Operations Bridge Reporter

Software Version: 10.22
Windows® and Linux operating systems

Performance, Sizing, and Tuning Guide

Hewlett Packard Enterprise

Legal Notices

Warranty

Restricted Rights Legend

Copyright Notice

Trademark Notices

Documentation Updates

To check for recent updates or to verify that you are using the most recent edition of a document, go to: https://softwaresupport.softwaregrp.com/.

This site requires that you register for an HPE Passport and to sign in. To register for an HPE Passport ID, click **Register** on the HPE Software Support site or click **Create an Account** on the HPE Passport login page.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HPE sales representative for details.

Support

Visit the HPE Software Support site at: https://softwaresupport.softwaregrp.com/.

Most of the support areas require that you register as an HPE Passport user and to sign in. Many also require a support contract. To register for an HPE Passport ID, click **Register** on the HPE Support site or click **Create an Account** on the HPE Passport login page.

To find more information about access levels, go to: https://softwaresupport.softwaregrp.com/web/softwaresupport/access-levels.

HPE Software Solutions Now accesses the HPE SW Solution and Integration Portal website. This site enables you to explore HPE Product Solutions to meet your business needs, includes a full list of Integrations between HPE Products, as well as a listing of ITIL Processes. The URL for this website is https://softwaresupport.softwaregrp.com/km/KM01702731.

Contents

Chapter 1: Introduction

HPE Operations Bridge Reporter (OBR) is a cross-domain performance reporting solution. OBR uses SAP Business Objects Enterprise for business intelligence and reporting. The performance metrics that are used for reporting, are stored for longer periods in the Vertica database. OBR consists of several collectors and data processors. The collectors gather performance metrics from various data sources. The data processors transform and enrich the data that is received from the data sources.

The key objective of this guide is to provide steps to arrive at sizing of the hardware needed to deploy the OBR in your environment and to modify various applications, databases, and operating system parameters to achieve optimal performance.

Read the following sections to understand more about the sizing parameters:

Sizing Approach provides the guidelines to determine the size of the deployment and the hardware and software required for various deployments.

Sizing Calculator helps determine the size of the hardware required for various deployments. It is recommended that this calculator is used to arrive at the hardware sizing for an enterprise OBR deployment.

General Recommendations and Best Practices provides general guidelines and best practices to obtain optimal performance from the OBR application, Vertica database, and the operating system.

Benchmark provides details of various performance benchmark tests conducted on OBR. You can use the results of these tests to choose a system configuration for specific OBR loads. Note that these tests were conducted in a controlled environment and should only be used as an indication of the capacity of the system.

Chapter 2: Sizing Approach

Sizing aims to estimate the system resource required to ensure that the deployed system meets the performance objectives.

Sizing the Deployment

The factors that affect hardware sizing are,

- The content you will deploy
- The load for each of the content pack
- The retention period for each of the content

Calculating Content Load

This section provides guidelines to calculate the load for some of the out-of-the-box OBR content. The load is computed based on certain assumptions and approximations. So, while choosing hardware include enough headroom to handle the actual load.

> **Note:** If you are using additional content packs that are not part of OBR media or custom content, ensure that you factor the additional throughput incurred from those content packs. See corresponding content documentation for guidance.

System Content

The size of the environment for system content is determined by

- the total number of physical and virtual nodes (n)
- the average number of file systems per node (fs)
- the average number of disks per node (disk)

- the average number of CPUs per node (cpu), and
- the average number of network interfaces per node (n/w if).

OBR extracts 5-minute summarized data for system content. So total number of records extracted per hour per CI is 60/5 = 12. Hence throughput requirement is computed as,

Total number of CIs (t) = n + n * (fs + disk + cpu+ n/w if)

Throughput requirement is ~ (t * 12) records per hour

Network Content (by NPS integration)

The Network Node Manager i (NNM i) SPI Performance for Metrics determines the size of the environment for network content (by NPS integration) using the following parameters in your deployment:

- number of performance polled network nodes (n)
- performance polled interfaces (n/w if)

OBR extracts hourly summarized data from the network data source. So the throughput requirement is computed as,

Total number of CIs = n + n/w if

Throughput requirement is ~ (n + n/w if) * 1 if per hour

Note: Network content calculations mentioned above do not factor NNMi direct load.

Network Content (by Direct NNMi Integration)

The Network Node Manager i (NNMi) SPI Performance for Metrics determines the size of the environment for network content (by Direct NNMi integration) using the following parameters in your deployment:

- number of performance polled network nodes (n),
- performance polled interfaces,
- number of polled components (c),

- polling time interval (t), and

- the number of policies per interface (p) in your deployment.

Usually, NNMi pushes data from the network data source every 5 minutes (polling time interval). So the throughput requirement is computed as,

Total number of CIs = n + i + c

Interface throughput per hour (IT) = (p*i)*60 /t

Component throughput per hour (CT) = (c+(2*n))*60/t

Throughput requirement is ~ IT + CT per hour

RUM/BPM Content

In the case of RUM/BPM content, the size of the environment is determined by

- the number of transactions (t),

- applications(a),

- locations(l), and

- MAX EPS.

Refer to the *BSM/APM Administration guide* for details on calculating MAX EPS for your environment.

Totals number of CIs ~= applications(a) + transactions(t) + locations(l)

Throughput requirement is ~ (RUM MAX EPS + BPM MAX EPS) * 60 * 60

OMi Content

In the case of OMi content, the size of the environment is determined by the OMi Events Per Second (EPS). Refer to the *OMi Guide* for details on calculating MAX EPS for your environment.

Throughput requirement is ~ EPS * 60 * 60 * 138.889

Service Health Content

In the case of Service Health content, the size of the environment is determined by

- the number of KPIs configured for logging (k), and

- HIs configured for logging (h).

Refer to the *BSM/OMi Administration Guide* for details on calculating MAX EPS for your environment.

Totals number of CIs ~= k + h

Throughput requirement is ~ (k + (h * 139)) * 12

Throughput Calculation and Custom Content in OBR

In addition to the content packs mentioned earlier, you need to calculate the throughput for other out-of-the-box content packs and custom content packs. Throughput for each content pack is calculated based on the number of records added to the database per hour. This in turn is calculated from the CSV files collected per hour.

For example, assume that a content pack having 2 rate tables is deployed in a smaller environment. Every hour, it collects CSV files with 1000 and 2000 records for each table respectively. Summing up the number of records pumped in by each CSV, the throughput for the content pack would be 3000 without considering any customer procedure or aggregation.

In the case of OMi even though the number of events, and the number of HI and KPI instances are not very high, and although the number of lines in the CSV file is relatively less, the throughput seems to be much higher. This is because for calculating the throughput many other parameters are taken into consideration, especially the parameters influencing the custom procedure execution.

