

# **EMC Sizing Symmetrix Platform Data Collection Procedures**

**Version 2.06**

Copyright © 2012 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

EMC2, EMC, EMC ControlCenter, AlphaStor, ApplicationXtender, Captiva, Catalog Solution, Celerra, CentraStar, CLARAlert, CLARiiON, ClientPak, Connectrix, Co-StandbyServer, Dantz, Direct Matrix Architecture, DiskXtender, DiskXtender 2000, Documentum, EmailXaminer, EmailXtender, EmailXtract, eRoom, FLARE, HighRoad, Input Accel, Navisphere, NetWorker, nLayers, OpenScale, Powerlink, PowerPath, Rainfinity, RepliStor, ResourcePak, Retrospect, Smarts, SnapShotServer, SnapView/IP, SRDF, Symmetrix, TimeFinder, VisualSAN, VSAM-Assist, WebXtender, where information lives, Xtender, and Xtender Solutions are registered trademarks and EMC Developers Program, EMC OnCourse, EMC Proven, EMC Snap, EMC Storage Administrator, Acartus, Access Logix, ArchiveXtender, Authentic Problems, Automated Resource Manager, AutoStart, AutoSwap, AVALONidm, C-Clip, Celerra Replicator, Centera, CLARevent, Codebook Correlation Technology, Common Information Model, CopyCross, CopyPoint, DatabaseXtender, Direct Matrix, EDM, E-Lab, Enginuity, FarPoint, Global File Virtualization, Graphic Visualization, InfoMover, Infoscape, Invista, Max Retriever, MediaStor, MirrorView, NetWin, OnAlert, PowerSnap, RepliCare, SafeLine, SAN Advisor, SAN Copy, SAN Manager, SDMS, SnapImage, SnapSure, SnapView, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, UltraPoint, UltraScale, Viewlets, and VisualSRM are trademarks of EMC Corporation.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on [EMC.com](http://EMC.com)

All other trademarks used herein are the property of their respective owners.

# Table of Contents

## Contents

- Introduction to Data Collection ..... 4
  - Recommended 3-Day Minimum at 10 Minute Intervals ..... 4
  - Minimum Data Collection Interval..... 4
- Step 1 – Collect Performance Information ..... 5
  - Collection Considerations ..... 5
  - Data Collection Procedure ..... 5
    - Option 1 - Collecting Performance Information with Symmetrix Performance Analyzer ..... 5
    - Option 2 – Collecting Performance Information with Solutions Enabler ..... 8
- Step 2 – Collect Configuration Information ..... 10
  - Collection Considerations ..... 10
  - Solutions Enabler ..... 10
- Step 3 - Copy the Files to EMC Anonymous FTP Server ..... 11

# Introduction to Data Collection

EMC's modeling tools require host or array based statistical data to determine data change activity. This enables modeling based on real world data rather than a best guess, usually based on highly averaged data.

## Recommended 3-Day Minimum at 10 Minute Intervals

The minimum data capture interval should be 3 consecutive days, while a week is quite common (this gives greater certainty). EMC requires collecting data for 3-7 peak days. The actual times for the peaks should be the responsibility of the customer as they know when their peak processing occurs. Ideally, the data capture will include the peak workload period such as month-end or year-end processing. By capturing at least 24 hours, a distinct online and batch processing profile is observed. This is important because the average IO size of write I/Os tends to be smaller for online workloads and larger for batch, which has a direct impact on bandwidth requirements.

Data capture should be in 10 minute intervals or less, but no less than 3 minutes. If longer intervals are used you may be impacted by the flattening out of the true peaks due to longer averages, thus skewing the results. Even at 10 minute intervals your true peaks could be twice as much as the average peak recorded in the interval. This can cause a significantly under-configured solution.

## Minimum Data Collection Interval

The minimum data collection interval is 1 minute. Using a data collection interval of less than 1 minute will often result in corrupted data and invalid modeling results.

