

# Tu Mai Anh Do

## CONTACT INFORMATION

---

FULL BIRTH NAME:	Tu Mai Anh Do	EMAIL:	<a href="mailto:tudo@isi.edu">tudo@isi.edu</a>
DATE OF BIRTH:	21 May 1993	ADDRESS:	Room 1220, 4676 Admiralty Way, Suite 1001
NATIONALITY:	Vietnamese		Marina del Rey, California 90292, USA

## RESEARCH INTERESTS

---

In situ Processing, Big Data Analytics, Scientific Workflows, Workflow Managment System, High Performance Computing, Distributed Systems

## EDUCATION

---

AUG. 2017	UNIVERSITY OF SOUTHERN CALIFORNIA (USC), Los Angeles, California, USA
:	<i>Viterbi School of Engineering</i>
:	Ph.D. Candidate in Computer Science
Now	<i>Advisor:</i> Ewa Deelman
SEP. 2011	HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY (HCMUT), Ho Chi Minh, Vietnam
:	<i>Faculty of Computer Science and Engineering</i>
:	B.E. in Computer Engineering, Honors Program
APRIL. 2016	<i>GPA:</i> 8.55/10.00

## EXPERIENCE

---

AUG. 2017	INFORMATION SCIENCES INSTITUTE, Marina Del Rey, California
:	<i>Graduate Research Assistant</i>
:	<i>Advisor:</i> Ewa Deelman
:	Science Automation Technologies Group
:	Working on In Situ Data Analytics for Next Generation Molecular Dynamics Workflows project ( <a href="#">Analytics4MD</a> is funded by NSF (National Science Foundation))
NOW	Tools: DataSpaces, ADIOS, Decaf, Pegasus, Cmake, Shippable, Gromacs, Plumed2, A4MD, TAU
MAY. 2018	LAWRENCE LIVERMORE NATIONAL LABORATORY, Livermore, California
:	<i>Student Intern</i>
:	<i>Advisor:</i> Ming Jiang
:	Computation Directorate/Institute for Scientific Computing Research
:	Enabled data analytics workflows that couple high-performance computing simulations with Big Data analytics using node-local storage
AUG. 2018	Tools: Spark, Ascent, Pegasus
OCT. 2014	HIGH PERFORMANCE COMPUTING CENTER, Ho Chi Minh City University of Technology
:	<i>Research Assistant</i>
:	<i>Advisor:</i> Nam Thoai
:	Studied and developed abnormal behavior detection techniques for large-scale parallel applications in message-passing programming model
JUL. 2017	Tools: C, C++, MPI, Makefile, Shell Script
MAR. 2017	NOVOBI VIETNAM, Ho Chi Minh City
:	<i>Part-time Software Engineer</i>
:	Built automated system for deploying, testing and delivering package of health care application
JUL. 2017	Tools: Batch Script, Powershell Script, Atlassian Bamboo, Visual Studio, SQL Server Management Studio
MAY. 2015	DEK TECHNOLOGY VIETNAM, Ho Chi Minh City
:	<i>Summer Intern</i>
:	Built and automate deploy small high availability cluster
AUGUST. 2015	Tools: Shell Script, Git

## TEACHING EXPERIENCE

---

FALL 2020	UNIVERSITY OF SOUTHERN CALIFORNIA (USC), Los Angeles, California, USA <i>Teaching Assistant</i> Courses: <i>CSCI 585 Database Systems</i> <i>Instructor: Saty Raghavachary</i>
SEP. 2016	HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY, Ho Chi Minh City <i>Teaching Assistant</i>
⋮	Courses: <i>Parallel Programming and Distributed Systems, Fundamentals of Programming</i>
MAY. 2017	<i>Instructors: Nam Thoai, Sach Le</i>

## NOTABLE PROJECTS

---

- **HCMUT B.E Graduation Thesis**

*Title:* Developing Methods To Help Large-Scale Parallel Applications More Reliable

*Advisor:* Nam Thoai

*Description:* This thesis explored the idea that scalability is the main cause of suffering high probability of error in large-scale parallel applications. To help the applications more reliable for better performance goal, the technique was proposed by detecting abnormal behaviors, which are considered as highly vulnerable from errors.

