

Short-lived Datacenters

Michael Wawrzoniak¹, Ingo Müller¹, Rodrigo Bruno², Ana Klimovic¹, Gustavo Alonso¹

¹Systems Group, Dept. of Computer Science, ETH Zürich

{michal.wawrzoniak,ingo.mueller,aklimovic,alonso}@inf.ethz.ch

²INESC-ID / Técnico, ULisboa Lisbon, Portugal

rodrigo.bruno@tecnico.ulisboa.pt

Abstract

Serverless platforms have attracted attention due to their promise of elasticity, low cost, and fast deployment. Instead of using a fixed virtual machine (VM) infrastructure, which can incur considerable costs to operate and run, serverless platforms support short computations, triggered on demand, with cost proportional to fine-grain function execution time. However, serverless platforms offer a restricted execution environment. For example, functions have limited execution times, limited resources, and no support for networking between functions. In this paper, we explore what it takes to treat serverless platforms as short-lived, general purpose datacenters which can execute unmodified existing applications. As a first step in this quest, we have developed Boxer, a system providing an execution environment on top of existing functions-as-a-service platforms that allows users to seamlessly migrate conventional VM-based cloud services to serverless platforms. Boxer allows generic applications to benefit from the fine-grain elasticity of serverless platforms without having to modify applications to adopt a restrictive event-triggered programming model or orchestrate auxiliary systems for data communication. We implement Boxer on top of AWS Lambda and extend it to transparently provide standard network interfaces. We describe its implementation and demonstrate how it can be used to run off-the-shelf cloud applications with a degree of fine-grained elasticity not available on traditional VM-based platforms.

1 Introduction

Serverless computing platforms such as AWS Lambda [4], Google Cloud Functions [13], and Azure Functions [23] offer highly elastic computing and fine-grained resource billing (as low as milliseconds [5]), while raising the level of abstraction for user interaction with the cloud [27]. Although serverless platforms cost more per hour than resources on virtual machines [8], the fine-grained cost model still makes serverless computing cost-effective for jobs with short execution times and sporadic invocation patterns [11, 24, 25].

Today’s serverless offerings are, however, impractical for a wide variety of applications [10, 14, 24, 25, 31]. Current systems limit the execution time of tasks; some restrict the rate at which functions can be created [23] or the number of concurrent functions [22]; and none of the platforms support heterogeneous hardware [6]. However, the most critical limitation is the lack of direct communication among functions which has wide-ranging implications for application design and performance. Current attempts at generalizing commercial serverless platforms implement inter-function communication by exchanging data through remote storage systems such as Amazon S3, Anna, and Pocket [3, 20, 35]. Doing so adds considerable overhead and also increases cost [24, 25]. For instance, parallelism can be used to hide the latency overhead of exchanging data through Amazon S3, with Peron et al. reporting an S3 latency of 14ms to read 256KB of data [25]. The consequences are higher costs (each S3 GET and PUT request incurs cost) and a more complex system design which requires parallel reads to a storage system and additional network round trip times.

In this paper, we explore how serverless computing platforms can be leveraged to run generic applications. Our goal is to provide a general-purpose solution that turns a serverless function into the equivalent of the Linux execution environment that most cloud applications expect. We identify networking as a key challenge, since generic cloud applications rely on direct communication between nodes while serverless platforms do not support it. We propose Boxer, a system that provides a familiar Linux execution environment on top of functions-as-a-service (FaaS) platforms. Boxer exposes a Linux socket API for networking inside functions and supports direct, transparent, TCP/IP communication. Boxer uses a NAT-punching technique to provide bi-directional TCP/IP communication to functions without the external proxies required by existing work [10, 11, 31] or specializing to a single use case [32].

In addition to the design of Boxer, we make several contributions. First, we demonstrate that Boxer can run applications that require inter-function communication on AWS Lambda

and achieve performance comparable to using similarly sized VMs (§ 5). Second, since Boxer provides the application with a standard network interface, we show in § 6 that it supports running unmodified applications, such as the DeathStarBench microservice benchmark suite[12]. Third, we show that Boxer allows cloud users to leverage the fine-grain elasticity of serverless platforms to quickly absorb load bursts in applications and optimize cost by migrating applications seamlessly between VM and serverless environments depending on the load pattern (§ 7). Through these capabilities, Boxer opens up many new opportunities for serverless computing (§ 8.1). By providing conventional networking capabilities using standard interfaces, Boxer enables a general-purpose, on-demand, short-lived datacenter that can run unmodified, distributed applications and is not limited to trivially parallel jobs. We have implemented and evaluated Boxer on top of AWS Lambda and we plan on making it available as open source.

2 Background

Leveraging serverless platforms as short-lived, general purpose datacenters requires bridging the gap between the requirements of distributed applications and the functionality provided by today’s serverless platforms. To provide the necessary background, we describe the execution environment of serverless platforms, focusing on AWS Lambda as the most advanced serverless platform in the market.

2.1 Serverless Execution Environments

Packaging and invoking functions: To run a serverless function, users register the function and specify its deployment package, invocation triggers, and memory resource allocation. For the deployment package, AWS Lambda allows users to supply source code or application binaries. AWS Lambda also supports container images of up to 10 GB in size [7]. In addition to the deployment package, users also configure events to trigger function invocations. Examples of event sources include storage and message queue service notifications, timers, and HTTP requests. Finally, users specify a memory requirement for the function. The cloud provider allocates CPU resources for the function proportionally to its memory allocation. As of October 2021, AWS Lambda supports functions with 128 MB to 10 GB of memory and up to 6 vCPUs (virtual CPUs). Every invocation of a function is executed in its own isolated environment, running a kernel derived from Linux 4.14 and a file system image based on the Amazon Linux [2] distribution. Serverless functions execute as unprivileged userspace processes with restricted capabilities in a secure and isolated runtime environment.

Dynamically loading dependencies: When a function process is invoked and begins executing, a dynamic linker in the AWS Lambda execution environment loads the shared libraries required by the executable. The dynamic linker uses

a set of rules to locate the objects that have to be loaded to satisfy the dependency list. It is also possible to direct the dynamic linker to load additional shared libraries before any other shared libraries are loaded. Objects and functions exported by the additional libraries can be then used by the executable. If the exported names are the same as the exported names from other libraries loaded later, the executable will use the function or objects from the additional library that is loaded before others. This provides a mechanism to intercept function calls to dynamically linked system libraries such as libc. We use this mechanism to embed Boxer’s interposition library in the system when deploying an application so that network and other operations are selectively routed through Boxer.

2.2 Serverless Networking

AWS Lambda functions are assigned private IP addresses from private subnets. They are able to send network traffic to external addresses as the traffic from a function to the public internet is routed through Network Address Translation (NAT) gateways. A NAT gateway forwards the traffic to the destination address but changes the source address to a publicly routable address available to the gateway. The assigned mapping between the internal source address and the externally routable source address is stored in the gateway state. As the network traffic arrives at the destination host, its source address is the address assigned by the NAT gateway. When the destination host sends traffic back to the source address it received the traffic from, the traffic arrives back at the gateway having the previously assigned mapping to an internal destination address. Based on the mapping, the gateway modifies the destination address of the traffic and routes it towards the original private source address assigned to the function. For the function, the traffic appears to come directly from the destination host. This results in AWS Lambda functions being able to transparently access external services and internet resources.

