StreamSets
A SOFTWARE AG COMPANY

How do you enable innovation, prototyping, and experimentation with centralized guardrails?

The challenge:

Fragmented data supply chains are hard to understand, govern, and manage

The diversity of modern infrastructures that span on-premises and multiple cloud environments, combined with decentralization that puts analytics tools and projects in the hands of business users, has created a data "wild west."

However, while this increases agility within individual pockets of the organization, it also creates challenges and risks for the enterprise as a whole. The sheer complexity of connecting varied data sources and ensuring consistency of definitions makes deployments and subsequent updates an arduous process. It's difficult to enforce consistent security measures to protect data as it flows between on-premises and cloud sources.

And regulatory requirements that restrict access to certain data can be inadvertently violated by users who are not compliance experts.

As more and more apps, systems, data sources, and tools get added to the ecosystem, it becomes much more difficult to maintain visibility and control. All this creates a management nightmare that ultimately hampers innovation and puts you at risk of breaches and fines.

But what if you could **enable innovation, prototyping, and experimentation with centralized guardrails?**

Here's how you'll make it possible...

Bridge the gap between new and legacy environments easily and securely

Enterprises have complex infrastructures comprising on-premises, public cloud, and private cloud environments.

Many organizations are subject to different regulatory requirements in different geographies, requiring different cloud providers, and as a result, different pipeline builds. Each environment has its own deployment requirements and challenges, and it can be difficult and time-consuming to move an application from on-premises to the cloud, for example, or to enable a cloud-based application to securely leverage data from a different environment.
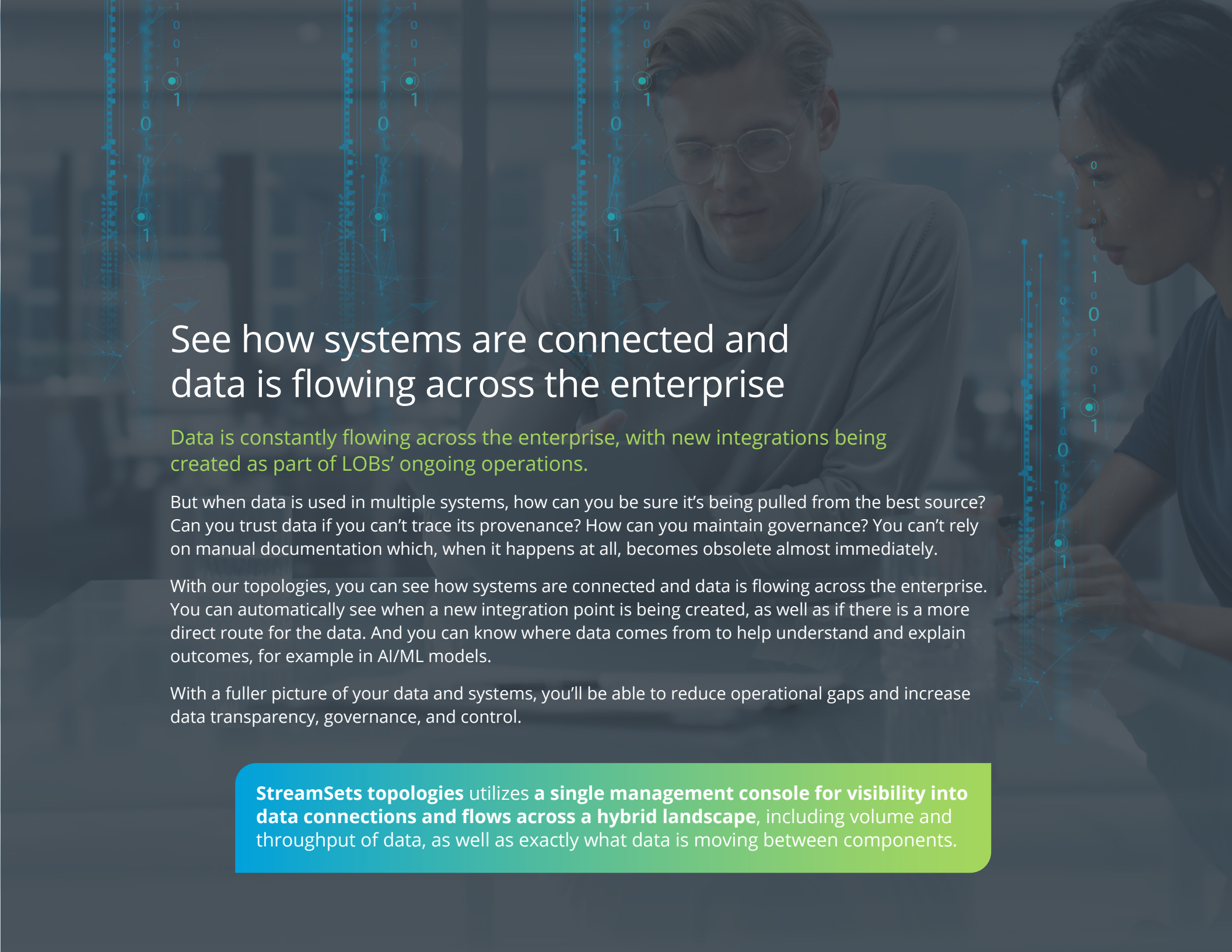
With StreamSets' hybrid deployment with centralized engine management, you can bridge the gap between new and legacy environments easily and securely. With a data "mission control" across all your environments, you can easily move between clouds and on-premises.

