



Environment Computing at HCMUT

Nam Thoai

namthoai@hcmut.edu.vn

HPC Lab (<http://www.hpcc.hcmut.edu.vn/>)

Faculty of Computer Science and Engineering

HCMC University of Technology (Bach Khoa University - BKU)

<http://www.cse.hcmut.edu.vn/>



Contents

- Motivation
- What we do





Motivation

- HCM city
 - Big plan for Smart city
 - Problems
 - (1) Traffic jam
 - (2) Urban flooding
 - (3) Pollution: waster & air
- Mekong delta
 - Flooding
 - Salinization
 - Healthy

Traffic in HCMC

■ Big data

- Motorbike: 5M
- Car: 5K
- 3.800 roads with the total length of 3.670km
- Realtime

■ Complexity

- Many motorbike
- Safety space
- Traffic coming from small routes



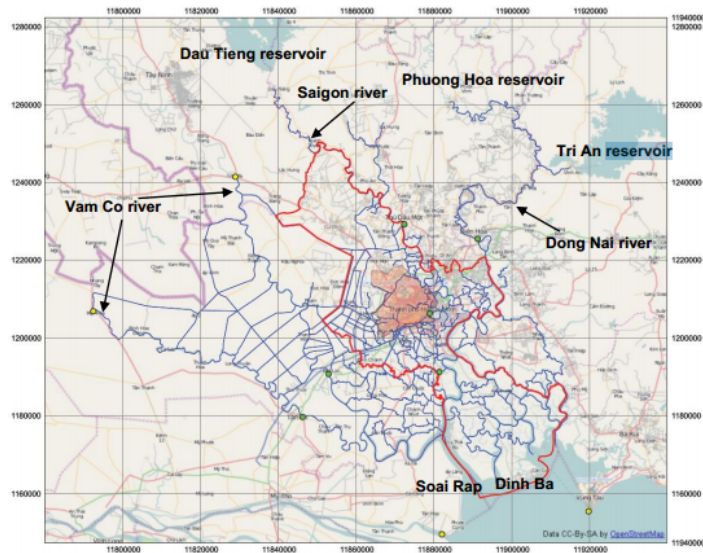
Urban flooding in HCMC



Street or River?



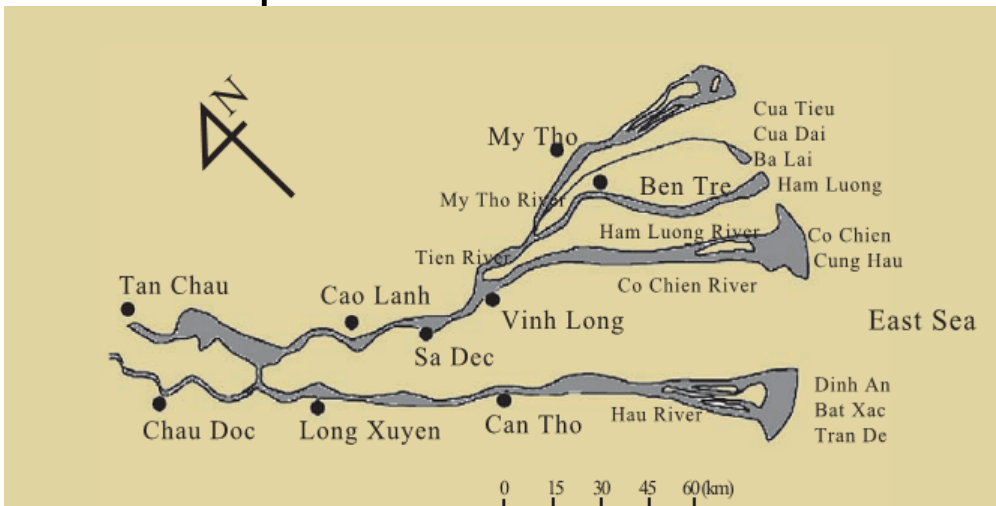
Airport?



Safety?

Problems in Mekong Delta

- Flooding
- Salinization
- Bank erosion
- Navigation hazards
- Pets
- Invasive exotic plants
- Healthy risks: water-borne, skin, mosquito-borne diseases





Activities at HCMUT



High Performance Computing

Partners

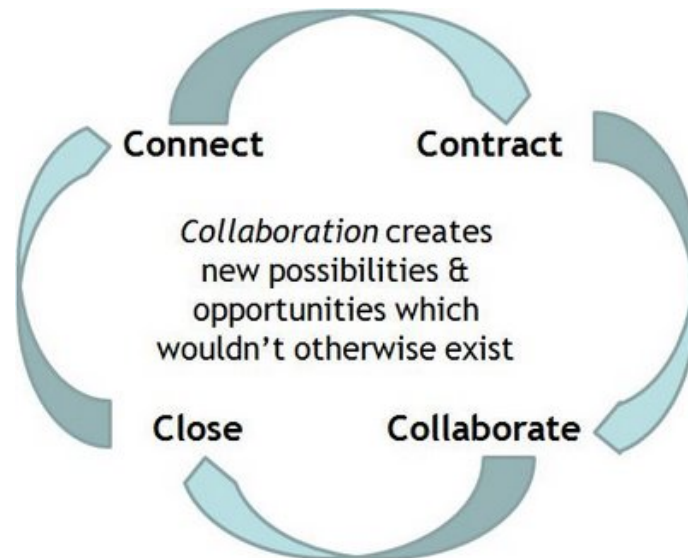
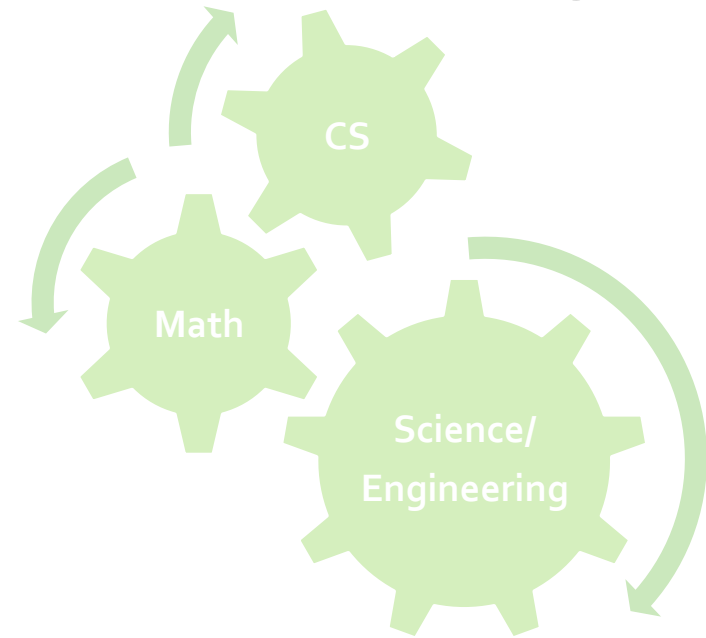
- HCMUT - VNU
- Intel
- Ho Chi Minh city

Plan: 2012-2022

- HPC research Lab: Set up in 2013
- HPC Center
- Strengthen HPC in Vietnam
- Solving big problems
- Leading in technology

