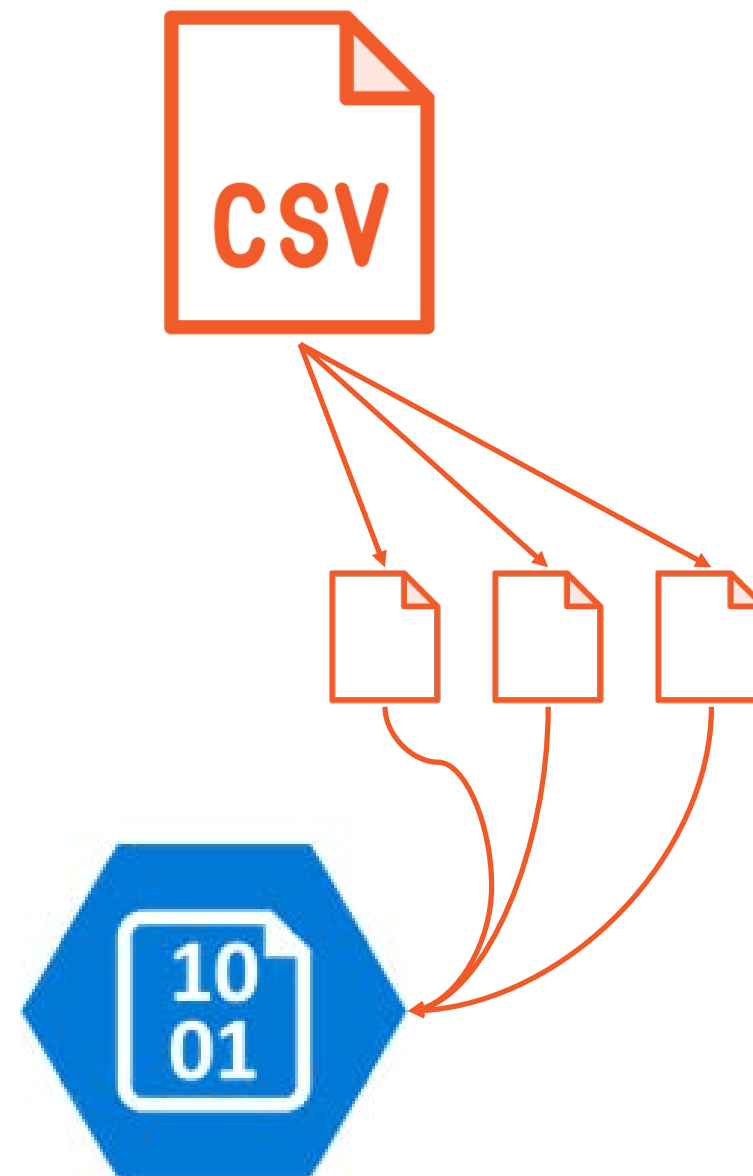
