Remote Procedure Call

Stephan Djurhuus, Pernille Lørup

December 16, 2020

Abstract

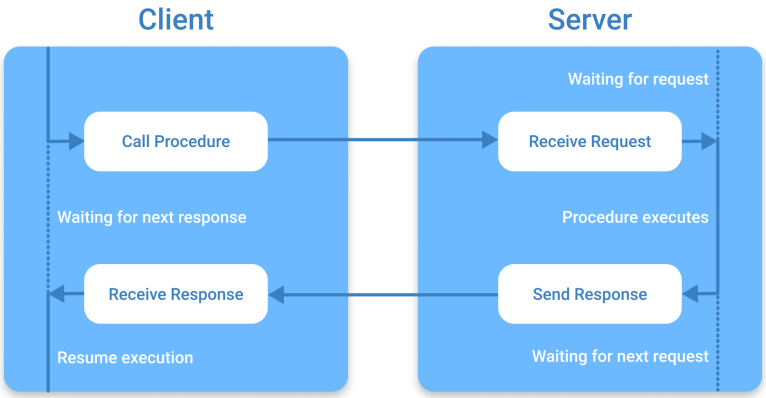
In large systems, optimal performance can be difficult to achieve because of limited resources. Most systems today demand high performance due to the increase of the desired amount of operations. A solution to this matter is the use of Remote Procedure Call which we investigate further in this article. Large systems with heavy operations can use Remote Procedure Calls to distribute workload which can lead to an increase of performance.

Introduction to Remote Procedure Call

Remote Procedure Call, also known as RPC, is a protocol which during runtime provides a connection between two or more systems. This protocol is a client-server model in which the client has knowledge of the server credentials. A connection between the two systems are obtained through a shared interface and the use of an RPC library. RPC offers the possibility to request a remote or local service while hiding the internal functionality from the user.

When making a remote procedure call it starts at the client environment where a request, containing parameters if necessary, is transferred through a network layer to the server. When the server receives the request, a method is executed in the remote environment that returns a response which is then transferred back to the client; this action is the procedure. This is visualized in *Figure 1*.

Figure 1: Procedure Flow



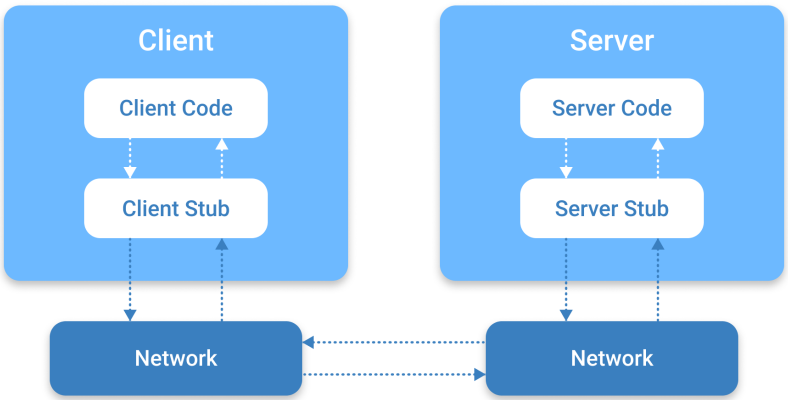
Functionalities

This section describes how the server and client communicates and how the general functionality of RPC works.

A contract between the two systems is established through the shared interface to give mutual constraints upon the requests and responses which is handled by a stub.[11]

A stub is the result when the shared interface is provided to the RPC library. After providing the interface, the stub contains the implemented methods which can be called in the application. Calling the stub methods transforms the parameters to a suitable format for transportation, which is then received at the other system as a request. This formatting is also referred to as marshaling.[10] When the request is received, the system unmarshals it to make it readable. Marshaling and unmarshaling happens between both systems every time data is transferred. The operational flow can be seen in *Figure 2*.

Figure 2: RPC Model



Remote procedure call can be synchronous or asynchronous depending on the implementation. The traditional implementation of RPC is synchronous which means the client system is suspended when making a request until a response is received.[8]

Advantages/ Disadvantages

Everything has advantages and disadvantages and RPC is not an exception. We will do our best to present some of the ones we found most important below.

