How to use the Batch Model performance tool

Shauna B. Rollings
Dave Legler
Shan Gao

August 07, 2014

This article is an introduction on how to use the Batch Model performance tool available in IBM® Navigator for i in version 7.2. Batch Model uses Collection Services performance data to predict batch workload run times, resources used, and the duration of your batch window. It predicts the changes in throughput that result from hardware upgrades (processor or disk) or from a workload increase. It also can help you optimize workloads by providing a timeline view of your batch window, so that you can locate the times when more efficient job scheduling can improve total system throughput.

Introduction

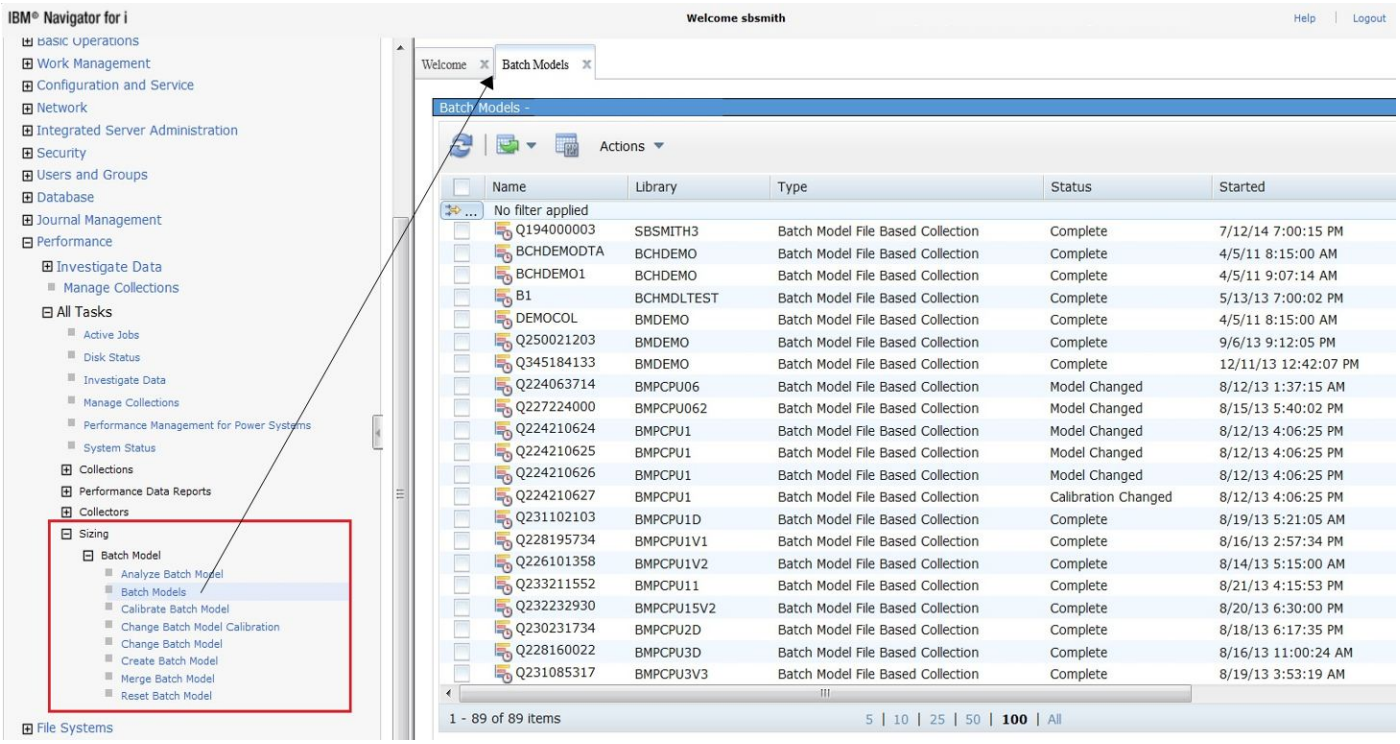
Batch Model is a new function in IBM i 7.2 under the performance task in IBM Navigator for i. To use it, you must have the Manager feature (option 1) of the Performance Tools Licensed Program Product (5770PT1) installed. Batch Model uses Collection Services performance data to predict batch workload run times, resources used, and duration of the *batch window*. The batch window (typically) refers to overnight batch runtime requirements for a particular application or applications to complete while interactive use of the system is at a minimum.

Batch Model helps optimize workloads by locating times during the batch window when more efficient job scheduling can improve total system throughput. It can predict changes in throughput that results from disk or processor upgrades. It can also predict run times for individual workloads and the overall batch window. It can model increases in batch workload growth, and can model workloads that are processor or disk intensive.

Batch Model links individual workloads together to create an ordered series of workloads. It is powered by an iterative analytic model that attempts to converge on a solution. It can predict utilization, throughput, and response time for each workload.

This article explains how to use this new tool and how to interpret the results.