# Step 1 – Collect Performance Information

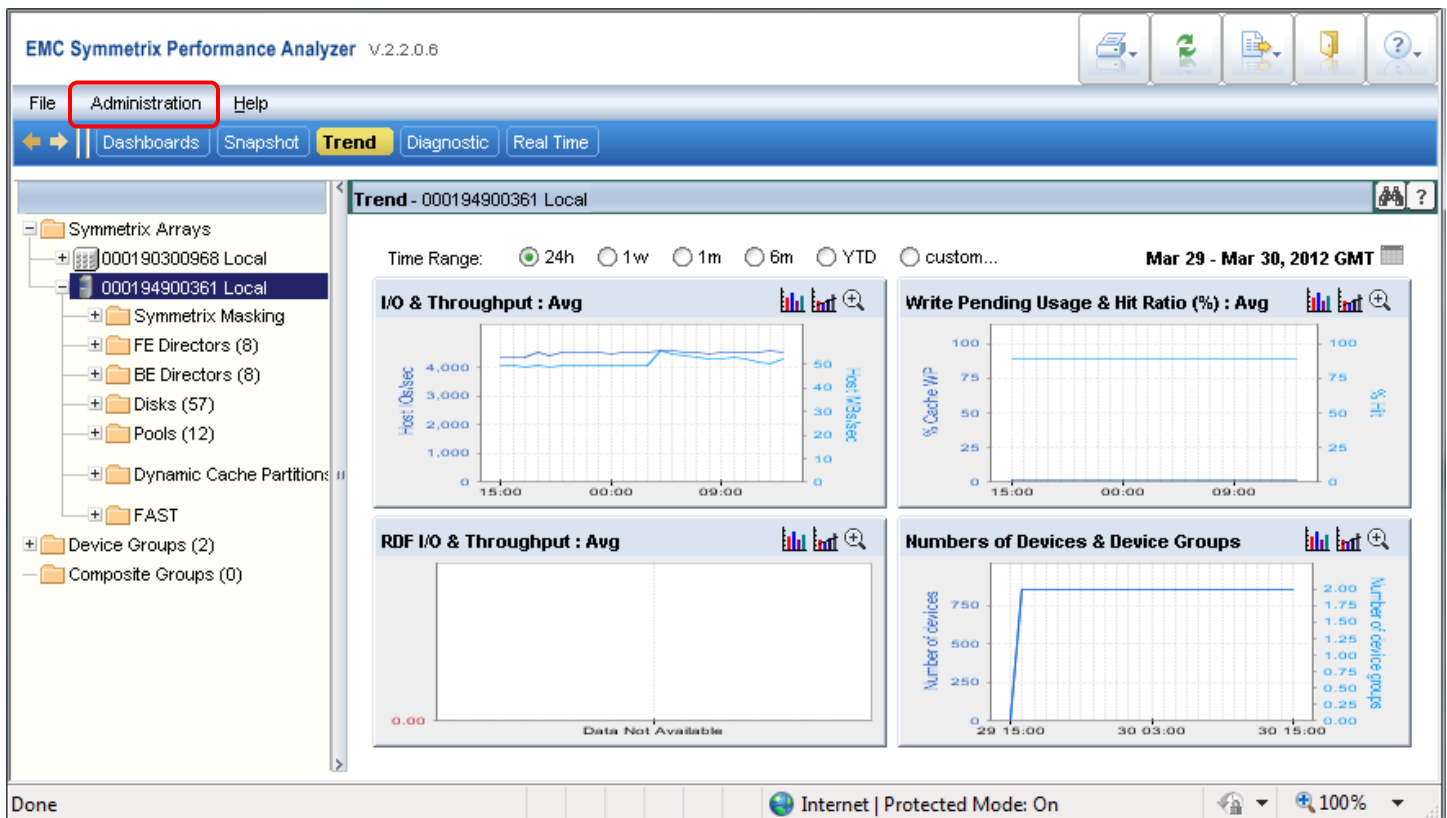
## Collection Considerations

If the workload to be modeled is a subset of the complete array, then please supply a list of the volumes that should be included in the sizing analysis. Access to the Symmetrix BIN file or symapi\_db.bin file is needed to accurately identify the meta-members. Use the Symmetrix Meta Member Expansion & Validation tool to expand your meta-head only include or exclude list.

