Raftian: Pygame Meets Raft

MARCO BELLÒ and TULLIO VARDANEGA, University of Padua, Italy

The problem of shared consensus in a distributed system is both older than a millennium and more relevant than ever: while in the Aegean island of Paxos the challenge was to keep track of the many laws passed in a parliament where legislators had better to do than attend sessions full-time, nowadays the ubiquity of web-based architectures and applications (from a simple cloud storage to the whole banking system) gives daily headaches to developers and system administrators alike. To reduce ibuprofen consumption in the IT sector, Ongaro and Ousterhout devised an easily understandable and implementable alternative to the Paxos algorithm, namely the Raft consensus algorithm, which we employ in this work to implement communication via XML-RPC between multiple Pygame applications, each running a game instance that aims to simulate a simplified version of a real-time strategy game.

CCS Concepts: • **Computer systems organization → Fault-tolerant network topologies**; • **Computing methodologies → Concurrent algorithms**; **Distributed algorithms**; **Concurrent programming languages**.

Additional Key Words and Phrases: Python, Raft, Shared Consensus, Gaming, Multiplayer, Multithreading, RPC, Pygame

ACM Reference Format:
Marco Bellò and Tullio Vardanega. 2025. Raftian: Pygame Meets Raft. 1, 1 (September 2025), 25 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

From sharing spreadsheets between a handful of laptops in a small basement office, through large-scale rendering on a supercomputer, to the entire global finance system, distributed computing has become an essential component of the modern world: nowadays, what most people need a computer for can be done in the browser thanks to services like email clients, cloud calendars, media streaming platforms and web-based office suites (like Google Docs) that expose word editors, spreadsheets managers, presentations programs and more, all while being constantly synchronized to the cloud, which not only ensures data persistence and availability, but also enables sharing and collaboration between users.

It does not end here: other examples of distributed applications include cloud storage services like Dropbox, Google Drive or OneDrive, streaming services like Netflix, YouTube or Spotify, distributed computing like blockchain technologies or AWS, online banking services (the banking system itself is distributed since way before), social networks, and even maritime and aircraft traffic control systems. Moreover, the rise of the gaming industry played a significant role in pushing distribution forward: in 2024 the gaming market revenue was estimated to be 187.7 billion U.S. dollars [3], making it a hefty slice of the pie that is the entertainment industry [14], with 111 billions generated by free-to-play games [4] (70 billions from social and casual games alone [9]), which interests us since their business model often relies on cosmetics, game passes and advertisements, forcing them to be constantly on-line.

Authors' address: Marco Bellò, marco.bello.3@studenti.unipd.it; Tullio Vardanega, tullio.vardanega@unipd.it, University of Padua, Padua, Italy.

But enough with examples, let's now define what distribution *is*: a distributed system is a computer system whose inter-communicating components are located on different networked computers [1, 18], which coordinate their actions via message-passing to achieve a common goal. There are three significant challenges to overcome: maintaining components' concurrency [1], eliminating global-lock reliance and managing the independent failure of components, all while ensuring scalability (often the purpose is scaling itself) and transparency to the user, meaning interactions with any exposed interface should be done while being unaware of the complexity behind them.

Moreover, shared consensus must be guaranteed: it does not require much thought to see that all servers in a cluster should agree on one or more shared values, lest becoming a collection of unrelated components that have little to do with collaboration (thus distribution). In the most traditional single-value consensus protocols, such as Paxos [11], cooperating nodes agree on a single value (e.g., an integer), while multi-value alternatives like Raft [12] aim to agree on a series of values (i.e., a log) growing over time forming a sort-of cluster's history. It is worth noting that both goals are hindered by the intrinsically asynchronous nature of real-world communication, which makes it impossible to achieve consensus via deterministic algorithms, as stated by Fischer, Lynch and Paterson in their FLP impossibility theorem [6]. Thankfully this can be circumvented by injecting some degree of randomness.

1.1 Project's Contributions

The concepts and examples we mentioned so far allow us to finally present the goal of this project: we will create a simplified clone of Travian [2], an old real-time [3] player-versus-player [4] strategy game [5], where players build their own city and wage war on one another (less wrinkly readers may be more familiar with the modern counterpart Clash of Clans [6]), built with Pygame [7], a Python library that creates and manages all necessary components to run a game such as game-engine, graphical user interface, sounds, player inputs and the like, where each player resides in a separate server (or node) that communicate with the others via an algorithm modelled after Raft's specifications.

This choice follows the authors' interest in exploring Raft capabilities and ease of implementation in a fun and novel way, using a language that while extremely popular is seldom used in such a fashion and, as far as we can tell, never with this mixture of libraries.

Both game and algorithm implementations have been reduced to a reasonably complex proof of concept to keep the project scope manageable: it is possible to instantiate games up to five players, each of which is restricted to the only action of attacking the others, while Raft's functionalities are limited to log replication and overwriting.

Experiments were conducted to evaluate both game responsiveness and the communication algorithm correctness.

All source code is visible at the following link: https://github.com/mhetacc/RuntimesConcurrencyDistribution/blob/main/raftian/raftian.py.

[1]Concurrency refers to the ability of a system to execute multiple tasks through simultaneous execution or time-sharing (context switching), sharing resources and managing interactions. It improves responsiveness, throughput, and scalability [7, 13, 15, 17, 21].

[2]Travian: Legends is a persistent, browser-based, massively multiplayer, online real-time strategy game developed by the German software company Travian Games. It was originally written and released in June 2004 as "Travian" by Gerhard Müller. Set in classical antiquity, Travian: Legends is a predominantly militaristic real-time strategy game. Source: https://www.travian.com/international

[3]Real-time games progress in a continuous time frame, allowing all players (human or computer-controlled) to play at the same time. By contrast, in turn-based games players wait for their turn to play.

[4]Player-versus-player (PvP) is a type of game where real human players compete against each other, opposed to player-versus-environment (PvE) games, where players face computer-controlled opponents.

[5]Strategy video game is a major video game genre that focuses on analyzing and strategizing over direct quick reaction in order to secure success. Although many types of video games can contain strategic elements, the strategy genre is most commonly defined by a primary focus on high-level strategy, logistics and resource management. [16]

[6]Clash of Clans: https://supercell.com/en/games/clashofclans/

[7]Pygame: https://www.pygame.org/docs/

2 PYTHON

We will now discuss the technologies employed in the development of the project (section 2) as well as presenting our implementation of the Raft's algorithm.

Python is a high-level, dynamically typed and interpreted programming language that is often used for scripting, data analysis and small application development, making it a non-obvious choice for this project, which does not fall into any of these categories.

As a language, it has two main advantages compared to others: first of all it is undoubtedly the most popular and widely used in the world (figure 1) [5, 8], meaning abundant documentation and resources. Secondly it has a huge ecosystem of libraries that implement all the functionalities we need for this project, namely: *XML-RPC* for the remote procedure calls (RPCs), *threading* to handle local concurrency and *Pygame* to manage everything game-related.

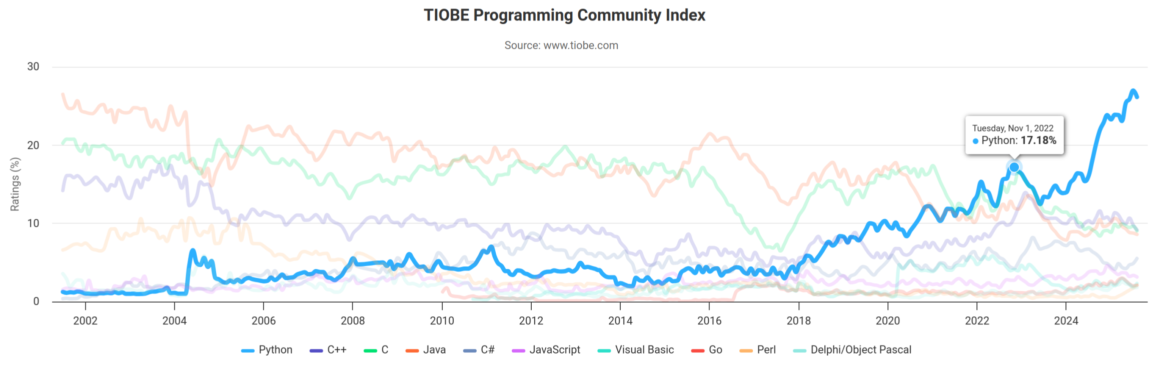


Fig. 1. TIOBE Programming Community Index, focus on Python statistics, 2025. (https://www.tiobe.com/tiobe-index/)

2.1 Remote Procedure Calls

In Raft's specifications it is stated that nodes communicate with each other via remote procedure calls [12], which in distributed computing is when a program causes a procedure (or subroutine) to execute in another address space (commonly on another computer on a shared network) calling it as if it were local (that is, the programmer writes the same code whether the subroutine is local or remote).

