Computer Network Specification for Distributed Spatial Data System

Agus Pribadi¹, I Wayan Astika², M. Arief Syafi'i³ ¹Student of MIT SEAMEO BIOTROP, Bogor Agriculture University ²MIT lecturer, ³Bakosurtanal staff

ABSTRACT

In recent years, spatial data and database become very important thing. Spatial data and database are stored into data and database server called data center. The last trend of spatial data and database management utilizes the concepts of data clearinghouse (DCH) and data warehousing (DWH). The Geographic Information System (GIS) can contain a wide variety of geographic data types originating from many diverse sources. Data source providers also come from various locations separately. The main issue of GIS project concerning to the data and information is currentness. In many cases, GIS projects consume a lot of the time to collect data inputs due to the distributed location of data. This condition decreases the currentness of GIS data used in the project. Another major development of recent years is in computer networks. The computer network can make many GIS data providers (server) interconnected. The most current utilization of computer network in the GIS is still used for data transfer and data visualization. The terminal simply provides a means for entering and displaying data for the computer.

The objective of this research is to specify the computer network specification to support distributed spatial data processing based on GIS application. This system will be developed based on computer network infrastructure in Indonesia. The research methodology is field study, especially for computer network at Bakosurtanal and Biotrop – MIT as case. The result of this research is computer network specification that is used for spatial data transmission on distributed spatial data especially is in the internet connection.

Based on current condition of physical computer network connection, the internet connection by peer-topeer communication mode was be found unstable. The requested timed out signals were gotten 41.8% - 75.2% in the ping test of internet connection. The required time to transfer a GIS data file about 500 MByte was around 9 hours. Thereby, 100 KByte of GIS data needing data transfer time was around 6 seconds. The current physical connection technique for distributed spatial data using 128 Kbps bandwidth of internet connection, network cable category 5, is not recommended because the network connection is dissatisfactory. The new computer network specification has been arranged to solve the internet connection problem. It characteristics are: a) the intranet backbone uses 10 gigabit mode of fiber optic, b) using at least category 6 of twisted-pair cable for local computer network connection, c) the internet connection capacity (bandwidth) is recommended to apply at least 2 Mbps with stable connection, d) the recommendation of physical computer network topology is tree, e) utilizing a computerized network switch system to manage the computer network resources easier.

Chapter I. INTRODUCTION

1.1. Background

Data source providers for Geographic Information System (GIS) come from various locations separately. Another major development of recent years is in computer networks (De Bay, 2000). GIS application developers usually do GIS project on the network terminal as workstation. The most common problem people have when creating or viewing a map is currentness of the geographic information (Callahan, 2002). The GIS projects consume a lot of the time to collect data inputs due to the distributed locations of data.

The conventional computer network utilization on the GIS was only applied for transfer data and also data visualization. In Indonesia, common computer network transmission applies 10 Mbps until 100 Mbps using cable connection, and the network-to-network connection is around 56 Kilo bit per second (Kbps) to 2 Mega bit per second (Mbps). That is so small and low speed for spatial data transmission requirement.

Data Clearinghouse and Data Warehousing implementation are only in a group supported by locally computer network, which data / database could not be shared for another user outside of the group. The spatial data collection and management (in Indonesia) still have character of sectorally and a little bit egoism, every institution does it to support their function and task itself (Syafi'i, 2002). In the GIS construction, around 60% - 80% of total expense and time are for data collection preparation. The activities are enough cleanse energy, timeconsuming, human resources elements and need many expenses.

1.2. Problem Statement

The problem statement of this research is the computer network specification to serve distributed spatial data system for conducting the product of GIS application easier that utilizes peer-to-peer communication model.

1.3. Objective

The objective of this research is to specify a computer network specification to support distributed spatial data processing based on GIS application. This system will be based on computer network infrastructure in Indonesia.

1.4. Scope and Limitation

The scopes and limitations of this research are :

- a. The computer network study was based on data transmission time (rate).
- b. The location of local computer network (case study) was done at Badan Koordinasi Survei dan Pemetaan Nasional Bakosurtanal (National Coordinating Agency for Survey and Mapping). This is an internal computer network calls as intranet.
- c. The internet case study was using the computer network of Bakosurtanal and laboratory of Master of Science in Information Technology for Natural Resource study program (MIT).
- d. The remote spatial data access and utilization are based on physical connection technique (peer-to-peer communication model).

