Career Episode 3

Design of an Enhanced Cloud Storage Solution

A) **Introduction**

[CE 3.1] The following career episode is based on my work experience for Sugon Information Industry Co., Ltd. (hereinafter referred to as the "Sugon") in Beijing China. The duration of the career episode is from 9th of October 2017 to 25th of January 2018.

> Project Title: Design of an Enhanced Cloud Storage Solution
> Duration: 9th October 2017 – 25th January 2018
> Location: Beijing, China
> Organization: Sugon Information Industry Co., Ltd.
> Position: Senior Engineer

B) **Background**

[CE 3.2] Sugon is one of the vendors which provides cloud storage solution in China. The company competes with other well-known vendors (such as Dell EMC, Huawei, and Hewlett Packard Enterprise) in the China market. In a cloud storage solution, new machines can be added and configured as the need arises. When the capacity reaches its limit, another storage machine can be added to expand the system capacity. It can employ the extra storage added across machines and use attached devices to increase storage capacity, enhance performance and address the requirement for additional storage. A classic case of cloud storage solution is shown below.
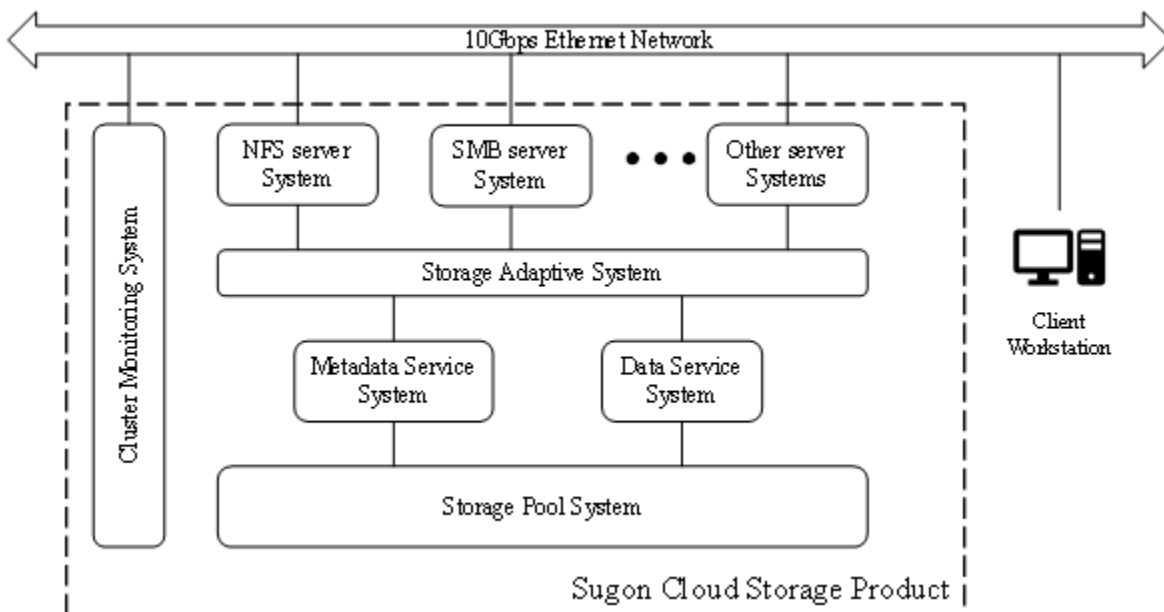
[CE 3.3] The overall aim of the project was to bridge the gaps with the designing of an enhanced cloud storage solution. I split the objectives into a more detailed set of aspects provided as follows:
- To identify the gaps and thus to define the dimensions of the project.
- To analyze and research the situation and specific requirement of the customers.
- To define and deliver an enhanced cloud storage solution to bridge the gap and meet customer needs.

[CE 3.4] The nature of my work area involved two aspects. One aspect concerned project management. It was mainly referred to as defining the dimensions of the project and directing team members to complete the project. The other aspect was as a technical leader in the project team. It began by analyzing customer requirements and then designing, developing, and testing the system to determine whether those needs were met. I applied mathematics, computer science, advanced technical knowledge and skills to the activities during the project.