Every content pack has aggregates and custom procedures that create more records. You need to factor these records as part of the hourly throughput. While preparing for an OBR setup with multiple content packs, you need to calculate the throughput based on the records coming in from the aggregates and custom procedures. Add the calculated throughput to the additional throughput section of the sizing calculator to get more accurate and realistic configuration requirements.

Retention Period

You should determine the retention period for each of the content. Out-of-the-box retention period for different summary tables is shown in Table 1. You should plan for more disk space on the Vertica database server if you increase the retention period.

Table 1: Out-of-the-box Retention Period

Table Type	Default Retention (Days)
Raw	90
Hourly	365
Daily	1,825

Deployment Size

In OBR, the deployment size is categorized as small, medium, large, and enterprise, based on the number of CIs collected from data sources. Small, medium, large and enterprise deployment correspond to 500, 5000, 20000, 40000 nodes respectively. Total number of CIs and throughput requirement for these deployments is shown in the following tables:

Table 2: Total CIs and Throughput requirement in Operations Bridge Deployment (with APM and Network)

Deployment Size	System Nodes	Network Nodes	Network Interfaces	Application (RUM + BPM)	RUM Event Rate	Total Number of CIs	Throughput Requirement (records/hour)
Small	500	5,000	10,000	~100/sec	10/sec	~30K	~600K
Medium	5,000	10,000	50,000	~300/sec	20/sec	~220K	~3,200K

Table 3: Total CIs and Throughput requirement in Operations Bridge Deployment (Systems and Events)

Deployment Size	System Nodes	OMi Events (per sec)	Total Number of CIs	Throughput Requirement (records/hour)
Medium	5,000	3	~160K	~3,400K (3.4 million)
Large	20,000	6	~640K	~10,700K (10.7 million)

Table 4: Total CIs and Throughput requirement in OM Deployment

Deployment Size	System Nodes	Network Nodes	Network Interfaces	Total Number of CIs	Throughput Requirement (records/hour)
Small	500	5,000	10,000	~30K	~200K
Medium	5,000	10,000	50,000	~220K	~2,000K
Large	20,000	20,000	70,000	~730K	~8,000K (8 million)

Table 5: Total CIs and Throughput requirement in Enterprise Deployment (System)

Deployment Size	System Nodes	Total Number of CIs	Throughput Requirement (records/hour)
Enterprise	40,000	~1280K	~16,000K (16 million)

For High Availability environment, the sizing for the shared and local storage can be calculated as follows:

- Shared storage for SAP BusinessObjects : *<space allocated for disk as mentioned in this document>*

- Storage for HPE Vertica: *<dbpace allocated for a 3-node cluster as mentioned in this document>*

- Local storage, CPU, and memory must be of the same specification on both the OBR servers in high availability environment. Similarly, both the BusinessObjects high availability servers (if OBR and BusinessObjects components are installed separately) must have identical configuration.

- In case of Vertica clustering, each of the nodes in the Vertica cluster must have identical configuration.

Table 6: CI Distribution Details

Data Source/Content			Small	Medium	Large	Enterprise
Agent		System Node	500	5,000	20,000	40,000
		File System	5,000	50,000	200,000	400,000
		Disk	5,000	50,000	200,000	400,000
		Network	2,500	25,000	100,000	200,000
		CPU	3,000	30,000	120,000	240,000

Data Source/Content		Small	Medium	Large	Enterprise
BPM	Applications	20	50	1,000	1,500
	Transactions	100	500	5,000	7,500
	Locations	10	50	1,000	1,500
	Trx-Loc Combinations	500	5,000	200,000	300,000
	Max EPS	1	10	220	300
RUM	Applications	5	20	100	150
	Transactions	150	500	5,000	7,500
	End User groups	100	500	10,000	15,000
	Locations	50	500	10,000	15,000
	Servers	5	15	100	150
	Events	10	50	100	150
	Trx-Loc Combinations	2,000	25,000	200,000	300,000
	Max EPS	100	300	900	1,500
NNM iSPI Performance for Metrics (integration with NPS database)*	Polled addresses	5,000	10,000	20,000	30,000
	Polled interfaces	10,000	50,000	70,000	100,000
OMi	OMi EPS	1	3	5	10

*Here we assume that network nodes consume hourly data from the NPS database.

NNMi direct integration is not certified for scale out scenarios in large enterprise deployments.

Above calculations include only the content that contributes the largest load to OBR. You should provide enough headroom for other content like KPI, HI, and so on.

> **Note:** You should also do a similar exercise for the custom content that you deploy in OBR.

Hardware Configuration

This section lists hardware configuration for Operations Bridge Reporter, based on throughput handled. For optimal performance of OBR in enterprise deployment scenarios, ensure that you supply optimal hardware resources (CPU, RAM, and Disk Space) as recommended by the *Sizing Calculator*.

> **Note:** Points to note:
>
> 1. OBR sizing calculations factor Vertica recommendations as suggested at the below given links:
>
>   https://my.vertica.com/documentation/vertica/configuring-hardware-and-virtual-machines/
>
>   http://my.vertica.com/docs/Hardware/HPE_Vertica_HW_planning.pdf
>
>   The lab benchmark tests done on Vertica are suggestive of 1:4 ratio for CPU: Memory. See table for other recommendations of Vertica.
>
>   In the unlikely event of any issue, Vertica recommendations for such cases would need to be honored for enterprise deployments.
>
> 2. SAP BusinessObjects requires a minimum of 16 GB RAM and 16 GB disk space.
>
> 3. Deployment sizes recommended in following tables are based on possible estimated throughput incurred in each scenario. For sizing enterprise deployments, it is recommended to arrive at the right throughput across expected content pack deployments and the corresponding sizing using the *Sizing Calculator*.

Table 7: Hardware and Software Configuration for a Single-System Deployment

Managed Environment Size				System (All-In-One) Configuration			
Deployment Type	No. of Content Packs	Throughput Requirement (records/hour)	No. of additional DPS*	CPU(64 bit) x-86-64	RAM (in GB)	Disk Space for DB	Disk space for s/w **
Small	4	~200K	0	24 CPU Cores	80	1.7 TB	900 GB

Data Processing Server (DPS) - Each additional DPS should be of 4 CPU, 16 GB RAM, 250 GB Hard Disk Drive (HDD).

