

ISILON FOR GEOSPATIAL SOLUTIONS

Navagis, Open GEE and Dell EMC Isilon empower geospatial solutions for the public sector

Abstract

The whitepaper describes the solution architecture of Dell EMC Isilon solution for Geospatial Information System (GIS) in both government and organizations with Open Google Earth Enterprise and Navagis. This total solution is designed to help our customers to manage the exponential growth of GIS data, enable more efficient use of IT assets and reduce of TCO, and improve data access and workflow efficiency.

May 2018



Revisions

Date	Description
May 2018	Initial release by Frances Hu

Acknowledgements

This paper was produced by the following members of the Dell EMC storage engineering team and Navagis:

Author: Frances Hu (Frances.Hu@dell.com)

Support: Joel Martens (joel.martens@navagis.com), Tim Trahan (timothy.trahan@navagis.com)

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

© 2018 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

TABLE OF CONTENTS

- Executive summary 4
- Audience 4
- Solution Overview 5
 - GIS Solution Challenges 5
 - Reference Architecture 5
 - Solution Key Components 7
 - Dell EMC Isilon Overview* 7
 - Navagis Overview* 8
 - Open Google Earth Enterprise Overview* 9
- Design Considerations 10
 - Server Role Design Considerations 10
 - Server Design Considerations 10
 - Storage Design Considerations 11
- Key Summary 13
- Technical support and resources 13
 - Related resources 13

Executive summary

There is increasing demand for Geospatial Information System (GIS) services in both government and public sectors. The GIS technology sector has witnessed significant growth in disaster management, including flood management, forest fire management, carbon management and climate change. Modern geographic information systems provides real-time information on geography, and any changes in the information leads to rapid assessment by experts. The importance of geospatially correlating any and all resources, equipment, personnel, facilities, data etc. is growing rapidly. In order to effectively and efficiently deploy these GIS solutions in organizations and agencies, it requires a high-performance, scalable storage system while simplifying management and reducing costs.

Organizations with existing investments in GIS technology and data realize the full potential of these existing investments by using Open Google Earth Enterprise (GEE) as a cost-effective means of making GIS data usable by the organization's end users. It allows end users to access anytime, anywhere with any device.

Navagis has developed the tools necessary to integrate GEE with Dell EMC's Isilon platform. This integration allows for data in the Isilon platform to be visualized and consumed in GEE. The integration enables data discovery on any georeferenced data, turning data it into actionable intelligence.

Dell EMC's partnership with Navagis and GEE allows organizations and customers to take advantage of a best of breed approach with world leading GIS software solutions and Dell EMC's market leading unstructured data platform.

Audience

This whitepaper is intended for customers who plan to design and implement GIS solutions in the organization and public sectors. The whitepaper can also be referenced by sales and services team of Dell EMC and our partners to create a design for customer's unique requirement and environment.

Solution Overview

The whitepaper describes the solution architecture of Dell EMC Isilon solution for Geospatial Information Systems (GIS) in both government and organizations with Google Earth Enterprise and Navagis. It includes the solution overview, key components and design considerations. The solution has been architected, integrated, tested and optimized by a joint team from Dell EMC and Navagis.

GIS Solution Challenges

GIS solution are prominent in the government sectors, and the importance of geospatially correlating any and all resources and data is growing explosively. Most agencies already know the power of GIS applications to transform raw data into interactive graphical maps. According to a May 2017 report from P&S Market Research, the global GIS industry will continue to grow at a Compound Annual Growth Rate (CAGR) of 10.1%. In addition, the development of smart cities in China, India and other South East Asian countries has driven demand of 3D GIS.

There are several key challenges of GIS solutions observed from customers and public sectors:

- **Massive Geospatial Map Data Growth**

Geospatial map data ranges from gigabytes to terabytes, and imagery and video files ranges from terabytes to petabytes. In fact, according to Daratech CEO Charles Foundyler, GIS data is today the fastest growing segment of geospatial business. GIS data has grown at a compound annual rate of 15.5 percent for the last eight years – about twice the rate of growth for GIS or geospatial software and services. Consequently, as more location-related data becomes available, the use and scope of geospatial analysis is sure to grow dramatically.

