FogROS2: An Adaptive Platform for Cloud and Fog Robotics Using ROS 2

Jeffrey Ichnovski1A3, Kaiyuan Chen1, Karthik Dharmarajan1A, Simeon Adebola1A, Michael Danielczuk1A, Víctor Mayoral-Vilches1A5, Nikhil Jha1A, Hugo Zhan1A, Edith LLontop1A, Derek Xu1A, Camilo Buscaron, John Kubiatowicz1, Ion Stoica1, Joseph Gonzalez1, and Ken Goldberg1,2A

Abstract—Mobility, power, and price points often dictate that robots do not have sufficient computing power on board to run contemporary robot algorithms at desired rates. Cloud computing providers such as AWS, GCP, and Azure offer immense computing power and increasingly low latency on demand, but tapping into that power from a robot is non-trivial. We present FogROS2, an open-source platform to facilitate cloud and fog robotics that is included in the Robot Operating System 2 (ROS 2) distribution. FogROS2 is distinct from its predecessor FogROS1 in 9 ways, including lower latency, overhead, and startup times; improved usability, and additional automation, such as region and computer type selection. Additionally, FogROS2 gains performance, timing, and additional improvements associated with ROS 2. In common robot applications, FogROS2 reduces SLAM latency by 50%, reduces grasp planning time from 14s to 1.2s, and speeds up motion planning 45x. When compared to FogROS1, FogROS2 reduces network utilization by up to 3.8x, improves startup time by 63%, and network round-trip latency by 97% for images using video compression. The source code, examples, and documentation for FogROS2 are available at https://github.com/BerkeleyAutomation/FogROS2 and is available through the official ROS 2 repository at https://index.ros.org/p/fogros2/

I. INTRODUCTION

It is difficult for the onboard computing resources of robots to keep up with advances in robot algorithms and new computing hardware. Cloud computing offers on-demand access to immense computing resources and new and power-hungry computing platforms, such as GPUs, TPUs, and FPGAs. Prior work [1] showed that using the cloud for intensive computing in robotics can be practical and cost-effective. However, gaining access to evolving cloud computing resources requires expertise with many new and emerging software packages, and experience handling data security and privacy. In prior work [2], we introduced FogROS (henceforth FogROS1), a framework that extends the Robot Operating System (ROS) (henceforth ROS 1) to enable quick access to the cloud.

*Equal Contribution
1Department of Electrical Engineering and Computer Science
The AUTOLab at UC Berkeley (automation.berkeley.edu)
2Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
3Robotics Institute, Carnegie Mellon University
4Acceleration Robotics, Ecuador 3, 1 I, Vitoria, Álava, Spain,
5System Security Group, Universität Klagenfurt, Universitätsstr. 65-67
9020 Klagenfurt, Austria
jeffi@cmu.edu, kych@berkeley.edu

However, FogROS1 has limitations in latency, usability, and automation. In this paper, we introduce FogROS2 to reduce latency, improve usability, and automate additional components of launching robot code in the cloud, while extending to additional potential robot applications. Furthermore, we rewrote FogROS2 from scratch to fully integrate with ROS 2 to benefit from improvements in networking, launch configurability, and its command line interface; and we added integration points to Foxglove [3] to enable remote monitoring from anywhere in the world.

Latency, the time between when an event occurs and when the robot reacts to the event, is a critical factor in many
applications. FogROS1 and FogROS2 demonstrate that using the cloud can reduce latency of complex computations. (It is important to note that this is not a universal statement—some computation, such as feedback control loops, time-bounded, or safety-critical computations are not always suitable for the cloud due to unpredictable network round-trip times.) FogROS1 suffered from long cloud computer startup times (around 4 minutes) and high round-trip network times, particularly for images (e.g., 5 seconds per image). FogROS2 lowers these latencies by using application-specific cloud-computer images, using a Kubernetes backend to avoid overhead associated with creating new cloud computers, switching from TCP to UDP secured networking, and adding transparent H.264 video compression for image topics.

FogROS2 includes several usability and automation improvements over FogROS1 to facilitate adoption, including: (1) a command-line interface (CLI) to interact with FogROS2 cloud computers, (2) integration with Foxglove to enable remote monitoring, (3) a new launch process to allow for automating cloud-computer specification and region, and (4) creation of custom cloud-computer images.