## Data Collection Procedure

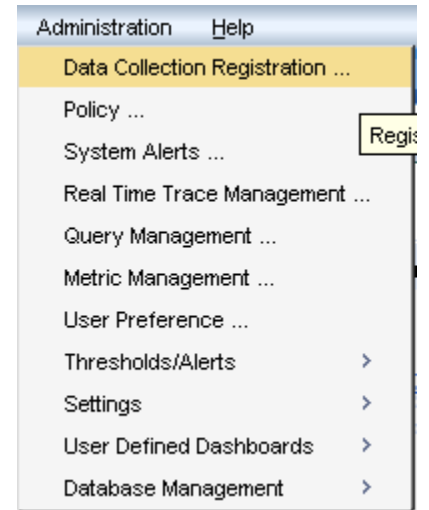
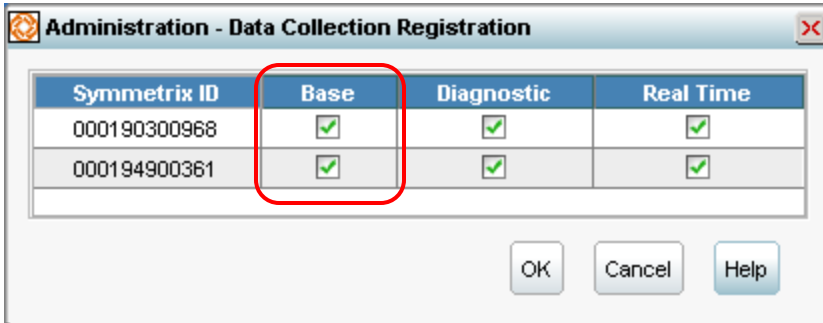
### Option 1 - Collecting Performance Information with Symmetrix Performance Analyzer

Symmetrix Performance Analyzer (SPA) must be configured and licensed in your environment to use this option. SPA is accessed via the Symmetrix Management Console (SMC) application.



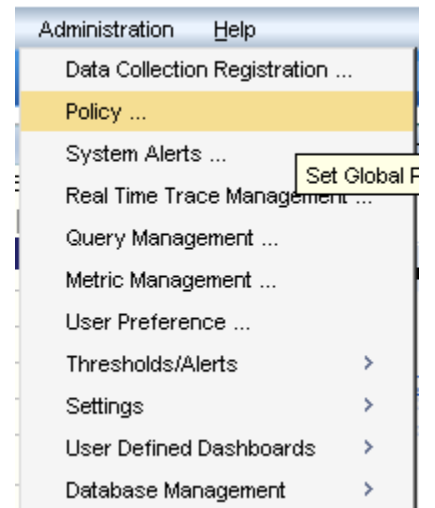
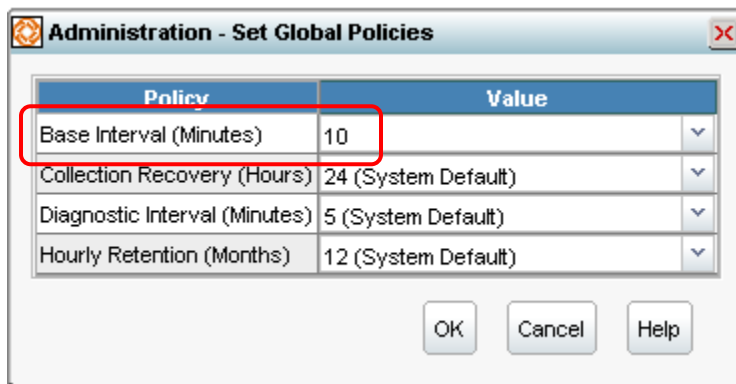
### Setup Data Collection Registration

- 1) From the Administration Menu, select *Data Collection Registration...*
  - a. The following dialog box opens, showing all arrays known by SPA
- 2) Verify that *Base* is checked for the array(s) that are to be analyzed. The checkboxes for *Diagnostic* and *Real Time* are irrelevant for this effort. Click "OK".



