



KISACO INTERVIEW WITH JOHN OVERTON

Kove founder and CEO John Overton took the stage for a keynote presentation at MemCon 2024, sharing how Kove:SDM™ is helping businesses do what they could never do before by eliminating memory limitations using only software. We were lucky enough to catch up with Overton afterwards to dig deeper on his company's first-of-its-kind technology and the considerable impact it is already having on partners such as Red Hat and Swift.

Q1: HOW ARE SHIFTING COMPUTE FUNCTIONS AND WORKLOADS INFLUENCING THE DEMAND FOR MEMORY, AND WHAT ADJUSTMENTS TO SYSTEMS ARCHITECTURE ARE NECESSARY TO ACCOMMODATE THESE EVOLVING NEEDS?

John Overton: Memory demand is exploding, and infrastructure struggles to keep up. AI, ML, any kind of artificial intelligence—the only way the edge happens is by manipulating memory. It's too expensive to strand memory across server stamps, across the edge, and software-defined memory will allow the memory to move to where it's needed. And then, in the data center, it gives you the ability to scale memory linearly across all the resources that you have as opposed to merely inside of individual servers. You can move memory to compute – across the datacenter with Kove software.

Q2: WHAT STRATEGIES SHOULD ENTERPRISES AND HYPERSCALERS ADOPT TO IMPLEMENT COMMON MEMORY POOLING AS A SCALABLE SOLUTION, AND HOW CAN THEY DEVELOP A GROWTH ROADMAP TO NAVIGATE THIS EVOLVING LANDSCAPE?

John Overton: Memory pooling is available as a scalable solution now, through SDM. Serialized interconnects need to be made available. These exist in InfiniBand and RoCE and new ones are already being worked on. You need SDM, not hardware, to achieve this. Some consortiums like CXL are proposing you need a hardware roadmap for this, but this is not the case. Software-defined approaches have repeatedly demonstrated their contribution to storage, compute, and networking. We at Kove think applying software-defined to memory will similarly contribute to the marketplace. If memory virtualization follows the pattern of benefits from other virtualization approaches, software will lead the way and subsume hardware virtualization efforts for the full benefits of pooled memory.

Q3: WHEN EVALUATING SOFTWARE-DEFINED MEMORY, WHAT CHANGES IN MINDSET ARE ESSENTIAL, AND WHICH ORGANIZATIONAL CONTEXTS OR SECTORS ARE PARTICULARLY WELL-SUITED FOR THIS TRANSFORMATIVE APPROACH?

John Overton: There are three categories: One is, you need memory-centric computing, which can be across any kind of field. Two is, going to be cost efficiencies that you need to have to supply memory at scale. And three is the redistribution of memory and that infrastructure you need to be able to define what the needs are for.

As far as the organizational context or sectors that are particularly well-suited to this approach, I think again edge is particularly well-suited for this approach because it's the only way to make it work. So, if you want smart cities, you want to avoid the latency of going up and down to the cloud. Any kind of data center—which could be hyperscale or it could be enterprise—would be well suited to the economies of having the right memory meet the right compute where and when it's needed. And when done, redistribute. The dispositional shift will evolve from always being about how much “compute you have”, to creating the right memory-compute relationship, where and when needed.

Q4: IN WHAT WAYS DO IT PROFESSIONALS NEED TO RECONSIDER THEIR INFRASTRUCTURE INVESTMENTS IN LIGHT OF ADVANCING MEMORY TECHNOLOGIES, AND WHAT STEPS CAN BE TAKEN TO MAXIMIZE THE UTILIZATION OF EXISTING INFRASTRUCTURE?

John Overton: We need to quit thinking in terms of maximizing CPU memory on a per server basis and start thinking about maximizing each of them as an independently virtualized surface area. We already understand the value of managing storage in aggregate. We can manage compute, GPUs, and CPUs in the aggregate. And how do we manage memory in its aggregate that would then apply both to performance and cost inverses? More performance means you get more efficiency, and therefore lower cost per performance unit. If you get more efficiency and performance, you can buy less memory per server, saving money and still achieving the same amount or more. Finally, in a software-defined memory approach, you can build servers much larger than physical servers – dynamically and cost-effectively – simply by applying memory from unused equipment to a composable server need at a moment in time.

Q5: WHAT ARE THE PRIMARY OBSTACLES HINDERING ENTERPRISES FROM DEPLOYING AND INTEGRATING EMERGING MEMORY INNOVATIONS, AND HOW CAN ORGANIZATIONS OVERCOME THESE CHALLENGES TO DRIVE VALUE AND COMPETITIVENESS?

John Overton: We've got to quit thinking it's not possible, because it is. That's really all there is to it. We've got to change the mindset that memory is subject to the speed of light, distance of cable problems, etc. Memory is now a virtualized resource like any other, and should be thought of and deployed accordingly. InfiniBand is currently the best serialized interconnect, but we expect many others to be competitive soon. And Kove can service and work with any hardware that's available.

Q6: WHAT DID YOU LIKE MOST ABOUT MEMCON 2024? HOW IS THIS BRINGING VALUE TO YOUR ORGANIZATION?

John Overton: It's one of the greatest densities of people who are thinking hard about memory and leaders of the next generation of computing infrastructure. We want these people especially to know that SDM is available now, so it's great to be surrounded by so many like-minded individuals who see the future and are thinking about it as something that's real and approachable. Everyone I spoke to and met at MemCon was like that, and it was exciting to meet so many potential future partners and collaborators. It enabled us to forge potential alliances that can be groundbreaking for the future of computing.

REGISTER YOUR INTEREST FOR MEMCON 2025



MEMCON



KOVE™