Managing Large Imagery Databases via the Web

UWE MEYER, Dortmund

ABSTRACT

The terramapserver system is a perfect platform for handling geospatial information as a basis for successful business solutions of any related field. terramapserver principally stores full-coverage up-to-date geodata of any type, quality and topicality. Geodata and applications are directly available.

As a central geodata and applications portal terramapserver serves the full range of possible customer demands by enabling both data acquisition as a retail model and transaction / ASP (application service providing) GIS solutions. The internal concept describes terramapserver as a set of components. All components throughout the entire system are fully configurable and scalable.

1. THE TERRAMAPSERVER CONFIGURATION - COMBINED HIGH-END POWER

The terramapserver configuration is a well reflected system of open components. The implementation of each component follows clearly defined guidelines which guarantee most optimal solutions concerning performance, scalability and availability. Project managers can individually configure each component according to the respective product and service policy.

1.1. Data Management

Data management is the actual key issue to terramapserver. It consists of upload management, update management and metadata management. Geodata storage capacities are exceptionally high and handling requires instant reaction on business model and data availability changes by new providers.

The metadata core model describes an optimized instrument for fast and efficient data access and minimum maintenance effort by all management tools. It is extensible according to common international metadata standards to account for vector geodata and also any other related data sets such as navigational, demographic or third party branch product specific information.

terramapserver' s upload management provides a highly automated solution for pre-processing, archiving and generating both geodata and metadata entries for heterogeneous mass data. Local and remote upload is possible.

Data management tools and the metadata core model account for XML (extensible markup language) interfacing connectivity. terramapserver is thus technologically ready for remote operations with heterogeneous tools.

1.2. Hardware Structure – High Availability and Scalability

terramapserver' s hardware concept guarantees fast access to geodata via any mapping and GIS application. All components are implemented consequently redundant so that no single point of failure occurs. The basic architecture is streamlined according to optimized modern internet solutions. The layered structure reflects principle information flows.

The web farm handles large amounts of external requests to the web applications as a balanced cluster. Load balancing is done by dedicated hardware devices which are implemented as a fail-over cluster. The number of servers, the number of CPUs per server and the RAM dimension are the most important configurable parameters.

Practically no down time of the system occurs for any change of these parameters. Each geodata request is forwarded to a database server for metadata and to a file server for the image information. Both servers are set up as fail-over clusters with at least two machines each. The geodata and the metadata cluster have separate dedicated RAID storage systems. Disk volumes can easily be extended and replaced.

1.3. Internet Access and Security Concept

The web server farm handles all external requests. Valid requests to the web servers are mainly http web mapping requests. A redundant set of firewall machines filter, alert and protocol any non desired request to the web servers. Dedicated point-to-point connections and VPN (Virtual Private Network) connectivity route all other requests such as upload management, remote site administration, and also e-commerce transactions.

1.4. Web Mapping / GIS – High Performance

The web mapping / GIS component provides fast query functionality on the geodata stored in the file and metadata servers. Mapping / query on one hand and purchase on the other are the distinguished geodata use cases. For mapping and query, the component transfers compressed data to the user. For the purchase case the actual download file is produced on the web servers and referenced by other components.
A geodata request generally consists of a request to the actual image file which is handled by the file server cluster and a request on navigation / GIS and e-commerce metadata which is handled by the metadata cluster.

Retail and transaction / ASP (application service providing) business cases have separate functionality implemented in the component. The retail case includes a graphical user interface (GUI) visualizing terramapserver' s current geodata content and product selection facilities. The web mapping / GIS component for the transaction / ASP case is based on a parameterized co-ordinate oriented image query.

The component additionally feeds further web mapping based business solutions, both internal (in-house or hosted partner) and external (partner managed) over the internet.

1.5. E-Business and E-Commerce Integration

E-Commerce components extend the high performance web mapping set-up into a fully integrated geodata retail solution. Central E-Commerce components are E-Shop and electronic fulfillment. Customers can directly generate an electronic order out of any selected geodata product or product combination. Electronic orders are fully configurable with respect to

- geo-spatial (e.g. coordinate system for raster/vector data)
- IT related (e.g. file format and compression)
- commercial (e.g. licensing)

parameters. The electronic fulfillment component tailors the chosen products according to these parameters and generates result files ready to be downloaded by consumer or business customers.

The e-commerce components have XML interfaces enabling them to interact with established bookkeeping accounting and integrate into the respective enterprise workflow.

1.6. Interoperability

Interoperability is an internal and external terramapserver feature. All internal components build a modular structure so that any existing components can be integrated. Existing components may use heterogeneous middleware or operating systems.

From an external point of view terramapserver interacts with other servers on geodata and metadata as well as on application level and thus enables B2B E-Business solutions including electronic procurement, pricing and billing.