Advantages

- **The possibility to develop two or more concurrent systems**
  By dividing the development into smaller systems, the possibility of distributing workload to multiple internal or external developers is obtained, which can result in an increased efficiency.[7]

- **The contract creates full transparency**
  The contract provides a specified description of all methods and necessary parameters in the common interface. This gives all developers an understanding of their constraints and limitations regarding the connection between the systems. These constraints give the possibility to obtain a better collaboration from external developers.

- **Possibility for distributed subsystems**
  When distributing the development into smaller subsystems, each subsystem has the ability to use the desired amount of resources available to it. This can optimize CPU and memory consumption by expanding it over multiple machines an thereby increase the performance significantly. [1][6]

- **Suitable for developing large systems**
  Because of the possibility of distributing multiple system features across different teams and servers, the use of RPC is convenient for developing large systems.[7]

- **Underlying procedures are hidden from client**
  The client is only provided with methods from the shared interface which gives the developer knowledge of what response the client receives. Because of this, the developer doesn't have to concern about how the server's implementation handled the request, which will provide more focus and avoid a bloated system.

- **Subsystems are more maintainable**
  Smaller systems are easier to develop and maintain due to the better overview of the system, instead of navigating through a large system and waste time understanding the surrounded functionality. A subsystem limits the amount of functionalities and thereby increases the level of understanding so more time can be spend on developing.

Disadvantages:

- **Response time is dependant on network bandwidth**
  If the server and client are located on different machines, an outgoing network connection needs to be established. The user's response time will be increased in case of a low network bandwidth on one of the systems.[7]

- **Distributed RPC systems are vulnerable to failure**
  These types of distributed systems are more vulnerable to failure in comparison to other systems where all implementation is gathered in one place. This is because the multiple distributed systems rely on each other in production and therefore if one of the systems where to disconnect, the others would be affected as well.[7]

- **Not a flexible connection**
  The interaction between the server and client is limited to the contract methods. Only methods with the required parameters can be used and nothing else.

- **It isn't easy to update the contract.**
  The contract defines how the two systems can interact with each other. If the contract needs to be modified all implementations need to be implemented accordingly. Depending on the degree of the modification a significant amount of code can end up being rewritten.

- **Integration testing of the system is limited during development**
  The uncertainty of proper integration between systems throughout the development is difficult to test, because different systems can be developed in separation.

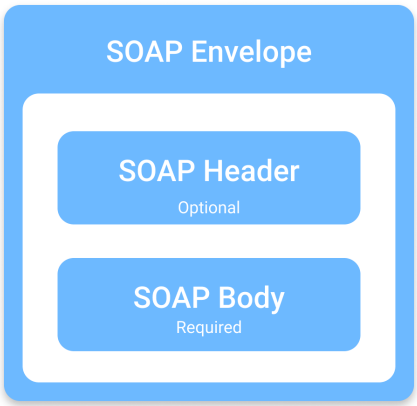
Alternatives

Connecting systems can be obtained through many technologies apart from RPC. In this section we will describe the usage and functionalities of some of the ones we know.

Simple Object Access Protocol

Simple Object Access Protocol also known as SOAP is a messaging protocol for exchanging data via XML. The message being exchanged is referred to as an envelope, which is mandatory because it indicates the beginning and the end of a message. The envelope consists of an optional header and an obligatory body as shown in *Figure 3*.

