DataBase Management Systems and the Internet

Panel Chair:
Susan Malaika, IBM Hursley, UK

Panel Members:
Mariano Consens, consens@niobe.uwaterloo.ca, University of Waterloo, Canada
Alistair Dunlop, and@ecs.soton.ac.uk, University of Southampton, UK
Daniel Ford, daford@almaden.ibm.com, IBM Research Almaden, USA
S Ramani, ramani@saathi.ncst.ernet.in, National Center for Software Technology, India
Chander Sarna, chander@netscape.com, Netscape Communications, USA
Gio Wiederhold, gio@cs.stanford.edu, Stanford University, USA

In the last few years, major changes have occurred in computing, as a result of the fast adoption of the Web (at least 300,000 sites excluding Intranets in July 1996) and the popularity of Java. One of the side effects has been conferences that attract thousands of attendees which do include some database related topics, e.g. Hyper-G and Java Database Development Toolkit[1, 2, 3, 4, 5]. However, the content of database conferences remains almost unaffected by the Internet. Thus one must assume that database research has not been swayed by the dramatic changes although most commercial database management systems have incorporated some Web interfaces, gateways and Java constructs.

In traditional database management systems, it is generally assumed that responsible people are in charge of naming, creating directories, defining datatypes, indexing, integrating sites, managing applications etc. The conglomerations of data and applications on the Internet (and on Intranets) are often created, maintained in bottom-up fashion, with few people in control. Indeed, this is one of the aspects that made rapid adoption possible.

Proceedings of the 22nd VLDB Conference
Mumbai(Bombay), India, 1996

Entities from individuals to whole countries can participate, on almost equal terms as illustrated in July 1993 New Yorker magazine: "On the Internet, nobody knows you are a dog". In some cases individuals are in a better position than countries, e.g. with respect to communications bandwidth, as described in Carl Malamud's "Exploring the Internet - A Technical Travelogue"[6].

Aspects of searching, navigation and visualisation have become much more significant than traditional directories for locating people, data and programs. This is due to the very rapid change in the data accessible on Internet and Intranets, the quantity of information and the lack of control. The sites on the Web that contain indexes built by crawlers are very popular, e.g. 14 million accesses per day for AltaVista[7], in July 1996, and with a high rate of increase. The content of most traditional database management systems is invisible to the Internet and Intranet worlds, because it cannot be found easily.

Internet protocols[8] include provision for caching on the network, e.g. through intermediate proxies and on client systems, thereby improving user response time by reducing the number of accesses to data and programs across the network. The communication protocols are supported by firewall systems (that protect intranets) and incorporate encryption based on public key. Heterogeneous data, e.g. text, structured data, images

and programs, can be downloaded by users and programs using simple interfaces that are consistent across platforms. Infrastructures are being devised to manage heterogeneous datatypes and programs, and the links between them. Servers dedicated to specific tasks are evolving, e.g. link management[9], end user link manipulation[10], Internet business applications[11], Java program execution[12]. Performance benchmarks[13] for accessing Internet based data are becoming established.

In "An Introduction to Database Systems"[14] first published in 1975, Chris Date described the benefits of centrally controlled databases containing structured data. The advantages include reduced redundancy and inconsistency, sharing data with security and integrity, balancing conflicting requirements. The emphasis was mainly on programming interfaces that provide security and control, in accordance with meta-data defined by an expert.

In a seminar in 1991, Tim Berners-Lee and Robert Cailliau described the fundamental Web constructs[15]:

- Common user interfaces (based on hypertext)

- Standard communications protocols to heterogeneous data and programs

- Standard naming conventions

The emphasis was on straight-forward end-user access to diverse data formats from a variety of platforms in different locations. It was referred to as "universal readership".

At the JavaOne[16] conference in May 1996, James Gosling described the history of the Java programming language, and its sudden success when applied to the Web, e.g. the introduction of the idea of Java programs (applets) associated with Web pages. The applet notion is now evolving into a method for distributing client software for client-server systems.

Is it time to create new data management services that embody constructs suited to the current popular environment? If the database community doesn't originate them - someone else will!

References

[1] The WWW5 Conference, Computer Networks and ISDN Systems, Vol 28, Numbers 7-11 May 1996
http://www.w3.org/hypertext
/Conferences/WWW5/

[2] Tutorial on Web access to legacy data, Leon Shklar, WWW5 conference
http://www.cs.rutgers.edu/?shklar/
web-legacy/www5_tutorial.html

[3] JDBC Java DataBase Connectivity, Rawn Shah, Java World. May 1996
http://www.javaworld.com/jw-05-1996/
jw-05-shah.html

[4] JavaScript For Database Connectivity, Chander Sarna, Netscape Developers conference, March 1996
http://developer.netscape.com/
/index.html

[5] Hyper-G from Graz University, Austria
http://fstgds08.tu-graz.ac.at/
C2127CC2/ROOT

[6] Exploring the Internet, A Technical Travelogue, Carl Malamud, Prentice Hall, 1993, ISBN 0-13-296898-3

[7] AltaVista from Digital
http://www.altavista.digital.com/

[8] IETF HTTP working group
http://www.ics.uci.edu/pub/ietf/http/

[9] Microcosm and the WWW: A Distributed Link Service G.J. Hill et al, Southampton University
http://www.ecs.soton.ac.uk/rj95/
mm/gjh/rj.html

[10] aqui from IBM - End user link manipulation
http://www.aqui.ibm.com/aqui
/index.html

[11] OpenMarket - Internet software for business
http://www.openmarket.com

[12] Jigsaw Java server from CERN
http://www.w3.org/pub/WWW/Jigsaw/

[13] Webstone Benchmark from Silicon Graphics
http://www.sgi.com/Products/
WebFORCE/WebStone/

[14] A Introduction to Database Systems, C.J. Date, Addison Wesley

[15] World Wide Web Seminar, 1991,
Tim Berners-Lee, Robert Cailliau
http://www.w3.org/hypertext/WWW/
Talks/General.html

[16] The JavaOne Developers Conference - Keynote presentations
http://java.sun.com/javaone/