Figure 1. Batch Model functions



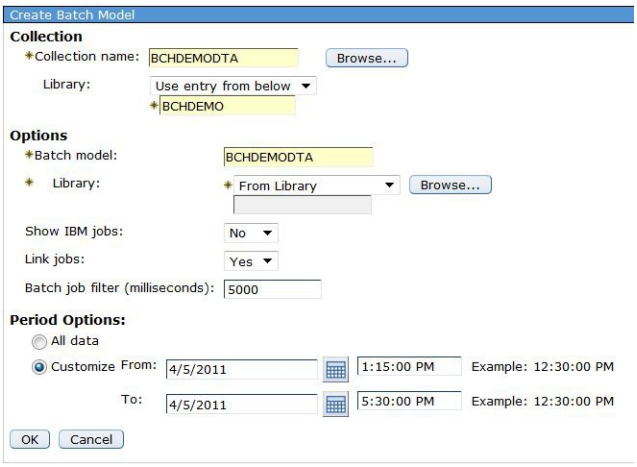
Batch Model is divided into the following different functions:

- **Create Batch Model** – Create a model from Collection Services data.
- **Change Batch Model Calibration** – Adjust workload characteristics and disk configuration for a more accurate model.
- **Calibrate Batch Model** – Re-create the model results after making changes to the calibration.
- **Change Batch Model** – Set the properties for the scenario you want to model: workload growth, processor upgrade, disk upgrade, and changes to workloads.
- **Analyze Batch Model** – Run the iterative analytic model to create model results.
- **Investigate Results** – View the modeled results: workload start/stop times, dependencies between workloads, and amount of resources used.
- **Merge Batch Model** – Merge two different Batch Model collections into one. This function allows you to merge batch models created from measured data collected on multiple different systems. This is useful if you want to model consolidating workloads from multiple systems into one. You also might need to merge batch models if the batch jobs you want to model run across two or more Collection Services file-based collections.
- **Reset Batch Model** – Set the status of a Batch Model collection to *Reset*. If your Batch Model collection is set to an *Error* status, you must reset your Batch Model collection to set the status to *Reset.* This allows you to make changes and then retry the operation that caused your Batch Model collection to end with an error.

Create Batch Model

This is the first step when using Batch Model. A Batch Model collection is created based on the existing performance data that is collected by IBM i Collection Services. This is referred to as the measured data. During this process, the measured data will be analyzed and a model will be built.

Figure 2. Create Batch Model



Collection

- **Collection name** – The name of the Collection Services file based collection on which this Batch Model collection will be based.
- **Library** – The library of the Collection Services file based collection on which this Batch Model collection will be based.

Options

- **Batch Model** – It is the name of the Batch Model collection to be created.
- **Library** – It is the library of the Batch Model collection to be created.
- **Show IBM jobs** – If you specify **No**, all system batch jobs (whose user starts with a Q) will be aggregated into one workload named QSYSTEM with other system threads and tasks of the same priority. If you specify **Yes**, each system batch job will be listed individually.
- **Link jobs** – Specify **Yes** to allow Batch Model to automatically create links between workloads. Creating links between related workloads is important when modeling. Ideally, all workloads during the batch window will be linked together so that when the processor or disk is upgraded, each workload runs faster and ends sooner. This means that the next linked workload can start sooner, resulting in a shorter batch window. Specify **No** if all workloads must always start at their fixed start time.
- **Batch job filter** (milliseconds) – Batch jobs with an accumulated CPU time below this value will not be listed individually. Rather, they will be summarized into one workload named QSYSBCH with other threads and tasks of the same priority. If your collection has thousands of jobs that consume more CPU time than the filter, then Batch Model takes a long time to run its analysis and the amount of workloads generated in the results will be difficult to investigate. To decrease the time it takes Batch Model to run its analysis and decrease the amount of results produced, set the value for job filter to the number of seconds or minutes that your shortest batch job runs.

Period Options

This option allows you to control by time of day the measured data that is part of the Batch Model collection. By selecting **All Data**, you can use all of the measured data. If you select **Customize,** you specify the starting and ending times (and dates) of the measured data. If you are modeling your batch window, you should select Customize, and enter the start and end time of your batch window.

The **Create Batch Model** action submits a batch job to create the model. Creating a batch model can sometimes be a long-running operation. The creation process has ended when the status of the Batch Model collection is in the **Complete** state.