Serverless functions cannot accept connections initiated by external sources or by other functions. AWS Lambda restricts the incoming network traffic for functions to traffic that directly corresponds to previously outgoing traffic. Even if an external host learned the private IP address assigned to a function, it could not route traffic to it. Network traffic destined for an internal address must be routed through an external gateway that will then appropriately route it through the private network. If the external gateway is a NAT, then just knowing the route is also not sufficient. A NAT gateway will drop any incoming traffic not matching its previously configured address translation map. Hence, in the case of TCP, a function process can initiate a connection to an external address and then receive traffic on that connection but it cannot accept new incoming TCP connections.

3 Related Work

3.1 Communication Through Proxies

Prior work has circumvented the lack of networking by using proxy or coordinator servers to initiate connections and relay messages between serverless functions. Fouladi et al. proposed *mu*, a framework for orchestrating parallel computation and communication across serverless workers, as part of their work on ExCamera [11]. The *mu* framework uses a long-lived coordinator and rendezvous server for inter-function communication. Workers running short-lived lambda function invocations establish a connection with the *mu* coordinator, which can instruct workers to communicate between each other through a rendezvous server. The *mu* rendezvous server buffers and relays messages from source to destination lambdas. The authors point out that the rendezvous server’s connection to workers can become a bottleneck and recommend implementing direct communication between lambdas via a hole-punching NAT-traversal strategy, but leave this as future work. In a similar approach, Wang et al. propose using serverless functions to build InfiniCache, a distributed in-memory cache [31] that uses a proxy to which functions connect and is used to relay messages to and between serverless functions. A number of other projects, e.g. [10], use a similar approach to enable communication.

Using an fixed infrastructure service that runs parallel to the serverless functions defeats some of purpose of using serverless. Initializing the supporting infrastructure and the proxy/coordinator add overheads, causing starting delays that serverless purportedly removes. Leaving the infrastructure running incurs costs and requires the maintenance, again defeating one of the main goals of serverless.

We are not the first one to point out the limitations of existing approaches. There was an open-source Serverless Networking SDK that allowed functions to communicate over UDT [36], a UDP-based protocol (the project is no longer active). It relied on a custom API instead of supporting conventional sockets [28]. Solutions have also been proposed to address the networking overhead of starting thousands of functions at the same time [30] or the overhead of sustaining many RPC calls [18]. These solution propose alternative architectures to current commercial deployments and rely on conventional containers or VMs to avoid the limitations of serverless system. In contrast, we are interested in solving these same issues but using the infrastructure provided by cloud vendors. It has been claimed that, in such settings, adding networking makes no sense [27] because it would remove the ability of the cloud vendor to optimize the deployment. As Boxer shows, it is possible to add networking to FaaS without affecting any of its underlying properties or restricting the possibility of optimizing the deployment.

3.2 Communication Through Storage

Applications that require to move larger amounts of data typically resort to an alternative design. Instead of using a coordinator or a proxy, they use ephemeral storage systems, such as Pocket [20], Anna [34], and Locus [26], to improve the performance of inter-function communication through remote storage. In such systems, communication is implemented by writing and reading to a specialized storage layer built for serverless functions. In other cases, rather than using a specially built systems, Amazon’s S3 is used directly. For instance, Lambada [24] and Starling [25] are query analytical systems that implement exchange operators as read/write patterns on Amazon S3 as a way to allow functions running parts of a query to exchange data. In both cases, a relatively complex design is needed to deal with the overhead of communication through storage.

While communication through storage solves the problem, it has both a high cost both monetary and in performance. Writing and reading to storage services like S3 is not for free. Implementing communication through storage induces a large amount or read and writes, increasing the cost of using serverless. Performance-wise, using storage to exchange data not only suffers from the higher latency of storage but also adds communication rounds to the exchange. This overhead becomes significantly visible since now AWS Lambda is charged per millisecond and the observed delay is potentially in the order of tens of milliseconds [25, 32].

3.3 General Purpose Serverless

It has been noted that serverless platforms can be used as a supercomputer on-demand to run highly parallel jobs as it is done in *gg*, a framework and collection of command line tools to help users run everyday applications – such as software compilation and unit tests – seamlessly on serverless platforms [10]. Users express their applications as a composition of lightweight, functional containers using *gg*’s intermediate representation, while the framework takes care of instantiating containers as serverless functions, loading dependencies, and dealing with function failures and stragglers. This idea of a *supercomputer-by-the-second* also appears in previous work of the same authors[11]. Recently, serverless has been presented as the next generation of cloud computing, reinforcing the idea of general purpose use instead of the narrow uses cases considered today [27].

In this paper we pursue the same notion but focus on generalizing serverless to support unmodified, distributed applications. We aim to turn serverless into a true sub-second datacenter that can be used by off-the-shelf applications rather than just for trivially parallel jobs.

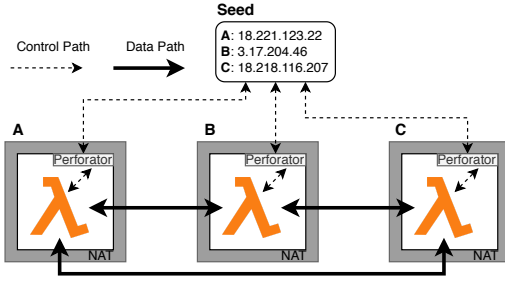


Figure 1: Networked serverless functions use a Seed process to connect functions during the startup. After the startup phase, the seed process is no longer needed.

4 Boxer

As a first step towards exploring the notion of the sub-second datacenter, we have developed Boxer, a system that transparently enables networking between serverless functions. We choose to focus on datacenter applications that use the conventional socket networking API in Linux and the TCP networking protocol, as this represents a large class of datacenter applications.

4.1 System Overview

Boxer uses two key mechanisms to allow external hosts to initiate connections with serverless functions and functions to communicate with each other transparently.

First, to allow incoming traffic to be routed to a serverless function, Boxer uses NAT punching techniques that resemble those proposed in the past for generic NAT services [9, 17]. However, our solution is tailored to the AWS Lambda environment and uses functionality specific to AWS Lambda. Through these techniques, Boxer provides the sending host with the appropriate gateway address and configures the gateway with the appropriate state for the traffic to traverse the NAT and arrive at the private IP address of the destination function. Second, to make the NAT traversal transparent to applications, Boxer intercepts some dynamically linked libc library function calls by leveraging the dynamic linker that is invoked in the Linux-based AWS Lambda execution environment. For example, when a serverless function calls `connect`, the Boxer `connect` implementation will execute instead of the original libc implementation – this allows Boxer to implement NAT traversal without requiring application modifications.

Boxer’s networking environment consists of 1) Boxer processes running inside each serverless function and each node that needs to initiate connections to a function, and 2) a Boxer ‘seed’ process (Figure 1), which runs on an externally routable host and is used for network initialization. Each Boxer process consists of a *networking service* to establishes TCP connections, a *coordination service* to send and receive network

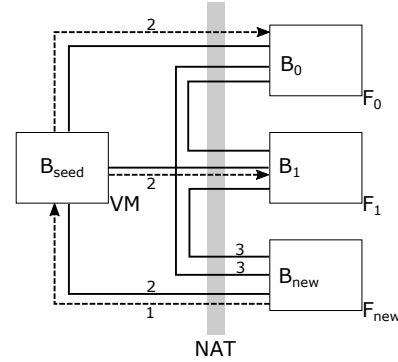


Figure 2: New Boxer node B_{new} in function F_{new} joins already connected nodes B_0, B_1 running in functions F_0, F_1 and seed node B_{seed} running in a virtual machine VM . Solid lines represent control connections, dashed lines messages sent, numerical labels the stage of the join protocol when message is sent or connection is established.

membership updates from other Boxer processes, and a *transparent execution service* to selectively intercept library calls and run unmodified Linux TCP sockets-based applications. We describe each of these services in more detail below.