Many of these improvements are facilitated by re-writing FogROS2 for ROS 2. ROS 2 [4], a rapidly growing replacement for ROS 1 [5], is a standard for developing robot applications. FogROS1 and FogROS2 enable moving computationally intensive parts (or nodes) of a computational graph to the cloud and securing communication channels for messages, all with a few small changes to the launch script and without changing a line of the robot code. By migrating to ROS 2, the FogROS2 launch system gains additional capabilities, such as: detecting the cloud server region nearest to the robot and automatically selecting computers and images based computational requirements. FogROS2 is now part of the official ROS 2 ecosystem, and installable with standard Ubuntu commands (apt install ros-humble-fogros2).

In 3 example applications, visual SLAM, grasp planning, and motion planning, we evaluate the ability of FogROS2 to reduce total computation times. We find that using cloud computers via FogROS2 speeds up the computation and can reduce compute times by 2x to 45x. Comparing to FogROS1, FogROS2 improves startup times by 63%, and network latency by 97% for images.

FogROS2 contributes 9 improvements over FogROS1 [2]: (1) FogROS2 extends the ROS 2 launch system introducing additional syntax in launch files that allow roboticists to specify at launch time which components of their architecture will be deployed to the cloud, and which ones on the robot. While FogROS1 existed outside the official ROS ecosystem, FogROS2 directly integrates with it. (2) FogROS2 provides launch script logic that allows robots to automate selection of cloud-computing resources, such as nearest region, computer image, and computer type. (3) FogROS2 adds support for streaming video compression between robot and cloud nodes—significantly improving the performance of image processing in the cloud, and potentially enabling new applications. (4) The architecture of FogROS2 is extensible, making it easy to plug in support for additional cloud computing providers, Data Distribution Service (DDS) providers (Sec. III), and message compression. (5) FogROS2 integrates with ROS 2 tooling and provides ROS 2 command-line interfaces to query and control cloud-robotics deployments. (6) FogROS2 interfaces with the new Foxglove web-based robot visualization software [3] to allow remote (anywhere-in-the-world) monitoring of FogROS2 applications. (7) FogROS2 supports a new backend based on Kubernetes that allows for faster warm starts and broader cloud-service provider support. (8) FogROS2 automates the building of cloud-based virtual machine images for faster startup time. (9) FogROS2 is part of the ROS 2 ecosystem and is accessible with the standard apt install command.

II. RELATED WORK

Robots have limited onboard computing capabilities and as the computing demands of robotics grow, the cloud has become an increasingly necessary source of computing power. Kehoe et al. [6], [7] surveys the landscape of cloud robotics, including capabilities, potential applications, and challenges.

Cloud-robotics platforms facilitate offloading computation and data to the cloud. A notable example is RoboEarth [8], which shared information between robot and cloud. The main use case was to use the cloud to share databases between robots, but it did not leverage the cloud for offloading computing. Rapyuta [9] emerged from RobotEarth to become a platform for centralized management of robot fleets. In Rapyuta, robot nodes or Docker images are built on the cloud and pushed to the registered robots. A similar approach is taken by AWS Greengrass [10]. Using proprietary interfaces, Rapyuta and Greengrass allow building and deploying an entire pipeline for robotics applications [11], [12], [13], [14] from a centralized cloud interface. In contrast, FogROS2 approaches cloud deployment from the opposite perspective—instead of pushing applications from the cloud to a robot, FogROS2 pushes robot nodes from robot to the cloud. It uses an interface familiar to ROS 2 developers, allowing developers and researchers to access cloud resources without learning or conforming to an additional framework.

Researchers have explored using the cloud for grasp planning (e.g., Kehoe et al. [7], Tian et al. [15], and Li et al. [16]), parallelized Monte-Carlo grasp perturbation sampling [17], [18], [19], motion planning services (e.g., Lam et al. [11]), and splitting motion plan computation between robot and cloud (e.g., Bekris et al. [20] and Ichnowski et al. [21]). Researchers also have explored using new cloud computing paradigms as they emerge, such as serverless computing [22], [23], in which algorithms run (and are charged) for short bursts of intensive computing; while others have explored using the cloud to gain access to hardware accelerators such as FPGAs [24]. Others have explored some of these challenges, such as preserving privacy [25] and sharing models between robots [26]. In many of these examples, using the cloud requires a custom one-off implementation or interfacing with a proprietary library. FogROS2 and ROS 2 reduces this complexity.