[CE 3.5] The following chart illustrates the organizational structure highlighting my position during the career episode.
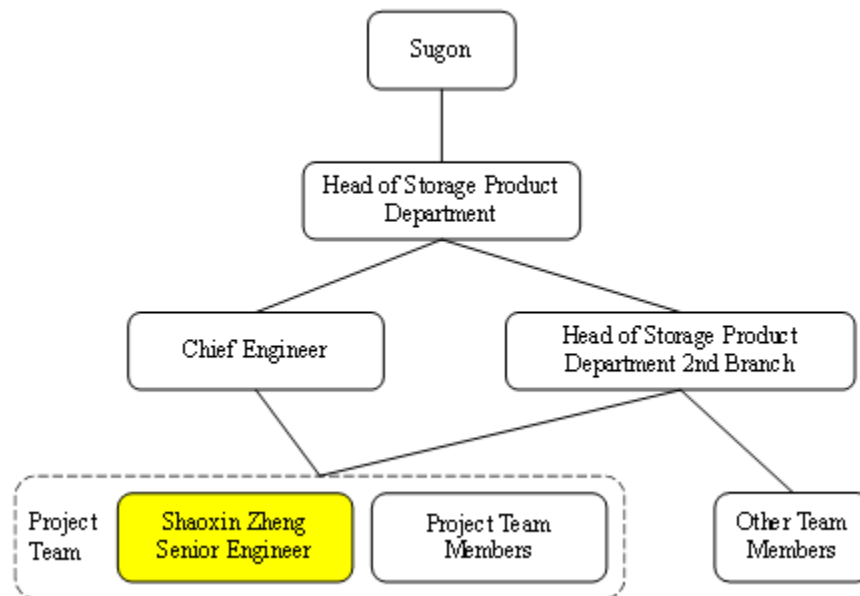
[CE 3.6] Below were the duties and responsibilities that I undertook for this project:
- Identifying and assessing scope, schedule, cost, and risk influenced by the customer requirements involving expected outcomes and desired performance indicators.
- I understood and defined the design requirements and functionality according to design principles, user experience, and technology to enhance the competitiveness of the existing solution in the market.
- I reviewed and refined the design solutions to determine the final design specification.
- I managed a short-term development team to complete the project.

- I directed technical activities to design, develop, test, and migrate solutions into production.

C) **Personal Engineering Activity**

[CE 3.7] I started the project by conducting the seminar to discuss the crucial gaps with outside critics and in-house experts. The former said a great solution should have full function and superior performance. They insisted almost 90 percent of the differences should be within the scope of the project. On the contrary, in-house experts said the solution of Sugon had its market positioning, and the goal of Sugon was not ever to be number one in the storage market. From their perspective, Sugon had delivered enough function and performance at a reasonable cost in a limited-resource setting, and thus it was not required of Sugon to change anything on the existing cloud storage solution. Indeed, though both sides told the truth, they were a little bit extreme, and their suggestions were not constructive.

[CE 3.8] I created a questionnaire on the desired features of a cloud storage solution, and clients could select up to three features from the differences between Sugon and other three vendors. I visited the clients and distributed the questionnaire to them, asking them to complete it. I thought this was a trade-off between outside critics and in-house experts. From all selected features, I determined the top three features provided as follows in the table. Considering limited human resources and project deadline, I defined the three features as the key gaps, and thus the scope of the project was determined.

Table - The gaps between Sugon and the market-leading vendors.

	Vendor A	Vendor B	Vendor C	Sugon
Data access control (DAC) with interoperability	Available	Not Available	Not Available	Not Available
High concurrent video streams on video rendering using Final Cut Pro (FCP)	Available	Unknown*	Unknown	Not Available
Server Message Block (SMB) transparent failover	Available	Available	Unknown	Not Available

* No information could be used to determine availability.

[CE 3.9] For DAC with interoperability, its emphasis was placed on the interoperability. The existing cloud storage solution had already supported data access control when connecting with either Unix or Microsoft Windows platform. For example, when user exported remote file system on their Unix operating system using networked file system (NFS) protocol, they were able to control well-defined permission using mode and access control list whose behaviors were defined by NFS protocol. By contrast, users were able to employ SMB protocol to share a remote file system on Windows system, and thus a more well-defined security descriptor would be used to manipulate object permission than those of Unix. When a file was shared by both a Unix user and a Windows user, with the help of DAC with interoperability one party's behavior would be restricted in terms of the permission rule set by the other party. However, the semantics of DAC

in NFS was different and incompatible with those of SMB. It was the core challenge that I needed to overcome.

[CE 3.10] I read the official documentation of Microsoft about SMB security descriptors. I found that there were thirteen types of permission to restrict user's access operations. I also read the NFS version 4 protocol standardized as Request for Comments (RFC) 7530. The protocol had a chapter to describe access control attributes, and it introduced seventeen types of bitmask to represent different permission. The definitions of both protocols seem to be quite similar, but there are some subtle differences in semantics. To verify my theoretical understanding, I conducted practice on both Windows SMB server and Linux NFS server. I designed the test cases for each permission and recorded actual results against expected results described in the official documentation. The experiment indicated that my understanding of both protocols accorded with the fact.

[CE 3.11] I proposed a design solution of a new DAC mechanism that would provide advanced access control lists (AdvACLs) which was merged to the cloud storage systems. The AdvACLs replaced the legacy permission mechanisms existing for NFS and SMB respectively. The core thought of the design solution was that the AdvACLs provided whole semantics for both NFS and SMB, and thus it was able to be employed to actualize access control attributes of NFS as well as security descriptors of SMB. When different users accessed certain one resource (file or directory) via different protocols, in fact, they interacted with the same permission mechanism that was AdvACLs instead of two separate systems like before. I summarized the experiment results and design thought as a high-level design specification and distributed the following task to one member of the project team.

