Service-Based Architecture

March 14, 2022

Richard Thomas

Presented for the Software Architecture course
at the University of Queensland

THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

Service-Based Architecture

Software Architecture

March 14, 2022

Richard Thomas

1 Introduction

Service-based architecture is one of the simpler, but still flexible, distributed architectural styles. It provides good support for delivering quality attributes of modularlity, availability, deployability, and simplicity (in the context of a distributed system). The key characteristic of a service-based architecture is that it uses domain partitioning, and each domain becomes its own distributed service. This partitioning provides high-level modularisation that helps ensure the domain partitions are independent of each other. The distribution of domains means that multiple instances of a service can be made available through a load balancer, providing better availability and some level of scalability.

Many medium-sized bespoke systems[1] are built using a service-based architecture. For example, an on-line store might be partitioned into services such as ProductBrowsing, ProductPurchasing, ProductFulfilment, InventoryManagement, and CustomerAccountManagement. Each of these could be independent distributed services that use a shared database.

Definition 1. Service-based Architecture

The system is partitioned into business domains that are deployed as distributed services. Functionality is delivered through a user interface that interacts with the domain services.

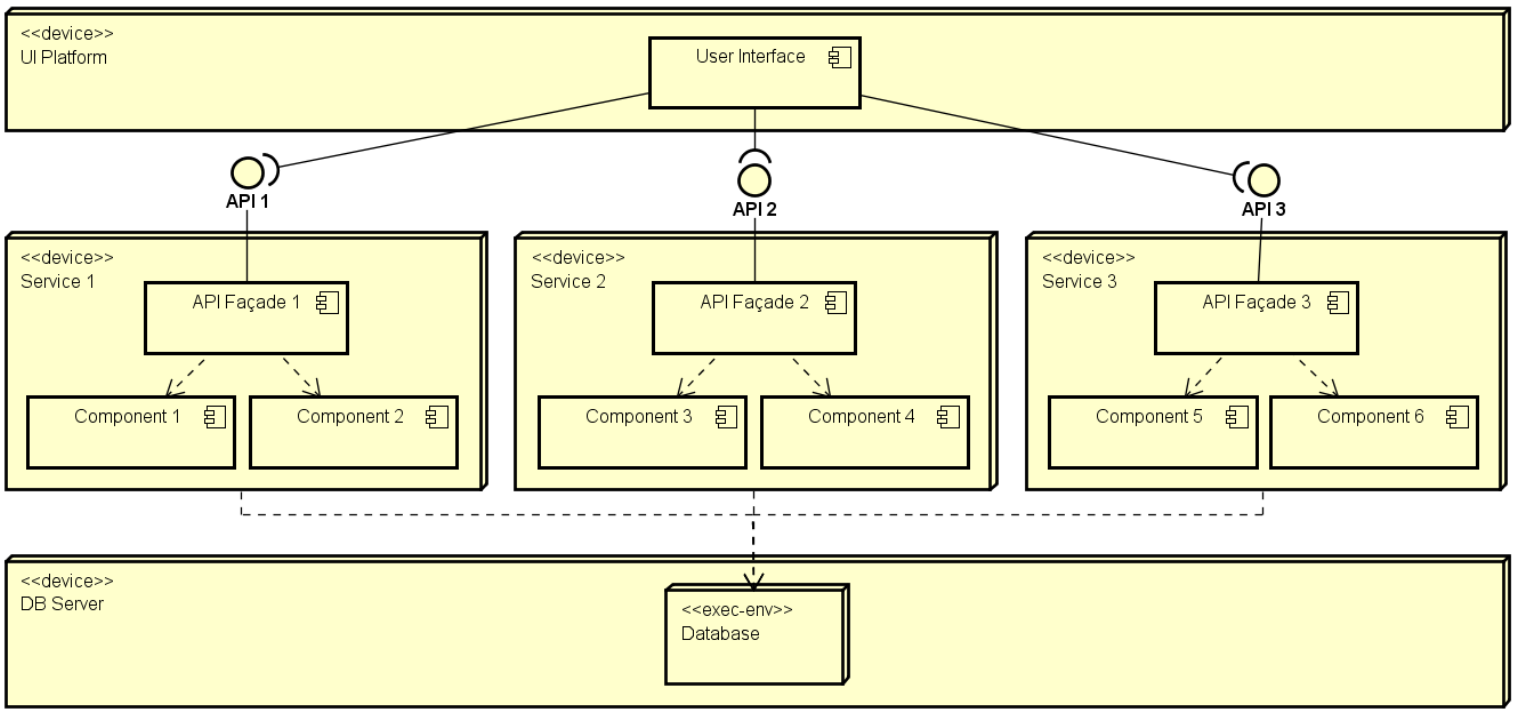

Figure 1: Service-based architecture – general structure.