- **Facilitate anytime, anywhere access GIS information**

In many organizations, there is a recognized need for GIS data consolidation and sharing to improve decision making, deliver better and more agile services, and provide consistent results. Organizations are looking for GIS solution to leverage tablets and smartphones to empower their employees and decision makers with enterprise information and location in the field. Large amount of imagery and data can be interacted via a fast and intuitive interface that requires little or no training.

- **Lower Total Cost of Ownership**

The high initial setup cost of GIS is still one of the major factors restraining the growth GIS market. Huge capital is needed for development and management of the overall GIS system. As a result, agencies and governments are looking for cost effective GIS solutions and compatibility with current IT enterprise.

- **Reduce GIS Data Silos and Security**

Many agencies and stakeholders use their own systems during the planning and operational phases of a special event. Those GIS systems and data are often disconnected, which makes exchanging and sharing information virtually impossible. Customers are looking for a GIS solution to allow users to access data in a single namespace, but also require the data to run on secure networks and keep sensitive data protected.

Reference Architecture

To address these challenges of GIS solution and facilitate easier access across the entire public sectors, Dell EMC partners with Navagis and GEE to provide our customers a total cost effective and market leading GIS solution.

Figure 1 shows the solution architecture of Dell EMC GIS solution with Navagis and Open GEE.

In production environment, Navagis web server and Google Earth Enterprise server are installed with Dell EMC Isilon hosting all the publishing flyable databases. Navagis web server allows authorized end users in the network to access with 3D or 2D databases. By integrating with Navagis, users can access visualize 3D globes or interactive 2D maps in real time with any devices. Authorized users can easily collect geo-located and time-stamped data and photographs by any operator without any training required.

In the development environment, Google Earth Enterprise fusion server is installed to integrate with Dell EMC Isilon hosting all source GIS data. It allows GIS specialist or admin to use workstations to remotely login the fusion server and do all the customization for organization, output a flyable globe to publish for end internal users.

For cost effective consideration, Isilon SmartPool feature allows GIS data to be aligned to the optimal price/performance storage tier, depending on the stage of the environment. For example, critical data, such as published flyable databases (GEE globes or maps) can be kept on a higher performance tier, while less critical data, such as source GIS data (imagery, vector, terrain) can be assigned to a more cost-effect, higher-density archive tier.

GIS data is a crucial component to truly understand underlying patterns and trends. GEE admin can integrate with ESRI software or other GIS tools to run spatial analysis of all GIS data on Isilon through HDFS (Hadoop Distributed File System) protocols and get the results in minutes.

