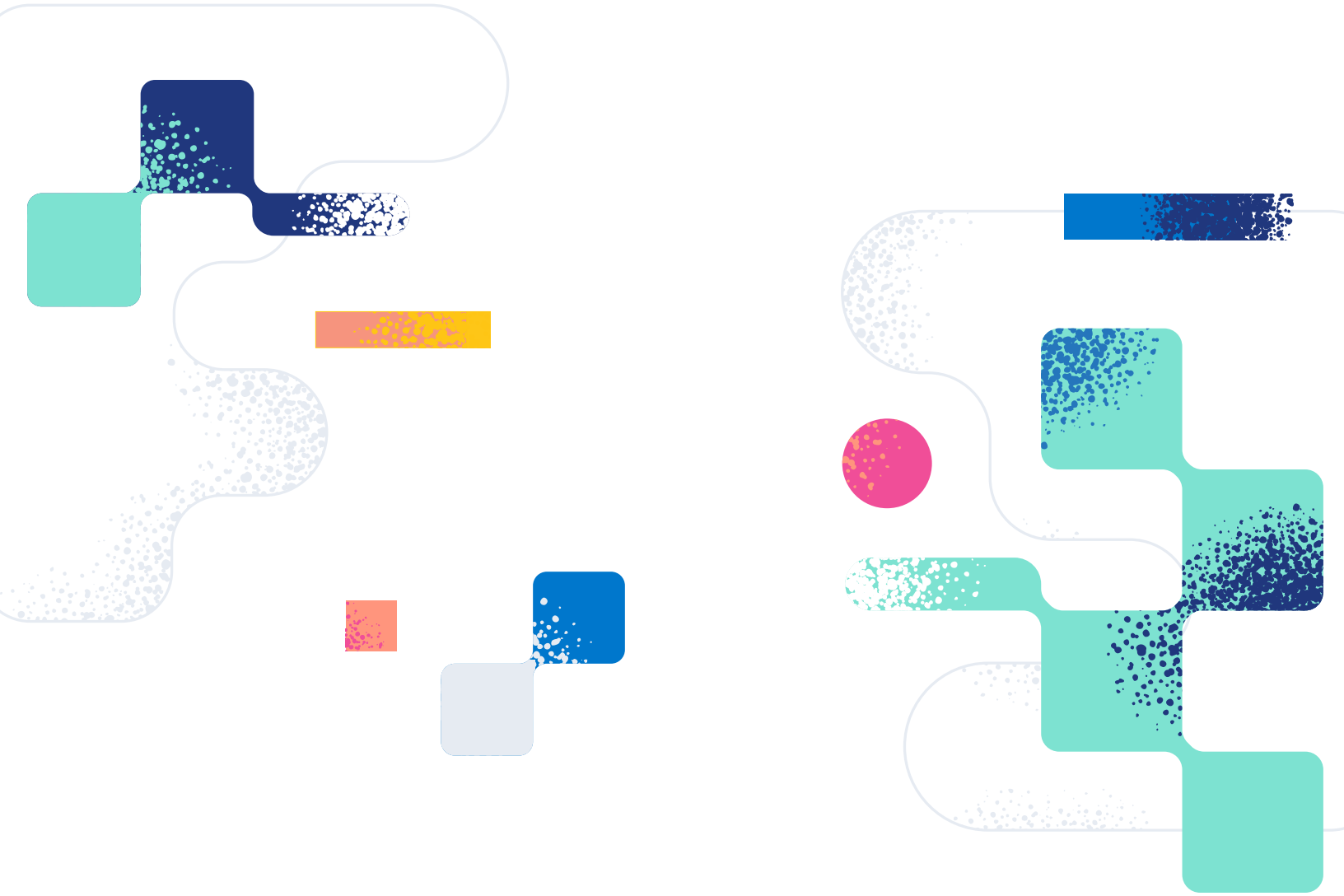
elastic



Simplifying and accelerating defence missions with a global data mesh

Table of Contents

Introduction

Outthinking and outfighting adversaries in an information-enabled battlefield calls for superior exploitation of data. Accessed, applied, and secured from wherever it resides, data becomes a force multiplier — delivering speed, agility, and actionable intelligence for military forces while preventing corruption and exfiltration by adversaries.

