SWE Assignment-3

Table of Contents

Table of Contents

Team Members

Architectural Patterns

1) Micro Services

High level architectural diagrams

Architectural Diagrams

Overview of System

Instructions for Running

Quality

2) Publisher Subscriber Pattern

High level architectural diagrams

Architectural Diagrams

Overview of the System

Instructions for Running

Quality

Comparison

Average Response Time

Requests per Second

Explanation

Team Members

- 1) Vineeth Bhat (2021101103)
- 2) Ishwar B Balappanawar (2021101023)
- 3) Swayam Agrawal (2021101068)
- 4) Mitansh Kayathwal (2021101026)
- 5) G.L.Vaishnavi (2023204009)

Architectural Patterns

1) Micro Services

Microservices Explanation

Microservices architecture breaks down the application into small, independent services, each responsible for specific functionalities.

- This allows for better scalability, flexibility, and maintainability.
- Each microservice can be developed, deployed, and scaled independently, which aligns well with the requirement of supporting up to 1000 requests/second.
- Additionally, microservices can help manage resources efficiently, enabling graceful degradation in case of sensor power outage or failures.

High level architectural diagrams

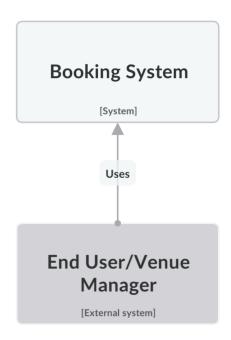
The diagrams can be viewed here as well with description for all the apps (containers):

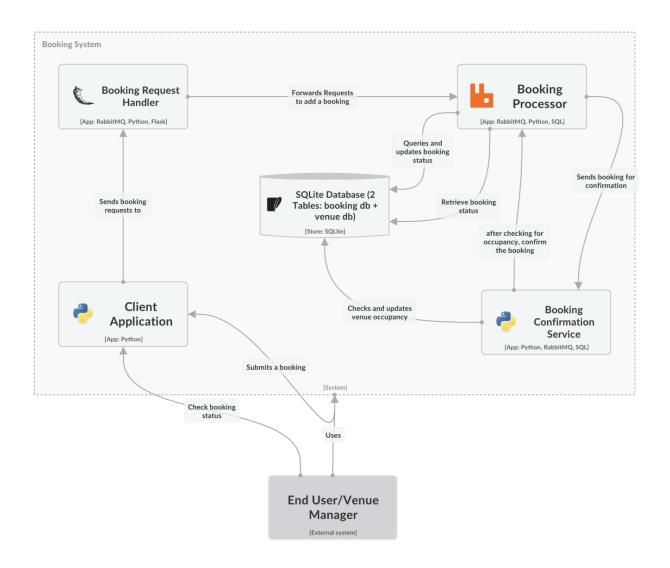
Explain complex software systems

IcePanel is a collaborative diagramming tool that helps software engineering and product teams align on technical decisions. Create an interactive map of your software systems and give your teams full context about how things work.

https://s.icepanel.io/aBtDn9Z0T0o6iY/MSqD

Architectural Diagrams





We can further decompose the containers to create a component diagram, but since the assignment mentions to create only high level diagrams, that hasn't been included.

An example: Consider "Booking Processor" container. We can use component blocks to represent internal components like "Request Receiver," "Booking Manager," "Confirmation Handler," and "Status Manager." We then connect these components with arrows to illustrate the flow of information, like how "Request Receiver" sends data to "Booking Manager."

Overview of System

Overview:

The seat booking system prototype is a microservices-based architecture built to handle booking requests for various venues. It consists of several components including a client interface, booking request handler, booking processor, and booking confirmation service. The system utilizes Flask for creating RESTful APIs and RabbitMQ for inter-service communication via message queues.

Client Code:

The client code provides a simple command-line interface for users to make bookings and check booking statuses. It interacts with the backend services via HTTP requests using the requests library.

Booking Request Handler (booking_request_handler.py):

This component receives booking requests from clients and forwards them to the booking processor via RabbitMQ. It exposes two endpoints '/book' for making bookings and '/status/<request_id>' for checking booking statuses.

Booking Processor (booking_processor.py):

The booking processor is responsible for processing booking requests and managing booking statuses. It listens for incoming requests from the booking request handler via RabbitMQ, processes them, updates the status, and sends back responses. It also handles status check requests and updates via separate queues.

Booking Confirmation Service (booking_confirmation.py):

Upon receiving confirmation messages from the booking processor, this service checks if the requested seats can be booked without exceeding the venue's maximum occupancy. It updates the venue's occupancy accordingly and sends status updates back to the booking processor.

