Software Architecture

Final Exam July 14, 2018

PRIVATE AND CONFIDENTIAL

1. **Closed book/notes.**
2. **No personal items including electronic devices (cell phones, computers, calculators, PDAs).**
3. **No additional papers are allowed. Sufficient blank paper is included in the exam packet.**
4. **Exams are copyrighted and may not be copied or transferred.**
5. **Restroom and other personal breaks are not permitted.**

BE VERY CAREFUL WITH THE GIVEN 2.5 HOURS AND USE YOUR TIME WISELY. THE ALLOTED TIME IS GIVEN FOR EVERY QUESTION.

Write your name and student id at the top of this page.

Write your answers clearly. If I cannot read it, you don't get points for it.

Question 1 [5 points] {10 minutes}

Explain how we can accomplish load balancing in Kafka?

If we have only one topic on one broker: by using partitions.
If we have a topic on multiple broker, we can also do load balancing over the different brokers

Question 2 [10 points] {15 minutes}

Suppose we have a small microservice architecture, and we implement logging with Logstash. Logstash is a service that collects the logging output from all services at a central point. In our current architecture all services send their logging to Logstash using HTTP REST.
This works very fine for the small number of services we have, but now we have to scale up to hundreds (or maybe thousands in a year from now) of services, and all these services write their logging to Logstash using HTTP REST. The characteristic for logging is that we want to log every important aspect in our system, and don't want to skip or lose logging information. Explain what kind of problem(s) we can expect with hundreds of services.

To many HTTP REST calls to the same logstash server. Load balancing the logstash server will not solve the problem. We need the logging in one place, not is multiple places.

Explain how we can solve these problem(s).

By publishing the loging into Kafka which is made for large streams of (log) data. Logstash can subscribe to these streams and handle the logs.

Question 3 [10 points] {15 minutes}

    a. You have to design a Microservice Architecture for a car rental system for a car rental company.
        Which domain/business services would you create:

Cars
Reservations
Rentals
Customers
Maintenance
Locations

    b. Suppose you are going to use REST for communication between services.
        Next to the domain/business services you also need supporting services/techniques that make sure your microservice architecture flexible, maintainable, secure, available, etc. Give the name of the supporting services/techniques that you need and for every supporting service/technique, write in one or two sentences what this service/technique does

API gateway	Loosely coupling between clients and services and crosscutting concerns
Config server	One central place to store config data for all services
Registry	Registring and lookups of service locations (URL,portnumber)
Authorization server	Token based security with OAuth
Resilience framework (Hystrix)	Implement resilience patterns like circuit broker and bulkheads
Tracing server (Zipkin)	Centralized tracing server
Sleuth	Add unique ID's to trace info
ELK stack	Centralized collection, searching and monitoring of logging data
Client side load balancing(Ribbon)	Load balancing between service instances

c. Suppose that instead of using REST, you are going to apply the Event Driven Architecture style to your microservice system, and you are going to use **Kafka for ALL communication between the business/domain services**. Which of the supporting services/techniques that you wrote down in part b would you still need in this architecture, and which supporting services/techniques are not necessary anymore in an EDA architecture. For every supporting services/techniques, explain in one or two sentences why you need it , or why you don't need it.

Still needed:

API gateway: we still have clients that need to access services
Config server: we still need centralized configuration
Resilience framework (Hystrix): we still want to know if we cannot access Kafka
ELK stack: we always need logging
Sleuth: for logging we need message id's
Registry: the API gateway still needs a registry

| |
|---|
| Tracing server (Zipkin): tracing is not that important with asynchronous publish-subscribe Only needed for API gateway |
| Client side load balancing(Ribbon) : not needed if we communicate with Kafka. Only needed for API Gateway |
| Authorization server: we do not use token based security with Kafka Only needed for API gateway |

Not necessary anymore:

Question 4 [15 points] {15 minutes}

Suppose in our Bookshop microservice architecture we have a ProductService and a SupplierService. The ProductService allows us to manage products. The SupplierService allows us to manage suppliers.
In production we find that the SupplierService calls take an exceedingly long time when doing a lookup of supplier information from the ProductService. We also see that supplier data rarely changes.
The developers decided to cache the supplier data in the ProductService.
If a supplier record changes via an update in the SupplierService, the SupplierService will call the ProductService to invalidate the cached supplier data, and retrieve the updated supplier data from the SupplierService.
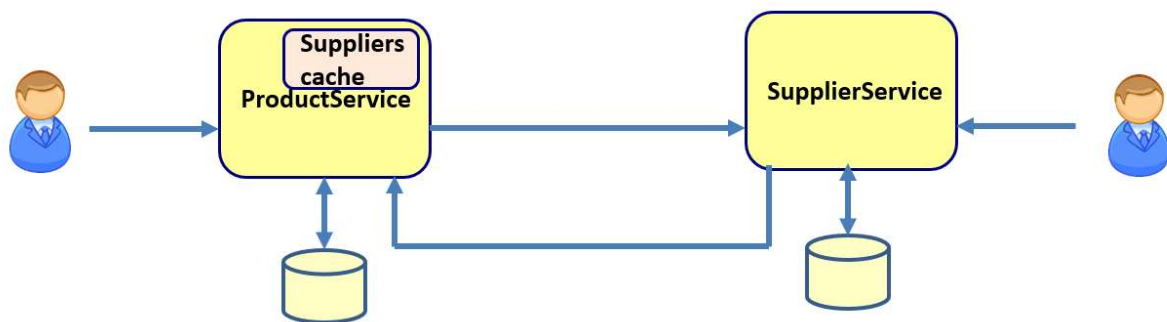
If a customer requests product data, the ProductService first checks its cache for supplier data. If the supplier data is not in the cache, the ProductService calls the SupplierService to retrieve it.
If an employee updates de data from a supplier, the SupplierService calls the ProductService to invalidate the cached record of this supplier.

Suppose you are hired as an software architecture expert, and you have to advice this company about the proposed solution as described above.

   a. What are the problems with this proposed solution? Explain in detail what the problems are.

Tight coupling between SupplierService and ProductService
If the ProductService is down or slow, the SupplierService is infected
If another Service also wants to know that data from a supplier changed, the SupplierService needs to change.
If we have multiple instances of the ProductService running, the SupplierService needs to call all instances to invalidate the cached record. If we start another instance of the ProductService, the SupplierService needs to know that.

b. How would you solve the problems described in part a?
   For this part, we want to keep the suppliers cache in the ProductService, so removing the suppliers cache is not an option.

Use a publish-subscribe message channel between the SupplierService and the ProductService.

If the ProductService is down or slow, the SupplierService is not infected
If another Service also wants to know that data from a supplier changed, they only need to subscribe
If we have multiple instances of the ProductService running, the SupplierService does not need to know that. All instances subscribe to the channel.

Question 5 [5 points] {5 minutes}

Explain how Kafka solves the synchronization problem that 2 services write to the same partition of the same topic at the same time.

Kafka uses event sourcing. Messages are immutable. Messages are only added, nothing is deleted or changed

Question 6 [15 points] {20 minutes}

Suppose we have 2 services A and B that both have a REST interface, and service A calls B. The user calls service A using postman.
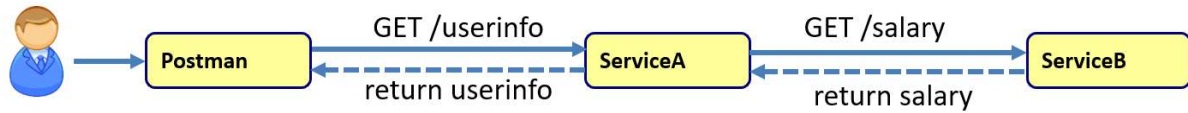