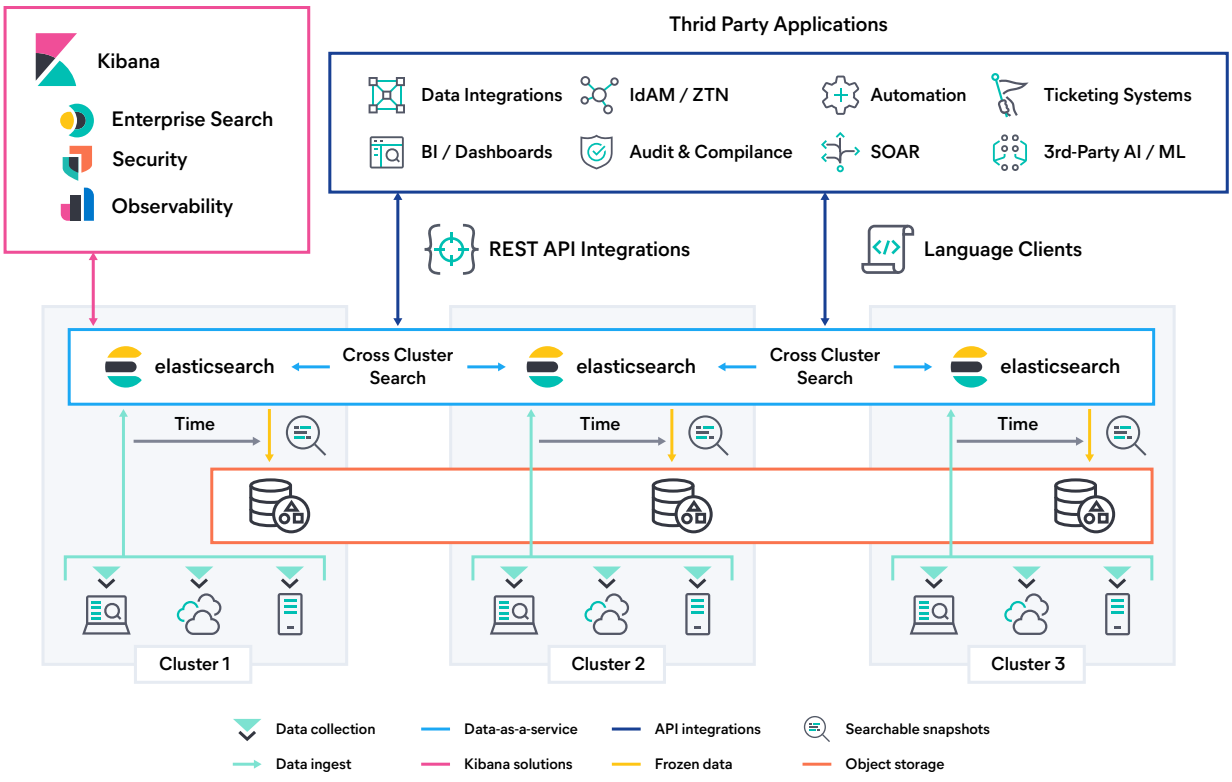
The sheer amount of data used and produced by military forces today poses a significant challenge when teams need to find and access specific information quickly. That challenge is compounded by the fact that data is most often stored in different formats (e.g., images vs. documents, vs. maps) and environments (e.g, cloud, on-prem). Not only that, but there are also significant variations in where data is collected, ranging from ground sensors used by soldiers to data gathered by unmanned aerial vehicles (UAVs), each presenting unique data sources. Moreover, unconventional data storage environments like floating/submersible data centres on aircraft carriers, submarines, combat aircraft and satellites require specialised approaches to data management. This inevitably leads to information silos, lack of interoperability, and generally complex environments.

Without a way to correlate and analyse all of the available data, the insights extracted from that data will inherently be incomplete and possibly inaccurate. But copying and moving terabytes — or even petabytes — of data to a central location for access and analysis is time- and bandwidth- consuming, and leads to delays and version control issues. The end result is that it's harder than ever to isolate the insights from the troves of available information.

