5G in Practice: Measuring Emerging Rural Wireless Technology for Edge Devices in Distributed Computation Workloads

Zack Murry
University of Missouri
Columbia, MO, USA
zjmfrr@missouri.edu

Alicia Esquivel Morel [advisor]
University of Missouri
Columbia, MO, USA
ace6qv@mail.missouri.edu

Kate Keahey [advisor]
Argonne National Laboratory
Lemont, IL, USA
keahey@mcs.anl.gov

Abstract

Edge computing, the notion of moving computational tasks from servers to the data-generating network edge, is an increasingly popular model for data processing. 5G wireless technologies offer an opportunity to enable complex distributed edge computing workflows by minimizing the overhead incurred in transmitting data to peer devices. In this work, we demonstrate the use and performance of edge devices in distributed computation workloads using Hadoop MapReduce on a cluster of six 5G-connected Raspberry Pis. Specifically, we first determine the network capabilities (i.e., latency and throughput) across millimeter wave (mmWave) 5G links and then analyze the scalability and performance of our cluster. Our experiment uses 5G radios at the Agricultural and Rural (ARA) Wireless Living Lab, spanning over six miles in diameter.

CCS Concepts: • **Networks → Network experimentation**.

Keywords: 5G, distributed computing, edge computing

ACM Reference Format:
Zack Murry, Alicia Esquivel Morel [advisor], and Kate Keahey [advisor]. 2024. 5G in Practice: Measuring Emerging Rural Wireless Technology for Edge Devices in Distributed Computation Workloads. In *The International Conference for High Performance Computing, Networking, Storage, and Analysis.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Edge computing enables on-site processing of tasks usually handled by remote servers, reducing network overhead. This approach leverages "edge" technologies, often Internet of Things (IoT) devices with modest compute capabilities, to enhance their utility [?]. Edge devices can often generate large amounts of data, which would traditionally be offloaded to a server to be processed. However, issues arise once the data collected by an edge device is too large, or the computation is too expensive, to be processed by a single edge device. A common approach would be to transmit the data to the cloud and forgo edge computing altogether. Another approach would be to leverage the computational capabilities of these networked edge devices to process the data in a distributed architecture. 5G wireless technology offers a promising method to optimize this form of edge-based distributed computing. By accelerating the transmission speed and offering miles of connectivity, 5G can extend the scope of these computation workloads in an efficient manner. In this work, we conduct experiments over a distributed network of Raspberry Pi 4 Single Board Computers (SBCs) to determine network capabilities and to perform standard data processing becnhmarks using Apache Hadoop [?]. We measure 5G rural wireless technologies and implement MapReduce [?], a versatile and widely used server technology, making our results broadly applicable. Our analysis, conducted with equipment from the NSF-supported ARA Wireless Living Lab [?], ensures reproducibility of our 5G experiments through open-source Docker containers, supported by the FLOTO [?] fleet management software for large-scale field deployments.

Supercomputing '24, November 17–22, 2024, Atlanta, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 Approach

To allow a distributed computation workload, we deployed six FLOTO instruments within the ARA testbed. ARA provides an extensive 5G radio infrastructure. This infrastructure is extended through several locations, including research farms, and city buses in Ames, Iowa accross the Iowa State University campus. FLOTO devices are located in a six-mile radius, connected to Aviat, Skylark, and USRP 5G radios, and supporting millimeter wave (mmWave) and microwave transmission. **Figure 1** shows our deployment distribution.

2.1 Experimental setup

Each devices was equipped with GT-U7 GPS receivers for precise time synchronization. This allows a precise one-way

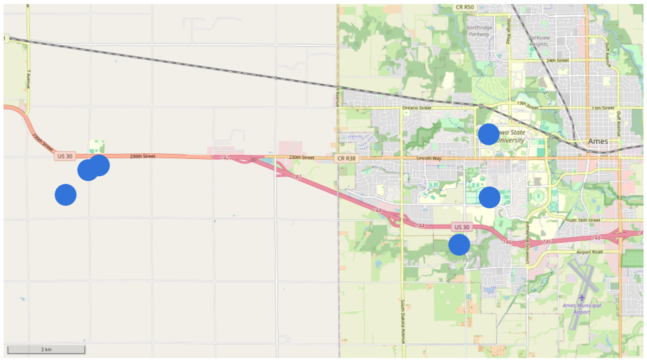

Figure 1. Map of deployment of devices in Ames, Iowa. The leftmost devices are at an Iowa State research farm and the rightmost in the city of Ames.

Path	Dist. (mi)	Throughput (Mbps)	Latency (ms)
1 ⟶ 2	6.55	752	0.642
3 ⟶ 2	6.55	32.6	14.792
4 ⟶ 2	1.08	928	0.571
5 ⟶ 2	6.55	62.7	17.525
6 ⟶ 2	0.91	9.21	15.124
1 ⟶ 4	6.44	70.9	0.702
3 ⟶ 4	6.44	32.5	12.893
5 ⟶ 4	6.44	60.7	22.182
3 ⟶ 1	0.12	30.3	18.882
5 ⟶ 1	0.12	44.8	18.081

Figure 3. 5G network link speeds. Devices 1, 3, and 5 are located at the research farm and devices 2, 4, and 6 are at Iowa State University's main campus.

latency measure between 5G-connected devices. An alternative method without hardware accessories would be to synchronize the system clocks using Network Time Protocol (NTP) [?]. However, NTP is less suitable for our latency experiments due to its reliance on the clocks of other computers on the network, which will To remotely manage and deploy applications to our devices, we leverage the FLOTO fleet management software, which is based on OpenBalena and *k3s*. face a latency delay in transmitting to any device. Our architecture also leverages CHI@Edge [?], a Chameleon edge device testbed, which allows us to compare the 5G system with a wired baseline.