Boxer is packaged as a Linux executable and a shared library object, as we describe in Section 4.6. Users include the Boxer package in each function deployment package when registering functions on the cloud platform. This allows a Boxer process to start executing in the serverless function environment when the function is invoked. Users are also responsible for spinning up a Boxer seed process on an externally routable host, such as an EC2 instance.

4.2 Networking Service

The networking service in Boxer provides TCP connectivity and manages NAT configuration. We start by describing how nodes join the network and then describe how TCP connections are established.

Joining the network: The Boxer seed process is responsible for network initialization (Figure 2). The user sets a parameter to specify the address of the seed process. The seed can be any other Boxer process that is reachable by the process being started. Since functions cannot establish connections to other functions before Boxer is initialized, the seed process must be running outside of AWS Lambda. We currently run it in an AWS EC2 instance, but it could be run anywhere routable from Boxer processes that are attempting to join the network.

The seed process performs three functions. First, it informs the connecting process of its observed external address. If the process is connecting from AWS Lambda, this is the external address assigned by the NAT gateway used. Second, the coordinator service of the seed process informs all other

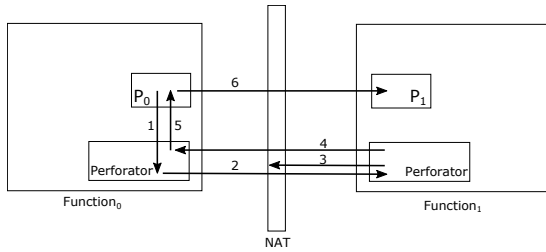


Figure 3: Opening TCP connection by process P_0 to process P_1 running in a remote function.

Boxer processes in the network of a new node, and it supplies the joining process with a list of all other nodes that already joined. Third, the seed process ensures that each external address only exists once. This deals with the fact that two concurrent function invocations may be handled by the same machine behind the same NAT gateway and would thus be indistinguishable. If that happens, all but the first node with a given address are rejected from joining the network, so the corresponding Boxer process exits immediately and all of the function resources are released.

When a new node joins the network, its coordinator service learns the addresses of all other nodes in the network. Concurrently, coordination services of all other nodes in the network are updated with the address of the newly joined node. The Boxer networking service of every node listens to membership updates and treats these events as signals that it is time to establish direct control connections with the new node. Control connections are used to exchange commands between Boxer processes. The membership updates contain the expected external address of the Boxer network control service. At this point, all nodes have enough information to establish new control connections through the NAT. Boxer nodes establish the control TCP connections between the new node and the rest of the nodes in the network. This procedure is repeated for every joining node resulting in $N \times N$ connectivity through TCP connections, one for each pair of nodes in the network.

Providing TCP connections: After the Boxer networking service is initialized as described above, every node in the network is ready to provide TCP connectivity to its local Boxer process. The Boxer networking service provides NAT setup service for its local Boxer process and to Boxer processes running on remote nodes. In the current version, the Boxer networking service does not open TCP connections on behalf of processes; instead, it sets up the appropriate state in the NAT so that when the process attempts to open a TCP connection to another function, the network traffic will traverse through the NAT and the connection can be established.

Figure 3 illustrates the TCP connection process. A Boxer process wishing to establish a TCP connection to a remote Boxer process cannot immediately open a connection to the remote address because the NAT would block it. Instead, (1) the

process sends a request to the local Boxer networking service to request that the NAT state be configured for the specified connection. If the connection is to an address that is a part of the Boxer network, then (2) the request is forwarded to the appropriate remote Boxer network service using the control connections configured during the initialization (described above). The remote Boxer network service (3) performs the NAT configuration and then (4) sends an acknowledgment back to the requesting Boxer service that the NAT is ready. (5) The confirmation is forwarded back to the process that made the original request. At this point, the requesting process can (6) attempt opening a TCP connection to the previously specified remote address. Because the remote Boxer configured the NAT, opening the connection will traverse the NAT and reach the remote process. Assuming the remote process is listening for incoming connections at the specified address, the TCP connection is established.

4.3 Coordination Service

Every Boxer node runs a simple coordination service. The service exposes an interface for local processes to stream membership updates as new nodes join the network. Every membership service instance allows other nodes to become children and propagate the updates. Currently, the seed node is used as the root of the propagation tree. The coordination service is used during the initialization process (described above) by the networking service to establish the control connections. The execution service (described below) uses the coordination service to determine when to proceed with the application execution and to provide the application with the list of peer addresses.

4.4 Transparent Execution Service

The Boxer execution service provides functionality to schedule the execution of applications in a transparent manner. To use Boxer, the programmer specifies a command for Boxer to run when the Boxer process starts executing. The command can either be run immediately or after a specified number of members have joined the Boxer network. For example, `perforator -s SEED_ADDR -n 5 zk-start.sh` will run the `zk-start.sh` command once the network has five members. This mode of execution is particularly useful to start datacenter applications that require a static network configuration. If a barrier for a target number of workers is specified, Boxer runs the specified command and sets its environment variables to indicate the local node address, a unique node ID assigned to the current node, and a file name of a file containing the list of addresses for all other nodes. If no barrier is specified and the command runs immediately, Boxer will still learn of network membership changes during execution via the coordination service. To run datacenter applications, users may write small shell scripts that generate the necessary

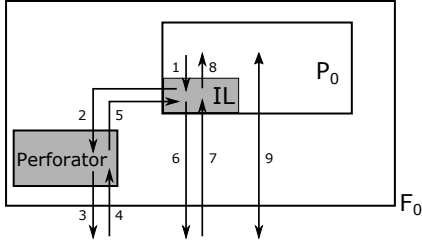


Figure 4: Transparent application execution via the interposition library (IL). Application process P_0 uses socket `connect` function to open a TCP connection to a remote function.

configuration files based on the environment variables and then start the application.

4.5 Interposition Library

To enable applications to establish connections using Boxer without requiring changes to application code, when Boxer runs commands, it instructs the dynamic linker to load the Boxer interposition library before all other shared libraries are loaded and the application execution begins. The Boxer interposition library is designed selectively to intercept functions provided by other shared libraries that are loaded by the dynamic linker.

We design Boxer to provide transparent networking support for applications that use the socket interface, which is generally provided by a version of the `libc` system library on Linux systems. Applications use the `connect` function on stream sockets to open TCP connections. When the interposition library is loaded, applications use the `connect` function exported by the interposition library, instead of the default version provided by `libc`. This provides the primary mechanism for intercepting the necessary function calls to provide transparent sockets semantics using Boxer networking to establish TCP connections.

Figure 4 illustrates the process of transparently opening remote TCP connections by an application running in a function with Boxer. Every time an application process calls the `connect` function, it is intercepted by the interposition library (1) If the call is not associated with a TCP socket, the interposition library immediately forwards the call to the default implementation of `connect`. Otherwise, if the socket is not already bound to a local address, it is bound to a local address assigned by the kernel, and then the local and remote addresses associated with the socket are used to construct the appropriate connection request command. The command is then sent to the local Boxer networking service. (2) The networking service forwards the request to the appropriate remote node (3) and, when the result is received, (4) it reports it back to the library. (5) If a successful acknowledgment is returned from Boxer networking, that indicates that the NAT is configured, so the interposition library (6) establishes a TCP

connection using the default version of the `connect` function. When the system `connect` function completes, (7) the result of the final call is returned to the calling application. (8) The application does not observe any functional difference from the default `connect` call. If the connection was successful, the application continues to use the TCP connection in the same way as if Boxer was not involved in the process. No function calls for sending and receiving data over the socket are intercepted; instead, (9) default library functions are used directly.

