

# Roshan Dathathri

---

CONTACT	<i>Website:</i> <a href="https://roshandathathri.github.io/">https://roshandathathri.github.io/</a> <i>Email:</i> <a href="mailto:roshan.dathathri@gmail.com">roshan.dathathri@gmail.com</a>
RESEARCH INTERESTS	My goal is to build <b>programming languages, compilers, and runtime systems</b> for efficient and secure <b>AI systems</b> that make it easy to develop and deploy end-to-end AI pipelines on large-scale <i>distributed</i> clusters, while utilizing <i>heterogeneous</i> architectures.
EDUCATION	<p><b>The University of Texas</b>, Austin, Texas, USA.      <i>August 2014 to May 2020</i></p> <p><b>Ph. D.</b>, Dept. of Computer Science</p> <ul style="list-style-type: none"><li>• <i>Thesis Advisor:</i> Dr. Keshav Pingali</li><li>• <i>Thesis Title:</i> Programming Systems for Graph Analytics and Homomorphic Encryption on Distributed and Heterogeneous Architectures</li><li>• Designed and built graph analytics systems that exploit domain knowledge to partition graphs and optimize communication, while providing application-specific fault-tolerance. Existing shared-memory graph analytics frameworks or applications can use our system to scale out to distributed CPUs and GPUs.</li><li>• Designed and helped build a high-level programming language and runtime system that can execute graph pattern mining on a multi-core CPU or a GPU.</li></ul> <p><b>Indian Institute of Science</b>, Bangalore, Karnataka, India.      <i>August 2011 to July 2014</i></p> <p><b>M.Sc. (Engg.)</b>, Dept. of Computer Science and Automation</p> <ul style="list-style-type: none"><li>• <i>Thesis Advisor:</i> Dr. Uday Bondhugula</li><li>• <i>Thesis Title:</i> Compiling for a Dataflow Runtime on Distributed-Memory Parallel Architectures</li><li>• Designed and built compiler techniques using the polyhedral model to automatically extract tasks from sequential affine loop nests and dynamically schedule them to run on distributed CPUs and GPUs with efficient data movement code.</li></ul>
PROFESSIONAL EXPERIENCE	<p><b>Katana Graph</b>, Austin, Texas, USA      <i>June 2020 to present</i></p> <ul style="list-style-type: none"><li>• <i>Designation:</i> Principal Engineer</li><li>• Led the graph engine team to build a distributed cloud platform for computing AI, analytics, and queries efficiently on large-scale graphs.</li><li>• Interacted with customers, product, and leadership to help drive the vision.</li><li>• Designed and built the distributed graph querying engine that minimizes latency of business intelligence queries and scales well on distributed hosts.</li><li>• Designed and built an interface for a general, sparse collective communication while supporting a protocol to recover from control divergence among distributed hosts.</li><li>• Led the design of a high-level graph API for orchestration of graph neural networks, graph recommendation systems, graph analytics, graph queries, and graph updates.</li><li>• Helped design an in-memory and on-storage log-structured representation for labelled property graphs that is compact for sparse node and edge properties, while being efficient for both reading and updating the graph topology and properties.</li><li>• Designed a distributed node property map that is optimized for highly concurrent sparse reductions by using a graph-partition-aware sparse representation and by avoiding thread conflicts.</li></ul> <p><b>Microsoft Research</b>, Redmond, Washington, USA      <i>Summer 2018 and Summer 2019</i></p> <ul style="list-style-type: none"><li>• <i>Designation:</i> Research Intern</li><li>• Designed and developed an optimizing compiler for translating tensor programs like neural network inferencing to run on encrypted data using fully-homomorphic encryption (FHE) efficiently, while guaranteeing security and accuracy of the computation.</li><li>• Designed a new encrypted vector arithmetic language for developing general-purpose FHE applications and built an optimizing compiler that generates correct and secure programs, while hiding all the complexities of the target FHE scheme.</li></ul> <p><b>National Instruments</b>, Bangalore, Karnataka, India.      <i>July 2008 to July 2011</i></p> <ul style="list-style-type: none"><li>• <i>Designation:</i> Staff Software Engineer</li></ul>

- Designed and developed **compiler optimizations** on the high-level intermediate representation and the LLVM IR for the graphical programming language, LabVIEW.

