

F5® BIG-IQ® Centralized Management: Data Collection Device Sizing Guide

Version 5.3.0



Table of Contents

Table of Contents.....	2
Overview.....	3
What resources do I need if I'm just collecting statistics?	5
What resources do I need to collect ASM event data?	6
What resources do I need to collect APM event data?	7
Using the APM data script	7
What resources do I need to collect FPS alert data?.....	9
What resources do I need if I collect data for multiple modules?	11
Revision History.....	12
Legal notices.....	13
Trademarks.....	13
Usage disclaimer.....	13

Overview

Use this guide to determine the resources that are required to handle the data generated by the BIG-IP® devices you manage. Requirements vary according to the type and amount of data you generate. Refer to the table(s) that correspond to the type of data that you want to collect from your BIG-IP devices.

Each table includes a Data Redundancy column to help you consider the additional data collection devices (DCDs) necessary to make sure that your data exists on more than one device.

In addition to the alert rate, consider the number of devices you plan to manage, the data retention policy you plan to use and the average size of the alerts you plan to store on the DCD. When you configure the retention and rotation settings for your alert data, you must consider the available storage on the DCD. Although DCD VE devices allow for expansion of the alert partition, be aware that if you exceed 50% of the available disk space (which is not the same as raw disk space), you will not be able to upgrade your BIG-IQ DCD cluster. After deployment, a certain amount of the disk space is consumed by the BIG-IQ system software and is unavailable.

If you are planning to upgrade a deployed DCD, you can use the following steps to determine the amount of available disk space:

1. Use SSH to log in to your DCD as root.
2. Check the disk space available by running the `vgdisplay` command:

For example:

```
vgdisplay
--- Volume group ---
VG Name                vg-db-sda
System ID
Format                 lvm2
Metadata Areas         1
Metadata Sequence No   179
VG Access               read/write
VG Status               resizable
MAX LV                 0
Cur LV                 16
Open LV                 7
Max PV                 0
Cur PV                 1
Act PV                  1
VG Size                498.50 GiB
PE Size                 4.00 MiB
Total PE                127617
Alloc PE / Size         27936 / 109.12 GiB
Free PE / Size          99681 / 389.38 GiB
```

In this example, the total disk space (VG Size) is 498.50 GB

Overview

3. Identify how the disk space is allocated across your partitions by running the `lvs` command:

For example:

```
[root@vBIGIQ-LN1:Active:Standalone] config # lvs
LV          VG          Attr      LSize  ...
dat.log.1   vg-db-vda   -wi-ao--- 7.00g
dat.maint.1 vg-db-vda   -wi-a---- 300.00m
dat.share.1 vg-db-vda   -wi-ao--- 28.00g
dat.swapvol.1 vg-db-vda   -wi-ao--- 1.00g
set.1._config vg-db-vda   -wi-a---- 3.17g
set.1._usr   vg-db-vda   -wi-a---- 3.29g
set.1._var   vg-db-vda   -wi-a---- 10.00g
set.1.root   vg-db-vda   -wi-a---- 440.00m
set.2._config vg-db-vda   -wi-ao--- 3.17g
set.2._usr   vg-db-vda   -wi-ao--- 3.29g
set.2._var   vg-db-vda   -wi-ao--- 10.00g
set.2.root   vg-db-vda   -wi-ao--- 440.00m
```

There are four entries to note in this response:

`dat.log.1` (highlighted in yellow) displays the size of the `/var/log` partition that is used to store the collected log data.

`dat.share.1` (highlighted in aqua) displays the size of the `/shared` partition that is used for ISOs and backups on the BIG-IQ system.

`set.1._var` (highlighted in green) displays the size of the `/var` partition for HD1.1 that is used for the initial software installation.

`set.2._var` (also highlighted in green) displays the size of the `/var` partition for HD1.2 that is used for the upgrade installation. Note that these two partitions on the active and upgrade volumes must be the same size (as in this example).