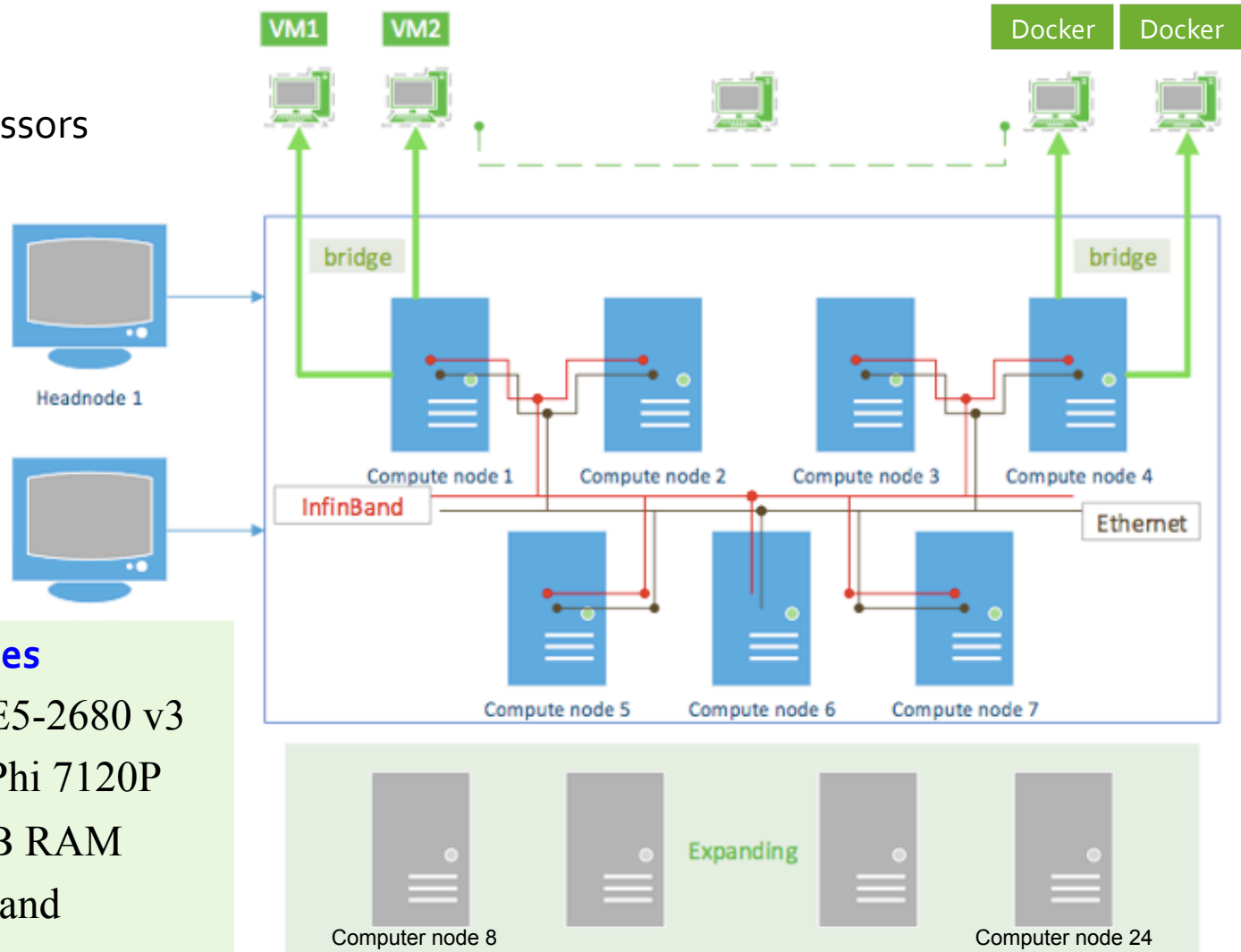
Applications: 2013-2022

- Traffic analysis
- Urban flooding
- Big data analytics



SuperNode-XP : 50 TFlops

- Vendors: HPE
- Intel Xeon processors
- Intel Xeon Phi
- Infiniband



24 Computing nodes

- 2 x Intel Xeon E5-2680 v3
- 2 x Intel Xeon Phi 7120P
- 512/256/128 GB RAM
- 50 Gbps Infiniband
- 2 TB hard disk/SSD



Libraries in SuperNode-XP

- ANSYS
 - Available 128 cores
- Cadence
- Open Telemac
- OpenFOAM
- Gromacs
- Gaussian
- BLAST
- Hadoop
- MPI
- Intel@ Parallel Studio XE
- Machine/Deep learning tools

Performance Evaluation of Environmental Applications using TELEMAC-MASCARET on Virtual Platforms

EWC workshop - The IEEE 12th International Conference on eScience, 23-27 Oct 2016, Baltimore, Maryland, USA

Minh Thanh Chung¹, Manh-Thin Nguyen¹, Nhu-Y Nguyen-Huynh¹,

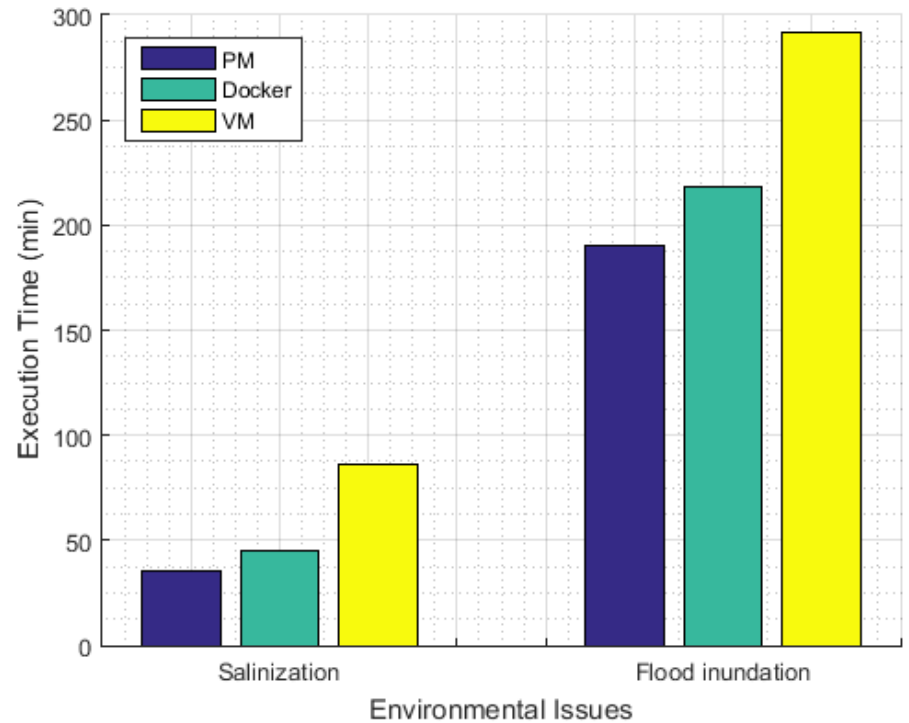
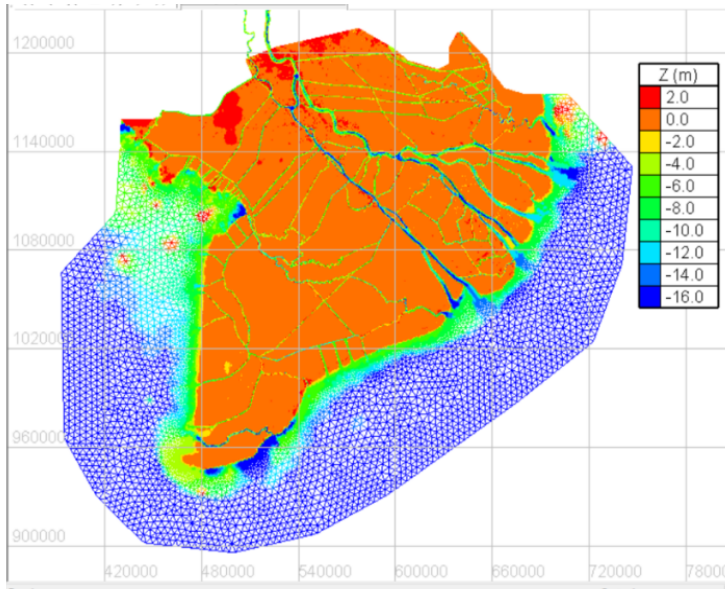
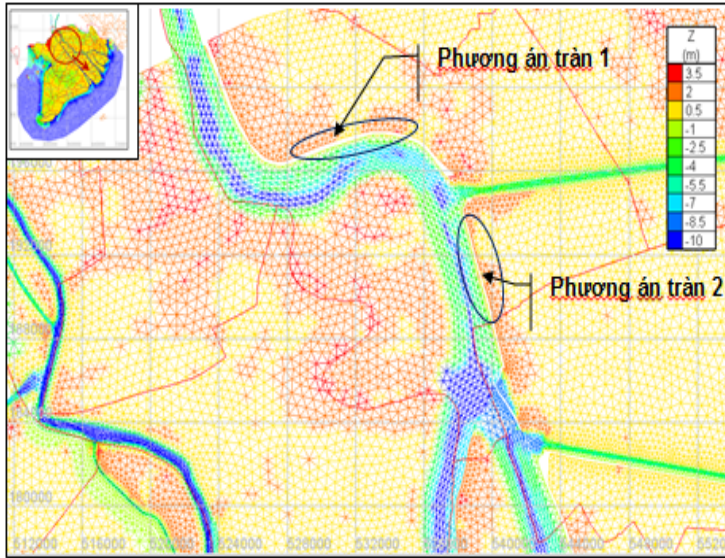
Nguyen Thong², Nam Thoai¹

