Demo Abstract: Decomposing Data Analytics in Fog Networks

Ta-Cheng Chang
National Chiao Tung University
Hsinchu, Taiwan 30010
edward.ee05g@nctu.edu.tw

Liang Zheng
Princeton University
Princeton, NJ 08544
liangz@princeton.edu

Maria Gorlatova
Princeton University
Princeton, NJ 08544
mariaag@princeton.edu

Chege Gitau
Princeton University
Princeton, NJ 08544
dgitau@princeton.edu

Ching-Yao Huang
National Chiao Tung University
Hsinchu, Taiwan 30010
cyhuang@mail.nctu.edu.tw

Mung Chiang
Purdue University
West Lafayette, IN 47907
chiang@purdue.edu

ABSTRACT

Fog computing, the distribution of computing resources closer to the end devices along the cloud-to-things continuum, is recently emerging as an architecture for scaling of the Internet of Things (IoT) sensor networking applications. Fog computing requires novel *computing program decompositions* for heterogeneous hierarchical settings. To evaluate these new decompositions, we designed, developed, and instrumented a fog computing testbed that includes *cloud computing* and *computing gateway* execution points collaborating to finish complex data analytics operations. In this interactive demonstration we present one fog-specific algorithmic decomposition we recently examined and adapted for fog computing: a multi-execution point linear regression decomposition that *jointly optimizes operation latency, quality, and costs*. The demonstration highlights the role fog computing can play in future sensor networking architectures, and highlights some of the challenges of creating computing program decompositions for these architectures. An annotated video of the demonstration is available at [5].

CCS CONCEPTS

•**Hardware** → *Emerging architectures;* •**Computer systems organization** → **Sensor networks;** *n-tier architectures;* •**Computing methodologies** → **Distributed algorithms;** •**Networks** → *Network economics;*

KEYWORDS

Fog computing, edge computing, distributed systems, data analytics, heterogeneous architectures, Internet of Things.

ACM Reference format:
Ta-Cheng Chang, Liang Zheng, Maria Gorlatova, Chege Gitau, Ching-Yao Huang, and Mung Chiang. 2017. Demo Abstract: Decomposing Data Analytics in Fog Networks. In *Proceedings of SenSys '17, Delft, Netherlands, November 6–8, 2017,* 2 pages.
DOI: 10.1145/3131672.3136962

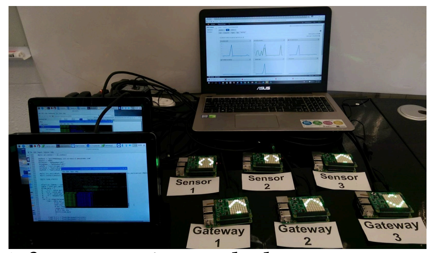

Figure 1: A fog computing testbed setup we are presenting in this interactive demonstration [5].

1 INTRODUCTION

Fog computing and edge computing, which distribute traditionally centralized datacenter operations closer to the end users, have been receiving considerable attention as enablers of the next level of interactivity and cognition in the Internet of Things (IoT) [6, 7]. In the last year a partner at Andreessen Horowitz, a top venture capital fund, called such "intelligent edge" developments *the next multi-billion dollar tech market* [4], while Microsoft CEO Satya Nadella referred to them as *the most interesting part of cloud computing* [8]. For the IoT sensor networking applications, an important aspect of fog computing architectures are the capabilities of the *computing gateways* [12]. Executing parts of computing operations on local gateways can reduce bandwidth consumption and cloud computing costs, and hence architectures that allow executing services in multiple points have been recently made available [1, 9]. However, decomposing computing programs between the cloud and the computing gateways is not straightforward, particularly for complex algorithms in data analytics applications. In particular, traditional parallelization methods, developed for distributing operations to homogeneous nodes, are insufficient for fog and edge computing settings due to the heterogeneity of capabilities of gateways and cloud nodes. To support our research on program decompositions for fog computing, we created a testbed for evaluating the performance of fog-specific decompositions, and developed an interactive demonstration that shows different elements of a heterogeneous fog computing system in action. The testbed setup we present in this demonstration is shown in Fig. 1.

2 FOG COMPUTING TESTBED FOR DATA ANALYTICS DECOMPOSITIONS

We implement the fog computing testbed using Raspberry Pi 3-based nodes [10] and Amazon cloud computing services [2, 3]. We

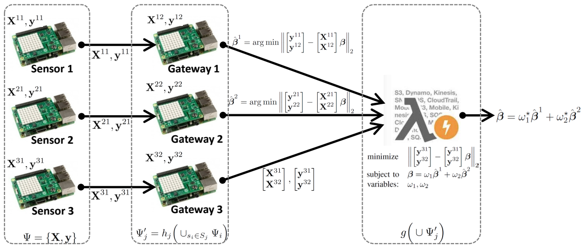

Figure 2: Decomposing data analytics between the computing gateways and the cloud services in fog networks: a linear regression case study.

use Raspberry Pis to emulate sensor nodes and computing gateways. For capturing environmental parameters and for display purposes, we outfitted Raspberry Pis with Sense HAT add-on boards that measure temperature, humidity, and air pressure, and display sensor status graphically using 64-element LED arrays. Cloud computing elements are implemented via Amazon Lambda (a computing service) [3] and DynamoDB (a data storage service) [2]. All computing operations are implemented in Python. Sensor nodes communicate with computing gateways over Bluetooth. Computing gateways reach cloud services over WiFi or Ethernet via standard HTTP request/response mechanisms.