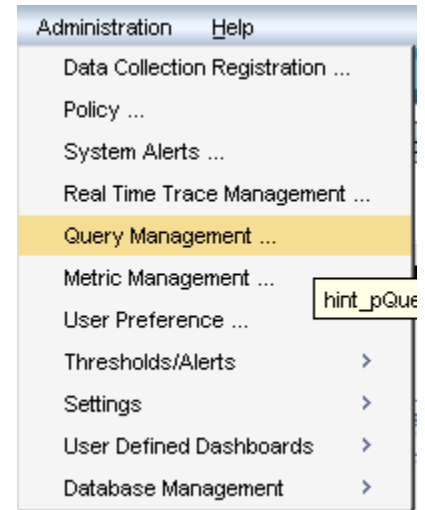
### Setup the Data Collection Policy

- 1) From the Administration Menu, select *Policy ...*
  - a. The following dialog box opens, showing the data collection policies and their respective settings.
- 2) Verify that that *Base Interval (Minutes)* is set to a value of 10. Click "OK".
  - a. The other values are irrelevant for this effort.



## Schedule the Data Collection

- 1) From the Administration Menu, select *Query Management ...*
  - a. The following dialog box opens, showing the existing queries and their respective settings.
- 2) Verify that that *btp file* entry is defined with the *Time Range* set to *Last 4 Hours*. Verify that this entry is enabled by placing a check in the *Scheduled* column. Click "OK".
  - a. The other values are irrelevant for this effort.
- 3) Let the data collection continue for at least three days.



**Administration - Query Management**

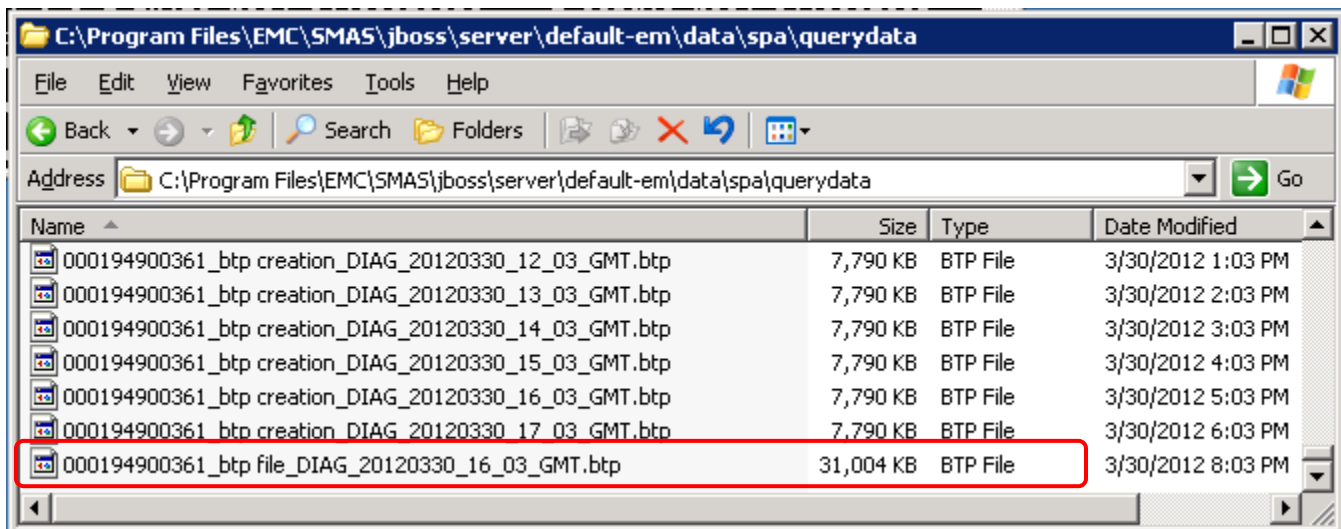
Name	Symmetrix Id	Type	Format	Time Range	Scheduled	Start Date/Time
btp creation	000194900361	Diagnostic	BTP	Last Hour	<input type="checkbox"/>	N/A
btp file	000194900361	Diagnostic	BTP	Last 4 Hours	<input checked="" type="checkbox"/>	Fri Mar 30 20:03:00
high vols in a tier	000194900361	Diagnostic	CSV	Last Hour	<input type="checkbox"/>	N/A
high write volumes	000194900361	Diagnostic	CSV	Last 4 Hours	<input type="checkbox"/>	N/A
IODensity Pool	000194900361	Diagnostic	CSV	Last 4 Hours	<input type="checkbox"/>	N/A