*** This column captures disk space requirement for software (about 100 GB) and the requirements for shared storage (or install directory if there is no DPS scale out) for run-time data.*

Table 8: Hardware Configuration for a Distributed Deployment

Managed Environment Size				OBR + BO System Configuration			Vertica System Configuration 3-Node Cluster (per node)		
Deployment Type	No. of Content Packs	Throughput Requirement (records/hour)	No. of additional DPS*	CPU (64 bit) x-86-64	RAM (in GB)	Disk space for s/w **	CPU (64 bit) x-86-64	RAM (in GB)	Disk Space
Medium	4	~2,000K	2	24 CPU Cores	48	1.2 TB	20 CPU Cores	80	2.4 TB
Large***	4	~8,000K (8 million)	5	24 CPU Cores	48	1.7 TB	24 CPU Cores	96	4.6 TB
Enterprise***	4	~16,000K (16 million)	8	24 CPU Cores	48	2.4 TB	24 CPU Cores	96	7.1 TB

*Data Processing Server (DPS) - Each additional DPS should be of 4 CPU, 16 GB RAM, 250 GB Hard Disk Drive (HDD).

** This column captures disk space requirement for software (about 100 GB) and the requirements for shared storage (or install directory if there is no DPS scale out) for run-time data.

*** For large and enterprise deployments of OA/SiS collections, additional remote collectors should be deployed on separate systems.

Table 9: Remote Collector Configuration

Multiple Remote Collectors are recommended in the cases of Operations Agent (OA) collections for larger deployments. An OBR collector component instance is certified for data collection from a maximum of 10,000 Operations Agent nodes (~320K CIs). An additional Remote Collector of the same configuration as specified below should be added for every 10,000 managed nodes.

Deployment Size (Number of Nodes)	System Configuration (per remote collector)			Collector Configuration	
	CPU(64 bit) x-86-64	RAM (in GB)	Disk Space (in GB)	Threads	Max Heap Size(in GB)
10,000	4 CPU Cores	8	250	2500	6

In the case of SiteScope collection, you need to consider the following crucial factors when deciding the number of additional remote collectors:

- Number of instances monitored

- Polling frequency

Table 10: Remote Data Processor Configuration

System Configuration (per remote data processor)		
CPU(64 bit) (x-86-64)	RAM (in GB)	Disk Space (in GB) 4 CPU
4 CPU Cores	16	250

Note: The distributed scenario for Data Processor on Windows operating system is not supported.

Chapter 3: General Recommendations and Best Practices

This section provides the guidelines and best practices to achieve optimal performance of OBR.

Hardware and Software

Processor

You can deploy OBR on systems with Intel 64-bit (x86-64) or AMD 64-bit (AMD64) processors. It is recommended to use Intel processors.

For Intel 64-bit (x86-64), the latest Gen 9 processors are recommended.

Disk for OBR Server

Disk performance is important for high scale environments that are medium tier or higher. For OBR server, it is recommended to use RAID 5+0 (50) with battery-backed write cache on disks of 15,000 rpm or high performance SAN storage. Disk configurations with lower level of performance, are not adequate.

For information about disk space recommendations for HPE Vertica, see "Vertica Database" on page 23.

Operating System

Linux

The Linux kernel provides a system to limit the number of file descriptors and other resources on a per-process basis. OBR uses sockets and file-system files extensively. So the OBR service start scripts, sets the files limit to 65,536.

Windows

OBR establishes a connection to various data sources to collect monitoring data. When a connection is established, the client side of the connection uses a port number. The ephemeral port range configured

on a Windows system limits the maximum number of connections from one system to another. You should increase this range to approximately 60,000, by executing the steps mentioned in http://support.microsoft.com/kb/319502

You should configure the virtual memory to at least twice the size of the RAM.

Software

To see the list of supported operating systems, see the *Operations Bridge Reporter Release Notes*.

Operations Bridge Reporter Server

This section provides information on the best practices for tuning OBR application. Operations Bridge Reporter (OBR) implements an Extract, Transform, and Load (ETL) layer to collect, transform, and load data into its data warehouse. The collector component in OBR communicates with the data sources and extracts data. The data processor component transforms and enriches the data. The data warehouse is implemented in a Vertica column store database. Based on the size of your deployment, you can deploy the following components on separate systems:

- The reporting server (SAP BusinessObjects)
- Collector
- Data processor
- Vertica components

The deployment on separate systems enable you to distribute the load of the central server. You can also choose to deploy the collector, close to the data sources, to reduce network bandwidth usage.

Data Extraction

Initial Data Collection

OBR collectors have the ability to collect historical data when it starts collecting data from a data source. The default settings for various data sources are shown in Table 11: Initial history collection period.

Table 11: Initial history collection period

Table Type	Initial History Collection period
Agents	15 Days
BSM Profile DB & Network DB	15 Days
OMi (HIs)	7 Days
OMi (Events and KPIs)	15 Days

These default settings may be changed to get additional historical data. However, increase in duration affects the RAM usage and increases the time taken for completing the operation.

To collect additional historical data from the Operations Agent, increase the `collector.initHistory` parameter in the `config.prp` file that is present in the {PMDB_HOME}/data folder in your collector system. The number of Operations Agents polled for data concurrently is controlled by the number of threads configured in the OBR collection. The `org.quartz.threadPool.threadCount` parameter in the file {PMDB_HOME}/config/ramscheduler.properties identifies the maximum number of threads that may be generated; and thereby determines the maximum number of Operations Agents that may be polled simultaneously. If the requested historical data is huge, decrease the number of threads. This ensures that the memory capacity of OBR is not breached and the `OutOfMemory` error is not displayed. With 5,000 hosts and 15 days of initial history collection, the recommended thread count is 50 for initial history collection. For deployments with OA/SiS collections having over 10,000 agents, it is recommended to run remote collectors only and disable the 'local' collector.

A large volume of data is extracted from the Profile and Network databases. If more than 15 days of data is required, modify the `dbcollector.initHistory` parameter in the file {PMDB_HOME}/data/config.prp. If more historical data is required, set the thread count in the file {PMDB_HOME}/config/ramscheduler.properties to a very low value. This can slow down the Operations Agent collection. However, it will allow the Profile database data to be collected. This in turn may increase the heap memory consumption of OBR.

After the collection is complete, set the thread count to the default value.

Missing Data Collection

If OBR is down for some period for maintenance or other reasons or if data source is unreachable for some period, OBR collects the missing data from the data sources. If the collection stops for some reason, `collector.maxHistory` parameter defined in the file {PMDB_HOME}/data/config determines the maximum amount of historical data that may be collected by OBR from the Operations Agents. The default value is set to 15 days (360hrs). For profile and network database collection, `dbcollector.maxHistory` parameter determines the maximum amount of historical data that may be collected by OBR from the BSM Profile and Network databases. The default value is 15 days (360 hrs).

If OBR collects missing data for a number of data sources, you may reduce the `org.quartz.threadPool.threadCount` value as in the case of initial history collection.

