Group 6

Software Architecture 22-23

Pablo Argallero Fernández UO283216

Adrián Vega Sanchez UO282365

Sara Fernández Arias UO269546

Introduction to topics

In this presentation we are talking about the 9 fallacies in distributed computing that were first introduced (first 4) and then were expanded by L.Peter Deutsh (4 more fallacies) and that at the end the general public ended up adding a ninth one.

What is distributed computing?

Distributed computing refers to the method of computing or solving a problem by means of the use of more than one computer in a network. These computers can be closely geolocated or be placed in two completely different locations on the world but end user will interact with this systems as if they were a single computer.

The purpose of distributed computing is to increase the computing power and efficiency of a system by dividing the workload among multiple machines. By doing so, it can handle larger and more complex tasks that a single machine cannot accomplish on its own.

Now we know the definition of distributed computing we can glance the main concept on each of the fallacies introduced in the podcast.

Fallacy 1 – "The network is reliable."

It refers to the assumption that the network that connects the different computers or elements of the network is reliable in the sense that no error or delay will occur while communication is being performed.

In the podcast L. Peter Deutsch argues that this issue had more weight on the early days of distributed computing when network technologies were not that advanced. Nowadays technology stack on low levels of the network layers hand packet loss or data corruption in a way application layer is more likely to not care on network failures. On the other hand a problem that is not yet solved due to the nature of the networks is the different outages that can occur while a communication is performed. L. Peter WDeutsch argues that application must be prepared for this kind of network problems and give users information on what is going on.

Fallacy 2 – "Latency is zero."

This fallacy refers to the common thought that network latency is not perceivable by users. This misconception on network speed had a huge impact on the first worldwide distributed network applications. At the beginning of distributed networks all connections were performed in a local environment (universities / governments) which led to that 'zero latency' fallacy most of the developers now believe.

To mitigate this the following recommendations should be taken into account:

1. Acknowledge the reality of network latency: Network communication between distributed components does take time, even if it is just a few milliseconds. Acknowledging this reality is the first step in mitigating the fallacy.
2. Design for asynchronous communication: Instead of assuming that components can communicate with each other synchronously and in real-time, design your system to use asynchronous communication where possible. This allows components to communicate with each other without blocking, which can help mitigate the impact of network latency.

3. Use caching: Use caching to reduce the need for components to communicate with each other unnecessarily. Caching can help reduce the amount of network traffic and can also help mitigate the impact of network latency by allowing components to access frequently-used data without having to wait for network communication.
4. Optimize network communication: Optimize network communication by reducing the amount of data being transmitted, using compression where possible, and using efficient communication protocols. This can help reduce the impact of network latency and improve the overall performance of your distributed system.
5. Use distributed tracing and monitoring: Use distributed tracing and monitoring tools to gain visibility into the performance of your distributed system. These tools can help you identify bottlenecks and other issues related to network latency, allowing you to take proactive steps to mitigate the impact of latency on your system.

Fallacy 3 – "Bandwidth is infinite."

This fallacy is quite auto descriptive since it refers to the believing that network bandwidth can handle any kind of payload that it is given. This was a serious issue when L.Peter Deutsh first introduced these fallacies since bandwidth on that time was a serious issue. Nowadays bandwidth in networks is not a huge issue but is a possible issue that must be taken into account when designing the infrastructure of the distributed system.

Some techniques to mitigate this is to use balanced payload sizes that can be handled and also to monitor network traffic to identify possible bottlenecks in the system.

Fallacy 4 – "The network is secure."

This fallacy is one of the most known by people. The network is not secure, and in fact, nowadays is less secure in certain ways.

We have much more users since the access to internet is more achievable than 20 years ago. That makes that the people who might be interested in attacking certain network is greater than before. There is also a relation between the value of the information and the threats, greater the value greater the threats and its level of sophistication. Fortunately, people are much more concerned about security than before, and some changes such as HTTPS and hop to hop encryption have been developed, and threats are not that much about the network itself, but to the people, such as phishing attacks.