Elastic solves this data interoperability challenge by serving as a unifying "data mesh" layer that breaks down silos while allowing data to remain in its original, decentralised location at the edge. Ultimately, leveraging and harnessing data as an enduring strategic asset can drive sustainable battlespace advantage and enhance military efficiency. Armed forces can unlock the full potential of their data resources by overcoming the challenges associated with diverse data formats and storage environments and implementing robust data management strategies. This enables informed decision-making, improved situational awareness, optimised resource allocation, and streamlined operations.

What is a data mesh?

Data mesh is an approach where data is managed in a distributed model, rather than a centralised model such as in a single central data lake. An effective data mesh layer for defence should provide the ability to not just collect, but also to fuse, a growing amount of data sets with minimal latency. It should also be able to work on the tactical edge, no matter the scope and speed of operations or how austere the environment. Put another way, a global data mesh is a high-quality, feature-rich approach that includes all the capabilities a defence department would need to run at enterprise scale – including functionality built in that you may not even realise you need at the outset of implementation.



The diagram above depicts multiple distributed clusters collecting local data streams and making them available (both locally and through cross-cluster search) to any application or use case they may be useful for, including within the many prebuilt Kibana solutions. As data ages out of the real-time data access patterns, it can be moved as searchable snapshots to object storage and made available to searches on demand.

What are the benefits of a data mesh approach?

The data mesh approach is valuable to defence agencies for a variety for reasons, including:

1

Ownership

Instead of relinquishing control to a centralised team or industry partners, in a data mesh model, the people who are responsible for and understand the nuances of the specific data and its applicable domain and functional area are in control of its use and distribution.

2

Data democratisation

By implementing a data mesh, defense departments can democratise access to data within their teams, enabling faster, more-informed decision-making processes.

3

Accuracy

Information remains at the source without needing to be correlated or translated, which promotes proactive data cleansing, validation, and monitoring practices, resulting in improved data accuracy and trustworthiness. Accessing data in-situ is also inherently faster, which makes data more relevant, accurate, and valuable.

4

Flexibility

Individual teams and cross-functional areas have the flexibility to use, collaborate, and manoeuvre their data based on their specific needs.

5

Comprehensive view

A data mesh layer unifies disparate data and provides a single source of truth that reduces reliance on incomplete viewpoints from disconnected systems and teams.

6

Simplified planning and training

A data mesh approach empowers various departments and units to take ownership of data. This decentralisation promotes data accountability and encourages subject matter experts to manage and govern their data effectively. And having a unified platform to acquire, learn and deploy simplifies planning, reduces training time and bridges the data aptitude gap.

Integrating legacy IT

When pulling together data sources from different military branches, data sources and enterprise architectures can vary significantly. There can be a number of accredited and multi-classification yet legacy systems involved, which represents an edge to cloud challenge. Nonetheless, mission speed cannot be sacrificed as interoperability is sorted out or legacy systems are phased out.

A data mesh can be layered on top of existing enterprise architectures so that structured, semi-structured, and unstructured data is pulled into the more modern environment at blazing fast speed. Having the ability to ingest, visualise, and analyse a wide variety of data – emails, open-source intelligence, instant messages, geo-temporal coordinates, IoT logs, imagery, and more – gives stakeholders the flexibility to use all types of data to make more informed decisions. In addition, developers can use this data along with APIs to build new parallel applications that will eventually replace the legacy applications.

Scalability is also key. More data can be pulled into the data speed layer over time, and as more applications are built on top of it, decommissioning the legacy environment and moving toward a fully modernised infrastructure becomes easier. The legacy system still exists, doing its day-to-day job. The goal is not to replace an entire system in a single go; instead, the aim is to make data available in an extremely fast, secure data speed layer and then have developers build applications and eventually shut down the legacy systems.

What differentiates Elastic's data mesh capabilities?