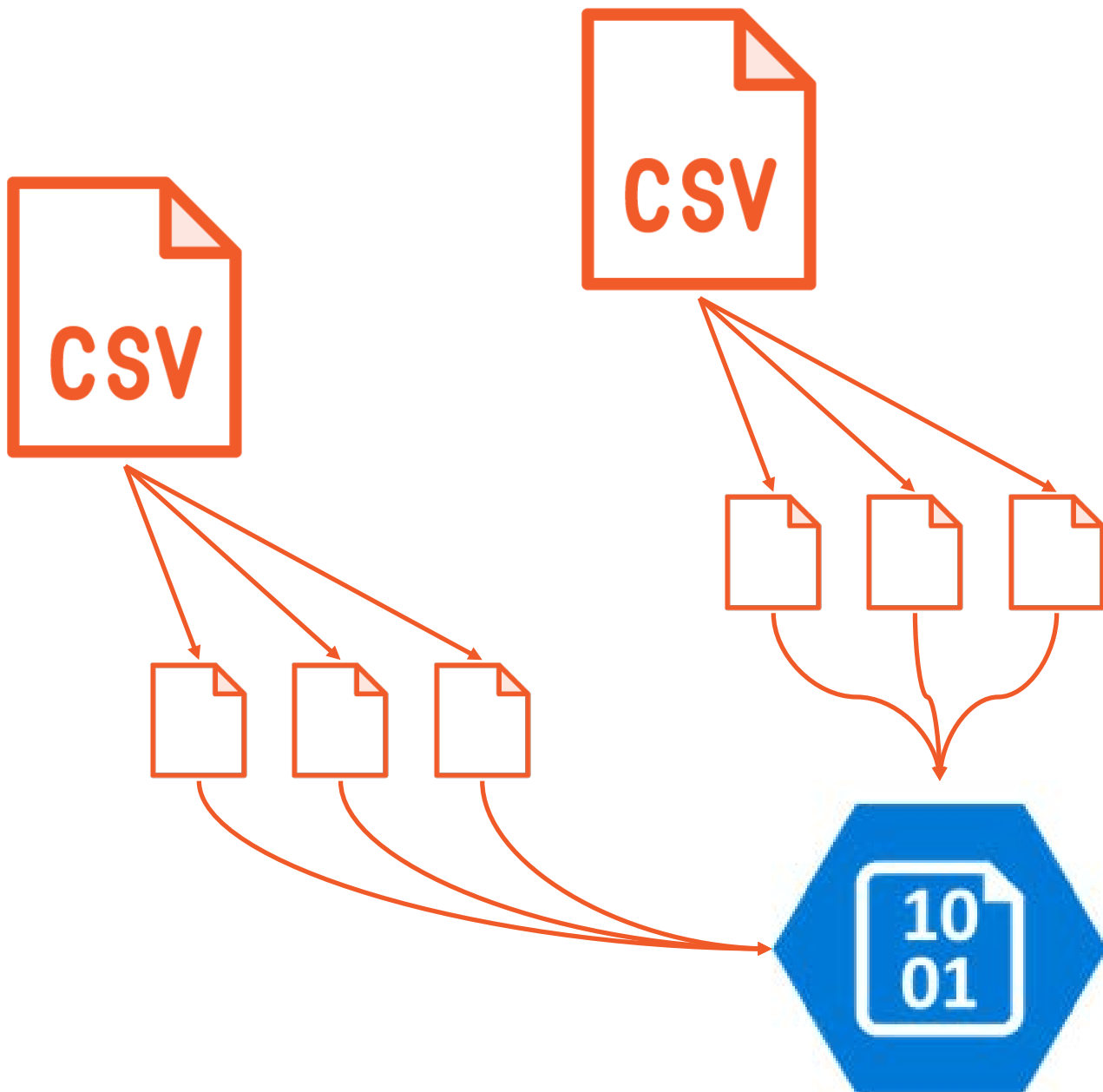