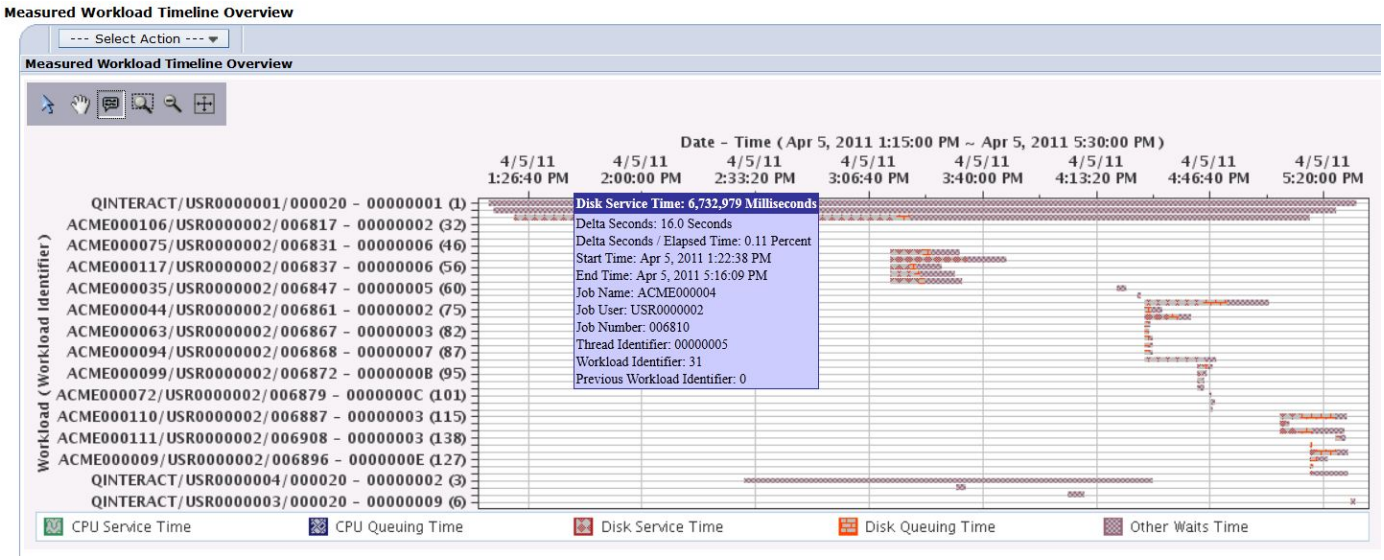
Calibrating a Batch Model

After creating your model, the next step you should take is to validate whether the model accurately modeled your measured data. This helps you determine whether you need to calibrate your batch model. Calibration is needed only when the model was unable to accurately model the measured data. The **Change Batch Model Calibration** action allows you to make corrections to the model if measured data was modeled incorrectly.

Investigate measured results

Perform the **Investigate Results** action on your recently created batch model. Investigate the data shown on the *Measured Workload Timeline Overview* chart to validate that the measured data was modeled correctly. Activate the tool tips to view the characteristics of each workload. Use the *Delta Seconds / Elapsed Time* field to validate workload runtime accuracy and use the *Previous Workload Identifier* field to validate that Batch Model correctly linked related workloads.

Figure 3. Measured Workload Timeline Overview chart



The *Measured Workload Timeline Overview* chart shows the model results that Batch Model calculated for the measured data. This chart is not showing the actual measured data from Collection Services. Rather, it is showing the resource utilization and workload runtime results

calculated by Batch Model for the measured data. The *Measured Workload Timeline Overview* chart will only change due to changes made during the **Change Batch Model Calibration** action.

The *Delta Seconds / Elapsed Time* field shows the percent difference between the measured run time of the workload as reported in Collection Services and the modeled run time of the workload as calculated by Batch Model. If *Delta Seconds / Elapsed Time* is a positive value, then Batch Model calculated a shorter run time for the workload than the actual measured run time reported in the Collection Services data. If *Delta Seconds / Elapsed Time* is a negative value, then Batch Model calculated a longer run time for the workload than the actual measured run time. If any workload has a large *Delta Seconds / Elapsed Time* percentage (positive or negative), you should investigate further to find out why these workloads were not modeled well. An easy way to view the workloads that have the greatest *Delta Seconds / Elapsed Time* is to view the *Measured Workload Timeline Overview* chart as a table and sort the data by the *Delta Seconds / Elapsed Time* column.

The *Previous Workload Identifier* field identifies the previous workload that must run before this workload. This identifier is used to build links between related workloads. Workload links are automatically created by Batch Model if you specified **Yes** on the *Link Jobs* parameter when creating your batch model. Creating links between related workloads is important when modeling. Ideally, all workloads during the batch window will be linked together so that when the processor or disk is upgraded, each workload can run faster and end sooner. This means that the next linked workload can start sooner, resulting in a shorter batch window. You should confirm that workloads were correctly linked by Batch Model, especially for the set of workloads that are your primary set of batch processing jobs. If a workload should not be linked to any other workload and should always start at the same start time, set *Previous Workload Identifier* to 0 (zero). You can also use the **Change Batch Model Calibration – Workloads** tab as another way to visually see the links that have been established between workloads (see Figure 5 below).