Fallacy 5 – "Topology doesn`t change."

Topology is the physical and logical arrangement of nodes and connection. This fallacy is not that much considered. Nowadays, the network accesses are mostly done from mobile devices, such as laptops and cellphones, making this fallacy have greater importance than it had before.

For example the connection of a user in a train is not always from the same place, since it is moving from a tower to another, so this is extremely important to take this into account. He also mentions that, when he was working in Sun there where some people that had problems when a client changed from an office to a new one, because the addresses changed, and now the problem is the same as the mentioned but increased. The solution is rather complicated, and more related to an architectural problem, so the thing is that your system must have to be able to maintain the requirements when the topology does change.

Fallacy 6 – "There is one administrator."

Multiple administrators, as subnets for rival companies, may have conflicting policies that avoids them to complete their desired paths.

In small projects you may achieve this "one administrator" in some degree, but in big projects it's near impossible. The solution suggested is having good standards, and people implementing them. This is usually not possible because companies want you to use only their product, making it unique and not cooperating with the rest. In fact, he claims that the internet has become this big because of good defined standards. The thing that must be achieved is administration based on standards, and open standards are more important than open source.

Fallacy 7 – "The transport cost is zero."

Some of the costs associated with any computer network are very clear to see: The servers, the cabling, the racks, the load balancers, power equipment…

Nowadays, with cloud computing, we delegate those costs to cloud providers we hire. But still, those costs are there and there is nothing we can do about them.

What we can manage are the costs associated to the usage of the network: Bandwidth, load balancing…These providers have used successful marketing strategies to make everyone think their services are just a few pennies an hour. However, if you sit down and think about how many hours of services you will need per month, and how many nodes they will require…. Prices might become hundreds of dollars per hour.

Fallacy 8 – "The network is homogeneous."

What Deutsch wanted to address is that, if standards are not used in an appropriate and consistent way, it would be like having no standards at all (or even worse).

At this point is natural to ask yourself "if we have standards, why does this fallacy even exist Then?" The reason is very well known to almost anyone in touch with web development: there are lots of standards to choose from. So sometimes it is hard to tell which standard is actually "the one" we should be using. For example: JSON or XML? While JSON is indeed considered the standard nowadays, many services functions still with xml.

Another question might arise now: how is any of this related to networking? Well, essentially in the way that when we are working in heterogeneous environments, the semantic model might change from one to another. That means, not being able to make an automated conversion of data since one of the formats you might be working with will probably contain information that the other one lacks.

It's fair to mention, though, that, while standards can be quite overwhelming nowadays, back then, the main reason was the lack of standardization. Still, even if the situation is not perfect now, L. Peter Deutsch states that this issue has improved a lot.

Fallacy 9 – "We all trust each other."

This fallacy wasn't originally on the list, but L. Peter Deutsch believes it does belong.

Authored by James Gosling, and it's straightforward. It states that we should not assume that the party we are communicating with is trustworthy.

Nowadays, there are two main security problems we do not have so much control over yet: phishing and malware. Especially nowadays, where some of the parties we could work with might perform malicious actions, sometimes even without the intention to do so. So, we should always interact assuming the other party is potentially dangerous.

Some of the costs associated with any computer network are very clear to see: The servers, the cabling, the racks, the load balancers, power equipment…

References

IBM. (n.d.). *What is distributed computing*. Retrieved from https://www.ibm.com/docs/en/txseries/8.2?topic=overview-what-is-distributed-computing

Particular Software. (n.d.). *Fallacy 1: Network is reliable*. Retrieved from https://www.youtube.com/watch?v=8fRzZtJ_SLk&ab_channel=ParticularSoftware

Software Engineering Radio. (n.d.). *L.Peter Deutsch*. Retrieved from Episode-470:Fallacies of Distributed computing.: https://www.se-radio.net/2021/07/episode-470-l-peter-deutsch-on-the-fallacies-of-distributed-computing/