Figure 3: SOAP Envelope Structure



It is based upon schemas that define the constraints and structure of the request. When a request is received on the server, the schema validates if the format is correct. If the request is accepted, it will be processed. However if the envelope body doesn't match the schemas' requirements, the request will be rejected.[9]

Representational State Transfer

Representational State Transfer or REST is a broadly supported architectural style for client-server communication. REST sends information by using JSON or XML data structure via multiple predefined endpoints instead of sending information through methods. This gives the client a more flexible interaction to the server. However it is still constraint to the HTTP methods and headers together with the required body content and structure. REST doesn't use any interface which means the client needs to read the API documentation before using it.[2][3]

Graph Query Language

Graph Query Language also called GraphQL is a newer query language standard for APIs. This implementation uses the client-server model where the server exposes a single endpoint to the client. Strongly typed schemas are used to define the possible structures of a query that a client can request to the server. These schemas, also referred to as contracts, give the possibility to create advanced queries, consisting of the desired response data structure. Queries like these will be executed as a request to the server which sends a response only containing the requested data. One of the advantages of using these queries is to prevent under- and over-fetching which often results in faster response time and reduced data consumption. However if the query is too complex it can end up increasing the response time.[4]

Our Implementations

As we prior to writing this article have had experience working with Remote Procedure Calls, we will shortly describe previous assignments and how we chose to solve them. In continuation of this, we will reflect on one of the newer RPC frameworks as a possibility to use in future work.

Project References

Our first assignment related to RPC was in the System Integration course where the main focus of the assignment was to get acquainted with and illustrate the use of RPC. We solved the assignment by making use of an RPC library for Node.js together with a strongly typed programming language, which we in our case decided would be TypeScript.

After getting acquainted with RPC we chose to use it as the foundation in our project from the Large System Development course. This new project gave us a broader understanding of the protocol and how to optimize the implementation. One of the key features when using RPC is the possibility to create two concurrent systems that work independently until production. Because the client and server share an interface we can implement a fake on the client-side to be used during development. The fake substitutes the server with hard-coded implementations to test the client system's functionality.

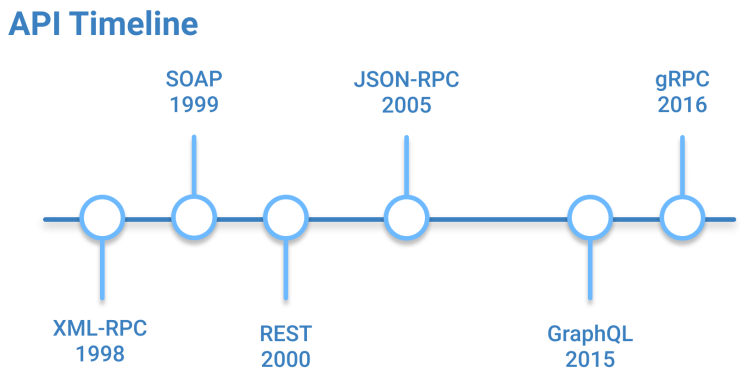
Because the criteria of the Large System Development assignment was to build two systems from a contract, we decided to use the RPC protocol that is build around the common interface which in our case was the contract.

Potential future work

An alternative to RPC we didn't have knowledge of before writing this article is the newly developed framework made by Google called gRPC.
One of the benefits that comes with gRPC is that the framework is language agnostic which means the systems can be written in multiple languages and still communicate with each other. The framework also provides load balancing which is favorable as the system automatically distributes workload.[5] This

framework would be interesting to look more into if we were to use RPC in future work.

Figure 4: API Timeline



In *Figure 4*, a timeline is presented with some of the previously mentioned API options along with their creation dates to clarify that Remote Procedure Call has been in the game for a long time and is still today a strong contestant.

Analysis of Response Times

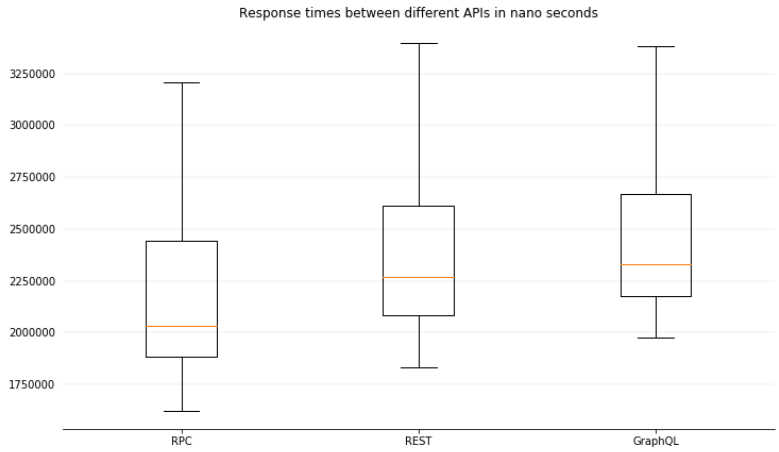
We created a simple Hello World example for RPC, REST and GraphQL to compare the different response times on a local machine.[12] The results are based upon 500 iterations of 'Hello World' requests that responds with a simple string. The analysis resulted in RPC being the best performer, which is presented in *Table 1*.