[CE 3.12] For high concurrent video streams on video rendering using FCP, it focused on the number of concurrent video streams. FCP was a proprietary video editing software running on Apple's Mac Pro workstation. I did not have the ready resources to conduct study, development, and testing activities, so I had to request a budget for expensive Apple products. To persuade budget commissioners, I collected previous proof of concept (PoC) cases that our storage solution failed to gain the orders because of the lack of such a feature. I found that the sum of expected revenues of the orders was much higher than the cost of Apple's software and workstation. As a result, the budget was approved, and I purchased the workstation; installed FCP on it. Now this, all the preconditions were met. I executed testing on the workstation, and I found the existing solution can only offer one playback video stream. That was unacceptable for regular applications. The leading vendor could provide six concurrent video streams. The gap was very huge, and this was the key challenge that I needed to overcome.

[CE 3.13] I reviewed the existing data path from video editing software to the hard disk drive. I found the NFS server system distributed I/O requests evenly on 8 threads, but the scheduling model made cache mechanism failed. Without the help of the cache system, it is impossible to gain high

concurrent performance. I proposed a new scheduling model that I/O requests on a stream can be only dispatched to one thread, and thus the thread could use a cache system to accurately predict the following access. When many requests to different streams entered the NFS system, the new scheduling model would assign a dedicated thread for each stream. As a result, it allowed concurrent threads to take care of their requests on certain stream. I recorded the above outcome as a design specification.

[CE 3.14] To ensure the quality of detailed design specifications, I made a specification template that involved key procedures the team members should follow during the process of design. This was a great method to produce adequate outcomes for inexperienced engineers. In addition to that, I invited professional colleagues and chief engineer to attend a review meeting for the specification of each feature. At the meeting, they provided many great advice and suggestions. Most of those were reasonable and I accepted them so that the whole design became more thorough and effective.

[CE 3.15] After the intense work of many months, the project was complete. Users from different operating systems could gracefully manipulate their storage resources according to the same semantics. The user could gain up to six video streams on an Apple Mac Pro workstation using FCP. I introduced two new systems as well as the enhanced NFS server system with a new scheduling model into our storage solution, and it evolved to that as shown below:
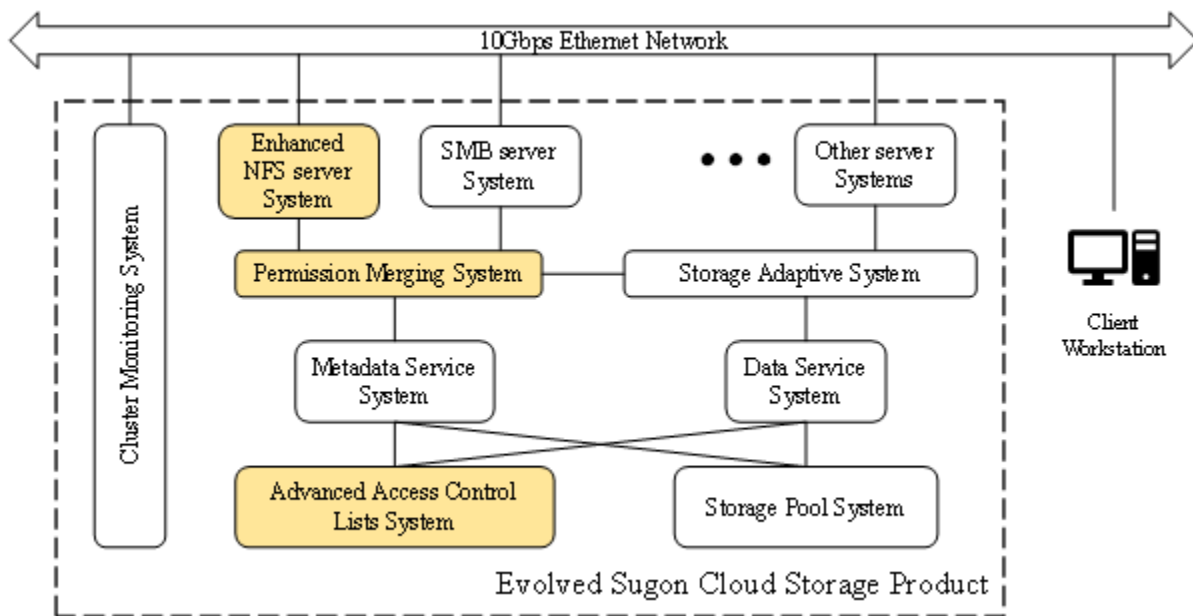
D) Summary

[CE 3.16] In this project, I applied a questionnaire to requirement analysis for the first time. I found this form of the investigation had a better effect than other methods. Also, I defined product

requirements and functionality in terms of customer requirements. This allowed me to accumulate more experience in product design.

[CE 3.17] With the evolved product, Sugon was able to meet almost 95 percent of key customer requirements. I heard positive voices from sales teams because they could get out of a passive situation of the lack of key features. Though I left Sugon later, I received the good news that Sugon earned more order because of the project. To sum it up, I identified scope, schedule, cost, and risk of the project and I understood customer needs and defined product requirements. I made a high-level product design and I managed the team to complete the project.