For a robot to gain access to cloud resources, it must provision a cloud computer and establish a network connection to it.
As robots operate in the physical world, the connection to the cloud must be secured. However, setting this up is an involved process, in some cases requiring 12 steps for configuration and 37 steps for verification [27]. Hajaj et al. [27] explored using SSH tunnelling for communication with ROS nodes running in the cloud. However, SSH tunnels do not support UDP which is needed when using ROS 2 Data Distribution Service (DDS) over UDP (while some DDS implementations support TCP, using TCP can introduce performance issues, and add unnecessary overhead for local communication). Crick et al. proposed rosbridge [28]. Pereira et al. [29] proposed ROS Remote, and Xu et al. proposed MSA [30] as alternate ROS communication stacks with varying degrees of security and modifications required for their use in ROS applications. Wan et al. [31] and Saha et al. [32] propose unifying robot-cloud communication. Lim et al. proposed using VPNs [33], and FogROS2 builds on this approach. FogROS2 allows ROS 2 applications to easily use the cloud without code modification, and with secured communication.

III. BACKGROUND ON ROS

ROS 2 [4], the successor to Robot Operating System (ROS 1), includes many substantial improvements. A core improvement in ROS 2 is its change from a proprietary publication/subscription (pub/sub) system to the industry-standard middleware Data Distribution Service (DDS) [34]. DDS addresses robotics concerns such as providing real-time, high-performance, interoperable, and reliable communication [34]. As DDS is a specification, there are several implementations, and ROS 2 is agnostic to DDS implementation.

In ROS 2, computational units are abstracted into nodes that communicate with each other via a pub/sub system. Nodes subscribe to named topics and receive messages (data) as other nodes publish them. In an example application (Fig. 1), a camera node publishes images, a Simultaneous Location And Mapping (SLAM) node processes the images and publishes a location and map, a Motion Planner node receives the map and then computes and publishes a collision-free path, and a path following node drives the wheels to reach a target.

When orchestrating a robot application, often multiple nodes must be launched simultaneously. The ROS 2 launch system facilitates this by providing the ability to specify all required nodes, topic mappings, and relations between nodes in a single python script file. Launching the robot application is then a matter of running the command:

```
ros2 launch <package> <script>.
```

Listing 1 without the circled FogROS2 extensions, shows an example launch script that launches two nodes, a “grasp_motion” and a “grasp_planner” simultaneously.

IV. APPROACH

At the front end of FogROS2 is the launch system that specifies what nodes to launch and where. Unlike FogROS1, FogROS2’s launch system is scriptable—allowing for launch-time logic to automate parts of the launch process. Listing 1 shows an example in which a grasp planner needs a GPU to run efficiently. The script first defines a cloud machine with a GPU, then adds an attribute to the grasp planner node to tell the FogROS2 launch process to run it on the cloud machine. In more extensive use cases, multiple nodes can run on the same machine, and FogROS2 can launch multiple machines.

The steps FogROS2 takes are (bold items are new to FogROS2): (1) trigger launch from integration with the command-line interface; (2) process the launch script logic (e.g., to automate cloud selections, see Sec. IV-B), (3) connect to the cloud provider through its programmatic interface to create and start a new instance along with setting up access control to isolate from other cloud computers, and generating secure communication key pairs, using the new Kubernetes backend when needed; (4) install the ROS libraries and dependencies on the cloud machine needed for the robot application to run in the cloud (skipped if using a pre-built custom image, see Sec. IV-C); (5) set up Wireguard virtual private networking (VPN) on robot and cloud machine to secure the ROS 2 DDS communication between them; (6) copy the ROS nodes from the robot to the cloud machine; (7) configure the DDS vendor’s discovery mechanism to peer cloud and robot across the VPN.

