BASIC UNDERSTANDING ON DISAGGREGATED DATA STRUCTURES

In the traditional server-centric data centres, fixed amounts of processor and memory resources were attached within the boundary of a mainboard tray. This arrangement could not handle diverse workloads efficiently. So to maximize the resource utilization and flexibility, the concept of Disaggregated data centres was introduced.

The main idea is to replace the server-centric model (where CPU co-exist with memory as a single unit) with a resource-centric model (function block as a unit). This kind of network-centric model could enable new workflows with few resource boundaries.

The main challenges of such a system are[1] :

a) latency overheads, compared to the current direct-attached model, should be minimized,

b) system should support substantially higher bandwidth and bandwidth density at very low cost and power consumption,

c) network system should offer specific performance and services according to communication type (e.g., compute-to-memory, compute-to-storage) on the same substrate for maximum flexibility

d) orchestration of computing, memory and network resources should maximize resource utilization and workload performance at minimum cost.

Optical fibre communication is a promising technique to offer high capacity and low latency, but it is still very challenging for state-of-the-art optical transmission technologies to meet the requirements of the fully disaggregated data centers.

From a very recent paper [2], an experimental study was done to evaluate the performance enhancement using different levels of disaggregation. It concluded that the bandwidth provided by the state-of-the-art optical transmission technologies is not always sufficient for the fully disaggregated data centers and states that research on cost-efficient short-reach optical transmission with higher bandwidth (e.g., over 1 Tb/s) is required to fully utilize the advantages of function disaggregation in data centres.

I Read about **Intel** Rack Scale Design from their whitepaper[3]. From this document, I got the basic idea of how they implemented the Composable Disaggregated Infrastructure (CDI).

The key concept of CDI:

The fixed ratios of computing, storage and networking resources are broken into separate resource pools that can be interconnected when required into nodes or logical systems. When a particular node is no longer needed, the resources could be released back into the resource pool for other uses.
Intel has relied on the advancing speed of 100G ethernet, Intel Omni-path and optical interconnects to implement this without sacrificing the performance.

At **Facebook**, they build the data centres with fully open and disaggregated hardware which allows them to replace the hardware or the software as soon as better technology becomes available. Because of this, they see compute, storage, and networking gains that scale with their business. Some of their networking hardware and software includes Wedge 100, Backpack, Voyager, FBOSS and OpenBMC.

To scale Facebook's infrastructure and run the network at the required speeds and with the flexibility needed, they developed their own 32x100G port ToR switch, **Wedge 100**, and their own 128x100G port modular switch, **Backpack**. The Wedge and Backpack platforms enable a complete disaggregation of software and hardware with clearly separated data, control, and management planes. They also applied innovations in chassis design to support higher power and high-speed 100G interconnects.

I shall read in more depth to understand the functionalities and the architecture of these technologies and try to identify the further scope of work.

References :
[1]**Disaggregated Compute, Memory and Network Systems: A New Era for Optical Data Centre Architectures**
G.Zervas1,F.Jiang2,Q.Chen1,2,V.Mishra1,H.Yuan1,K.Katrinis3,D.Syrivelis4,A.Reale3, D.Pnevmatikatos5, M.Enrico6, N.Parsons6
1University College London, United Kingdom, 2University of Bristol, 3IBM Research – Ireland, 4UTH, 4FORTH-ICS, 5Huber-Suhner Polatis
g.zervas@ucl.ac.uk
[2] **Disaggregated Data Centers: Challenges and Tradeoffs**
Yuxin Cheng1, Rui Lin1, Martlet De Andrade2, Lena Wosinska3, and Jiajia Chen1
School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden. Ericsson Research, Sweden. Department of Electrical Engineering, Chalmers University of Technology, Sweden
[3]**Intel Rack scale architecture** - White paper
https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/rack-scale-design-architecture-white-paper.pdf