A new way to analyze historical performance data on IBM i

Using Performance Data Investigator Graph History function

Lora R. Powell

May 18, 2016
(First published May 18, 2016)

As you cannot predict the future, the next best thing is to study the past. The Graph History function in IBM® Navigator for i helps you do that. With Collection Services historical data and the Graph History function as part of the Performance task on the web, you can understand how your system performance compares to last week, last month, or even last year. Collection Services now provides a new collection type which aggregates a subset of data over long time periods. Performance Data Investigator can analyze and display this new historical data collection type. Historical data combined with the power to view and analyze it across days, weeks, and years provides new insight about system performance. This article describes how the Graph History function in IBM Navigator for i can help you understand your data, and how system performance changes over time.

Graph History in IBM Navigator for i

Collection Services and Performance Data Investigator have previously allowed you to look at a day's worth of performance data in graphical views. System Monitoring and Performance Data Investigator allow you to monitor live performance data graphically. Now in IBM i 7.3 release there is historical data support with Collection Services and graphical display of that data with Graph History in the Performance task of *IBM Navigator for i.*

Historical data can help you identify trends and investigate performance issues from the past or to compare today's performance with a time period in the past that you select.

Historical data

Collection Services keeps up to 24 hours of data in each collection. By default, the collections are preserved on a system for 10 days. To look at performance from week to week, month to month, or even year to year, it was necessary to provide a way to keep performance data for longer periods, without taking up too much storage and also to allow an easy way to view data across multiple days.

Collection Services now does this through the new **Historical Data** collection. The new performance collection type *HSTFILE contains historical data in a file-based performance collection.

The Historical Data collection is data collected by Collection Services that has been reduced and summarized to keep the data over a long period of time and conserve space. The data is intended to be just a subset of the original Collection Services data.

The Historical Data collection will be populated with data from management collection objects in the configured collection library (for example, QPFRDATA) when the Collection Services collector is cycled or ended.

To enable the concept of aging or decaying data, there are two levels of historical data – detail and summary. As you can see from the following table, as the maximum collection duration increases, the number of metrics available decreases.

Table 1. Collection data duration and variety

| Data	Maximum duration of collection	Number of metrics available (variety of data)
Collection Services data in *CSFILE	24 hours	most
Historical Detail data in *HSTFILE	60 days	fewer
Historical Summary data in *HSTFILE	50 years	least

The **summary** historical data is the system-level and summarized metrics. This data is useful in identifying trends or detecting changes in a system over a long period of time.

The **detail** historical data is the collected data saved for top contributors of each interval, as well as other relevant supplementary data. This is the supporting data for the summarized metrics. The detail data can be used when looking deeper into a problem identified from summary historical data. Only the top contributors for each metric will be stored as historical detail data.

Create historical data on IBM i 7.3

To turn on the creation of historical data on an IBM i 7.3 partition, use the Collection Services configuration panel. Under the **Performance** task on *IBM Navigator for i*, click **Configure Collection Services**.
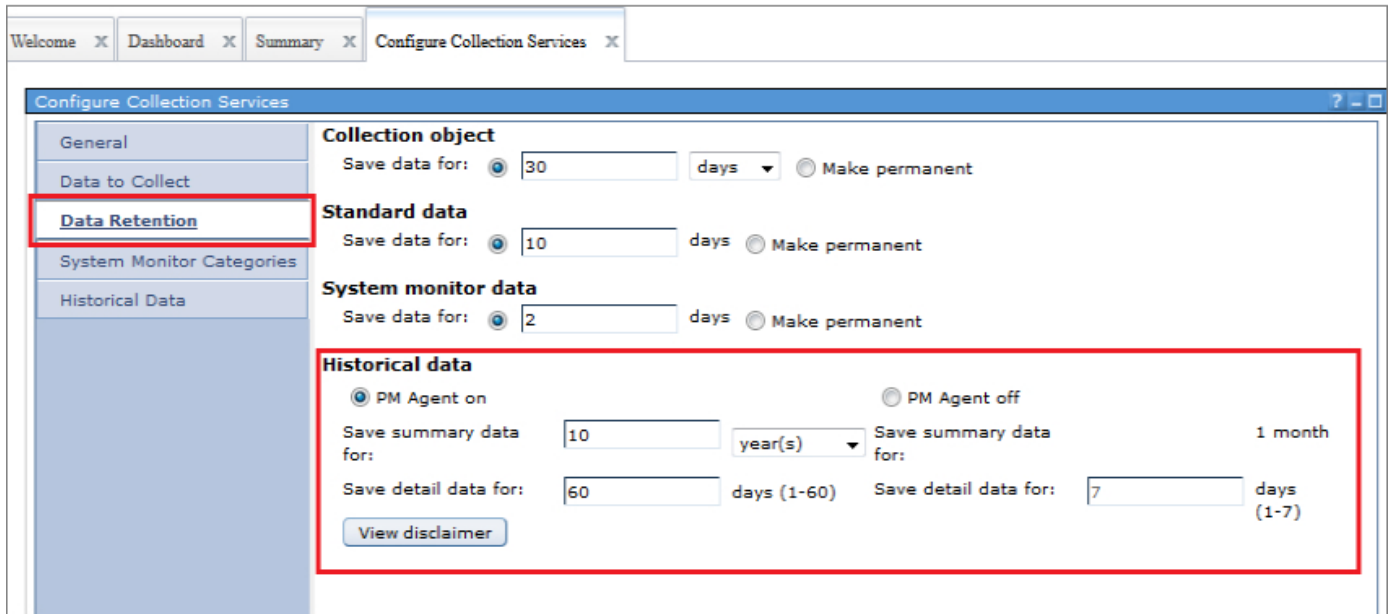
Figure 1. Configure Collection Services
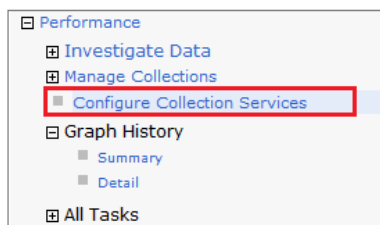
On the **General** tab, select the **Create historical data when collection is cycled** check box. This check box is cleared by default.

Figure 2. Option to create historical data
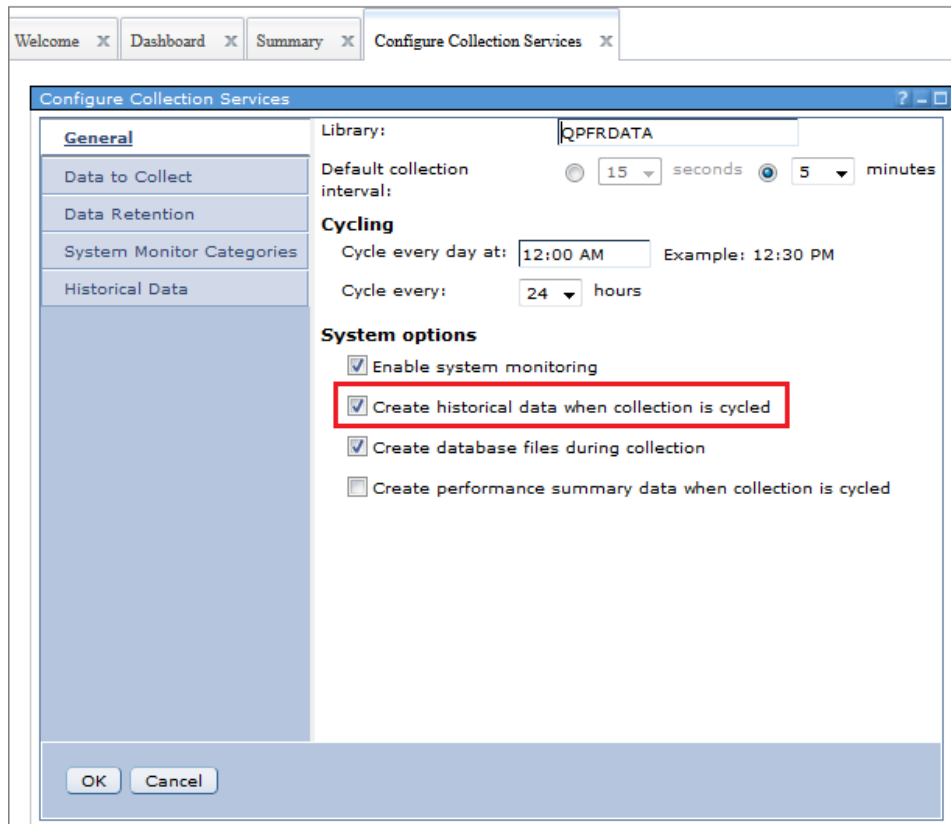
To set the historical data retention options, click the **Data Retention** tab.

Figure 3. Historical data retention settings

On this panel, you can specify the length of time that historical data will be kept. The time span that historical data can be saved depends on whether or not PM Agent is turned on. If PM Agent if set to *off*, Collection Services will only preserve 7 days of detail historical data and 1 month of summary historical data. With PM Agent set to *on*, Collection Services will preserve up to 60 days of detail historical data and 50 years of summary historical data. The following table shows the maximum and default settings.

Table 2. Historical Data retention

Historical Data Retention	Maximum	Default
Summary Data with PM Agent off	1 month	1 month
Detail Data with PM Agent off	7 days	7 days
Summary Data with PM Agent on	50 years	1 year
Detail Data with PM Agent on	60 days	7 days

For more historical data configuration options, click the **Historical Data** tab.

Figure 4. Historical Data settings
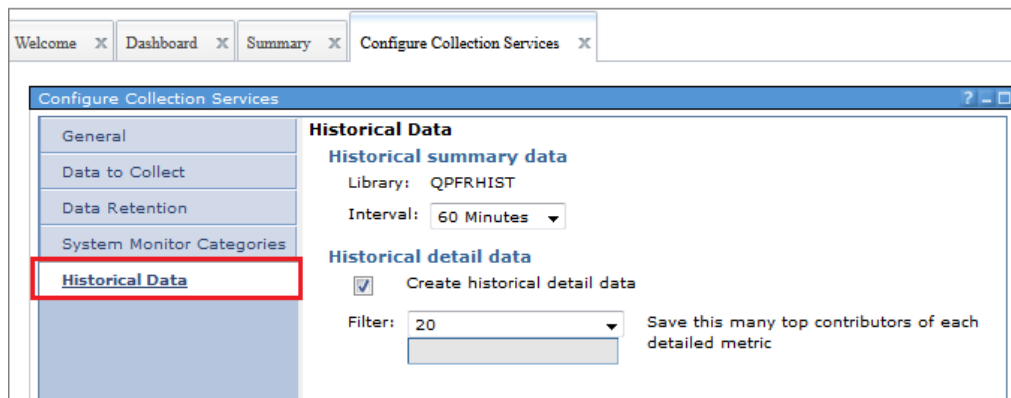
On this panel you can specify the historical data interval (default is 60 minutes). The summary and detail historical data will be saved in the Historical Data collection at this requested interval.

You can use any the following options to configure historical detail data.

- **Create historical detail data**
  You may create only summary historical data or both summary and detail historical data. The default is to create historical detail data when summary historical data creation has been selected.

- **Detail data filter**
  The detail data filter determines what number of top contributors for each metric will be stored in the detail data files. The **top contributors** for each metric means the job or task, disk, memory pool, or communication line that had the most significant values (CPU %, response time, utilization, and so on) during that interval. Only data for these top contributing elements are saved as historical detail data for that interval. The default is that 20 top contributors for each metric at each interval will be stored.
  After you have set the options, click **OK** at the bottom of the panel. If you have turned PM Agent *on* when it was previously *off*, you will need to read and agree to the disclaimer displayed regarding the use of PM Agent.

Note: Turning PM Agent *on* does not cause data to be transmitted to IBM. The user must still set the contact information and have an Electronic Service Agent (ESA) configured to begin actively transmitting data to IBM.

With these configuration options in place, you can immediately cycle Collection Services to create historical data for any *MGTCOL objects in your configured library. To find this function quickly in *IBM Navigator for i*, type **Cycle** in the **Search Task** field on the left navigation panel and pause.
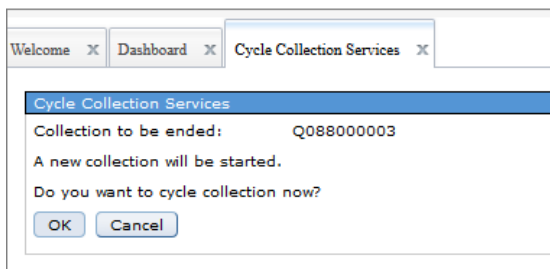
Figure 5. IBM Navigator for i Search Task field
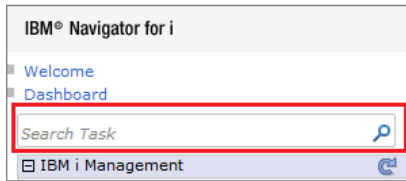
The search shows a list of options with matching text. Select **Cycle Collection Services** and press **Enter** to invoke the action.

Click **OK** in the Cycle Collection Services panel.

Figure 6. Cycle Collection Services panel

This provides you with a starting set of data to view using the Graph History function.

## Creating historical data from existing collections

When you turn on historical data creation, the next cycle of Collection Services will cause all available data in *MGTCOL objects within the configured library to be used to create the Historical Data collection. You can also manually generate historical data from other *MGTCOL objects by using the **Create historical data** action in **Manage Collections**.

Figure 7. Create historical data manually
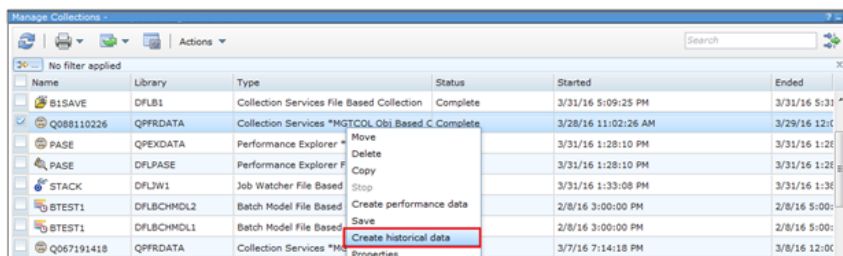
It is also available on the left navigator under **Performance** → **All Tasks** → **Collections**.

Figure 8. Create Historical Data task
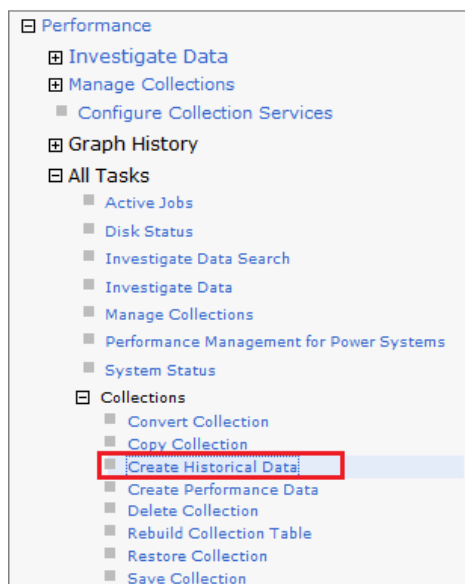
## Visualizing historical data with the Graph History function

After you have historical data created, the new **Graph History** function in the Performance task on *IBM Navigator for i* allows you to view this historical system performance data. As mentioned earlier, this enables you to view data over long periods of time, from days and weeks to months and years. The Graph History function provides access to the new Historical Data collection, both summary and detail performance data.

Graph History Summary charts

Historical Summary charts utilize the historical summary data to provide single metric charts to view and analyze performance over the selected time span. The metrics provided in the summary charts focus on the system monitoring data saved into historical data. From the Summary chart, the top contributor information from the historical detail data for that metric is available by selecting an interval. From that Top Contributor chart, the property data for a top contributing element is available by selecting the element.

To view the historical data, use the new Graph History task section in Performance. To start, select **Summary** under **Performance → Graph History**.

Figure 9. Summary action
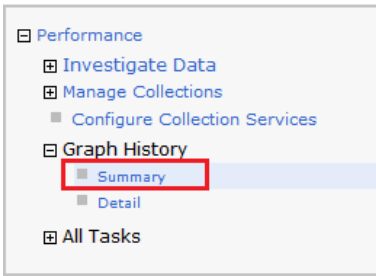
The majority of the new function for the 7.3 release is in the Summary charts. There is one Detail chart available now, and more are expected to be made available in the future.
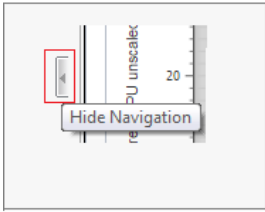
Figure 10. Historical Summary chart
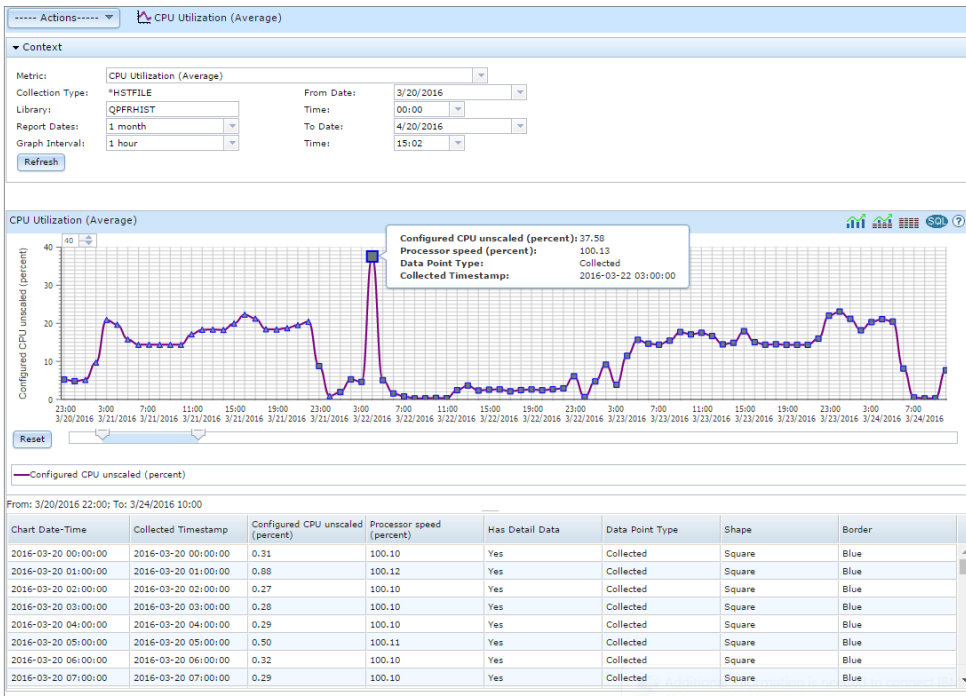
The Historical Summary chart in Figure 10 shows the CPU utilization (average) for 4 days with 1 hour intervals.

To maximize the screen space for charting, click the **Hide Navigation** icon to close the left navigation panel.

Figure 11. Hide Navigator icon

Zoom in

When the chart is first displayed, it displays the data for the past month at one hour intervals. There are too many points to clearly distinguish each interval on the chart, so it is shown as a line. As you zoom in, the individual data points will be shown.

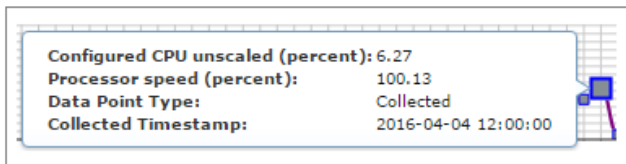To zoom in a specific point or time interval, you can use the slider or the mouse wheel.

To use the mouse wheel, click the data point on the graph for the interval area that you want to zoom in, then use the mouse scroll wheel to zoom in.
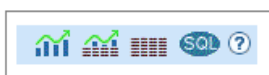
Figure 12. Slider
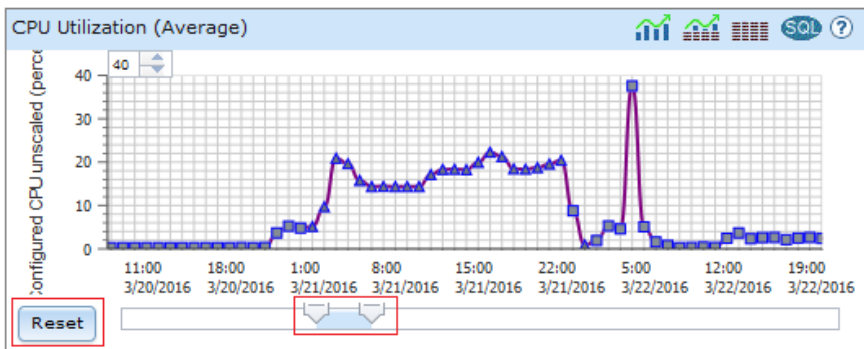
To return to the chart as it was originally presented, click **Reset**.

Figure 13. Tooltips

As you move the mouse pointer over the plotted points on the graph, tooltips appear with the value charted and other detailed metrics about each interval data point.

Figure 14. Icons

The five icons at the top-right corner of the Summary chart allow you to modify the presentation of the panel. The options are as follows:

- **Chart** only – Select the chart icon to view the data on the chart. This is the default.
- **Chart & Table** – Select the combined chart and table icon to split the screen between the graph and the table.
- **Table** only – Select the table icon to show the dataset only in table format.
- **Display SQL** – Select the SQL icon to display the query used to obtain the data set charted on the graph. This is shown as a pop-up panel and the chart panel will not be affected.
- **Help** – Select the '?' icon for help text.

Figure 15. Maximum graph value
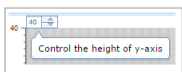
The dial at the top of the chart allows you to zoom in the chart by lowering the graph range (maximum graph value) shown on the chart. Points that occur above this setting will not be seen on the chart.

Context panel

When the Summary chart is first displayed, it shows data starting from today's date in the previous month up to today. The Context panel allows you to select any other available metric to be displayed.

Figure 16. Summary chart
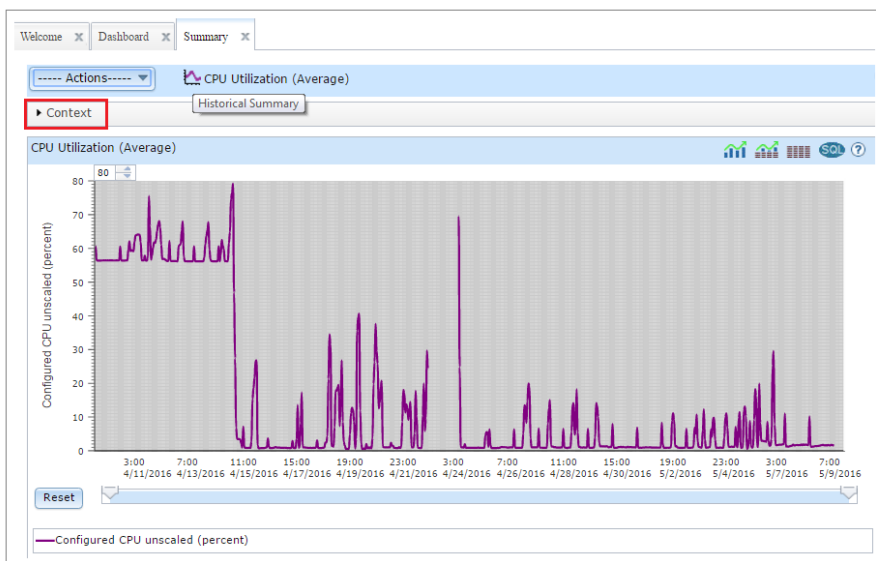
The Context panel is closed when the chart is displayed or updated to provide more room for the chart. Expand **Context** to open it.

Figure 17. Context panel
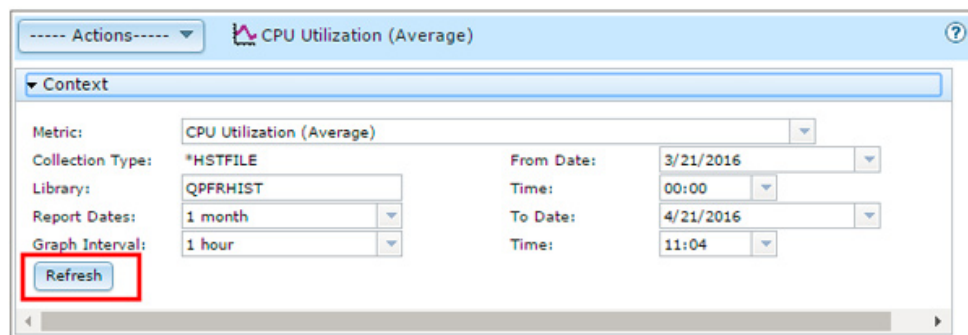
Summary metrics

- Batch Logical Database I/O Rate
- CPU Utilization – Average, Interactive Jobs, SQL, Uncapped
- Communication Line Utilization – Average, Max
- Disk Arm Utilization – Average, Max
- Disk Arm Utilization for Independent ASP(s) – Average, maximum
- Disk Arm Utilization for System ASP – Average, maximum
- Disk Arm Utilization for User ASP(s) – Average, Maximum
- Disk Response Time – Read, Write
- Disk Storage Utilization – Average, Maximum
- Disk Storage Utilization for Independent ASP(s) – Average, Maximum
- Disk Storage Utilization for System ASP – Average, Maximum
- Disk Storage Utilization for User ASP – Average, Maximum
- Interactive Response Time – Average, Maximum
- Interactive Transaction Rate
- LAN Utilization – Average, Maximum
- Machine Pool Faults Rate
- Shared Processor Pool Utilization – Physical, Virtual
- Spool File Creation Rate
- Temporary Storage Utilization
- User Pool Faults Rate – Average, Maximum

For more information on summary metrics, see Historical data summary database files in IBM Knowledge Center.

Use the Context panel to view and modify the chart specifics such as metric, report dates, graph interval, and dates.

After any changes to the Context panel, click **Refresh** to update the chart with the new settings.

The **Metric** drop-down list enables you to switch to any of the other available Historical Summary charts. See Summary Metrics sidebar for the list of metrics. Selecting another metric to display will preserve other settings of the Context section such as graph interval and dates.

The **Collection Type** is *HSTFILE.

The **Library** defaults to QPFRHIST. This can be changed to another library where you might have copied or moved a Historical Data collection.

The **Report Dates** field determines the time span to be displayed on the graph.

The **Graph Interval** values available are based on the Report Dates range. For a larger range of time, the graph interval will be larger. For a smaller range of time, the graph intervals available will be smaller. For example, if one month of data is requested, you will be able to set the graph interval to 1 hour, 6 hours, 12 hours, or 1 day (24 hours). If a report date greater than one year is chosen, then the graph interval values available will be one month or one year. Table 3 shows the Graph Interval settings available based on the selected Report Date range.

Table 3. Report dates and graph interval settings

Report date (range of time selected)	Graph intervals available
Today	5 min, 15 min, 30 min, 1 hour
1 week	30 min, 1 hour, 6 hours, 12 hours
1 month (default)	1 hour, 6 hours, 12 hours, 1 day
1 year	6 hours, 12 hours, 1 day, 1 week, 1 month
Custom < 6 hours	5 min, 15 min, 30 min, View All Points
Custom > 1 year	1 month, 1 year

The **From Date** and **Time** fields and **To Date** and **Time** fields allow you to select a custom range (report dates) to be displayed.

The Historical Summary chart initially displays data starting from one month ago and ending with the collection last ended (often at the default midnight cycle time). To modify the dates and times displayed, first modify the Report Dates field or change the **From Date** and **Time** fields and **To Date** and **Time** fields. Then, select a **Graph Interval** value and click **Refresh**.

To view historical data collected today, you must manually generate historical data for the active Collection Services *MGTCOL collection by using the Create historical data action in Manage Collections.

Understanding the Historical Summary chart

The Summary chart is created for one metric over the specific time range selected in the Context section. Each point on the graph is shown with a symbol and color which indicates information about the data charted at that time interval.

Data point symbols

Circle data points (summary data only):
🔵 Blue border – collected or aggregated
⚪ White border – extended

Square data points (detail data available):
🟦 Blue border – collected
⬜ White border – extended

> Triangle data point (detail data available):
> ▲ Blue border – aggregated

A **circle** represents a data collection point that only has historical **summary** data available.

A **square** or **triangle** both represent data collection points that also have historical **detail** data available. You can click on these to see the top contributors data for that interval. The Top Contributors chart appears at the right side of the Graph History window.

The outline of the data point symbol also indicates information about the data being charted. The data point may be collected, aggregated, or extended data point.

- A **collected** data point is a value taken directly from a historical database file.
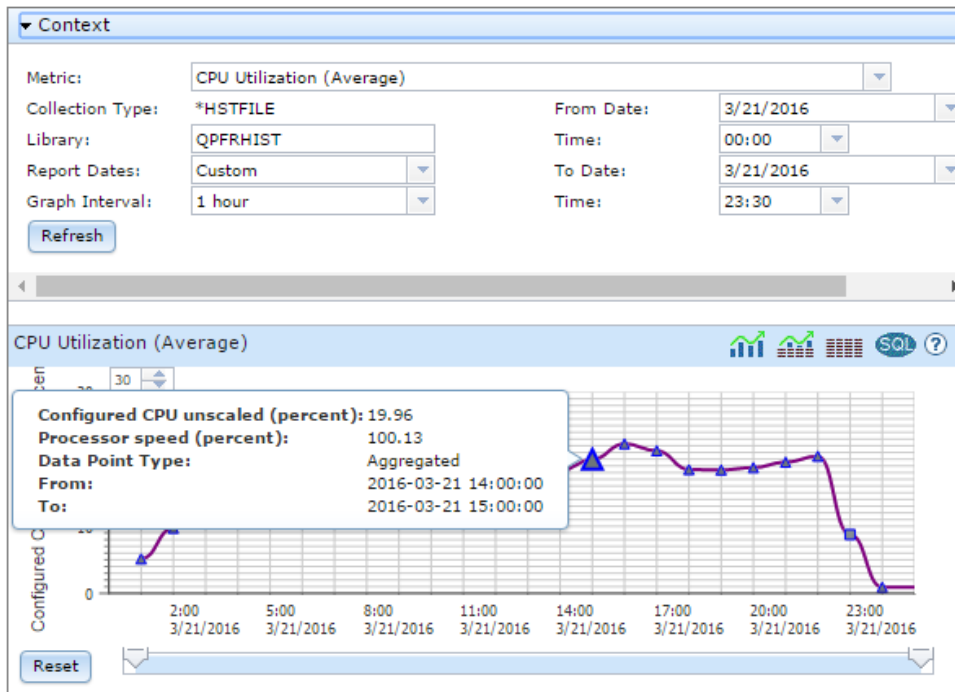
Figure 18. Collected data point
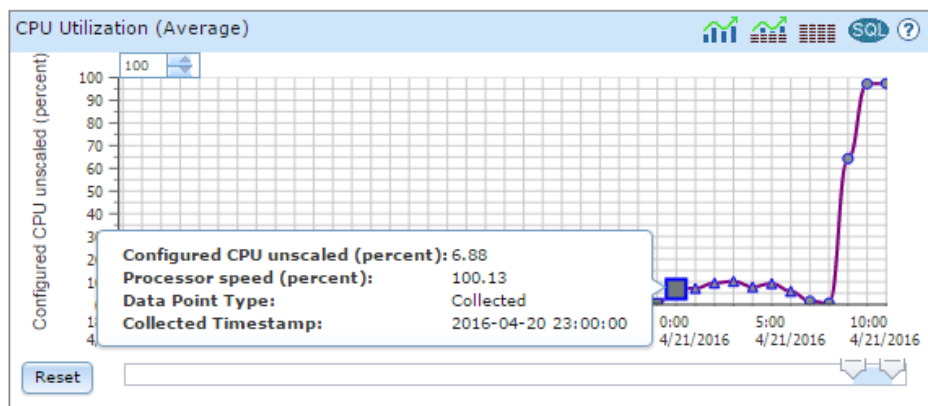
The historical data requested interval and graph requested interval match at the highlighted data point.

- An **aggregated** data point is a calculated value from the metric at multiple intervals in the historical database files. Aggregation is done when the collected interval is less than the requested graph interval. Calculations such as sum, average, or max are performed depending on the metric and what it represents.
  For example, if data is collected at 15-minute intervals and the chart is showing 1 hour graph intervals, the data point will be a calculation of metrics from 4 time intervals.
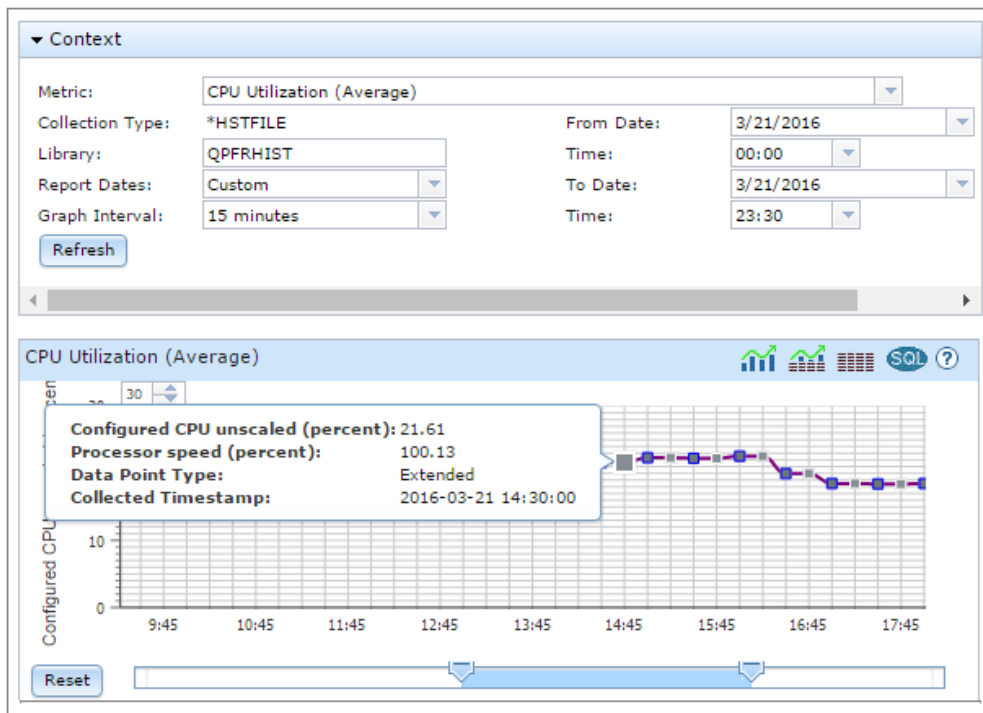
Figure 19. Aggregated data point

The historical data requested interval is smaller than the graph requested interval at the highlighted data point.

- An **extended** data point is a value for that time interval which is not available in the historical database file. The data value is extended from the last available interval value. This is needed to provide integrity to the chart for the graph interval requested.
  For example, if data is collected at 30-minute intervals and the chart is showing 5-minute graph intervals, the data points between the collected intervals will be extended.

Figure 20. Extended data point

The graph request interval is smaller than the historical data creation requested interval at the highlighted data point.

Top Contributors chart

To view top contributor information for an interval, the detail data needs to be available. This is indicated by a square or triangle plot point. Select the data point. The Top Contributors chart is displayed at the right side of the panel.