Utilities (utilities.py):

This module contains utility functions for managing SQLite databases, including

creating tables, inserting data, updating statuses, and retrieving occupancy information. It is used by both the booking processor and the booking confirmation service.

Instructions for Running

You need RabbitMQ running for the code to function. Please use

```
sudo apt-get update
sudo apt-get install rabbitmq-server
sudo systemctl start rabbitmq-server
sudo systemctl enable rabbitmq-server
sudo systemctl status rabbitmq-server
```

After this, you may run the files as:

1. Initial Setup:

Run

clean_init.py to initialize the SQLite database and create necessary tables.

2. Start Services:

- Run booking_request_handler.py to start the booking request handler service.
- Run booking_processor.py to start the booking processor service.
- Run booking_confirmation.py to start the booking confirmation service.

3. Client Usage:

- Execute the client code (client.py) to interact with the system.
- Choose the action:
 - [1] Make Booking: Enter venue ID and number of seats to make a booking.
 - [2] Check Status: Enter request ID to check booking status.

4. Monitoring:

- Monitor the console output of each service for logs and messages.
- Check the SQLite database (request_booking_db.sqlite) for status updates and venue occupancy.

Quality

requirements

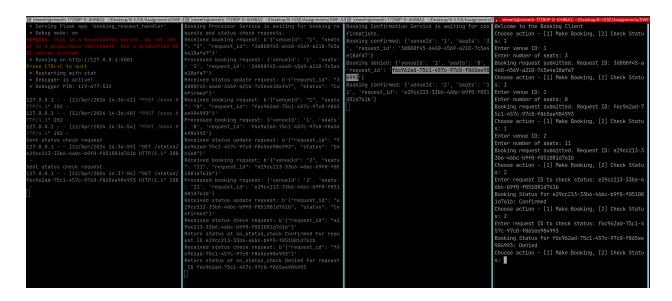
1. Performance:

- NOTE: Added load testing results below this section
- SQLite typically exhibits high performance for single-threaded, low to
 moderate concurrency workloads. It offers fast read and write operations,
 with transactional support ensuring data integrity. For our operation, we
 understand that SQL gaurantees upto a million requests/second satisfying
 the requirement of a 1000 requests/second
 (https://stackoverflow.com/a/57846740). This is subject to us having
 access to actual servers rather than just running it on our laptops.
- RabbitMQ is known for its low-latency message delivery and high throughput, capable of handling thousands to millions of messages per second depending on hardware and configuration. For our operation, we note that it can handle queue sizes of upto 1MiB, showing that it is sufficient for maintaining enough messages in the queue to counteract if the SQL server ever slows down.

2. Reliability:

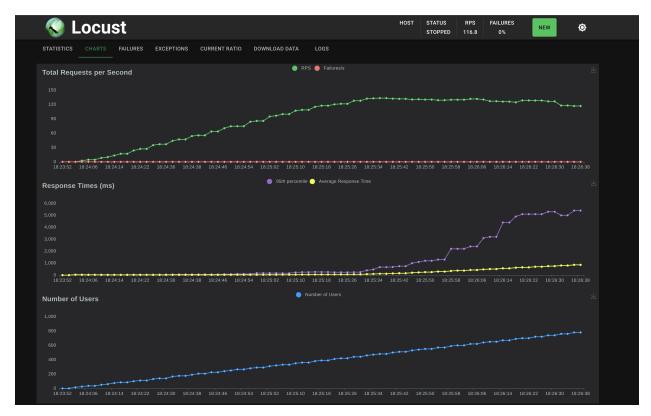
- SQLite ensures reliability through ACID transactions and crash recovery mechanisms. It provides data durability by persisting changes to disk before committing transactions.
- RabbitMQ ensures reliability by persisting messages to disk and employing message acknowledgments to prevent message loss. Clustering and mirrored queues enhance fault tolerance and data redundancy, further improving reliability.