There are many libraries that implement this functionality, like gRPC (https://grpc.io/), which is a high performance open source RPC framework used by many big players, such as Netflix [8] and Cockroach Labs [9], available for many languages (Python included), but we opted for the standard library *XML-RPC* [10] thanks to its promised simplicity and ease of use.

The library provides both server and client implementations, encapsulating the former in its own loop, while the latter can be fired as needed allowing a bit more flexibility in its usage.

In code 1, *client* is an instance of *ServerProxy*, which acts as the client-side interface for XML-RPC, allowing to call the remote procedure *test_foo* as if it were a local function, even though it executes on a server in a different networked location.

[8]Netflix Ribbon is an Inter Process Communication library built in software load balancers: https://github.com/Netflix/ribbon
[9]Cockroach Labs is the company behind CockroachDB, a highly resilient distributed database: https://www.cockroachlabs.com/
[10]XML-RPC is a Remote Procedure Call method that uses XML passed via HTTP as a transport: https://docs.python.org/3/library/xmlrpc.html

Listing 1. Client as server proxy

```
with xmlrpc.client.ServerProxy('http://localhost:8000', allow_none=True) as client:
    print(client.test_foo(42)) # print returned value
```

The server must be instantiated and kept running by calling its event loop (e.g., using *serve_forever*), and all remote procedure calls must be registered using the *register_function* method of *SimpleXMLRPCServer* (code 2).

Listing 2. Server

```
with SimpleXMLRPCServer (('localhost', 8000)) as server:
    def test_foo(number):
        return f'The number is {number}'

    server.register_function(test_foo)
    server.serve_forever() # keep server alive
```

For this project, we extended *SimpleXMLRPCServer* to create a class that implements the Raft protocol (more details in section 3)

2.2 Concurrency

In this project the need for concurrent programming arises from two challenges: every server has an internal timer that fires at certain intervals, and every node has to run a game engine and the server itself at the same time, both of which can be simplified as two *"while true"* loops.

Most Raft implementations achieve concurrency with asynchronous programming, using libraries such as *asyncio* [11], which while powerful and efficient, makes writing code a bit awkward and cumbersome. We thus opted for a more traditional approach using multithreaded programming: in computer science, a thread of execution is the smallest sequence of programmed instruction that can be managed independently by the scheduler [10], and multiple threads may be executed concurrently sharing resources such as memory, which is directly counterpointed to multiprocessing where each process has its own storage space (moreover, processes are typically made of threads).

In Python there are modules in the standard library for both of them, respectively *threading* [12] and *multiprocessing* [13]. It is fundamental to note that the former does not provide real multi-threading since, due to the Global Interpreter Lock of CPython (the, for want of a better word, official Python implementation), only one thread can execute bytecode at once. To cite directly from the documentation: *"[GIL is] The mechanism used by the CPython interpreter to assure that only one thread executes Python bytecode at a time. This simplifies the CPython implementation by making the object model (including critical built-in types such as dict) implicitly safe against concurrent access. Locking the entire interpreter makes it easier for the interpreter to be multi-threaded, at the expense of much of the parallelism afforded by multi-processor machines."* [14].

[11]Asyncio is a library to write concurrent code using the async/await syntax: https://docs.python.org/3/library/asyncio.html
[12]The threading module provides a way to run multiple threads (smaller units of a process) concurrently within a single process: https://docs.python.org/3/library/threading.html
[13]The multiprocessing module is a package that supports spawning processes using an API similar to the threading module: https://docs.python.org/3/library/multiprocessing.html
[14]Global Interpreter Lock: https://docs.python.org/3/glossary.html#term-global-interpreter-lock

Thankfully, this does not apply with the *multiprocessing* module, which creates separate processes instead, offering both local and remote concurrency effectively side-stepping the Global Interpreter Lock, allowing programmers to fully leverage multiple cores. As previously stated, processes are much heavier than threads and thus more expensive to create, but do not incur the risks of shared memory.

2.2.1 Comparison. To evaluate which of the two modules is more suited for our purposes, we devised a simple experiment: we created two game instances with one hundred and one thousands coloured dots respectively (figure 2), that move around by offsetting their position each frame of a random amount between minus five and plus five pixels (pseudocode 3).

Then we ran both of them in three scenarios: with the game instance alone (baselines), with a server alive in a thread and with a server alive in a process, and we measured the *frames per second* (FPS) [15] in each case, since it is the most common metric to evaluate game performance. Higher FPS-count translates to a smoother and more responsive, i.e., better, gaming experience.

Listing 3. Pygame graphical dot offset

```
# create random offsets for both x and y coordinates
xmov = random.randint(-5,5)
ymov = random.randint(-5,5)

# move the dot by a certain offset
dot.move_by(xmov, ymov)
```

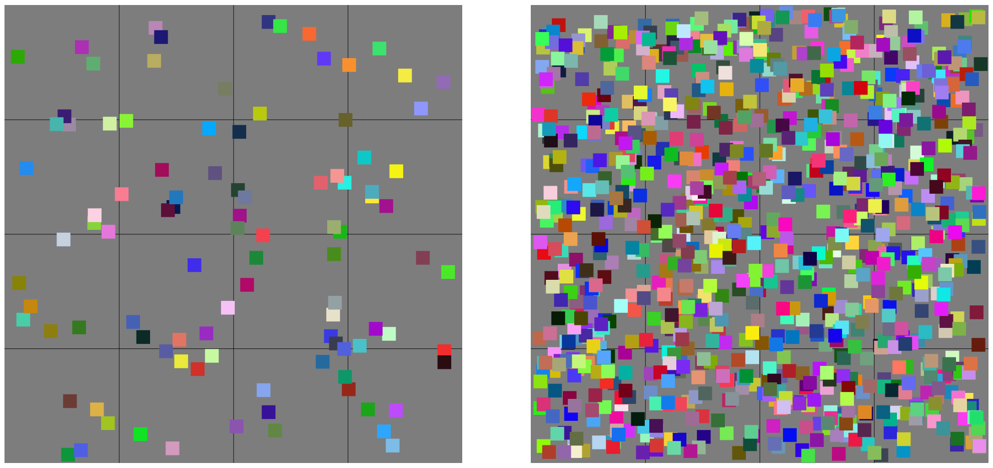


Fig. 2. Two game instances made with Pygame, with respectively 100 and 1000 dots that randomly move around

Results, shown in the graph at figure 3, tell us us that:

- Increasing the number of dots from 100 to 1000 halves the FPS count;

[15]Frame rate, most commonly expressed in frames per second or FPS, is typically the frequency (rate) at which consecutive images (frames) are captured or displayed. This definition applies to film and video cameras, computer animation, and motion capture systems, while in the context of computer graphics is the rate at which a system, particularly the graphic card, is able to generate frames. Source: https://en.wikipedia.org/wiki/Frame_rate

- Adding a server in a separate thread halves performance;
- Using *multiprocessing* yields worse performance than *threading* in the 100-dots scenario (about -30%), while performing similarly in the 1000-dots one.

This leads us to conclude that, for our specific purposes, the *threading* module is the best choice, especially since the final game will be way less computationally expensive from a graphical standpoint, hence using a lighter weight alternative should be even more beneficial than tested.