2. GEODATA DRIVEN BUSINESS

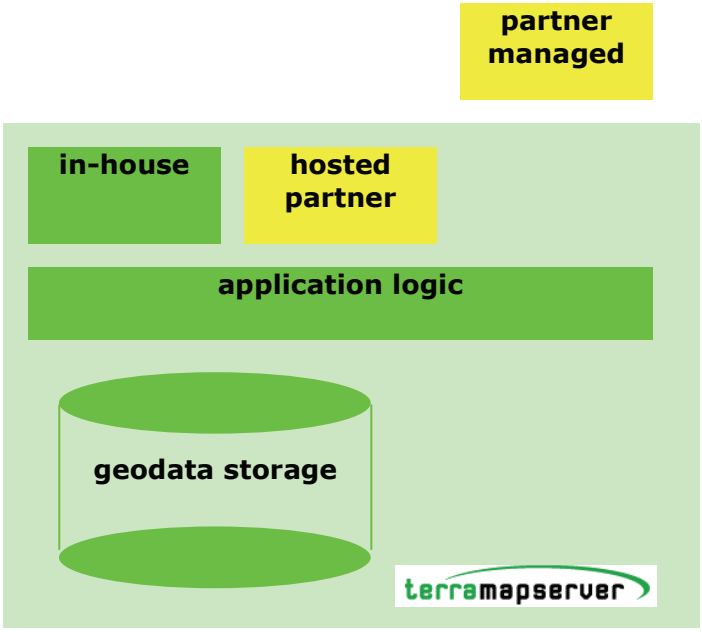
terramapserver entrepreneurs purchase the package as a basis for serving their own or any other geodata solutions to customers. By adapting the concept of strategic partnerships the server offers a three fold cooperation issue. *terramapserver data providers* update the server with geodata of optimal topicality, quality and completeness of coverage. *terramapserver technology partners* ensure the permanent update on technology of all components. *terramapserver business partners* cooperate by

- in-house solutions where terramapserver offers applications directly to the customers
- hosted partner solutions where partner geodata applications run on the server
- partner managed solutions where the application is run on a partner's server using terramapserver' s geodata

2.1. terramapserver partner concept

In-house solutions are favorable for partners who like to have the desired application logic or parts of it realized by terramapserver.
Hosted partner solutions are the optimal solution for partners with an already realized application which will use terramapserver' s geodata and infrastructure.
Partner managed solutions are particularly interesting for those partners who have also the system infrastructure ready for running the application and use the geodata of terramapserver.

All concepts are fully scalable with respect to the number of applications and the desired capacity and performance.
Running a terramapserver is thus interesting for any business model in the internet / GIS environment dealing with mass geodata content and high performance requirements on the applications plus the demand for complete E-Business integration.



3. A SUCCESS STORY

The following example implementation of the described components and services shows a successful proof of concept. The illustration focuses the final installation of the components.

The web farm consisted of eight hardware load balanced web servers running Windows 2000 Server®. The redundant load balancer devices allow the routing of each request to particular servers. Each two servers for the geodata, metadata and commerce database were set up as fail-over clusters running Windows 2000 Advanced Server®. The commerce and the meta database used SQL Server 2000®.

Web mapping was based on Intergraph's GeoMedia Web Map®. The E-Commerce part is a Navision Web Shop® solution. The metadata RAID 1 system of 360 Gigabyte has shared partitions for the Navision built-in database and for the two SQL Server databases. The second 2.1 Terabyte RAID 5 system stores the geodata.
The redundant firewall devices allow exactly the necessary external access modes to the system.

Several extensive test scenarios have proven to be necessary to guarantee the optimal layout. Particular tests were load stress test for the load balanced web farm and tests for the commerce, metadata and geodata fail-over clusters.
Table 1 summarizes stress test parameters and results. Load was generated by concurrent map server requests. Three phases distinguished LAN (phase A), load balanced local access (phase B), and internet routed access using a 100 BaseT connection to the web farm (phase C). For phase A load balancing was switched off so that tests were performed concurrent but as direct requests to each web server.
Phase B tests were operated inside the firewall protected zone. Phase C tests corresponded to external user connection.

parameter / result	phase A	phase B	phase C
no. of web servers	**8**	**8**	**8**
no. of concurrent requests (users)	**48**	**48**	**48**
load balancing	**off**	**on**	**on**
access mode	**local**	**local**	**external**
average response time [sec]	**1,647**	**1,707**	**1,696**
no. maps per server [1/hour]	**12507**	*12641*	*12695*
no. maps [1/hour]		**101130**	**101562**
internet network utilization		**33,0%**	
database server CPU utilization		**40,0%**	
database server RAM utilization		**50,0%**	
geodata server CPU utilization		**30,0%**	
geodata server RAM utilization		**20,0%**	
internet connection utilization	**-**	**-**	**25,0%**

Table 1: Load stress test results

4. THE TERRAMAPSERVER PACKAGE

The terramapserver package provides the complete foundation for setting up an internet based geodata business. Highly scalable and available components make use of most modern computer and networking standards. Terra Map Server GmbH provides expert know-how for any configuration detail including

- support concept
- data management
- internet content and layout management
- legal advisory aspects
- market leading web mapping / GIS software consulting

The terramapserver package means

- fundamental applications serving high performance
- professional project management during consulting installation and set-up
- stand-by in the start-up phase

Optional services include

- connectivity to any ERP or accounting systems
- professional multi-location upload management
- billing service interfaces