

Scientific Study of CAP Theorem and Understanding its Different Implementation Methods

Prof. Barry Wiling
Professor, Department of Computer Science,
U.S.M.N Oman
barrywiling@usmn.edu

Article Info

Page Number: 133 – 137

Publication Issue:

Vol 71 No. 1 (2022)

Article History

Article Received: 18 November 2021

Revised: 03 December 2021

Accepted: 19 December 2021

Publication: 28 January 2022

Abstract

In this paper the CAP theorem will be discussed. The paper will explain that in any kind of system, only any two properties of the CAP can be implemented. In real world scenarios, in a database or any other system implementing CAP theorem, any one or two characteristics will be implemented. The paper will explain each component of CAP theorem and will see what are the challenges in each component. The paper will explain why CAP is important and why not all the components can be implemented together in one system.

Keywords: - CAP, Consistency, Availability, Partition, Importance, Challenges.

Introduction: - CAP is consistency, availability and partition tolerance features which are expected from a system. The CAP theorem suggests that only two of these properties can be expected from the system. Either CA, AP or CP can be available in a system at a time. In real world, the developers find it difficult to implement any one of these properties. CAP theorem has its own set of challenges. One has to be very careful and should have sound technical knowledge while the implementation of the theorem. Following are the components of CAP theorem: -

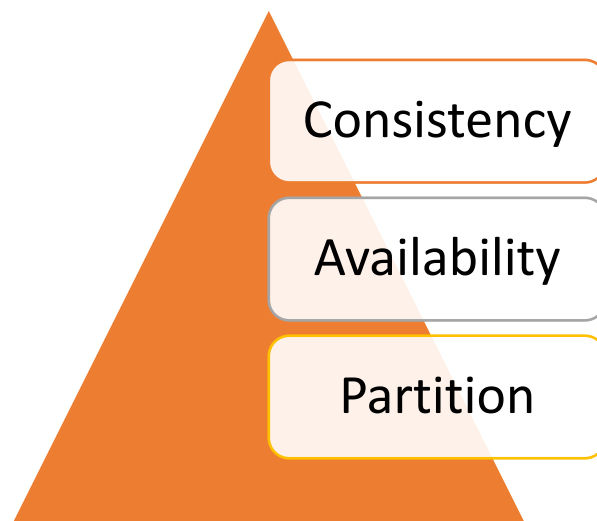


Figure 1 Components of CAP theorem.

- Consistency: - Consistency means that during the exchange or access of data, all the users requesting for the data should have same result on their screens. In a network many devices which acts as clients may be connected to the network. If the client requests something from the server, then all the clients connected to any node of the network may receive same response.
- This can be understood by taking example of a banking site. If the user is logging in the banking sites in his account from a laptop or a phone, then he should receive same result in both the devices which are laptop as well as phone.
- Some of the database requires update so the consistency of such that should be in such a way that every time the user requests the data, the most recent updated data should be received by all the clients in the network.
- Following are the types of various consistency in distributed systems: - [1]