Buttons: Execute, New ...

## Collect the BTP Files

- 1) Open a Windows Explorer window and go to the data directory where the BTP files are stored:
 

```
C:\Program Files\EMC\SMAS\jboss\server\default-em\data\spa
```
- 2) Locate the BTP files associated with the "BTP file" query identified in the last step. Add them to a compressed folder or a ZIP file.
- 3) Place the ZIP file into a convenient location so it can easily be transferred to EMC.



## Option 2 – Collecting Performance Information with Solutions Enabler

Symmetrix performance files can be generated on a host with Solutions Enabler 7.x or greater installed directly attached to the arrays that need to be analyzed.

### Configure the storstpd Daemon

Verify the settings in the daemon\_options file. The file can typically be found at the following location, depending on the platform of your host:

UNIX: /var/symapi/config/daemon\_options

Windows: C:\Program Files\EMC\SYMAPI\config\daemon\_options

Example of the daemon\_options file header:

```
#####
# The daemon_options file contains parameters that can be set to control
# the behavior of the various EMC Solutions Enabler daemons. Each daemon reads
# this file as it starts and applies any settings within it that apply.
#
#
# CAUTION: These parameters are intended for experienced users. In
# most scenarios, the default values used by the daemons will be sufficient.
#
# CAUTION: This file should be protected so that only authorized users can
# make changes to it.
#
# Lines in this file are allowed to have one of the following formats:
#
# NAME = VALUE Set the parameter 'NAME' for all
# daemons that understand this parameter.
#
# stororad:NAME = VALUE Set the parameter 'NAME' for only the
# stororad daemon.
#
# storora*:NAME = VALUE Set the parameter 'NAME' for all
# daemons whose name begins with
# storora. The asterick (*) character is
```

Search the file for the following options and set their values accordingly, then save the file:

- dmn\_run\_ttp = enable
- dmn\_symmids = *your Symm Serial #(s)*
- dmn\_max\_arrays = *some value, default is 6*
- storstpd:ttp\_collection\_interval = 10
- storstpd:ttp\_retention\_days = 7 (or greater)
- storstpd:ttp\_tp\_dev\_metrics = enable
- storstpd:sync\_vp\_data = enable

More information about the daemon\_options file can be found in the *EMC Solutions Enabler Symmetrix CLI Command Reference* documentation found on EMC's PowerLink website: <http://powerlink.emc.com>.



### *Start the storstopd Daemon*

- 1) First, check the hosts' process table to see if storstopd is already running. If the daemon is already running, it should be stopped. To cleanly shutdown the daemon from the command line, type:

```
stordaeomon shutdown storstopd
```

- 2) To start the daemon, type the following on the command line:

```
stordaeomon start storstopd
```

- 3) In order to setup the daemon to start automatically after a system boot, type the following command:

```
stordaeomon install storstopd -autostart
```

- 4) Let the data collection continue for at least three days.