1.5. Output

The output of this research is computer network specification that is used for spatial data transmission on distributed spatial data especially is in the internet connection.

1.6. Benefit (Outcome)

The benefit of this technique is to prepare and provide GIS applicators / producers for generate the thematic GI easier, than conventional technique and technology, and it will be helpful. The other benefits are :

- a. connection of (all) related data provider;
- b.reducing cost consumable, resources problems and time consuming;
- c. the simplicities to acquire, process, and analyze the spatial data;
- d.increases the effectiveness, efficient-ness and currentness;
- e.easy to in finding and processing the spatial data;
- f. standardizations are on the institution in charge of data, and format;
- g.effectiveness and efficiency are in the global data management's;
- h.easy to control the data utilization by data custodian;

Chapter II. LITERATURE REVIEW

2.1. Distributed System and Databases

A distributed system consists of independent computers connected to one another (Gallo, 2002). Post (1999) also explains in his book that, the databases are usually in different physical locations; each database is controlled by an independent Database Management System (DBMS), which is responsible for maintaining the integrity of its own database. In the extreme condition, the database might be installed on different platform (hardware, software and operating system), but could even use DBMS by utilizing software from different vendor. In the example shown in Figure 1, an institution could have offices in three different locations.

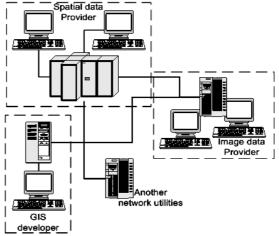


Figure 1. Distributed database (Post, 1999)

The most popular method today involves a client/server approach (Post, 1999). Another technique is to utilize the technology what called Data Clearinghouse (DC) and Data Warehousing (DW). Based on the GIS science namely spatial data, the DC is the core module which direct users to the relevant data resources (Wonjun, 2002).

2.2. Data Warehousing

A data Warehousing (DW) is an organized collection of database and processes for information retrieval, interpretation and display (Green, 2002). Data source layers from various sources are combined to produce a global subject oriented view of a geographic region. The process includes assembling and managing data from various sources for the purpose of gaining a single detailed view or part or all of an organization (Green, 2002).

The applied model of the data-warehousing concept is distributed data by single centralized index as shown on the Figure 3 below. DW systems are most effective when data can be combined from many operations systems (Green, 2002).

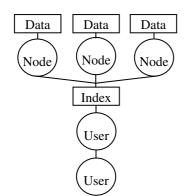


Figure 3. Datawarehouse model using distributed data configuration (Green, 2002)

2.3. Data Clearinghouse

A Data Clearinghouse (DC) is the core module which direct users to the relevant data resources (Wonjun, 2002). The functional design of the DC database is based on the supporting architecture established in the functional requirements document. One important thing in the data clearinghouse is the core module which direct users to the relevant data resources (Wonjun, 2002).

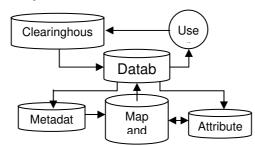


Figure 4. Clearinghouse configuration (Wonjun, 2002)

Figure 4 illustrates the data clearinghouse configuration. User can access GIS data easily after contacting the clearinghouse first directly. The user can access GIS data to the data sources through the clearinghouse after order by query. The clearinghouse will connect user to a suitable database based on user requirement.

2.4. Computer Network

Computer network is a collection of computers and other devices (nodes) that use a common network protocol to share resources with each other over network medium (Gallo, 2002). Another understanding about computer network is the group of computers and related devices that have connection by common medium, and can communicate each other (Pribadi, 2002).

If a collection of computer networks is connected to one another, then we have what is known as a network of networks, which is formally called an **internetwork** (Gallo, 2002). A wide area network interconnects geographically separated LANs using a point-to-point topology (Gallo, 2002).

Chapter III. METHODOLOGY

3.1. Study Arrangement

The core methodology of this research is a field study to computer networks. The field study was dealing with activities to observe the computer network operations.

3.1.1. Location

The research locations were carried out at :

- a) Bakosurtanal, that was focused on the computer network of "Center for Networking System and Spatial data Standardization" (PSJSDS) and "Center for Marine Base Mapping and Aeronautical Charting" (PPDKK).
- b) The computer laboratory of MIT campus, which is a part of wider computer networks under Southeast Asian Regional Center for Tropical Biology of South East Asia Ministry of Education Organization (Biotrop).