Distributed search & Cross-Cluster Search

Elastic uses a distributed search approach to power its data mesh. With Elastic's Cross-Cluster Search functionality, data can be globally exploitable yet stored at the edge, for a frictionless viewpoint in that value or kill chain. Within seconds of data being ingested from sensors and systems, Elastic can normalise and index all data in optimised ways to allow for extremely fast query analytics that don't shy away from the native characteristic of the data. This means that data is normalised for unified querying and analytics.

Unlike other solutions, this approach to data indexing can leverage both "Schema on Write," which provides both scale and performance, as well as "Schema on Read," which offers flexibility and speed to value.

Users can use a single search query to find and analyse data stored across clusters, which can be in different data centres and/or clouds. The data resides in its compliant environment but is queried as remote clusters.

This means that all relevant data across formats, including time, space, geography, compliance level or other attributes can be analysed in seconds. As a result, data is assured, discoverable, and interoperable – and can be used as an enduring asset beyond siloed programs.

These queries can also be re-used for additional operational efficiency. In this way, Elastic helps users bring questions to the data, even if silos exist — enabling compliant inter-departmental data sharing through the power of search.

Cross-Cluster Replication

The ability to natively replicate data to an Elasticsearch cluster from another Elasticsearch cluster is known as **Cross-Cluster Replication (CCR)**. This functionality is especially valuable for defense agencies when co-locating data to and from the edge.

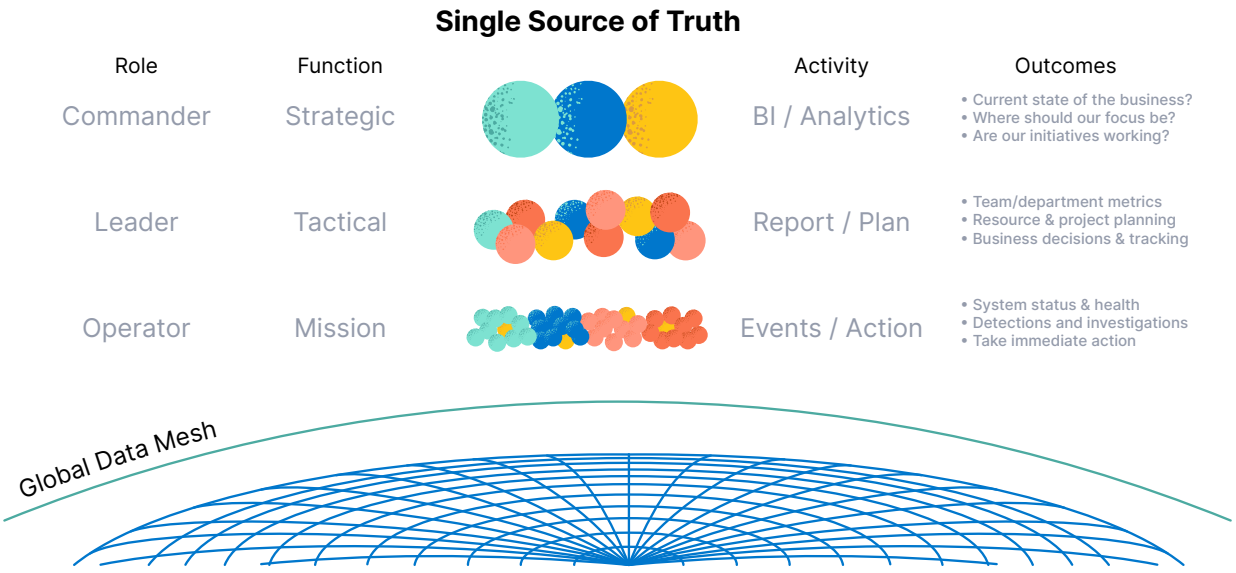
CCR enables a variety of mission-critical use cases, such as data recovery and high availability – enabling teams to withstand an outage in a region or data centre. It also allows you to colocate data to and from the edge, allowing data to be closer to the user or application server, which reduces latencies that could affect your mission. CCR can also be used to replicate data from a large number of smaller clusters back to a centralised reporting cluster. This is useful when it may not be efficient to query across a large network.