Change Batch Model Calibration

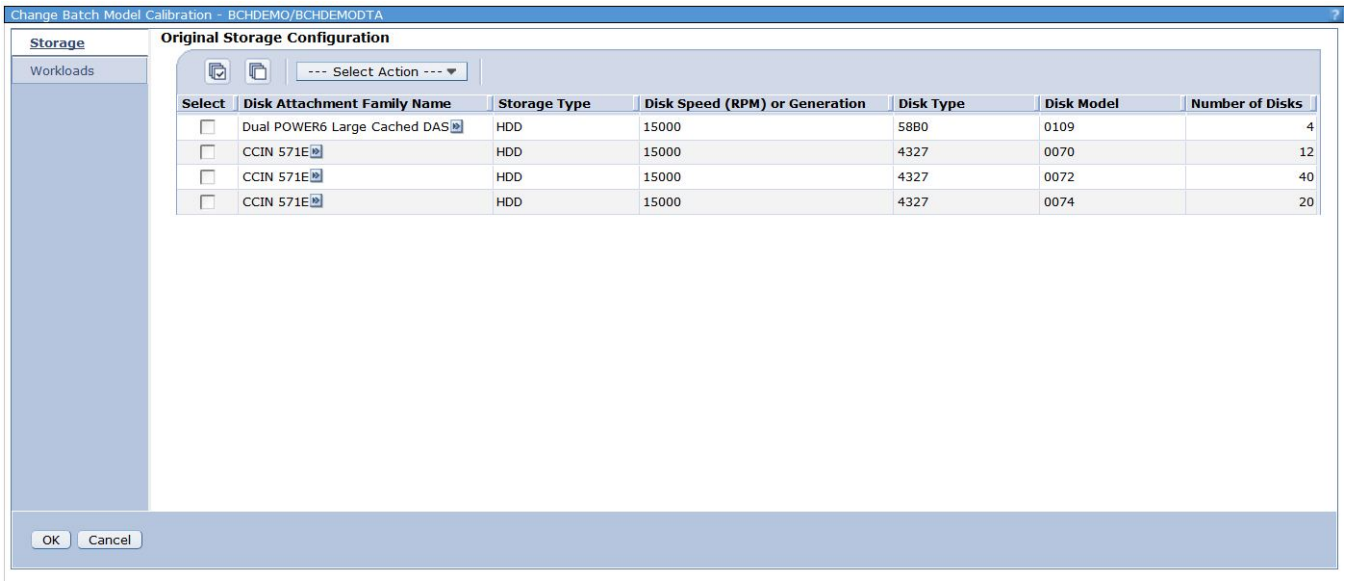
Use the **Change Batch Model Calibration** action to continue investigating calibration changes and make corrections if measured data was modeled incorrectly.

Change Batch Model Calibration – Storage

Use the Storage tab to investigate and fix the storage configuration that was incorrectly modeled. Ensure that the storage type and disk speed/generation is correctly assigned for each disk type and model. Also, be sure to set the correct disk attachment family name for any disk attachment families that were unknown to Batch Model.

Refer to Disk attachment type for more information about it.

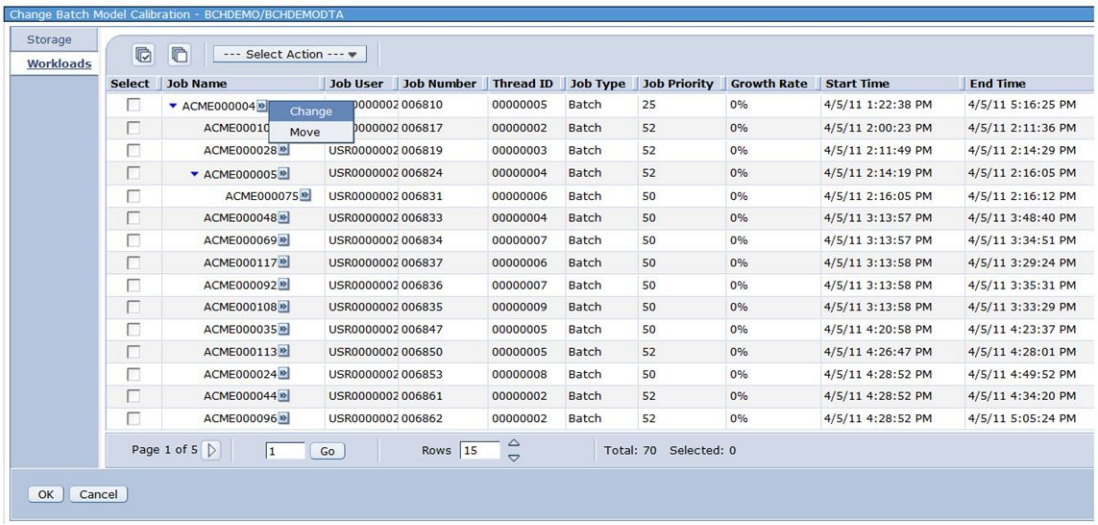
Figure 4. Change Batch Model Calibration – Storage tab



Change Batch Model Calibration – Workloads

Use the Workloads tab to investigate and fix any changes to workloads that were incorrectly modeled. You can easily see how workloads are linked, based on the indentation in the Job Name column. The workload characteristics that can be changed are: priority, job type, start time, and previous workload.

Figure 5. Change Batch Model Calibration – Workloads tab



Those workloads for which the priority changes during their run times, must be validated to ensure that Batch Model is using the priority you expect.

Interactive workloads that behave like batch workloads must be changed to job type *Batch* so that they are treated as batch workloads.