3.1.2. Time

The intranet observation was carried out on July 2005. The internet observation was carried out on August to September 2005. The observation has been done between two sides that are MIT through Biotrop – Bakosurtanal (in PSJSDS) and PSJSDS (Bakosurtanal) – Biotrop.

3.2. Material and Tool

3.2.1. Data

- The supporting data fulfilled this observation are:a) 3 bands of Landsat satellite image for Bogor city in format 'ers'.
- b) Vector data of Road and Stream for Bogor city in format 'erv'.
- c) 5 visible bands of Landsat satellite image for Riau area in format 'ers'.
- d) 3 bands of Ikonos satellite image for Bogor city in format 'tiff'.
- e) Vector data of topo-map for Depok municipality seperated in 9 NLPs and in format 'dxf'.

3.2.2. Software

The application software, installed on computer (terminal) in Bakosurtanal :

- f) ER Mapper version 6.4 (Earth Resource Mapping),
- g) ArcMap GIS 9.0 (Environmental System Research Institute Inc.

The computer network and personal computer system softwares are Microsoft Windows, Read Hat Linux 9 and CISCO Router and firewall management (hardware installer).

3.2.3. Computer Network Infrastructure

The Bakosurtanal computer network specifications are :

- completing by central server system,
- 12 CISCO (multi speed system) switches,
- 3 hub devices 2-proxy servers (128 Kbps leased line service by Indosat & broadcast by 512 Kbps downstream and 128 Kbps upstream of Indosat),
- fiber optic cable backbone.

The Biotrop – MIT Laboratory computer network and specifications are :

- central server (4 servers & 2 shared servers),
- 2 switched-hubs,
- 256 Kbps leased line dedicated connection,
- 10/100 Mbps Ethernet connection.

3.3. Observation

This network observation is activities to observe real condition in situ of the computer network infrastructures that used by Bakosurtanal and added with MIT Laboratory networks. Figure 5 depicts the research methodology structure.

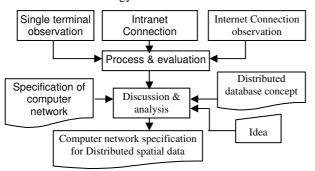


Figure 10. Research methodology diagram

3.3.1. Network Observation

To get the bit rate of each data can be derived from time operation to file capacity. Green (1991) proposed a formula to calculate data signaling speed is :

Data signaling speed = $\frac{Log 2^n}{T}$1

and, simplify to be :

Data signaling speed = $\frac{filesize}{t} x8$ 2

where :

• Log 2^n = file size = data stream size (bit)

• T = t = time (second)

Operated formula is on each node :

Capacity =
$$\frac{(Bandwidth - 10\% bandwidth)}{n - node} \dots 3$$

where :

- Bandwidth = data transmission rate (bit/second)
- n-node = number of node in the network

3.3.2. Data

Based on the results in the table can be calculated the difference, trend, and average of data load and process by personal computer, local computer network (intranet-work) and internet-work. Mangkuatmodjo (2003) proposed a formula to calculate geometric mean is :

Average_(g eo) = $\sqrt[n]{X_1 x X_2 x \dots x X_n}$ 4

Average application is geometric mean. The geometric mean is a useful summary when we expect that changes in the data occur in a relative fashion (Simon, 2005).

Chapter IV. RESULT AND DISCUSSION

4.1. Correlation

The Geographic Information System (GIS) can contain a wide variety of geographic data types originating from many diverse sources (Longley, 2001). The distributed data system for spatial database management system (SDBMS) was applying the Data Clearinghouse (DC) concept and or Data Warehousing (DW) concept. Computer network supports the basic operation of SDBMS. Separated spatial data storage can be accessed by user remotely.

Distributed Geographic Information Systems (DGIS) element applies the dynamic client/server concept in performing GIS analysis tasks (Peng, 2003). Client computer can access and request data to process it utilize analysis tools from server and the server will process job itself and sends the result back to the client, or the server sends data and analysis tools to the client for processing. Peer-to-peer communication model of computer networks servers for DGIS using physical spatial data access gives opportunity to user for utilizes spatial data that was stored in remotely server.