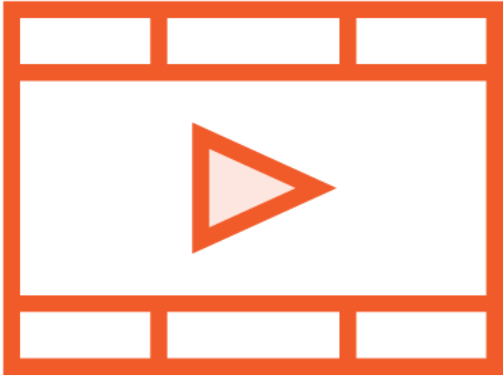






Storage Account
Max 500TB



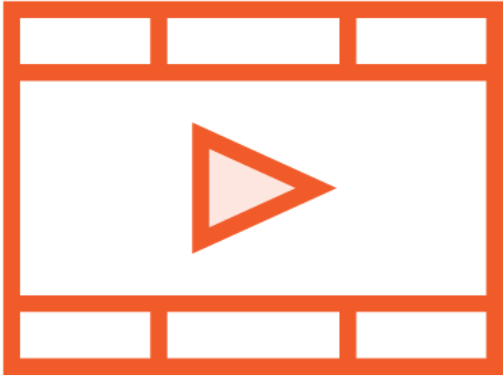


Real World Big Data in Azure



Elton Stoneman

@EltonStoneman blog.sixeyed.com



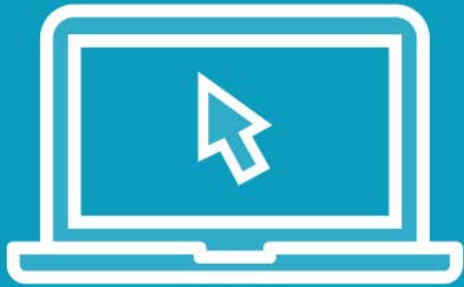
HDInsight Deep Dive: Storm, HBase and Hive



Elton Stoneman

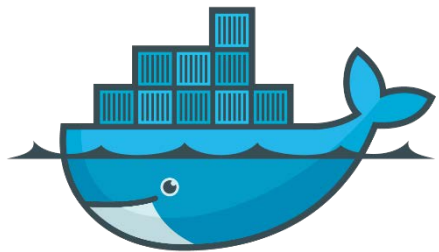
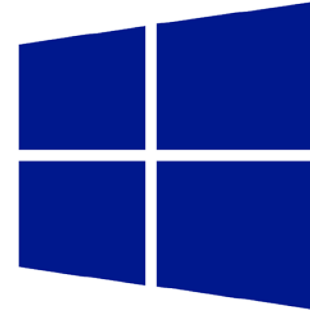
@EltonStoneman blog.sixeyed.com

Demo

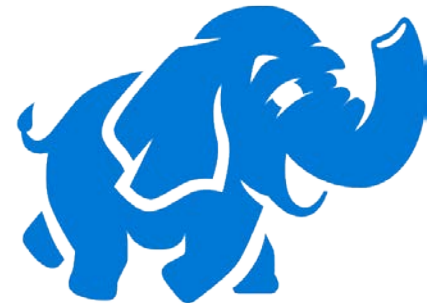
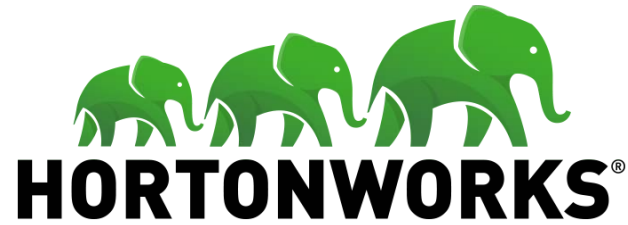