¹HPC Lab – Faculty of Computer Science and Engineering

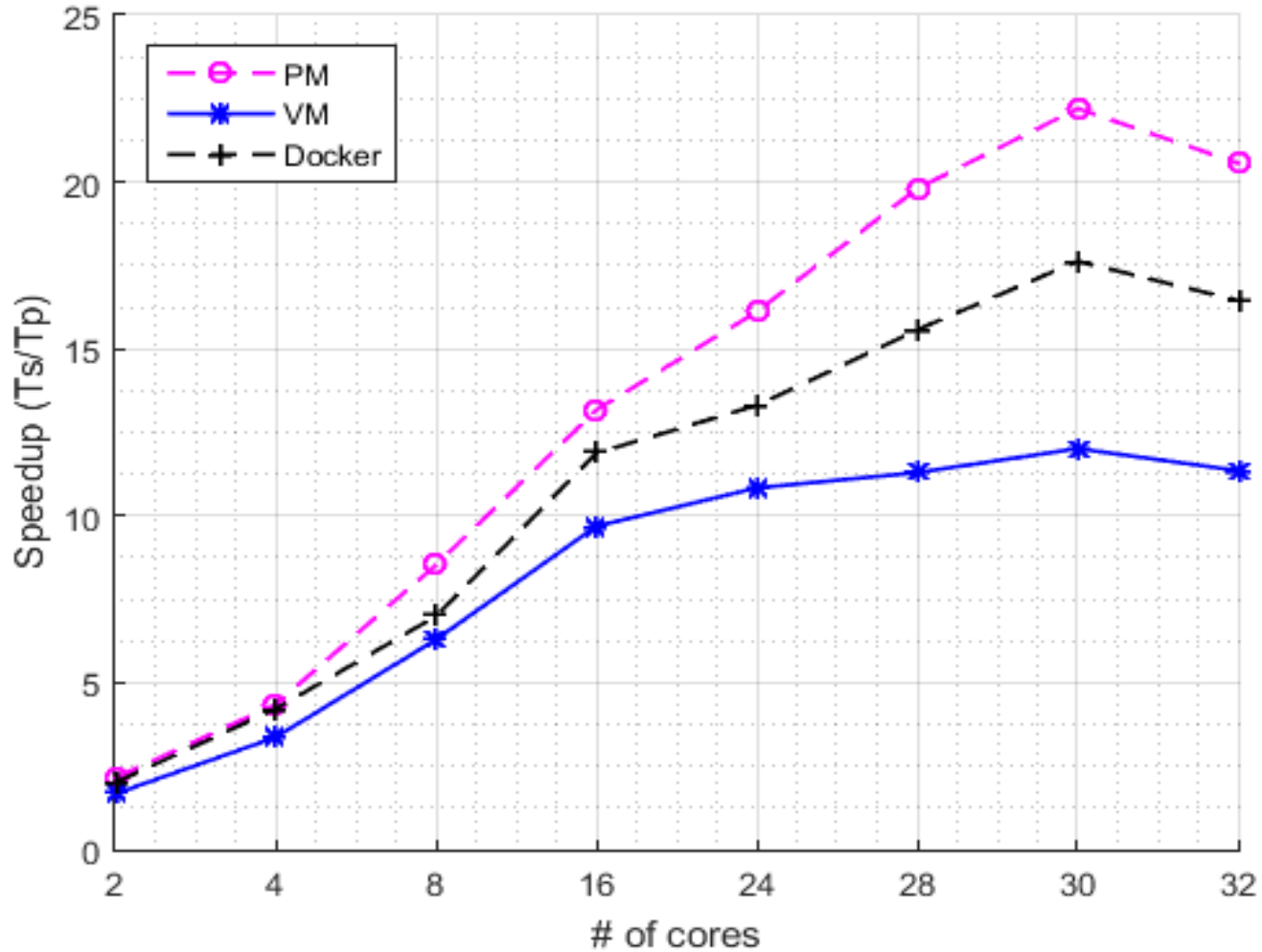
²Faculty of Civil Engineering

HCMC University of Technology, Vietnam

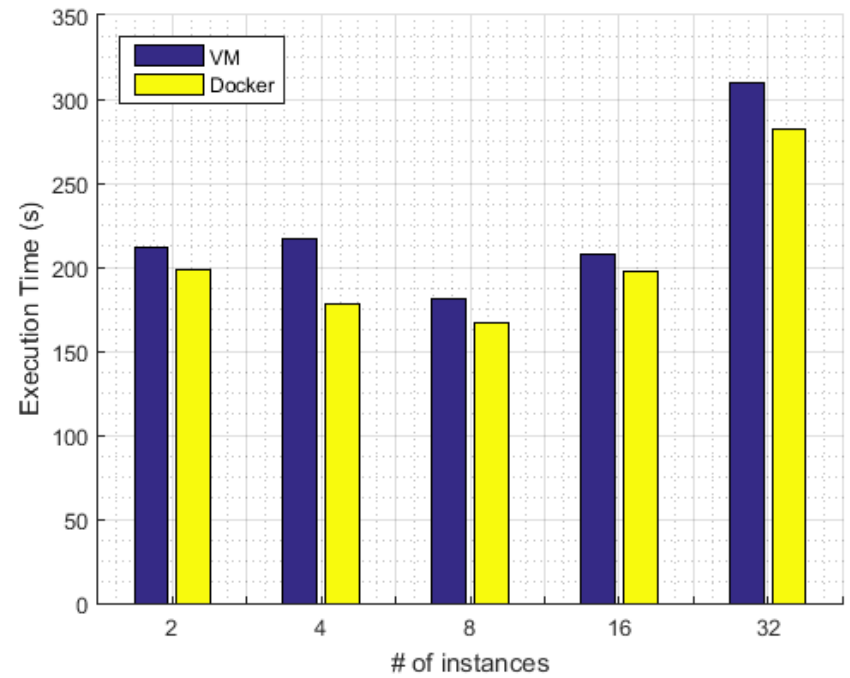
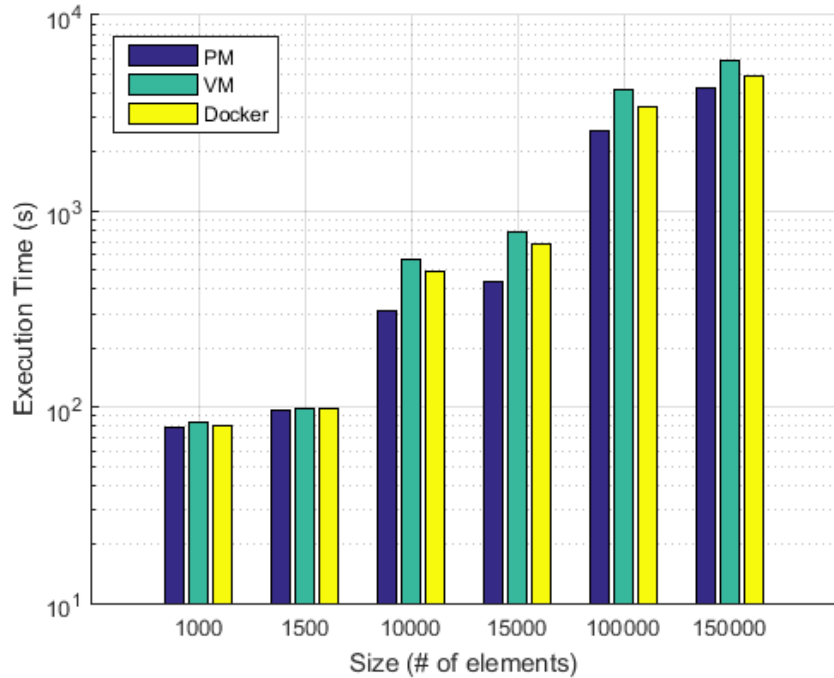
Real data of Mekong Delta



Speedup



Problem size and scalability



ITS Lab

Pham Tran Vu



Smart BK Traffic

traffic.hcmut.edu.vn

Provide the current state of the traffic in Ho Chi Minh City in real-time



Road Maester

Help the people living in Ho Chi Minh monitor the traffic condition in real-time and avoid bad traffic areas



