

# SUDIPTO DAS

---

6750 El Colegio Rd Apt 328, Goleta, CA 93117, USA.

+1 (805) 403-1146 • sudipto@cs.ucsb.edu • sudiptdas@gmail.com • http://www.cs.ucsb.edu/~sudipto

---

## RESEARCH INTERESTS

---

Cloud data management, elastic and autonomic database systems, large scale distributed systems, transaction processing, scalable data analysis systems.

## EDUCATION

---

### PhD in Computer Science, UC Santa Barbara

• Graduated Dec 2011 • GPA: **3.98**

**Dissertation:** “*Scalable and Elastic Transactional Data Stores for Cloud Computing Platforms*”.

**Advised by:** Prof. Divyakant Agrawal and Prof. Amr El Abbadi

### Bachelor of Computer Science and Engineering, Jadavpur University, Kolkata, INDIA

• (First Class, Honors), Graduated June 2006 • GPA: **9.21/10**

---

## RELEVANT WORK EXPERIENCE

---

• **Microsoft Research, Redmond, WA** (June - September 2010)

**Research Intern, Mentor/Manager: Dr. Philip A. Bernstein**

Worked on the Hyder research project—a scale out transactional system for shared flash. Designed an end to end simulation for the entire system for thorough evaluation and performance analysis. Advanced the design with new algorithms and optimizations.

• **IBM Almaden Research Center, San Jose, CA** (June - September 2009)

**Research Intern, Mentor: Dr. Yannis Sismanis, Manager: Dr. John McPherson**

Designed a framework to enable *deep statistical analysis* over large amounts of data. Implemented the scalable analytics and modeling framework using R, Jaql, and Hadoop.

• **Google Inc., Mountain View, CA** (June - September 2007)

**Software Engineering Intern, Mentor: Shilpa Kolhar, Manager: Dr. Johnny Chen**

Worked as a Software Engineering Intern in team **LocalGIS** working towards importing Geo-Spatial data sets like Crime, Traffic Accidents, Parcels, Health, and Pesticide to name a few.

• **UC Santa Barbara - Department of Computer Science** (Fall 2006 - Present)

**Graduate Research Assistant and Teaching Assistant**

• **IBM India Pvt. Ltd., Kolkata, INDIA** (June - July 2005)

**Summer Intern, Mentor: Dr. Anup K. Ghosh, Manager: Dr. Amitava Mukherjee**

Worked towards integrating an SIP phone into the SAP Enterprise portal and worked in J2EE based architecture.

---

## SELECTED PUBLICATIONS

---

• “*Albatross: Lightweight Elasticity in Shared Storage Databases for the Cloud using Live Data Migration*”

**Sudipto Das**, Shoji Nishimura, Divyakant Agrawal, Amr El Abbadi. *In the 37<sup>th</sup> International Conference on Very Large Databases (VLDB) 2011.*

• “*Zephyr: Live Migration in Shared Nothing Databases for Elastic Cloud Platforms*”

Aaron Elmore, **Sudipto Das**, Divyakant Agrawal, Amr El Abbadi. *In the ACM International Conference on Management of Data (SIGMOD) 2011.*

• “*Hyder – A Transactional Record Manager for Shared Flash*”

Philip Bernstein, Colin Reid, **Sudipto Das**. *In the 5<sup>th</sup> Biennial Conference on Innovative Data Research (CIDR) 2011 [Recipient of Best Paper Award].*

• “*MD-HBase: A Scalable Multi-dimensional Data Infrastructure for Location Aware Services*”

Shoji Nishimura, **Sudipto Das**, Divyakant Agrawal, Amr El Abbadi. *In 12th International Conference on Mobile Data Management (MDM) 2011[Recipient of Best Paper Runner-up Award].*

• “*G-Store: A Scalable Data Store for Transactional Multi key Access in the Cloud*”

**Sudipto Das**, Divyakant Agrawal, Amr El Abbadi. *In the 1<sup>st</sup> ACM International Symposium on Cloud Computing (SoCC) 2010.*

• “*Ricardo: Integrating R and Hadoop*”

**Sudipto Das**, Yannis Simanis, Kevin S. Beyer, Rainer Gemulla, Peter J. Haas, John McPherson. *In the ACM International Conference on Management of Data (SIGMOD) 2010.*

