Chapter LXXXIII
Querical Data Networks

Cyrus Shahabi
University of Southern California, USA

Farnoush Banaei-Kashani
University of Southern California, USA

INTRODUCTION

Recently, a family of massive self-organizing data networks has emerged. These networks mainly serve as large-scale distributed query processing systems. We term these networks *Querical Data Networks (QDN)*. A QDN is a federation of a dynamic set of peer, autonomous nodes communicating through a transient-form interconnection. Data is naturally distributed among the QDN nodes in extra-fine grain, where a few data items are dynamically created, collected, and/or stored at each node. Therefore, the network scales linearly to the size of the dataset. With a dynamic dataset, a dynamic and large set of nodes, and a transient-form communication infrastructure, QDNs should be considered as the new generation of distributed database systems with significantly less constraining assumptions as compared to their ancestors. Peer-to-peer networks (Daswani, 2003) and sensor networks (Estrin, 1999, Akyildiz, 2002) are well-known examples of QDN.

QDNs can be categorized as instances of "complex systems" (Bar-Yam, 1997) and studied using the complex system theory. Complex systems are (mostly natural) systems hard (or complex) to describe information-theoretically, and hard to analyze computationally. QDNs share the same characteristics with complex systems, and particularly, bear a significant similarity to a dominating subset of complex systems most properly modeled as large-scale interconnection of functionally similar (or peer) entities. The links in the model represent some kind of system-specific entity-to-entity interaction. Social networks, a network of interacting people, and cellular networks, a network of interacting cells, are two instances of such complex systems. With these systems, complex global system behavior (e.g., a social revolution in a society, or food digestion in stomach!) is an emergent phenomenon, emerging from simple local interactions. Various fields of study, such as sociology, physics, biology, chemistry, etc., were founded to study different types of initially simple systems and have been gradually matured to analyze and describe instances of incrementally more complex systems. An interdisciplinary field of study, the complex system theory[a], is recently founded based on the observation that analytical and experimental concepts, tools, techniques, and models developed to study an instance of complex system at one field can be adopted, often almost unchanged, to study other complex systems in other fields of study. More importantly, the complex system theory can be considered as a unifying meta-theory that explains common characteristics of complex systems. One can extend application of the complex system theory to QDNs by:

1. Adopting models and techniques from a number of impressively similar complex systems to design and analyze QDNs, as an instance of engineered complex systems; and
2. Exporting the findings from the study of QDNs (which are engineered, hence, more controllable) to other complex system studies.

This article is organized in two parts. In the first part, we provide an overview, where we 1) define and characterize QDNs as a new family of data networks with common characteristics and applications, and 2) review possible database-like architectures for QDNs as query processing systems and enumerate the most important QDN design principles. In the second part of the article, as the first step toward realizing the vision of QDNs as complex distributed query-processing systems, we focus on a specific problem, namely the problem of effective data location (or search) for efficient query processing in QDNs. We briefly explain two parallel approaches, both based on techniques/models borrowed from the complex system theory, to address this problem.

BACKGROUND

Here, we enumerate the main componental characteristics and application features of a QDN.

Componental Characteristics

A network is an interconnection of nodes via links, usually modeled as a graph. Nodes of a QDN are often massive in number and bear the following characteristics:

- **Peer functionality:** All nodes are capable of performing a restricted but similar set of tasks in interaction with their peers and the environment, although they might be heterogeneous in terms of their physical resources. For example, joining the network and forwarding search queries are among the essential peer tasks of every node in a peer-to-peer network.
- **Autonomy:** Aside from the peer tasks mentioned above, QDN nodes are autonomous in their behavior. Nodes are either self-governing, or governed by out-of-control

Related Content

Mining in Spatio-Temporal Databases
Junmei Wang, Wynne Hsuand Mong Li Lee (2005). *Spatial Databases: Technologies, Techniques and Trends (pp. 272-293).*
www.irma-international.org/chapter/mining-spatio-temporal-databases/29668

Optimization of Multidimensional Aggregates in Data Warehouses
Russel Pearsand Bryan Houliston (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications (pp. 2324-2347).*
www.irma-international.org/chapter/optimization-multidimensional-aggregates-data-warehouses/8040

Goal-Oriented Requirement Engineering for XML Document Warehouses
Vicky Nassis, Tharam S. Dillon, Wenny Rahayuand R. Rajugan (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications (pp. 570-595).*
www.irma-international.org/chapter/goal-oriented-requirement-engineering-xml/7932

Antecedents of Online Game Dependency: The Implications of Multimedia Realism and Uses and Gratifications Theory
Kaunchin Chen, Jengchung V. Chenand William H. Ross (2010). *Journal of Database Management (pp. 69-99).*
www.irma-international.org/article/antecedents-online-game-dependency/42086

Situational Method Engineering to Support Process-Oriented Information Logistics: Identification of Development Situations
Tobias Bucherand Barbara Dinter (2012). *Journal of Database Management (pp. 31-48).*
www.irma-international.org/article/situational-method-engineering-support-process/62031