Agent Response Timeout

If the agents in your environment do not respond after connection is established, you will see socket read connection timeout errors in logs. This slows down the data collection from other data sources. To overcome this issue, set the socket read timeout to a lower value for agent communication by executing the following commands:

```
ovconfchg -ns bbc.cb -set RESPONSE_TIMEOUT <timeout in secs>
```

```
ovc -restart
```

However, if you set this to a very low value, then the socket closes before the agent responds. This abrupt closure will result in data loss.

Collection Interval

OBR uses OMi/APM Run-time Service Model (RTSM), Operations Management (OM), or VMware vCenter as its topology source. The default collection interval for topology sources is set to 24 hours. This is the recommended minimum period. However, you can change this value through the OBR Administration Console. You should set this parameter value depending on the frequency with which the topology sources are updated. If RTSM or OM is updated at a lower frequency, you may increase the collection interval. Thereby, you can avoid expensive dimension updates of all Content Packs. Performance of OBR decreases if you reduce the collection interval.

The collection interval for data collection from Operations Agent, Profile Database, and Network database is set to one hour, by default. You can change this parameter from OBR Administration Console. Increasing the collection interval results in latency increase.

Data Retention Period for the Collector

The OBR server pulls (copies if collector coexists with server) data from collector and archives it in the `{PMDB_HOME}/extract/archive` folder on the collector system. You can configure the retention period for the archive folder using the `archivefilecleanup.job.freq` and `archive.retention.period` parameters in the `{PMDB_HOME}/config/collection.properties` file. The `archivefilecleanup.job.freq` parameter indicates the frequency of the cleanup job in minutes and `archive.retention.period` indicates the retention period in hours.

Data Processing

Data Processor

The Data Processing layer (ETL) in OBR extracts and transforms tthe data before it enters Vertica.

The Collector extracts data from the data sources and writes it to the collector file system as CSVs. The OBR server pulls the data from the collector periodically, writes it to the OBR server file system. The data processor reads the data, transforms and loads to Vertica.

> **Note:** The distributed scenario for Data Processor on Windows operating system is not supported.

In case of scale-out scenario, data transformations in ETL layer scale out when ETL layer scales out. You should have a shared file system that is accessible from all the systems on which the data processor is installed. The OBR server stores the data fetched from the collector on the shared file system. The data processor reads the CSVs from shared file system, processes and loads the data to Vertica. The ETL layer transformations are always triggered only from the OBR server system. The lab benchmarking was done with NFS as sharing mechanism on Linux OBR/DPS systems and CIFS as sharing mechanism for Windows OBR/DPS systems. Its recommended to have a SAN storage disk for the shared storage, and to ensure that a network speed of 10G is available between the machines in the deployment.

The data processor includes Orchestrator, Job Manager and Task Manager services. These services are present on the OBR server system. On the Data Processing system, only the Task Manager service is present. The status of Orchestrator and Job Manager can also be checked from the OBR Administration Console Dashboard.

The Orchestration service controls and monitors all the Vertica operations (Stage, Aggregate and custom SQL steps) and also monitors non-Vertica transformations (Mapper and Reconcile). OBR uses Apache Flink, an open source distributed processing framework to manage the data streams. When the Orchestration service starts, it triggers an Apache Flink job which submits the ETL steps (like Mapper and/or Reconcile tasks) to Job Manager. Job Manager in turn distributes these to different Task Managers for execution.

OBR uses Apache Flink, an open source distributed processing framework to manage the data streams. The Orchestrator creates data processing pipelines corresponding to the transformations defined in the Content pack. These pipelines are submitted as a single job in Apache Flink. The parallelism for these pipelines is configured in the Flink configuration file. The Apache Flink Job Manager distributes the tasks to different task managers in a sequential manner on the same system, where each task takes input from the previous task. This is to avoid data distribution across task managers.

Data Processing Servers (DPS) Parameters to be modified

Depending on the number of Data Processing Servers (DPS) used, some of the Apache Flink parameters have to be modified.

> **Tip:** In order to have a continuous and unaffected data flow if any of the DPS goes down, it is

> recommended to have one more additional DPS along with the recommended numbers of DPS with the parallelism same as before.

Make sure you modify these parameters on the OBR server and the Data Processor server(s).

The Apache Flink configuration file is in the following location:

On Linux: `$PMDB_HOME/../Flink/conf/flink-conf.yaml`

On Windows: `%PMDB_HOME%\..\Flink\conf\flink-conf.yaml`

Modify the following parameter in the `flink-conf.yaml` file:

`parallelism.default: 30 * number of task managers`

where, `number of task managers = number of additional DPS + 1`

`taskmanager.heap.mb: <allocated value>`

<allocated value>*: The task manager heap space by default is set at 2048 in OBR.

*This value has to be calculated to be the high enough value, so that memory requirements of other OBR components (Vertica, BO, Orchestration, custom scripts, etc) are met, and at the same time, the largest possible memory allocation is given to Apache Flink.

For example, `taskmanager.heap.mb: 3072`

Suggested formula to calculate the above parameter in different scenarios:

1. For Typical, all-in-one machine, 1/16 times the available RAM (in GB) * 1024

   [Example: for a 48 GB RAM machine, 1/16 * 48 * 1024 = 3 * 1024 = 6144]

2. For machines with OBR alone or DPS, 3/4 times the available RAM (in GB) * 1024

   [Example: for a 48 GB RAM machine, 3/4 * 48 * 1024 = 36 * 1024 = 36864]

3. For machines with BO and OBR, 3/8 times the available RAM (in GB) * 1024

   [Example: for a 48 GB RAM machine, 3/8 * 48 * 1024 = 18 *1024 = 18432]

Also, update the parameter `taskmanager.network.numberOfBuffers`. The default value for this parameter is 12000. Calculate the value using the following formula:

`numberOfBuffers = (number of task slots per DPS)^2 * (number of DPS) * 8`

For Example: If we have 3 DPS and each DPS has 30 task slots, then
`taskmanager.network.numberOfBuffers: 21600`

Number of OBR Processes

Content Packs installed in OBR deploy data processing streams as shown in Figure 1, to audit and control the data flow. These streams consist of steps that implement various ETL tasks and also control their sequence of execution. Each Content Pack deploys one or more streams in OBR. These streams are launched periodically and each step launches a process that runs the specified task. To keep the performance overhead of idle Content Packs low, it is recommended to install only those Content Packs that have data sources configured.

Figure 1: OBR Stream
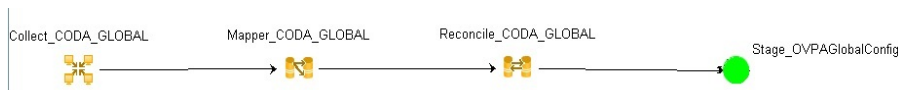
All data movement within OBR is controlled through a data processing framework. This framework allows you to control the total number of orchestration services that run at any given time. It is also possible to control the number of processes per step type. If the OBR system has limited resources or is consuming a lot of CPU resources, then you should provide a limit for the total number of OBR data processes and limit the per step type process to reduce resource utilization.This, however, can slow down the movement of data into OBR. Similarly if there is high latency with data movement then you can increase the limits of OBR processes depending on the hardware resources available to OBR.