In addition to the `connect` function, to provide transparency to the applications accepting TCP connections, the `bind` socket function is also intercepted by the interposition library. The NAT configuration process performed by the networking service requires binding a temporary socket to the same local address as the socket that an application accepts connections on. To make this possible, the application's listening socket must have the `SO_REUSEPORT` socket option set. To support unmodified applications, Boxer modifies this socket option for the application by intercepting `bind` calls (the same effect could be achieved by intercepting other functions as well). It should be noted that this is an example where perfect transparency is not preserved. The application can detect that the option was changed on the socket, and if the application relies on this option not being set, then it could have an impact on the application. However, we do not aim to cover unusual corner cases, and typical datacenter applications are functionally not affected.

4.6 Packaging Applications with Boxer

In order to run existing applications in serverless functions, they have to be packaged together with Boxer using the deployment process of the cloud provider. A typical application package consists of four components: (1) the event handler the function service calls when the function is invoked, (2) the Boxer executable and interception library, (3) the application itself including templates of its configuration files, and (4) a dynamic configuration script. In AWS, we have used dependency layers for the Boxer and application code and deployed the handler and configuration scripts as the function code. Boxer introduces a small function deployment package size overhead resulting from the Boxer executable and interception library. These components account for 3.9 MB and 2.8 MB before compression, and 873 KB and 545 KB after compression (Boxer and interception library, respectively). In total, the package size increases by 1.4 MB, a reasonable overhead considering the current deployment package size limit of 50 MB.

The event handler typically starts the networking service such that the newly invoked function joins the network. The address of the seed process, which is required for that process, can be passed as a function parameter in the invocation or read from some pre-defined place in the cloud. The exe-

TCP connection type	Mean		Median		Std.		Min		Max	
	Forward	Reverse	Forward	Reverse	Forward	Reverse	Forward	Reverse	Forward	Reverse
Function-to-Function	622.57	622.63	626.88	628.59	24.32	25.81	564.25	561.16	680.63	678.36
VM-to-Function	428.28	426.03	429.03	428.09	1.69	4.01	420.31	415.62	429.16	429.18
Function-to-VM	410.31	427.05	422.64	428.74	26.58	3.61	335.70	414.39	429.06	430.33
Function-to-VM-native	427.77	426.67	429.04	428.62	3.16	4.93	412.74	399.35	429.07	429.07
VM-to-VM	428.89	428.96	428.96	428.95	0.55	0.29	424.78	427.50	429.01	430.39
VM-to-VM-native	429.02	429.06	429.03	429.07	0.15	0.02	428.43	429.02	429.65	429.09

Table 1: TCP throughput in Mbit/s for different connection types measured using iperf3[16] Forward and Reverse modes refer to throughput achieved by the client side generating traffic or server side respectively. VMs are EC2 m4.large instances.

TCP connection type	Mean		Median		Std.		Min		Max	
	Forward	Reverse	Forward	Reverse	Forward	Reverse	Forward	Reverse	Forward	Reverse
Function-to-Function	621.48	621.44	628.42	627.95	23.87	24.28	560.61	560.99	674.14	677.28
VM-to-Function	622.98	624.10	610.70	623.97	26.59	32.31	595.92	566.07	680.64	681.22
Function-to-VM	623.18	621.78	637.01	640.82	26.99	32.22	564.00	570.34	658.91	664.03
Function-to-VM-native	624.13	622.39	639.62	622.06	27.75	31.24	566.61	571.08	657.34	668.00
VM to VM	4684.53	4706.22	4744.35	4735.99	144.29	103.36	4034.14	4263.13	4789.13	4787.11
VM-to-VM-native	4770.42	4695.97	4786.23	4740.02	39.46	159.41	4571.73	3903.98	4787.20	4789.87

Table 2: TCP throughput in Mbit/s for different connection types measured using iperf3[16] between 32 pairs of hosts for each type. Connection types are the same as in Table 1. The measurements were conducted in AWS eu-west-3 region, VMs are EC2 m5.large instances.

cution service then starts the dynamic configuration scripts; depending on the selected start-up mode, this happens either immediately or after enough nodes have reached the barrier. The configuration script consults the coordination service for the current list of nodes in the network and their addresses and generates the runtime configuration of the application using the configuration template based on that list. When everything is set up, the script runs the application itself.

5 Benchmarking Boxer

To evaluate Boxer in practice, we start by characterizing its throughput and latency.

5.1 Throughput analysis

We summarize the networking characteristics we observe in AWS Lambda using Boxer. We show what a typical application running in AWS Lambda with Boxer can achieve in terms of TCP throughput and latency. Unless noted otherwise, all measurements were performed in AWS us-west-2 region with Lambda functions with 3008MB of memory and m4.large EC2 VM instances.

Boxer enables TCP connectivity between AWS Lambda serverless functions and allows outside hosts, running outside of AWS Lambda to initiate TCP connections to serverless functions. As described in Section 4.2 this is achieved by traversing the NAT gateways between hosts. During development we have seen unfavorable network conditions when Boxer used different methods to establish connectivity. This

is because the cloud applications we want to enable usually expect symmetric network properties between end-points. However, given that Boxer traverses an unknown network of middle-boxes that may impose arbitrary network filtering or throttling rules, we must verify the properties of the various scenarios the middle-boxes could differentiate, based on ordering of packets, timing or types of end-hosts. Thus, in the evaluation we distinguish six connection types; (1) Function-to-Function connections are established by Boxer between a pair of AWS Lambdas, (2) VM-to-Function are connections initiated by Boxer running in an EC2 VM to an AWS Lambda also running Boxer, (3) Function-to-VM are initiated by Boxer running in AWS Lambda to a EC2 VM also running Boxer, (4) Function-to-VM-native are connections initiated from AWS Lambda to EC2 VM without the use of Boxer (this scenario is allowed by default on AWS Lambda) as it is used a baseline, (5) VM-to-VM are connections established using Boxer between a pair of EC2 VMs and (6.) And as a baseline, VM-to-VM-native are vanilla connections established between a pair of EC2 VMs without Boxer. The evaluation also proves the versatility of Boxer and the many configurations in which it can be deployed.

We benchmark TCP throughput of different connection types between pairs of hosts (VMs or functions) by running unmodified iperf3[16] tool as a Boxer application (or natively for the native connection types). We instantiate 32 non-overlapping pairs of functions for a 60 seconds period for each scenario. In each pair, one function runs iperf3 in server mode, and one in client mode. The iperf3 client function connects to the listening server to begin the configured

TCP connection type	Round-trip latency of 1k byte message (μs)					Connection establishment time for time-to-first byte (μs)				
	Mean	Median	Std.	Min	Max	Mean	Median	Std.	Min	Max
Function-to-Function	694.23	758.00	289.52	202.00	2769.00	2735.21	2625.00	10001.00	890.00	1033112.00
VM-to-Function	547.84	457.00	194.29	244.00	2471.00	1981.38	2153.00	7909.74	821.00	1011171.00
Function-to-VM	520.10	436.00	189.53	244.00	2372.00	2086.03	2239.00	6124.57	882.00	1015205.00
Function-to-VM-native	622.53	686.00	175.57	241.00	2434.00	1378.56	1244.00	8027.04	382.00	1012935.00
VM-to-VM	193.69	188.00	43.34	150.00	2165.00	1067.24	1034.00	166.09	894.00	7402.00
VM-to-VM-native	197.62	194.00	28.18	153.00	1862.00	407.81	345.00	284.37	258.00	6416.00

Table 3: Round-trip latency and connection establishment times for different TCP connection types.

benchmark. When configured in the forward mode, the client side generates TCP traffic, when configured in the reverse mode, the server side generated the TCP traffic. We measure both to verify that the underlying network does not apply different network policies in different directions (we have seen this during development.) The achieved throughput reported on the receiving side at 1 second interval.