4. If you determine you need additional disk space, increase the disk space (for both volumes) to the size you need. Refer to: *K16103: Extending disk space on BIG-IQ Virtual Edition a:t support.f5.com/kb/en-us/solutions/public/17000/400/sol17406.html* for details.
 - [What resources do I need, if I'm just collecting statistics?](#)
 - [What resources do I need to collect ASM event data?](#)
 - [What resources do I need to collect APM event data?](#)
 - [What resources do I need to collect FPS alert data?](#)
 - [What resources do I need if I collect data for multiple modules?](#)

What resources do I need if I'm just collecting statistics?

We've created a spreadsheet to help you determine how many DCDs you need (and how much storage you need on each one) to handle the data generated by the BIG-IP devices you manage. You can download these tools from downloads.f5.com.

The spreadsheet uses metrics that you provide to calculate your resource requirements.

To use the spreadsheet, you must first download it.

1. Log in to downloads.f5.com.
2. Navigate to BIG-IQ Centralized Management, v5.3.0, and click **I Accept** to accept the EULA.
3. Click **F5_Networks_BIG_IQ_DCD_Sizing_Tools.zip**.
After a brief pause, the Download Locations screen opens.
4. Click a download location to start the download.
5. When the download completes, unzip the file. Two files will extract:
`F5_Networks_BIG_IQ_DCD_Sizing_Tool.xlsx` and
`F5_Networks_BIG_IQ_Apm_Dcd_Sizing_Metrics.sh`.
6. Open the spreadsheet to the **Statistics DCD Sizing Tool** tab, and type values that describe the environment, policies, and modules for which you want to collect statistics.
7. After you fill in all of the fields, the recommended number and configuration of DCDs is calculated and displayed under **Sizing Recommendation for DCD Nodes**. You can see the impact of your storage decisions by noting how the values under **Time Layer Calculations** change when you enter different values.

The calculations performed for these elements are based on some assumptions. Bear these assumptions in mind as you complete the spreadsheet.

- The amount of device data increases marginally as you increase the number of CPUs and interfaces on your BIG-IP devices.
- The amount of pool and pool member data is based on an assumed average number of pool members per pool. For pools that have a larger than average number of pools, the amount of data increases as well.
- The spreadsheet assumes you have an average of 5 events per iRule. If your average number of events per iRule is lower, storage requirements go down; if your average is greater, storage requirements will increase.

What resources do I need to collect ASM event data?

What resources do I need to collect ASM event data?

Use this table to estimate the number of DCDs necessary to collect and display ASM event data from BIG-IP devices.

- If you are using an off-box logging solution to log ASM events, review a week's worth of alert logs to identify the logging load you need to support. Take note of both the average and maximum alerts per second.
- If you are logging ASM events locally on the BIG-IP device, you can also use those logs to identify your peak event logging rates. Compare the peak rates against the alerts per second (Alerts/Sec column) to estimate the number of DCDs you need.
- If you have not yet deployed an ASM solution in your environment, but you have an estimate of your application's peak HTTP/HTTPS traffic requirements, you can estimate your logging needs by projecting that 10% of your peak traffic could generate logging events. You can then use that value to estimate the number of DCDs you need.

BIG-IP device ASM sizing detail

Alerts/Sec	Data Redundancy	Recommended DCD count
10000	No	1
14000	Yes	2
17000	Yes	3
20000	Yes	4
23000	Yes	5
26000	Yes	6

Note:

- The sizing recommendations in this table assume that you provision your DCDs with 8 cores/CPU's and 32 GB of memory.
- When you choose an alert rate, consider both the average and peak loads that you anticipate. If your peak loads are frequent, consider provisioning additional data-collection nodes for the increased load. Also, when you determine your peak rates, plan for growth by considering the rate at which you expect your traffic load to increase.

What resources do I need to collect APM event data?

For APM®, we provide a spreadsheet and a script to help you determine the number of DCDs required to handle the data generated by the BIG-IP® devices you manage. You can download these tools from downloads.f5.com.