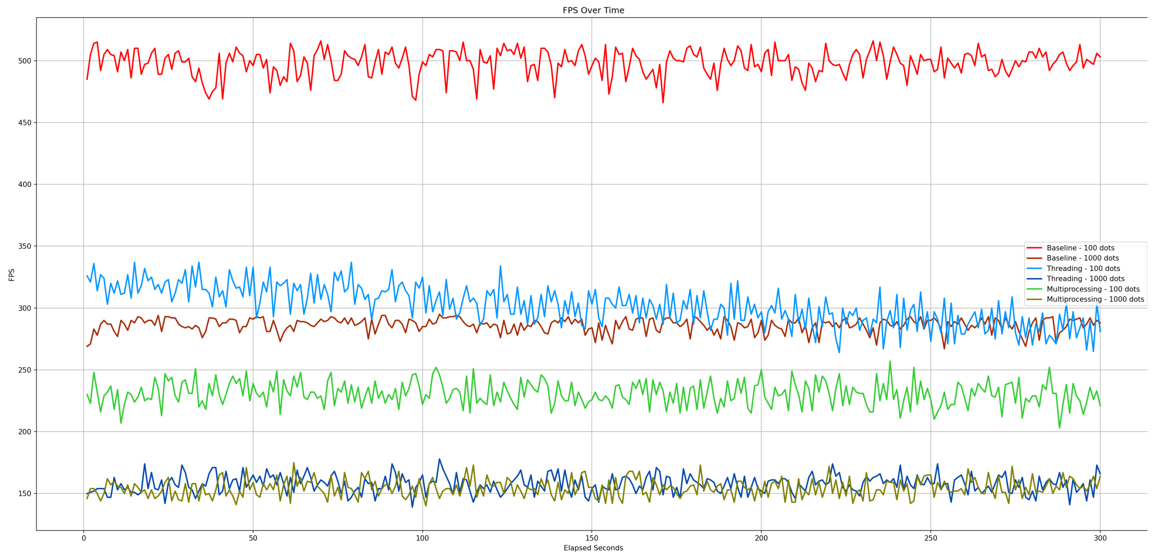


Fig. 3. Performance evaluation graph: red hues for baselines, blue hues for threading and green hues for multiprocessing. Darker shades for 1000 dots and lighter shades for 100 dots game instances

All tests have been performed with the following machine:

- OS: Ubuntu 24.04.1 LTS x86_64;
- Kernel: 6.8.0-52-generic;
- Shell: bash 5.2.21;
- CPU: 13th Gen Intel i7-13620H;
- GPU: NVIDIA GeForce RTX 4050 Laptop GPU;
- Memory: 15610MiB;
- Python version: 3.12.3;
- Power Mode: Balanced;
- Power Supply: 100W via type C.

2.3 Game Engines

There are many ways to implement a graphical user interface: from clever shell tricks like htop [16], to full-fledged game engines like Unity [17] or Godot [18] that often come with their own editor and a *top-down* approach, meaning build the UI first and then go down to code as needed for scripting and refining.

Unfortunately, our needs are quite opposite: what we want is a code-only, mono-language framework that while slowing down game development should simplify merging Raft with it. The choice thus boiled down to two alternatives: tkinter and Pygame.

2.3.1 Comparison. Let's list strengths and weaknesses of the two.

- **tkinter:**
  - 👍 Module of the standard library;
  - 👍 Few lines of code to make simple UIs;
  - 👎 Low flexibility;
  - 👎 No game loop;
  - 👎 Not a game engine;
- **Pygame:**
  - 👍 Extreme flexibility;
  - 👍 Direct access to game loop;
  - 👍 APIs to access many kinds of user inputs;
  - 👎 Verbose to obtain simple UIs;
  - 👎 Non-standard community-made framework.

We ultimately decided to opt for Pygame for four reasons: it is extremely flexible, exposes many useful functions (for example to catch different user inputs), gives direct access to the game loop and it is a novel and fun framework that has never, to the best of our knowledge, been used in such a fashion.

3 RAFT

It is not our intention to plagiarize Ongaro and Ousterhout's excellent work *"In Search of an Understandable Consensus Algorithm"* [12] by presenting the Raft algorithm's specifications. Instead, we will discuss how we molded it to our own use case.

The algorithm uses three types of nodes, namely *leader*, *follower* and *candidate*, and revolves around three core functionalities: leader election, log replication and cluster membership change. Log compaction is also mentioned, while a byzantine fault tolerant variant is never explored by the original authors.

One last component, instrumental to the functioning of the algorithm, is the *term*: everything happens in a certain term, which divides time logically and increments every election. This is necessary to recognize out-of-date leaders: if some follower has a term greater than the leader's, said leader is outdated.

Our Raft class directly extends *simpleXMLRPCServer* from XML-RPC module, as shown at code 4.

[16]Htop is a cross-platform text-mode interactive process viewer: https://htop.dev/
[17]Unity is a cross-platform game engine developed by Unity Technologies: https://unity.com/
[18]Godot is a cross-platform open-source game engine: https://godotengine.org/

Lastly, to fire off non-blocking concurrent RPCs on the cluster, we leverage the *concurrent.futures* module using *ThreadPoolExecutor*. To avoid creating and destroying pools every time a server needs to communicate with the cluster, we embedded a finite amount of workers as class attributes (code 5).

Listing 4. Class Raft definition

```
class Raft(SimpleXMLRPCServer):
    def __init__(self,
                 addr: tuple[str, int],
                 allow_none: bool = True,
                 # ...
                 last_index_on_server: list[tuple[int, int]] | None = None
                 ):
        SimpleXMLRPCServer.__init__(self, addr=addr, allow_none=allow_none)
```

Listing 5. ThreadPoolExecutor created with as many workers as there are servers in the cluster

```
class Raft(SimpleXMLRPCServer):
    def __init__(self,
                 # ...
                 )
        # start executors pool
        self.executor = concurrent.futures.ThreadPoolExecutor(max_workers=len(self.cluster))
```

3.1 Node Types

As previously stated, there are three node types: leader, follower and candidate (code 6). In this section we are going to show their characteristics and similarities. Note that all nodes have a timer: it is randomized for each of them and has been implemented by extending *threading.Timer*, thus making it thread-safe (code 7)

Listing 6. Node modes

```
class Raft(SimpleXMLRPCServer):
    class Mode(Enum):
        LEADER = 1
        CANDIDATE = 2
        FOLLOWER = 3

    def __init__(self,
                 # ...
                 mode: Mode = Mode.FOLLOWER,
                 )
```

Listing 7. Threadsafe looping timer

```
class LoopTimer(Timer):
    def __init__(self, interval, function, args=None, kwawrgs=None):
        Timer.__init__(self, interval, function, args, kwawrgs)
        self.was_reset : bool = False
    # ...

class Raft(SimpleXMLRPCServer):
    def __init__(self,
                 # ...
                 )
        # start timer
        self.timer = LoopTimer(timeout, self.on_timeout)
        self.timer.start()
```

3.1.1 Leader Node. The algorithm revolves around one leader node, whose job is to synchronize all servers' logs to ensure data consistency. It does so by replicating its own log on all followers (the non-leader nodes) by sending new or, if needed, old entries via remote procedure calls.

To make sure all nodes believe the leader's alive, periodically sends an empty remote procedure call called *heartbeat* (every 150-300ms).

3.1.2 Follower Node. All nodes, except for the leader, are classified as followers. They are not allowed to replicate their own log, and they have to forward any request to the leader.

To make sure the cluster never remains without a leader, every follower has an election timeout (between 150ms and 300ms) which resets every time an RPC from the leader is received. If it times out, the follower changes its state to *candidate*, increments its current term and starts a leader election.

Followers become candidates in another scenario: whenever they receive an entry from the leader, they compare it with their own last log entry. If the leader's term is smaller, it is out of date and a new election is started.

3.1.3 Candidate Node. When a follower's election timeout times out, it becomes a candidate, increments its own term and starts an election. Votes for itself and then waits for either of two outcomes: wins, thus becoming a new leader, or loses (either another leader gets elected or the old one manifests itself) thus reverting back to being a follower.