3 Evaluation

We conducted performance evaluations for the GPS-based time synchronization, the 5G network links, and Pi Hadoop clusters of varying sizes.

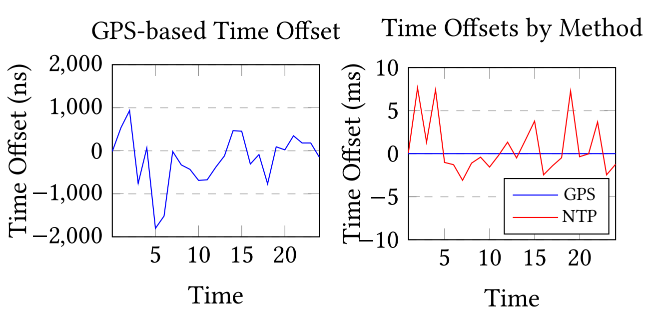

Figure 2. Precision of time synchronization mechanisms in broad daylight with a clear line of sight to the sky.

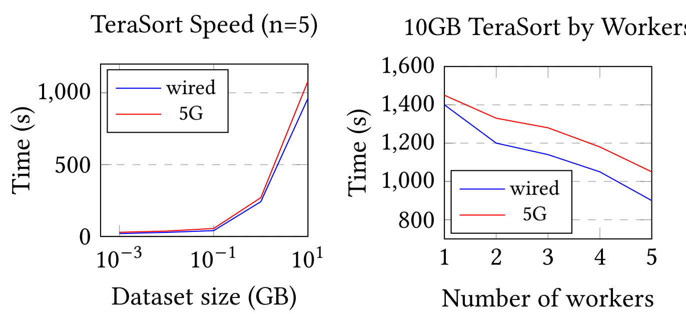

Figure 4. Performance of the 5G distributed network using TeraSort, a standard Hadoop benchmark that sorts a dataset with arbitrary size.

4 Conclusion

Our experimental results indicate that the use of 5G in edge-based distributed computation does not significantly hinder the performance of the TeraSort benchmark. However, the network overhead in this architecture may be greater with more worker nodes due to the requirement to synchronize the data across devices using the Hadoop Distributed File System (HDFS). Additionally, we measured the latency and throughput of wireless connectivity over millimeter radio wave connections and compared the efficacy of network- and GPS-based time synchronization methods. Opportunities for future work include comparing different 5G link technologies for this use case, as well as modifying our architecture to work with other distributed computation tools, such as MR-Edge [?].

Acknowledgments

Results presented in this paper were obtained using the FLOTO project (Award 2213821) supported by the National

Science Foundation. This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357.

References

- [] Jeffrey Dean and Sanjay Ghemawat. 2008. MapReduce: simplified data processing on large clusters. *Commun. ACM* 51, 1 (jan 2008), 107–113. https://doi.org/10.1145/1327452.1327492
- [] Najmul Hassan, Saira Gillani, Ejaz Ahmed, Ibrar Yaqoob, and Muhammad Imran. 2018. The Role of Edge Computing in Internet of Things. *IEEE Communications Magazine* 56, 11 (2018), 110–115. https://doi.org/10.1109/MCOM.2018.1700906
- [] D.L. Mills. 1991. Internet time synchronization: the network time protocol. *IEEE Transactions on Communications* 39, 10 (1991), 1482–1493. https://doi.org/10.1109/26.103043
- [] Alicia Esquivel Morel, William Fowler, Kate Keahey, Kyle Zheng, Michael Sherman, and Richard Anderson. 2023. Chi@Edge Education. https://chi-education.gitbook.io/chi-edge-or-education/ [Accessed: 2024-3-3].
- [] Alicia Esquivel Morel, Mark Powers, Kate Keahey, Zack Murry, Tomas Javier Sitzmann, Jianfeng Zhou, and Prasad Calyam. 2024. IBIS: An Infrastructure Mananagement Framework for Adaptable, Multi-Sensor Data Collection in Scientific Research. *WOCC'24 - Second Workshop on Converged Computing on Edge, Cloud, and HPC* (2024).
- [] Konstantin Shvachko, Hairong Kuang, Sanjay Radia, and Robert Chansler. 2010. The Hadoop Distributed File System. In *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*. 1–10. https://doi.org/10.1109/MSST.2010.5496972
- [] Qian Wang, Brian Lee, Niall Murray, and Yuansong Qiao. 2019. MR-Edge: a MapReduce-based Protocol for IoT Edge Computing with Resource Constraints. In *2019 16th IEEE Annual Consumer Communications Networking Conference (CCNC)*. 1–6. https://doi.org/10.1109/CCNC.2019.8651855
- [] Hongwei Zhang, Yong Guan, Ahmed Kamal, Daji Qiao, Mai Zheng, Anish Arora, Ozdal Boyraz, Brian Cox, Thomas Daniels, Matthew Darr, Doug Jacobson, Ashfaq Khokhar, Sang Kim, James Koltes, Jia Liu, Mike Luby, Larysa Nadolny, Joshua Peschel, Patrick Schnable, Anuj Sharma, Arun Somani, and Lie Tang. 2021. ARA: A Wireless Living Lab Vision for Smart and Connected Rural Communities. In *Proceedings of the 15th ACM Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization* (New Orleans, LA, USA) *(WiNTECH '21)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/3477086.3480837