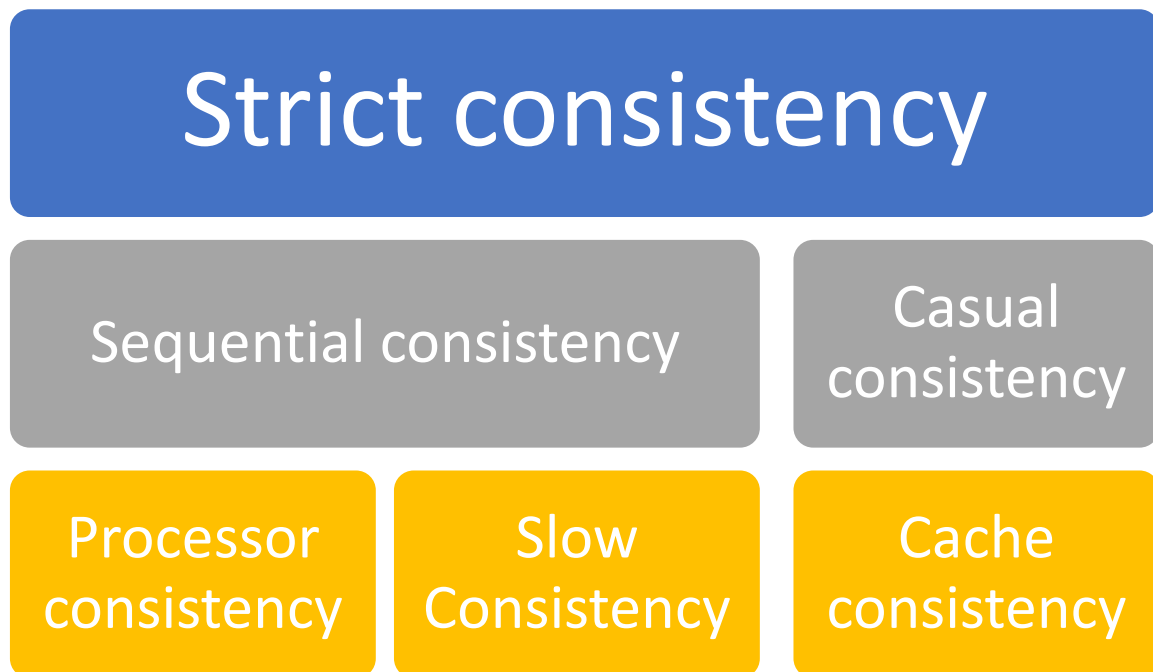


Figure 2. Various types of consistency.

1. Strict Consistency: -This is the strictest model where if anything is written on a processor then all the processor should read the same data. It is realistic model as every time the user requesting the data will be replied with the same data in all nodes. The exchange of message or data at once is not possible in this model.
2. Sequential Consistency Model: -This is the model which is majorly used by the CAP theorem. This type of model is less strong than the strict consistency model. In this model, there is no need of instant response on all the processors. What is important in this type of consistency model is that the sequence with which the operations are carried out should be reflected in the same sequence in all the processors at the same time only then the consistency will be maintained. The message should appear instantly on all the processors in the same sequence by which the operations takes place. It is not necessary that the whole message is written instantly but what is important is that as soon as the operation is executed, the message should be written in all the processors instantly.

3. Casual Consistency: - In this type of model, the events are divided in to two parts events which are casually related or which are not casually related. The events which are casually related are instantly written in the same sequence with which it appears in all the processors.
 4. Processor Consistency: - In order to maintain the consistency, processor consistency model was derived. It is necessary that the processors should be consistent by which it sees the write in itself and by other processors at the same location. It is not necessary that the processor should see all the writes in all the processors at different location.
 5. Slow Consistency: - It is weaker consistency model where the seeing of writes is slow which means that if an event reads a value from a location which was previously written then it cannot read any other values from that particular location.
 6. Cache Consistency: -Cache consistency means that the read of the write done at one location should be done sequentially. This is also a weaker consistency.
- Availability of CAP: - This feature means that every request should receive the reply which should be non-error response. This means the sources should be available whenever the request is made. For every query, there should be response from the sever which means that data is available. For example, if a user has login to his facebook account through his phone or laptop to check likes on his comment, then he should be able to see it through both the devices. It might be possible that the number of likes which he can see through his phone is less or more than the number of likes through his laptop. But one thing is sure that the data is available on both the devices whenever a request is made.
- Partition Tolerant: - In a distributed system, many clients and servers are connected to the network. The client sends requests and the server in the distributed network which has the information about that particular request will respond with the reply to that client. The P of the CAP theorem suggests that if one node of the network is down in a distributed system, still the other nodes should continuously respond and should stop responding to the queries. This will help to maintain the consistency of the system regardless of if any node of the network is down and the system can continue to perform its tasks.

The theorem suggests that only one or combination of above- mentioned components of CAP can be implemented in the system. The developers find it very difficult to implement one component so due to this it is very difficult to implement two components of CAP in real time systems.