Operator batch workloads must be changed to job type *System*. Operator batch workloads are workloads that have large amounts of wait time, but not for processor or disk. Instead, they wait for operator intervention or hardware devices such as tape drives. Some will just wake up periodically and look for work. These types of workloads do not have throughput driven by the speed of the processor or disk, and therefore, they are not candidates for modeling changes with Batch Model.

Note: If you specified **No** for **Show IBM jobs** when creating your batch model, changing the job type to *System* will hide the details for the individual workload and group it into a system workload, named QSYSBCH, with other workloads of the same priority. Do not change the system type if you do not want to hide the workload details in this case.

Use *Move Workload* if a workload is not properly linked. Again, this is especially important to have correct for the set of workloads that are your primary set of batch processing jobs. If a workload should not be linked to any other workload and should always start at the same start time, set the value of *Previous Workload Identifier* to *0* (zero).

Calibrate Batch Model

After making changes to the batch model calibration, you must run the **Calibrate Batch Model** action. Calibrate re-creates the model results, taking into account the changes you made to the calibration. After making changes during the **Change Batch Model Calibration** action, you will be prompted to run **Calibrate.**

The **Calibrate Batch Model** action submits a batch job to your system to re-create the model results. Calibrating a batch model can sometimes be a long-running operation. The calibration is done when the status of the Batch Model collection is in the **Complete** state.

Change Batch Model

After creating or calibrating your model, use the **Change Batch Model** action to do a prediction. The status must be **Complete** before you are allowed to perform the **Change Batch Model** action.

Batch Model can analyze and predict the performance of batch jobs on IBM i for various *what if* conditions. For example, what will the performance of batch jobs be if you change the processor or the disk in the system? The **Change Batch Model** action allows you to change the system configuration or workloads and the **Analyze Batch Model** action then rebuilds the model and gives the new result affected by your changes.

You can change a batch model to model the following scenarios:

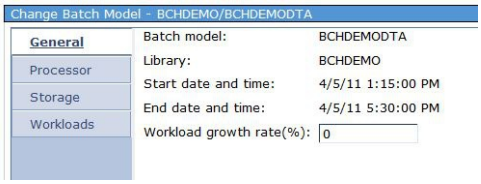
- Overall system workload growth
- Processor configuration changes
- Disk configuration changes
- Changes to workloads

Change Batch Model – General

On the General tab, you can increase or decrease the workload growth rate, which specifies the growth rate to apply to all the workloads in the Batch Model. To increase all workloads by 5%,

specify a workload growth rate of 5. To decrease all workloads by 5%, specify a workload growth rate of -5.

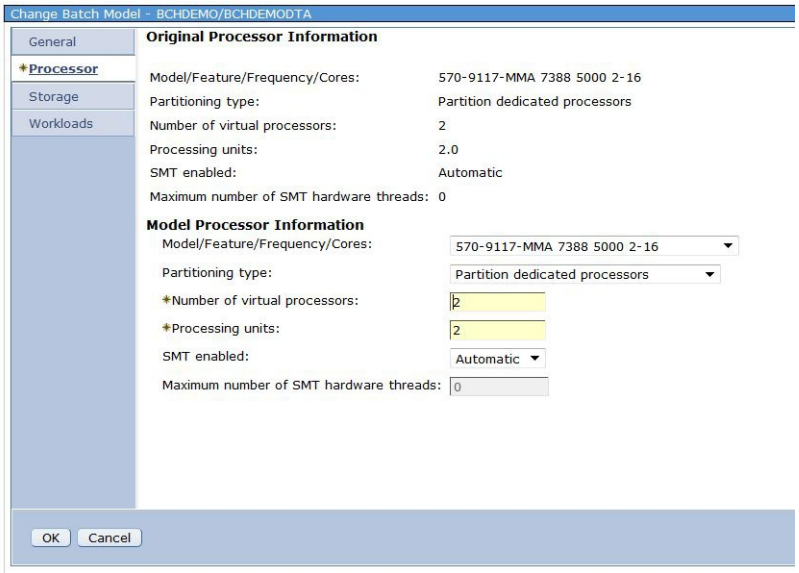
Figure 6. Change Batch Model – General tab



Change Batch Model – Processor

Processor changes are common when upgrading or changing the system configuration. Batch Model can model processor changes, including processor type and the number of processors.