Table 1: Mean & Standard Deviation

	Mean	**Standard Deviation**
RPC	2.334ms	1.190ms
REST	2.529ms	1.576ms
GraphQL	2.636ms	2.520ms

We have plotted the results in a box-plot featured in *Figure 5*, to show the descriptive statistics calculated through the observed observations.

Figure 5: Response Times



We can see that RPC has the highest performance for this type of test with about 0.2 millisecond faster than REST. This test could be optimized by creating a more complex request where the server and client would be located on separate machines.

GraphQL has the lowest performance which might be an unfair statement due to the limitation of the test. Because RPC and REST are limited to the predefined responses, multiple requests may occur. If the test handles these types of scenarios, GraphQL would probably be a better performer.

Specifications

The code performing the analysis was executed on a MacBook Pro (Retina, 13-inch, Early 2015) with a 2,9 GHz Dual-Core Intel Core i5 Processor. The MacBook has a memory of 8 GB 1867 MHz DDR3.
The code was executed with Node.js version 12.16.2 and Python version 3.7.4.

Conclusion

Remote Procedure Call is a powerful protocol with many advantages. Working with it in multiple projects along with research and our previous knowledge of its alternatives in SOAP, REST and GraphQL, we have come to learn that RPC is a protocol well suited for developing large systems. If the contract is correctly specified between the systems, it creates a strong base for development.

A perfect tool for everything doesn't exist. The compromise you give by developing a distributed RPC system is your response time, which can be reduced while your operational performance can be increased. A system with heavy operations would prefer to use a distributed system like RPC, while a fast responding system wouldn't.

References

[1] Open Cirrus. What is distributed computing, its pros and cons? https://opencirrus.org/distributed-computing-pros-cons/.

[2] Todorka Stoyanova Dimitrova. System integration. https://datsoftlyngby.github.io/soft2020fall/resources/0f9c7c38-Session4RESTplus.pdf.

[3] Roy Thomas Fielding. Architectural styles and the design of network-based software architectures - chapter 5 representational state transfer (rest). https://www.tutorialspoint.com/soap/soap_envelope.htm.

[4] GraphQL. Learn graphql. https://graphql.org/learn.

[5] gRPC Authors 2020. About grpc. https://grpc.io/about/.

[6] Guru99. Remote procedure call (rpc) – remote procedure call (rpc) protocol in operating system. https://www.guru99.com/remote-procedure-call-rpc.html/.

[7] IONOS. Remote procedure call (rpc) – efficient communication in client-server architectures. https://www.ionos.com/digitalguide/server/know-how/what-is-a-remote-procedure-call/.

[8] Microsoft. Asynchronous rpc. https://docs.microsoft.com/en-us/windows/win32/rpc/asynchronous-rpc.

[9] Tutorials Point. Soap - envelope. https://www.tutorialspoint.com/soap/soap_envelope.htm.

[10] Margaret Rouse. Marshalling. https://whatis.techtarget.com/definition/marshalling?_ga=2.99920477.2106886235.1608112504-176318560.1601985231.

[11] Margaret Rouse. Remote procedure call (rpc). https://searchapparchitecture.techtarget.com/definition/Remote-Procedure-Call-RPC.

[12] Pernille Lørup Stephan Djurhuus. Github repository to benchmark source code. https://github.com/pernillelorup/UFO-Exam-Article/tree/main/apis.