Combination of components of CAP theorem: -

There are three possible combinations of CAP theorem which are: -

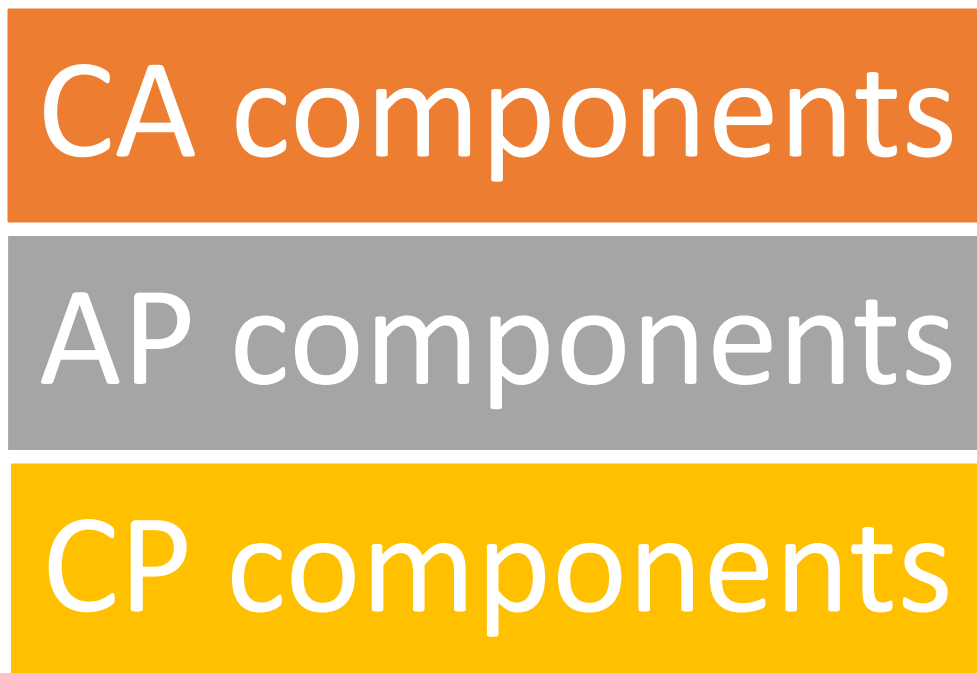


Figure 3. Various combinations of CAP Theorem.

a. Implementing CA component of CAP Theorem: - [2]

This type of implementation will help in providing the consistency and availability of CAP theorem together in one database. Hence the data provided by this type of implementation will be consistent as well as always available. The example of such data base will be MySQL. There is no duplication of this server and only single node is available in this type of server. The data will always be consistent and available if the server is up and running always. But if the server is down then neither the data will be consistent nor it will be available for the users. This is the main challenge of this type of implementation that either it can be consistent in providing the data or the data can only be made available. The other drawback of this type of data base is that it cannot handle error or fault which might arise during the request and response of the query.

b. AP implementation of CAP theorem: -

This implementation facilitates the availability as well as partition tolerance in the network. The data will be available through and also in case of one node is down, the data can be available from other nodes in the network. No matter what the nodes will respond but it is not guaranteed that the correct data is being responded. This means all the nodes are not updated with the latest data. Hence, it is not feasible to implement these two properties together in one server. The data will not be consistent. Although it is available always but it is not the updated version of the last written data. Example of this type of implementation is Cassandra.

c. CP implementation of the CAP theorem: -

This implementation is used to implement the consistency and partition tolerance in the network. The data which is accessed on request will be consistent in nature and also it can handle the partition in the distributed network. In this type of architecture, there is one primary node where the data will be written and the other nodes will be secondary nodes. The secondary node will copy

the data from the primary node. If in case of partition and the primary node is not working and is down then the data can be read from the secondary node.

Challenges of CAP theorem: -

- The major challenge of CAP theorem is that, it says that either of the two components of the CAP can be implemented but in reality, it is not possible.
- The technicians and developers find it very difficult to implement one feature properly and it requires high knowledge and skills to achieve efficiency.
- In many systems there are consistency and availability issues. It is hard to implement either consistency and availability of the data.
- There is no guarantee that the system will behave as per the requirements. It is difficult to achieve accuracy while implementing and monitoring the consistency and availability of the data.
- This theorem is now considered out dated as it is impossible to achieve accuracy and efficiency while implementing it.

Conclusion: - Hence it is concluded that the components of CAP theorem cannot be implemented in real scenarios. It will require sound technical skills if a person wants to implement such system. The data can either be consistent and available, available and partition tolerance or consistent and partition tolerance. There are many challenges involved in the implementation of the CAP theorem in real world scenarios. It has become outdated and there are other methodologies to improve the consistency etc of the data.

References: -

- [1]. https://en.wikipedia.org/wiki/Consistency_model#Strict_consistency
- [2]. <https://www.rga.com/futurevision/pov/the-cap-theorem-in-real-life>
- [3]. AGYEI , I. T. . (2021). Simulating HRM Technology Operations in Contemporary Retailing . International Journal of New Practices in Management and Engineering, 10(02), 10–14. <https://doi.org/10.17762/ijnpme.v10i02.132>
- [4]. Kumar, P. P. . “Creating One Time Password Programming for the Security of Data for Remote Location Treatment in Body Area Network”. International Journal on Recent and Innovation Trends in Computing and Communication, vol. 9, no. 10, Oct. 2021, pp. 01-02, doi:10.17762/ijritcc.v9i10.5497.