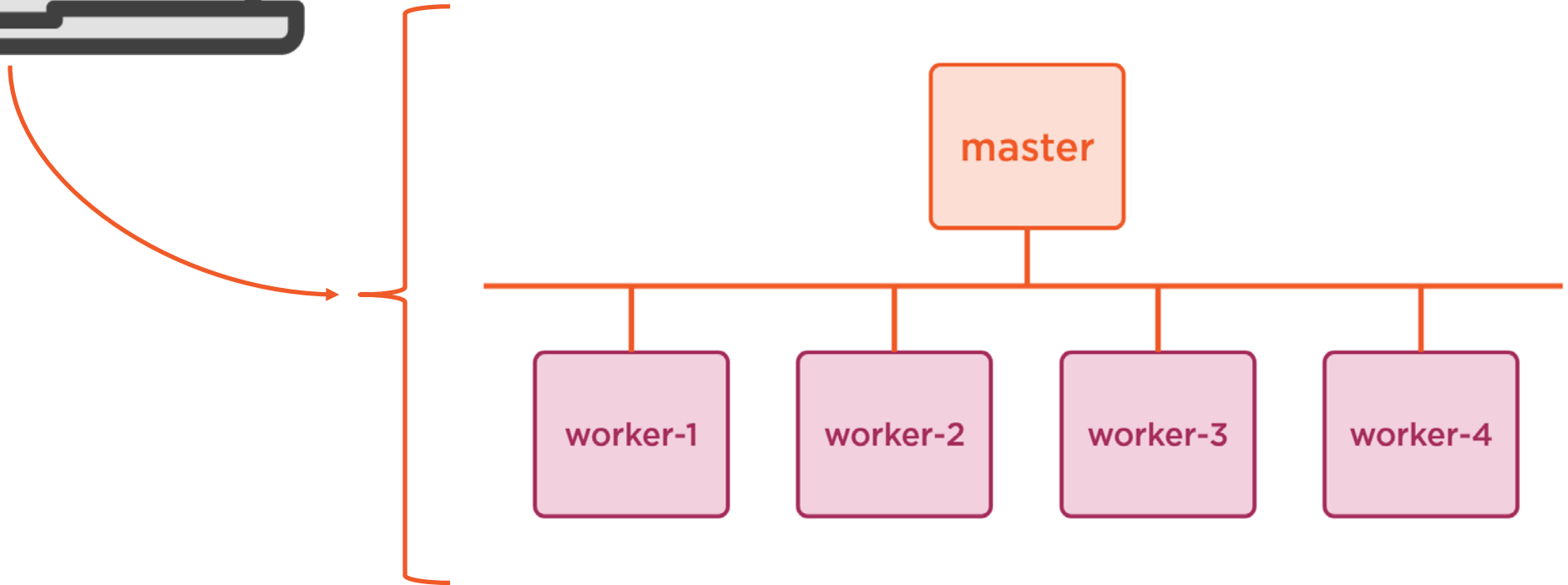
Hadoop on HDInsight