Role-based access control (RBAC)

Elastic fully supports securing digital data at creation, curation, and when handling, storing, and transmitting data. Additionally, Elastic respects your data access controls with integrated role-based access control (RBAC) security, so only users with the appropriate security credentials can access data, whether it's coming from local or remote clusters.

Using Cross-Cluster Search (CCS), RBAC security permissions are applied locally, where the cluster resides.

This allows you to create secure dynamic data access policies that span domains and functional areas. Each role can have its own view of only the data that's relevant to them, and this kind of cross-sectioning can be used to support whatever ad hoc or task-oriented teams you want to create. When that ad hoc mission is complete, you simply remove that RBAC role and access is also removed.



A global data mesh becomes a single source of truth for every use case and each level of the organisation:

- **Operators** who are monitoring the infrastructure or protecting the systems from attack interact with the latest, real-time information as it's being generated. At the individual event level, they typically need immediate access and as much automation to help them sift through that information as quickly as possible.

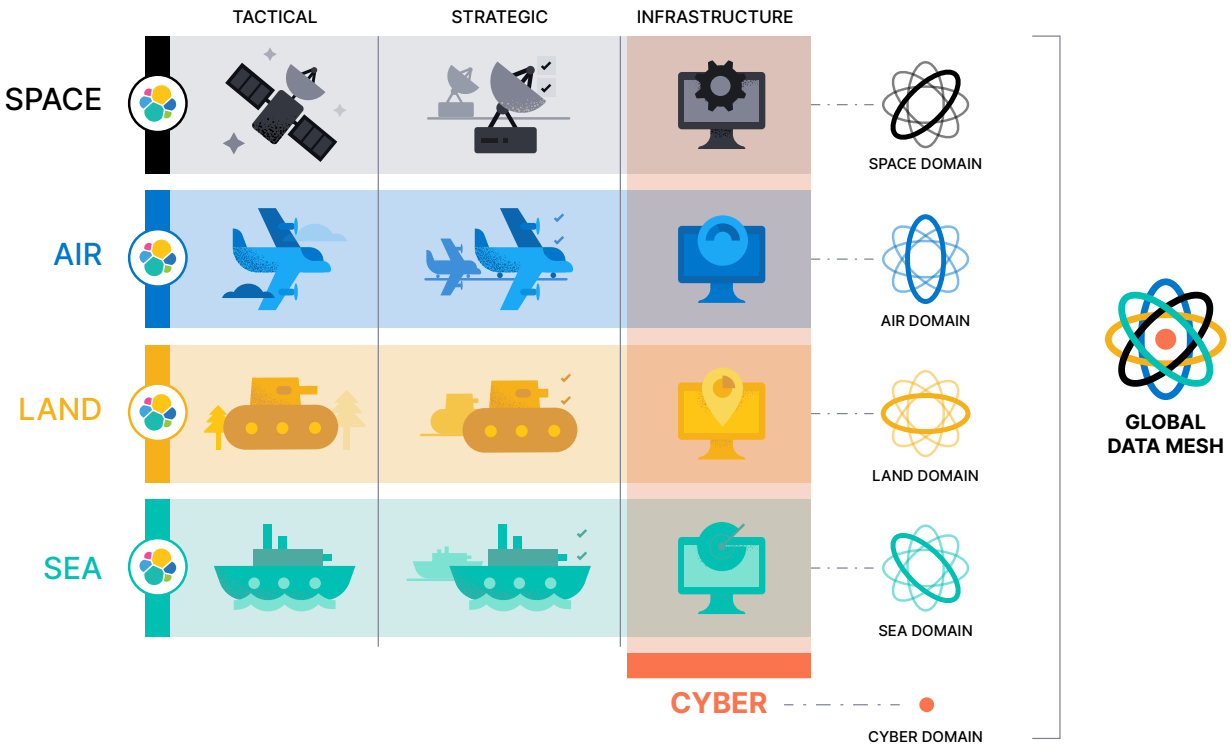
- **Company/Platoon/Squadron Leaders** need a slightly higher level aggregate view of the events so they can make sense of the operational activities, direct teams to ensure efficiency and continuity of operations, and then report on those things.

- **Commanders** have to straddle both the real-time status of the mission and use long-term trends analysis to set the course for the future.