## PUBLICATIONS

- X. Chen, R. Dathathri, G. Gill, L. Hoang, K. Pingali, “*Sandslash: A Two-Level Framework for Efficient Graph Pattern Mining*”, 35th ACM International Conference on Supercomputing (ICS), June 2021.
- G. Gill, R. Dathathri, S. Maleki, M. Musuvathi, T. Mytkowicz, and O. Saarikivi, “*Distributed Training of Embeddings using Graph Analytics*”, 35th IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2021.
- H. Lee, D. Wong, L. Hoang, R. Dathathri, G. Gill, V. Jatala, D. Kuck, and K. Pingali, “*A Study of APIs for Graph Analytics Workloads*”, Proceedings of the 2020 IEEE International Symposium on Workload Characterization (IISWC), October 2020.
- A. Azad, M. M. Aznaveh, S. Beamer, M. Blanco, J. Chen, L. D’Alessandro, R. Dathathri, T. Davis, K. Dewese, J. Firoz, H. A Gabb, G. Gill, B. Hegyi, S. Kolodzie, T. M. Low, A. Lumsdaine, T. Manlaibaatar, T. G Mattson, S. McMillan, R. Peri, K. Pingali, U. Sridhar, G. Szarnyas, Y. Zhang, Y. Zhang, “*Evaluation of Graph Analytics Frameworks Using the GAP Benchmark Suite*”, Proceedings of the 2020 IEEE International Symposium on Workload Characterization (IISWC), October 2020.
- R. Dathathri, B. Kostova, O. Saarikivi, W. Dai, K. Laine, M. Musuvathi, “*EVA: An Encrypted Vector Arithmetic Language and Compiler for Efficient Homomorphic Computation*”, 41st ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI), June 2020.
- V. Jatala, R. Dathathri, G. Gill, L. Hoang, V. K. Nandivada, K. Pingali, “*A Study of Graph Analytics for Massive Datasets on Distributed GPUs*”, 34th IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2020.
- X. Chen, R. Dathathri, G. Gill, K. Pingali, “*Pangolin: An Efficient and Flexible Graph Pattern Mining System on CPU and GPU*”, 46th International Conference on Very Large Data Bases (PVLDB), April 2020.
- G. Gill, R. Dathathri, L. Hoang, R. Peri, K. Pingali, “*Single Machine Graph Analytics on Massive Datasets Using Intel Optane DC Persistent Memory*”, 46th International Conference on Very Large Data Bases (PVLDB), April 2020.
- R. Dathathri, G. Gill, L. Hoang, H. Dang, V. Jatala, V. K. Nandivada, M. Snir, K. Pingali, “*Gluon-Async: A Bulk-Asynchronous System for Distributed and Heterogeneous Graph Analytics*”, 28th IEEE International Conference on Parallel Architectures and Compilation Techniques (PACT), September 2019.
- L. Hoang, V. Jatala, X. Chen, U. Agarwal, R. Dathathri, G. Gill, K. Pingali, “*DistTC: High Performance Distributed Triangle Counting*”, 23rd IEEE International Conference on High Performance Extreme Computing (HPEC), September 2019.
- R. Dathathri, O. Saarikivi, H. Chen, K. Laine, K. Lauter, S. Maleki, M. Musuvathi, T. Mytkowicz, “*CHET: An Optimizing Compiler for Fully-Homomorphic Neural-Network Inferencing*”, 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI), June 2019.
- L. Hoang, R. Dathathri, G. Gill, K. Pingali, “*CuSP: A Customizable Streaming Edge Partitioner for Distributed Graph Analytics*”, 33rd IEEE International Parallel and Distributed Processing Symposium (IPDPS), May 2019.
- R. Dathathri, G. Gill, L. Hoang, K. Pingali, “*Phoenix: A Substrate for Resilient Distributed Graph Analytics*”, 24th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), April 2019.
- L. Hoang, M. Pontecorvi, R. Dathathri, G. Gill, B. You, K. Pingali, V. Ramachandran, “*A Round-Efficient Distributed Betweenness Centrality Algorithm*”, 24th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP), February 2019.
- G. Gill, R. Dathathri, L. Hoang, K. Pingali, “*A Study of Partitioning Policies for Graph Analytics on Large-scale Distributed Platforms*”, 45th International Conference on Very Large Data Bases (PVLDB), 12(4), 2018.