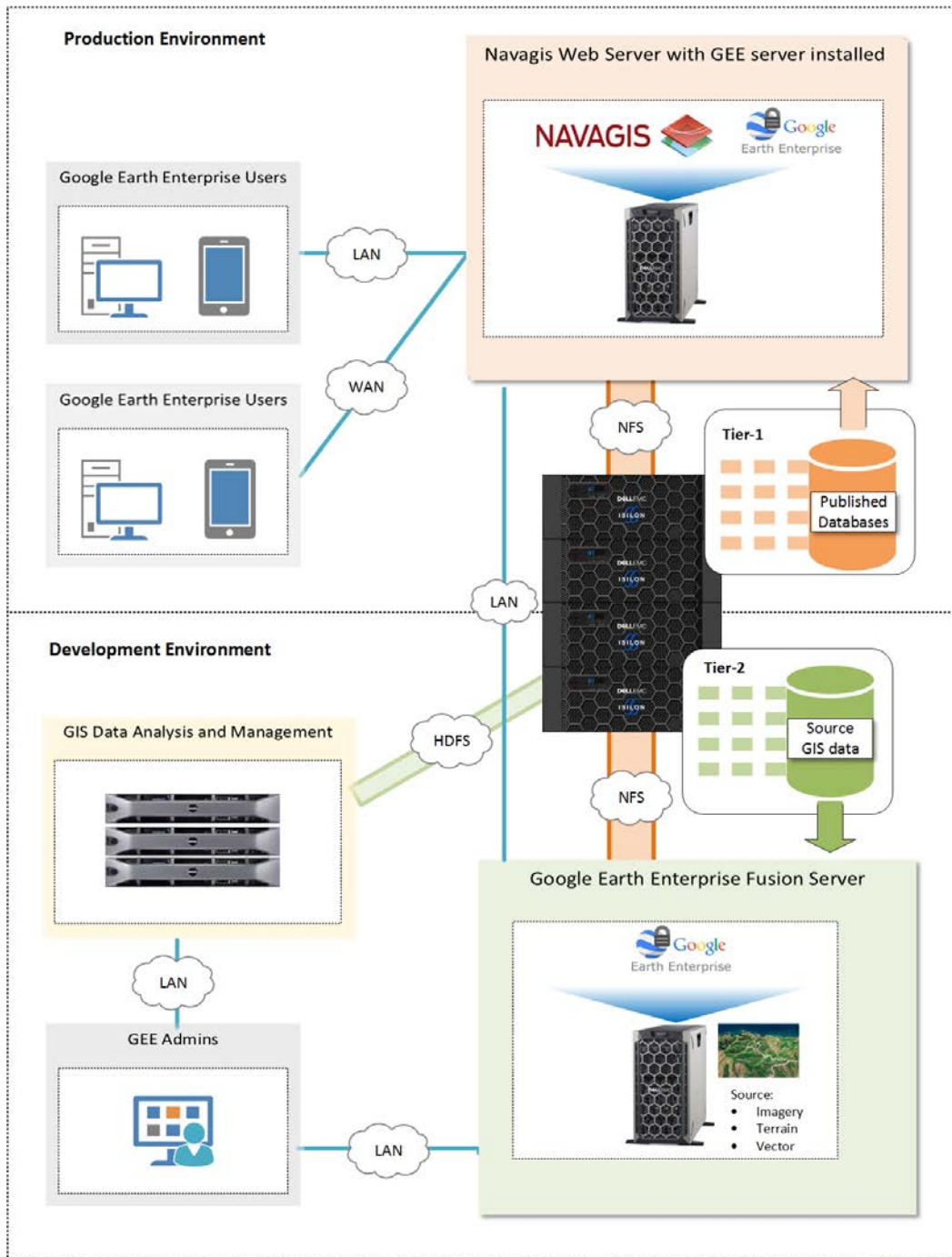


Figure 1 Solution Architecture of Dell EMC Isilon GIS Solution with Navagis and Open GEE

Here are some key factors that how agencies can benefit from an Isilon GIS solution with Navagis and Open GEE:

- **Massive scalability of capacity and performance in one single file system**

Dell EMC Isilon meets the GIS solution requirement to scale from terabytes to petabytes of capacity in a single file system. Capacity can be added to existing Isilon volumes in minutes, with no downtime or interruption to GIS data availability and no reduction in performance. The Isilon file system, OneFS, is a node-based architecture that allows up to 144 nodes, with only a single file system to administer. The scale-out architecture of Isilon and its simplicity makes it unique in the market, reduce the GIS data silos in the public sector.

- **Improve the data accessibility with multi-protocol support**

With large dataset, agencies and public sector can maximum the investment by running spatial analysis using the power of distributed processing with Hadoop technology. Isilon OneFS operating system provides access to the data not only through typical NAS protocols like NFS3, NFS4, SMB2, SMB3 and FTP, but also HDFS. By combining the highly efficient storage platform with native Hadoop integration, Isilon enables GIS specialist to accelerate their analytics efforts and get results in minutes. Native HDFS support eliminates the time and expense of moving large data sets between File and Hadoop storage.

- **Lower total cost of Ownership**

With open source Google Earth Enterprise (GEE) software and Isilon storage tiering technology, government agencies and organizations can benefit lower pricing and cost effective GIS solution, while keeping all sensitive data behind the firewall with access restricted to authorized employees. Isilon CloudPool feature allows GIS data to be aligned to the optimal price/performance storage tier, depending on the stage of the environment.

- **Access anywhere, anytime with any device without training**

By integrating with Navagis, it allows authorized users to access 2D maps or 3D globes anywhere, anytime with any device. Make useful geospatial information accessible to all of your organization's employees with the same easy-to-use technology used in Google Earth. With Navagis Mobile and Web software, it allows end users to access data in real time and collect data. Collected data is automatically converted to KML and instantly viewable on the globe.

- **Robust security and compliance options**

This solution allows government agencies to build a private globe from their own geospatial data with Navagis, GEE and Dell EMC Isilon technology. Admins run the total solution on their own datacenter and secure networks to keep all sensitive data protected. Meeting compliance and governance requirements is essential for most organizations and agencies today. To address these needs, Isilon provides robust security options including file system auditing and Data at Rest Encryption (DARE) with self-encrypting drives (SEDs).