4.2. Observation

4.2.1. Intranet Observation

Table 1 depicts the summary result of the observation procedures for local computer and local computer network. The mean of time access / process duration was calculated using geometric mean.

The general result is vicinity to satisfy. The local computer and network operation has the same trend. It was fast for small raster and vector data size. The network operation is more time consuming than the local operation.

 Table 1. Result of observation in the Intranet and local computer

No	Itama	Average of process		
	Items	Local (mm:ss)	Intranet (mm:ss)	
1	Loading the satellite image using ER Mapper 6.4	00:04	00:07	

No	Items	Average of process		
110	items	Local (mm:ss)	Intranet (mm:ss)	
2	ISOCLASS Unsupervised classification (15 class) processing of satellite image using ER Mapper 6.4	00:41	01:46	
3	Loading the 5 bands of satellite image using ER Mapper 6.4	00:02	00:04	
4	Loading the multi-taks bands of satellite image using ER Mapper 6.4	00:02	00:05	
5	Loading the satellite image using ArcMap GIS 9	00:02	00:09	
6	Image display by 1 level fix zoom using ArcMap GIS 9	00:03	00:10	
7	Loading a scene (NLP) of vector data using ArcMap GIS 9	00:04	00:06	
8	Loading 4 scenes (NLPs) of vector data using ArcMap GIS 9	00:05	00:10	
9	Loading 9 scenes (NLPs) of vector data using ArcMap GIS	00:08	00:23	

4.2.2. Internet observation

The special condition in the internet connection observation did not get observation result because was found connection problem. The observation was changed to be network connection test.

Table 2 depicts the summary result of the internet connection observation by temporal ping. Two ways test procedure was done. Consider to table 2 and 3, the general result is not satisfied. The ping test procedure has proven that the internet connection between Bakosurtanal and Biotrop was unstable.

 Table 2. Result of observation on internet using temporal ping

	Connection test		Result of ping test			
	from MIT-lab	Number of signal				
No	through Biotrop to				Loss	
	Bakosurtanal by	Send	Receive	Loss	(%)	
	temporal ping					
1	Day 1	120	49	71	59.2	
2	Day 2	120	8	112	93.3	
3	Day 3	120	20	100	83.3	
4	Day 4	120	64	56	46.7	
5	Day 5	120	72	48	40	

 Table 3. Result of observation on internet using continues ping

	Connection test activity	Result			
No		Number of signal			Loss
		Send	Received	Loss	(%)
1	Bakosurtanal to Biotrop server.	2199	545	1654	75.2

		Result			
No	Connection test	Number of signal		Loss	
	activity	Send	Received	Loss	(%)
2	MIT-lab through Biotrop to Bakosurtanal data server.	2073	1206	867	41.8

4.2.3. Data Transmission 4.2.3.1. Intranet Connection

Intranet data transmission rate will be influenced by internal bit-rate capacity. Bakosurtanal applied 1 Giga bit per second (Gbps) of backbone capacity by fiber opticThe bandwidth quota each Center (9 centers) is :

B/W cap = (Bandwidth - 10 % bandwidth) bps n - nodeB/W cap = (1000 Mbps - 10 % 1000 Mbps)= 100 Mbps

Based on calculation above, the bandwidth quota of every Center is equal 100 Mbps. The 100 Mbps quota corresponded with fast-ethernet by category 5E of twisted-pair cable. That condition could handle the data transmission that file-size is likes in table 1 relatively quick.

4.2.3.2. Internet Connection

Based on the internet connection, tries to calculate the spatial data namely 543.4 MB of Ikonos satellite image for Bogor city. It assumed that there is not disturbance on data transmission through internet connection.

The byte to bit conversion is :

Bit-s = 543,400 Kbyte x 8 = 4,347,200 Kbit

note : 8 = bit - byte conversion factor

The time consuming is :

Time = file size : bandwidth capacity5

Time = 4,347,200 Kbit : 128 Kbps = 33962.5 s

Time = 9 hour 26 minute 2.5 second

The result is very consuming loads time.

Based on the calculation above, it is tried to implement for 100 Kbyte of spatial data. The calculation is :

Bit-s = 100 Kbyte x 8 = 800 Kbit

The time consuming is :

Time = 800 Kbit : 128 Kbps = 6.25 second

Time access = 6.25 second

6.25 second is not so consuming time for loads the GIS data file.