To limit the number of OBR data processes, see the *Managing data processes* section in the *HPE OBR Online Help for Administrators* section . To limit the number of process per step type, execute the following command:

abcAdminUtil -setResourceCount -resourceType *<resource type>* -value *<pool count>*

where,

<resource type> is type of the step.

For example: STAGE_PROC, LOAD_PROC, AGGREGATE_PROC, and EXEC_PROC_PROC

<pool count> is the limit on the number of processes of type *<resource type>*.

Default values set for each of these steps are listed in the table below:

Step Type	Default Process Limit
STAGE_PROC	20
LOAD_PROC	30
AGGREGATE_PROC	20
EXEC_PROC_PROC	20

Each data movement step that is processed in OBR has a maximum time limit. By default, this limit is set to 60 minutes. In certain cases where a large amount of data is being processed, steps like pre-

aggregation and forecasting might exceed this limit. This causes the data processing stream to display an error state. In such cases, you must wait until the data processing is complete.

Disk space usage

Increase in number of files in OBR folders affects performance of disk operations. The OBR components move the files to the failed folder if it encounters errors while processing the data in the file. These files contain data rejected by OBR's ETL layer and may need to be corrected manually. Accumulation of files in these folders can increase disk space usage and may affect other disk operations. You are required to manually process the data in the `{PMDB_HOME}/stage/failed_to_transform, {PMDB_HOME}/stage/failed_to_stage, {PMDB_HOME}/stage/failed_to_load` and `{PMDB_HOME}/stage/failed_to_reconcile` folders.

After the data is loaded into the stream tables, the collected data is archived as CSV files in the folder `{PMDB_HOME}/stage/archive`. These files are deleted periodically by OBR. An increase in the number of files increases disk space usage and may affect other disk operations.

Increasing log file size from the default settings uses more disk space. Before increasing log file size, make sure you have adequate disk space.

For more information, see the *Operations Bridge Reporter Troubleshooting Guide*.

Vertica Database

Traditional OLTP databases store data row-wise, which is the preferred mechanism for transaction processing. Vertica stores data by column which is suitable for queries that extract few fields from a table. Vertica performance is generally limited by the CPU and Memory.

Vertica supports scale up and scale out as per your need. For scale up, more hardware resources such as CPU and RAM need to be added. For Scale out, Vertica clustering requires minimum of 3 nodes, and all the nodes must have the same hardware configuration.

Clustering lets you scale out your database cluster easily by adding more hardware.

Columns are duplicated across cluster nodes. If one machine goes down, you still have a copy:

- Data warehouse log-based recovery is impractical. Instead, stores enough projections for K-safety

New cluster node queries existing nodes for the data it needs:

- Rebuilds missing objects from other nodes, another benefit of multiple sort orders

The following hardware configurations provide optimal performance for your Vertica database.

Processor

See Vertica Hardware Planning Guide for information about processor requirements for Vertica.

Memory

HPE Vertica requires a minimum of 8 GB of memory per physical CPU core on each server. However, in high-performance applications, you should run 12-16 GB of memory per physical core. The lab benchmark tests done on Vertica recommend 1:4 ratio for CPU: Memory. In the unlikely event of any issue, Vertica recommendations for such cases would need to be honored for enterprise deployments. The memory should be at least DDR3-1600 (preferably DDR4-2133), and should be appropriately distributed across all memory channels on each server

Storage

Vertica requires a minimum read/write speed of 40 MB/s per physical core of the CPU. However, for the best performance, you should have 60–80 MB/s per physical core. Each node should have 1–9 TB of storage post RAID. In a production setting, Vertica recommends RAID 10. If RAID 10 is not available, you can use RAID 50 as a viable alternative.

Due to the heavy compression/encoding that Vertica performs, you do not need to use solid-state drives (SSDs). To satisfy Vertica requirements, a RAID array of more, cost effective hard disk drives (HDDs) works just as well as a RAID array of fewer SSDs.

HPE Vertica recommends that you use enterprise grade direct attached storage (DAS) instead of storage area network (SAN) or Network Attached Storage (NAS). Doing so usually results in faster data retrieval.

> **Note:** If you intend to use RAID 50 for your data partition, keep a spare node in every rack. This allows for manual failover of a Vertica node in the case of a drive failure. (Recovering a Vertica node is faster than rebuilding a RAID 50.To keep node recovery times at an acceptable rate, never put more than 10 TB compressed data on any node.)

Network

Vertica recommends 10G networking over 1G networking in almost every situation.

Configuring and Monitoring Virtual Machines

Vertica recommends that the virtual servers reach at least these performance goals:

Networking

- 100 MB/s of UDP network traffic per node on the private network (as measured by vnetperf)

- 20 MB/s per core of TCP network traffic on the private network (as measured by vnetperf)

- Independent public network

I/O:

- Measured by vioperf concurrently on all Vertica nodes:

- 25 MB/s per core of write

- 20+20 MB/s per core of rewrite

- 40 MB/s per core of read

- 150 seeks per second of latency (SkipRead)

- Thick provisioned disk, or pass-through storage

> **Note:** Vertica does not support VMware Vmotion and Logical Volume Manager (LVM) on any drive where database (catalog and data) files are stored.

For best performance

- Disable CPU scaling on the physical hosts.

- Configure the disk blocks to align with the blocks that ESX creates. Unaligned blocks may cause reduced I/O performance during high load.

All virtual machines in a virtualized Vertica cluster must be configured with the same specifications. Vertica recommends that you configure your virtual machine as per the following instructions:

- One socket per virtual machine and 4 GB of memory per core in that socket

- Configure all volumes attached to each virtual machine as:

    - Thick Provisioned Eager Zeroed

    - Independent

    - Persistent

SAP BusinessObjects

OBR reports are web intelligence documents. The Web Intelligence (WebI) Report Server in SAP Business Objects is responsible for generating Web Intelligence documents. The maximum memory available to the WebI server is only 6 GB. For further sizing of reports based on the number of users, you can distribute the Adaptive Processing Servers (APS). The confirmation options are available under OBR Central Management Console (CMC) home under **System Configuration Wizard**.

The load on the APS server depends on the number of simultaneous connections to the server, and the complexity and size of the report documents accessed. If the server is not configured correctly, while accessing reports you may get errors like "Web Intelligence Server is busy" and "Server reached maximum number of simultaneous connections."