3.2 Log

As stated, the leader's job is to accept requests (in our specific case they are player inputs) and then forward them to the followers. Let's talk about the structure of the log.

The log is basically a *list* (or an *array*) of entries, where *entry* is an element that encapsulates data (like an integer or a string), has an index (unique for each entry) and the term of its creation (figure 4). We defined entries as *Data Classes* [19] (decorators that simulate C's structures) as seen in code 8.

[19]Data Classes module provides a decorator and functions for automatically adding generated special methods to user-defined classes: https://docs.python.org/3/library/dataclasses.html

Listing 8. Dataclass Entry definition

```
@dataclass
class Entry:
    term: int
    index: int
    command: str
```

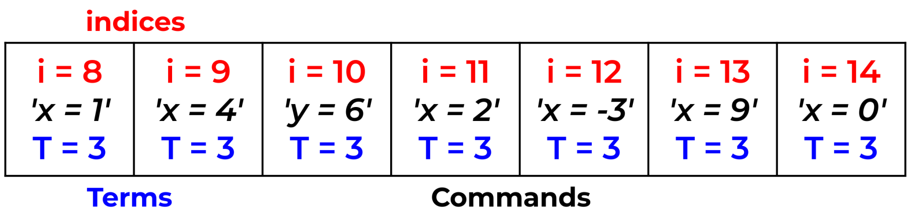


Fig. 4. Raft's log is fundamentally an array made of entries

3.3 Log Replication and Overwriting

All log propagation revolves around one remote procedure call named *append_entries_rpc*, which the leader calls on a list of server proxies that connects it to the followers, which in turn call on a proxy of the leader. It must be registered in the server to be callable, as seen in listing 9.

Listing 9. Register, thus making it callable, the remote procedure call *append_entries_rpc*

```
def handle_server():                                # enclose server in a callable function
    with Raft(...) as server:                       # creates SimpleXMLRPCSever
        def append_entries_rpc(entries, term, commit_index, all_log):
            # ...
        server.register_function(append_entries_rpc)    # makes function callable on the other side
        server.serve_forever()                      # keeps server alive
```

3.3.1 *Leader Propagates Entries.*

The leader (each server as a matter of fact) periodically checks whether there are new commands to propagate (always stored in queue *pygame_commands*, more details in section 4), by overriding *SimpleXMLRPCServer*'s method *service_actions* (listing 10).

Then, translates them into entries by giving each of them the current term and a log index that starts from $lastLogEntry(index) + 1$ and increases by one for each entry. To clarify: if $lastLogEntry(index) = 7$ and we have three new commands, their indexes will respectively be eight (8), nine (9) and ten (10). The translation can be seen at listing 11.

At this point, it propagates *new_entries* to the whole cluster, updating the commit index (necessary for applying log to state) as soon as propagations are successful on at least half of the cluster, like so: $commitIndex = lastNewEntry(index)$.

What happens if the *append entries* gets rejected? The leader adds to *new_entries* its own last log entry: $new\_entries = lastLogEntry + new\_entries$ (figure 5). Then repeats the propagation procedure, for each reject a new *last log entry* gets added, progressively traversing the log backwards. If, at a certain point, $new\_entries == allLog + new\_entries$ (i.e., all leader's log gets propagated) the flag *all_log* is set to *True*.

Since every server may reject or accept different sets of entries, depending on their own local log, every propagation must be "local" for each follower.

The flow of execution for the log propagation is: *Raft: service_actions* → *Raft: propagate_entries* → *propagate_entries :encapsulates_proxy: append_entries_rpc*. The last one gets called as many times as needed on every single follower.

Of course, all propagation happens concurrently using a *ThreadPoolExecutor*, and the code for entries propagation (leader's side) can be seen at listing 12.

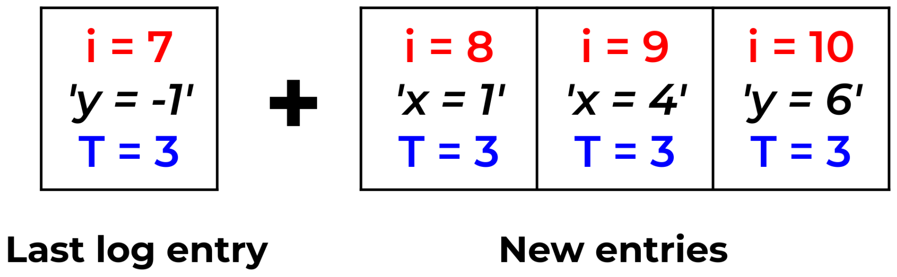


Fig. 5. New entries with the last log's entry added to the beginning of the list

Listing 10. Periodically checks whether there are new commands

```
def service_actions(self):                      # RUN method of the server, override
    if time.time() - self.countdown >= .005:    # do actions every .005 seconds
        global pygame_commands

        if not pygame_commands.empty():         # propagate entries to cluster
            self.propagate_entries()
```

Listing 11. Translates commands into new entries

```
def propagate_entries():
# ...
    while not pygame_commands.empty():
        command = pygame_commands.get()
        log_index += 1                          # lastLogEntry(index) + 1
        self.new_entries.append(Raft.Entry(     # append to support list 'new_entries'
            term=self.term,
            index=log_index,
            command=command
        ))
```

Listing 12. Leader propagation procedure, for complete code refer to project's repository

```
def propagate_entries(self):
    # ...
    if self.log:                            # travel backwards through self.log for each reject
        entries: list[Raft.Entry] = []
        entries.append(self.log[-1])
        entries.extend(self.new_entries)
        log_iterator: int = -2              # log_iterator soft resets for each follower
    else:
        entries: list[Raft.Entry] = self.new_entries
        log_iterator: int = -1

    # ...
    # inner function necessary for concurrent execution
    def encapsulate_proxy(self, follower, entries, log_iterator):
        # ...
        with xmlrpc.client.ServerProxy(complete_url, allow_none=True) as proxy:
            while not propagation_successful:
                # send new entries (local for each follower)
                # ...
                result = proxy.append_entries_rpc(entries, self.term, self.commit_index, all_log)
                if result[0] == False:
                    # add another entry from self.log to new entries
                    entries = [self.log[log_iterator]] + entries
                    log_iterator -= 1
                elif result[0] == True:
                    propagation_successful = True

        return propagation_successful # to propagate_entries, make propagation counter increase
        # encapsulate_proxy ends

    results = []

    # fires RPCs concurrently using ThreadPoolExecutor
    future_result = {           # clever python syntax trick
        self.executor.submit(
            encapsulate_proxy,  # function
            self,               # function's parameter
            follower,           # function's parameter
            entries,            # function's parameter
            log_iterator        # function's parameter
            ): follower for follower in self.cluster}
    for future in concurrent.futures.as_completed(future_result):
        # results of RPCs
        data = future.result()
        results.append(data)

    # finally counts if propagation was successful enough
    if results.count(True) >= len(self.cluster) / 2:
        self.log.extend(self.new_entries)              # add new entries to log
        self.new_entries.clear()                       # clear new entries list
        self.commit_index = self.log[-1].index         # ensure log gets eventually applied
    else:
    #   new entries are not cleared, so they will be propagated again
```

3.3.2 Follower Receives Entries. When a follower receives an *append entries* request from the leader, first checks whether leader is up to date. If it's not, i.e., $leaderTerm < followerTerm$, rejects by answering with the tuple $(False, followerTerm)$. In this context, *answering* is done via the remote procedure call's return value.

On the other hand, if the leader's term is equal or greater than its own (i.e., $leaderTerm \geq followerTerm$), the follower updates its commit index and, if $leaderEntries \neq \emptyset$, checks the *all_log* flag. If it's *True*, clears all its own log to overwrite it with the leader's (fundamental to log forcing, listing 13). Otherwise ($all\_log \neq True$), the leader did not send all its log, so the follower searches through its own log for an entry equal to the leader's previous one (i.e., the entry preceding the new ones). Let's make an example:

- Leader's log = $[1, 2, 3, 4, 5]$;
- Leader's new entries = $[6, 7]$;
- Thus leader's prev = $[5]$.