Listing 1: FogROS2 Launch Script Example. This example launches two nodes. Unlike ROS 1 and FogROS1, which used an XML launch file, FogROS2’s launch files are python scripts. In this example, the FogROS2 launch extensions are circled. The first extension defines a machine on which to launch nodes. The second tells FogROS2 to launch the grasp planner node on that machine.

```
def generate_launch_description():
    ld = FogROSLaunchDescription()
    # configure cloud machine with a GPU (g4dn)
    machine1 = AWSCloudInstance(
        region="us-west-1",
        instance_type="g4dn.xlarge")
    # launch grasp_motion node on robot
    grasp_motion_node = Node(
        package="fogros2_examples",
        executable="grasp_motion",
        output="screen")
    # launch grasp_planner node in on the cloud
    grasp_planner_node = CloudNode(
        package="fogros2_examples",
        executable="grasp_planner",
        output="screen",
        machine=machine1)
    ld.add_action(grasp_motion_node)
    ld.add_action(grasp_planner_node)
    return ld
```
(8) optionally configure streaming video compression (Sec. [IV-E]); (9) launch docker instances; (10) optionally launch Fogxglove for monitoring (Sec. [IV-F]); (11) launch cloud-based nodes; (12) launch nodes on the robot; (13) once launched, users can use FogROS2 CLI integrations interact with FogROS-related cloud computers (Sec. [IV-A]).

Once the launch process is complete, the nodes running on the robot and on the cloud machine(s) securely communicate and interact with each other—and the only change needed was a few lines of the launch script.

A. ROS 2 Command Line Integration

FogROS2 integrates with the ROS 2 Command Line Interface (CLI), offering an intuitive way to interact with FogROS2 cloud instances not available in FogROS1 or ROS 1. To use the CLI, a user types into a terminal window:

```
ros2 fog <command> [args...]
```

where `command` specifies an interaction with FogROS2 along with additional arguments. For example, `command` can be `list`, which lists cloud instances, `delete`, to delete existing instances that are no longer in use, `image`, to create and manage cloud images, or `connect`, to connect via SSH [35] to running instances.

B. Launch Script Extensions

The new launch system in FogROS2 enables custom logic during launch, which was not possible with the launch system in FogROS1. We include several options:

- Since the distance between robot and cloud can dramatically affect network latency, the launch script can select nearest cloud computer based on the robot location.
- Since different computers (e.g., Intel vs Arm and with or without GPU) and regions require different images. The launch script can automate selecting the correct image.
- Selecting the best or most cost-effective cloud computer for a ROS 2 node can require significant effort [36]. By integrating ideas from Sky Computing [37], the launch script can select a machine type based on a specification of requirements (e.g., CPU type and core count, memory size, GPU type and memory, and more).

As an example, changing Listing 1 line 5 to call `region= find_nearest_aws_region()`, automates region selection. This function uses the robot’s IP and a geolocation API to determine the nearest cloud data center. FogROS2 also offers a function to automate finding the cheapest instance type that matches a user’s computing specifications.

C. Cloud Computer Virtual Machine Image Management

Cloud computers with FogROS1 suffered from long startup times, sometimes approaching 4 minutes. A significant portion of this time is due to installing ROS and dependencies on the cloud computer. Advanced users could address this problem by creating computer images with software pre-installed. FogROS2 adds a tool to automate this process using the command-line interface, allowing it to start up pre-installed instances in a fraction of the time of FogROS1.

D. New Kubernetes Backend

FogROS1 supported AWS only. To support additional cloud service providers, FogROS2 integrates a new Kubernetes [38] backend. Kubernetes (see Fig. [1]) is a system that orchestrates running containers, or units of software packages and their dependencies—e.g., a robot ROS node, FogROS2, ROS 2, and underlying operating system components. With Kubernetes, computers can already be on and waiting to run a new container. This allows for significant speedup in startup time. There is a trade-off—machines managed by Kubernetes must already be on, and there can be significant initial delay when starting Kubernetes the first time.

E. Streaming Image Compression

Many robot algorithms depend on fast processing of image and video data, and these algorithms increasingly require hardware acceleration e.g., via GPUs. However, images and videos are data intensive in ROS, and the time to transmit data to the cloud can reduce the advantage of cloud-based acceleration. Processing images in the cloud was possible with FogROS1, but with high latency.

To address this, at launch time FogROS2 can setup transparent streaming compression between robot and cloud. The video compression we use is H.264 [39] from the open-source libx264 [40] library. Using H.264 allows FogROS2 to greatly reduce the latency of processing video in the cloud. We implement a ROS 2 `image_transport plugin` [41] to make the compression transparent to the application—publishing nodes still publish a sequence of images and subscribing nodes still receive a sequence of images. Fig. [3](b) shows how FogROS2 implements transparent streaming compression.

H.264 compression is often hardware accelerated, reducing the CPU utilization required to compress and decompress.
which a cloud node subscribes. The cloud node responds
This setup differs from the examples in FogROS1, thus we
The result (right) is grasps poses (green) and their relative quality score
video compression between robot and cloud. In this experi-
visualize the robot application in a browser.
A. Streaming Video Compression
We evaluate the performance of using streaming H.264
we compare to additional baselines of (a) raw/uncompressed,
TABLE II: ORB-SLAM2 results on FogROS2
Table I: Streaming FogROS2 video compression
example robot publishes images to a ROS topic that FogROS2 transparently
charged and with a 10 Mbps
computations using the cloud, to lower startup latency, to
cloud nodes deployed to AWS unless specified otherwise. This setup differs from the examples in FogROS1, thus we re-run the experiments in FogROS1 [2] to compare FogROS1 and FogROS2 on equivalent hardware.
A. Streaming Video Compression
We evaluate the performance of using streaming H.264
and show it in Table I. We compare to Uncompressed, which is the raw pixel arrays native to ROS 1 and ROS 2, Compressed, which uses (non-streaming) image compression, Theora [42] streaming image compression, and FogROS2’s H.264 compression. From the table, we observe the benefit of streaming video compression between the robot and the cloud, as the cloud can receive images 13× faster FPS, while shortening the latency by 97%. We observe performance improvement of H.264 over Theora, the previous best available compression, with H.264 shortening the latency by 54%. The reduced latency from 1401 ms to 38 ms may enable some real-time cloud-robotics applications not possible without compression.
However, there may be a tradeoff in some applications. Theora and H.264 are both lossy compression algorithms, meaning they are designed to compress videos by discarding some image information. This information loss is tailored to human perception [39], and thus may adversely affect computer vision algorithms.
B. Cloud Robotics Benchmark Applications
We evaluate FogROS2 in a benchmark on 3 example robot applications: SLAM with ORB-SLAM2 [43], Grasp Planning with Dex-Net [44], and Motion Planning with Motion Planning Templates (MPT) [45]. Refer to FogROS1 [2] for further details on these benchmarks. We compare to a baseline of robot-only computing and FogROS1 using equivalent cloud computers. For examples with cloud-based image processing, we compare to additional baselines of (a) raw/uncompressed, (b) PNG compressed, and (c) Theora [42] compressed, where Theora is an open-source video compression library with an existing image transport plugin [41].
In SLAM experiment (Table II), H.264 compression lowers per-frame latency from 0.82 s to 0.24 s, a 3.4× improvement, allowing FogROS2 to achieve lower latency than robot-only
When we test to deploy the robot on the US east coast, the AWS backend must create a new cloud computer each time, and the launch script selects the nearest to them. We test deploy a robot on the US west coast, and the launch script selects the cloud data center that is nearest to them. We test if FogROS2 can shorten startup times compared to FogROS1. Short startup times benefit software development cycles and robots that intermittently operate (e.g., vacuuming robots). In this experiment, we use the new image command in FogROS2 to generate a custom computer image, and measure the time between launch and first robot-cloud ROS2 node interaction with and without the custom image. For comparison, we manually create a custom image for FogROS1 by using the AWS web console.

In Table IV, we observe that the custom image in FogROS2 reduces AWS startup times by 63%.

Using Kubernetes on local cluster or Google Cloud Platform (GCP) reduces startup times further, due to having the computers in the Kubernetes cluster already running. The AWS backend must create a new cloud computer each time, accounting for approximately 40 s of delay.

There is a tradeoff to be made in startup times. Kubernetes requires starting up a cluster of computers, which can take on the order of 10 minutes. If one is willing to spend this time upfront, Kubernetes allows users to redeploy ROS nodes over and over, which may be beneficial when rapidly prototyping changes to robot code.