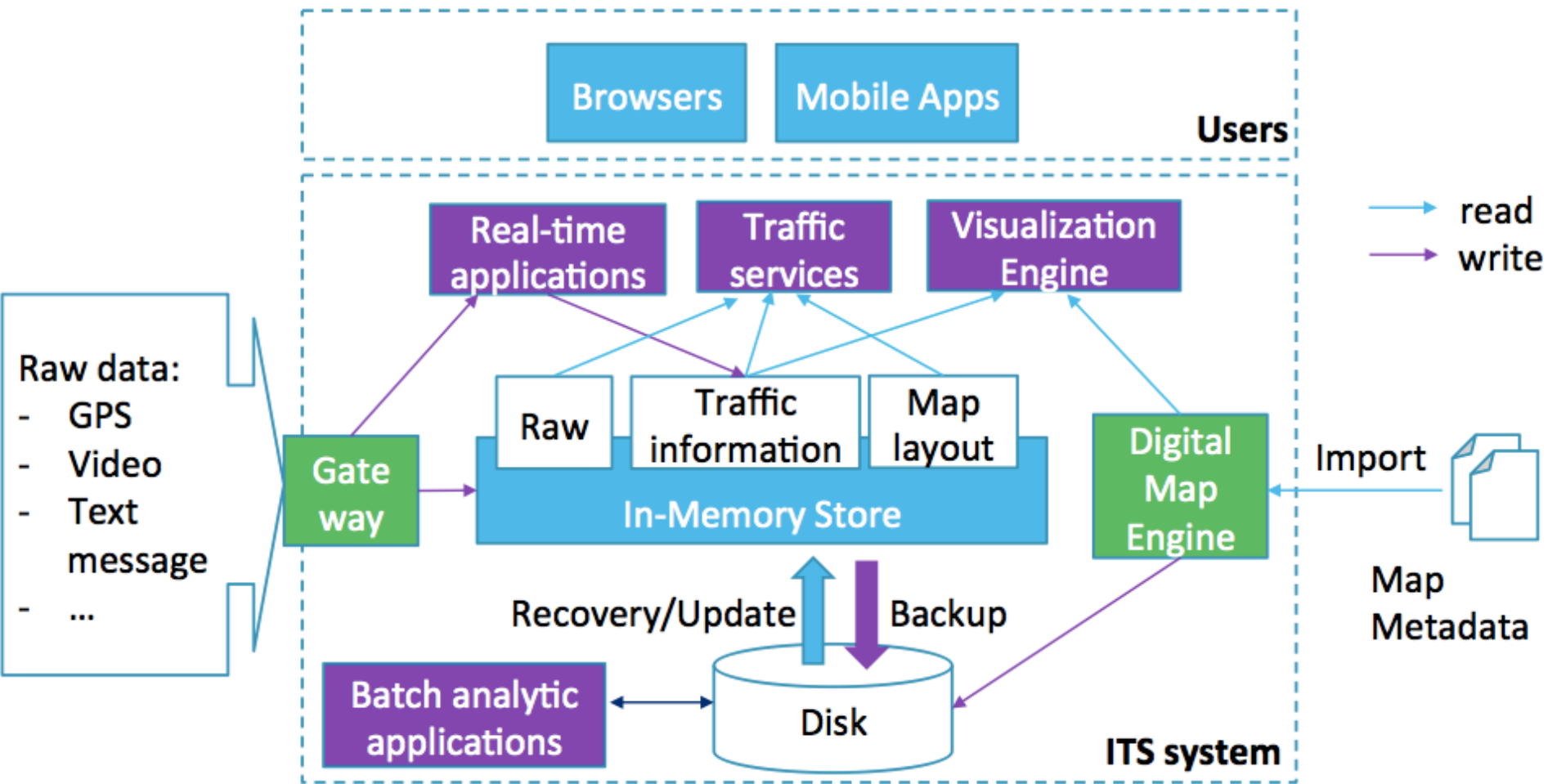
Smart Bus

Provide useful information and utilities for bus travellers

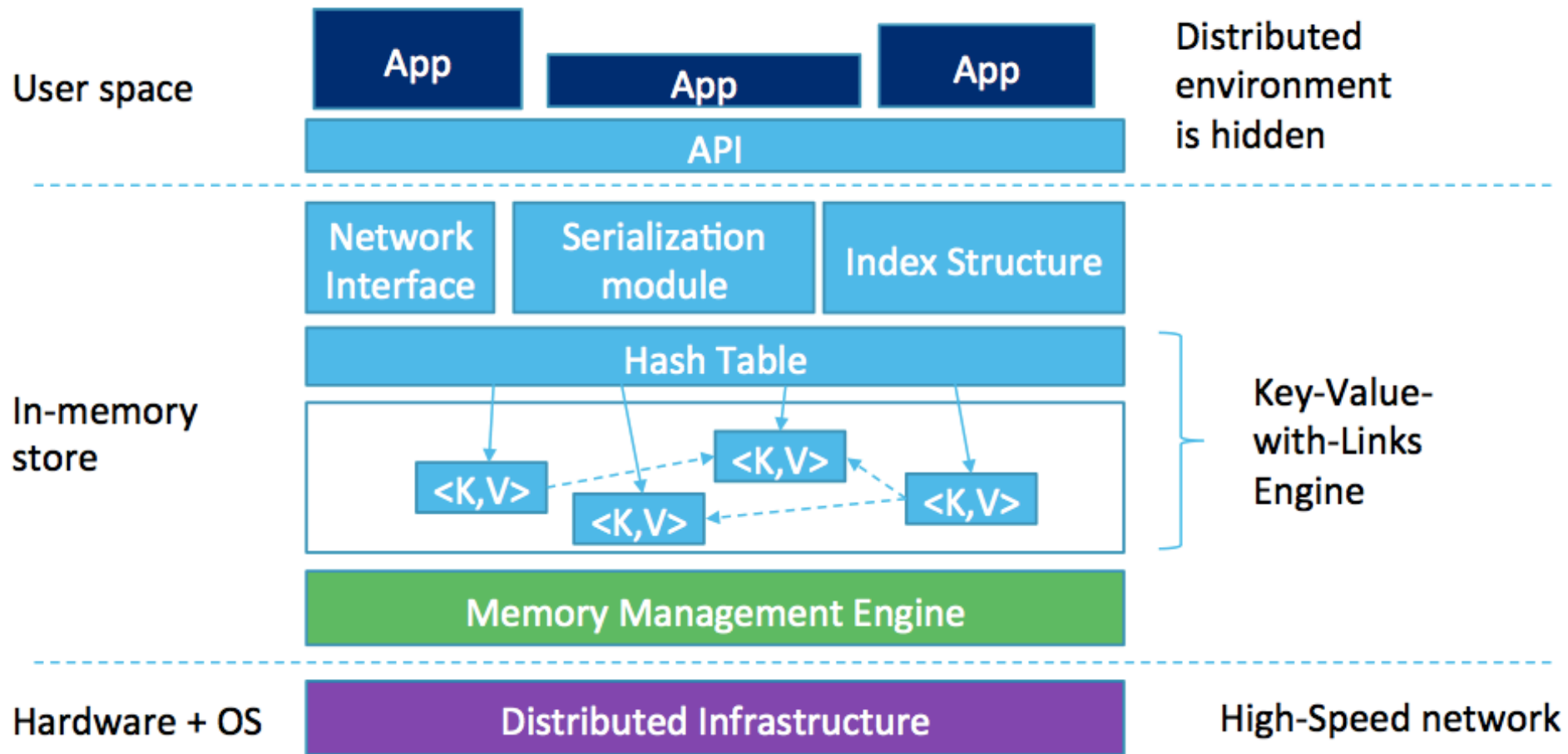
Challenges

- The volume of sensing data keeps increasing over time.
 - More than 10M records generated per day
- Extracting useful traffic information from a massive amount of data in real-time.
 - The processing must be done within 1 minute
- Low stability of the network infrastructure, The quality and heterogeneity of sensing devices, and the accuracy of the digital map
 - Only ~17% of raw data could be used directly
- Unique (and difficult) traffic problems
 - Motorbike is the most popular mean of transportation
 - Transportation development does not keep up with economic and population growth.

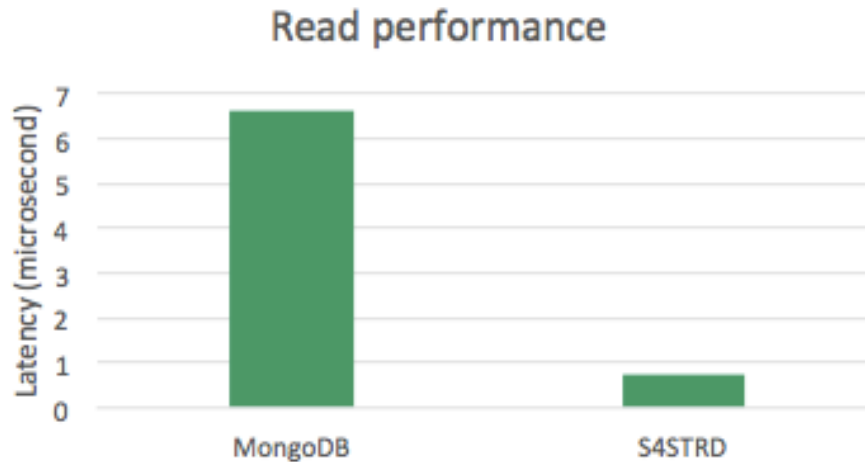
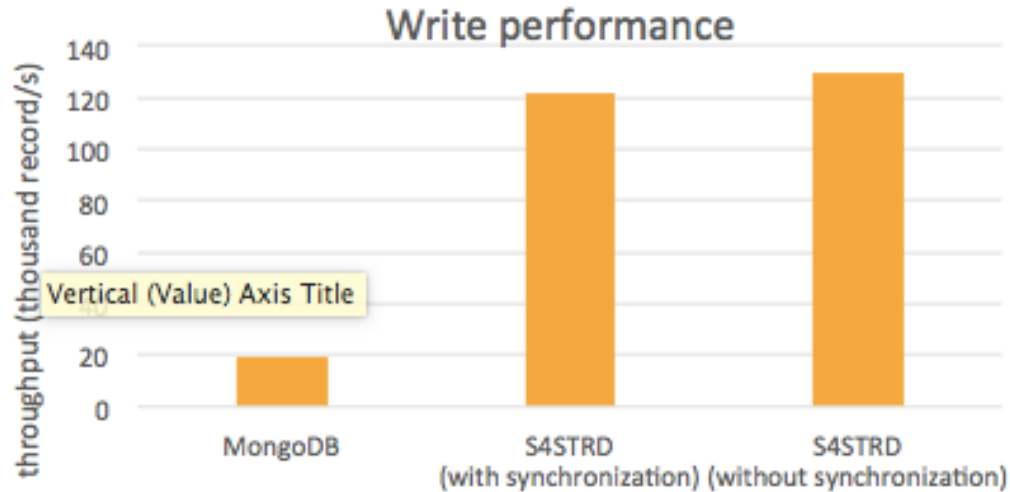
ITS System Architecture



Distributed In-memory Storage System for Big Data



Performance of Storage System





Publications

Nguyen, D.H., Le, V.A., Chung, T.M., Pham, T.V., Thoai, N., *Accelerating Range Query Execution of In-Memory Stores: A Performance Study*, IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems, 2016.

Cao, T.D., Hoang, H.H., Huynh, X.H., Nguyen, B.M., Pham, T.V., Tran, M.Q., Tran, V.T., Truong, H.L., *IoT Services for Solving Critical Problems in Vietnam: A Research Landscape and Directions*, IEEE Internet Computing 20 (5), 2016, pp.76-81.

Huynh, N., Le, T.V., Pham, T.V., Huynh, T.N., *Predicting Traveler Movement Based on a Hybrid Model of Hierarchical Clustering and Bayesian Network*, Advanced Science Letters, 22 (9), 2016, pp.2086-2090.

Nguyen, D.H., Doan, K., Pham, T.V., *SIDI: A Scalable in-Memory Density-based Index for Spatial Databases*, Proceedings of the ACM International Workshop on Data-Intensive Distributed Computing, ACM, 2016, pp.45-52.

Cao, T.D., Pham, T.V., Vu, Q.H., Truong, H.L., Le, D.H., Dustdar, S., *MARSA: A Marketplace for Realtime Human Sensing Data*, ACM Transactions on Internet Technology (TOIT) 16 (3), 2016, p.16.

Vo, D.K., Pham, T.V., Huynh, T.N., Tran, V.H., *Least Expected Time Paths in Stochastic Schedule-Based Transit Networks*, Mathematical Problems in Engineering, 2016.

Vo, D.K., Vu, H.L., Pham, T.V., Nguyen, H.T., Tran, V.H., *The α -Reliable Earliest Arrival Paths in Stochastic Public Transit Networks*, in Proc. of the 95th Transportation Research Board (TRB) Annual Meeting, Washington, DC, Jan 2016.

Pham, T.V., Nguyen, D.H., Doan, K., *S4STRD: A Scalable in Memory Storage System for Spatio-temporal Real-Time Data*, IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity 2015), IEEE, 2015, pp.896-901.

Vo, D.K., Pham, T.V., Huynh, T.N., Nguyen, N., Tran, V.H., *Finding alternative paths in city bus networks*, International Conference on Computer, Control, Informatics and its Applications (IC3INA), IEEE, 2015, pp.34-39.

Vo, D.K., Pham, T.V., Huynh, T.N., Tran, V.H., *Multi-criteria Route Planning in Bus Network*, IFIP International Conference on Computer Information Systems and Industrial Management, 2014, Springer, pp. 535-546.

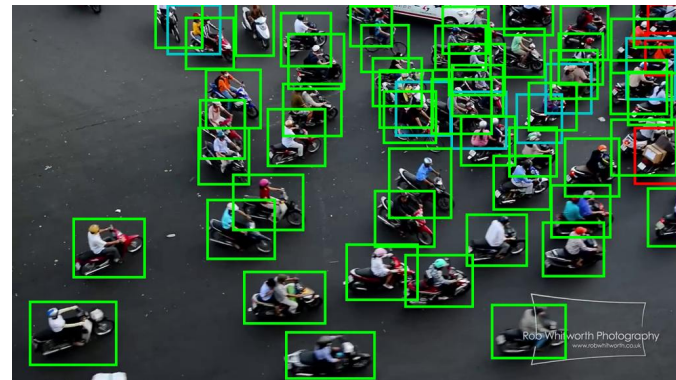
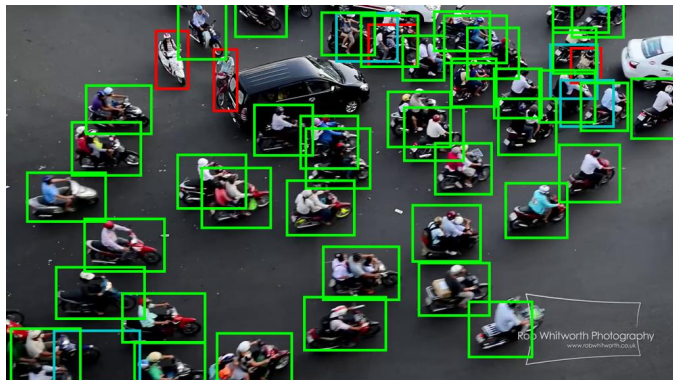
Graphics & Vision Lab

Le Thanh Sach

Automatic vehicle detection

- *Deep learning, Machine learning* -

- Object of interest: motorbike, car, bus, etc.
- Achieve an accuracy up to 95% in ideal conditions.
- Can predict correctly if there are around 60 vehicles per image or fewer.
- Can be reused in other problems (human counting, forest density estimation, etc.)



Graphics & Vision Lab

Le Thanh Sach

Motorbike count estimation

- *Machine learning* -

- Counting motorbike by applying regression over global image features (do not know the position of each vehicle).
- Fast (18 FPS) and accurate (93%)
- Can predict accurately up to 100 motorbikes per image.



→ Motorbike count: 63



→ Motorbike count: 87

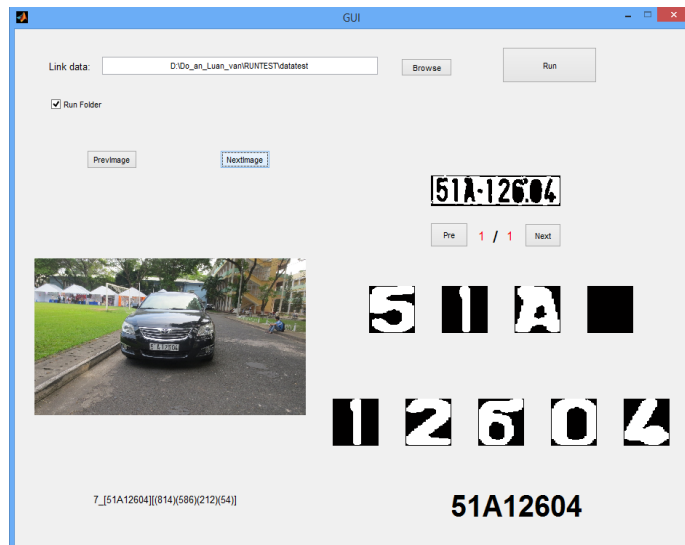
Graphics & Vision Lab

Le Thanh Sach

License plate detection

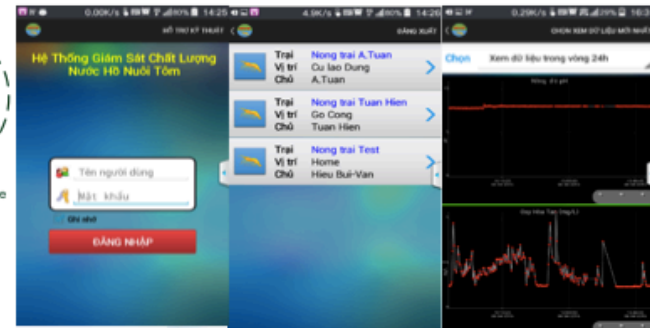
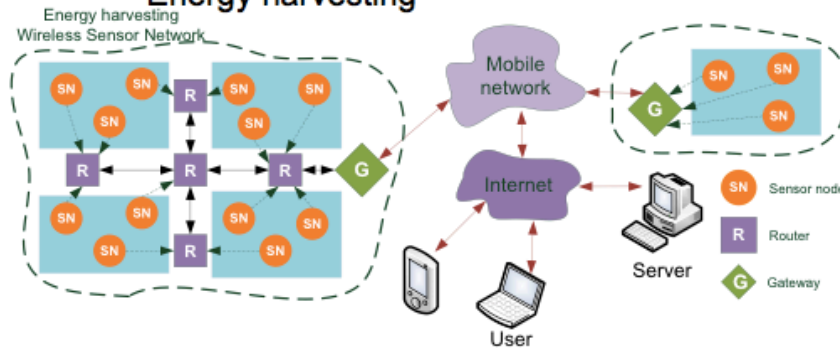
- Image processing, Machine learning -

- Both license plate detection and character recognition are integrated into the same workflow
- Can detect and recognize license plate in various lighting condition
- Detection accuracy: 96%, character recognition accuracy: 87%



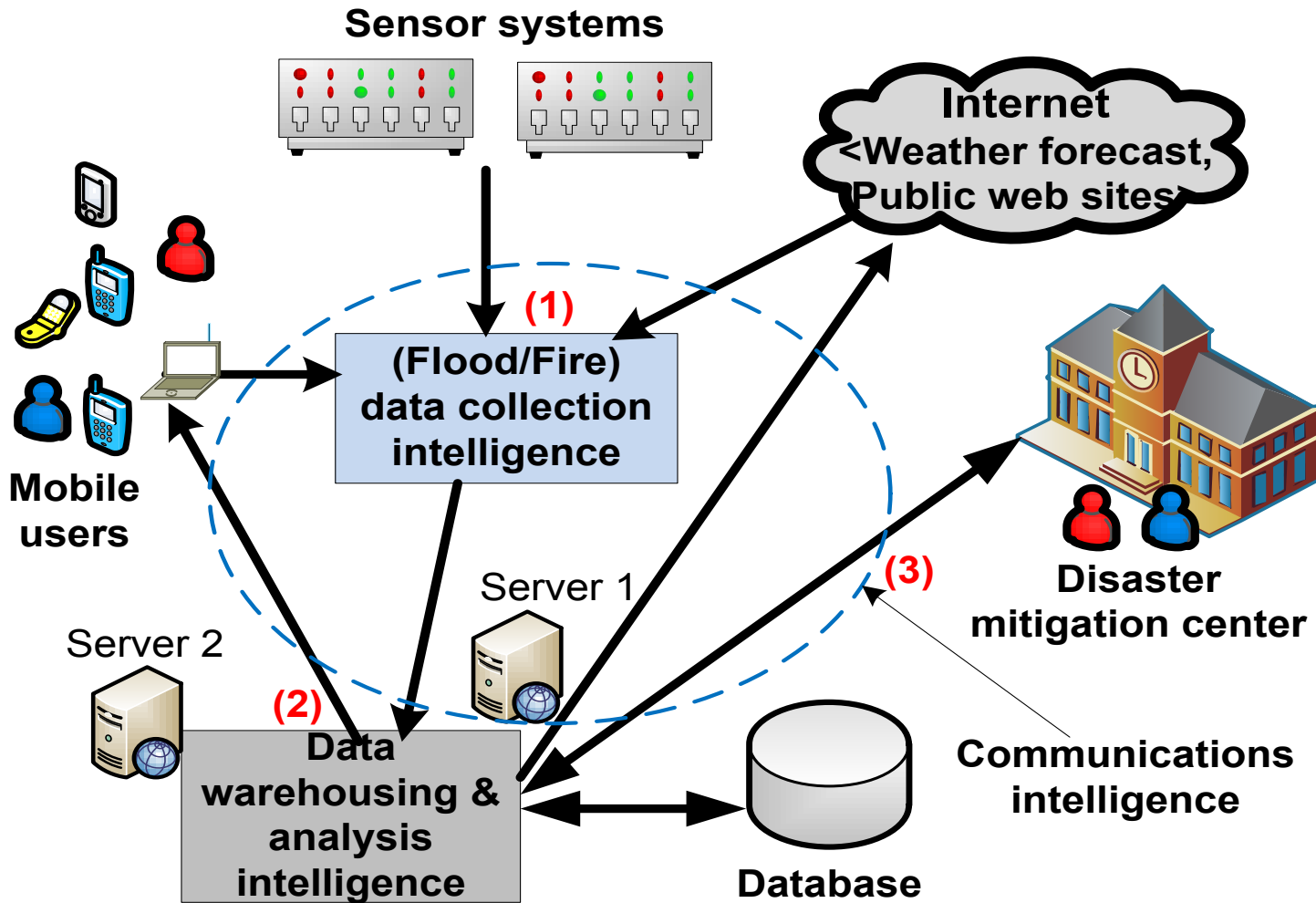
Low-power wireless water quality monitoring system

- Water quality monitoring system for shrimp farming
- Energy harvesting
- Low-power network protocol
- Data management



