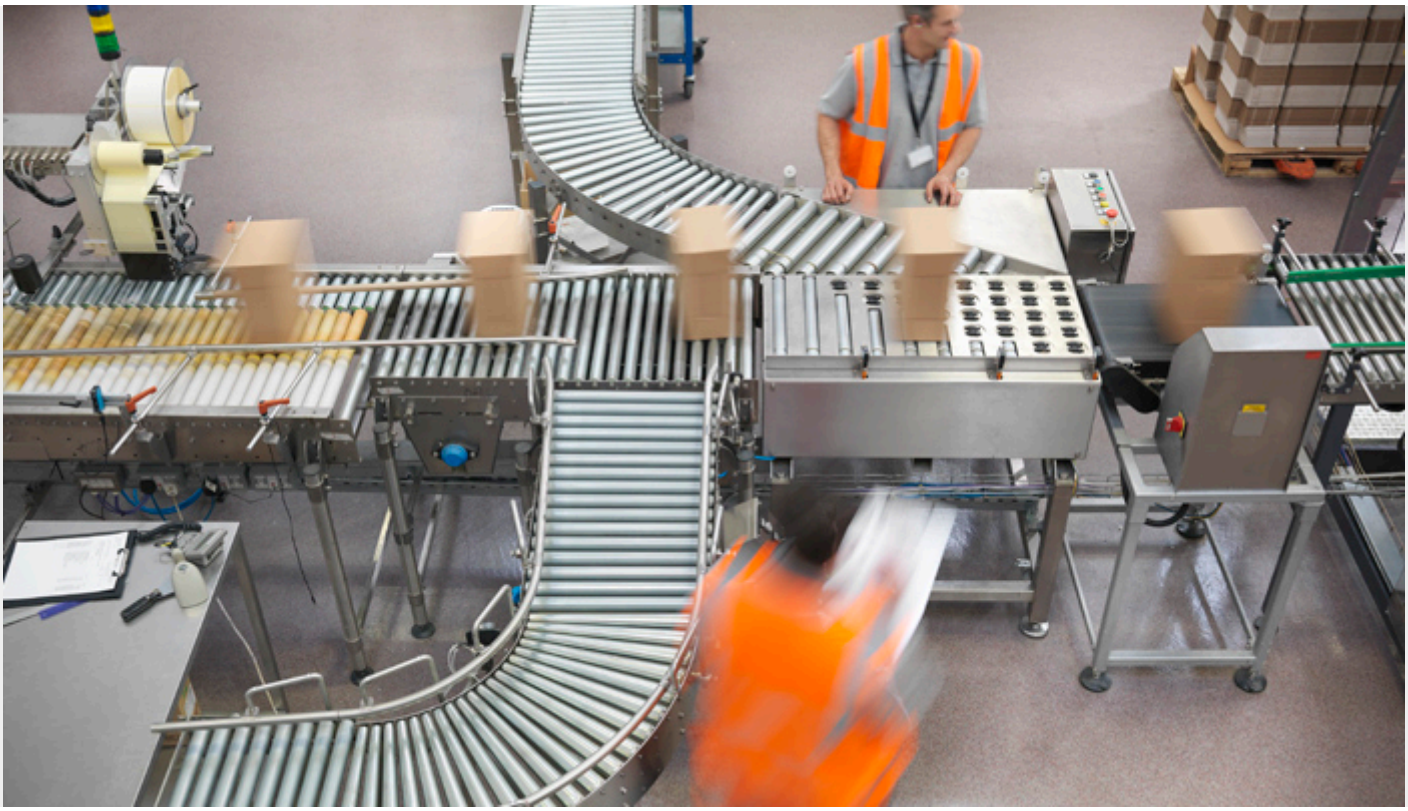
# BATCH PROCESSING IN 2020: AN INTRODUCTION



Put simply, batch processing is the process by which a computer completes batches of jobs, often simultaneously, in non-stop, sequential order. It's also a command that ensures large jobs are computed in small parts for efficiency during the debugging process.