### *Collect the TTP Files*

- 1) Open a Windows Explorer window and go to the data directory where the TTP files are stored:

```
C:\Program Files\EMC\SYMAPI\stp
```

- 2) Locate the TTP files within the date range that needs to be analyzed. Add them to a compressed folder or a ZIP file.
- 3) Place the ZIP file into a convenient location so it can easily be transferred to EMC.

## Step 2 – Collect Configuration Information

### Collection Considerations

Historically, BIN files refer to the file containing the configuration that has been applied to a Symmetrix array. EMC's sizing tools also accept a more complete "BIN" file: the SYMAPI database file, which also has a .bin extension. This file contains dynamic configuration data that is not contained in the configuration BIN file, and having this information greatly helps the analysis process.

### Solutions Enabler

Symmetrix configuration files are generated on a host with Solutions Enabler installed directly attached to the arrays that need to be analyzed.

#### Step 1 - Locate the SYMAPI Database File

By default, you can find the BIN file in the following locations when Solutions Enabler is installed on the host:

UNIX: `/var/symapi/db/symapi_db.bin`

Windows: `C:\Program Files\EMC\SYMAPI\config\symapi_db.bin`

If using the storstopd daemon to collect the performance information, the daemon can also be setup to store the active SYMAPI database file at the same time by setting the "*storstopd:ttp\_config\_db*" value to "*enable*" in the daemon\_options file. In this case, you will find the database file at the same location as the performance data.

#### Step 2 - Update the SYMAPI Database File

To update the SYMAPI database file, use the following command:

UNIX: `/usr/symcli/bin/symcfg discover`  
`/usr/symcli/bin/symcfg sync -local`  
`/usr/symcli/bin/symcfg sync -fast`  
`/usr/symcli/bin/symcfg sync -vpdata`

Windows: `C:\Program Files\EMC\SYMCLI\bin\symcfg discover`  
`C:\Program Files\EMC\SYMCLI\bin\symcfg sync -local`  
`C:\Program Files\EMC\SYMCLI\bin\symcfg sync -fast`  
`C:\Program Files\EMC\SYMCLI\bin\symcfg sync -vpdata`

If your server is connected to multiple arrays, you can add the "-sid" parameter to reduce the discovery time for each of these commands.

This process will ensure that the latest running configuration is collected.

#### Step 3 - Collect the SYMAPI Database File

If using data collected by the storstopd daemon AND the configuration files were collected at the same time, include these in your zip file that was created in the first section, Symmetrix BTP and TTP.

Otherwise copy the symapi\_db.bin file into the same directory as the zip file with the performance data so everything is in one place for easy transfer to EMC.

## Step 3 - Copy the Files to EMC Anonymous FTP Server

Note: EMC has a new Anonymous FTP server implementation. You will not be able to see the files that are transferred.

- 1) Verify that all SPCollect and Performance data files are located in one directory on your system.
- 2) From a command line interface:
  - a. Change your working directory to be where the SPCollect and Performance data files are located.
  - b. Initiate the ftp connection
    - `prompt> ftp ftp.emc.com`
    - `username: anonymous`
    - `password: your.email`
  - c. Change to a binary file transfer.
    - `ftp> bin`
  - d. Turn on the Hash feature so that you can see the progress of the file transfer.
    - `ftp> hash`
  - e. Create a unique subdirectory for your data files to be copied
    - `ftp> mkdir mysubdir`
  - f. Change to the directory that was just created
    - `ftp> cd mysubdir`
  - g. Copy the files to the ftp server
    - `ftp> put filename`
  - h. Close your ftp session
    - `ftp> quit`
- 3) Notify your EMC Sales Team of the location of your data files.