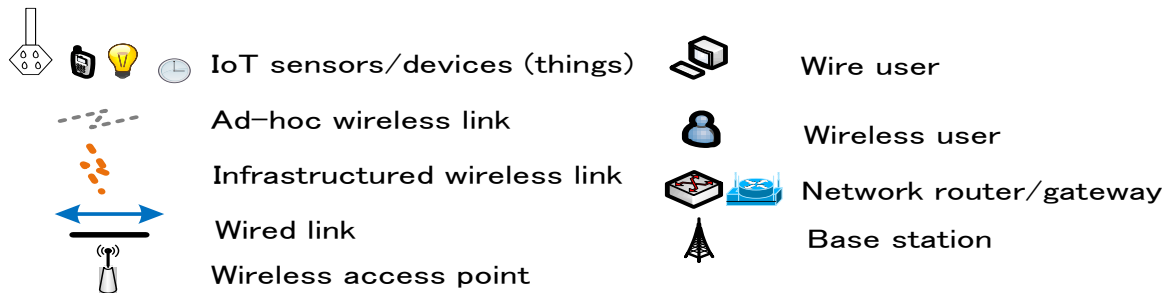
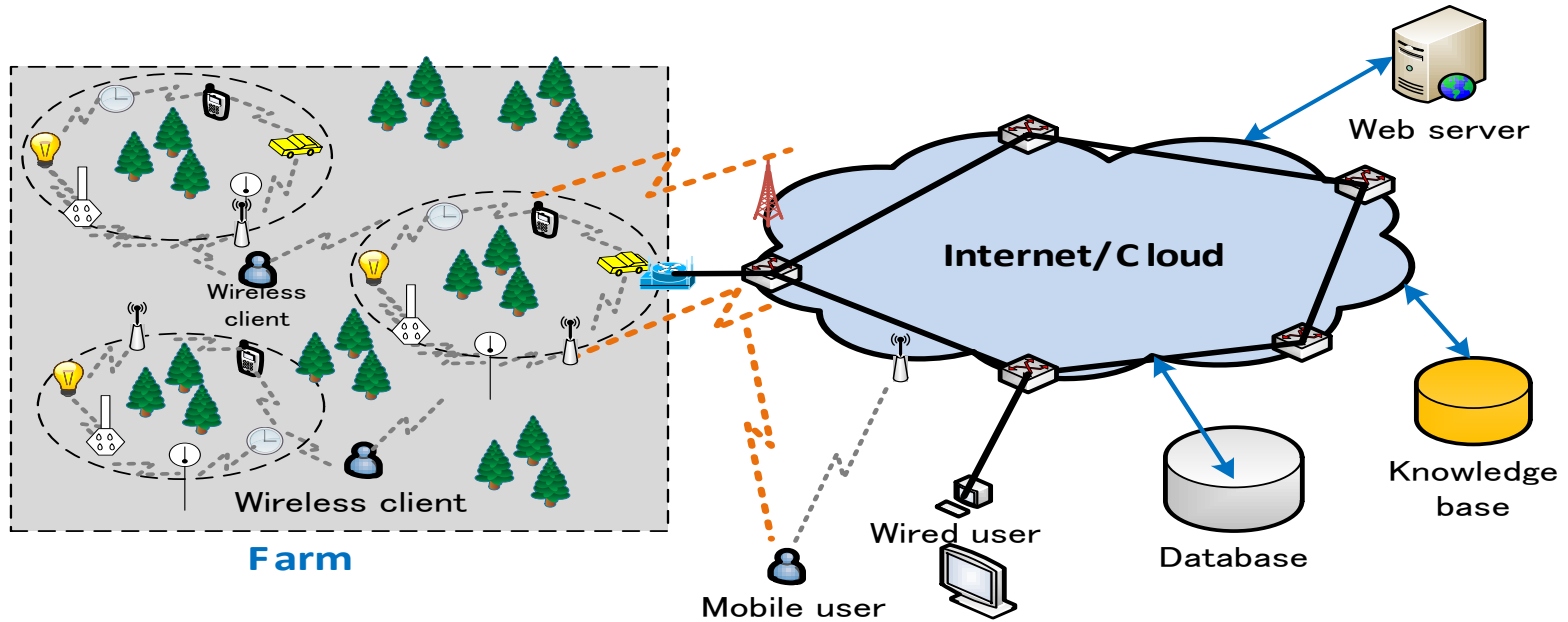
IoT Lab

Tran Minh Quang



Urban flooding mitigation

IoT - Smart farm designs



Selected projects and publications

■ Selected projects

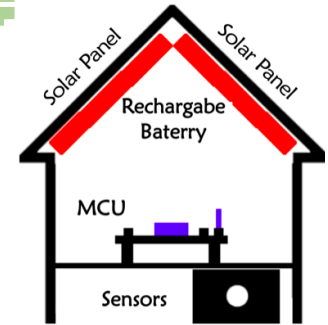
1. Context-aware multi-tier architecture for the Internet of Things, sponsored by National Foundation for Science & Technology Development (Nafosted), 2017 - 2019.
2. Mobile crowdsourcing system for flood mitigation in Ho Chi Minh City, sponsored by VNU-HCM, 2017-2018.

■ Selected publications

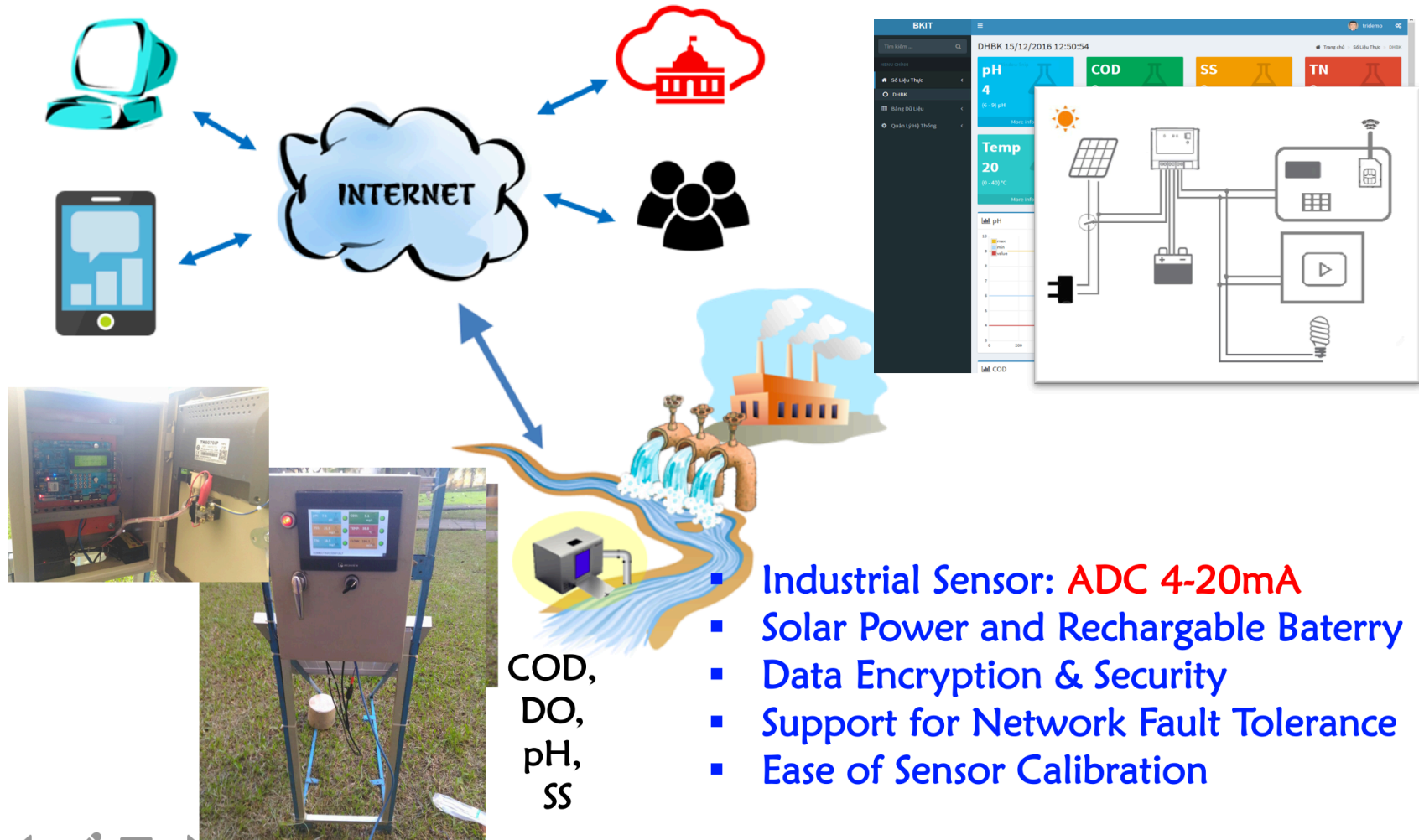
- [1] **T.M. Quang**, B. Cristian, Y. Shibata, S. Yamada, “On-site Configuration of Disaster Recovery Access Networks Made Easy, ” Elsevier Ad-hoc Networks Journal, 40, pp. 46-60, 2016.
- [2] Cao T.D., Hoang H.H., Huynh X.H., Nguyen B.M., Pham T.V., **Tran M.Q.**, Tran V.T., Truong H.L., “ IoT Services for Solving Critical Problems in Vietnam: A Research Landscape and Directions, ” IEEE Internet Computing, 20(5), pp. 76-81, 2016.
- [3] **T.M. Quang**, Shigeki Yamada, “ Evaluations of commodity WiFi-based multihop access network for disaster recovery, ” International Journal of Space-Based and Situated Computing, 6, pp. 3-11, 2016.

