NETWORK GIS PERFORMANCE

Phil Yang[1, 3] John Evans[2, 3]

[1]joint Center for Intelligent Spatial Computing
College of Science, George Mason University, Fairfax, VA, 22030-4444

[2]Global Science & Technology Inc., 7855 Walker Dr., Suite 200, Greenbelt, MD 20770

[3]NASA Geosciences Interoperability Office, Goddard Space Flight Center, Greenbelt, MD, 20771

SYNONYMS

Computing Performance, Quality of Services

DEFINITION

Network GIS Performance (NGP) refers to the level of Quality of Services (QoS) of a Network GIS. NGP includes both the efficient use of Network GIS resources (CPU, memory, massive storage, geospatial data and geospatial services) and the perception of speed of a Network GIS [1]. Network GIS is a GIS where the geospatial data and geospatial processing are distributed across a computer network [2]. The performance [3] can be broken down into different factors, such as user perceived system response time, system reliability, system extensibility, and system service quota. Different practitioners with different roles for a network GIS have different views of NGP (table 1).

Table 1 Different Network-GIS Practitioners' Perspectives of NGP

Roles	Different Perspectives	Mainly Rely on	Restricted by
Owner	How much have to pay for serving a certain number of users?	The designer and developer to provide solution, and the end – users to provide feedback	Available budget, resources to support the procurement of a Network GIS
Designer	What strategy/architecture should be chose to meet the owner's NGP requirements?	Owner requirements and designer's experience	Available technologies, architecture, COTS
Developer	How to implement a reliable and stable system with designed functions?	Designer's design and developer's experience	Available developing tools and developer's knowledge
User	How much time is needed to wait for a system response?	System experienced and the expectation	Available Graphical User Interfaces and users' experiences
Host	How much resources are needed to provide needed service?	Infrastructure and the design requirements	Outgoing Network bandwidth, server capacity

All these perspectives are somehow depended on the different user's perceived performance, for example, how much time should be spent on waiting for a task to be accomplished (user)? Could a system support prioritizing different applications according to different types of users (designer or developer)? Is it possible to guarantee a certain level of performance if a critical application needed to be supportted, for example, police car routing or battle field supports (host/owner)?

An experience study in 2001 on user's waiting time for web browser response discovered an 8-second rule: the amount of time for a general user to wait for a web response without giving up is about 8 seconds [4]. Nevertheless, different users will have different requirements.

Performance Bottlenecks of a Network GIS

The bottlenecks of NGP includes: 1) availability of geospatial data or services[5], 2) server capabilities to respond to massive users' simultaneous accesses, 3) network bandwidth when big volume of data need to be transmitted, 4) software efficiency in processing geospatial functions. The first bottleneck is being researched as interoperability and supported by purchasing more data and services. The second bottleneck is tackled by utilizing super computer or massive cluster computing [6]. The third bottleneck is addressed by improving the bandwidth of computer network, for example, the former kilo bps phone connection is upgraded to mega bps for our home connection and the mega bps backbone is upgraded to over gega bps and tera bps. The fourth bottleneck is being solved by applying different techniques in developing and configuring geospatial software to improve the NGP.

Software Techniques for Improving Performance

Software techniques can be used to tackle all four bottlenecks, especially the fourth bottleneck (table 2).

Table 2 Methods Applicable to Improve NGP (adopted from [7, 8])

Techniques	**Bottlenecks addressed**	**Implementation location**	**Overhead**
Pyramid, Cut, Hash index	Software efficiency on Large image management	Data management server	Maintain metadata
Cache	Network Bandwidth, Data I/O	GUI and GIS server	Memory Usage
Dynamic Request	User perceived response time	Client	Smart Client
Multithreading	Software Efficiency	GUI and GIS server	Threads scheduling
Cluster/Grid Computing	Server computing capacity	GIS server	User session management
Compression	Network bandwidth	GIS server and client	Compression, decompression
Interoperability	Availability of Data and Services	GIS server and client	Reformatting data or reprogram system interface

Mechanisms for Ensuring Performance

Performances are also needed in a different manner. For example, for a routing system integrates different user requirements based on a developed Spatial Data Infrastructure (SDI): 1) travelers may need responses within minutes or hours when they plan to drive; 2) postmen may need responses within minutes before driving to next stop every time a package is delivered; 3) policemen may need responses within second to drive to a crime scene. The best case in this scenario would be everyone can get their responses within allowable time limit. In this scenario, how to guarantee the policeman get real time response needs a performance ensuring mechanism.