The script reads the last 10 days of logging data and provides the metrics you need for the spreadsheet. The spreadsheet uses those metrics to calculate your resource requirements. You enter information about the amount of data you expect to generate each day, and the spreadsheet calculates the number of DCDs required to handle that data.

Using the APM data script

F5 provides a script that you can run to help you determine how much APM data is generated. The APM data script collects metrics for the data your APM devices generate and stores the data in a file on your BIG-IP device.

***Note:** You can use the script only if you have enabled local db-logging. For information about enabling local-db logging, refer to [Overview: Configuring remote high-speed APM and SWG event logging on support.f5.com](#). If you use another logging mechanism, you can run reports (such as log reports or sessions reports) to get this data.*

***Note:** Unless there has been a recent HA failure event (within the last 3 days), run this script on the active device. In this case, the active device may not have enough data and it is better to run the script on the standby device.*

Even with the script, it might be tedious to collect data from each of your BIG-IP devices, so we recommend that you take the values from one of the BIG-IP devices with the most users and access profiles to account for the worst-case scenario.

To run the script, you must first download it.

8. Log in to downloads.f5.com.
9. Navigate to BIG-IQ Centralized Management, v5.3.0, and click **I Accept** to accept the EULA.
10. Click **F5_Networks_BIG_IQ_DCD_Sizing_Tools.zip**.
After a brief pause, the Download Locations screen opens.
11. Click a download location to start the download.
12. When the download completes, unzip the file. Two files will extract:
F5_Networks_BIG_IQ_DCD_Sizing_Tool.xlsx and
F5_Networks_BIG_IQ_Apm_Dcd_Sizing_Metrics.sh.
13. Copy the script (.sh) file to the
BIG-IP device on which you want it to run.
14. In the folder where you copied the script file, type:
Run F5_Networks_BIG_IQ_Apm_Dcd_Sizing_Metrics.sh.

What resources do I need to collect APM event data?

The script runs and collects the data. The output is stored in a file named `BIG_IQ_Apm_Dcd_Sizing_Metrics.txt`

15. After the script has run, navigate to the `/var/tmp` folder on the BIG-IP device and open the output file to see the collected metrics.

A sample output file is provided below:

```
NUMBER_OF_EVENTLOGS_PER_DAY
400,000
-----
NUMBER_OF_SESSIONS_PER_DAY
20,000
-----
LOGS_PER_SESSION
20
-----
NUMBER_OF_ACCESSSS_PROFILES
10
-----
NUMBER_OF_ACLs
50
```

16. Open the spreadsheet to the **Access DCD Sizing Tool** tab and type the values from the script output file into the appropriate fields on the spreadsheet.
For example, for the sample output above, you would type 20,000 in the **Number of sessions per day** field.
17. After you fill in all of the required fields, the recommended number of DCDs is calculated and displayed in the **Recommended number of DCD nodes** field.

What resources do I need to collect FPS alert data?

Use this table to estimate the number of DCDs required to collect and display FPS alert data from BIG-IP devices.

If you are using an off-box logging solution to log FPS events, review a week's worth of alert logs to identify the logging load you need to support. Take note of both the average and maximum alerts per second.

To determine the peak alerts/second rate to use in the table, first estimate the peak user load for the protected application over your highest period of production hours, and then plug that value into the following calculation:

$$\left(\text{< peak user load >} \frac{\text{users}}{\text{peak duration (h)} \times 60\text{m} \times 60\text{s}} \right) \approx \text{number of logins/sec} \approx \text{number of alerts/sec}$$

For example, if you estimate that there are, at most, 1 million users logging in to access your FPS environment over a 4-hour timespan, the calculation results in a login rate of:

$$\left(1,000,000 \frac{\text{users}}{4\text{h} \times 60\text{m} \times 60\text{s}} \right) = \approx 70 \text{ logins/sec} \approx 70 \text{ alerts/sec}$$

Although each user login does not generate an alert, even single user logins or application transactions can generate multiple alerts depending on the protection policy's configuration. For determining the number of DCDs needed to operate during peak conditions, 70 alerts per second is a reasonable assumption for this login rate.