Air Quality Monitoring System

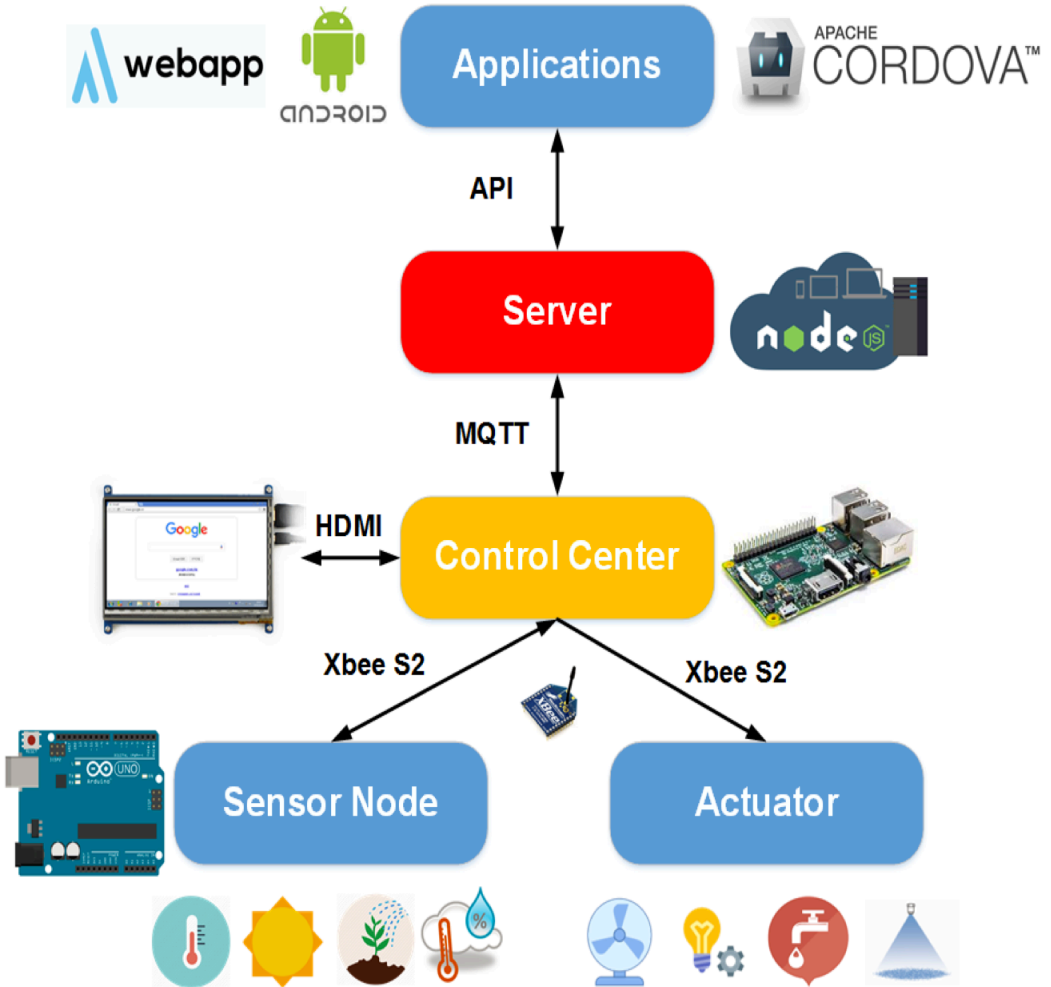
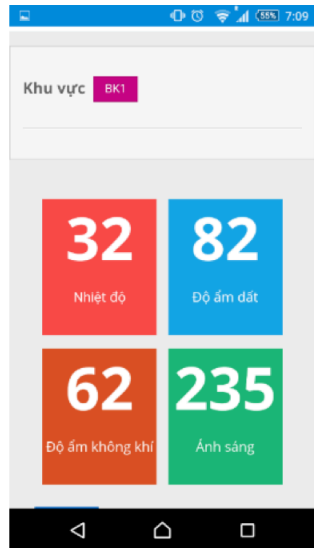
- CO, Temperature, Humidity, GAS
- Solar Power and Rechargeable Battery.



Waste Water Monitoring System

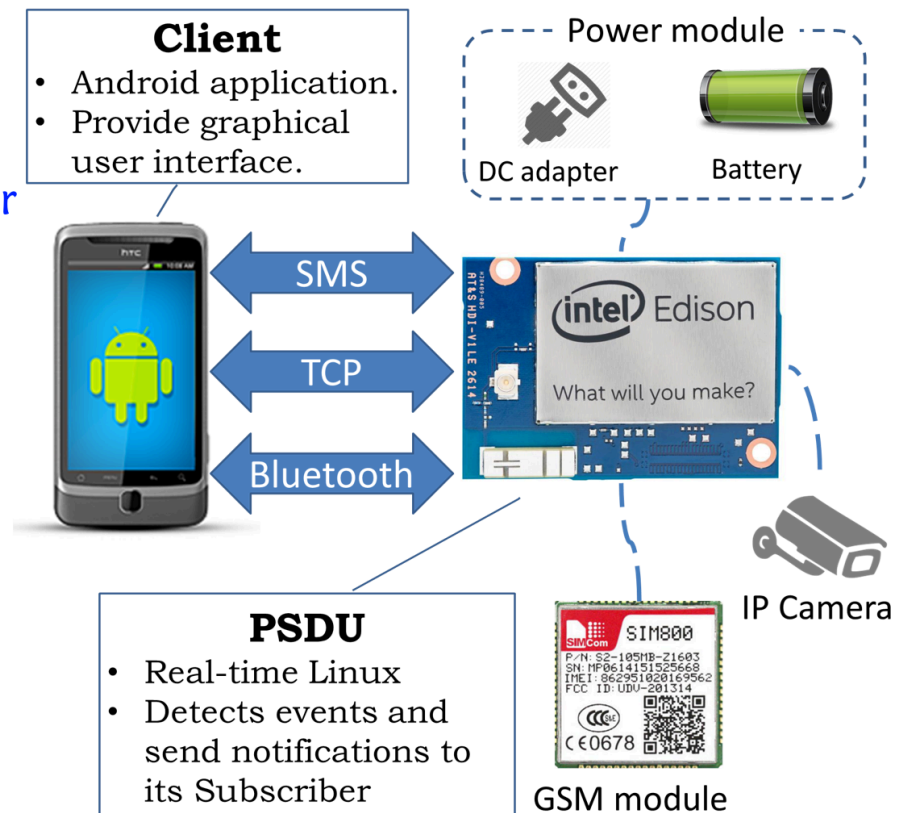


Prototype of Smart Farming



Framework for Event Management in IoT

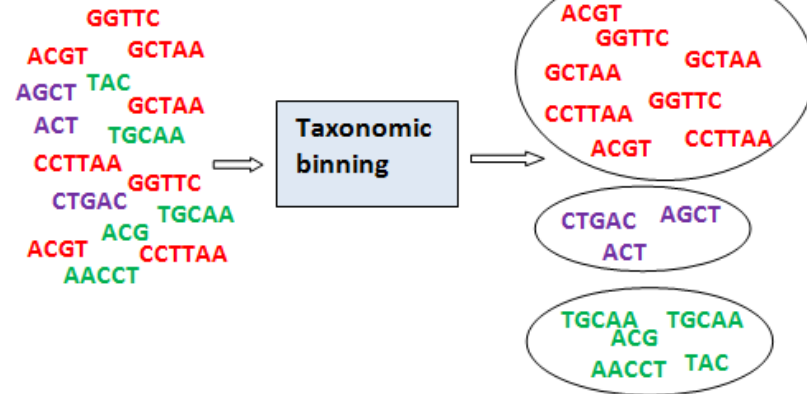
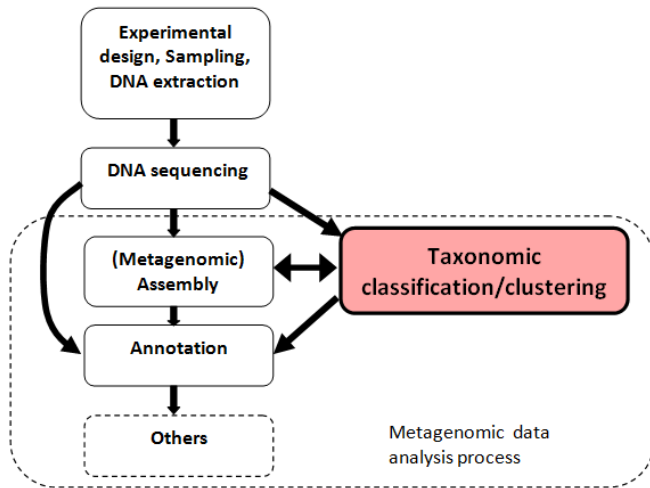
- Publish-Subscribe Framework for Event Management in IoT Applications. (**SEATUC 2017**)
- Demo Application
 - Power Signal Detection Unit for Home Security Application



Metagenomic data analysis

BiMeta: Clustering of metagenomic reads

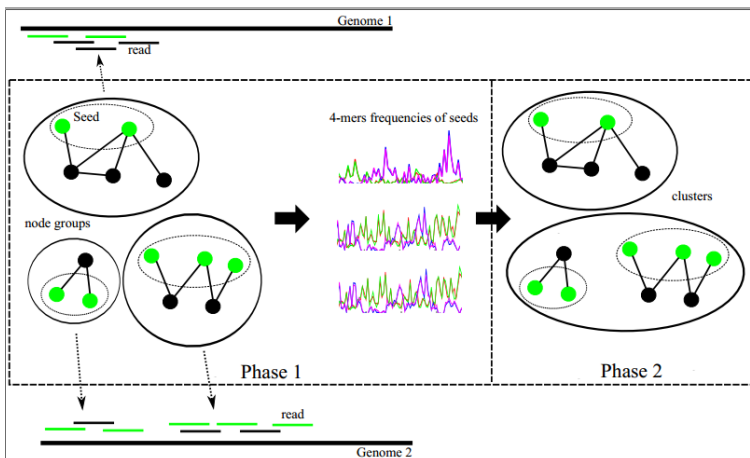
Tran Van Hoai



bioinformatics.vnu.edu.vn/bioinfo/metapro/BiMeta.html

24 cuốn sách sẽ giúp

Bioinformatics



- Home
- MetaAB
- BiMeta
- SeMeta
- Publications

BiMeta

BiMeta is a new software for binning metagenomic reads. It supports both single-end and paired-end reads. The software is implemented in C++.