Solution Key Components

Dell EMC Isilon Overview

Dell EMC Isilon allows you to scale out geospatial solutions with simple, powerful storage and take advantage of the Data Lake. Isilon can store, manage and protect unstructured GIS data with efficiency and massive scalability. It is the industry #1 family of scale out network-attached storage systems, designed for demanding enterprise file workloads.

GIS solutions in public sector organizations face static or shrinking budgets while being expected to provide more capabilities and services. Public sector agencies, unlike their commercial counterparts, are subject to additional transparency, security, and regulatory requirements

The Isilon Data Lake Solution revolutionizes data storage and is characterized by four pillars to the GIS solutions. As the recognized leader in enterprise network-attached-storage (NAS), Dell EMC's Isilon helps customers address these challenges by providing superior scalability, ease of use, and ease of management. Isilon is powerful yet simple to install, manage, and scale to virtually any size. Unlike traditional enterprise storage, Isilon remains simple to use and manage no matter how much storage capacity is added or how much performance is required.

- **Flexibility:** Support for all major protocols and data access methods including NFS, SMB, HDFS, HTTP, FTP etc.

- **Scalability:** Up to 68 petabytes (PB) of capacity per cluster lets Isilon take on more projects, perform more in-depth research, and serve up more data.
- **Ease of use:** tools to dramatically increase workflow productivity and maximize the value of your applications and Big Data assets.
- **Analytics:** Support for Hadoop, Pivotal, Spark, Kafka, machine learning, etc.
- **Compliance:** SEC 17a-4 - system auditing, Data at Rest Encryption (DARE). Write once, read many (WORM)
- **Access Control:** role-based access control (RBAC), Security and Technical Implementation Guide (STIG) , CAC/PIV Smartcard authentication, FIPS OpenSSL support
- **Data Protection:** fast and efficient data backup and recovery, snapshots as frequently as needed, Erasure coding, fast data replication with push-button failover and failback simplicity