That means you'll extract maximum value from your data faster while lowering the cost and effort of managing your disparate toolchains.
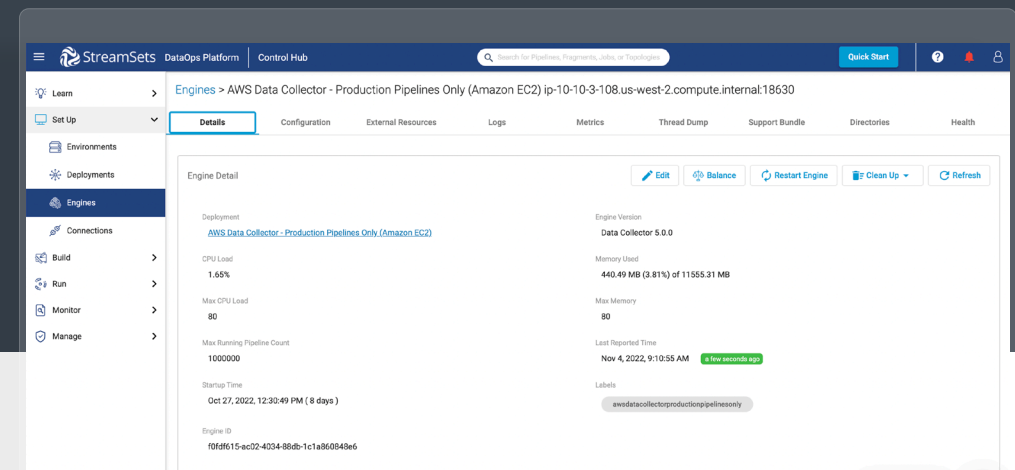
> Through StreamSets' hybrid deployment with centralized engine management, you'll be able to use a **single management console** for the secure deployment of data processing components **in any environment**.

# See how systems are connected and data is flowing across the enterprise

Data is constantly flowing across the enterprise, with new integrations being created as part of LOBs' ongoing operations.

But when data is used in multiple systems, how can you be sure it's being pulled from the best source? Can you trust data if you can't trace its provenance? How can you maintain governance? You can't rely on manual documentation which, when it happens at all, becomes obsolete almost immediately.

With our topologies, you can see how systems are connected and data is flowing across the enterprise. You can automatically see when a new integration point is being created, as well as if there is a more direct route for the data. And you can know where data comes from to help understand and explain outcomes, for example in AI/ML models.

With a fuller picture of your data and systems, you'll be able to reduce operational gaps and increase data transparency, governance, and control.

