Hyper Historian – Estimating Disk Space for Data Logging

APPLICATION NOTE

July 2011

Description: Guide to estimating the required disk space for logging historical data with Hyper Historian.
OS Requirement: Windows Server 2003 x64/Vista x64/ Server 2008 x64/Windows 7 x64/ Server 2008 R2 x64
General Requirement: Installation of Hyper Historian v10.6

Introduction

When estimating the amount of disk space that is required to run Hyper Historian, you should consider the following:

- The available disk space to install required software components and files
- The disk space required to store Hyper historian data files
- The disk space required to store the Hyper Historian archive data files

Table 1 shows the minimum disk space required to install the software (besides the operating system) for Hyper Historian.

Table 1 - Software Disk Space Requirement

Software	Minimum Disk Space Required
MS SQL Server	Refer to Microsoft SQL Server Documentation
Common Components	30 MB or more
Hyper Historian Program files, including Workbench and documentation	400 MB

Disk Requirements for Configuration Database Files

All historian configuration data and event data are stored within a separate database. The configuration data in the database file remains relatively static and usually never causes the file size to go above tens of megabytes. However, if you set up events, the records of event detections and the results of any data summaries or snapshots will increase the size of the configuration database file because the tables are filling up. Also, entries are created in the log file for event-related transactions. If the database files are set to auto-size, the configuration database file expands to accommodate event-related data until the hard drive is full.

A temporary database stores tag definitions being imported from integrated HMI software.

> **NOTE:** Historical plant data is not stored in the database files. This type of data is stored in special archive files.

The Minimum disk space for the configuration database files is shown in Table 2.

Table 2 - Configuration Database Size

File	Minimum Disk Space Required
Configuration Database (*.mdf)	10 MB
Configuration Log (*.ldf)	10 MB

The configuration database is set to automatically expand at a rate of 10% (the default).

You cannot change these defaults during the installation. The database can be resized later using Microsoft SQL Server utilities. For more information on sizing databases, see your Microsoft SQL Server documentation.

Disk Requirement for Historical Data Files

Hyper Historian stores historical plant data to the hard disk in special archive files. By default during installation, the Hyper Historian is set up to log data to the following path: *ProgramData*\ICONICS\HyperHistorian\Data. You may specify a different storage location (directory) after installation, where these files will be created and then filled. This can be done using the Hyper Historian Configurator Loggers Settings screen. Be sure that you have allocated sufficient disk space for storing plant data for your desired length of time.

The amount of disk-based storage that will be needed depends on the number and frequency of stored tag values. The more tags you have, the fewer values that can be stored per tag before archiving the oldest data becomes necessary. The higher the specified tag storage rate, the faster it consumes disk space.

Data Type Sizes:

Data Type	Size
Int8, UInt8	1 byte
Int16, UInt16	2 bytes
Int32, UInt32	4 bytes
Int64, UInt64	8 bytes
Float32	4 bytes
Float64	8 bytes
Boolean	0 bytes

Copyright 2011 ICONICS, Inc.

Data Type	Size
String	String length * 2 bytes
Timestamp	4 bytes
Quality	1 to 3 bytes (depends on presence of sub-status code and/or history bits)
Trigger identification*	4 bytes

*This is optional settings uses with Calculations. We are not counting with these 4 bytes in the calculations bellow.

Analog and Discrete Storage Requirements

For analog and discrete tags, each value that is stored uses Storage size + 4 bytes of disk space, plus approximately 5% overhead (reserve). Use the following formula to estimate the disk usage:

*Estimated disk usage per day = (Size + Timestamp + Quality + 1) * (1 + reserve in percent / 100) * Number of tags * [Approx. number of samples per day]*

Calculation Examples

Example 1: Analog Tags
The disk usage per day for 10,000 Float32 analog tags (that is, Storage Size = 4 bytes) that are stored at 10-second intervals (for example, 6 per minute * 60 * 24 = 8,640 samples/day) would be: (4+4+1+1)*(1+(5/100))*10000*8640 which is approximately 907 MB per day.

Example 2: Discrete Tags
The disk usage per day for 10,000 discrete tags (that is, Storage Size = 0 bytes) that are stored at 10-second intervals (for example, 6 per minute * 60 * 24 = 8,640 samples/day) would be (0+4+1+1)*(1+(5/100))*10000)*8640, which is approximately 544 MB per day.

String Storage Requirements

The requirement for logging strings depends on the length of the string; a longer string consumes more disk space. For strings, the number of bytes required for storage is: the actual number of characters being sent from the data source times 2, plus 5% overhead (reserve). Therefore, the average disk usage per day for a variable-length string that is receiving an average of 'N' characters per sample could be calculated using the following formula:

*Estimated disk usage per day = (N * 2) (1 + reserve in percent / 100) * Number of tags * [Approx. number of samples per day]*

Calculation Example

The disk usage per day for 1,000 string tags that are changing, on average, every 60 seconds where the average incoming value is 50 characters in length would be: (50*2)*(1+(5/100))*1000*1440, which is approximately 151 MB per day.

Advanced Considerations

To achieve the best write/read performance, data slot sizes should be a multiple of disk allocation units. On most Microsoft operating systems the default disk allocation unit is 512KB. Data slot sizes are configurable in whole kilobytes between 1 and 255; the default value is 4KB. Since 4KB divides evenly into 512KB, this is perfectly acceptable for most configurations.

Each data tag has an allocated standalone data slot for data. To be precise in your calculations, you should consider the size of a single data slot. Count the number of values in (single slot + reserve for future updates), then multiply it by the number of data slots.

For example: Let's assume a data slot size of 4KB (the default) and an application with 10,000 tags. The initial update of each tag will create a file of size of 10000 * 4KB + 4KB (header size), which is approximately 40MB. This size remains unchanged until the data slot becomes full of logged data, and then the Hyper Historian will create additional data slots as needed.

Most users will not need to modify the data slot size; however, this is important information to understand if you are closely monitoring the growth and performance of your Hyper Historian application.