Table 1 presents throughput statistics for different connection scenarios. The sustained average throughput is 622Mbit/s in forward and reverse direction between a pair of AWS Lambda functions running Boxer (Function-to-Function). The variance level is low; the throughput is steady and sustained throughout the connection. In an additional experiment, we verified that the throughput can be sustained throughout the maximum lifetime of an AWS Lambda function (currently 15 minutes). The observed throughput between a pair of VMs is 429Mbit/s if Boxer is used or not, demonstrating that Boxer adds no data-plane overhead (after a connection is established). Throughput between functions and VMs is similar (410-428Mbit/s) and symmetric in all connection scenarios. In this case it is limited by the throughput of the VMs (m4.large) network. Table 2 shows the same benchmarks but using a higher-bandwidth VM network (m5.large instances): the upper-bound on the throughput between AWS Lambda and VMs is the same as the throughput of the AWS Lambda internal network of 621Mbit/s (the experiment is performed in a different AWS region, and AWS Lambda function-to-function throughput is the same). The achieved TCP throughput between VMs and functions matches that observed by others [24] between functions and AWS services such as S3.

To gain further insight about the bandwidth limits enforced, we conduct a load testing experiment by concurrently sending data from multiple functions to one. The server executing in one function listens for connections from clients executing in N other functions. Each client establishes one TCP connection to the server and attempts to saturate the connection by sending data in a tight loop. At 1 second interval, the server records aggregate bytes received from all of the clients. We vary the number of clients from 1 to 256 functions and run each configuration over a 5 minute interval. Figure 5 presents the averages of the aggregated received throughput at the server after removing the initial and final 30 seconds of the experiment measurements. The maximum observed through-

put is 621.69Mbits/s with 1 sending function and minimum of 607.04Mbit/s at 128 sending functions. We attribute the degradation of less than 3% to the overhead associated with handling multiple connections. This leads us to the conclusion that the ingress TCP bandwidth limits imposed on AWS Lambda functions do not depend on the number of sending function and that the available bandwidth is comparable to that of regular instances.

The only resource parameter that can be adjusted for AWS Lambdas is the amount of memory, which then proportionally determines the vcpu share allocated to the function. To determine if the memory setting also influences network properties we varied memory allocated to functions and measured the achievable throughput. The throughput did not vary with memory settings of 512MB, 3008MB, and 10240MB.

5.2 Latency analysis

We measure TCP latency of the six connection types described above. For each connection type, we instantiate 32 non-overlapping pairs of hosts (AWS Lambda functions or VMs). Each host executes a benchmarking program, and every pair opens a single TCP connection with Nagle’s algorithm disabled by both hosts. The host assigned the client role implements a TCP echo client that initiates a connection to the assigned server host. After accepting the connection from the client, the server function initiates 128 rounds of ping-pong exchanges of a 1024-byte message and measures the total time, this measurement is repeated 1024 times. Table 3 list the summary statistics of observed round-trip latencies for different connection types and Figure 7 shows the eCDF of the latencies measured.

Boxer’s TCP connections between pairs of functions have a mean round-trip latency of $694\mu s$. These connections also show significant variability in latencies, with a range between $202\mu s$ and $2769\mu s$, as can be seen in the eCDF plot. The main source of the observed variance is not due to the variance within each TCP connection but due to the variance between different TCP connections. This suggests that the network between different Lambda functions, from the latency perspective, is not uniform, and different instantiation patterns result in different latencies. This makes sense as the distance between the machines where functions are deployed plays a

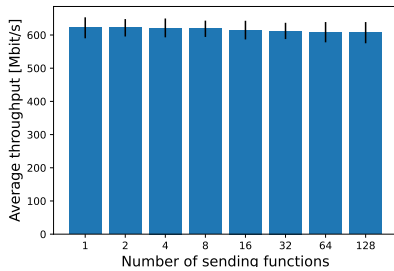


Figure 5: Comparison of aggregate receive throughput as number of sending functions varies. Maximum is 621.69Mbits/s at 1 sending function and minimum at 607.04Mbit/s at 128 sending functions, black lines represent standard deviation.

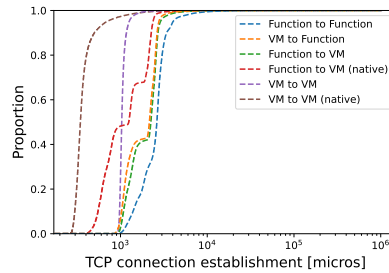


Figure 6: Empirical CDF of TCP connection establishment times. Time-to-first-byte (TTFB) for different connection types measured in microseconds between 32 distinct pairs of hosts establishing 1024 TCP connections each.

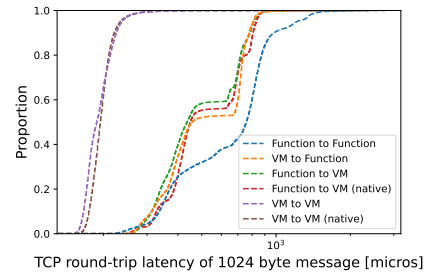


Figure 7: Empirical CDF of TCP round-trip latencies between 32 distinct nodes pairs (of VMs and AWS Lambda functions) echoing 1024 byte message, repeated for 6 connection scenarios.

role in determining the latency.

The latency measurements of the VM-to-VM-native connections, without the use of Boxer, and VM-to-VM connections that are established by Boxer show very close round-trip latencies of $198\mu s$ and $194\mu s$, and follow a very similar distribution which can be seen in the eCDF plots of the two connection types. This further shows that there is no data plane overhead of Boxer provided connections once they are established. The observed latency of connections between Lambda functions and VMs have similar mean round-trip latencies ranging from $520\mu s$ to $622\mu s$, and all follow a similar distribution as can be seen in the eCDF plot. Perhaps surprising, the connections provided by Boxer have a slightly lower round-trip latency, which could be attributed to a different processing of the traffic by the network due to its different initial packet signature (we may investigate this further in the future.)

We observe that the mean round-trip latency of Boxer’s TCP connections between pairs of functions is $3.51\times$ greater than latencies observed on AWS (m4.large) VM-to-VM native TCP connections. We consider this acceptable given the added functionality of Boxer.

5.3 Connection establishment

We measured the time required for application to have an established TCP connection for the six connection types described above. For each connection type we instantiate 32 host pairs (AWS VMs or Lambda functions depending on the scenario) and assign one host in each pair to be the client and the other to be the server. The time-to-first-byte(TTFB) measurement is recorded by the client. The client starts a timer, attempts to connect to its server and waits to read 1 byte of data from the server. Once it receives the 1 byte it stops the timer, records the duration and closes the connection. The

server accepts connections and replies with 1 byte right after it accepts a new connection. This is repeated 1024 times by each pair, the measurements are in Table 3.

Unlike TCP throughput and latency, Boxer does have overhead compared to native TCP connection times. Mean time to establish a function to function connection is $2735\mu s$. The overhead can be seen in comparison of establishment time for VM-to-VM-native connections of $408\mu s$ vs VM-to-VM connections that are established by Boxer of $1067\mu s$. This is due to additional round trip necessary to contact the destination Boxer to request that the connection setup and wait for acknowledgement before proceeding locally, or in case of an error response to forward to the error to the application (for example in case when there is nothing listening at the destination address.) This additional signaling adds to the connection setup times, and can also be observed in the connection times between functions and VMs. However, given that there are no alternative native connections for function-to-function and VM-to-function connection types, we consider the latency acceptable.