Download source code here:

- Version 1.1: [BiMetal.1.tar.gz](#) (updated date: 08/03/2015)
- Version 1.2: [BiMetal.2.tar.gz](#) (updated date: 16/04/2015) (Support an additional function to separate output file into different files of different species)

Publication:

Vinh, L. V., Lang, T. V., Binh, L. T., & Hoai, T. V. (2015). A two-phase binning algorithm using 1-mer frequency on groups of non-overlapping reads. *Algorithms for Molecular Biology*, (1), 2.

Datasets:

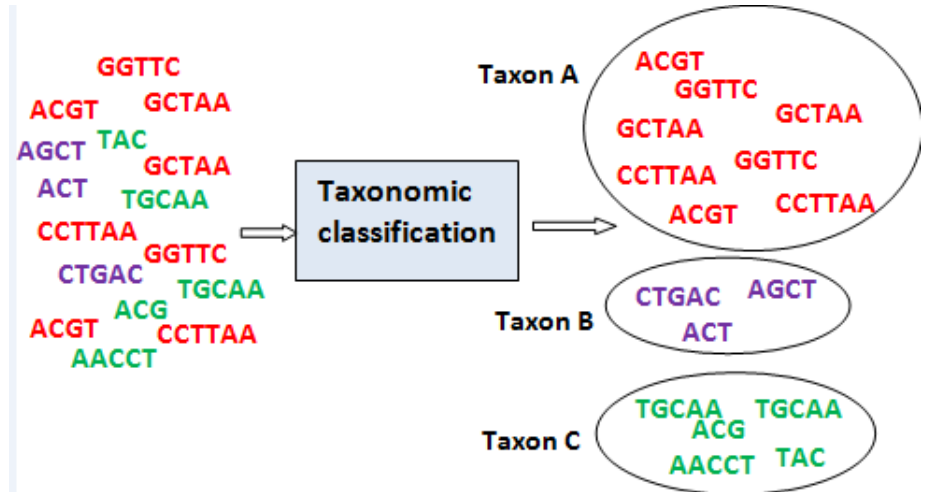
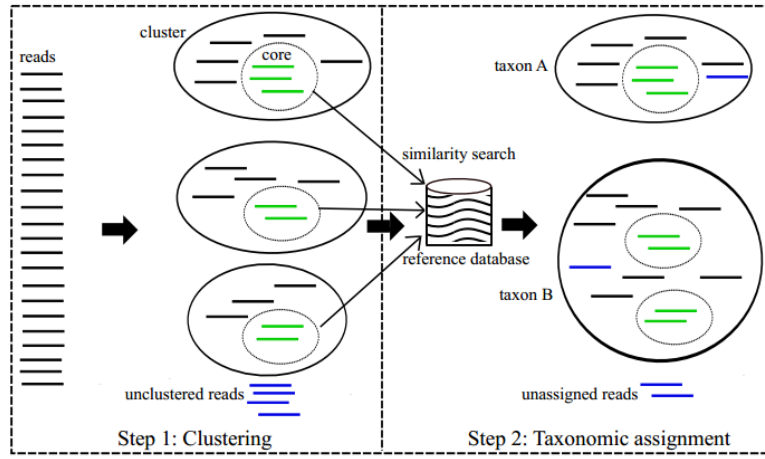
The application was tested with various datasets. The datasets can be download here:

- Short read datasets:
S1.fna.gz, S2.fna.gz, S3.fna.gz, S4.fna.gz, S5.fna.gz, S6.fna.gz, S7.fna.gz, S8.fna.gz, S9.fna.gz, S10.fna.gz, L1.fna.gz, L2.fna.gz, L3.fna.gz, L4.fna.gz, L5.fna.gz, L6.fna.gz
- Long read datasets:
R1.fna.gz, R2.fna.gz, R3.fna.gz, R4.fna.gz, R5.fna.gz, R6.fna.gz, R7.fna.gz, R8.fna.gz, R9.fna.gz
- Acid mine drainage: <http://www.ncbi.nlm.nih.gov/books/NBK6860/>

Metagenomic data analysis

SeMeta/ParSeMeta: Classification of metagenomic reads

Tran Van Hoai



edu.vn/bioinfo/metapro/SeMeta.html

24 cuốn sách sẽ giúp:

Bioinformatics



- Home
- MetaAB
- BiMeta
- SeMeta
- Publications

SeMeta

SeMeta is a new software for taxonomic assignment of metagenomic reads. It supports both single-end and paired-end reads. The software is implemented in C++

Source codes:

Download source code here:

Version 1.0 (release 02/10/2015): [SeMeta1.0.tar.gz](#)

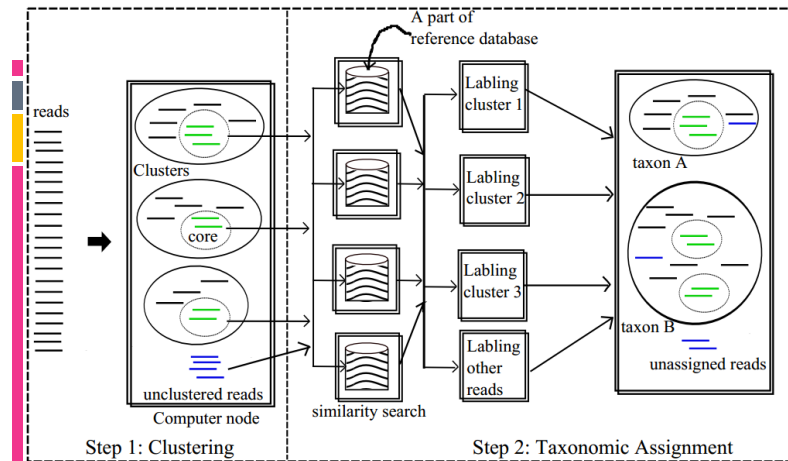
This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

The application was tested with various datasets. The datasets can be downloaded here:

- + Dataset 1: [ds1.fna.gz](#)
- + Dataset 2: [ds2.fna.gz](#)
- + Dataset 3: [ds3.tar.gz](#)
- + Acid mine drainage: <http://www.ncbi.nlm.nih.gov/books/NBK6860/>
- + Human gut metagenome: <http://public.genomics.org.cn/BGI/gutmeta/>

How to compile the source code, and prepare input files:

Please read the README.txt file in the source code folder.



Publications

Journals

- [1]. Vinh V. L., Lang V. T., and Hoai V. T. (2016). A novel semi-supervised algorithm for the taxonomic assignment of metagenomic reads. *Journal of BMC Bioinformatics*, 17 (22) (**SCIE index, IF=2.58**).
- [2]. Vinh, L. V., Lang, T. V., Binh, L. T., and Hoai, T. V. (2015). A two-phase binning algorithm using l-mer frequency on groups of non-overlapping reads. *Journal of Algorithms for Molecular Biology*, 10(1) (**SCIE index, IF=1.46**).
- [3]. Vinh V. L., Lang V. T, and Hoai V. T.(2014). A novel l-mer counting method for abundance based binning of metagenomic reads, *Journal of Computer Science and Cybernetics*, ISSN 1813-9663, V.30, N.3, 267-277.
- [4]. Vinh, L. V., Lang, T. V., Binh, L. T., and Hoai, T. V. (2014). Performance of clustering approaches for metagenomic reads. *Journal of Science and Technology*, ISSN 0866-708X, 52 (1B) , pp.28-36

Conferences

- [1]. Van Vinh, L., Van Lang, T., and Van Hoai, T. (2014). MetaAB-A Novel Abundance-Based Binning Approach for Metagenomic Sequences. *In Nature of Computation and Communication* (pp. 132-141). Springer International Publishing.
- [2]. Vinh L. V., Nhut D. H, Hoai T. V. and Lang T. V. (2014), A combination of genomic signatures for the binning of metagenomic sequences, Proceedings of The 2nd International Conference on Green Technology and Sustainable Development, HCM City Oct 30-31, ISBN 978-604-732-817-8, 662-668.
- [3]. Vinh L. V., Lang T. V., Hoai T. V. (2014), An abundance-based binning approach for metagenomics read using a fuzzy k-medoids methods, FAIR'7, Thai Nguyen, ISBN: 978-604-913-300-8, Natural Science and Technology Publishing House.



Environment Computing Session – ACOMP 2016

23-25 Nov 2016, Can Tho city, Vietnam

- Nguyen Thong & [Ho Tuan Duc](#)
Civil Engineering Faculty - HCMUT
Application of Telemac to Flood and Salinity Management in Vietnam
- [Le Thanh Sach](#)
CSE Faculty - HCMUT
Recently Developed Techniques for Vehicle Detection from Traffic Video
- [Doan Chi Hieu](#)
HP Vietnam
HPC solutions
- [Thoai Nam](#)
HPC Lab - HCMUT
Open Computing Platform on SuperNode-XP

Environmental Computing Session

in

ACOMP 2017

International Conference on Advanced
Computing and Applications

Ho Chi Minh City, Vietnam, Nov 29-30 to Dec 01, 2017

Submission deadline: June 30, 2017



ACOMP 2017 Proceedings will be published by the [Conference Publishing Services \(CPS\)](#). All the presented papers are to be submitted to the [IEEE Xplore](#) and [IEEE CS Digital Library](#). ACOMP proceedings started to be indexed by the DBLP bibliography since 2015.

HCMUT support

- Open Computing Platform at HPC Lab – HCMUT
<http://www.hpcc.hcmut.edu.vn>
- Short course training in IoT , Big Data analytics, HPC
- Master program in “Data Science”, 2018

Summary

- There are very critical issues that need to be solved in Vietnam
 - Environmental problems are big problems
 - Environmental Computing is very hot
- Open Computing Platform at HPC Lab – HCMUT
<http://www.hpcc.hcmut.edu.vn>
- Labs in applications & infrastructure computing
 - Looking for collaborations



Thank you!

More information:

namthoai@hcmut.edu.vn

<http://www.cse.hcmut.edu.vn/>

<http://www.hpcc.hcmut.edu.vn/>