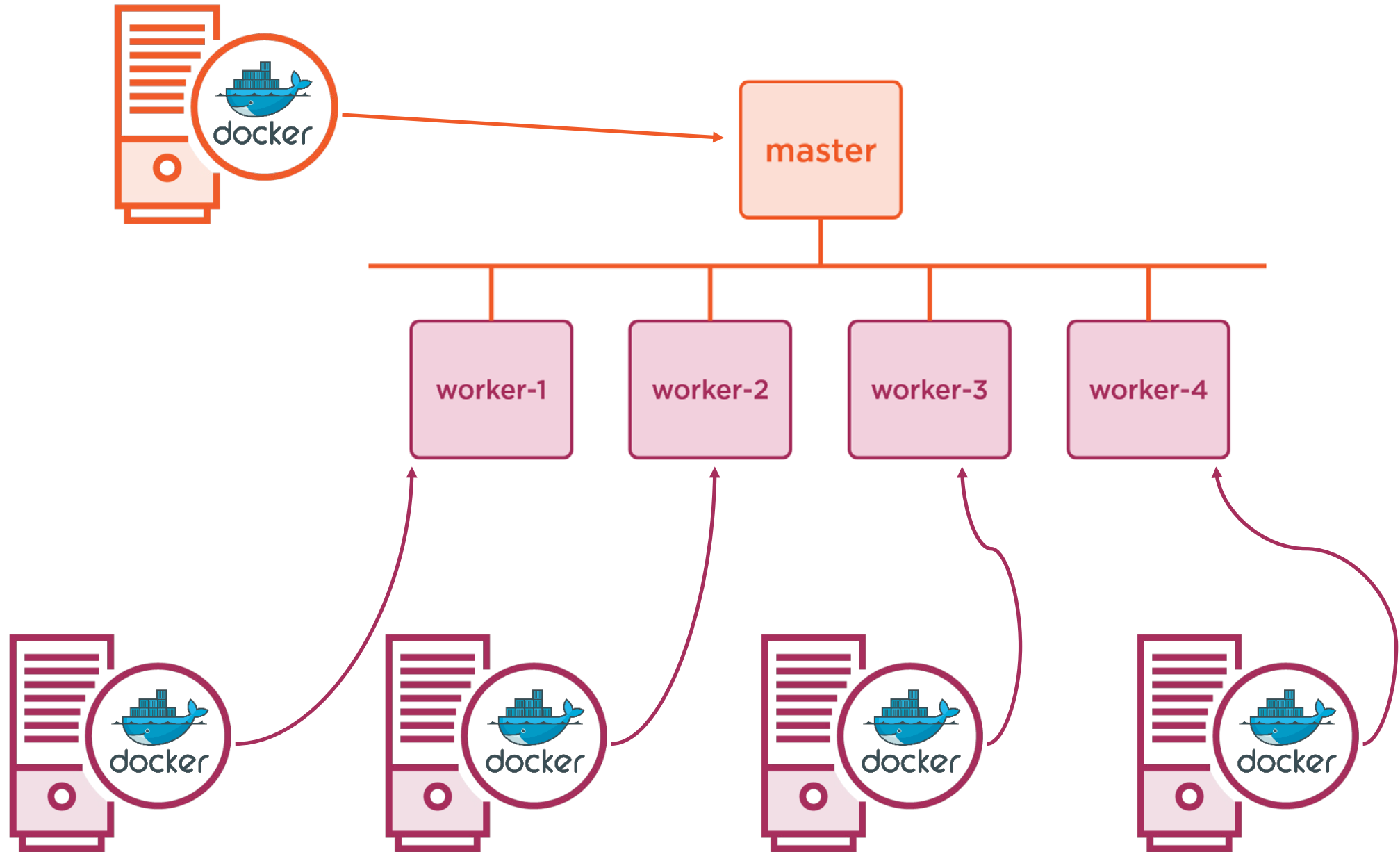
- Creating the cluster
- Uploading data
- Querying with .NET