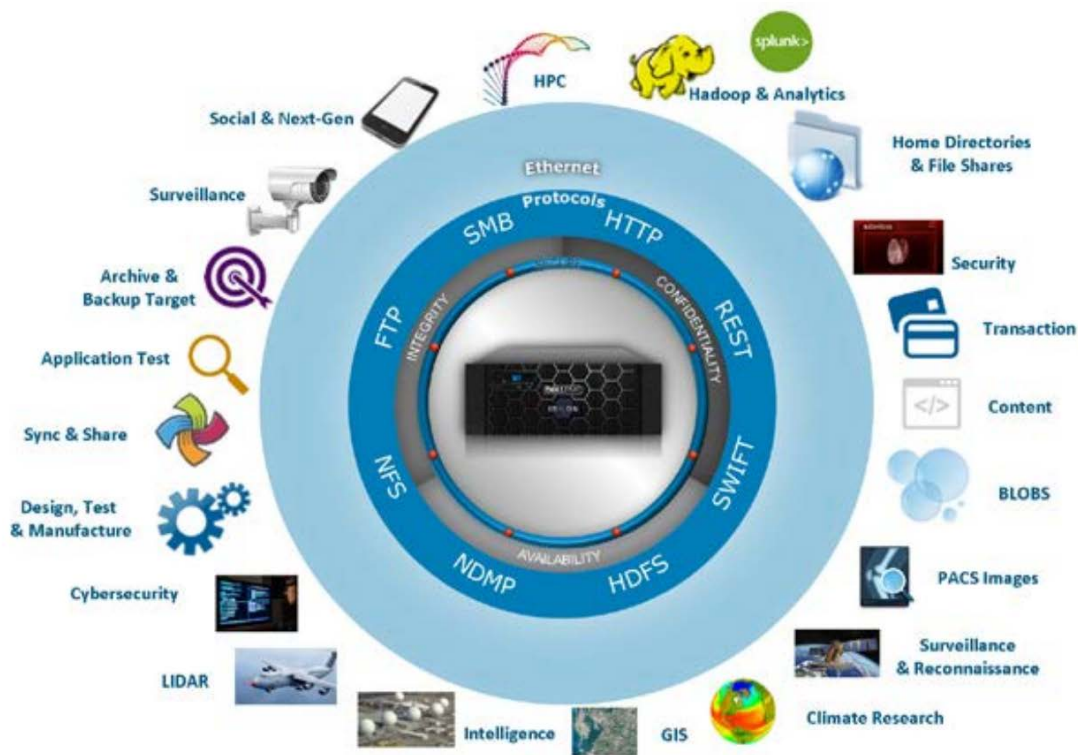


Figure 2 Isilon Data Lake Solution

For more detail information about Isilon Overview, refer to article [Dell EMC Solution Scale-Out NAS Product Family](#).

Navagis Overview

Navagis was founded with the purpose of helping defense organizations better utilize Google Earth Enterprise. Our two founders originated from the Army Corps of Engineers and Air Force, and our CEO was on the Google Earth Enterprise development team at Google. Since then, Google chose Navagis to help open source Google Earth Enterprise at Google and now Navagis maintains an active development team improving the Open Google Earth Enterprise ecosystem.

Navagis has a highly successful history of working with complex and highly sensitive data for visualization and controlled dissemination. Working with petabytes of data and stream this data in real-time to be visualized in 3D or 2D as a collaborative common operating theater is a key feature of our technology. This includes LIDAR and submarine topography data. Navagis' has worked directly for the US Army, US Army Corps of Engineers, US Air Force, Department of Justice, and other Department of Defense organizations.

Navagis has developed the tools necessary to integrate GEE with Dell EMC's Isilon platform. This integration allows for data in the Isilon platform to be visualized and consumed in GEE. The integration enables data discovery on any georeferenced data, turning data into actionable intelligence.

For more information refer to Navagis' website: www.navagis.com

Open Google Earth Enterprise Overview

Open Google Earth Enterprise (GEE) is an open source geospatial solution with a vibrant community and loads of features. It is an enterprise production ready geospatial stack with over 12 years of development behind it which allows to store and process petabytes of imagery, terrain and vector data on your own server infrastructure.

Open GEE is software that is installed on an organization's on premise servers. As such, it consumes data typically from a NAS or other storage solution. Open GEE installations on which Navagis has been involved, contain well over 100TB of information.

Open GEE helps organizations with imagery and other geospatial data make information accessible and useful to all employees via an intuitive, fast application. Open GEE transforms an organization's data into a Google Earth/Map format to visualize, explore, and understand information on a fully interactive 3D globe or 2D browser based map. It enables users to collaborate, improve decision-making, and take faster, more informed action based on geospatial information.