Example

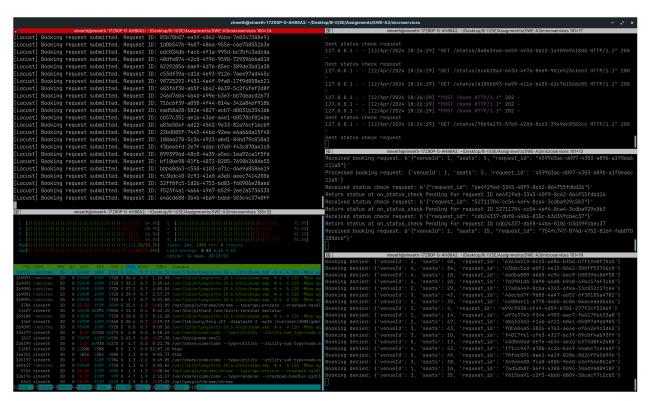


- In this demonstration, we run the handler, the processor, the confirmation and the client in separate terminals and run the following operations
 - Book 3 seats in venue 1 this is accepted
 - Book 8 seats in venue 1 this is denied since we define the max occupancy of the venue as 10
 - Book 11 seats in venue 2 this is accepted as the max occupancy of venue
 2 is 12
 - Check status of the operation of booking 11 seats in venue 2 this returns
 Confirmed to the client
 - Check status of the operation of booking 8 seats in venue 1 this returns
 Denied as explained earlier

Load Testing Results



Load Testing on Locust



Results shown on local system

So, we successfully demonstrate that our system is scalable and can handle uptop 100 requests/seconds on our local machine. Increasing compute speeds will increase this as demonstrated by the image of system resources using <a href="https://example.com/https://examp

2) Publisher Subscriber Pattern

Publisher-Subscriber Explanation

Pub-Sub architecture is a messaging pattern where senders of messages, called publishers, do not program the messages to be sent directly to specific receivers, called subscribers. Instead, messages are published to channels, without knowledge of what (if any) subscribers there may be. Subscribers express interest in one or more channels and receive messages only if they are subscribed to the relevant channel(s).

- Publishers and subscribers can be added or removed independently, allowing for scalable and flexible systems.
- loose coupling components don't need to know about each other, promoting flexibility and easier modifications.
- Messages are sent asynchronously, improving system responsiveness.
- **Example Services**: Services like sensors publish data, while display services subscribe to receive and display relevant data.
- Subscribers control their own message consumption, enabling decentralized decision-making.
- Load Balancing and Scalability- Load balancers distribute messages among subscribers for scalability.

High level architectural diagrams

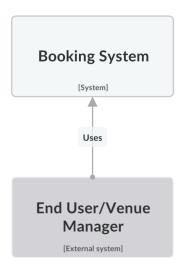
The diagrams can be viewed here as well with description for all the apps (containers):

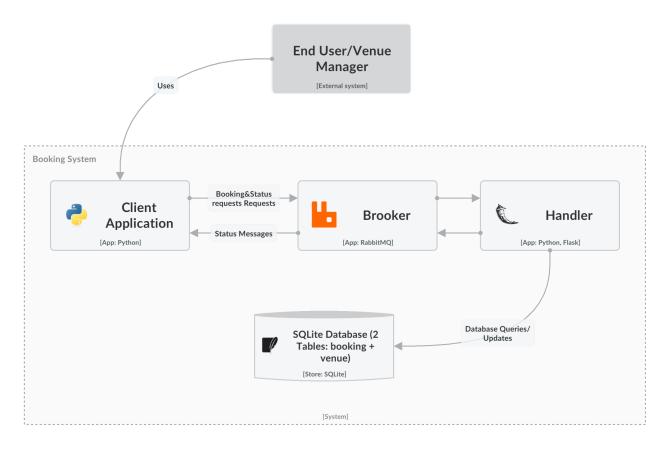
Explain complex software systems

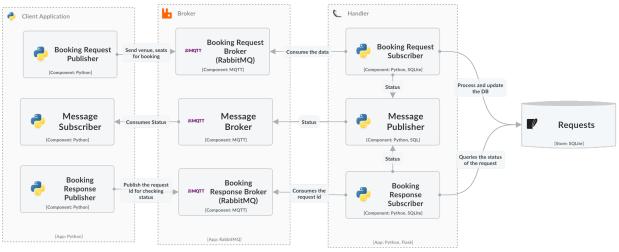
IcePanel is a collaborative diagramming tool that helps software engineering and product teams align on technical decisions. Create an interactive map of your software systems and give your teams full context about how things work.

https://s.icepanel.io/aBtDn9Z0T0o6iY/MSqD

Architectural Diagrams







Overview of the System

Overview:

This prototype employs a publisher-subscriber (Pub-Sub) architecture to distribute event information to subscribers interested in specific topics. It consists of distinct components, including publishers, subscribers, and a message broker for facilitating communication. The system utilizes RabbitMQ as the message broker for efficient message delivery.

Publishers:

booking_request_publisher: Responsible for publishing booking requests, including venue and seat details, to the message broker.

booking_response_publisher: Publishes request IDs to the message broker for querying booking request statuses.

message_publisher: Publishes status updates and other messages to the message broker for distribution to subscribers.

Subscribers:

booking_request_subscriber: Subscribes to booking requests, processes them, updates the database, and publishes status updates.

booking_response_subscriber: Subscribes to request IDs, queries the database for statuses, and publishes them.

message_subscriber: Subscribes to general messages and status updates published by the system.

Client:

Utilizes booking_request_publisher to submit booking requests and booking_response_publisher to inquire about the status of requests. Additionally, it subscribes to messages relevant to its requests via message_subscriber.

Handler:

Orchestrates the flow of data within the system. It includes logic for receiving requests from clients, directing them to the appropriate processing components

(e.g., booking_response_subscriber), and managing the response flow back to clients.

Utilizes message_publisher.py for distributing status updates.

Message Broker:

RabbitMQ: Acts as the message broker for the system, facilitating communication between publishers and subscribers. It efficiently routes messages from publishers to subscribers based on topics or routing keys, ensuring reliable message delivery.

Utilities:

This module contains utility functions for managing SQLite databases, including creating tables, inserting data, updating statuses, and retrieving occupancy information. It is used by both the booking processor and the booking confirmation service.

Instructions for Running

You need RabbitMQ running for the code to function. Please use

```
sudo apt-get update
sudo apt-get install rabbitmq-server
sudo systemctl start rabbitmq-server
sudo systemctl enable rabbitmq-server
sudo systemctl status rabbitmq-server
```

After this, you may run the files as:

1. Initial Setup:

Run

clean_init.py to initialize the SQLite database and create necessary tables.

2. Client Usage:

- Execute the client code (client.py) to interact with the system.
- Choose the action:
 - [1] Make Booking: Enter venue ID and number of seats to make a booking.
 - [2] check status: Enter request ID to check booking status.

3. Publishers/Subscribers:

- Run request_subscriber.py to start the listening to messages sent by booking_request_publisher (Client)
- Run response_subscriber.py to start the listening to messages sent by booking_response_publisher (Client)
- Run <u>message_</u> <u>subscriber.py</u> to start the listening to messages sent by message_publisher (Handler)

4. Monitoring:

- Monitor the console output of each service for logs and messages.
- Check the SQLite database (request_booking_db.sqlite) for status updates and venue occupancy using testdb.py

Quality

requirements

1. Performance:

- NOTE: Added load testing results below this section
- SQLite typically exhibits high performance for single-threaded, low to
 moderate concurrency workloads. It offers fast read and write operations,
 with transactional support ensuring data integrity. For our operation, we
 understand that SQL gaurantees upto a million requests/second satisfying
 the requirement of a 1000 requests/second
 (https://stackoverflow.com/a/57846740). This is subject to us having
 access to actual servers rather than just running it on our laptops.

 RabbitMQ is known for its low-latency message delivery and high throughput, capable of handling thousands to millions of messages per second depending on hardware and configuration. For our operation, we note that it can handle queue sizes of upto 1MiB, showing that it is sufficient for maintaining enough messages in the queue to counteract if the SQL server ever slows down.

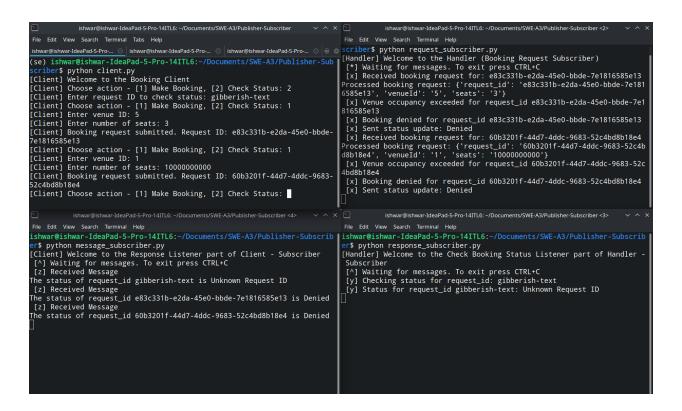
2. Reliability:

- SQLite ensures reliability through ACID transactions and crash recovery mechanisms. It provides data durability by persisting changes to disk before committing transactions.
- RabbitMQ ensures reliability by persisting messages to disk and employing message acknowledgments to prevent message loss. Clustering and mirrored queues enhance fault tolerance and data redundancy, further improving reliability.