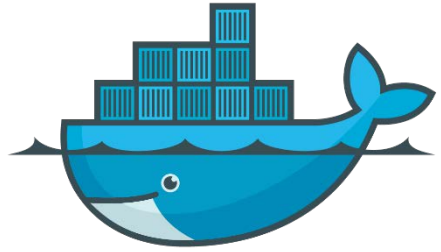


docker









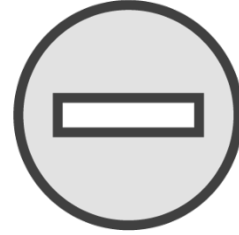
docker



Simple

Pure Hadoop

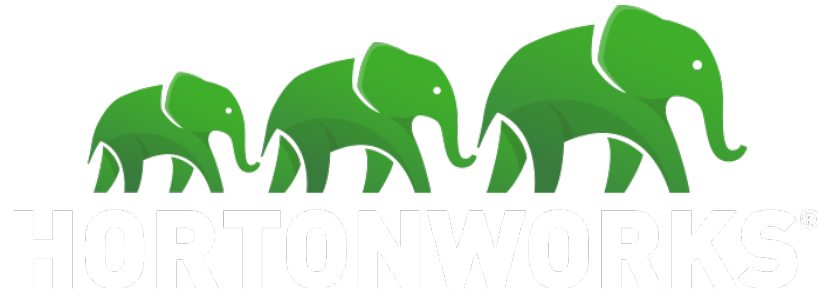
Portable



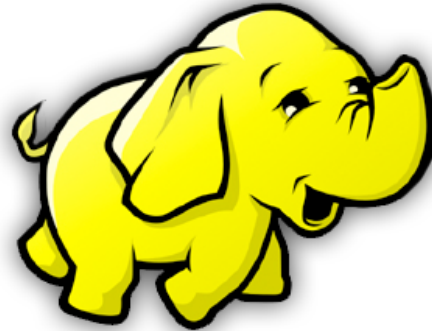
Not Windows-Native

.NET Core only

HDP



APACHE
HBASE

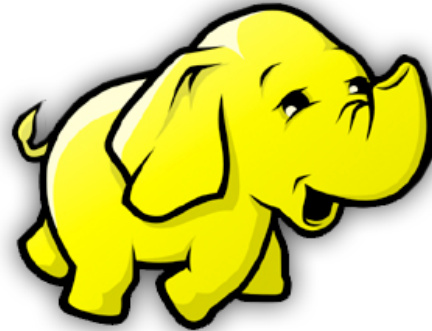


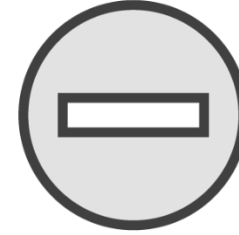
Apache Ranger

Big Data Platform



Spark





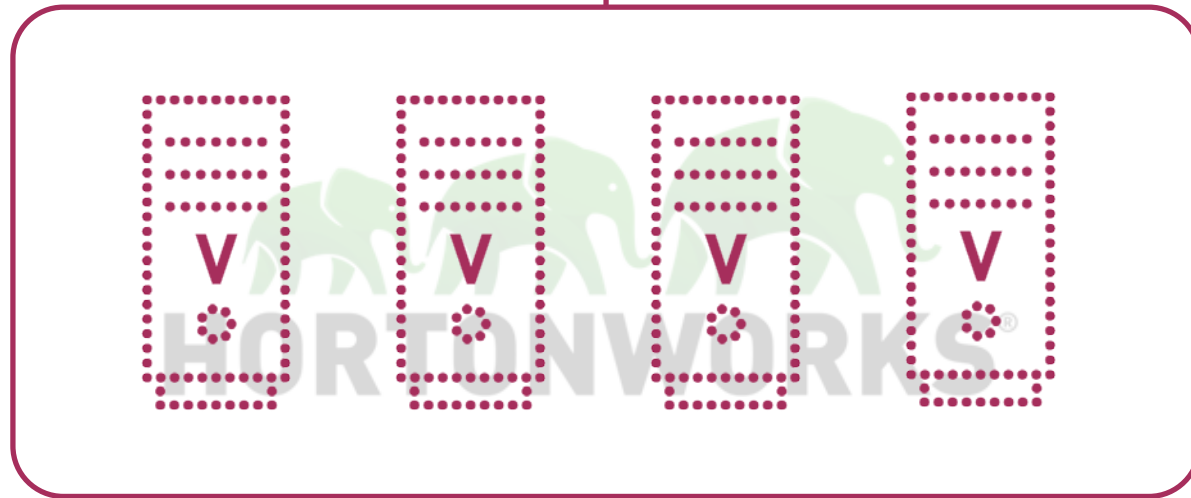
Windows-Native

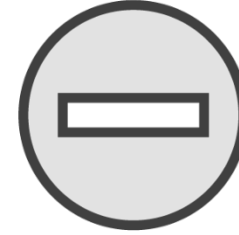
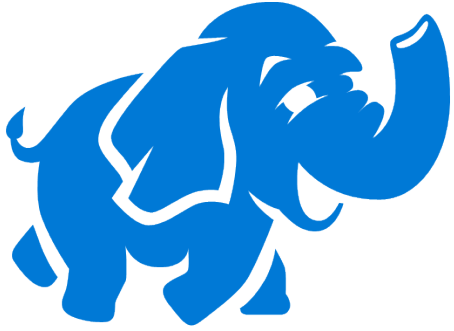
Complex Setup

Full .NET Framework

Hardware Investment

Cluster Management





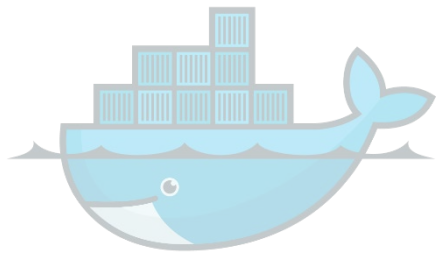
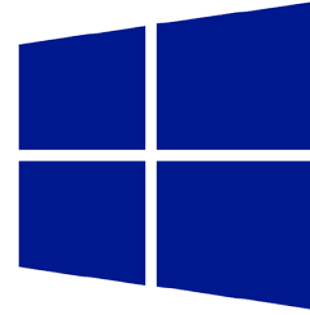
Cheap Storage