StreamSets topologies utilizes **a single management console for visibility into data connections and flows across a hybrid landscape**, including volume and throughput of data, as well as exactly what data is moving between components.
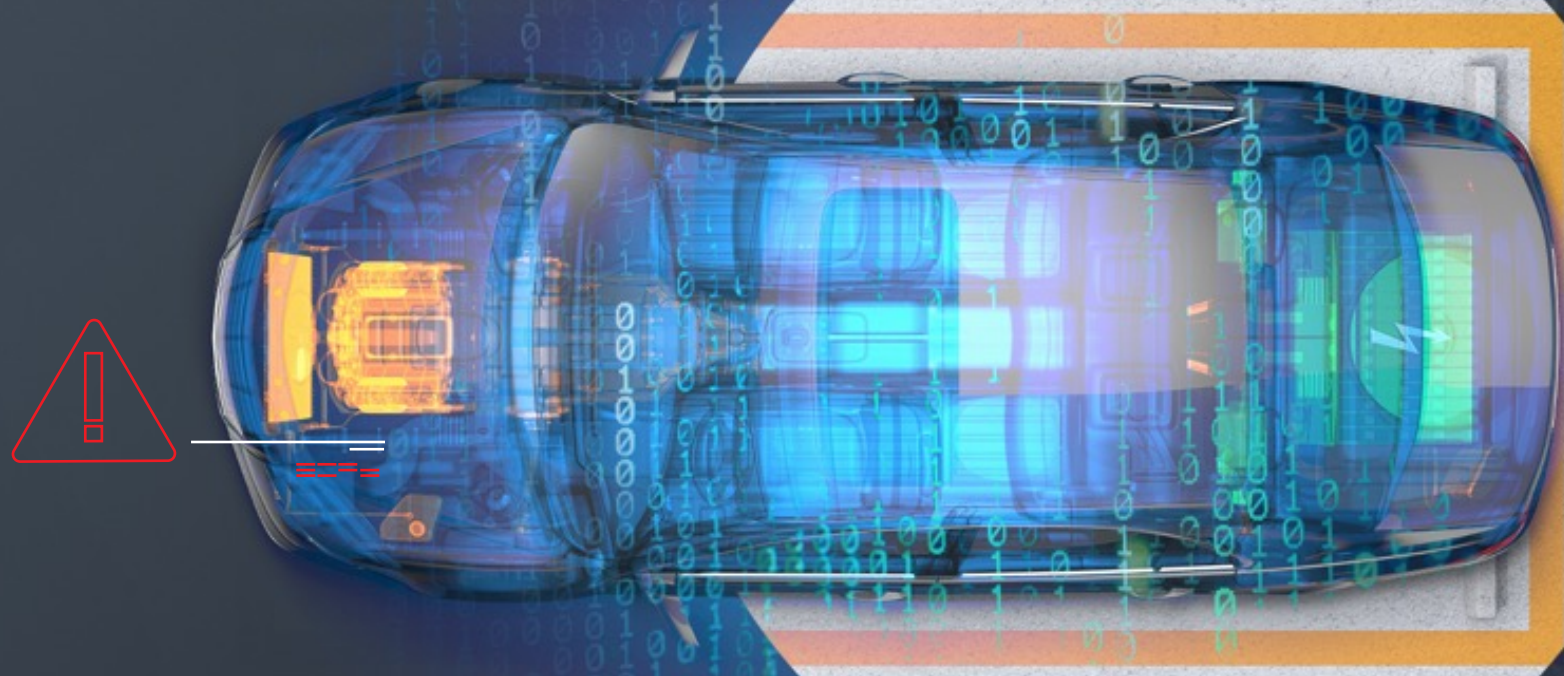
Expose hidden problems in your data flows

Performance, quality, security, and compliance are all critical requirements you need to meet, but different stakeholders prioritize these to varying degrees.

For example, engineers are focused on speed while someone in governance is concerned with control and compliance. It's hard to ensure that all of these considerations are appropriately maintained throughout all of your data pipelines.

With our data SLAs and rules, you can expose hidden problems in your data flows. You can create guardrails and quality checks, and then manage by exception. And with issues more clearly defined, that means you'll address performance, data leakage, quality, and other risks before they affect the business.

StreamSets' data SLAs and rules enable **auto-notification of problems based on user-defined triggers** throughout data pipelines with **alerts that can be sent via email, Slack, or through other system messages**.

Eliminate superfluous activities

While taking a "wild west" mentality toward decentralized data can contribute to agility in some parts of your business, it can also create significant risk. Connecting varied data sources is complex—and it's difficult to enforce consistent security measures as data flows between on-premises and cloud sources. Meanwhile, the more systems and users you add to your data ecosystem, the greater your risk of security breaches and costly fines.

With StreamSets you'll be able to...
Bridge the gap between new and legacy environments easily and securely, get a clearer picture of how systems and data are flowing across the enterprise, and expose hidden problems in your data flows.

And that means you'll...

Extract maximum value from your data faster, while lowering the cost and effort of managing your disparate toolchains.

Reduce operational gaps and increase data transparency, governance, and control.

Address performance, data leakage, quality, and other risks before they affect the business.

To learn more,
visit us online at www.streamsets.com.

StreamSets
A SOFTWARE AG COMPANY