*Score:* 9.96/10.00 – Highest score among concurrent defended theses

- **[Analytics4MD](#)**

*Name:* In Situ Data Analytics for Next Generation Molecular Dynamics Workflows

*Advisor:* Ewa Deelman

*Description:* Molecular dynamics simulations studying the classical time evolution of a molecular system at atomic resolution are widely recognized in the fields of scientific computing. Moving to exascale requires the simulations to analyze data as it is generated and store only necessary data. The analysis need to perform in-situ. We propose a paradigm for transforming the centralized, off-line nature of the molecular dynamics analysis into performing in-situ processing via in-memory staging area or in-transit processing via optimized parallel file systems or emerging burst buffer. This effort provides the ability to interleave simulations and analytics to improve data analyzing performance and processing more data.

## ACTIVITIES

---

- *Attended* Student Volunteers program of 2019 SUPERCOMPUTING (SC19), November 16th, 2019 - November 22nd, 2019
- *Attended* 2019 ESCIENCE CONFERENCE, September 24th, 2019 - September 27th, 2019
- *Presented* a poster titled "Enabling Data Analytics Workflows using Node-Local Storage" at the USC ISI Graduate Student Symposium 2019, Marina Del Rey, California, March 26, 2019
- *Attended* 2018 SUPERCOMPUTING (SC18) as a Student Volunteer and *presented* an accepted poster titled "Enabling Data Analytics Workflows using Node-Local Storage", Nov 10, 2018 - Nov 16, 2018
- *Visited* RUTGERS DISCOVERY INFORMATICS INSTITUTE (*RDI*<sup>2</sup>), New Brunswick, New Jersey, May 13th, 2018 - May 17th, 2018
- *Presented* a poster titled "In Situ Data Analytics for Next Generation Molecular Dynamics Workflows" at the USC ISI Graduate Student Symposium 2018, Marina Del Rey, California, April 5th, 2018

## AWARDS AND HONORS

---

- 2017 | ISI Distinguished Top-Off Fellowship
- 2016 | Recommended Candidate of [2017 VEFSTA Fellowship Program](#)
- 2016 | 18th Eureka Scientific Research Student Award Finalist
- 2015 | 7th HCMC Information and Communication Technology Award for Student
- 2014 | DATALOGIC Vietnam's Scholarship, CSC Vietnam's Scholarship

## PUBLICATIONS

---

- **T. M. A. Do**, L. Pottier, S. Thomas, R. Ferreira da Silva, M. A. Cuendet, H. Weinstein, T. Estrada, M. Taufer, and E. Deelman, "A Novel Metric to Evaluate In Situ Workflows," in International Conference on Computational Science (ICCS), 2020.
- S. Thomas, M. Wyatt, **T. M. A. Do**, L. Pottier, R. Ferreira da Silva, H. Weinstein, M. A. Cuendet, T. Estrada, E. Deelman, and M. Taufer, "Characterization of In Situ and In Transit Analytics of Molecular Dynamics Simulations for Next-generation Supercomputers," in 15th eScience Conference, 2019.
- R. Ferreira da Silva, S. Callaghan, **T. M. A. Do**, G. Papadimitriou, and E. Deelman, "Measuring the Impact of Burst Buffers on Data-Intensive Scientific Workflows," Future Generation Computer Systems, vol. 101, p. 208–220, 2019.
- **T. M. A. Do**, M. Jiang, B. Gallagher, A. Chu, C. Harrison, K. Vahi, and E. Deelman, "Enabling Data Analytics Workflows using Node-Local Storage," in The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC18), Poster, 2018.
- **T. M. A. Do**, D. Diep, and N. Thoai, "Message Leak Detection in Debugging Large-Scale Parallel Applications," 2015 International Conference on Advanced Computing and Applications (ACOMP), 2015
- **T. M. A. Do**, D. Diep, and N. Thoai, "Race Condition and Deadlock Detection for Large-Scale Applications," 2016 15th International Symposium on Parallel and Distributed Computing (ISPDC), 2016.