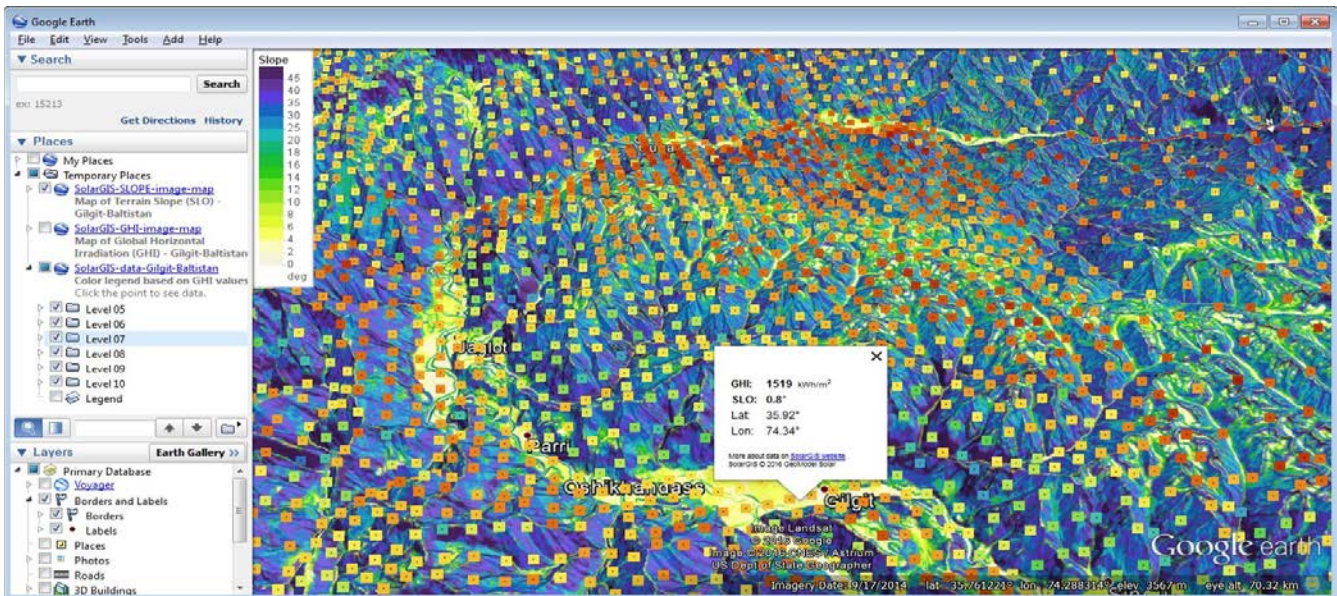


Figure 3 Layer your organization's data on top of the earth – and now the sea floor – as seen in the image of Marianas Trench.

For more information about Open Google Earth Enterprise, refer to website <http://www.opengee.org/>.

Design Considerations

This chapter lists some key design considerations of this GIS solution.

Server Role Design Considerations

Open GEE brings the same technology used to power the public Google Earth to your organization's data center. Three key components of computing resources – Client, Navagis web server and GEE server, and GEE Fusion server – bring full Google Earth capabilities to internal data.

- **Client** – Authorized users can use any device to access their organization's private globe through Navagis web server integrates with Isilon, shown in Figure 4.

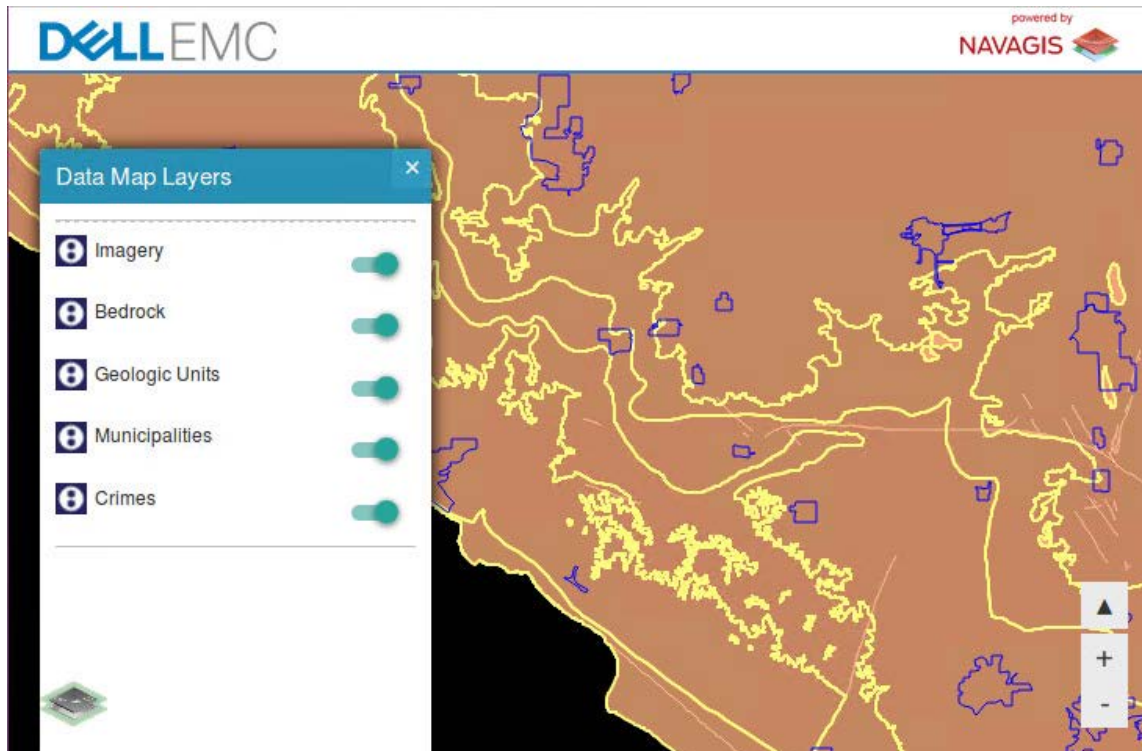


Figure 4 Data Map Layers accessed through browser

- **Navagis Web and GEE server** – The Google Earth Enterprise Server hosts the data and integrates with Navagis web server, the data is presented to end user through the connection from different clients.
- **Google Earth Fusion** – Combines all of the organization's geospatial data into a flyable globe that is made available to the users via the Google Earth Enterprise Server. Google Earth Fusion can integrate raster data – including imagery and terrain – vector data, KML, 3D models, and even data stored in traditional models.