The eCDF of the TTFB times for the different scenarios in Figure 6 shows that the comparable scenarios with and without Boxer connection setup follow similar distribution, but are shifted by additional delay of the round trip times. Function-to-function connections, and connection types that cross the AWS Lambda and EC2 networks have maximum TTFB times observed to be over 1 second, including a connection scenario that does not involve Boxer to establish TCP connections (Function-to-VM-native.) This suggests that there may be packet loss or network congestion in the network and it is not a direct consequence of the procedure Boxer uses to establish the connections. As it can be seen from the eCDF such extreme connection times are rare, in the case of Function-to-Function connection type, 99.9 quantile for TTFB is $20654\mu s$. Compared to communicating through storage, such latencies

are acceptable, especially considering that in all of the above experiments performed, all connections were successfully established.

6 Short-lived data center

To demonstrate the ability of Boxer to provide a short-lived datacenter, we run an unmodified, complex distributed application: the DeathStarBench benchmark [12]. DeathStarBench is a suite of cloud microservice benchmarks deployed using container networks which mimic how real applications are deployed in production. We show that Boxer allows us to transparently deploy stateless microservices using AWS Lambda instead of AWS EC2 (VMs) or Fargate (Containers).

6.1 Using Boxer with the DeathStarBench

We focus on DeathStarBench’s *socialNetwork* application, which offers a social network service to users and is organized using three microservice layers: i) front-end layer (implemented using an NGINX webserver); ii) logic layer (implemented using stateless Thrift services that communicate through RPCs); iii) caching and storage layer (implemented with MongoDB and Memcached instances). In *socialNetwork*, user requests are received by the front-end layer (NGINX web server) and then routed to one of the services in the logic layer. Depending on the user request, the logic layer may perform one or multiple requests to the caching and storage layers. Since the logic layer is stateless (i.e., it contains no internal persistent state), it can be deployed on AWS Lambda. However, since functions cannot receive connections from external components, (e.g., the front-end NGINX layer in this case), Boxer is required to establish such connections.

We did not have to make any modifications to the application code to deploy DeathStarBench on AWS Lambda with Boxer. The benchmark was only modified to i) use names instead of fixed local IPs (for example, replace `127.0.0.1` by `nginx-thrift`), and ii) wrap the invocation of all components of the front-end and logic layers with the Boxer binary (for example, replace `SocialGraphService` with `Boxer -s <seed> -n <number of components> SocialGraphService`). Wrapping the invocation of the front-end and logic layer components with Boxer ensures that the creation of new connections goes through the Boxer network, allowing not only the front-end layer to establish connections to services running in the logic layer (VM to Lambda), but also components in the logic layer to establish connections between them (Lambda to Lambda).

6.2 Methodology

To evaluate the performance impact of moving the logic layer to Lambda using Boxer, we use three types of deployments.

First, an *EC2-only* deployment in which all components are deployed as VMs in EC2 is used as our baseline. Second, an EC2 VM-only deployment in which all components are still deployed as VMs in EC2 but the components of both the front-end and logic layers use Boxer. This deployment, *EC2-only (w/ Boxer)*, is used to measure the performance overhead of using Boxer. Third, a mixed deployment where the front-end, and caching and storage layers are deployed as VMs, and the logic layer is deployed using Lambdas, *EC2 + Lambda (w/ Boxer)*.

To measure the throughput and latency of the end-to-end system we use two workloads included in the DeathStarBench suite. A read workload that issues requests to read a user timeline in the social network, and a write workload that creates follow relationships between users. Both workloads are generated using the `wrk` [33] tool which builds and issues requests to the front-end layer. The performance of both workloads (read and write) is reported separately as each workload stresses the network and Boxer in a different way. The read workload mostly transfers data from the caching and storage layer (VMs), to the logic layer (VMs or Lambdas), and then to the front-end layer (VMs). The write workload operates in the opposite direction.

All experiments in this section were conducted in AWS Ohio (us-east-2) region. All VMs use a base Amazon Linux 2 [2]. For front-end, and caching and storage layers, we use `t3a.micro` instances due to the memory requirements of the services included in these layers. For logic layer, when deployed in VMs, we use `t3a.nano` instances. Lambdas are deployed using the Python3.7 runtime (which is used to launch Boxer and the service binaries). Each Lambda is given 2048 MB of memory size (we experimentally determined that in us-east-2, the performance of a 2048 MB Lambda is similar to a `t3a.nano` VM instance).

6.3 Stateless Services in Lambdas

We analyze the throughput and latency impact of using Boxer to deploy the DeathStarBench social Network application and move the application logic layer to AWS Lambda. Figure 8 shows the results for both read and write workloads across the three different types of deployments. For each workload, we collect the average throughput and 90th percentile latency with an increasing load in the system.

Results show that Boxer introduces only a low overhead. For the read workload, the EC2-only deployment becomes saturated at 3270 ops/s while the EC2-only (w/ Boxer) becomes saturated at 3070 ops/s. For the same data points, the 90p latency of a single request for the EC2-only and EC2-only (w/ Boxer) deployments are 3.18 ms and 5.07 ms, respectively. Note that these latencies are measured end-to-end and therefore include multiple internal microservice to microservice requests. The write workload demonstrates similar results. The EC2-only and EC2-only (w/ Boxer) deployments be-

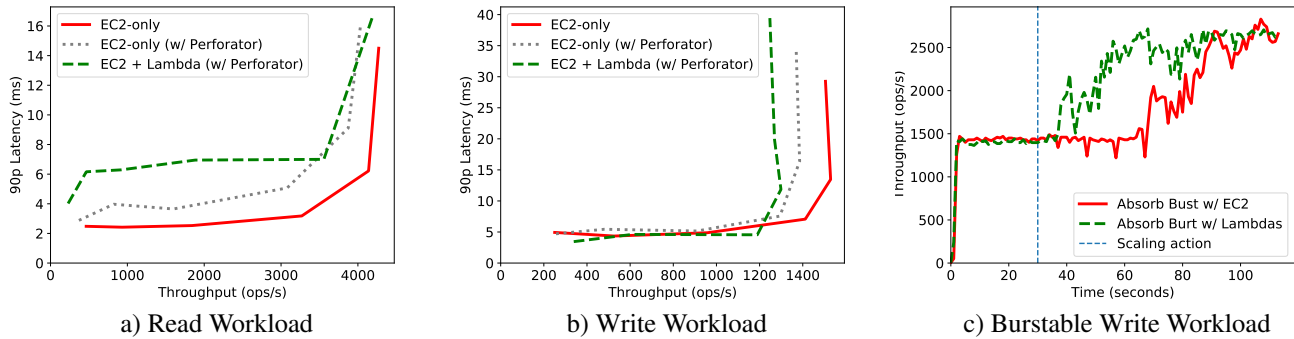


Figure 8: DeathStarBench experimental results for read and write workloads in a static deployment (left and center); burstable write workload in dynamic deployment (right).

come saturated at 1411 ops/s and 1294 ops/s, with latencies of 7.07 ms and 7.56 ms, respectively.

We use a similar analysis to measure the overhead of launching the logic layer services in AWS Lambda by comparing the EC2-only (w/ Boxer) and EC2 + Lambda (w/ Boxer) deployments. Figure 8 shows that for the read workload, the EC2 + Lambda (w/ Boxer) deployment saturates at a 3556 ops/s with a 90p latency of 7 ms. For the write workload, the same deployment saturates at 1189 ops/s and with a 90p latency of 4.55 ms.