The 100 Kbyte spatial data indicated satisfactory result although the big size (543,4 Mbyte) was dissatisfactory. The 6.25-second of time access for 100 Kbyte spatial data is not indicator to get satisfactory connection.

4.3. Computer Network Specification **4.3.1.** Specification Arrangement

According to Chapter 4 Sub 4.2.3.1, the 100 Mbps quota of every Center has been derived from 1 gigabit (Gbit) of Bakosurtanal backbone. It is used for 543.4 MB of satellite image data.

Bit-s = 543.4 MByte x 8 = 4347.2 Mbit

Time = 4347.2 Mbit : 100 Mbps = 43.472 s Time access = 43.472 second

The 543.4 MB of single satellite image data could be accessed in 43.472 second. The 100 Mbps quota will be limited to implement for more than 2 spatial data.

The 10 Gbit is biggest capacity of computer network connection cable; it uses fiber optic (Peng, 2003). If it was known that Bakosurtanal has 9 Centers. The bandwidth quota each Center is :

B/W cap =
$$(10.000Mbps - 10\%10.000Mbps)$$

$$= 1000 \text{ Mbps} = 1 \text{ Gbps}$$

The 1 Gbps per Center is used to transmit the 543.4 MB of satellite image data.

Bit-s = 543.4 MByte x 8 = 4347.2 Mbit

Time = 4347.2 Mbit : 1.000 Mbps = 4.3472s Time access = 4.3472 second

4.3472 second is enough to use for raster data processing. The 1 Gbps should be used twisted-pair cable category 6 is recommended. Its capability is 1.2 Gbps - 2.4 Gbps (Gallo, 2002).

The 1 Gbps of internet connection has not been available in Indonesia but it is in mega-bit. The access time is assumed in 1 minute. The calculation uses 543.4 Mbyte of satellite image data is :

Bit-s = 543.4 MByte x 8 = 4347.2 Mbit

Time = 1 minute = 60 second

Data transmission capacity = file size : time...6 where :

- file size = data stream size (bit)

- time = total of transmission duration (second) Data transmission capacity = 4347.2 Mbit : 60s

Data transmission capacity = 72.4533333 Mbps Based on the calculation, the data transmission capacity requirement is 72.4533333 Mbps but it is not available in Indonesia. 2 Mbps of peer-topeer interconnection was obtained by ITB (Plasa.com, 2001). Assumed, the quota is 2 Mbps used for the 543.4 MB satellite image data. Bit-s = 543.4 MByte x 8 = 4347.2 Mbit

Time = 4347.2 Mbit : 2 Mbps = 2173.6 second Time = 36 minute and 13.6 second

4.3.2. Technical Specifications

According to observation result and calculation, the recommended technical specifications of computer network for distributed spatial data system are :

- 10 gigabit mode intranet backbone;
- gigabit system cable with 1.2 Gbps 2.4 Gbps is recommended for local area network;
- recommended to apply at least 2 Mbps internet connection;

- physical computer network topology recommendation is tree;
- it is recommended to utilize 3.6 Giga-Hertz & 64-bit CPU system, 144 MByte cache memory, 512 GByte RAM, 2 Terra-Byte storage;
- completed with spatial data application system is on provider and user side.

Chapter V. CONCLUSIONS AND RECOMMENDATION

5.1. Conclusions

- 1. The internet connection by peer-to-peer communication mode was found unstable. The requested timed out signals were gotten 41.8% 75.2% in the ping test of internet connection.
- 2. By using 128 Kbps of internet connection, the required time to transfer a GIS data file about 500 MByte was around 9 hours. Thereby, 100 KByte of GIS data needing data transfer time was around 6 seconds. The current physical connection technique for distributed spatial data using 128 Kbps bandwidth of internet connection, network cable category 5, is not recommended because the network connection is unsatisfactory.
- 3. The new recommended of computer network specification for distributed spatial data system are:
 - a. intranet backbone uses 10 gigabit mode of fiber optic;
 - b.gigabit system cable with 1.2 Gbps 2.4 Gbps is recommended for local area network;
 - c.internet connection capacity (bandwidth) is recommended to apply at least 2 Mbps;
 - d.the recommended of physical computer network topology is tree;
 - e. the latest technology of system server that using high-capability of multi-processor at least 3.6 Giga-Hertz of dual-core server processor, 64-bit CPU system, 144 Mega-Byte cache memory, 512 Giga-Byte memory, 2 Terra-Byte storage, gigabit fiber channel and fast-ethernet of network terminal;
 - f. in the spatial data provider side and user side are completed with spatial data application system.

