HPTS Comes Full Circle

Over the years I've been attending HPTS, the meaning of "scale" has changed fundamentally. In the early days, "high-scale" were the transaction processing systems used by the largest enterprises, and we talked about how these systems were built and how to keep them scaling. Later, data warehouses and large-scale analytics became the HPTS definition of "large", and their 2x growth per year became our definition of rapid scaling. In 2001 Charlie Bell and Rick Dalzell spoke at HPTS about Obidos, the core of the Amazon.com ecommerce system. Again, our definition of scale had advanced – these mega-ecommerce systems made the large-scale enterprise systems look small.

Eight years later our definition of large had further advanced, and very large search and advertising systems had become our latest definition of "big." Urs Holzle and Luiz Borroso's "The Datacenter as a computer: An Introduction to the Design of Warehouse Scale Machines" focused on treating the entire datacenter as a computer. At this scale, custom hardware was possible and most of the hyperscalers were using custom server designs, with a few doing all-custom networking hardware. This revised definition of "high-scale" was a fundamental breakthrough in that we were no longer restricted to software innovation and the new canvas for high scale systems included the entire hardware and software stack.

High-scale cloud computing operates at a scale an order of magnitude beyond search and advertising systems and opens up yet another dimension for innovation. Building-level warehouse computing has given way to regional computing complexes of 10s of datacenter each hosting 80 to 100 thousand servers. The world-wide footprint of the largest operators is many 10s of millions of cores and growing rapidly. The best operators develop their own software, their own servers, and their own networking equipment. None are stuck accepting the slow-paced world of standards, and the scale of a single purchaser has continued to increase to the point where even custom semiconductors are now possible. In the networking world, companies have emerged whose entire business model is producing custom networking ASICs (Application Specific Integrated Circuits) for hyperscaler cloud computing operators.

The growth rate of the hyperscaler operators has continued to accelerate and the next layer of innovation has opened up: custom general-purpose compute processors. Processor design is incredibly expensive where a single part will require a several hundred-million-dollar R&D program and competitive processors require generations of investments. HPTS innovation now includes the entire software stack, the servers, the networking system, and also the processors that form the servers.

In some ways, we have come full circle where, in the early days of HPTS, we were looking at the crazy low-volume IBM mainframes, assembled into complex systems whose sole purpose was to allow the few very largest enterprises to scale their transaction systems. These wonderful systems that supported the world's largest companies in the early days of HPTS were made irrelevant by high-volume silicon, high-volume commodity servers, and open source systems software. The super-optimized server hardware systems were replaced by high-scale clusters of commodity systems available to all customers.

The scale of cloud computing where 10s of millions of cores are purchased each year by a single provider returns us to the days of hardware innovation. The volume is there to support the development of general-purpose server processors by a single customer for use by a single customer.

The pallet of optimization open to HPTS attendees now starts at semiconductors optimized for the use of a single customer, includes custom servers, custom networking designs, the entire system software stack, custom database systems that aren't even available on the open market, and all the control software that ties it all together. Just as early IBM mainframes had custom ASICs whose sole purpose was to control the cluster and specialized processors who sole purpose was to offload I/O from the central processing complex, all these concepts have returned and exist in modern high-scale cloud computing system.

This talk focuses on a high-scale cloud operator with 10s of millions of cores deployed annually and the layers of innovation from semi-conductor, through custom servers, custom control processors, and custom networking hardware. We'll look at what's made possible by having control of all components in the hardware and software stack and show how many of the innovations of early mainframe processor complexes have returned.

James Hamilton
Vice President & Distinguished Engineer
Email: james@amazon.com
Web: https://perspectives.mvdirona.com/
Home: https://mvdirona.com

James joined Amazon in January 2009 working on the Amazon Web Services team where he has focused on infrastructure efficiency, reliability, and service scaling. He’s now in a similar role with cross-Amazon scope.

In past roles James was at Microsoft for 12 years where he served GM for the Microsoft Exchange Hosted Services team, GM of the SQL Server WebData team, and held many other positions on Microsoft SQL Server. Prior to that, he was at IBM for 11 years where he started the IBM C++ Compiler team and was Lead Architect for IBM DB2. Before joining IBM, he worked as a professional auto mechanic servicing Italian cars and exotics such as Maserati, Ferrari, Lamborghini, and Alfa Romeo.

He has a Master of Math (Computer Science) from the University of Waterloo and Bachelors of Science (Computer Science) from the University of Victoria. James holds 209 patents in 22 countries in server and datacenter infrastructure, database, and cloud computing.

James and his wife Jennifer have spent a decade sailing around the world in a small boat. It was an incredible adventure full of life changing experiences but it did mean that a few HPTSs were missed :-).