This command goes by many names, including Workload Automation (WLA) and Job Scheduling. Like most things in programming, it's changed over time. Luckily, these changes have made batch job processing more sophisticated and efficient. For many businesses, it's a necessary component to their daily success.

In this article, we'll provide an overview of batch processing, look at use cases as well as benefits and drawbacks, and then offer guidance on determining whether batch processing is right for you.

## The evolution of batch processing

Today, a defining characteristic of batch processing is its lack of user interaction. There are few, if any, manual processes to kick it off. This is part of what makes it so successful and efficient, but that was not always the case. Batch processing began with the use of punch cards that were tabulated to tell computers what to do. Often decks, or batches, of cards would be processed at one time. This practice goes back to 1890 when Herman Hollerith created punch cards to process Census data. Working for the U.S. Census Bureau, he developed a system by which a card that he punched manually was read by an electromechanical device. Hollerith soon formed a small company that we know today as IBM.

The punch card revolutionized the way businesses operated—but that was then. Over the past two decades, batch processing has continued its evolution. Data entry professionals are no longer necessary. Most batch processing functions are enabled without interaction, and they are completed to meet specified timing needs. Some jobs are completed in real time with daily monitoring and reporting features, others are done immediately.

## Dependencies and monitors in batch processing

Today's batch processing uses exception-based management alerts to notify the right people if there are issues. This allows managers the freedom to work without regularly checking in on the progress of batches. The idea is that managers need not check in at all unless they receive an alert about a critical exception.

Exceptions are determined by a system of dependencies and monitors that are essential to the software:

- **Dependencies** are the events that trigger the batch processing to begin. This could be that a customer places an online order, or a user requested new supplies, triggering the system to generate a request.. The dependency has set the batch processing in motion.
- **Monitors** look for abnormalities in the batch. Perhaps, one job is taking longer than normal to complete. The next job can't begin until the one before it ends. If this is causing an unusual delay, the monitor will catch it, generating an exception and sending it to a manager.

## When to use batch processing

You might expect latencies when using batch processing. For many situations, however, this type of delay before the transfer of data begins is not a big issue—the processes that use this function are not mission critical at that exact moment.

Batch processing should be considered in situations when:

- Real-time transfers and results are not crucial
- Large volumes of data need to be processed
- Data is accessed in batches as opposed to in streams
- Complex algorithms must have access to the entire batch
- Tables in relational databases need to be joined
- The work is repetitive

## Advantages of batch processing

There are a number of reasons why companies deploy Batch Processing systems. Business owners should look at the total impact when deciding on new software for their organization.

### Speed, at lower costs

Because batch processing does not need Data Entry clerks to support its functioning, it helps to reduce the operational costs that businesses spend on labor. It also doesn't require any additional hardware to function, outside of a computer.

In fact, using batch processing can reduce a company's reliance on other expensive pieces of

hardware, making it a relatively inexpensive solution that helps businesses save money and time. Without the possibility of user error, batch processes are completed in the most efficient way possible. The result is fast, accurate processing and managers who have more time to spend on day-to-day operations.

### Offline features

Batch processing systems work offline, so when the workday ends, batch systems are still processing in the background, chugging along. This gives managers ultimate control over when to start processes.

Software can be set for overnight processing of certain batches, which provides a convenient solution for businesses that don't want a job like automatic downloads to disrupt daily activities.

### Easy hands-off management

Managers have enough to do without logging in every hour to check on their batches. The exception-based notification system of modern batch processing software makes it easy for managers to do their job without worrying about whether their software is functioning properly and if batches are being completed. If there's an issue, notifications are sent to the right people to solve it. Managers can take a hands-off approach, trusting that their batch processing software is doing its job.