By using the new computer network specification, the expected time to transfer a big file about 500 MByte is about 37 minutes by internet connection and less than 5 seconds using local computer network connection.

5.2. Recommendation

According to the observation result, the internet connection was unstable for the incircuiting computer. Data transmission improvement needs high-specification technology support of computer network capability. Client/server communication model is another alternative. The client/server communication model was supported by service application.

The internet connection is suggested to use client/server communication model of computer network communication technique. Some examples of client/server communication model application on spatial data transmission are Web Map Service, Web Feature Service, Web Coverage Service and Web Processing Service. Based on the computer network comprehension, it will be better to implement that communication model for spatial data protocol to ensure the system can run well for internet GIS system.

REFERENCES

- Callahan, Christina L, John A Callahan, Richard S Sacher, Michael B Mahaffie, William S Schenck, Shannon S Bain, Robert E Rinehart. 2002. Delaware DataMIL : A Web-based Mapping Collaboration, A Prototype for USGS's The National Map. *In* ArcNews. 2002. GIS in Action. Vol 24 no 2. ESRI.
- De Bay, Rolf A., Ricard A. Knippers, Yuxian Sun, Martin C. Ellis, Menno-Jan Kraak, Michael J. C. ewir, Yola Georgiadou, Mustafa M. Radwan, Cees J. van Westen, Wolfgang Kainz, Edmund J. Sides. 2000. *Principles of Geographic Information Systems : An introductory textbook.* The Netherlands, Enschede : The International Institute for Aerospace Survey and Earth Science (ITC).
- Gallo, Michael A., and William M. Hancock. 2002. *Computer Communications and Networking Technologies*. USA : Brooks / Cole, Thomson LearningTM.
- Green, David and Terry Bassomaier. 2002. *Line GIS and Spatial Metadata*. London : Taylor and Francis, 11 New Fetter Lane.
- Longley, Paul A., Michael F. Goodchild, David J. Maguire, David W. Rhind. 2001. *Geographic Information Systems and Science*. England, Chichester : John Wiley & Sons Ltd.
- Mangkuatmodjo, H. S., Prof. Drs. 2003. *Pengantar Statistik*. Jakarta : P.T. Rineka Cipta.
- Peng, Zhong-Ren, Ming-Hsiang Tsou. 2003. Internet GIS : Distributed Geographic Information Services for the Internet and Wireless Networks. New Jersey : John Wiley & Sons, Inc.
- Plasa.com [Internet]; Gateway Internet dan Link ke ISP; 2001 [cited 2005 October 16th]; Available from : http://www.plasa.com/instan/.
- Post, Gerald V. 1999. Database Management System
 : Designing and Building Business Applications.
 McGraw Hill Companies, Inc.
- Pribadi, Agus, Dyah Susilowati, B Reni Sari Dewi. 2002. Peningkatan Penguasaan Materi Mata

Kuliah Jaringan Komputer Melalui Pendekatan Praktek dengan Menggunakan Metode "Student Directed Strategic". Report of Action research. Mataram : SPMU TPSDP STMIK Bumigora and CPMU TPSDP of DGHE, Informatics Engineering Department, STMIK Bumigora.

- Simon, S., Ph.D. [Internet]; *Geometric Mean*. StATS; 2005 August 8th – [cited 2005 September 19th]; Available from : http://www.cmh.edu/stats/ definitions/geometric.htm.
- Syafi'i, Muhammad Arief. 2002.Spatial Data Standard, Sebuah Sarana untuk Memudahkan Pertukaran Data Spasial. *In* : Prosiding FIT ISI. ISBN : 979-95834-3-8.
- Wonjun, CHOI. 2002. Institutional Issues of the Spatial Data Exchange in Korea. FIG XXII International Congress, Washington, D.C.; Education & Research Centre, Korea Cadastral Survey Corp. [serial online]. Available : http://www.survey.ntua.gr/main/labs/photo/researc h/wg_33/National Report s/SDE - Korea.pdf via the INTERNET. Accessed 2005 April 30th.