If it finds an entry equal to leader's previous (i.e., $followerLog(someEntry) == leaderPrev$), deletes all log entries that follow it and appends the new ones, otherwise ($\nexists(followerLog(someEntry) == leaderPrev)$) rejects the request. Since the leader, when faced with a reject, adds a new *prev* and keeps repeating the send until it comprises all its log, at a certain point the follower will be forced to overwrite all its log, thus making it equal to the leader's. This overwriting is called *log forcing* and ensures that all logs are equal to the leader's.

The code can be seen at listing 14 (for the complete one refer to the repository).

Listing 13. Follower clears its own log to overwrite it with the leader's

```
if all_log == True:
    server.log.clear() # if leader sent all its log, clear and rewrite log (leader's log forcing)
```

Listing 14. Follower search in its own log for an entry equal to leader's prev

```
if commit_index is not None:
        server.commit_index = commit_index  # update commit index
```

```
if entries is not None:     # not an heartbeat
    if all_log == True:     # complete overwrite
        server.log.clear()

    if server.log:  # if follower's log not empty search for an entry equal to leader's prev
        entry_log_index: int | None = None              # save its log index (!= entry index)
        for i, my_entry in enumerate(server.log):
            if (my_entry.index == entries[0].index
                and my_entry.term == entries[0].term):
                entry_log_index = i
                break # no need to search further
        # here entry_log_index == (position of entry equal to leader.prev) | None

        if entry_log_index is None:         # entry equal to leader's prev not found
            return(False, server.term)      # rejects

        del server.log[(entry_log_index ):] # delete all log following leader prev

    server.log.extend(entries) # append new entries
```

3.3.3 Follower Sends Entries. Since every server is a Raftian node with a game instance and thus player inputs, followers have their own Pygame commands to propagate. Just like the leader, in their *service_actions* function they periodically check whether there are new commands to propagate and call *propagate_entries* accordingly. Then, they translate all Pygame commands into entries (same code as listing 11) and propagate them to the leader via *append_entries_rpc*. Nothing else.

As previously stated, followers are *passive*, meaning they do not apply their own player inputs when they register them, but only after the leader propagates them back to the whole cluster.

3.3.4 Leader Receives Entries. The leader does very little when receives entries from the followers: it just puts them into its own *pygame_commands* queue. They will get processed and propagated eventually, as stated in section 3.3.1.

3.4 Apply Log to State

Let’s first explain two key attributes: *commit index* and *last applied*. Both of these represent an index, but the former is the highest-index entry successfully propagated in the cluster, while the latter is the highest-index entry already applied to state.

Every node, whether leader or follower, applies entries to state in the same way: inside their function *service_actions* they periodically check if there is a discrepancy between *commit index* and *last applied* attributes (i.e., $commit\_index > last\_applied$). Then, starting from the last applied entry, they apply to state all successive entries up to and including the one with the same index as *commit_index*, updating *last_applied* as they go. To clarify: servers apply all entries between $log(entry.index == last\_applied)$ and $log(entry.index == commit\_index)$ as shown in figure 6.

To apply entries in our context means that they get appended to the queue *raft_orders*. The code can be seen at listing 15 (for the complete source refer to the repository)

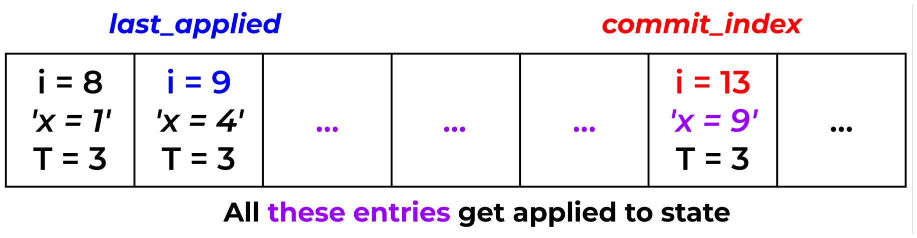


Fig. 6. All entries between *last_applied* and *commit_index* (included) get applied to state

Listing 15. All nodes apply entries to state based on *commit_index*

```
def service_actions(self):
        # ...
        if self.commit_index is not None and self.commit_index > self.last_applied:
            global raft_orders  # applying means appending entries to this queue

            #...

            last_applied_log_position: int = -1
            for i, my_entry in enumerate(self.log):
                if (my_entry.index == self.last_applied):
                    last_applied_log_position = i
                    break # found log position of last applied entry

            log_iterator = last_applied_log_position + 1    # improves code clarity

            while self.last_applied != self.commit_index:
                raft_orders.put(self.log[log_iterator])
                self.last_applied = self.log[log_iterator].index
                log_iterator = log_iterator + 1
            # here self.last_applied == self.commit_index
```

3.5 Log Compaction

This functionality and all the successive ones have not been implemented due to time constraints. We decided to include them anyway since they were still the product of careful consideration and could prove useful in future implementations.

Log compaction, also called *snapshot*, is a way to clean servers' logs and save them in persistent memory, necessary in long-running or message-heavy applications. Every node autonomously decides when to do it.

The idea follows: inside their function *service_actions*, servers check whether their log is larger than a certain size (to be determined by testing) and then call a *snapshot* method accordingly, which saves all entries in a JSON file progressively deleting them from the log, from the first one up to (but excluding) the *last applied*.

An alternative interpretation, closer to Ongaro and Ousterhout's idea, is saving the state of the cluster. When applied to our context, we can imagine a JSON file that describes all servers' (i.e., players') state, by saving for each of them *id*, *url*, *port*, and *hp* values. *Index* and *term* of the last snapshotted entry should also be saved, as well as the current configuration (which will be mentioned in section 3.7). An example of the JSON structure can be seen at listing 16.

This method requires more pre-processing to generate the JSON, and more post-processing to later check log's correctness, but is also more compact, which helps when the leader calls *install snapshot*, a remote procedure call that can sometimes be needed to bring up to speed new or outdated servers by sending them the leader's snapshot.

It goes without saying, but whichever method is ultimately chosen, every new snapshot must comprise all information contained in the old ones.

Listing 16. The JSON for a snapshot that saves cluster's state would look something like this

```
{
    servers:[
        {
            id : 1,
            url : "localhost",
            port : 8000,
            hp : 70
        },
        {}, {}
    ],
    lastIndex : 11,
    lastTerm : 3,
    lastConfig : 8
}
```

3.6 Leader Election

To grant consistency, Raft's design choice is to centralize all decisions on one node, called leader, that synchronizes all cluster's nodes. To make the protocol fault tolerant, the leader can dynamically change over time via a distributed election: whenever a follower finds out that the leader is either outdated or missing (i.e., internal follower's timer times out before receiving any call from it), said follower starts an election. It changes its internal state to *candidate*, increases its own term by one, votes for itself, and then propagates to the whole cluster a specialized remote procedure call named *request_vote_rpc* (code, removed in the final version, at listing 17). Votes are given on a first-come-first-served basis, and to prevent split votes each server's election timeout is randomized between 150ms and 300ms at the start of every election. This ensures that in most cases only one server will be candidate at a time.

At this point there are two possible outcomes: more than half of the cluster votes for the candidate (which we will call *"A"*), that therefore becomes leader and propagates a heartbeat to the whole cluster, or another candidate (which we will call *"B"*) is more up-to-date (i.e., *B's term* is greater than *A*'s or equal but with a greater *lastIndex*). In this last case, candidate *A* reverts back to follower and votes for *B*. The pseudocode for all the above can be seen at listing 18.

One last eventuality is that the old leader manifests itself. In this case, if the old one is equally or more up-to-date than the new one (both term and last index count), the latter reverts back to follower and the preceding monarch gets reinstantiated.