[1] *Bespoke systems* are custom designed and built for a single organisation.

2 Terminology

The service-based architecture consists of four elements. The *user interface*, *services*, *service APIs*, and *database*, as shown in figure 1.

User Interface provides users access to the system functionality.

Services implement functionality for a single, independent business process.

Service APIs provide a communication mechanism between the user interface and each service.

Database stores the persistent data for the system.

The user interface runs as a standalone process to manage user interactions. It communicates with the services through their service APIs to invoke system behaviour. This requires a remote access communication protocol, such as REST, a message transport service, remote method invocation, SOAP[2] or some other protocol.

To reduce coupling between the user interface and the service APIs, and to provide easier extensibility, the user interface often uses a service locator design pattern[3] to manage access to the services[4]. This provides a registry of the services and the details of the API of each service. The user interface uses the registry to retrieve an object that communicates with a service through its API. This also makes it easier to add new services to the application, as the details of the new service are encapsulated in the registry.

Services implement the application logic for independent business processes. These are often called "coarse-grained" services, as each one implements a significant part of the system's functionality. Each service is deployed on its own computing infrastructure. Commonly, there is a single instance of each service but it is possible to run multiple instances of services. Multiple instances of services improves availability because if one instance goes down, other instances can handle future requests from the user interface. To provide higher reliability, in the context of running multiple instances of a service, it should implement the stateless service pattern[5]. A system running multiple instances of a service, that does not implement the stateless service pattern, would still have higher availability if a service instance went down, as other instances could handle future requests. But, any user in the middle of performing a business process would need to restart their activity, thus lowering system reliability.

Services implement their own service API using the façade design pattern[6]. This defines the communication protocol used between the user interface and the service. For simplicity, usually all services use the same communication protocol. The façade design pattern reduces coupling between the user interface and the services, as the user interface does not depend on the implementation details of the services.

The service API provides the benefit that different user interfaces can all use the same services. For example, an application with web and mobile interfaces could use the same set of distributed domain services.

The database stores persistent data for the system. Often, a single database is shared by all the services as it is common that some data will be shared between services. A shared database makes it easier to maintain data integrity and consistency. This is because each service implements a single business process and can usually perform all transaction management related to the data involved in the process. For example, the ProductPurchasing service for an on-line store can manage the entire database transaction for making an order. If the product is no longer available or payment fails, the service can rollback the transaction to ensure data integrity.

[2] https://www.w3schools.com/xml/xml_soap.asp

[3] https://www.baeldung.com/java-service-locator-pattern

[4] Martin Fowler provides good commentary about using the service locator pattern at https://martinfowler.com/articles/injection.html#UsingAServiceLocator. He expands further on some tradeoffs of the pattern than other more superficial descriptions of the pattern.

[5] https://www.oreilly.com/library/view/design-patterns-and/9781786463593/f47b37fc-6fc9-4f0b-8cd9-2f41cb36xhtml

[6] https://refactoring.guru/design-patterns/facade

3 Design Considerations

A service-based architecture is typically used for medium-sized systems. This is because the user interface interacts with all services through their APIs. The user interface becomes more complicated when it has to deal with many services. If the services follow the common approach of using a shared database, it means the the greater the number of services, the more complicated the database design becomes. There is also a potential performance bottleneck if many services are using a shared database. Strategies to improve database performance, like replicated databases, defeat some of the benefits of a shared database (e.g. consistency). Typically a service-based architecture will have six to twelve domain services. There is no specific upper or lower limit on the number of services allowed, it is a tradeoff that architects make based on all the requirements for the system.

Coarsed-grained services will usually have some internal complexity that requires some architectural structure. This internal structure may follow either technical or domain partitioning. Technical partitioning will typically consist of three layers, the API façade, business logic and persistence. Domain partitioning will break the service domain into smaller components related to each part of the domain's behaviour. For example, the ProductPurchasing service domain may have components for the internal behaviours of checking out, payment and inventory adjustment. Payment would use an API to process payment through a financial service gateway. Figure 2 provides an example of the structure for both technical and domain partitioning of a service.