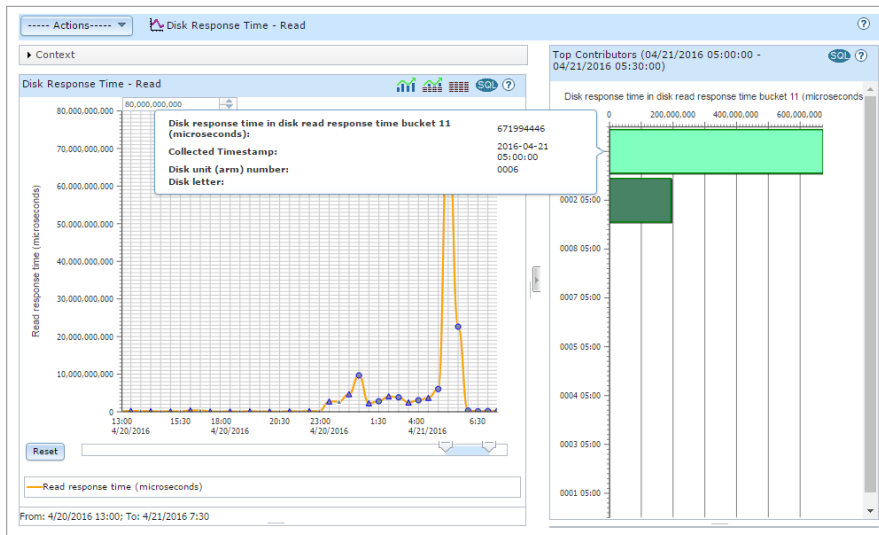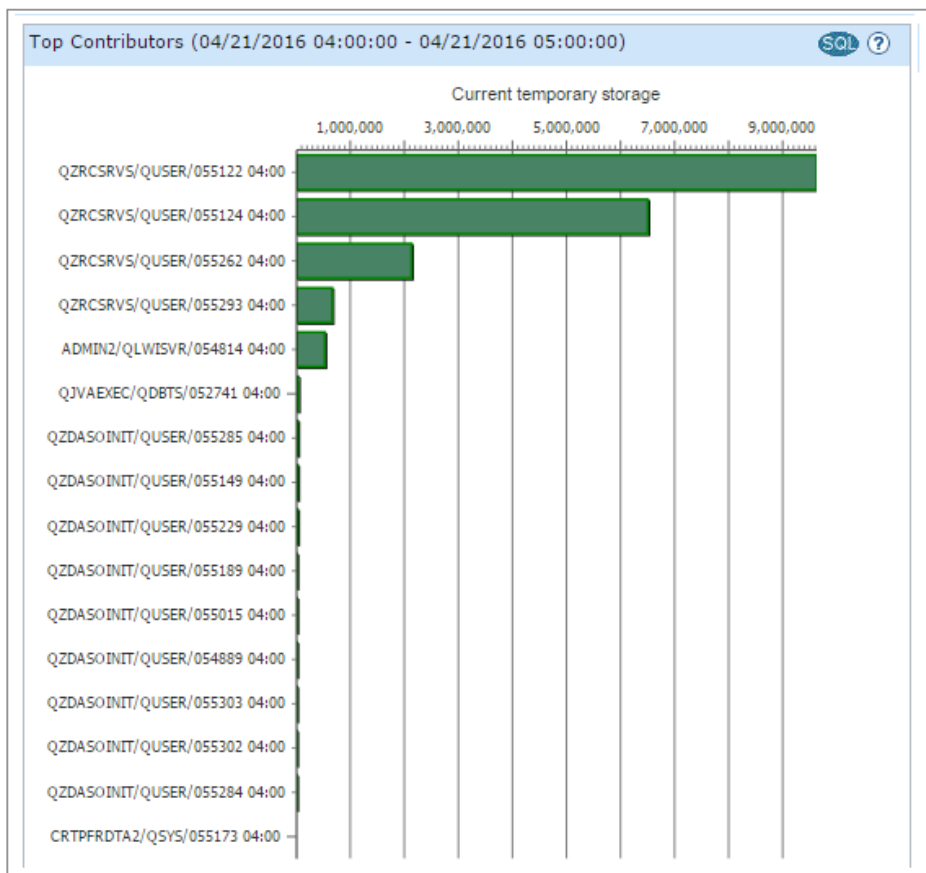## Figure 21. Summary chart with the Top Contributor panel

Figure 22. Top Contributors chart

This chart shows the top contributors for the temporary storage in the selected time period. A top contributor will be listed multiple times if it was a top contributor for more than one interval of an aggregated data point. (See aggregated data point.)

Figure 23. Top contributors with duplicate element
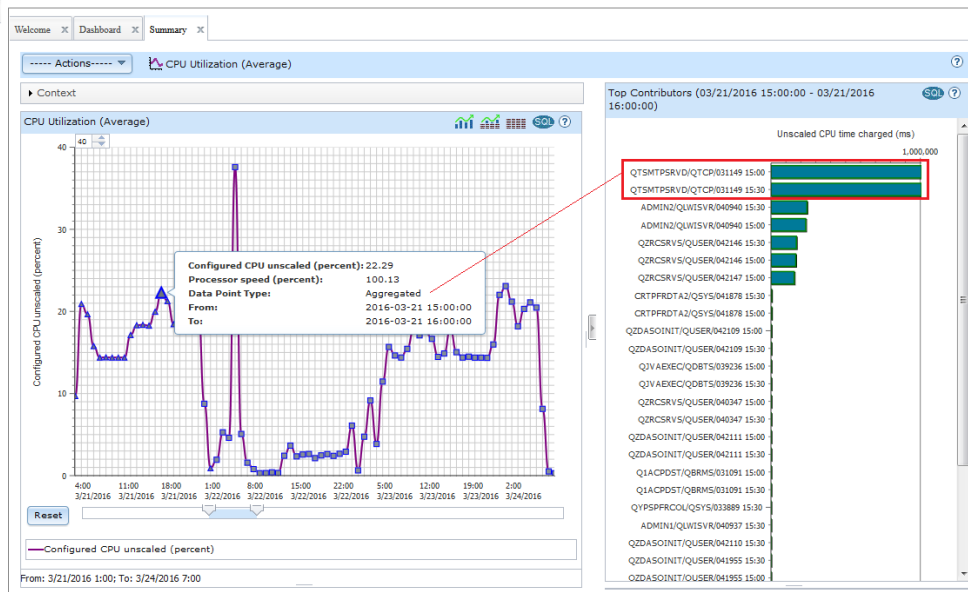
For example, in Figure 23, the triangle data points are those where the data was collected at 30-minute intervals and a 1-hour graph interval was requested. There are several top contributors that are duplicated because they happen to be top contributors for both intervals. For an aggregated data point, both values are included in the calculation for this specific requested interval. Note that the time stamp on the Top Contributors chart helps to differentiate them.