We conclude that using Boxer incurs a small performance overhead arising from the interception and management of connections between the multiple services. Moving services to Lambda also incurs a small overhead due to the different way CPU and Network are allocated to VMs and lambdas. One could further increase the memory budget given to lambdas to also increase their computational power and thus close the performance gap between EC2-only (w/ Boxer) and EC2 + Lambda (w/ Boxer).

7 Dynamic Load Adaptation

We now show that Boxer enables us to increase the scaling factor of stateless microservices running on VMs by leveraging the elasticity of serverless platforms to adapt to bursty loads. While in the previous experiments we deployed the logic layer of the DeathStarBench socialNetwork application entirely on Lambda or EC2, in this section we start by deploying all logic layer services on VMs. When the load increases, additional logic layer services are allocated to handle the increased load either on VMs or Lambdas. We compare two deployments: i) VM-based deployment that allocates extra VMs to handle increased load; ii) VM-based deployment that allocates lambdas to accommodate increased load. Our goal is to study the ability to adapt to a load burst.

Figure 8 (right-hand side) presents a throughput trace of both deployments. Throughput is measured using wrk [33] by looking at how many requests the front-end layer can handle

per second. After 30 seconds (dashed vertical line), a new set of VMs or lambdas are deployed to handle the increased load. After requesting a new set of VMs to join the logic layer ($t=30s$), the system throughput starts increasing after 40 seconds ($t=70s$) and only stabilizes (completely absorbs the burst) after 75 seconds ($t=105s$). There are several reasons to this. First, VM launch time is approximately 30 seconds (i.e., until the VM is connected to the network). After the VM launches, we also register its IP in our nameserver (we use DNS round-robin to distribute the load). At this point, when NGINX (front-end layer) needs to open a new connection to serve a new client, it will try to resolve the logic layer component name through the nameserver. Because load distribution is performed per-connection and connections are kept open to serve multiple requests, adapting to a burst request takes dozens of seconds. The time interval between nodes becoming available and being fully utilized by NGINX could be reduced by forcing NGINX to recycle connections more frequently. This would, however, add overhead for creating more connections.

Using Lambdas to absorb load bursts can significantly reduce the time to add new logic layer services, and therefore, reduce the time to reach a steady throughput using all new nodes. Note that, although lambdas take less than one second to launch (compared to approximately 30 seconds for VMs), throughput starts increasing 8 seconds after the new lambdas are launched ($t=38s$), and only stabilizes after 39 seconds ($t=69s$). In summary, using Lambda to accommodate bursts reduces the time for new components start receiving requests by $5\times$, and reduces the time to fully accommodate a burst by $1.9\times$.

8 Discussion

We have shown that supporting direct networking among functions enables running off-the-shelf distributed applications, which were designed for traditional Linux execution environments, on serverless platforms. By decoupling the serverless

resource management model from its event-driven programming model, Boxer enables cloud users to benefit from the resource elasticity of serverless platforms without having to change the programming model of their applications. The ability to run applications seamlessly between VM and serverless platforms allows users to leverage serverless platforms during bursty load periods and benefit from the lower cost per unit time of VM resources during steady load periods.

8.1 Opportunities

Enabling direct networking among functions opens up many opportunities to address other key limitations of serverless platforms. For example, to overcome limitations on function execution time, a function that is about to reach the maximum supported execution time on the platform could spawn a follower function and send its state to the follower directly to transparently continue execution. For instance, Infinicache [31] refreshes functions that cache data by having a function call itself and send the cached data to its newly spawned clone. With Boxer, state can be sent through a direct communication channel, without relying on a proxy. Workflows can be programmed into functions so that a sequence of steps can execute automatically with functions spawning the next step and passing the necessary data without having to rely on an external orchestrator [15, 37]. In addition, scatter-gather interactions become possible and more efficient. This would considerably simplify running applications such as those considered in ExCamera [11] and gg [10], or removing a variety of work-arounds built to bypass the lack of function-to-function communication to implement stateful functions [1, 29]. It also enables implementing truly distributed data processing operators such as joins instead of using today’s contrived solutions which need to communicate through storage [24, 25]. Similar ideas apply to ML over serverless, which today is expensive due to the lack of communication [19]. Boxer can also be used to implement a form of work stealing among functions of a serverless applications, since functions could communicate directly with one another to request additional work if they are idle. This mechanism adds another dimension of elasticity to serverless computing, which is particularly useful when the amount of work to be done is not easily determined upfront or may exhibit skew.

Another key limitation of serverless platforms is that they do not allow users to optimize the placement of functions to improve performance [27]. By supporting direct networking between functions, cloud providers can get a complete picture of the communication patterns of an application and use this information to optimize function placement [21]. Understanding the communication patterns of serverless applications is difficult for providers to do today as functions are forced to communicate through remote storage systems or proxies. While we have shown how users can circumvent the networking limitations in today’s serverless platforms with Boxer, we

hope that cloud providers will natively provide networking abstractions on serverless platforms. For example, providers could extend the serverless programming model to include communication collectives, which would allow them to optimize function placement policies for different communication patterns.

8.2 Current limitations

Boxer raises a number of issues that we intend to address in the future:

Boxer control network topology: Currently Boxer establishes a fully connected TCP control topology between all participating Boxer nodes. As the scale of deployment increase, this will likely be one of the bottlenecks. We are investigating solutions such as indirection through other Boxer nodes to route control messages, or using a datagram protocol to conserve the file descriptors.

Non-blocking I/O: Boxer’s transparent interposition layer does not support all non-blocking socket operations. Currently Boxer will block a calling thread for the duration of connect call, which can degrade performance of applications and can break the semantics of some.

Lambda execution environment: Applications running in Lambda do not have access to the same exact environment that would be available in an EC2 VM. For example, the local file system might be read-only, there might be only one network device with a local IP, etc. Boxer can circumvent these limitation and emulate an environment similar to the one available in VMs by further interception of the libc functions used to query local interfaces, local files, etc.

Multiple lambdas behind a single NAT address: Boxer currently does not support multiple lambdas deployed behind the same NAT address. When this happens, a single external IP address is visible and Boxer cannot distinguish between the two nodes. To support this scenario transparently to the user application, Boxer needs to virtualize the network address space.

8.3 Truly general purpose short-lived datacenters

Besides networking, other limitations that serverless platforms impose today include limited resources per function invocation (e.g., up to 10 GB of memory and 6 vCPUs in AWS Lambda), limited execution time for each function (e.g., 15 minutes in AWS Lambda), and the lack of support for heterogeneous hardware such as GPUs. Cloud providers have been steadily increasing resource limits and function execution time limits. We expect this trend to continue as serverless computing becomes increasingly popular for a broad range of applications. In particular, supporting higher resource limits per function gives developers the option to scale *up* in

addition to scale *out* their application to achieve higher performance. Both of these scaling dimensions are important for general datacenter computing. Adding support for running serverless functions on heterogeneous hardware resources, such as accelerators, will also become increasingly important as applications such as machine learning jobs, which rely on GPUs and ASICs, can also benefit from the elasticity, fine-grain billing, and higher level of abstraction to the cloud that serverless computing offers.

9 Conclusion

We presented Boxer, a system that transparently provides direct communication between serverless functions. Boxer uses a NAT traversal mechanism to enable serverless functions to accept incoming traffic from external sources. By leveraging dynamic linking to intercept function calls, Boxer provides direct networking functionality for serverless functions through the standard socket networking API and requires no modifications to Linux TCP-based datacenter applications. With Boxer, latency sensitive applications such as microservices can be run unmodified on AWS Lambda. Our system presents a major step towards treating serverless platforms as short-lived, instant, general purpose datacenters and allows a broader range of applications to benefit from the elasticity, fine-grain billing, and high level of abstraction to the cloud that serverless platforms offer.