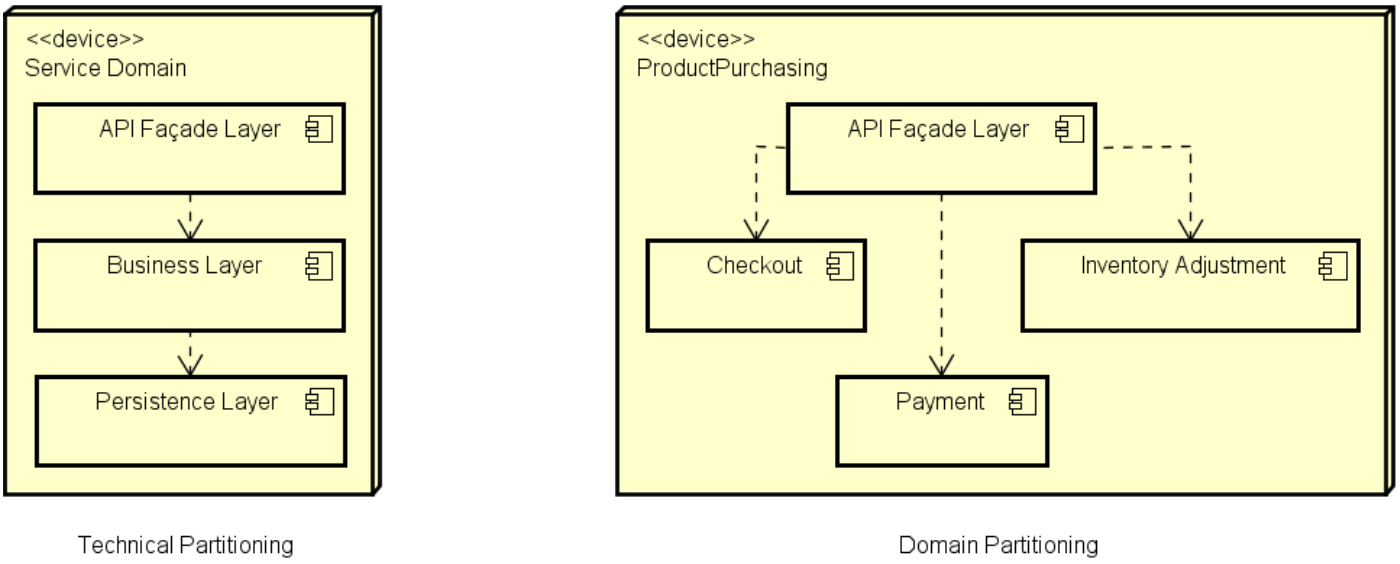

Figure 2: Partitioning options for a service domain.

Consequences of a shared database are increased data coupling between the services and lower testability. Increased data coupling means that if one service changes its persistent data, then all services that share that data need to be updated, as well as the tables storing the data in the database. Lower testability is the consequence of shared data and services implementing complete business processes. A small change to one part of the service requires the entire service to be tested, and all other services that share data with the service also need to be tested.

To mitigate data coupling, design a minimal set of shared (or *common*) persistent objects and their corresponding tables in the database. Implement a library containing the shared persistent classes that is used by all services. Restrict changes to the shared persistent classes and their database tables. Changes may only occur after consideration of the consequences to all services. A variation is to not only have shared persistent objects, but other persistent objects that are only shared with a subset of services.

Each service may have its own set of persistent objects and corresponding database tables. These are independent of other services, so there are no external consequences to changing these within a service. Figure 3 is an example of shared persistent objects and a service with its own persistent objects.