Listing 17. Pseudocode for *request_vote_rpc*

```
def request_vote_rpc(...):
    # if candidate less up to date -> reject
    if self.term > candidate_term:
        return (self.term, False)

    # if a candidate already exists
    if self.voted_for is not None and not candidate_id:
        return (self.term, False)

    # vote for candidate
    self.voted_for = candidate_id
    return (self.term, True)
#...
server.register_function(request_vote_rpc)
```

Listing 18. Pseudocode for *to_candidate*, gets fired on *election timer* timeout

```
def to_candidate(self):
    self.mode = Raft.Mode.CANDIDATE
    self.term += 1
    self.voted_for = self.id

    self.timer.reset() # reset election timer

    for server in self.cluster:
        server.request_vote_rpc(...)
        count votes

    if some_return.more_up_to_date:
        self.mode = Raft.Mode.FOLLOWER
        self.voted_for = some_return.id

    if votes > len(self.cluster) / 2:
        self.to_leader()    # handles mode change and heartbeat
```

3.7 What is Missing

There is one last functionality discussed by Ongaro and Ousterhout, which is gracefully managing changes in the cluster's members by leveraging a configuration attribute and keeping multiple configurations alive at the same time (figure 7). We never intended to include it due to time constraints, therefore there is nothing we can add beyond the original work.

Another concern, which was not considered in the Raft paper, is faults caused by bad actors that purposely send malicious information. This is a real problem for our use case, since players want to win and are therefore incentivized to act maliciously by cheating (a practice so widespread that has created its own multimillion-dollar market [2]).

The implementation of Byzantine fault tolerance was beyond the scope of this work. Therefore, we refer the reader to some example works for further details: *"A Raft Algorithm with Byzantine Fault-Tolerant Performance"* by Xir and Liu [20], and *"VSSB-Raft: A Secure and Efficient Zero Trust Consensus Algorithm for Blockchain"* by Tian et al. [19].

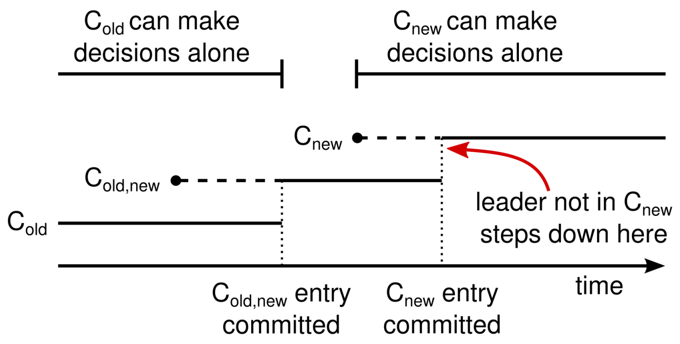

C_{old} can make decisions alone
C_{new} can make decisions alone
C_{new}
$C_{old,new}$
C_{old}
leader not in C_{new} steps down here
time
$C_{old,new}$ entry committed
C_{new} entry committed

Fig. 7. Cluster goes through a hybrid configuration to pass from the old to the new one. Source: Raft paper [12]

4 ARCHITECTURE

In the previous section (3) we explained in detail how nodes communicate with each other and handle their log. Now we will explain what actually happens inside a node. Before showcasing the architecture, we will briefly explain how Pygame works, thus taking the opportunity to present the user interface.

4.1 Pygame

Pygame's approach is very straightforward: first comes the declaration and set up of all the graphical components, such as game window, fonts, colors, variables and constants. Every element gets positioned on the main window by coordinates *(x,y)*, where *(0,0)* is the top left corner. Most items are made of two fundamental Pygame classes: *Rect*, which creates non-graphical objects that expose many useful methods, for example to position, move, and resize themselves, or to detect collisions and mouse clicks, and *Surface*, which is the most basic graphical component that has dimensions and can be drawn upon. Often we want to bind, thus constrain, surfaces with rects so that we use the latter for spatial operations. One last fundamental is the *blit* function, a method that draws one image onto another or, to be precise, that draws a source Surface onto the object Surface that calls it. We can give it an optional argument to specify a drawing destination, either with coordinates or a rect. To clarify: $baseSurface.blit(sourceSurface, destination)$ draws *sourceSurface* onto *baseSurface* at the coordinates specified by *destination*. An example of all the above can be seen at listing 19.

Pygame, being low-level in nature, is very flexible and allows us to do pretty much whatever we want. For example, we defined our players as dataclasses that encapsulate both the players' data (like *id* or *health points*) and their Rect and Surface objects, as in listing 20.

Finally, Pygame gives us direct access to the game loop, which is implemented as nothing more than a *while loop*. In it, we can process player inputs and refresh the screen, dynamically changing what is displayed. In short, we manage everything that happens while the game is running. In listing 21 we can see two types of player input, one for quitting the game and a left-mouse click, the latter of which causes a refresh of the header. Specifically, if *Player 2* gets clicked, the header will display *"Player 2 pressed"*, reverting back to its original state after a couple of seconds. The last command, *clock.tick(fps)*, allows us to limit the framerate, effectively slowing down or speeding up the game engine itself by constraining the amount of times per second the game loop repeats itself.

Listing 19. Pygame base components

```
pygame.init()                               # starts pygame
GREY = (125, 125, 125)                      # define a color
DISPLAY = pygame.display.set_mode((1000, 1200)) # creates game window 1000x1200 pixels in resolution
clock = pygame.time.Clock()                 # necessary to mange fps
font = pygame.font.Font(None, 60)           # creates default font
toptext = font.render("Top Text", False, BLACK) # header text
rect_header = pygame.Rect(0, 0, 1000, 100)  # creates rect for header
header = pygame.Surface((1000, 100))        # creates surface for header
header.fill(WHITE)                          # draw on surface

DISPLAY.blit(header, rect_header)           # draw on DISPLAY the header surface
                                            # position is given by rect_header
#...
# draw text on coordinates
DISPLAY.blit(toptext, (rect_header.centerx - xoffset, rect_header.centery - yoffset))
```

Listing 20. Players defined as dataclasses that encapsulate Pygame elements

```
@dataclass
class Player:
    id: int
    hp: int
    rc: pygame.Rect     # represents player position and size
    ui: pygame.Surface  # exposes UI of the player e.g., colour

player1 = Player(
    id=1,
    hp=100,
    rc=pygame.Rect(585, 685, 80, 80),  # x0, y0, width, height
    ui=pygame.Surface((80,80))
)
player1.ui.fill(RED)                    # colour player red
DISPLAY.blit(player1.ui, player1.rc)    # draw on display via rect
```

Listing 21. All interactions and frame-by-frame rendering happen in the game loop

```
while True: # game loop
    for event in pygame.event.get():    # process player inputs
        if event.type == pygame.QUIT:
            pygame.quit()

        if event.type == pygame.MOUSEBUTTONDOWN and event.button == 1:  # left mouse button click

            pos = pygame.mouse.get_pos() # gets mouse position

            if player.rc.collidepoint(pos): # rect allows us to detect collisions
                toptext = font.render(f"Player {player.id} pressed", False, BLACK)

                DISPLAY.blit(header, rect_header)   # erase previous text
                DISPLAY.blit(toptext, (rect_header.centerx - xoffset, rect_header.centery - yoffset))

    pygame.display.flip()  # refresh on-screen display
    clock.tick(60)         # limits framerate
```

4.2 Raftian User Interface

Figure 8 shows different phases of a normal Raftian's game session. Specifically, they demonstrate how the interface changes when the player repeatedly clicks on (thus *attack*) Player 3, the blue one on the top left (players are represented as four coloured squares on the board). Players' colours become progressively desaturated as their health points decrease, by modifying their alpha channels as shown in listing 22. When a player is dead, its colour changes to a darker shade.