TABLE III: Dex-Net results on FogROS2 We measure compute time in seconds for 10 trials on a robot with a CPU, and compute and network time using cloud computer with an Nvidia T4 GPU via FogROS2.

| Scenario | Robot | Cloud | FogROS1 | FogROS2 |
|----------|--------|--------|---------|---------|
| Uncompressed | 14.0 | 0.6 | 5.0 | 5.6 |
| Compressed | 14.0 | 0.6 | 1.3 | 1.9 |
| H.264 | 14.0 | 0.6 | - | 0.6 |

The best results are shown in bold. A Table III shows example round-trip network times in this experiment, with the best latencies in bold and automatically selected by the geolocation script.

### C. Cloud-Computer Startup Times

We test if FogROS2 can shorten startup times compared to FogROS1. Short startup times benefit software development cycles and robots that intermittently operate (e.g., vacuuming robots). In this experiment, we use the new image command in FogROS2 to generate a custom computer image, and measure the time between launch and first robot-cloud ROS2 node interaction with and without the custom image. For comparison, we manually create a custom image for FogROS1 by using the AWS web console.

In Table IV, we observe that the custom image in FogROS2 reduces AWS startup times by 63%.

Using Kubernetes on local cluster or Google Cloud Platform (GCP) reduces startup times further, due to having the computers in the Kubernetes cluster already running. The AWS backend must create a new cloud computer each time, accounting for approximately 40 s of delay.

There is a tradeoff to be made in startup times. Kubernetes requires starting up a cluster of computers, which can take on the order of 10 minutes. If one is willing to spend this time upfront, Kubernetes allows users to redeploy ROS nodes over and over, which may be beneficial when rapidly prototyping changes to robot code.

### D. Automating Region Selection

We test the launch script extensions for automatic region selection to allow robots to select the cloud data center that is nearest to them. We test deploy a robot on the US west coast, and the launch script selects the us-west-1 data center. When we test to deploy the robot on the US east coast, the robot selects the us-east-2 data center. Table V shows example round-trip network times in this experiment, with the best latencies in bold and automatically selected by the geolocation script.

### V. Conclusion

We present FogROS2, an adaptive cloud-robotics platform for running compute-intensive portions of ROS 2 applications in the cloud. FogROS2 addresses 9 shortcomings of FogROS1, integrates with the ROS 2 launch and communication systems, to provision and start cloud computers, configure and secure network communication, install robot code and dependencies, and launch robot and cloud-robotics code. As a redesigned and distinct successor to FogROS, FogROS2 supports ROS 2, transparent video compression, improved performance and security, access to more cloud computing providers, and remote visualization and monitoring. In experiments, we observe a significant performance benefit to using cloud computing, with the additional improvement from transparent video compression.

In future work, we will continue to improve performance and capabilities of FogROS2. We will explore additional models of computing, such as serverless, spot instances, and more. We will also explore extending the networking capabilities of FogROS2 to allow multiple robots to communicate, collaborate, and share data more easily [48].

### Acknowledgements

This research was performed at the AUTOLab at UC Berkeley in affiliation with the Berkeley AI Research (BAIR) Lab, the CITRIS “People and Robots” (CPAR) Initiative, the Real-Time Intelligent Secure Execution (RISE) Lab, and by the NSF/VMware Partnership on Edge Computing Data Infrastructure (ECDI) Secure Fog Robotics Project Award 1838833. We thank our colleagues for helpful discussions and testing. We thank Brandon Fan and Emerson Dove for their contributions and discussion related to the new Kubernetes backend.
planning framework for robots with low-power CPUs,” in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 2019.

[46] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, “A benchmark for the evaluation of rgb-d slam systems,” in Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), Oct. 2012.

[47] I. A. Şucan, M. Moll, and L. E. Kavraki, “The Open Motion Planning Library,” IEEE Robotics & Automation Magazine, vol. 19, no. 4, pp. 72–82, Dec. 2012. [Online]. Available: http://ompl.kavrakilab.org

[48] K. Chen, J. Yuan, N. Jha, J. Ichnowski, J. Kubiatowicz, and K. Goldberg, “Fogros g: Enabling secure, connected and mobile fog robotics with global addressability,” 2022. [Online]. Available: https://arxiv.org/abs/2210.11691