Figure 7. Change Batch Model – Processor tab



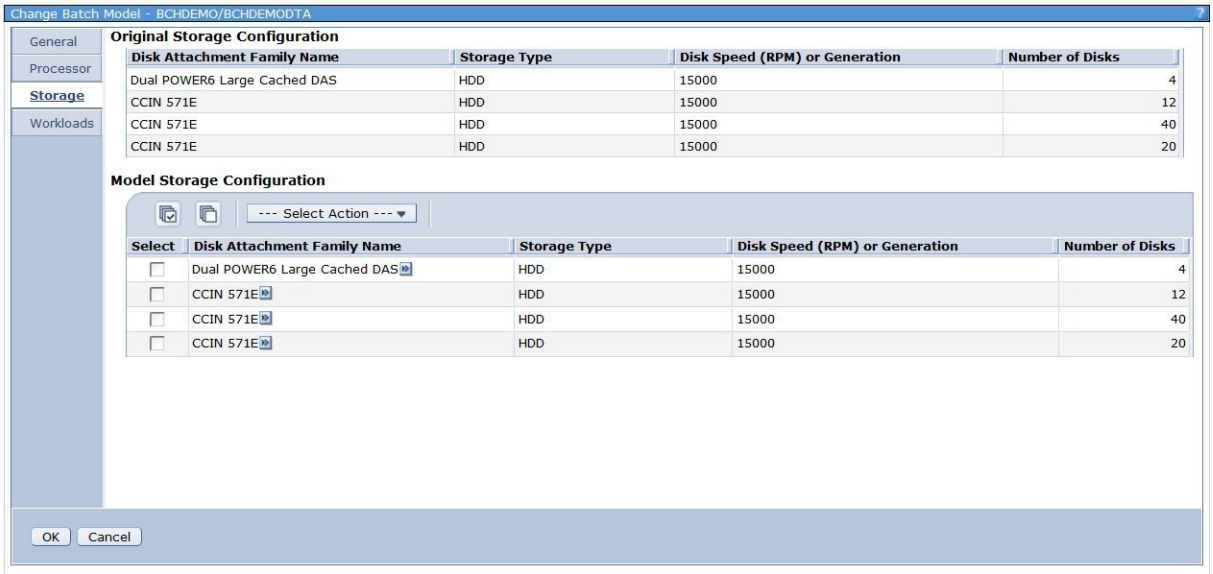
The value in the **Number of virtual processors** field should be an integer no less than the value in the **Processor units** field.

Change Batch Model – Storage

Use the Storage tab to model changes to your storage configuration. Using this tab, you can change, add, or delete disk configurations.

Refer to Disk attachment type for more information about it.

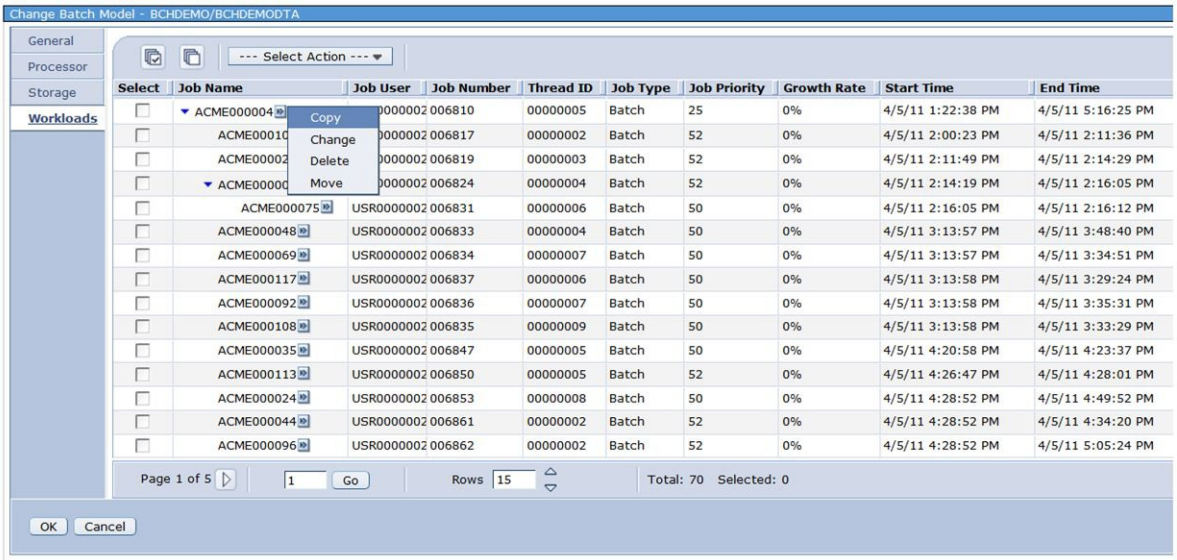
Figure 8. Change Batch Model – Storage tab



Change Batch Model – Workloads

Use the Workloads tab to copy, change, delete, and move existing workloads. You can easily see how workloads are linked, based on the indentation in the Job Name column.

Figure 9. Change Batch Model – Workloads tab

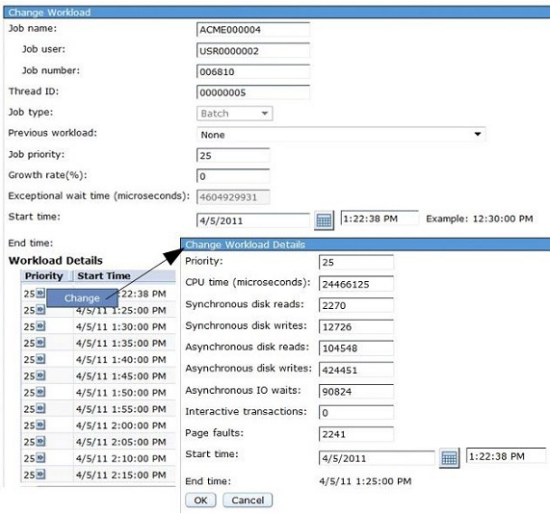


The following actions are available from the workloads tab:

- The **Copy** option is used to create a copy of an existing workload. Copying a workload can be useful when you want to model splitting a single workload into multiple concurrent running workloads. For example, to model changing a single job into two concurrent workloads, you can create a copy of the job and then change both workloads' growth rates to -50% to decrease both of them by half.

- The **Delete** option is used to delete an existing workload.
- The **Move** option is used to unlink a workload from all other workloads or change the link to another workload. This can be used to model the scenario when workload scheduling has changed.
- The **Change** option is used to change the characteristics of an existing workload. You can change the workload characteristics (see Figure 10) such as priority, growth rate, previous workload, start time, CPU time, reads/writes, I/O waits, transactions, and page faults.

Figure 10. Change Batch Model – Change Workload



Analyze Batch Model

The **Analyze Batch Model** action must be done to re-create the model results after changes were made to the batch model. After making changes during the **Change Batch Model** action, you will be prompted to run Analyze.

The **Analyze Batch Model** action submits a batch job to re-create the model results. Analyzing a batch model can sometimes be a long running operation. The analysis has successfully ended when the status of the Batch Model collection is in the **Complete** state.

Investigate Batch Model Results

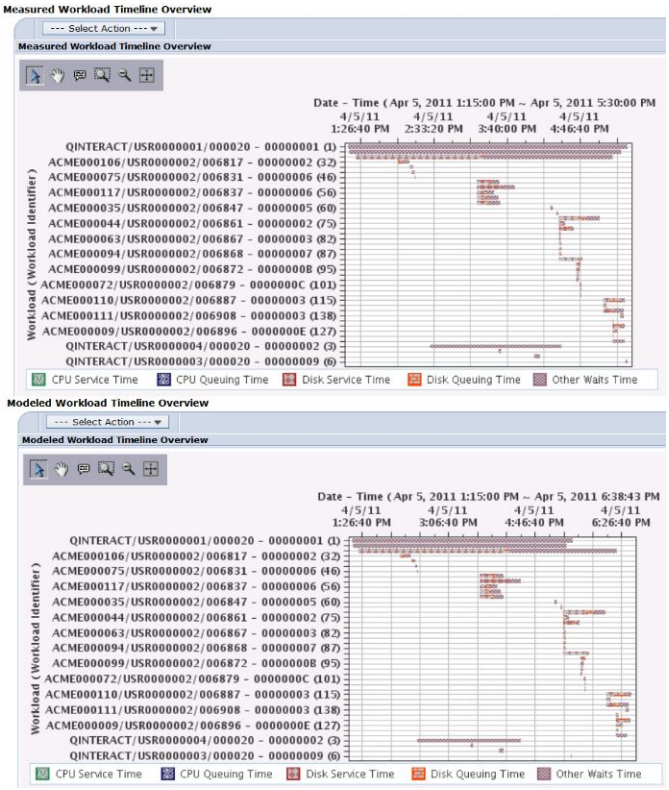
Perform the **Investigate Results** action on your recently analyzed batch model to view the batch window in a form that shows workload start/stop times, dependencies between workloads, and the amount of resources used.

Measured / Modeled Workload Timeline Overview

This perspective shows the run times for the workloads during the batch window for both measured and modeled workloads. The *Measured Workload Timeline Overview* chart shows the run times for the workloads using the measured data and measured system configuration. The *Modeled Workload Timeline Overview* chart shows the run times for the workloads using the modeled data and modeled system configuration.

The run-wait time signature is graphically displayed for each workload, breaking the workloads' run time into the portion of time spent using CPU, using disk, or waiting on something else. Activate the tooltips to view more characteristics of each workload.

Figure 11. Investigate Results – Measured / Modeled Workload Timeline Overview charts