---

- “*Big Data and Cloud Computing: Current State and Future Opportunities*”  
Divyakant Agrawal, **Sudipto Das**, Amr El Abbadi. *Tutorial presentation at the 14<sup>th</sup> International Conference on Extending Database Technology (EDBT) 2011.*
- “*Big Data and Cloud Computing: New Wine or just New Bottles?*”  
Divyakant Agrawal, **Sudipto Das**, Amr El Abbadi. *Tutorial presentation at the 36<sup>th</sup> International Conference on Very Large Databases (VLDB) 2010.*
- “*ElasTraS: An Elastic, Scalable, and Self Managing Transactional Database for the Cloud*”  
**Sudipto Das**, Shashank Agarwal, Divyakant Agrawal, Amr El Abbadi. *UCSB Technical Report CS 2010-04.*
- “*ElasTraS: An Elastic Transactional Data Store in the Cloud*”  
**Sudipto Das**, Divyakant Agrawal, Amr El Abbadi. *In the 1<sup>st</sup> Usenix Workshop on Hot topics on Cloud Computing (HotCloud) 2009.*
- “*Database Scalability, Elasticity, and Autonomy in the Cloud*”  
Divyakant Agrawal, Amr El Abbadi, **Sudipto Das**, Aaron Elmore. *In the 16<sup>th</sup> International Conference on Database Systems for Advanced Applications (DASFAA) 2011 (Invited Paper).*
- “*Data Management Challenges in Cloud Computing Infrastructures*”  
Divyakant Agrawal, Amr El Abbadi, Shyam Antony, **Sudipto Das**. *In the 6<sup>th</sup> International Workshop on Databases in Networked Information Systems (DNIS) 2010 (Invited Paper).*
- “*Anónimos: An LP based Approach for Anonymizing Weighted Social Network Graphs*”  
**Sudipto Das**, Ömer Eğecioğlu, Amr El Abbadi. *To appear in the IEEE Transactions on Knowledge and Data Engineering (TKDE).*
- “*Anonymizing Weighted Social Network Graphs*”  
**Sudipto Das**, Ömer Eğecioğlu, Amr El Abbadi. *In the 26<sup>th</sup> International Conference on Data Engineering (ICDE) 2010 (Short Paper).*
- “*Thread Cooperation in Multicore Architectures for Frequency Counting Over Multiple Data Streams*”  
**Sudipto Das**, Shyam Antony, Divyakant Agrawal, Amr El Abbadi. *In the 35<sup>th</sup> International Conference on Very Large Databases (VLDB) 2009.*
- “*CoTS: A Scalable Framework for Parallelizing Frequency Counting over Data Streams*”  
**Sudipto Das**, Shyam Antony, Divyakant Agrawal, Amr El Abbadi. *In the 25<sup>th</sup> International Conference on Data Engineering (ICDE) 2009 (Short Paper).*
- “*QUORUM: Quality of Service in Wireless Mesh Networks*”  
Vinod Kone, **Sudipto Das**, Ben Y. Zhao, Haitao Zheng. *In Mobile Networks and Applications (MONET) Journal 12(5-6), April 2008.*

---

## HONORS AND AWARDS

- 2011 Outstanding Student Award in Computer Science, UC Santa Barbara.
- Best Paper award at the 5<sup>th</sup> biennial Conference on Innovative Data Systems Research (CIDR) 2011.
- Best Paper Runner-up award at the 12<sup>th</sup> International Conference on Mobile Data Management (MDM) 2011.
- Best Reviewer Award for UCSB GSWC 2011.
- TCS-JU Best Graduating Student Award for 2006, Computer Science Dept., Jadavpur University, Kolkata.
- UCSB Graduate Division Dissertation Fellowship for Winter 2011.
- Outstanding Teaching Assistant for Graduate Course on Distributed Systems, Fall '08.
- UC Senate Travel Award for VLDB 2009 & ACM SoCC Student Travel Award 2010.
- UC President’s Work Study award.

---

## SELECTED TALKS

- “*Scalable and Elastic Transactional Data Stores for Cloud Computing Platforms,*” PhD Defense.
- “*Scalable, Consistent, and Elastic Database Systems for Cloud Platforms,*” MSR, IBM Almaden, IBM Watson, Yahoo! Research, NEC Labs America.
- “*Albatross: Lightweight Elasticity in Shared Storage Databases for the Cloud using Live Data Migration,*” At VLDB 2011.
- “*Zephyr: Live Migration in Shared Nothing Databases for Elastic Cloud Platforms,*” At SIGMOD 2011.
- “*Big Data and Cloud Computing: Current State and Future Opportunities,*” Tutorial at EDBT 2011.
- “*Big Data and Cloud Computing: New Wine or just New Bottles?,*” Tutorial at VLDB 2010.

- “*The Performance of Hyder: A Novel Transactional Record Manager for Shared Raw Flash*,” *Microsoft Research, Redmond, WA, September 2010*.
- “*G-Store: A Scalable Data Store for Transactional Multi key Access in the Cloud*,” *SoCC 2010*.
- “*Ricardo: Integrating R and Hadoop*,” *SIGMOD 2010*.
- “*RJaql: A Scalable Platform for Analytics and Modeling in the Cloud*,” *IBM Almaden Research Center, San Jose, CA, September 2009*.
- “*Anonymizing Weighted Social Network Graphs*,” *ICDE 2010*.
- “*Thread cooperation in multicore architectures for frequency counting over multiple data streams*,” *VLDB '09*.

## TECHNICAL SKILLS

---

- Proficient in software development in Java, C, C++.
- Familiarity with concurrent programming, distributed systems, database system implementations, SQL, HTML, Linux Shell Programming, Perl,  $\LaTeX$ .
- Experienced in Linux, Windows, and Solaris environments.
- Experienced with MapReduce, Hadoop, HBase, Hive, Jaql, Matlab, R.

## PROFESSIONAL ACTIVITIES

---

- Student member of the ACM.
- Reviewer for VLDB Journal, IEEE Trans. of Know. and Data Engg., Journal of Dist. and Par. Databases, IEEE Internet Computing, IEEE Computer, SoCC 2011.
- Program Committee member: ICDE 2012 Workshop on Data Management in the Cloud, UCSB Graduate Students Workshop 2010 & 2011.
- External Reviewer for SIGMOD 2008, CIKM 2008, ICDE 2009 and 2010, EDBT 2009, SSDBM 2009 and 2011, DaMoN 2009, SIGSPATIAL 2009, LADIS 2010, TODS, VLDB 2011 & 2012, EuroPar 2011.
- Helped organize the NSF Workshop “Science of Cloud” held in March 2011.
- Served as Graduate Student Association representative in departmental Colloquium committee '09 and '10, Faculty search committee '11.
- Mentored a high school student with UCSB Summer Sessions in Summer '08, four undergraduate projects, and one MS project.
- Conferences Attended: SIGMOD '08, '10, and '11, VLDB '09, '10, and '11, ICDE '10, SoCC '10, EDBT '11.

## REFERENCES

---

Available on request.