To ensure NGP for needed GIS applications, software techniques on scheduling could be applied to reserve or make sure high priority is given to urgent tasks. The QoS supporting protocols for computer networks can be applied here to ensure NGP. These protocols maintain a resource loading status and incoming requirements will be assigned a priority. Therefore, tasks with higher priority can be chosen for processing when geospatial resources are available. Some popular network protocols can be adopted to ensure NGP (table 3).

Table 3 Networking Protocol for Ensuring the NGP [9]

Methods	**Working Mechanism**	**Advantage**	**Disadvantage**
Best Effort	Requests are tried to be served at	Efficient System	No ensured performance

	the best capacity of the system	Usage	
FCFS	First come first serve	Based on the waiting philosophy	No priority added to specific tasks
Integrated Service	Resources are first reserved; Remaining resources can be allocated by FCFS.	Have priority and guaranteed performance	System resources may be wasted if reserved resources are not used
Differentiated Service	Requests are marked at different priorities. Resources will be allocated to the priorities first, and then could be FCFS.	Best used of the resources	Low priority requests may never be responded if there are too many high priority requests.

HISTORICAL BACKGROUND

Performance was used as a mechanism to measure time needed to accomplish a task since the inception of computer in 1941, and for scheduling tasks input to the computer [10]. Since then, performance has been widely used as a criterion for procuring an information system, designing and developing a system [11]. With the popularity of GIS and Network GIS, and the construction of the National and Global Spatial Data Infrastructure (NSDI, GSDI), performance becomes an important factor to measure the success of a Network GIS. For example, the FGDC Clearinghouse has a web interface to check the status of FGDC clearinghouse nodes [12]. Performance is also used in differentiating the geospatial information services provided by different Network GIS applications [13].

SCIENTIFIC FUNDAMENTALS

The scientific fundamentals of NGP include support from at least four different aspects: GIS, computer network research, high performance computing, and cartography principles.

Performance is reflected by a network GIS application, which requires performing a certain function (for example, routing a driving route from Washington DC to San Francisco) within an acceptable time (for example, under 8 seconds) and returning an accurate results (a correct routing map with needed details). The GIS applications also provide expectations to benchmarking the performance of a Network GIS.

The research on computer networks has lots of similarities with the research on network GIS. The methods and algorithms used for computer network routing, scheduling, assuring can also be sued to ensure the performance of a network GIS.

The high performance computing provides needed computing power to support a time consuming network GIS. For example, the google earth online mapping application is supported by thousands to millions of computers (or CPUs). An earth science simulation runs for a week on a desktop computer could be finished within minutes on a high performance computer

Cartography principles provide fundamental methods in separating computing objects into pieces to be calculated on different computers and abstracting computing objects into different levels to facilitate the needs of different users from global level to regional level, for example, visualizing the globe on a desktop and visualize the National Mall on a desktop from the same data source.

KEY APPLICATIONS

NGP can be used in many application domains, most notably the infrastructure and emergency management communities.

Spatial Data Infrastructure (SDI)

NSDI and GSDI have been researched, developed for over 10 years to share the geospatial information resources through computer networks. When massive users accessing the infrastructure, NGP will be an important mechanism to measure the success of a GSDI/NSDI and provide reliable services to urgent needs.

Decision Support Systems
NGP can be used in decision support systems to schedule the information integration and user access where a variety of geospatial information are integrated and a variety of user needs are supported.

Daily Life
NGP can be used to help our daily life by prioritizing our tasks in geospatial information processing, such as travel places, meeting locations, and other complex geospatial information needs.

Defense
NGP can be used to help schedule the various defense needs of global geospatial information network [14] to be constructed by the national spatial intelligence agencies for control, field battle, and fight plan.

Emergency Management
NGP can be used to differentiate GIS services provided to different personnel involved in emergency management, such as fire fighters, doctors, and policeman.

Geospatial computing infrastructure
NGP can help to prioritize the users' access to geospatial computing infrastructure if it is accessed by simultaneous users for services. For example, real time coastal management forecasting, weather forecasting, emergency management, and disaster management parameter calculations.

Popular geospatial services
NGP can be used for popular service, Google Earth and Virtual Earth [15], for example, to provide reliable services to paid customers and emergency needs while provide acceptable services to the general public.

National and Global Geospatial Applications
NGP can be used to provide reliable services to different users, communities, countries, and regions to strategize the usage and maximize the benefit of such national and global geospatial applications.