The user sends a GET request with the URL /userinfo (we ignore the remaining url) rom Postman to ServiceA. ServiceA then sends a GET request with url /salary to ServiceB. ServiceB returns the salary information, and ServiceA returns user information including the salary to Postman.

a. Now we are going to secure all our services with OAuth2. Draw in the given sequence diagram the flow of all calls between all parties involved. You can add more parties to the sequence diagram. In the sequence diagram show for every call the type of HTTP request (GET, POST, PUT, etc) and the url (like /userinfo). Also show the return of every call, and mention what is returned by this call..
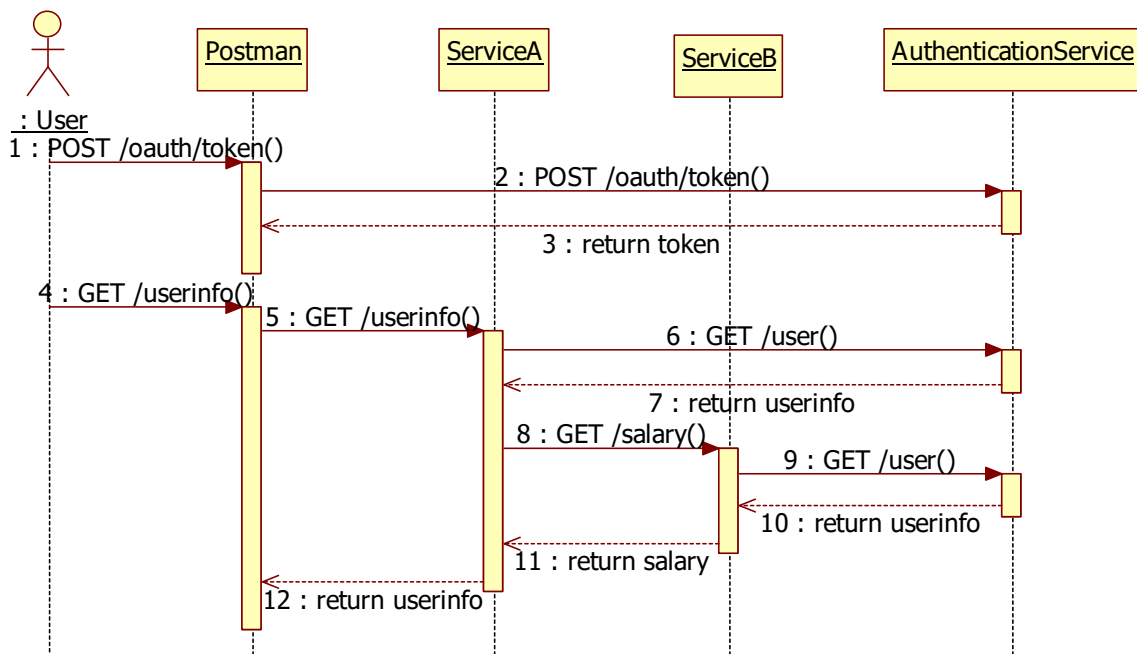
b. Now we are going to secure all our services with OAuth2 using JWT tokens. Draw in the given sequence diagram the flow of all calls between all parties involved. You can add more parties to the sequence diagram. In the sequence diagram show for every call the type of HTTP request (GET, POST, PUT, etc) and the url (like /userinfo). Also show the return of every call, and mention what is returned by this call.
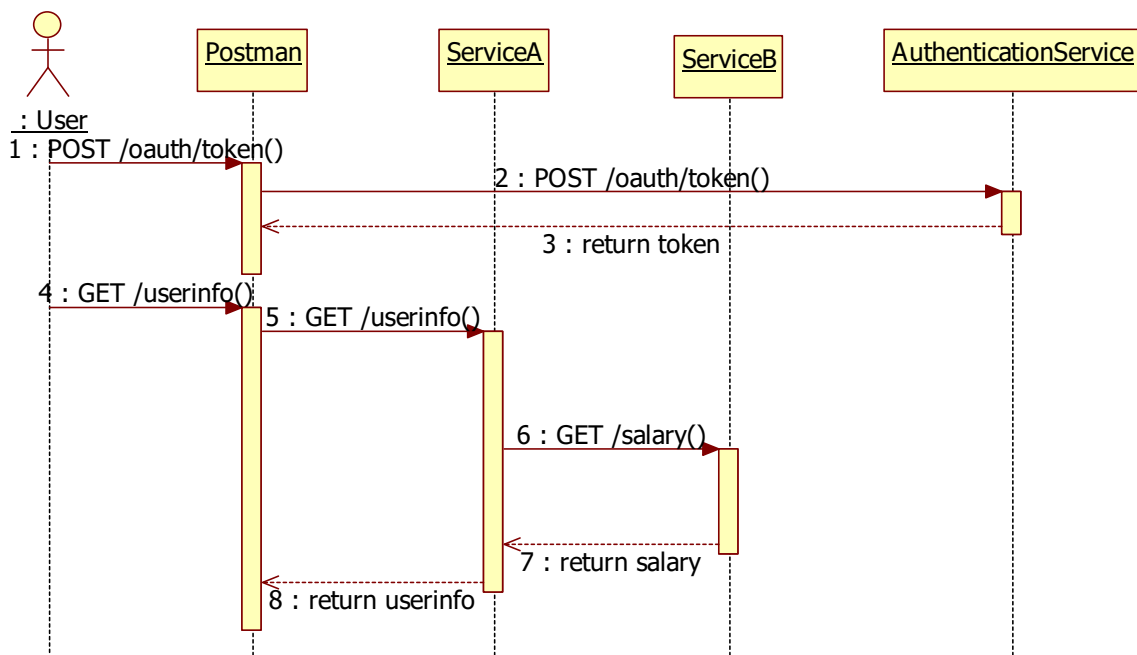
c. Explain what we need to configure in ServiceB to make this service secure. You don't need to write code. Just explain what you need to configure and give an example of this.