### Simplicity

Compared to real-time processing or stream processing, batch recessing is significantly less complex: it does not require constant system support for data input or unique hardware. Once installed and established, a batch system doesn't involve heavy duty maintenance, making it a relatively low-barrier-to-entry solution.

## Disadvantages of batch processing

While batch processing is great for many scenarios, owners should also consider these components, which might be disadvantages for certain companies.

### Deployment and training

As with any new technology, there's some degree of training involved in managing these systems. Managers who are not familiar will need to understand what triggers a batch, how to schedule them, and what exception notifications mean, among other things.

### Complex debugging

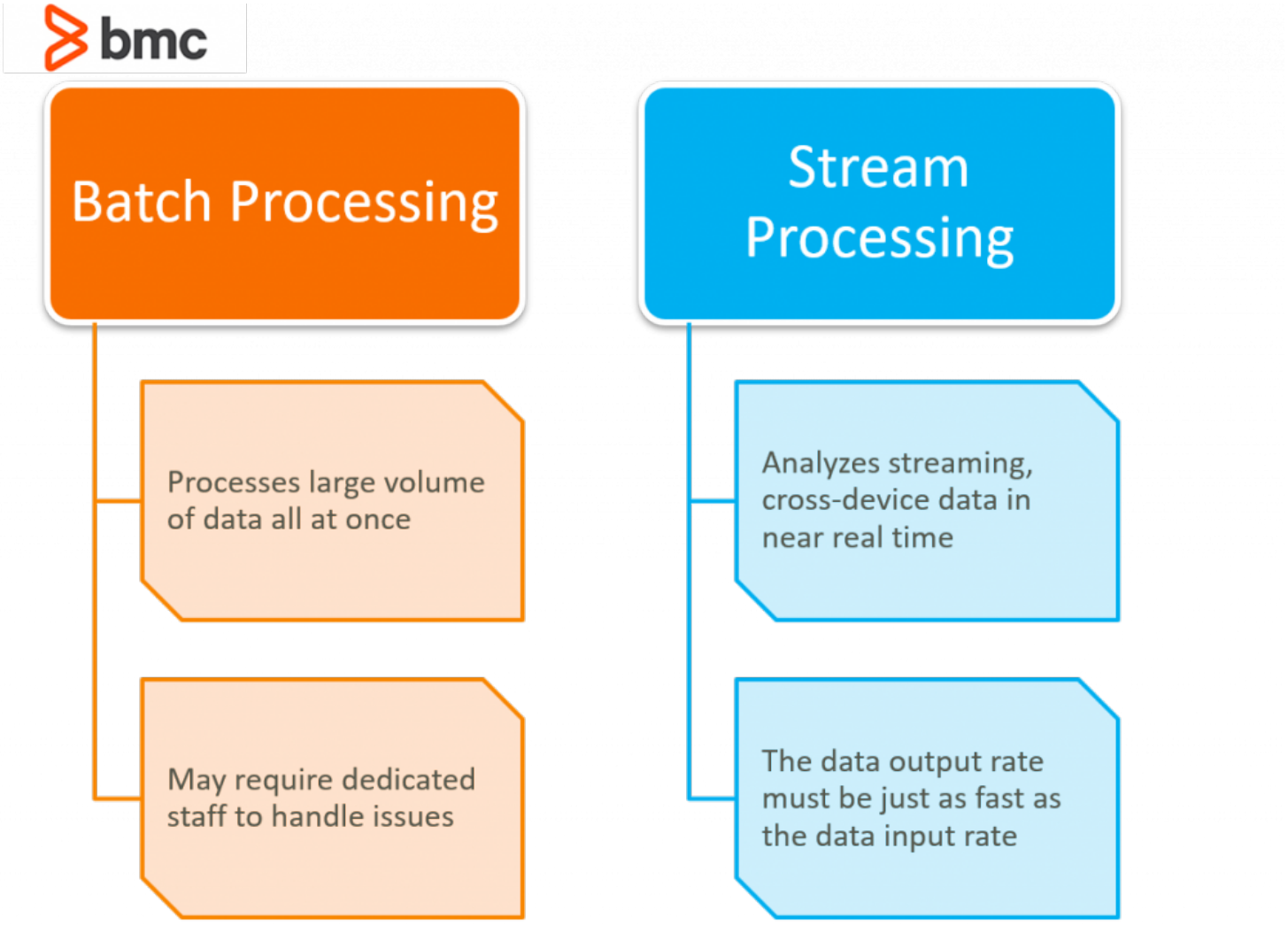
When an error occurs, managers also need to know how to fix it. Debugging batch processing systems can be understandably complex. You'll likely need an in-house employee who specializes in these systems; otherwise, expect an additional cost whenever you need an outside consultant to assist.