Some of the steps that can help you avoid these errors are:

- While accessing reports, default values for prompts can result in several thousands of records being loaded into the Web Intelligence server. You should specify appropriate values for prompts to avoid high load on the server. For example, when opening reports, you should specify values for business service or node group prompts, such that the number of nodes fetched from database is not more than 1000 to 2000.

- The default value for the maximum number of connections is 200. If the load on the server per request is high, then you may see the Server Busy error message. Consider reducing this parameter and adding more Web Intelligence Server to support additional connection requests. While adding additional servers, always add one Web Intelligence Processing Server per CPU core per machine.

To summarize, the objective is to arrive at the number of WebI servers and maximum connections per server so that the server allows all users to connect and open report documents without reaching the 4 GB limit.
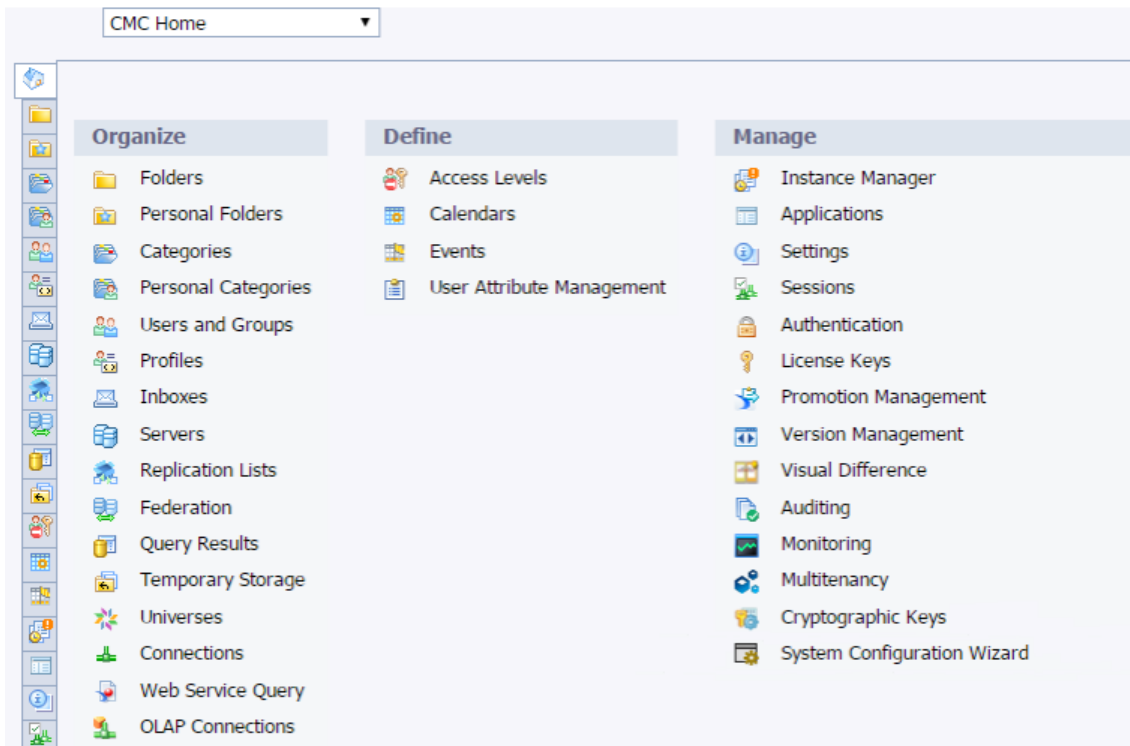
Adding Additional Web Intelligence (WebI) Servers

Perform the following steps to add additional WebI Servers:

1. Log in to Central Management Console (`https://<OBRSystem_FQDN>:8443/BOE/CMC`), where, *<OBRSystem_FQDN>* is the fully qualified domain name of the system where SAP BusinessObjects is installed.
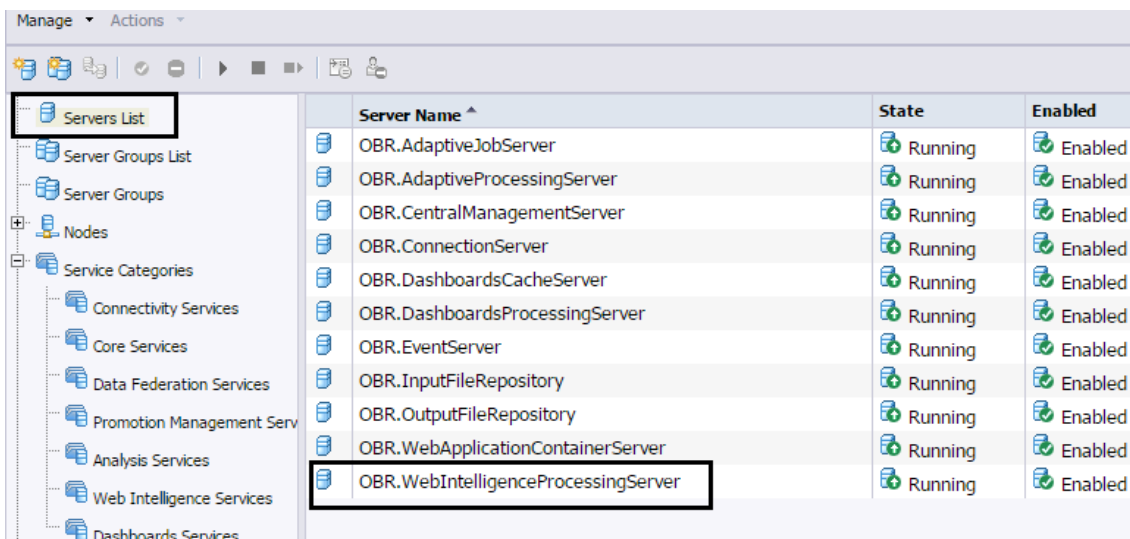
   The **Central Management Console** home page is displayed.
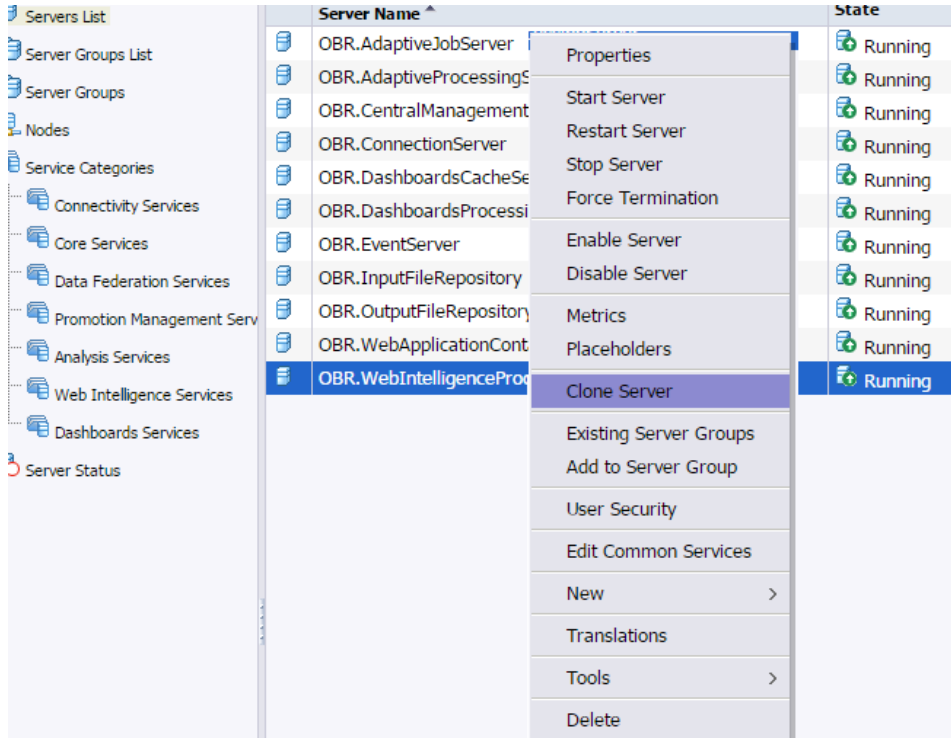
2. Click on **Servers** and then click **Servers List**.