BIG-IP device FPS sizing detail

Alerts/Sec	Data Redundancy	Recommended data collection device count
100	No	1
200	Yes	2
300	Yes	3
400	Yes	4
500	Yes	5
600	Yes	6

Note:

- The sizing recommendations in this table assume that you provision your DCDs with 8 cores/CPU's and 32 GB of memory.
- When you choose an alert rate, consider both the average and peak loads that you anticipate. If your peak loads are frequent, consider provisioning additional data-collection nodes for the increased load. Also, when you determine your peak rates, plan for growth by considering the rate at which you expect your traffic load to increase.

What resources do I need to collect FPS alert data?

- The outgoing data rate from your BIG-IP devices also depends on the FPS policy configuration, the number of simultaneous application users, and their expected transaction rate.
- The scale numbers provided assume the use of three forwarding rules (100% syslog forwarding, 10% SOC forwarding, 10% customer forwarding), and 17000 transform rules.

Note: Increasing the number of forwarding targets or the distance between your data collection devices increases the performance requirements for each node in the DCD cluster.

What resources do I need if I collect data for multiple modules?

Use this process to estimate the number of DCDs that you need to collect and display more than one kind of data from your BIG-IP devices.

1. Perform the estimate for each module independently.
2. Record the maximum number of DCDs recommended by any one estimate.
3. Sum the total storage needed by all estimates.
4. Divide the result in step 3 by the result of step 2.

For example:

You've performed the estimate for an ASM module, and the result is 2 DCDs and 60GB of storage.

You've performed the estimate for Statistics, and the result is 3 DCDs and 120GB of storage.

In this example, the maximum number of DCDs is 3 and the sum of the total storage is 180GB. Consequently, the recommendation is 3 DCDs (the maximum, from the statistics estimate) and 60GB storage (The sum of both, divided by the recommended number of DCDs).

Revision History

Revision History

Date	Change log
04/26/2017	Release Version
06/01/2017	Post release update

Legal notices

Copyright © 2017 F5 Networks, Inc. All rights reserved.

F5 Networks, Inc. (F5) believes the information it furnishes to be accurate and reliable. However, F5 assumes no responsibility for the use of this information, nor any infringement of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent, copyright, or other intellectual property right of F5 except as specifically described by applicable user licenses. F5 reserves the right to change specifications at any time without notice.

Trademarks

For a current list of F5 trademarks and service marks, see:
<http://www.f5.com/about/guidelines-policies/trademarks>.

All other product and company names herein may be trademarks of their respective owners.

Usage disclaimer

THE CONFIGURATION GUIDE IS PROVIDED "AS IS" AND POSSIBLY WITH FAULTS. F5 DISCLAIMS ANY AND ALL WARRANTIES AND GUARANTEES, EXPRESS, IMPLIED OR OTHERWISE, INCLUDING BUT NOT LIMITED TO THE WARRANTY OF MERCHANTABILITY, THE WARRANTY OF FITNESS FOR A PARTICULAR PURPOSE, AND ANY WARRANTY OF NON-INFRINGEMENT, WITH RESPECT TO THE CONFIGURATION GUIDE, YOUR USE THEREOF, OR ANY RESULTS. CUSTOMER WILL USE THE CONFIGURATION GUIDE AT ITS OWN RISK. F5 WILL NOT BE LIABLE TO CUSTOMER FOR ANY DIRECT OR INDIRECT DAMAGES INCURRED IN USING THE PRODUCT AND IN NO EVENT WILL F5 BE LIABLE FOR LOSS OF PROFITS, LOSS OF USE, LOSS OF DATA, BUSINESS INTERRUPTION, NOR FOR PUNITIVE, INCIDENTAL, CONSEQUENTIAL, OR SPECIAL DAMAGES OF ANY KIND, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.