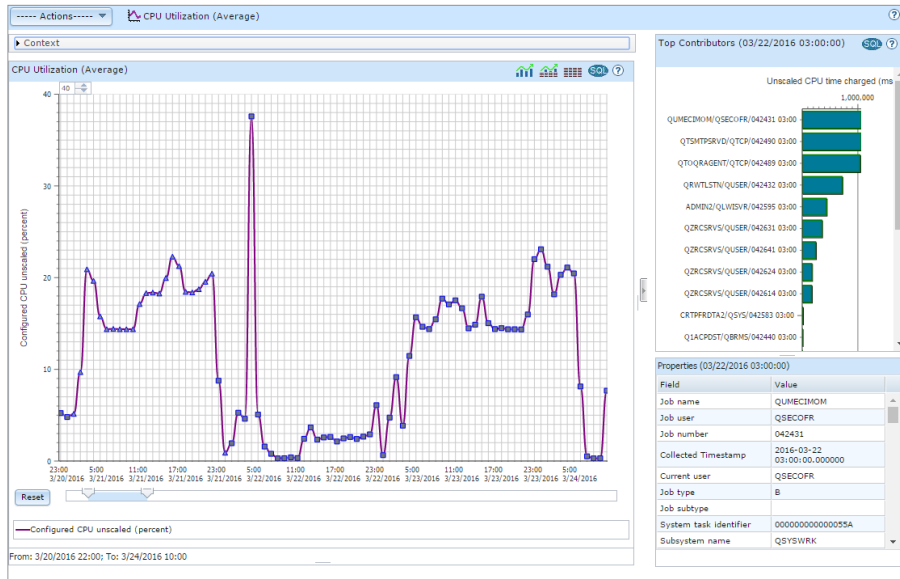To resize this section, click the Hide Top Contributors button.

Figure 24. Hide Top Contributors
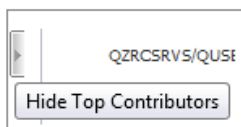
Property data

From the Top Contributors graph, you can view property data for a job or task, memory pool, disk unit, or communications line by selecting a bar from the Top Contributors graph.

Figure 25. Summary chart with Properties panel

The Properties panel appears below the Top Contributors chart.

Figure 26. Properties panel

The Properties panel contains all the information available for the job or task, communications line, memory pool, or disk pool. You can scroll down the panel to see all the fields.

Actions menu

The Actions menu allows you to open the Historical Detail charts. At this time, there is only one Detail chart (CPU Utilization Overview) available. More are expected to be created in future releases.

Figure 27. Actions menu
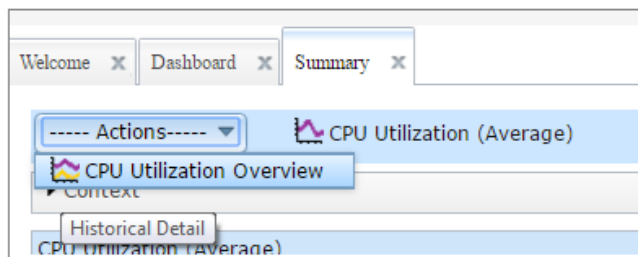
Graph History Detail charts

The Historical Detail charts provide a view of additional data beyond the single metrics provided in the Summary charts. The Detail chart utilizes multiple summary metrics to provide more information at each interval. From the Detail chart, the top contributor information for one metric is available by selecting an interval. From that Top Contributor chart, the property data for a top contributing element is available by selecting the element.

The CPU Utilization Overview Historical Detail chart provides the Partition CPU Utilization and CPU time used from the summary data. When drilling down to top contributors, the metric is CPU time charged to the job.

With the first release of Graph History, the focus was placed on the Summary charts. There is one Detail chart available now, and more are expected to be available in the future.

Requirements

Historical Data and the Graph History analysis tool is included with the base operating system in IBM i 7.3.

PTF requirements

For best results, make sure you have the latest level Group PTF for *IBM HTTP Server for i* installed on your system:

- IBM i 7.3 HTTP group SF99722 level 1 at GA.
- Refer to IBM System i Support to find the latest level available for 7.3.

User profile authority requirements

To access the historical data, a user profile needs to have *ALLOBJ authority or be included in the data authority list QPMCCDATA for access to the Collection Services data.

To use the performance commands in *IBM Navigator for i* (Configure and Cycle Collection Services), a user profile needs to have *JOBCTL authority or be included in the function authority list QPMCCFCN to access the Performance Tools functions.

Summary

We can't predict the future. We can't tell you today what your system performance will be tomorrow. However, with this new Graph History function, you can look back in time and better analyze your system performance over a long term. This will allow you to anticipate future performance. Turn on historical data now and in the future you can examine past system performance. This will give you hindsight about what is happening on your system today.

This 7.3 release with Graph History support and the new Visualize Monitor Data, which charts all metrics for a system monitor on one panel, provides another significant step in the replacement of Management Central function with *IBM Navigator for i* for a single system.

Resources

- Refer to IBM i Technology updates on **developerWorks** for Performance on the web.
- Refer to Knowledge Center for IBM i 7.3 – Performance.
- Follow Dawn May's blog, i Can… Investigate Performance Data and look for the latest blogs in the *Performance* category.