The move to serverless
What’s next for this simplified solution?

“I think there is a world market for maybe 5 computers”

- Thomas Watson, President of IBM, 1943

Just a couple decades ago, we couldn’t have imagined how ubiquitous computers would become, how they’d penetrate the market, and their effect on our daily lives – so what happened next is nothing short of a revolution. Computers are everywhere, at work, at home, in our pockets, robots, cars, the data centers (the cloud) that power businesses and so much more.

Fueling this growth was the concerted effort made to ensure that consumers would be able to use computers intuitively. Operating systems improved, evolving from DOS to Windows and MacOS, while a focus on the user interface and experience ensured consumers were delighted by these products’ instinctive ease of use. Similar transitions can be seen with smartphones and many other compute-powered consumer devices.

Unfortunately, a similar revolution didn’t occur for businesses – one main reason being that up until a decade or so ago, most users were technical employees or in technical fields. But today, compute-driven workloads have become core value drivers for businesses across all industries and employees spanning technical and non-technical roles, yet the skills required to set up, service and maintain the infrastructure powering these workloads are not held by the majority of users.
Before cloud was introduced, all compute resources were managed on-premise by company employees. While this allowed for full control of their machines, better security and potentially lower costs, it came at a great expense to their personnel. IT departments had, and still have, to buy and configure their own servers, connect them to the local network and internet (if the servers are meant for customer-facing functions), install and run any required software on it, and more. And in addition to procurement and set up, these teams are responsible for ongoing maintenance to ensure things are always up and running.

Managing computing locally means that in addition to all the processes IT has to handle, the organization also faces compute restrictions that make it very difficult, and in some cases, impossible, to scale. For growing companies with increasing demands for compute, these restrictions create bottlenecks that obstruct the path to success.
The rise of cloud was born from a need to remove some of the challenges presented from on-prem systems. Users no longer have to maintain and create space for physical equipment, but managing and configuring compute has remained a complicated task due to the sheer number of configuration options available. This creates several problems. One, it’s extremely difficult to set up a cloud computing cluster. Engineers spend an extensive amount of time cobbling together piecemeal solutions which leaves them with less time to focus on building the product they were hired to build. Their lack of expertise not only prolongs the time it takes to find a solution, but often results in poorly built solutions that are difficult to scale, lack flexibility, and run at higher costs due to inefficient resource management or over-provisioning.

Companies who recognize this inefficiency have formed or brought in DevOps teams, a profession that has taken off due to the rise in popularity of cloud computing. DevOps is often under a lot of pressure to build and maintain production environments within cloud platforms. DevOps teams are often understaffed and can’t keep up with the needs of teams across the enterprise. These teams are then slowed down as they “wait their turn”, or they repeat the cycle mentioned above, diverting time from their primary responsibilities to build their own suboptimal solution.
Taking advantage of the continual growth and changes in computing, serverless represented an evolution of the traditional cloud model. While serverless is still powered by servers, it eliminates the need to deal with server set up and configuration.

Serverless was created to solve the following issues:

**Server management**

The serverless model removes the need for system admins to spend time installing, configuring, monitoring and optimizing servers. With serverless, none of this burden falls to the user, they simply activate the service and gain access to compute.

**Availability**

Both cloud and on-prem servers have to constantly be monitored to make sure they’re always up and running. With serverless, this burden is eliminated. The provider checks for problems on an ongoing basis and devises the solutions required to resolve them. This decreases downtime significantly in addition to minimizing opportunities for such occurrences.
The move to serverless

Serverless addressed many of the problems created by the first generation of cloud computing, but wasn’t without a few issues of its own. Since users don’t manage the servers, they have little control over the types of computing resources\(^1\) that are deployed for them. What’s allocated are often lower performance machines which limits the types of workloads that can be executed. Performance is also diminished by long response times (latency), the distance between the data and compute, and multi-tenant architecture which has multiple tenants vying for use of the same resources therefore reducing performance. Lastly, code often has to be adapted to the framework of their serverless solution, resulting in vendor-lock that prevents easy migration between tools.

Serverless also promised developers that spend their time focusing on developing and deploying flawless code, but unfortunately, there are still some gaps that, even with serverless, derail developers from making the best use of their time. For companies that desire to connect on-prem resources, where most computing workloads are still executed, to their cloud, this continues to be an uphill battle requiring more time and expertise than many developers have.

---

\(^1\) In AWS Lambda for example, users can only choose how much memory they would like to be allocated for their function (up to \(\sim 3\)GB), which indirectly determine the processing power as well.
The next steps for serverless

Ultimately, computing solutions will have to address the growing need for accessible compute power that’s not subject to vendor lock (being locked into the ecosystem of only one provider). Multicloud offerings are needed to allow organizations to scale compute, mitigate compute, bring compute closer to the data, and seamlessly connect on-prem and cloud resources to optimize the power they already have, saving them money in the long run. The next gen of serverless also has to simplify the lives of developers. They shouldn’t be required to adapt and change their code for each service used. This requirement is also what enables vendor lock to remain a challenge, developers don’t have time to customize code to every solution, they need an easier way to maintain consistency across providers. By removing this requirement, systems, both cloud and on-prem, will become more unified freeing users to run jobs anywhere, on any service making multicloud more attainable.

dis.co is playing a part in paving the way for the next evolution of computing by seeking to close the previously mentioned gaps. It’s a serverless solution that allows developers to maintain a focus on nothing more than their code and their data. Developers gain access to infinitely scalable compute while being able to choose where their code runs - in their own cloud environment, in the dis.co cloud or on on-premise devices. Finally, dis.co provides the flexibility to optimize for costs, team productivity, and/or performance in the click of a button.
If you're interested in the next generation of serverless computing, email to learn more:

contact@dis.co