## Cost

While most businesses save money on labor and hardware when they switch to batching, some businesses don't have data entry clerks or expensive hardware needs to begin with. For these businesses, the cost of some batching systems may seem unfeasible.

# Batch Processing vs Stream Processing

Given the benefits of both, many organizations are facing the dilemma of which is better: batch processing or stream processing? Although a clear-cut answer might be ideal, there is no single option that is the perfect solution for every instance, rather the optimal method varies depending on needs, the company, and the specific situation.




*Benefits and drawbacks of common data processing types*

With its ability to process large volumes of data all at once, even millions of records that are stored and saved up throughout the day, Batch Processing is a good option for things that can wait until after-office hours. Examples are things like:

- All line item invoices
- Payroll processes
- All the single transactions a financial firm might submit over the period of any given week

Stream processing, on the other hand, is useful for processes or systems that depend on having

access to data in real-time. By being able to almost instantaneously analyze streaming data, this method is better suited for scenarios in which events are happening frequently and close together, and cannot wait until evenings for computers to be idle. This type of processing is best used for tasks like cybersecurity and fraud detection as these need to be found right away for risks to be mitigated.

Many organizations have found that a combination of batch and stream processing is the most beneficial for workflows, each method being utilized as needed.

## Do I need batch processing?

If you're still wondering if batching is the right solution for you, think about examples of where you might use this function in your business. Are there gaps that you could fill with automation?

To name a few, these might include:

- Payroll Processes and Employee Time Sheets
- Bank Statements
- Line Item Invoicing
- Stream, Analytics and Transaction Processing
- Research and Reporting
- Email Systems
- Supply Chain and Fulfillment
- Other Requests

As a rule of thumb, if you find yourself regularly doing large computing jobs manually, there's a good chance the right Batch Processing software could be the key to freeing up more time and money for your organization.

There are also a few questions to consider when deciding if your company needs Batch Processing:

- How do you ensure that manual jobs are done correctly? Do you have a system in place to determine that they have been submitted and processed in the correct order?
- Do you have jobs waiting to start that are contingent on another completing? Do you have a system in place to track each job to completion or that knows when the first job will be done?
- Do you manually check for new files? Does your script loop frequently enough to be efficient when checking for files?
- Do you have a current setup with job-level retries on your server? What impact does it have? Would you benefit from less stress being put on the server?
- How do you track dependencies across servers? How do you know the dependent server will be available when needed?

## Automate job scheduling for your company

By design, modern systems aim to free up managers from daily minutiae, prevent user error and promote fast, efficient completion of jobs with little oversight. Batch Processing has the added effect of saving money for nearly all the businesses that deploy it.

It isn't without its considerations, though, as noted above. Companies without the IT staff to implement a successful deployment and maintenance plan can still benefit from batch processing,

but they should prepare on the front end to invest time and money in partnering with IT experts to ensure a successful launch and end user experience.

## BMC for workload automation

For more on this topic, see our BMC Workload Automation Blog. Ready to see what BMC can do for your company? Explore workload automation, including batch processing, with Control-M.