

# Architecting Scientific Computing Systems

By Kemal A. Delic – Martin Antony Walker

The explosive evolution of the Internet over the past sixty years, from a restricted network connecting a few sites to today's ubiquitous infrastructure, has wrought profound changes [1] on the lives of an ever increasing fraction of the population of our planet. The arrival of the internet of things (**IoT** in the figure below) is, if anything, accelerating the impact of the internet.

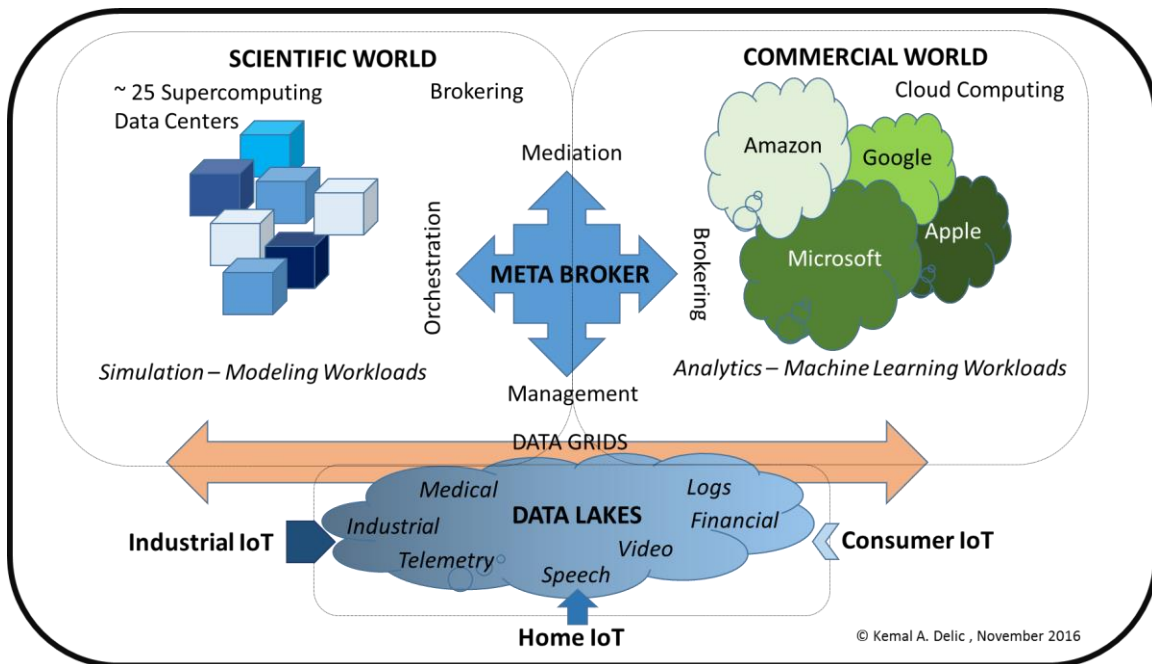
An essential concomitant of this evolution of the Internet is the arrival of hyperscale computing systems that have greatly facilitated the provision of, and access to, Internet services. We observe that science is an essential human activity that has not yet fully benefitted from the combination of the internet of things and hyperscale computing systems.

We suggest that architecting hyperscale computing systems for science will enable tackling grand challenges in science and engineering [2], including, for example, understanding weather and climate; improved fitness for use and more benign environmental impacts of products and services; better understanding of the fundamental laws of nature in physics, chemistry and biology; understanding human society, life, and the human brain; understanding ecosystems and diversity in the oceans and on the earth; and understanding the solar system, the galaxy, and the universe.

Recently, we have outlined a conceptual architecture for hyper scale scientific systems [3] which may represent a possible future scientific infrastructure based on current technologies and real-world circumstances. **Data lakes** are recently introduced entities containing vast quantities of a wild variety of data coming from, e.g. streaming from CCTV cameras; industrial telemetry data feeds; medical instrumentation creating torrents of data; astrophysical research producing petabyte volumes in seconds and minutes; and financial systems creating avalanches of data from high-frequency-trading. In the preceding technological era, we had data grids – which embodied computing grids spanning several data centers in various geographies, representing previous instance of 'data lakes'.

Data by themselves have limited value until they are transformed into information, knowledge, discovery, or wisdom. To achieve such a transformation, large computing facilities – known as **supercomputing data centers** have been established across the world. In parallel, in the commercial world, **cloud computing** has become a major business computing paradigm, in which a few vendors are dominating markets. They serve huge number of commercial clients, and increasingly include also scientists and scientific institutions among their customers.

We envision the future scientific backbone as the integrated, brokered infrastructure mediated by a **meta-broker**, which aligns and balances demand and supply of a spectrum of ICT resources. Of course, some big workloads will still require the engagement of many months of resources, but a large proportion of scientific workloads can be labeled as medium or small. This short note will not delve into intricate details of meta-broker functioning, but we wish to point out that this function is crucial for the bridging and integration of vast scientific computing resources.



To conclude, we advocated [4] academic computing clouds nearly ten years ago as an emerging new computational environment for the advancement of science, while today, with the arrival of new hyperscale technologies, entities and concepts, we are convinced more than ever that a new era of scientific computing is coming. Technology, the economy, and society itself have been transformed by the Internet, and our expectation is that the logical next step will be a transformation of science itself.

## References

- [1] Big Science Blog - <https://businessvalueexchange.com/blog/2016/08/28/big-science-will-require-big-different-infrastructure/>
- [2] On Big Science i-KNOW'15 - <http://dl.acm.org/citation.cfm?id=2809622>
- [3] Architecting Hyperscale Scientific Systems – vol. 49/p. 29  
<https://www.hipeac.net/publications/newsletter/>
- [4] Emergence of Academic Computing Clouds - <http://ubiquity.acm.org/article.cfm?id=1414664>

**Kemal A. Delic** is an associate editor for Ubiquity Magazine. He is also a senior technologist with Hewlett-Packard Co. He serves as an adjunct professor at PMF University in Grenoble, advisor to the European Commission FET 2007–2013 Programme, and expert evaluator for Horizon 2020. He can be found on Twitter @OneDelic.

**Martin Antony Walker** is a retired mathematician and theoretical physicist who has been engaged in high performance computing for the past thirty years. He advises corporations and international organizations on issues around scientific computing and technology.