

ORCHESTRATE AND AUTOMATE TO MAKE DATAOPS SUCCESSFUL



DataOps is intended to smooth the path to becoming a data-driven enterprise, but some roadblocks remain. This year, according to a new IDC InfoBrief sponsored by BMC, DataOps professionals reported that on average, only 58 percent of the data they need to support analytics and decision making is available. How much better would decision-making be, and how much business value would be created, if the other 42 percent of the data could be factored into decisions as intended? It seems logical to assume it would be almost twice as good!

That raises another question: Why can't organizations get the data that they already have where they need it, when they need it? In most cases, the answer comes down to complexity.

A [previous blog](#) by my colleague, Basil Faruqi, introduced why DataOps is important. This one follows up to highlight what is needed. Spoiler alert: The ability to orchestrate multiple data inputs and outputs is a key requirement.

The need to manage data isn't new, but the challenges of managing data today to meet business needs is changing very fast. Organizations now rely on more data sources than ever before, along with the technology infrastructure to acquire, process, analyze, communicate, and store the data. The complexity of creating, managing, and quality-assuring a single workload increases exponentially as more data sources, data consumers (both applications and people), and destinations (cloud, on-premises, mobile devices, and other endpoints, etc.) are included

DataOps is helping manage these pathways, but is also proving to have some limitations. The IDC InfoBrief found integration complexity is the leading obstacle to operationalizing and scaling DataOps and data pipeline orchestration. Other obstacles include a lack of internal skills and time to

solve data orchestration challenges, and difficulty using the available tooling. That means that for complex workloads like those shown above, organizations can't fully automate the planning, scheduling, execution, and monitoring because the complexity causes gaps, which in turn cause delays. This results in decisions being made based on incomplete or stale data, thus limiting business value and hampering efforts to becoming a data-driven enterprise.

Complexity is a big problem. It is also a solvable one. Orchestration, and more specifically, automating orchestration, are essential to reducing complexity and enabling scalability, unlike scripting and other workarounds. Visibility into processes, self-healing capabilities, and user-friendly tools also make complexity manageable. As IDC notes in its InfoBrief, "Using a consistent orchestration platform across applications, analytics, and data pipelines speeds end-to-end business process execution and improves time to completion."

Some of the most important functionality that is needed to achieve orchestration includes:

- Built in connectors and/or integration support for a wide range of data sources and environments
- Support for an as-code approach so automation can be embedded into the deployment pipelines
- Complete workflow visibility across a highly diverse technology stack
- Native ability to identify problems and remediate them when things go wrong

Tooling that is specific to a software product, development environment, or hyperscale platform may provide some of that functionality, but typically isn't comprehensive enough to cover all the systems and sources the workflow will touch. That's one reason so many DataOps professionals report that tooling complexity hinders their efforts.

[Control-M](#) can simplify DataOps because it works across and automates all elements of the data pipeline, including extract, transform, load (ETL), file transfer, and downstream workflows. Control-M is also a great asset for DataOps orchestration because:

- It eliminates the need to use multiple file transfer systems and schedulers.
- It automatically manages dependencies across sources and systems and provides automatic quality checks and notifications, which prevents delays from turning into major logjams and job failures further downstream.

Here are a couple quotes from Control-M users that illustrate its value. A professional at a healthcare company said, "Control-M has also helped to make it easier to create, integrate, and automate data pipelines across on-premises and cloud technologies. It's due to the ability to orchestrate between workflows that are running in the cloud and workflows that are running on-prem. It gives us the ability to have end-to-end workflows, no matter where they're running."

Another user, Railinc, said, "The order in which we bring in data and integrate it is key. If we had to orchestrate the interdependencies without a tool like Control-M, we would have to do a lot of custom work, a lot of managing. Control-M makes sure that the applications have all the data they need." You can see the full case study [here](#).

These customers are among the many organizations that have reduced the complexity of their DataOps through automation. The IDC InfoBrief compares enterprises that excel at DataOps orchestration to those that don't and found advantages for the leaders in multiple areas, including compliance, faster decision-making and time-to-innovation, cost savings, and more.

Can you orchestrate similar results at your organization? Learn more about Control-M for Data Pipeline Orchestration [here](#).