   The list of servers are displayed. You will find the **OBR.WebIntelligenceProcessingServer**.

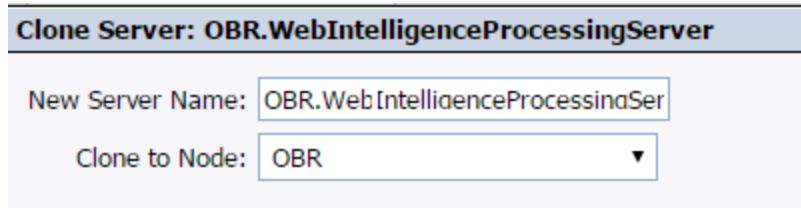3. Right-click on **OBR.WebIntelligenceProcessingServer** and then click **Clone Server**.

The **Clone Server** dialog box is displayed.

4. Enter the **New Server Name** as **OBR.WebIntelligenceProcessingServer1**.

The Clone to Node is selected by default as **OBR**.

An additional WebI Server is created to balance the load when the reports with data more than 4 GB is running simultaneously.
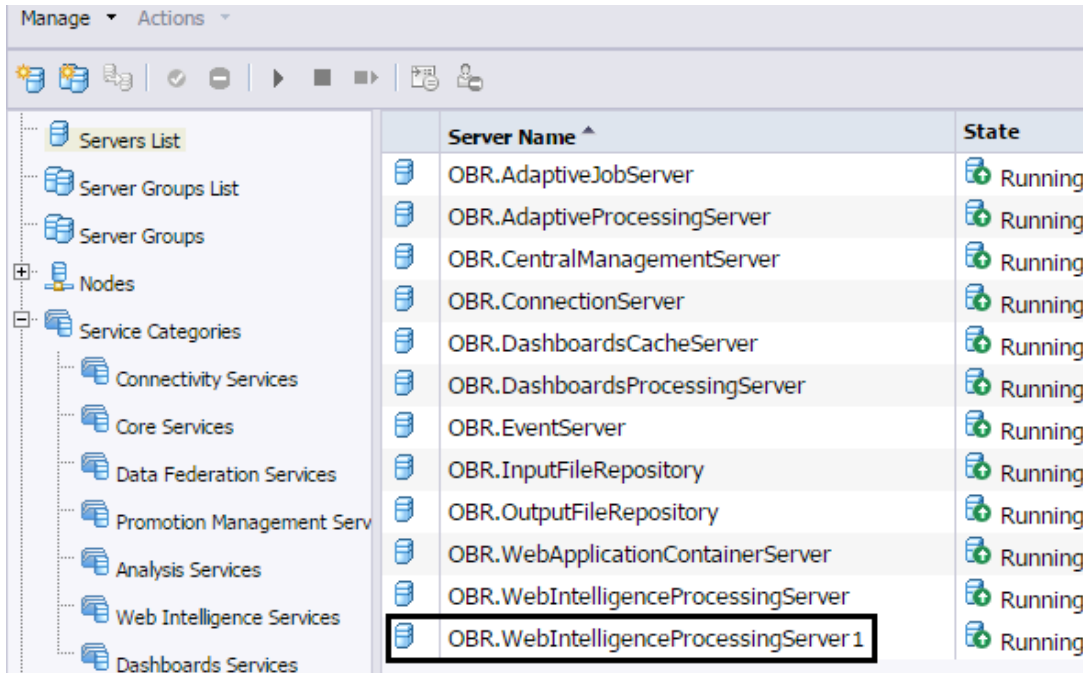
5. Click **OK**.

6. To enable the server, right-click on new **OBR.WebintelligenceProcessingServer1** and click **Enable Server**.

7. To start the server, right-click on **OBR.WebintelligenceProcessing server1** and click **Start Server**.

Chapter 4: Sizing Calculator

The key objective of the OBR Licensing and Sizing Calculator is to provide the following estimates:

- Node licenses required for OBR deployment based on the input parameters

- Sizing of hardware to deploy OBR.

Accessing the Licensing and Sizing Calculator

The Licensing and Sizing Calculator is available as an excel attached with this PDF. Follow these steps to access the calculator:

1. From the Bookmarks pane of this PDF, click icon.

2. The *OBR10.22_Sizing_License_Calc* attachment is listed. Double-click to open the excel sheet. The excel sheet appears.

3. Type the values to calculate the licensing and sizing requirements.

Chapter 5: Benchmark

This chapter describes the benchmark test scenarios and the test methodology used for performance tests.

> **Note:** The benchmark tests were performed with HPE OBR database as the only database on a Vertica instance.

Benchmark Scenario 1

OBR is installed with all Content Packs on different systems. This test was carried out using System content with a load of 10,000 hosts and SM, DB SPI, OMi, App SPI, Service Health content with medium load (6 million throughput), on an All-In-One system.

For Typical deployment, benchmark test was performed on an all-in-one linux system

Windows

For Windows operating system, benchmark test was performed with OBR server on the Windows and Vertica on Linux.