FUTURE DIRECTIONS

NGP is an emerging concept with the popularization of geospatial information services, interoperability, spatial information infrastructure, and geospatial computing infrastructure. It provides a promising mechanism to utilize limited resources to support unlimited and differentiated requirements according to the requirement nature, submit user, regions, and user community. However, to put it in practical use, more research, test, installment, evaluation, and benchmarking have to be conducted.

The evolution of NGP relies on future work of 1) architectural evolution of geospatial information services and performances research, 2) the sharing strategies of super computing and other computing research [16], 3) the user cognitive research, 4) priority scheduling evolution of computing sciences.

CROSS REFERENCE

Distributed Geospatial Computing, Efficient Indexing Techniques of Uncertain Spatial Data，
Spatial Information Retrieval

RECOMMENDED READING

1. Apple Developer Connection, *What Is Performance*? http://developer.apple.com/documentation/Performance/Conceptual/PerformanceOverview/DevelopingForPerf/chapter_2_section_2.html
2. C. Yang, D. Wong, M. Kafatos, R. Yang, Network GIS, *Earth Science Satellite Remote Sensing*, Qu, Gao, Kafatos, Murphy and Solomonson (Eds), Springer and Tsinghua University Press, Vol 2, pp. 252-269, 2006

3. J. Evans and F. Clarence, *Deploying IP and MPLS QoS for Multiservice Networks: Theory and Practice*, Morgan Kaufmann, 2007, ISBN 0-123-70549-5
4. Submit Corner, 2001, *8 Second Rule*, http://www.submitcorner.com/Guide/Bandwidth/001.shtml
5. C. Yang, V. Tao, 2005, Chap.5: Distributed Geospatial Information Service, *Frontiers of Geographic Information Technology*, Edited by Sanjay Rana and Jayant Sharma, London: Springer, pp. 113-130.
6. R. Healey, S. Dowers, B. Gittings, M. Mineter, *Parallel Processing Algorithms for GIS*. Bristol: PA: Taylor & Francis, 1998.
7. C. Yang, D. Wong, M. Kafatos, R. Yang, Implementing computing techniques to accelerate network GIS, Proc. SPIE Vol. 6418, 64181C, *Geoinformatics 2006: GNSS and Integrated Geospatial Applications*; Deren Li, Linyuan Xia; Eds. 2006
8. C. Yang, D. Wong, R. Yang, M. Kafatos, and Q. Li, WebGIS performance improving techniques, *International Journal of Geographic Information Science* (19):319-342, 2005
9. Evans J. and Filsfils C., 2007, *Deploying IP and MPLS QoS for Multiservice Networks: Theory and Practice*, Morgan Kaufmann. ISBN 0-123-70549-5.
10. Kempf, K., 1961, *Historical Monograph: Electronic Computers Within the Ordnance Corps*. http://ed-thelen.org/comp-hist/U-S-Ord-61.html
11. Barber S., 2004, *Beyond performance testing part 2: A performance engineering strategy*, http://www-128.ibm.com/developerworks/rational/library/4215.html
12. FGDC Clearinghouse Node performance measurement, http://www.fgdc.gov/clearinghouse/
13. Yang C., Evans J., Bambacus M., Cole M., Alameh N., and Marley S., 2006, Spatial Web Portal: Architecture and Example, *PE&RS* (in press).

14. G. W. Fuller, *Capabilities for a Future National Intelligence Enterprise*, http://www.garywfuller.net/pubs/capabilities_nie.html
15. V. Tao, *Microsoft Virtual Earth: New Horizons in on-line mapping*, keynote at GeoInformatics'2006, Wuhan, China October 28-29, 2006
16. G. Aydin, A. Sayar, H. Gadgil, M. S. Aktas, G. C. Fox, S. Ko, H. Bulut,and M. E. Pierce. Building and Applying Geographical Information System Grids 12 January 2006, *Special Issue on Geographical information Systems and Grids based on GGF15 workshop, Concurrency and Computation: Practice and Experience*, 2006

ACRONYMS

COTS, Commercial Off The Shelf

FCFS, First Come First Serve

FGDC, Federal Geographic Data Committee

GEOSS: the Global Earth Observation System of Systems

GIS, Geographic/Geospatial Information System

GSDI, Global Spatial Data Infrastructure

GUI, Graphical User Interface

NGP, Network GIS Performance

NSDI, National Spatial Data Infrastructure

PTL, Pervasive Technology Lab

QoS, Quality of Service

SDI, Spatial Data Infrastructure