Example Demonstration:

```
chwar@chwarldeaPadsPro-14TL6:-/Documents/SWE-AJPublisher-Subscriber
File Edit View Sarch Terminal Tabs Holp
Immar@chwarldeaPadsPro-
Immar@chwarldeaPad
```

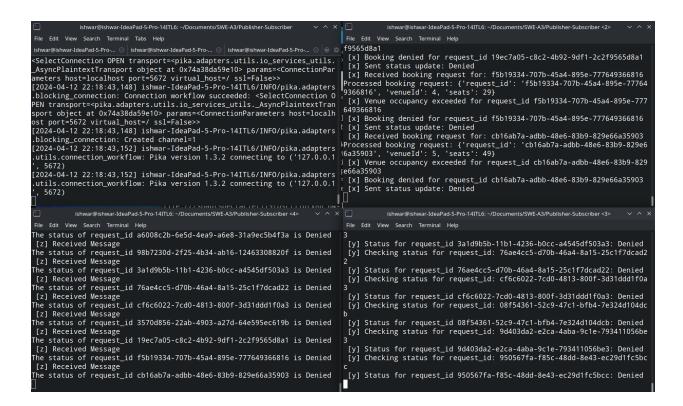
- Book 2 seats in venue 2 this is accepted, message sub receives the status.
- Check status of the above operation response sub get the request and the status is sent to message sub



- Checking status of a gibberish request id → Unknown Request id
- Similarly trying to book in unavailable venues/ booking too many seats →
 Denied

Load Testing Results





So, we successfully demonstrate that our system is scalable and can handle uptop 350+ requests/seconds on our local machine.

Comparison

Average Response Time

Micro-services starts with **20 ms** response time, whereas Pub Sub starts with **0.3** ms

Buy the time there are 500 users,

Response Time for Micro-services - **619.89 ms**, Response Time for Pub Sub - **0.53 ms**

Buy the time there are 1000 users,

Response Time for Micro-services - **2014.42 ms**, Response Time for Pub Sub - **0.62 ms**

Requests per Second

Buy the time there are 500 users,

RPS for Micro-services - **89**, RPS for Pub Sub - **334**

Buy the time there are 1000 users,

RPS for Micro-services - **101**, RPS for Pub Sub - **367**

• 0 fails in both

Explanation

Possible reasons for the better performance of Pub Sub Architecture

1. Decoupling in Pub Sub:

Pub Sub architecture inherently decouples the components involved in communication. In a micro-services architecture, each service typically communicates with others via synchronous HTTP requests, leading to tight coupling and potential bottlenecks as the system scales. In contrast, Pub Sub allows for asynchronous communication where publishers and subscribers are decoupled. This decoupling can lead to more efficient resource utilization and better scalability, especially under heavy loads.

2. Message queues are faster than HTTP requests:

Pub Sub architectures often rely on message queues for communication between components. Message queues are optimized for fast, asynchronous message delivery and can handle high throughput with low latency. In contrast, micro-services communicating via HTTP requests may incur overhead due to protocol negotiation, connection setup, and teardown. As the number of users increases, the overhead associated with managing HTTP connections in micro-services architecture can lead to longer response times

- and lower RPS compared to the more lightweight message queue approach in Pub Sub architecture.
- 3. Scalability and Elasticity: Pub Sub architectures are often more scalable and elastic compared to traditional micro-services architectures. With Pub Sub, it's easier to add or remove subscribers without impacting publishers or other subscribers. This scalability allows Pub Sub systems to handle spikes in traffic more gracefully by distributing the workload across multiple instances or nodes.
- 4. Reduced Latency Through Parallel Processing: In Pub Sub architectures, messages can be processed in parallel by multiple subscribers, leading to reduced overall latency. This parallel processing capability enables Pub Sub systems to handle high throughput without sacrificing responsiveness. On the other hand, in micro-services architectures, synchronous communication may lead to sequential processing of requests, which can introduce latency, especially under heavy loads.
- 5. Fault Tolerance and Reliability: Pub Sub architectures often provide built-in fault tolerance and reliability features. Messages published to a topic are typically persisted and can be redelivered to subscribers in case of failures. This ensures that messages are not lost even if some components in the system fail. In contrast, in micro-services architectures, ensuring fault tolerance and reliability often requires additional complexity and infrastructure, such as implementing retries, circuit breakers, and distributed tracing.
- 6. Optimized Resource Utilization: Pub Sub architectures can optimize resource utilization by dynamically adjusting the number of instances or nodes based on workload demands. For example, Pub Sub systems can automatically scale up or down the number of subscribers based on message volume or processing requirements. This dynamic resource allocation allows Pub Sub architectures to efficiently utilize computing resources and minimize costs, especially in cloud environments where resources are billed based on usage.