Expensive Compute

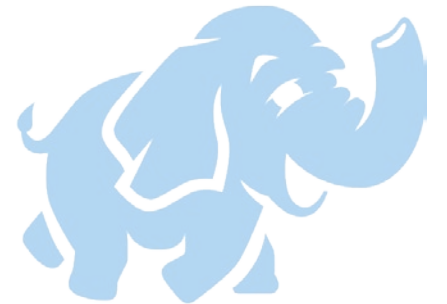
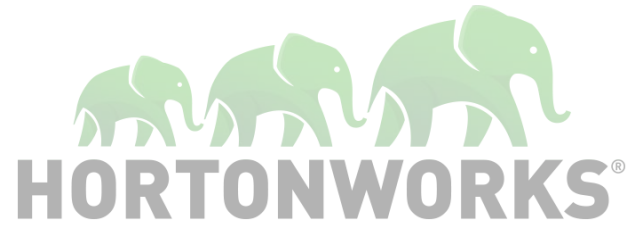
Easy expansion

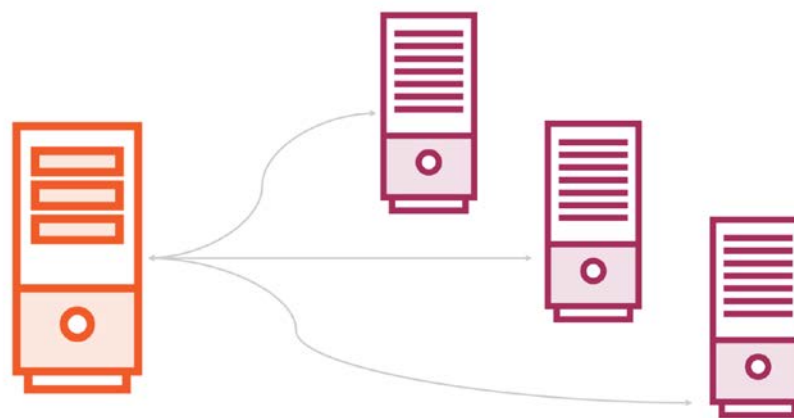
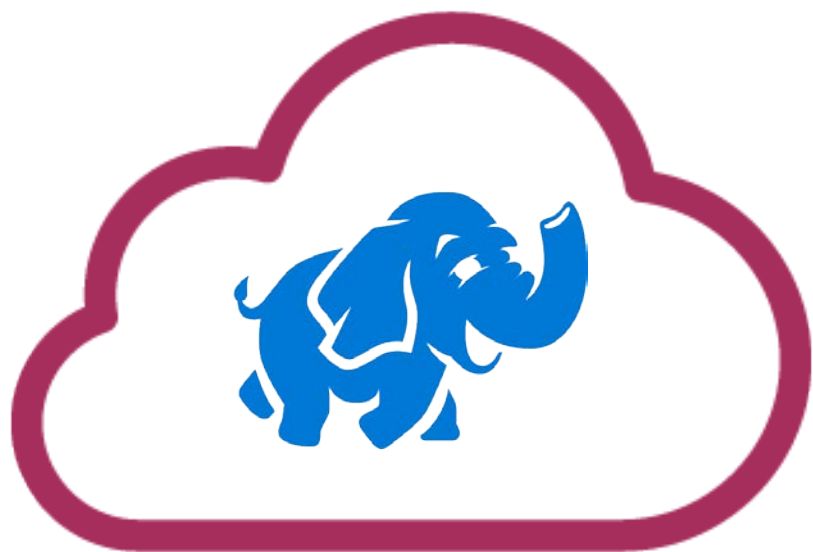
Latency to Cloud

Managed Platform



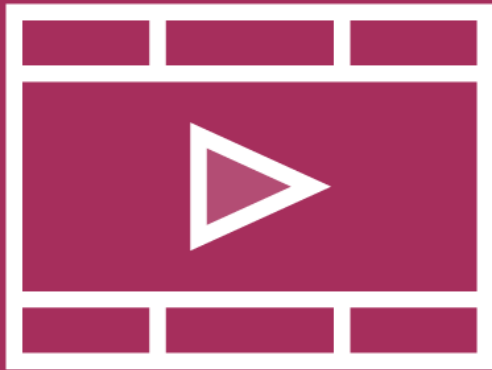
docker





HORTONWORKS®

Coming Next



Working with Hadoop in .NET

- Understanding Hadoop Streaming
- Writing MapReduce in .NET
- Staying portable with .NET Core