References

- [1] Adil Akhter, Marios Fragkoulis, and Asterios Katsifodimos. “Stateful Functions as a Service in Action”. In: *Proc. VLDB Endow.* 12.12 (2019), pp. 1890–1893.
- [2] *Amazon Linux 2*. URL: <https://aws.amazon.com/amazon-linux-2/> (visited on 12/10/2020).
- [3] *Amazon S3*. URL: <https://aws.amazon.com/s3> (visited on 08/17/2020).
- [4] *AWS Lambda*. URL: <https://aws.amazon.com/lambda> (visited on 08/17/2020).
- [5] *AWS Lambda changes duration billing granularity from 100ms down to 1ms*. URL: <https://aws.amazon.com/about-aws/whats-new/2020/12/aws-lambda-changes-duration-billing-granularity-from-100ms-to-1ms/>.
- [6] *AWS Lambda FAQs*. URL: <https://aws.amazon.com/lambda/faqs/>.
- [7] *AWS Lambda now supports container images as a packaging format*. URL: <https://aws.amazon.com/about-aws/whats-new/2020/12/aws-lambda-now-supports-container-images-as-a-packaging-format/>.
- [8] Paul Castro, Vatche Ishakian, Vinod Muthusamy, and Aleksander Slominski. “The Rise of Serverless Computing”. In: *Commun. ACM* 62.12 (2019), pp. 44–54.
- [9] Bryan Ford, Pyda Srisuresh, and Dan Kegel. “Peer-to-Peer Communication Across Network Address Translators”. In: *USENIX ATC*. 2005.
- [10] Sadjad Fouladi, Francisco Romero, Dan Iter, Qian Li, Shuvo Chatterjee, Christos Kozyrakis, Matei Zaharia, and Keith Winstein. “From Laptop to Lambda: Outsourcing Everyday Jobs to Thousands of Transient Functional Containers”. In: *USENIX ATC*. 2019.
- [11] Sadjad Fouladi et al. “Encoding, Fast and Slow: Low-Latency Video Processing Using Thousands of Tiny Threads”. In: *NSDI*. 2017.
- [12] Yu Gan et al. “An Open-Source Benchmark Suite for Microservices and Their Hardware-Software Implications for Cloud & Edge Systems”. In: *ASPLOS*. 2019.
- [13] *Google Cloud Functions*. URL: <https://cloud.google.com/functions>.
- [14] Joseph M. Hellerstein, Jose M. Faleiro, Joseph Gonzalez, Johann Schleier-Smith, Vikram Sreekanti, Alexey Tumanov, and Chenggang Wu. “Serverless Computing: One Step Forward, Two Steps Back”. In: *CIDR*. 2019.
- [15] Martijn de Heus, Kyriakos Psarakis, Marios Fragkoulis, and Asterios Katsifodimos. “Distributed Transactions on Serverless Stateful Functions”. In: *Proceedings of the 15th ACM International Conference on Distributed and Event-Based Systems*. 2021, pp. 31–42.
- [16] *iperf3*. URL: <https://software.es.net/iperf/> (visited on 12/10/2020).
- [17] Eppinger J.L. “TCP Connections for P2P Apps: A Software Approach to Solving the NAT Problem”. In: *Carnegie Mellon University, Tech. Rep, ISRI-05-104*. 2005.
- [18] Zhipeng Jia and Emmett Witchel. “Nightcore: Efficient and Scalable Serverless Computing for Latency-Sensitive, Interactive Microservices”. In: *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*. 2021, pp. 152–166.
- [19] Jiawei Jiang et al. “Towards Demystifying Serverless Machine Learning Training”. In: *Proceedings of the 2021 International Conference on Management of Data*. 2021, pp. 857–871.
- [20] Ana Klimovic, Yawen Wang, Patrick Stuedi, Animesh Trivedi, Jonas Pfefferle, and Christos Kozyrakis. “Pocket: Elastic Ephemeral Storage for Serverless Analytics”. In: *OSDI*. 2018, pp. 427–444.

- [21] Ashraf Mahgoub, Karthick Shankar, Subrata Mitra, Ana Klimovic, Somali Chaterji, and Saurabh Bagchi. “SONIC: Application-aware Data Passing for Chained Serverless Applications”. In: *USENIX Annual Technical Conference (USENIX ATC 21)*. 2021.
- [22] *Managing AWS Lambda Function Concurrency*. URL: <https://aws.amazon.com/blogs/compute/managing-aws-lambda-function-concurrency/> (visited on 08/28/2020).
- [23] *Microsoft Azure Functions*. URL: <https://azure.microsoft.com/en-us/services/functions> (visited on 08/17/2020).
- [24] Ingo Müller, Renato Marroquín, and Gustavo Alonso. “Lambada: Interactive Data Analytics on Cold Data Using Serverless Cloud Infrastructure”. In: *SIGMOD*. 2020.
- [25] Matthew Perron, Raul Castro Fernandez, David DeWitt, and Samuel Madden. “Starling: A Scalable Query Engine on Cloud Functions”. In: *SIGMOD*. 2020.
- [26] Qifan Pu, Shivaram Venkataraman, and Ion Stoica. “Shuffling, Fast and Slow: Scalable Analytics on Serverless Infrastructure”. In: *NSDI 19*. 2019.
- [27] Johann Schleier-Smith et al. “What Serverless Computing is and Should Become: The next Phase of Cloud Computing”. In: *Commun. ACM* 64.5 (Apr. 2021), pp. 76–84.
- [28] *Serverless Networking SDK*. URL: <http://networkingclients.serverlesstech.net/> (visited on 08/17/2020).
- [29] Vikram Sreekanti, Chenggang Wu, Xiayue Charles Lin, Johann Schleier-Smith, Joseph E. Gonzalez, Joseph M. Hellerstein, and Alexey Tumanov. “Cloudburst: Stateful Functions-as-a-Service”. In: *Proc. VLDB Endow.* 13.12 (2020), pp. 2438–2452.
- [30] Shelby Thomas, Lixiang Ao, Geoffrey M. Voelker, and George Porter. “Particle: Ephemeral Endpoints for Serverless Networking”. In: *Proceedings of the 11th ACM Symposium on Cloud Computing*. 2020, pp. 16–29.
- [31] Ao Wang et al. “InfiniCache: Exploiting Ephemeral Serverless Functions to Build a Cost-Effective Memory Cache”. In: *USENIX FAST*. 2020.
- [32] Michal Wawrzoniak, Ingo Müller, Rodrigo Bruno, and Gustavo Alonso. “Boxer: Data Analytics on Network-enabled Serverless Platforms”. In: *CIDR*. 2021.
- [33] *wrk - a HTTP benchmarking tool*. URL: <https://github.com/wg/wrk> (visited on 10/09/2021).
- [34] Chenggang Wu, Jose M. Faleiro, Yihan Lin, and Joseph M. Hellerstein. “Anna: A KVS for Any Scale”. In: *ICDE*. 2018.
- [35] Chenggang Wu, Vikram Sreekanti, and Joseph M. Hellerstein. “Autoscaling Tiered Cloud Storage in Anna”. In: *PVLDB* (2019).
- [36] Yunhong Gu, Xinwei Hong, and R. L. Grossman. “Experiences in Design and Implementation of a High Performance Transport Protocol”. In: *SC '04*. 2004.
- [37] Haoran Zhang, Adney Cardoza, Peter Baile Chen, Sebastian Angel, and Vincent Liu. “Fault-tolerant and transactional stateful serverless workflows”. In: *14th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2020, Virtual Event, November 4-6, 2020*. 2020, pp. 1187–1204.