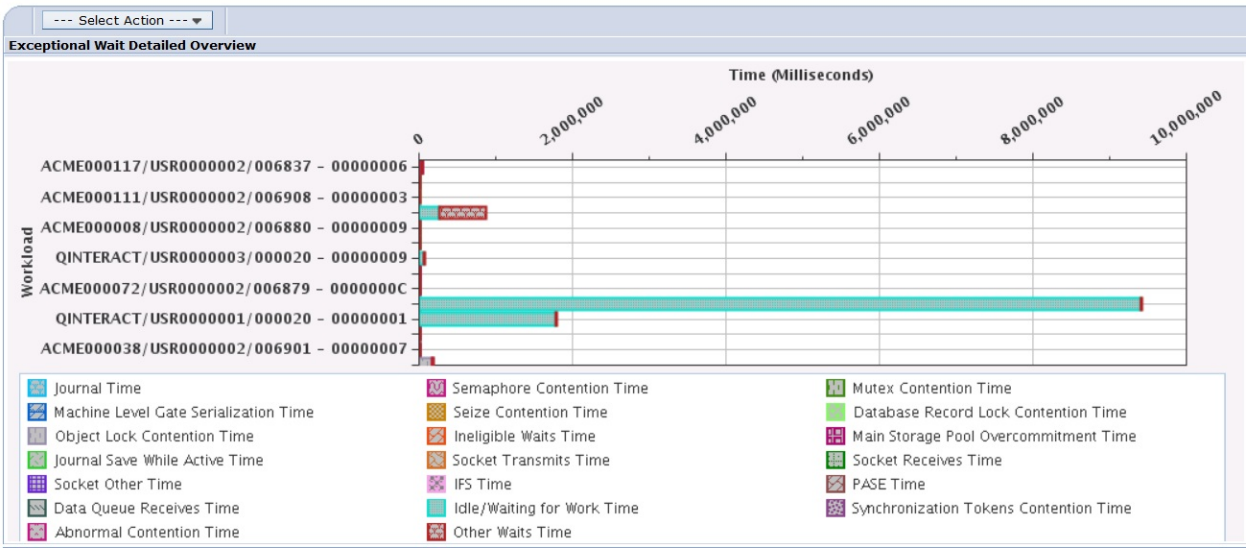


From either of the charts in Figure 11, you can drill into the *Exceptional Wait Detailed Overview* perspective or the *Measured/Modeled Resource Utilization Overview* perspective.

Exceptional Wait Detailed Overview

The Exceptional Wait Detailed Overview chart shows the time spent waiting for "other" waits that make up each workload's run-wait time signature. The "other" waits shown here are waits that are not directly related to CPU or disk.

Figure 12. Investigate Results – Exceptional Wait Detailed Overview chart

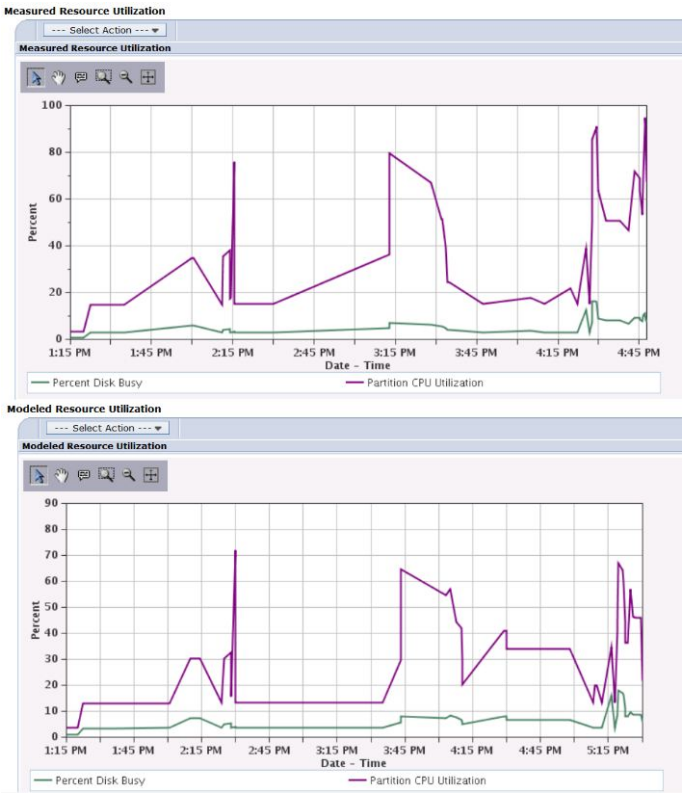


From the chart shown in Figure 12, you can drill into the *Measured / Modeled Workload Timeline Overview* perspective.

Measured / Modeled Resource Utilization Overview

This perspective shows the resource utilization percentage for both the measured and modeled workloads. The *Measured Resource Utilization* chart shows the CPU and disk utilization for the workloads using the measured data and measured system configuration. The *Modeled Resource Utilization* chart shows the CPU and disk utilization for the workloads using the modeled data and modeled system configuration.

Figure 13. Investigate Results – Measured / Modeled Resource Utilization Overview



From either of the charts shown in Figure 13, you can drill into the *Measured and Modeled Workload Timeline Overview* perspective.

Analyzing data from previous releases

Even though Batch Model is a new function in IBM i 7.2, it can be used on data collected from IBM i 6.1 or IBM i 7.1. To do so, use the SAVPFRCOL (Save Performance Collection) command on your IBM i 6.1 or IBM i 7.1 system. Move the *save file* that gets created to your IBM i 7.2 system and then run the RSTPFRCOL (Restore Performance Collection) command to restore the data. Finally, you must convert the data to the current release format by running the CVTPFRCOL (Convert Performance Collection) command.

Summary

Batch Model is a useful, new tool in IBM i 7.2 under the performance task in IBM Navigator for i. It enables you to create various *what if* scenarios where you predict resource usage based on changes to processors and disks, and the growth of your workload. In addition, it can estimate the size of your batch window and allow you to schedule your batch activity more efficiently.

Resources

The following references provide more information about Batch Model and IBM i Performance Tools.

- IBM Knowledge Center - IBM i Batch Model
- developerWorks - IBM i Performance Tools technology updates
- developerWorks - IBM i Performance Tools forum