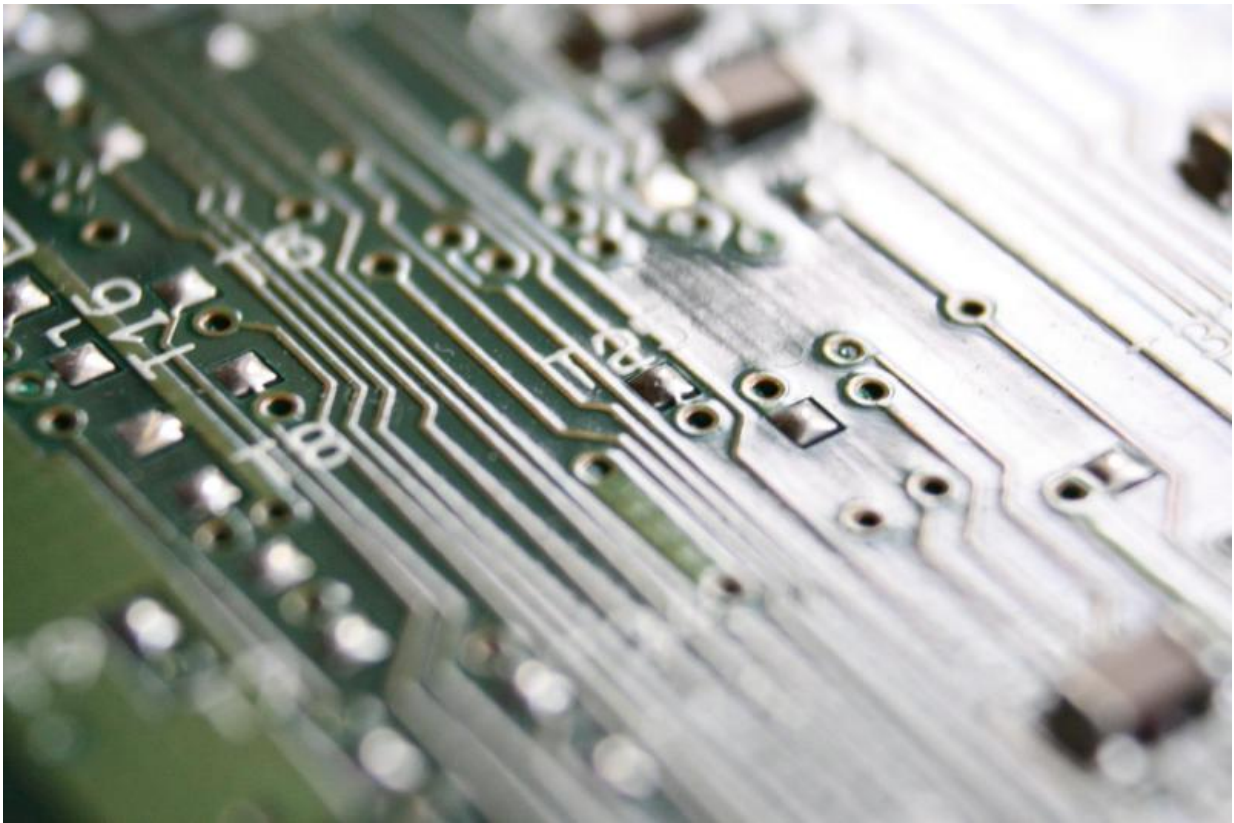# 'Memory disaggregation' for large-scale computing made practical

May 25 2017, by Nicole Casal Moore



Credit: Public Domain

For decades, operators of large computer clusters in both the cloud and high-performance computing communities have searched for an efficient way to share server memory in order to speed up application

performance.

Now a newly available open-source software developed by University of Michigan engineers makes that practical.

The software is called Infiniswap, and it can help organizations that utilize Remote Direct Memory Access networks save money and conserve resources by stabilizing memory loads among machines. Unlike its predecessors, it requires no new hardware and no changes to existing applications or operating systems.

Infiniswap can boost the memory utilization in a cluster by up to 47 percent, which can lead to financial savings of up to 27 percent, the researchers say. More efficient use of the memory the cluster already has means less money spent on additional memory.

"Infiniswap is the first system to scalably implement cluster-wide 'memory disaggregation,' whereby the memory of all the servers in a computing cluster is transparently exposed as a single memory pool to all the applications in the cluster," said Infiniswap project leader Mosharaf Chowdhury, U-M assistant professor of computer science and engineering.

"Memory disaggregation is considered a crown jewel in large scale computing because of memory scarcity in modern clusters."

The software lets servers instantly borrow memory from other servers in the cluster when they run out, instead of writing to slower storage media such as disks. Writing to disk when a server runs out of memory is known as "paging out" or "swapping." Disks are orders of magnitude slower than memory, and data-intensive applications often crash or halt when servers need to page.

Prior approaches toward memory disaggregation—from computer architecture, high-performance computing and systems communities, as well as industry—aren't always practical. In addition to the new hardware or modifications to existing applications, many depend on centralized control that becomes a bottleneck as the system scales up. If that fails, the whole system goes down.

To avoid the bottleneck, the Michigan team designed a fully decentralized structure. With no centralized entity keeping track of the memory status of all the servers, it doesn't matter how large the computer cluster is. Additionally, Infiniswap does not require designing any new hardware or making modifications to existing applications.

"We've rethought the well-known remote memory paging problem in the context of RDMA," Chowdhury said.

The research team tested Infiniswap on a 32-machine RDMA cluster with workloads from data-intensive applications that ranged from in-memory databases such as VoltDB and Memcached to popular big data software Apache Spark, PowerGraph and GraphX.

They found that Infiniswap improves by an order of magnitude both "throughput"—the number of operations performed per second—and "tail latency"—the speed of the slowest operation. Throughput rates improved between 4 and 16 times with Infiniswap, and tail latency by a factor of 61.

"The idea of borrowing memory over the network if your disk is slow has been around since the 1990s, but network connections haven't been fast enough," Chowdhury said. "Now, we have reached the point where most data centers are deploying low-latency RDMA networks of the type previously only available in supercomputing environments."

**More information:** Get Infiniswap at github: github.com/Infiniswap/infiniswap

Efficient Memory Disaggregation with Infiniswap. www.usenix.org/conference/nsdi … ions/presentation/gu

Provided by University of Michigan

Citation: 'Memory disaggregation' for large-scale computing made practical (2017, May 25) retrieved 20 April 2024 from https://techxplore.com/news/2017-05-memory-disaggregation-large-scale.html