Project Overview

The Cerebras Systems Wafer-Scale Engine 2 (WSE-2) is by far the largest silicon product available, with a total silicon area of 46,225mm². It utilizes the maximum square of silicon that can be made out of a 300mm diameter wafer. The square of silicon contains 84 die that are 550mm² each. These die were stitched together using proprietary layers of interconnect, making a continuous compute fabric. By developing this interconnect on a single piece of silicon, Cerebras were able to connect the equivalent of 84 die and significantly lower the communication overhead and physical connections within the systems.

Challenges

Giant models need massive memory, compute, and massive communication to tie it all together. Trying to provide this odelsee memoovide coms5 (e

Having the monitors and sensors distributed in large number across the device allowed Cerebras to measure variation across