G. Gill, R. Dathathri, L. Hoang, A. Lenharth, K. Pingali, “*Abelian: A Compiler and Runtime for Graph Analytics on Distributed, Heterogeneous Platforms*”, 24th International European Conference on Parallel and Distributed Computing (**Euro-Par**), August 2018.

R. Dathathri, G. Gill, L. Hoang, H. Dang, A. Brooks, N. Dryden, M. Snir, K. Pingali, “*Gluon: A Communication-Optimizing Substrate for Distributed Heterogeneous Graph Analytics*”, 39th ACM SIGPLAN Conference on Programming Language Design and Implementation (**PLDI**), June 2018.

H. Dang, R. Dathathri, G. Gill, A. Brooks, N. Dryden, A. Lenharth, L. Hoang, K. Pingali, M. Snir, “*A Lightweight Communication Runtime for Distributed Graph Analytics*”, 32nd IEEE International Parallel and Distributed Processing Symposium (**IPDPS**), May 2018.

A. Ulanov, M. Marwah, M. Kim, R. Dathathri, C. Zubieta, J. Li, “*Sandpiper: Scaling Probabilistic Inference to Large Scale Graphical Models*”, IEEE International Conference on **Big Data**, December 2017.

R. Dathathri, R. T. Mullapudi, and U. Bondhugula, “*Compiling Affine Loop Nests for a Dynamic Scheduling Runtime on Shared and Distributed Memory*”, ACM Transactions on Parallel Computing (**TOPC**), August 2016.

M. Ravishankar, R. Dathathri, V. Elango, L. Pouchet, J Ramanujam, A. Rountev, P Sadayappan, “*Distributed memory code generation for mixed irregular/regular computations*”, 20th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (**PPoPP**), January 2015.

R. Dathathri, C. Reddy, T. Ramashekar, U. Bondhugula, “*Generating Efficient Data Movement Code for Heterogeneous Architectures with Distributed-Memory*”, 22nd ACM International Conference on Parallel Architectures and Compilation Techniques (**PACT**), September 2013.

## AWARDS

### **INCITS Team Award,**

*Expert Group on SQL Extensions for Property Graphs*,  
International Committee for Information Technology Standards (**INCITS**), June 2023.

### **Best Paper Nominee,**

*Gluon-Async: A Bulk-Asynchronous System for Distributed Heterogeneous Graph Analytics*,  
IEEE Parallel Architectures and Compilation Techniques (**PACT**), September 2019.

### **Student Innovation Award,**

*DistTC: High Performance Distributed Triangle Counting*,  
IEEE High Performance Extreme Computing (**HPEC**) GraphChallenge, September 2019.

## SERVICE

*Program Committee (PC)*

**IPDPS** (International Parallel and Distributed Processing Symposium): 2021, 2022, 2023  
**PACT** (Parallel Architectures and Compilation Techniques): 2022

*External Review Committee (ERC)*

**ASPLOS**(Architectural Support for Programming Languages and Operating Systems): 2021

*Artifact Evaluation Committee (AEC)*

**CGO** (Code Generation and Optimization): 2018

*Reviewer*

**ASPLOS**(Architectural Support for Programming Languages and Operating Systems): 2018

**ISCA** (International Symposium on Computer Architecture): 2019

**PoPL** (Principles of Programming Languages): 2020

**PPoPP** (Principles and Practice of Parallel Programming): 2017, 2020, 2021

**IPDPS** (International Parallel and Distributed Processing Symposium): 2018, 2019, 2020

**PACT** (Parallel Architectures and Compilation Techniques): 2017, 2019

**CGO** (Code Generation and Optimization): 2018

**TACO** (Transactions on Architecture and Code Optimization): 2020

## TECHNICAL SKILLS

*Programming* : C, C++, OpenMP, MPI, Python

*Applications* : Visual Studio Code, Vim, Git,  $\LaTeX$  2 $\epsilon$