Server Design Considerations

Fusion Server can be both CPU and I/O intensive. For optimal and stable performance, it is strong recommended to run fusion server on physical machines with Dell EMC Isilon integrated to ensure the performance requirement.

For more detail information of System Requirements for Fusion and GEE Server, please refer to link below:

<http://www.opengee.org/geedocs/answer/4492599.html>.

The web server will require resources influenced by intended use cases. For typical use cases, following normal web server (Apache based) best practices based on number of users is recommended.

Storage Design Considerations

Google Earth Enterprise Server and Google Earth Enterprise Fusion asset root, source volumes, and publish root require large amounts of disk storage space. Google Earth Enterprise Fusion requires about three times as much storage space as Google Earth Enterprise Server. In our solution, these devices are configured into Isilon NAS through NFS protocol and presented to the operating system as volumes. Volumes can be several hundred gigabytes up to hundreds of terabytes.

There are several key factors to be considered for storage design in this solution:

- **Plan the share on Isilon for Fusion data and databases**

It is strongly recommended to plan the Isilon NAS share carefully before installing the Google Earth Enterprise software. Plan where to store fusion data and publish google earth enterprise fusion databases on Isilon as these information will need to provide during the installation.

Google Earth Enterprise Fusion data (resources, projects, and databases) require a local name and network path to resolve the locations of both source files and related Google Earth Enterprise Fusion data. For that reason, you cannot change the network naming convention you adopt for host volumes without invalidating Google Earth Enterprise Fusion data.

As Isilon is a single namespace which is very easily for user to manage, we may suggest the following mount point guideline shown in Table 1:

Server Name	Mount Point Naming Convection
Google Earth Fusion	<p>The asset root must be located on a single volume. It cannot be split across multiple volumes. Therefore, it is important to think ahead and allocate as much storage space as possible for the asset root.</p> <p>Isilon NFS Share: <i>/ifs/assets</i></p> <p><i>Assets Root</i> Location on Fusion server: <i>/gevol/assets</i></p>
Google Earth Enterprise	<p>If you specify the same volume as the asset root, when you publish a database, Google Earth Enterprise Fusion registers the database on the specified volume and sets symbolic links to the database files. It doesn't take time to do copying and doesn't require extra disk space.</p> <p>Isilon NFS Share: <i>/ifs/assets</i></p> <p><i>Publish Root</i> Location on GEE server: <i>/gevol/published_dbs</i></p>

Table 1 Recommended Mount Point Naming Convection for GEE

There are different options of mount point naming conventions which is recommended by GEE. For more information, refer to article [Configure Multiple Storage Devices](#).

- **Configure the user permission on the Isilon share**

As we are using GEE over a network with different workstations and client, it is strongly recommended to use a centralized network authentication system, such as LDAP, NIS, or one of the many commercially available systems.

It is important to add the following users to authentication system's user list and grant permissions of these account to Isilon NFS share:

Type	Name
User	<ul style="list-style-type: none"> • gefusionuser • geapacheuser • gegpuser
group	geggroup

Table 2 Open GEE users and group name

It is important to make sure that **UID of users and GID of group are same** on the Isilon NAS share and GEE and Fusion server to ensure the permission. If customer hits any permission error during the configuration, please refer to the article [Permission denied](#).

• **Enable Isilon Automated Storage Tiering to match performance and cost requirement**

Isilon SmartPools software simplifies management and lowers storage costs with a transparent, policy-based, automated tiering approach. It lets organizations optimize storage resources and automatically move older, unused data to economical archive storage. SmartPools is integrated with the Isilon OneFS operating system to allow a single point of management, with a single scalable file system that offer multiple tiers of performance—depending on the data.