Multi-domain data sharing for a global data mesh

What does all this mean for sharing data across functional areas and geographical locations? The tactical, strategic, and infrastructure functions within a single domain each have data sets and tools that are unique to their functions, which differ from data sets and tools in other domains. Using Cross Cluster Search, however, a functional team can share select data sets with another function using role-based access control. If the shared data set is augmented by the data owner, Elastic Cross Cluster Replication enables the data to be automatically refreshed with users authorised to see the shared data set.

Going a step further, a domain of all domains – or a **global data mesh** – can be created to enable true multi-domain operations. Machine learning that spans the entire data mesh can be incorporated to automatically pinpoint and alert on data anomalies, such as an unmanned system uplinking data from outside of its assigned zone.



Within the domains of Space, Air, Land, Sea, and Cyber there are many different functions. Some are unique to each domain, but many are functional areas that are common across all. Within each domain you can have a distributed Elastic environment that connects data wherever it makes the most sense for that function, within that domain. And spanning all clusters in each domain you can create a unified view of the data in that domain. You can then take it a step further and create a domain of all domains to enable true multi-domain operations.

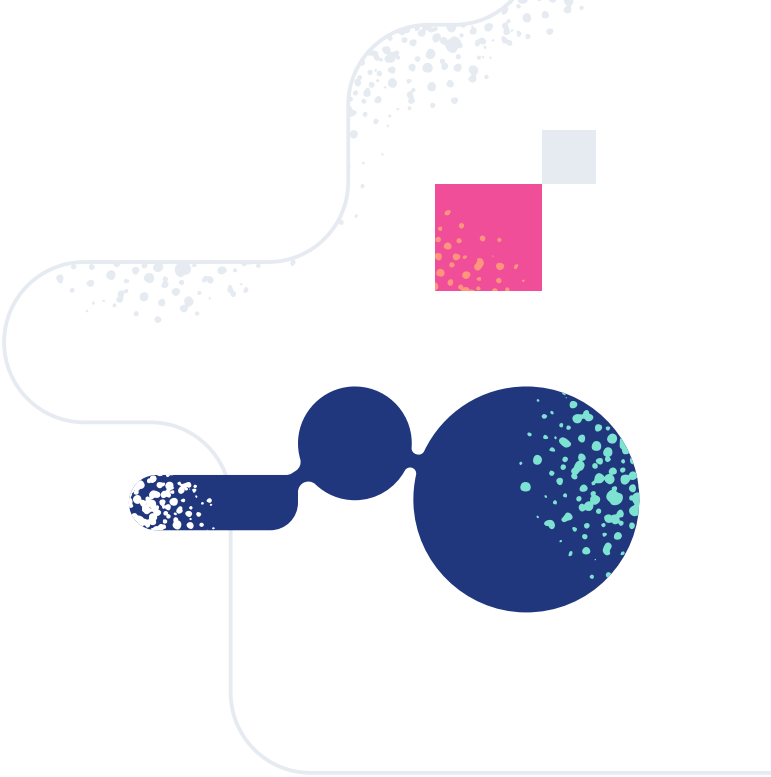
Operates in DIL environments

Elastic recognizes that disconnected, intermittent, and limited bandwidth (DIL) environments are common in the modern battlefield. Warfighters on the tactical edge can still use distributed search to query available systems, and certain remote clusters can be tagged as being more critical, ensuring that communications exist within these high value systems for the most accurate operational picture. Users are notified if remote clusters are unavailable to respond to queries, or can easily determine how long queries will take before a timeout occurs.

Elastic data mesh in action

Elastic serves as a data mesh layer in the United States© Continuous Diagnotics and Monitoring dashboard (CDM). The CDM program centralises security data from over 100 federal agencies in the US. Under the CDM program, federal agencies retain control of their own data, but the Department of Homeland Security (DHS) has central visibility across individual agency data, allowing DHS to holistically manage and remediate security vulnerabilities when needed.

Learn more about CDM's data strategy



elastic

Next steps

To learn more about a how an Elastic data mesh can accelerate your mission, get in contact with one of our experts:

www.elastic.co/contact/publicsector