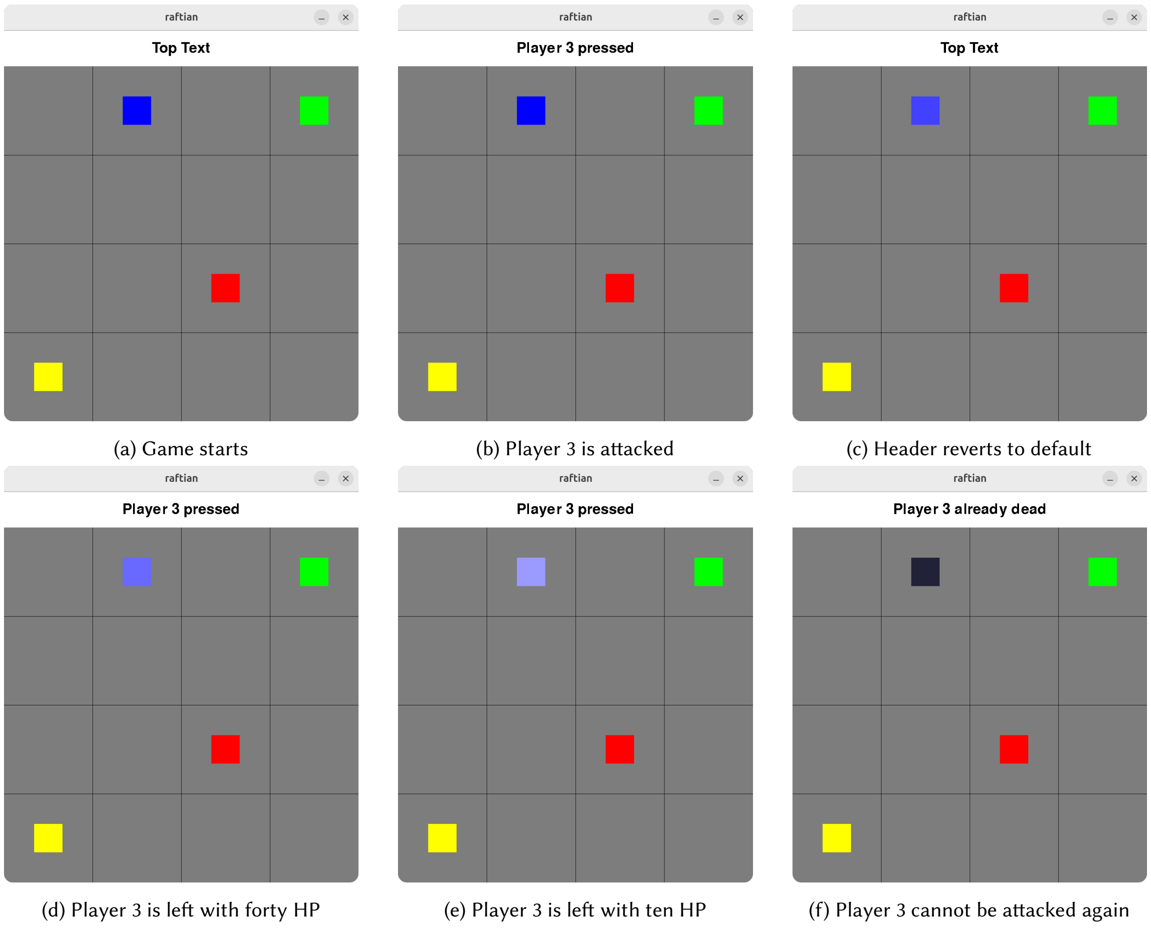
In the header an attack message gets written, reverting back after half a second. Said message changes depending whether the player is still alive; if not, damage is ignored.

Listing 22. Whenever a player gets damaged, its colour gets desaturated

```
for player in players:
    if player.id == order.command and player.hp > 0:
        player.hp -= 30  # apply damage to player:

        if player.hp < 90 and player.hp >= 60:
```

```
        player.ui.set_alpha(190)
        DISPLAY.blit(player_UI_cleaner, player.rc)  # clean player UI
        DISPLAY.blit(player.ui, player.rc)          # redraw player UI
```



(a) Game starts (b) Player 3 is attacked (c) Header reverts to default

(d) Player 3 is left with forty HP (e) Player 3 is left with ten HP (f) Player 3 cannot be attacked again

Fig. 8. Different phases of a normal Raftian's game session. Player 3 keeps getting damaged until it dies

4.3 Raftian Node Architecture

The architecture of a Raftian node can be seen at figure 9. Let's explain it: first of all, the game loop and the server are encapsulated in two functions to be handed over to two different threads, enabling concurrent execution (listing 23). Whenever a player clicks on (i.e., attacks) one of the four players, the game engine does not apply damage immediately. Instead, it generates a *command* which represent, if we want, the *intention* of attacking said player. This *command* is thus appended to a queue called *pygame_commands*, one of the two synchronized queues necessary to allow communication between server and Pygame's threads (listing 24). Both are instances of Python's standard library queue module[20], which implements thread-safe, multi-producer, multi-consumer queues.

At this point, Pygame does not concern itself anymore with said user input. The server, by itself, periodically checks the *pygame_commands* queue (as in listing 10) and, when not empty, removes elements from it (as in listing 11) and propagates them as entries to the Leader (or to the whole cluster if said server *is* the Leader, as in listing 12).

Then, the Leader propagates to the whole cluster the commands received, which we will now call *orders*. Each server adds received orders to its own log, as explained in section 3.3, so that they can later be appended to the *raft_orders* queue when entries get applied to state (as in section 3.4). This way, the original user input gets propagated back to the server that generated it in the first place.

Finally, Pygame checks (periodically) the *raft_orders* queue for orders. When it finds them, it removes them from the queue and updates the user interface accordingly (an example can be seen at listing 25).

The whole idea is to keep server and game engine as separated as possible: the former reads commands, propagates them and writes received orders, the latter reads orders, updates the UI, and writes commands, following a unidirectional cyclic communication pattern.

Listing 23. Start both Pygame and server's threads

```
def handle_pygame():
    pygame.init()
    #...
    While True:
         #...
def handle_server():
    with Raft(...) as server:
        #...
        server.serve_forever()

server_thread = threading.Thread(target=handle_server)
server_thread.start()
pygame_thread = threading.Thread(target=handle_pygame)
pygame_thread.start()
```

Listing 24. Queues for commands and orders, they allow inter-thread communication

```
# user inputs through Pygame which writes them here
# Raft reads them and propagates them to the cluster
pygame_commands = Queue()

# commands that have been applied to state are written here by Raft
# Pygame reads them and updates UI accordingly
raft_orders = Queue()
```

[20]Python's queue, a synchronized queue class: https://docs.python.org/3/library/queue.html

Listing 25. Pygame periodically checks whether there are new orders and updates the UI accordingly

```python
while True: # Pygame's main loop
    #...
    while not raft_orders.empty():
        order: Raft.Entry = raft_orders.get()

        for player in players:
            if player.id == order.command and player.hp > 0:
                player.hp -= 30  # apply damage to player:
                #...
```