With SmartPools, storage administrators can automatically match storage resources with GIS data and Open GEE application performance requirements. It also simplifies management by eliminating the need for manual data migrations. SmartPools moves data among tiers based on the enterprise requirements without sacrificing data protection, application performance, or uptime. Administrators can also use defined policies to move data based on age, type, owner, location, or other criteria, from one tier to another.

Figure 5 illustrates a mixed-node, single file system deployment in an Isilon cluster for online and archive data using SmartPools.

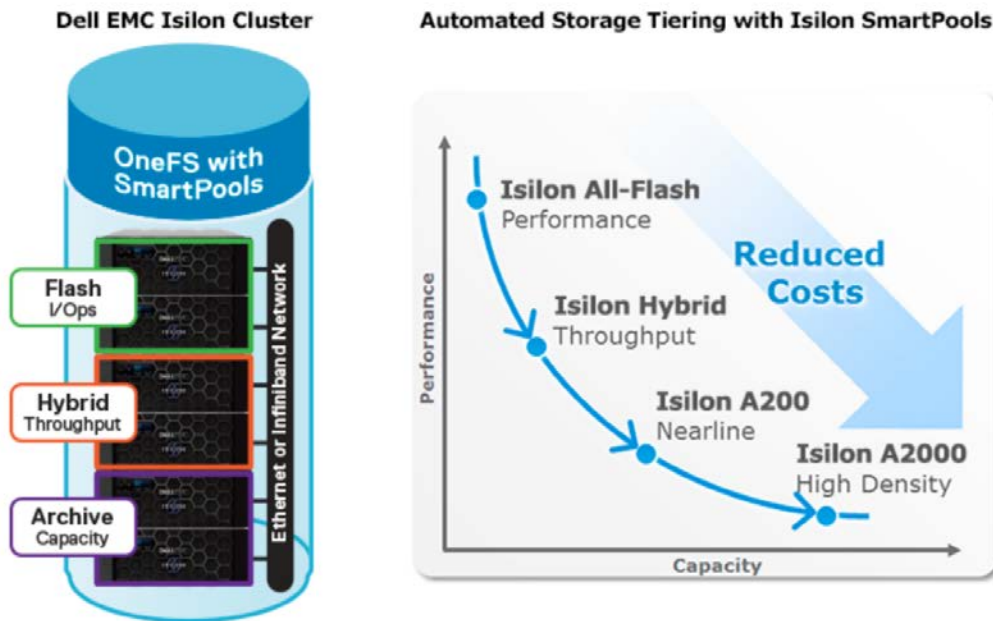


Figure 5 Isilon cluster with SmartPools for automated tiered storage

Key Summary

GIS is a dynamic IT environment. Hardware and software vendors are coming together to provide a better understanding of industry and workload patterns. They operate different workflows and processes, but what is consistent is the need to capture, analyze, and display exploding volumes of geographical information across multiple public sector departments. To aggregate, deduplicate, and make consistent our customer's multiple pools of GIS data into one single volume is the mandate for today.

Dell EMC's partnership with Navagis and GEE allows our customers to take advantage of a best of breed approach with world leading GIS software solutions and Dell EMC's market leading unstructured data platform. This total solution is designed to help our government and agencies to manage the exponential growth of GIS data and images, enable more efficient use of IT assets and reduce of TCO, and improve data access and workflow efficiency. The solution architecture provides to customer with different features and benefits of flexibility, scalability, efficiency, and manageability. The solution has been architected, integrated, tested and optimized by a joint team from Navagis and Dell EMC.

Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

[Dell TechCenter](#) is an online technical community where IT professionals have access to numerous resources for Dell EMC software, hardware and services.

[Storage Solutions Technical Documents](#) on Dell TechCenter provide expertise that helps to ensure customer success on Dell EMC Storage platforms.

[Dell EMC Solution Scale-Out NAS Product Family.](#)

[Archive Solutions with Dell EMC Isilon Scale-Out NAS](#)

[OneFS Multiprotocol Security Untangled](#)

Related resources

[Navagis Website](#)

[Google Earth Enterprise Documentation](#)



[Learn more](#) about Dell EMC Isilon solutions



[Contact](#) a Dell EMC Expert



[View more](#) resources



Join the conversation with [#DellEMCStorage](#)