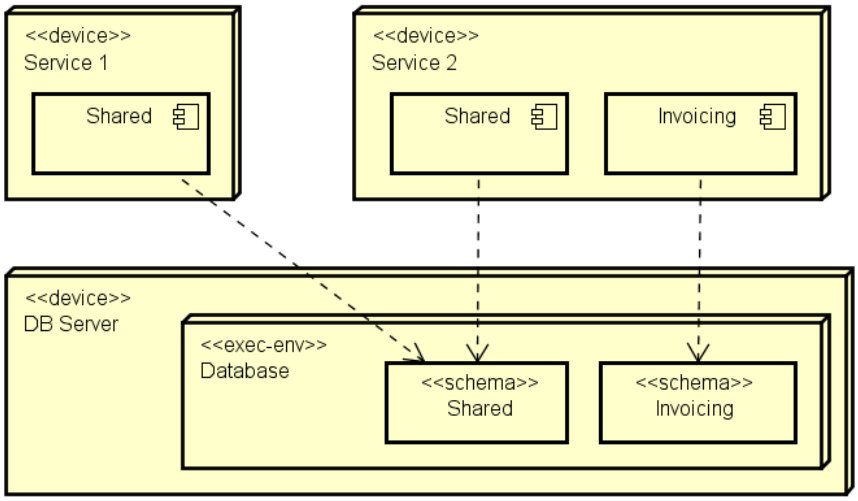

Figure 3: Database logical partitioning example.

4 Service-Based Principles

There are a couple of principles which should be maintained when designing a service-based architecture to produce a simple, maintainable, deployable and modular designs.

Definition 2. Independent Service Principle

Services should be independent, with no dependencies on other services.

Services should be independent of each other. If a service depends on other services they either cannot be deployed separately, or they require communication protocols between services, which increases the coupling and complexity of the system design.

Definition 3. API Abstraction Principle

Services should provide an API that hides implementation details

The user interface should not depend on implementation details of any services. Each service should publish an API that is a layer of abstraction between the service's implementation and the rest of the system. This provides an interface through which the service can be used and reduces coupling between the service and its users. In a service-based architecture, the user interface is the primary client of service APIs but it is not necessarily the only client. Auditing services may also need to use domain services. In more sophisticated environments, services may be shared across different systems.

5 Variations

TODO: Describe variations with unique UIs for domains and some shared across domains. Describe variation with independent databases for domains and a shared db.

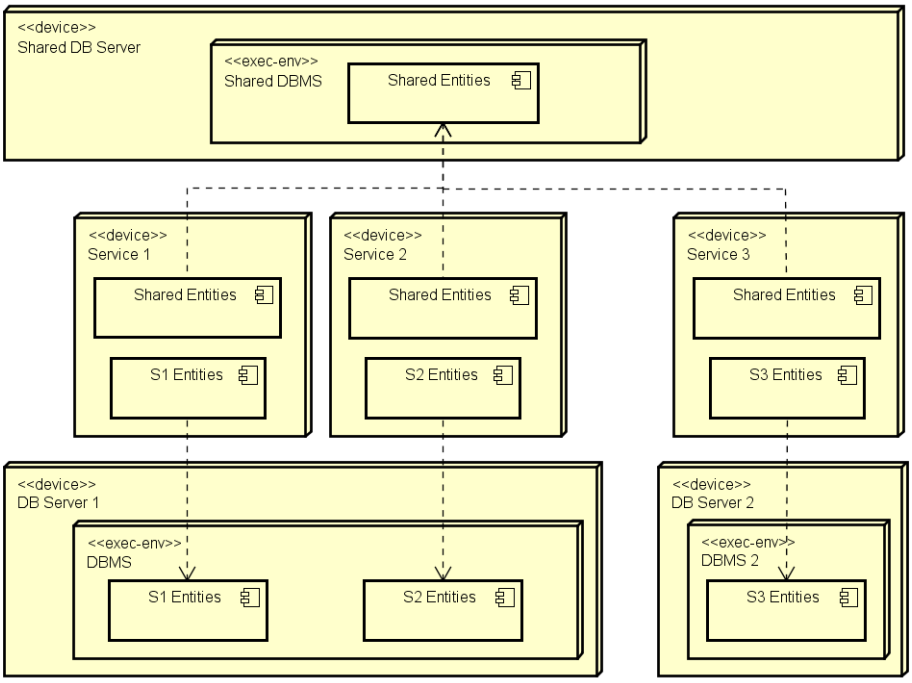

Figure 4: Separate databases example.

6 Extensions

TODO: Describe API layer between UI and domain services. Summarise the idea of it being a reverse proxy or gateway. https://www.l7defense.com/cyber-security/api-gateway-vs-reverse-proxy/ Summarise uses of this approach. Use the idea of service discovery to lead into an intro about SOA (and critique of its commonly poor implementations).

7 Conclusion