We examine scenarios where local IoT devices $\mathcal{S} = \{s_1, s_2, \ldots\}$ and the remote cloud services carry out data sensing, collection, and analytics. We select a subset of the IoT devices to act as *computing gateways* that perform computing operation h_j over the data Ψ_i received from each sensor node i connected to the gateway j, and obtain the result $\Psi'_j = h_j\big(\cup_{s_i \in S_j} \Psi_i\big)$. Here, S_j denotes the set of sensor nodes that send their sensed data to gateway j. The functional operations h_j can be different for the different gateways j. We focus on the arising tradeoffs in *latency* (as both computing and communications can be time-consuming), *quality* (as different decompositions lead to different results), and *costs* (as cloud services are charged for data transmission, storage, and computing). We instrumented our testbed to measure all related parameters.

3 LINEAR REGRESSION CASE STUDY

We chose a linear regression as a demonstrative case study due to its widespread use and well-understood mathematical properties. In the decomposition method we adapt from [11] for fog computing settings, the linear regression is solved jointly by the computing gateways and the cloud services, with a combining method of [11] used in the cloud service operations for calculating the final result based on the partial results calculated by the computing gateways.

In the demonstration, we calculate the linear regression using the data collected by the sensor nodes' Sense HAT add-on boards: we infer temperature from the time of the day, humidity, and air pressure. The network topology and the mathematical details of the method are shown in Fig. 2. Here, Ψ_1 and Ψ_2 on Gateways 1 and 2 run the gradient decent algorithm to calculate features $\hat{\boldsymbol{\beta}}^1$ and $\hat{\boldsymbol{\beta}}^2$, while Ψ_3 on Gateway 3 aggregates its' data without processing it. Cloud services combine the features received from Gateways 1 and 2 that minimize the least-square error on the data received from Gateway 3. The combined $\hat{\boldsymbol{\beta}} = \omega_1^\star \hat{\boldsymbol{\beta}}^1 + \omega_2^\star \hat{\boldsymbol{\beta}}^2$ is the final data analytics result.

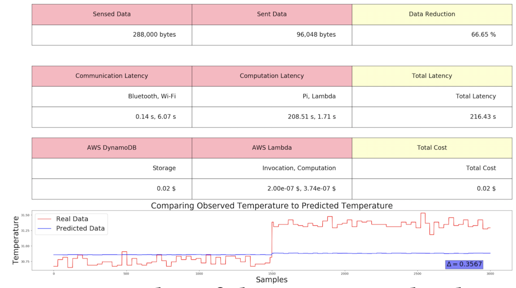

Figure 3: A screenshot of the custom graphical user interface (GUI) we developed for this demonstration.

4 INTERACTIVE DEMONSTRATION

An annotated video of the interactive visual demonstration is available at [5]. The demonstration follows the connectivity pattern shown in Fig. 2. It showcases the different operations of sensors, computing gateways, and cloud computing points in the linear regression case study described above. It also allows the participants to develop a feel for the different end-to-end system-level latency, quality, and cost tradeoffs.

The participants interact with the demo setup by setting different parameters (e.g., the number of data samples collected by the sensor nodes) in a graphical user interface (GUI) we designed. A sample screenshot of the GUI is shown in Fig. 3. We use two small SunFounder displays to show in-depth the activity of a representative sensor node and a representative computing gateway. We also use Sense HAT LED arrays to display the ongoing operations of all nodes in our demonstration (the arrays flash ↑, ×, ◯, and other symbols to show different ongoing system operations). The participants can also examine the end-to-end performance parameters captured by the testbed and shown in the GUI.

ACKNOWLEDGMENTS

The authors thank Dr. Andrew S. Lan for his useful suggestions on this work. This research is partially supported by DARPA DCOMP Program under Contract No. HR001117C0052 and No. HR001117C0048, and Comcast Innovation Fund Research Grant.

REFERENCES

[1] Amazon AWS. 2017. AWS Greengrass. (2017). http://aws.amazon.com/greengrass.
[2] Amazon AWS. 2017. DynamoDB. (2017). https://aws.amazon.com/dynamodb/.
[3] Amazon AWS. 2017. Lambda. (2017). https://aws.amazon.com/lambda/.
[4] J. Bort. 2016. The next multibillion-dollar tech market was quietly born this year, says A-list VC Peter Levine. Business Insider. (2016). http://www.businessinsider.com/edge-computing-is-the-next-multi-billion-tech-market-2016-12.
[5] Ta-Cheng Chang, Liang Zheng, Maria Gorlatova, Chege Gitau, Ching-Yao Huang, and Mung Chiang. 2017. Demo Video: Decomposing Data Analytics in Fog Networks. (2017). https://youtu.be/CBpCVyFvGXI.
[6] Mung Chiang and Tao Zhang. 2016. Fog and IoT: An Overview of Research Opportunities. *IEEE Internet of Things Journal* 3, 6 (2016), 854–864.
[7] OpenFog Consortium. 2017. OpenFog Reference Architecture. (2017). https://www.openfogconsortium.org/ra/.
[8] Tom Krazit. 2017. Microsoft CEO Satya Nadella Outlines its Cloud Computing Strategy for Wall Street. GeekWire. (2017). https://www.geekwire.com/2017/microsoft-ceo-satya-nadella-outlines-cloud-computing-strategy-wall-street/.
[9] Microsoft Inc. 2017. Azure IoT Edge. (2017). http://github.com/Azure/iot-edge.
[10] Raspberry Pi Foundation. 2017. Raspberry Pi. (2017). https://www.raspberrypi.org/.
[11] Yuhong Yang. 2003. Regression with Multiple Candidate Models: Selecting or Mixing? *Statistica Sinica* (2003), 783–809.
[12] T. Zachariah, N. Klugman, B. Campbell, J. Adkins, N. Jackson, and P. Dutta. 2015. The Internet of Things Has a Gateway Problem. In *Proc. ACM HotMobile'15*.