We need to configure which role can access which URL.

```
http
    .authorizeRequests()
    .antMatchers("/name").permitAll()
    .antMatchers("/salary").hasRole("MANAGER")
    .antMatchers("/phone").hasRole("USER")
    .anyRequest()
    .authenticated();
```

Question 7 [15 points] {20 minutes}

In the webshop we wrote in the labs, the services where calling each other using REST.
Suppose we want to change this webshop to an Event Driven Architecture. To make this
exercise short, we don't use the customer service.

Draw the sequence diagram of the following scenario:

```
// add products to the shoppingcart
restTemplate.postForLocation("http://localhost:8080/cart/cart/1/A33/
3", null);
restTemplate.postForLocation("http://localhost:8080/cart/cart/1/A34/
2", null);

//checkout the cart
restTemplate.postForLocation("http://localhost:8080/cart/cart/checko
ut/1", null);

//confirm the order and update the stock
restTemplate.postForLocation("http://localhost:8080/order/order/1",
null);
```
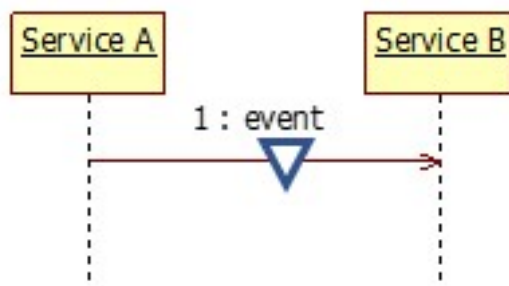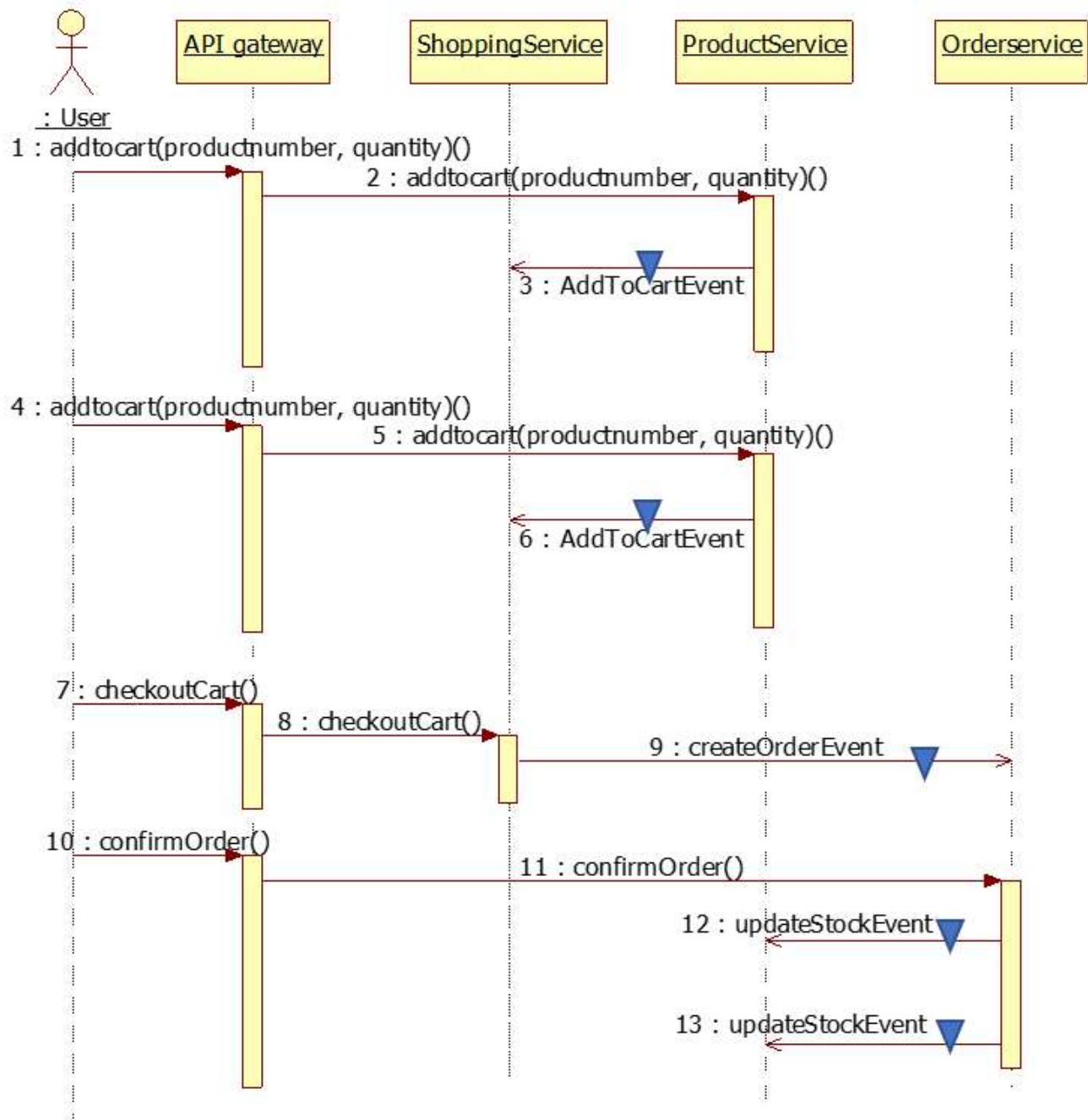
For synchronous calls, use this notation:
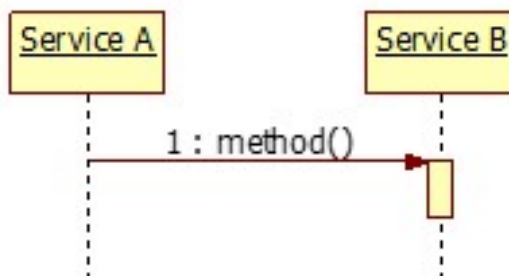
For asynchronous events, use this notation:

**Use a separate piece of paper for this sequence diagram**

Question 8 [20 points] {40 minutes}

Suppose you get a job at Tesla and you have to design the software for the next generation self-driving car. In the first week you learn everything about Tesla cars. You learn that every Tesla car has many sensors, and that there is an API (Application Programmable Interface) for steering, braking, light&signals, speed regulation and dashboard instrumentation. It is of course very important that the software is fast, reliable, available, fault-tolerant, and maintainable. Because this software will be used in many different car models, the software should be very flexible so we can plug and play with different modules.

You are asked to draw your design on **one piece of paper**, which should convince your boss that you the right person for this task. So draw and write as clearly as you can.

Draw on a separate piece of paper the design of your architecture. On your design, **write clearly the architectural styles, patterns and techniques you apply, and why you apply them**.

The number of points you get for this question depends on how well you explain your design on how well your design achieves the upfront requirements.
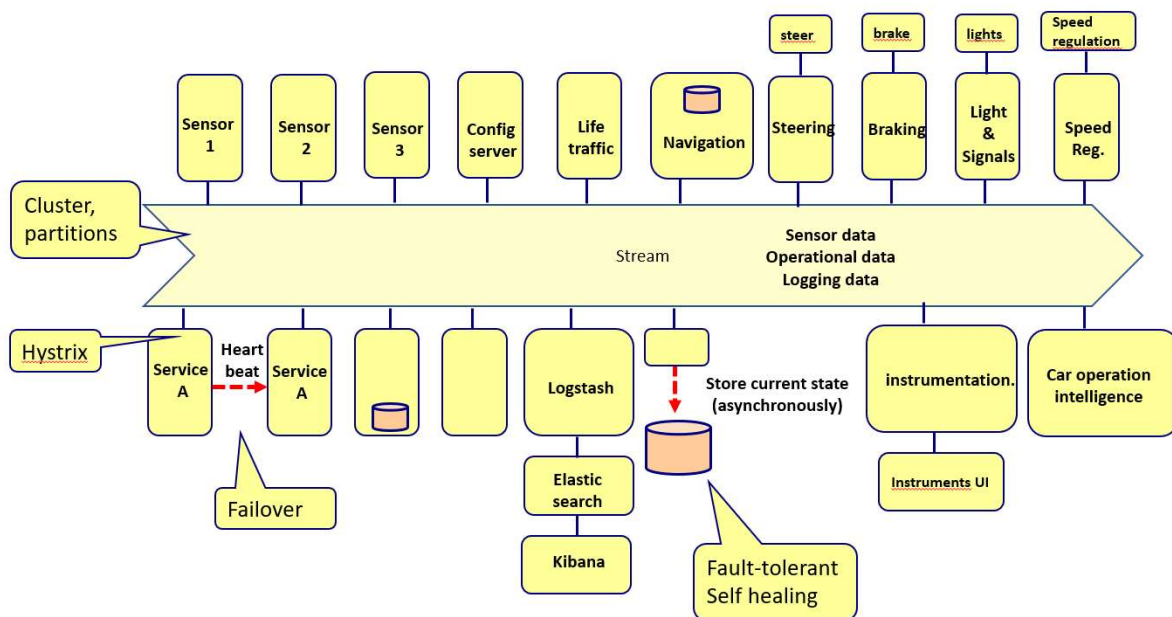
Question 9 [5 points] {10 minutes}

Describe how the possibilities of orchestration and choreography relates to one or more of the SCI principles you know. Your answer should be about half a page, but should not exceed one page (handwritten). The number of points you get for this question depends on how well you explain the relationship between orchestration/choreography and the principles of SCI.
Write clearly, I cannot give points if I cannot read it.

Your answer: