

Distr.
GENERAL

WP.6
14 April 2010

ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION
FOR EUROPE (UNECE)
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE EUROPEAN
UNION (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Meeting on the Management of Statistical Information Systems (MSIS 2010)
(Daejeon, Republic of Korea, 26-29 April 2010)

Topic (i): Developing common high-level architectures

A Study about Statistical ICT Architecture for National Statistical Office – The Case of China

Prepared by ZHENG Jingping, National Bureau of Statistics of China

I. Introduction

1. In this paper, based upon a general study about statistical ICT architecture for a national statistical office, the National Bureau of Statistics of China is used as an example, to discuss China's Statistical ICT System (SICTS), which covers all stages of statistics production – statistical questionnaires design and distribution, data collection, data processing and dissemination. The SICTS is composed of “a network”, “a software platform” and “a set of databases”.
2. The paper will be comprised of 4 parts. Part 1: Architecture is very important for building a Statistical ICT System, as well as for the NSO itself. Part 2: The basic principles for building up Statistical ICT architecture. Part 3: A case from China. Part 4: It is a common task to build up an efficient Statistical ICT System. There is a large room for international cooperation.

II. Architecture is very important for the building a Statistical ICT System, as well as for the NSO itself

A. The characteristics of statistical ICT system

3. Now human being is in the ICT time which is composed of the environment official statistics. To take the advantages of ICT time and enhance the statistical capability of the NSO it is necessary to build a statistical ICT system which has the following characteristics:
 - (a) Integrity – the system should cover all stages of statistics production---statistical questionnaires design and distribution, data collection, data processing and dissemination. There for, the system requires all part of it could link each other easily and smoothly.;
 - (b) Harmonization – indicators, statistical data and mate data in the system should come from one and only one source, have one and only one explanation, consistent each other;

- (c) Continuity/ long term project – the system should be used in NSO’s daily work continuously when it is set up, since the standard evaluation of a statistical ICT project is sustainable and continuously application. There for, the system is long term project or life long project of NSO;
- (d) Wider coverage – the system is related to each staff, each sector and each key point of data production and dissemination in NSO;
- (e) Huge investment – the system is a huge investment project, since the above characteristics having decided it.

B. Guarantee of the success of the statistical ICT system

4. Since the statistical ICT system is a key project for NSO, since the system is a long-term project with the characteristics of wider coverage, integrities, harmonization and high cost, we should have an architecture to build and control the project to guarantee the success.

III. The basic principles for building the Statistical ICT architecture

A. Consistent with the architecture of NSO

5. A good statistical ICT architecture should be consistent with the architecture of the NSO. Since there is no pure statistical ICT project, each ICT project has to serve the demand of statistical capacity building which essentially is a statistical application project.

B. Proven by law or having validity

6. Since the importance of the statistical ICT architecture, it is necessary to be proven by law, or by the authority of legal system.

C. Advancing relatively

7. This means the technology and guilds idea should consider the development of ICT and statistical capacity building in the future.

D. Practicability

8. The system should easy to use with friendship interface.

E. Communicating with users sufficiently.

9. To reach the practicability, we need communicating with users sufficiently and understanding their demand completely.

F. Considering international standards and experience carefully and sufficiently.

10. Since in the opening world, communicating and comparison are important. To reach the purpose, we need “common” language. Since statistical function and capacity building are quite similar from one country to another, learn from each other can avoid mistakes and wrong way.

G. Executive process of the architecture is inspecting able and auditable.

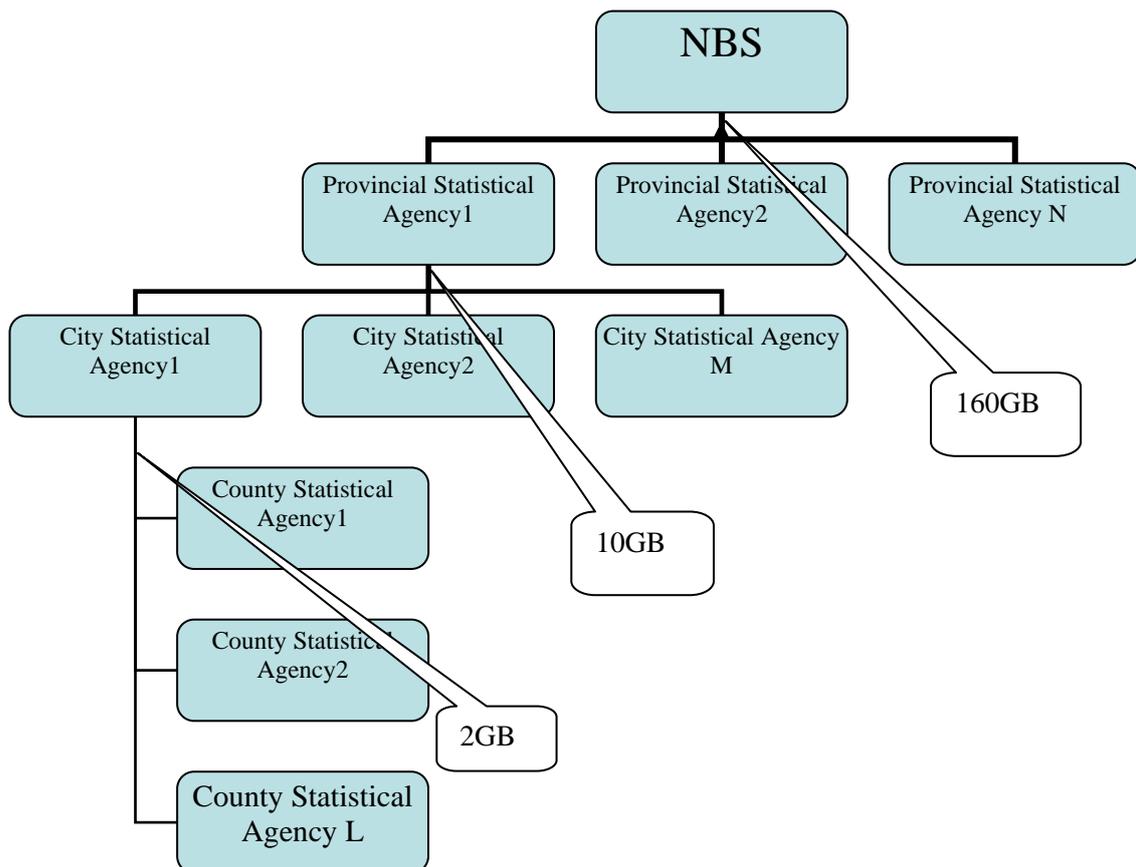
11. To keep a good architecture practicable it should be inspected and audited by users during its executive processing in order to test and revise our project timely and guarantee the project success.

IV. The Architecture of China Statistical ICT System (SICTS)

12. The Architecture of China Statistical ICT is called China's Statistical ICT System (SICTS), which covers all stages of statistics production – statistical questionnaires design and distribution, data collection, data processing and dissemination and all sectors (agriculture, manufacture, service sector etc.) of the National Bureau of Statistics of China. The SICTS is composed of “a network”, “a software platform” and “a set of databases”.

A. “A network”

13. A network is the data communication highway which connects raw data suppliers and Statistical Agencies, Statistical Agencies themselves, data users and Statistical Agencies. At present, the bandwidth of internet between provincial statistical agencies and National Bureau of Statistics of China is about 160GB; the bandwidth of internet between city statistical agencies and provincial statistical agencies is about 10GB; the bandwidth of internet between county statistical agencies and city statistical agencies is about 2GB. The data suppliers of companies and individuals are using VPN to connect NBS's servers (see the diagram below).



B. “A software platform”

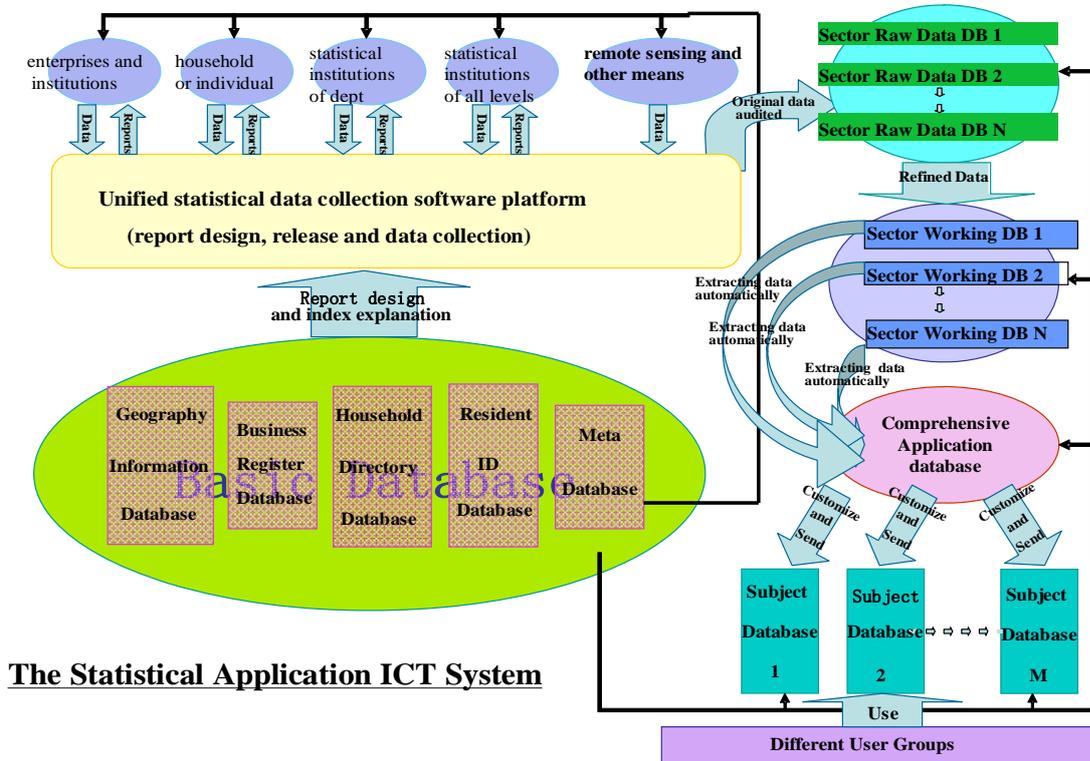
14. A software platform is the unified software platform for data collection which makes all sectors (including agriculture, manufacture, service sector etc.) and all means of data collection (including sampling survey, census, registration, administrative records and direct report from raw data suppliers) on line when the data supplier is ready to use electronic means.

C. “A set of databases”

15. A set of databases is composed of Basic Database, Sector’s Raw Databases, Sector’s Working Databases, Comprehensive Application Database and Subject Databases like Dissemination Database.

- (a) The Basic Database is composed of Meta Database, Business Register Database, Household Directory Database, Resident ID Database and Geography Information Database.
- (b) The comprehensive application database is a platform for comprehensive departments’ daily work and for data sharing among sector departments. It is a common platform for data sharing with NBS.
- (c) The different subject databases, which are generated by the comprehensive application database, serve the external users for different purposes.
- (d) All these databases, as well as the unified software platform, are supported by the meta database that is unique in the system. This ensures that the production and application of data have a unified, normative and unique standard.

D. The development strategy for NSO of China



16. The graph of Statistical Application Part of SICTS gives the basic idea of the system. The software platform could design and distribute questionnaire to data suppliers by using Basic Database. Data suppliers could report the questionnaires to different sector department in NBS through the software platform. When different sector department receive the raw data from the data suppliers, they store them in different Raw

Database. After cleaning these raw data, the refined data will be exported to different sector's Working Database. Then the Comprehensive Application Database will extract some aggregated data from different sector's Working Database automatically. Finally according to different regulations we can create different Subject Database for different users, for example, the Dissemination Database, and so on.

17. As the graph shows that the Meta Database connects to software platform, data suppliers, Raw Databases, Working Databases, Comprehensive Application Database, Subject Databases and users.

V. International Cooperation

Room for international cooperation

18. Because the official statistics production processes are so "standard", that means almost all countries use the same method to produce statistical products, there is possibility for the international statistical and ICT experts working together to create a (or a set of) "unified" software of statistical production processes. During the course, maybe there are some languages problems for the software, but the main principles and the frameworks are a good model for all countries around the world. The SDMX project is a very good start and example. By this way the statistical capacity and the application ability of all developing countries could be improved quickly, effectively and efficiently.

19. Now NBS of China is concentrating on the construction of statistical informationization: reconstructing the statistical production processes by using ICT, building up China's Statistical ICT System (SICTS). Any suggestion, comment and support is welcome.
