

Dept. of Computer Science.
Box 1910, Brown University.
Providence RI 02912-1910, USA.
Tel: 401-863-7698, Fax: 401-863-7657

Education

- 2004 - present Ph.D. candidate. Computer Science.
Brown University, Providence, RI, USA.
Thesis: Continuous-Time Data Stream Processing.
Advisor: Prof. Uğur Çetintemel.
- 2001 - 2004 Sc.M. Computer Science.
Brown University, Providence, RI, USA.
Thesis: Network-Aware Query Processing.
Advisor: Prof. Uğur Çetintemel.
- 1998 - 2001 B.Eng. Computing.
Imperial College of Science, Technology and Medicine, London, UK.
Thesis: CVE - A Collaborative Virtual Environment.
Advisor: Prof. John Darlington.

Research Experience

Pulse: A Framework for Continuous-Time Data Stream Processing

Pulse is a data stream framework for processing continuous queries over models of continuous-time data, which can compactly and accurately represent many real-world activities and processes. Pulse implements standard streaming operators that consume continuous-time models as inputs and produce the same type of model as outputs. Each continuous-time operator works by solving a simultaneous equation system, which, in many cases, is significantly cheaper than processing a stream of tuples. My contributions included deriving the high-level design of Pulse, and implementing an actual prototype of the framework in the Borealis stream processing engine. Furthermore I experimentally verified that Pulse is practical and demonstrates significant performance gains for real workloads from the New York Stock Exchange, and the Automatic Identification System for tracking naval vessels.

SenseWeb: A geo-based web interface for publishing and browsing live sensor data – <http://research.microsoft.com/nec/senseweb/>

SenseWeb is a peer produced sensor network that consists of sensors deployed by contributors across the globe, providing querying and tasking functionality that may be used to develop sensing applications. SensorMap is one such application that mashes up sensor data from SenseWeb on a map interface. In this project, my contributions included the design of COLR-Tree, a database abstraction layer to support efficient spatio-temporal queries in SensorMap. COLR-Tree uses a novel caching mechanism on multi-resolution spatio-temporal aggregates, and incorporates an efficient one-pass sampling algorithm that factors in cached data and occasional unavailability of sensors. I implemented and evaluated COLR-Tree on Microsoft SQL Server 2005 with a real workload from Windows Live Local to demonstrate significant improvements in query and network performance.

Scalable Adaptive Network Databases – <http://www.cs.brown.edu/research/sand/>

The SAND project investigates distributed algorithms and protocols to facilitate widely and densely distributed query processing on large scale networks. This project's key focus is to determine how to overcome structural mismatches between the application layer, and the network layer, resulting in a network-aware application deployment. My contributions included designing and implementing a distributed operator placement algorithm. I deployed our protocol across PlanetLab, and validated the

results with actual continuous queries, leveraging Borealis as an underlying query processing engine. I also investigated partitioning and replication mechanisms to distribute and parallelize an operator across a wide-area network by exploiting three types of locality, namely data, spatial and temporal locality.

Borealis Stream Processing Engine – <http://www.cs.brown.edu/research/db/borealis/>

The Borealis project extends first-generation SPEs semantically, and for distributed operation. My involvement on the Borealis prototype has included the implementation of several core distribution mechanisms, namely how to move operators between sites, stream processing under dynamic query modification and presently how to split operators for parallelization and localization. For applications, I am investigating a continuous query approach to a networked multiplayer game and have integrated an existing open-source game (Cube, <http://www.cubeengine.com/>) to employ Borealis queries to execute the game server application logic.

XPORT: Extensible Profile-Driven Overlay Routing Trees

XPORT presents the design and implementation of a data dissemination infrastructure that supports an extensible set of data types and profiles, and an extensible optimization framework. XPORT provides a core API that allows the user to fill in the dissemination application logic, and executes this logic on a tree-based overlay network. XPORT also allows users to customize the optimization metric, and proceeds to minimize this objective through a sequence of local, adaptive overlay transformations, triggered by changing resource conditions.

Conference and Workshop Publications

Simultaneous Equation Systems for Query Processing on Continuous-Time Data Streams.

Yanif Ahmad, Olga Papaemmanouil, Uğur Çetintemel, Jennie Rogers.

Proceedings of the 24th International Conference on Data Engineering, ICDE '08, to appear.

COLR-Tree: Communication Efficient Spatio-Temporal Index for a Sensor Data Web Portal.

Yanif Ahmad, Suman Nath.