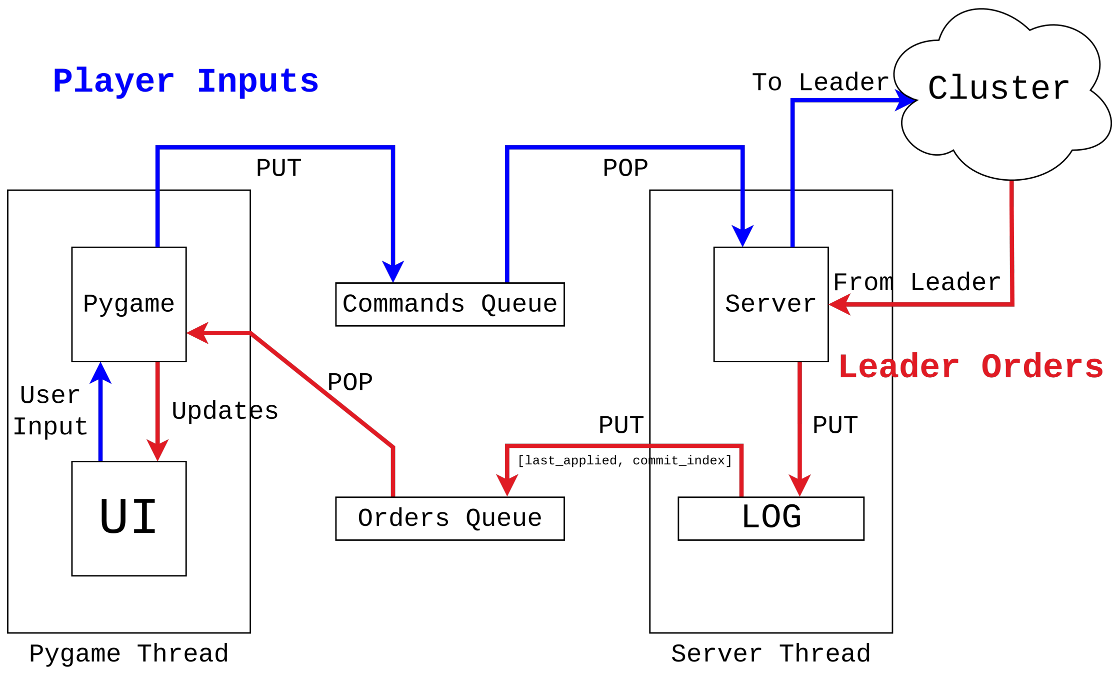


Fig. 9. Raftian node architecture

4.4 Evaluation and Results

While we did not have enough time to implement evaluation procedures for measuring things like response time or network latency, the game has been thoroughly tested. Responsiveness is good, with no noticeable input latency and a solid framerate way above sixty frames per second (which is still the preferred limit), and both log replication and overwriting functionalities have been confirmed working as intended. The proof of this fact can be observed in the logs at https://github.com/mhetacc/RuntimesConcurrencyDistribution/tree/main/logs, specifically by comparing logs among folders *bob1*, *bob2*, *bob3*, *bob4*, and *raftian*: all logs (meaning *Raft logs*) contain the same entries, in the same order, between all players.

5 REFLECTION

Let's now discuss problems, potential future expansions and learning outcomes of this project.

5.1 Self-Assessment

Using *XML-RPC* and *threading* libraries proved to be sub-optimal: the former has a very contrived syntax and makes writing procedure calls a bit unintuitive, since forces the programmer to think in the "opposite direction". When writing a remote procedure call (i.e., those functions that get registered by *register_function()*) is important to keep in mind that they are going to be used by the caller and not by sender (in whom code block they are written in). Regarding *threading*, a potentially easier approach would have been leveraging asynchronous programming instead (thanks to *asyncio* library).

Code could be less coupled: both server and game loop reside in the same file, and a lot of components are either internal classes or nested functions. Moreover, both command and order queues are global variables, which is generally a practice to be avoided.

On the other hand, code is well documented and as understandable as possible, even though following the flow of, for example, an input propagation requires jumping through it many times.

5.2 Future Works

Apart from the two features already discussed (leader election and log compaction), future expansions could implement cluster's membership change and a Byzantine fault-tolerant version of Raft. Adding new game functionalities, thus command types, should be easy since they can be propagated by the existing infrastructure, and the same is true for adding new players: provided that a new, bigger, user interface gets created, changing cluster's size should, in our testing, work without any issues.

5.3 Learning Outcomes

We started this project by having very limited Python competencies, having never written concurrent programming, never touched a game engine and never worked with network transmission protocols. All in all, we learned all of the above, in some cases going so far as trying different alternative solutions (we implemented Raft nodes mockups with threading, multiprocessing and asyncio libraries), making this project an invaluable learning experience.

REFERENCES

[1] 2009. *Distributed Programs.* Springer London, London, 373–406. https://doi.org/10.1007/978-1-84882-745-5_11
[2] Matt Besturgess. 2025. Inside the Multimillion-Dollar Gray Market for Video Game Cheats. https://www.wired.com/story/inside-the-multimillion-dollar-grey-market-for-video-game-cheats/
[3] Michiel Buijsman, Devan Brennan, Tianyi Gu, Lester Isaac Simon, Tomofumi Kuzuhara, Spyros Georgiou, Michael Wagner, Ngoc Linh Nguyen, Brett Hunt, Alejandro Marin Vidal, and Tiago Reis. 2024. Newzoo's Global Games Market Report 2024. https://newzoo.com/resources/trend-reports/newzoos-global-games-market-report-2024-free-version
[4] Jessica Clement. 2025. Free-to-play (F2P) games market revenue worldwide from 2018 to 2024. https://www.statista.com/statistics/324129/arpu-f2p-mmo/
[5] Stephen Crass. 2024. The Top Programming Languages 2024. https://spectrum.ieee.org/deep-brain-stimulation-depression
[6] Michael J. Fischer, Nancy A. Lynch, and Michael S. Paterson. 1985. Impossibility of distributed consensus with one faulty process. *J. ACM* 32, 2 (April 1985), 374–382. https://doi.org/10.1145/3149.214121

[7] George Coulouris, Jean Dollimore, Tim Kindberg, and Gordon Blair. 2012. *Distributed Systems: Concepts and Design* (5 ed.). Addison-Wesley.

[8] Paul Jansen. 2025. TIOBE Index for August 2025. https://www.tiobe.com/tiobe-index/

[9] Rida Khan and Zack Aboulazm. 2023. $300 Billion of Video Gaming Revenue, by Segment (2017-2026). https://www.visualcapitalist.com/sp/video-games-industry-revenue-growth-visual-capitalist/

[10] Lamport. 1979. How to Make a Multiprocessor Computer That Correctly Executes Multiprocess Programs. *IEEE Trans. Comput.* C-28, 9 (1979), 690–691. https://doi.org/10.1109/TC.1979.1675439

[11] Leslie Lamport. 1998. The part-time parliament. *ACM Trans. Comput. Syst.* 16, 2 (May 1998), 133–169. https://doi.org/10.1145/279227.279229

[12] Diego Ongaro and John Ousterhout. 2014. In search of an understandable consensus algorithm. In *Proceedings of the 2014 USENIX Conference on USENIX Annual Technical Conference* (Philadelphia, PA) *(USENIX ATC'14)*. USENIX Association, USA, 305–320.

[13] David A. Patterson and John L. Hennessy. 2008. *Computer Organization and Design, Fourth Edition, Fourth Edition: The Hardware/Software Interface (The Morgan Kaufmann Series in Computer Architecture and Design)* (4th ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[14] PWC. 2025. Global entertainment and media industry revenues to hit US$3.5 trillion by 2029, driven by advertising, live events, and video games: PwC Global Entertainment & Media Outlook. https://www.pwc.com/gx/en/news-room/press-releases/2025/pwc-global-entertainment-media-outlook.html

[15] Michael J. Quinn. 1994. *Parallel computing (2nd ed.): theory and practice.* McGraw-Hill, Inc., USA.

[16] A. Rollings and E. Adams. 2003. *Andrew Rollings and Ernest Adams on Game Design.* New Riders. https://books.google.it/books?id=Qc19ChiOUI4C

[17] Abraham Silberschatz, Peter Baer Galvin, and Greg Gagne. 2008. *Operating System Concepts* (8th ed.). Wiley Publishing.

[18] A.S. Tanenbaum and M. van Steen. 2017. *Distributed Systems.* CreateSpace Independent Publishing Platform. https://books.google.it/books?id=c77GAQAACAAJ

[19] Siben Tian, Fenhua Bai, Tao Shen, Chi Zhang, and Bei Gong. 2024. VSSB-Raft: A Secure and Efficient Zero Trust Consensus Algorithm&nbsp;for Blockchain. *ACM Trans. Sen. Netw.* 20, 2, Article 34 (Jan. 2024), 22 pages. https://doi.org/10.1145/3611308

[20] Ting Xie and Xiaofeng Liu. 2022. A Raft Algorithm with Byzantine Fault-Tolerant Performance. In *Proceedings of the 5th International Conference on Information Science and Systems* (Beijing, China) *(ICISS '22)*. Association for Computing Machinery, New York, NY, USA, 95–99. https://doi.org/10.1145/3561877.3561892

[21] Albert Y. H. Zomaya. 1996. *Parallel and distributed computing handbook.* McGraw-Hill, Inc., USA.

Received 05/09/2025; revised –; accepted –