Deployment (with Modules installed)	Hardware Configuration - Typical
All-In-One (Typical Deployment - OBR Server) **Configuration Parameters (edited on OBR 10.21) flink-conf.yaml:** `taskmanager.heap.mb: 12288`	OBR, DPS, BO & Vertica installed on same system
	Model: HP ProLiant DL380 G9
	CPU: 16 (Intel Xeon CPU E5-2630 v3 @ 2.40 GHz)
	RAM: 48 GB & Virtual Memory: 32 GB
	HDD size (type RAID 1) : 1.2 TB Storage Type: P6000 EVA Storage Systems Drive Type: SAS Rotational Speed: 15K RPM Transfer Speed PHY 1:3 GBPS Disk Cache Battery: 1 GB

Deployment (with Modules installed)	Hardware Configuration - Typical
	OS: RHEL 6.5

Benchmark Scenario 2

OBR and Vertica installed on different systems, with Vertica as a remote single node cluster. This test was carried out using System content with a load of 20,000 hosts and all content with medium load (12 million throughput).

For Medium deployment system, benchmark test was performed with OBR & BO server on 1 Linux host, an additional DPS server on another Linux host and Vertica on a 3rd Linux host.

Deployment (with Modules installed)	Hardware Configuration - Medium (Distributed)
OBR (Deployment module - OBR and BO Server)	OBR and BO installed on same system
Configuration Parameters (edited on OBR 10.21) flink-conf.yaml:	Model: HP ProLiant DL380 G9
taskmanager.heap.mb: 12288	CPU: 16 (Intel Xeon CPU E5-2630 v3 @ 2.40 GHz)
taskmanager.network.numberOfBuffers: 12000	RAM: 32 GB & Virtual Memory: 16 GB
parallelism.default: 60	OS: RHEL 6.5

Deployment (with Modules installed)	Hardware Configuration - Medium (Distributed)
DPS (Deployment module - DPS Server)	Only DPS is installed on this system
Configuration Parameters (edited on OBR 10.21) flink-conf.yaml:	Model: HP ProLiant DL380 G9
taskmanager.heap.mb: 12288	CPU: 4 (Intel Xeon CPU E5-2630 v3 @ 2.40 GHz)
taskmanager.network.numberOfBuffers: 12000	RAM: 16 GB & Virtual Memory: 8 GB
parallelism.default: 60	OS: RHEL 6.8

Deployment (with Modules installed)	Hardware Configuration - Medium (Distributed)
Shared Disk (to access run-time and configuration data)	OBR and BO installed on same system
	Model: HP ProLiant DL380 G9
	Share Type: NFS
	HDD size (type RAID 1) : 650 GB
	Storage Type: HP Flexible Smart Array P440ar
	Drive Type: SAS
	Rotational Speed: 15K RPM
	Transfer Speed PHY 1:3 GBPS
	Disk Cache Battery: 1 GB
	OS: RHEL 6.5
	Network speed: 1G

Deployment (with Modules installed)	Hardware Configuration - Medium (Distributed)
OBR (Deployment - Remote DB) – Vertica single node	OBR and Vertica installed on different systems
	Model: HP ProLiant DL380 G9
	CPU: 16 (Intel Xeon CPU E5-2630 @ 2.40 GHz)
	RAM: 64 GB & Virtual Memory: 8 GB
	HDD size (type RAID 1) : 850 GB
	Storage Type: HP Flexible Smart Array P440ar
	Drive Type: SAS
	Rotational Speed: 15K RPM
	Transfer Speed PHY 1:3 GBPS
	Disk Cache Battery: 1 GB
	OS: RHEL 7.2

Benchmark Scenario 3

OBR and Vertica are installed with SM Content Pack on the same system. It is then deployed in an OM environment. This test was carried out using System content with a load of 40,000 hosts and OMi content with enterprise load (19 million throughput). Benchmark test was performed on the Linux operating system (for OBR server, BO server, 4 DPS servers and 3 node Vertica cluster).

Deployment (with Modules installed)	Hardware Configuration - Enterprise (Distributed)
OBR (Deployment module - OBR Server)	OBR only istalled on the system

Configuration Parameters (edited on OBR 10.21) flink-conf.yaml:
 	
`taskmanager.heap.mb: 24576`
 	
`taskmanager.network.numberOfBuffers: 30000`
 	
`parallelism.default: 150`
	Model: HP ProLiant DL580 G8

	CPU: 16 (Intel Xeon CPU E7-4830 v2 @ 2.20 GHz)
	RAM: 32 GB & Virtual Memory: 24 GB
	OS: RHEL 7.2

Deployment (with Modules installed)	Hardware Configuration - Enterprise (Distributed)
DPS (Deployment module - DPS Server) (specs mentioned are per DPS node)	Only DPS is installed on this system

Configuration Parameters (edited on OBR 10.21) flink-conf.yaml:
 	
`taskmanager.heap.mb: 12288`
 	
`taskmanager.network.numberOfBuffers: 30000`
 	
`parallelism.default: 150`
	Model: HP ProLiant DL380p G8

	CPU: 4 (Intel Xeon CPU E5-26700 @ 2.60 GHz)
	RAM: 16 GB & Virtual Memory: 8 GB
	OS: RHEL 6.8

Deployment (with Modules installed)	Hardware Configuration - Enterprise (Distributed)
Shared Disk (to access run-time and configuration data)	OBR and BO installed on same system
	Model: HP ProLiant DL580 G8
	Share Type: NFS
	HDD size (type RAID 1) : 2 TB
	Storage Type: HP Flexible Smart Array P440ar
	Drive Type: SAS
	Rotational Speed: 15K RPM
	Transfer Speed PHY 1:3 GBPS
	Disk Cache Battery: 1 GB
	OS: RHEL 7.2
	Network Speed: 1G

Deployment (with Modules installed)	Hardware Configuration - Enterprise (Distributed)
OBR (Deployment module - BO Server)	OBR and BO installed on same system
	Model: HP ProLiant DL580 G8
	CPU: 16 (Intel Xeon CPU E7-4830 v2 @ 2.20 GHz)
	RAM: 32 GB & Virtual Memory: 24 GB
	OS: RHEL 7.2

Deployment (with Modules installed)	Hardware Configuration - Enterprise (Distributed)
OBR (Deployment - Remote DB) – Vertica 3 node cluster (specs mentioned, are per DB node)	All 3 nodes have identical configurations
	Model: HP ProLiant DL380p G8
	CPU: 24 (Intel Xeon CPU E5-26700 @ 2.60 GHz)
	RAM: 96 GB & Virtual Memory: 10 GB
	HDD size(preferably of type RAID 5) : 2 TB Storage Type: P6000 EVA Storage Systems Drive Type: SAS Rotational Speed: 15K RPM Transfer Speed PHY 1:3 GBPS Disk Cache Battery: 1 GB OS: RHEL 6.8 Network Speed: 1G

Send documentation feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on Performance, Sizing, and Tuning Guide (Operations Bridge Reporter 10.22)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to docfeedback@hpe.com.

We appreciate your feedback!