Proceedings of the 24th International Conference on Data Engineering, ICDE '08, to appear.

Declarative Temporal Data Models for Sensor-Driven Query Processing.

Yanif Ahmad, Uğur Çetintemel.

Proceedings of the 4th International Workshop on Data Management in Sensor Networks, DMSN '07.

XPORT: Extensible Profile-Drive Overlay Routing Trees.

Olga Papaemmanouil, Yanif Ahmad, Uğur Çetintemel, John Jannotti.

Proceedings of the 26th ACM SIGMOD Conference, SIGMOD '06.

Locality-Aware Networked Join Evaluation.

Yanif Ahmad, Uğur Çetintemel, John Jannotti, Alexander Zgolinski.

Proceedings of the 1st IEEE International Workshop on Networking Meets Databases, NetDB '05.

Network Awareness in Internet Scale Stream Processing.

Yanif Ahmad, Uğur Çetintemel, John Jannotti, Alexander Zgolinski, Stan Zdonik.

IEEE Bulletin of the Technical Committee on Data Engineering, March 2005, Vol 28, No. 1.

The Design of the Borealis Stream Processing Engine.

with the Borealis Team.

Proceedings of the 2nd Biennial Conference on Innovative Database Systems, CIDR '05.

Network-Aware Query Processing for Stream Based Applications.

Yanif Ahmad, Uğur Çetintemel.

Proceedings of the 30th International Conference on Very Large Data Bases, VLDB '04.

A Type System for Statically Detecting Spreadsheet Errors.

Yanif Ahmad, Tudor Antoniu, Sharon Goldwater, Shriram Krishnamurthi.

Proceedings of the 18th IEEE International Conference on Automated Software Engineering, ASE '03.

Exploiting Precision vs. Efficiency Tradeoffs in Symmetric Replication Environments (abstract).
Uğur Cetintemel, Peter Keleher, Yanif Ahmad.
Proceedings of the 21st ACM Symposium on Principles of Distributed Computing, PODC '02.

Demonstrations

The Borealis Distributed Stream Processing Engine.
with the Borealis Team.

Proceedings of the 2nd International Conference on GeoSensor Networks (GSN '06, invited).

XPORT: Extensible Profile-driven Overlay Routing Trees.

Olga Papaemmanouil, Yanif Ahmad, Ugur Cetintemel, John Jannotti, Yenel Yildirim.

Proceedings of the 26th ACM SIGMOD Conference (SIGMOD '06).

Distributed Operation in the Borealis Stream Processing Engine.

with the Borealis Team.

Proceedings of the 24th ACM SIGMOD International Conference on Management of Data, 2005.
(Best Demonstration Award)

Book Chapters and Articles

1. Data streams: Architectures and prototypes.

2. Data streams: Streaming applications.

Yanif Ahmad, Uğur Çetintemel.

Encyclopedia of Database Systems.

L. Liu, M. Tamer Ozsü (editors), Springer.

Load Management and High Availability in the Borealis Distributed Stream Processing Engine.

Nesime Tatbul, Yanif Ahmad, Ugur Cetintemel, Jeong-Hyon Hwang, Ying Xing, Stan Zdonik.

Advances in Geosensor Networks.

S. Nittel, A. Labrinidis, A. Stefanidis (editors), LNCS, Springer-Verlag.

The Aurora and Borealis Stream Processing Engines.

with the Borealis Team.

Data Stream Management: Processing High-Speed Data Streams.

M. Garofalakis, J. Gehrke, R. Rastogi (editors), Springer-Verlag, July 2006.

Development Tools Experience

Languages - OCaml, C++, Java, C, C#, Python, Perl, TCL, Prolog, Matlab.

Databases - Microsoft SQL Server, PostgreSQL, BerkeleyDB, MySQL.

Networking - PlanetLab, Emulab, OpenDHT, Vivaldi, NS-2, GT-ITM, XML-RPC, IMAP, RSS.

Academic Experience

Teaching Assistant CS161 - Building High Performance Servers. Fall 2004.

Advising: Jennie Rogers (Sc.M.), Alexander Zgolinski (Sc.M.), George Kong (Sc.M.).

External reviewer: TKDE, SIGMOD 2008, ICDE 2008, VLDB 2007, ICDCS '06, ICDE '06,
SIGMOD '04, ICDE '04, ICDCS '03, CIKM '02.

Student volunteer: ICDE '04, ICDCS '03.

Personal

Date, and place of birth: 11th July 1980, London, UK.

Citizenship: UK.

Visa status: F-1

References

On request.