# Smart Community CyberInfrastructure at the Speed of Life

A White Paper for BDEC2 by Glenn Ricart, US Ignite and U. Utah

**At the Speed of Life**

Smart Community cyberinfrastructure collects data at the speed of life and similarly often needs to respond in sync with the real world. If vehicles will communicate and coordinate their way through intersections without stopping, there must be reliable and timely information about the vehicles and their location, motion, and acceleration/deceleration fed into complex cyberphysical system decision-making which in turn is sent back to and verified by the vehicles in a trust-worthy way. All of that must occur in at most a few thousandths of a second. Smart community cyberinfrastructure must respond at the speed of life. (This is not unlike real-time data collection in HPC science experiments where the real-time results are coupled to and alter the experimental controls in real-time.)

Over time, smart and connected communities are finding a growing number of cyberphysical systems that improve their communities and its healthcare, public safety, education, economic development, and recreation, but which need to reliably operate at the speed of life in that community.

**Characterizing the Smart Community Applications of the Future**

As part of the NSF-sponsored *Looking Beyond the Internet* series of workshops held in 2015-2016, Glenn Ricart and Prasad Calyam held a workshop on *Applications and Services in the Year 2021*. That report noted that several of the sciences increasingly would rely on data-driven and AI-driven systems with strict **time constraints**. Data will most often be presented in streams. It often will be voluminous. Reliability and calibration may not be as high as in HPC scenarios due to commercial cost constraints. Context and metadata will be as important as the data itself. There are serious requirements for being immune to attack that don't arise in scientific instrumentation scenarios. While many lessons of HPC and distributed systems are highly useful, they are not sufficient in this environment.

In addition, sensor-driven data is often of use primarily in localized environments. In computer science terms, it has high **locality**. In addition, the data itself often is **perishable**. It may well be archived for future study, but its actionable information value fades as newer values from the same stream arrive.

**Enter Einstein**

In these circumstances of perishable data with high locality and time constraints for actions taken, the ability to use the currently dominant model of sending all data to huge and highly efficient datacenters or supercomputers is limited due to the speed of light in fiber. To achieve adequate response time and sufficiently reliable service, edge computing has become an important tool.

**The Edges and their Clouds**

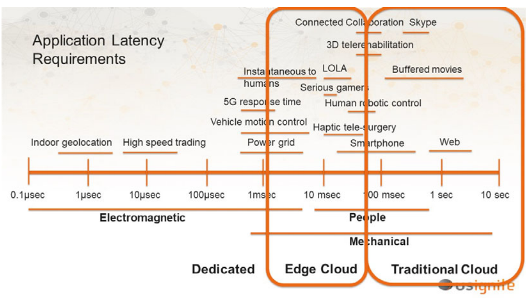
The leading edge of edge computing thinking is multiple edges. Different portions of an application may have differing requirements for response time, reliability, security, etc. Each microservice invoked has its own requirements which may be most efficiently handled through a network of edge clouds. It may often be better to move computation to the data than data to the computation. Computation and data need each other; they are cyberinfrastructure duals. It's a challenging and unsolved multi-dimensional optimization in time, space, and capability at the very least. Coordinated multi-edge and multi-service orchestration and hypervision will need to be part of the answer.

At the *Second National Research Platform Workshop: Toward A National Big Data Superhighway* (August 6-7, 2018), a major advance was recognition of the role of microservices and containers in high performance platforms. This means HPC and big data are moving to embrace distributed systems techniques while distributed systems are moving to embrace HPC and big data techniques.

Due to the current lack of suitable multiple-edge cyberinfrastructure, Rick McGeer, the author, and US Ignite are using an NSF grant to develop Edge-Net, a Kubernetes-based distributed container and VM driven viral edge cloud designed to service challenging smart community applications and other science.

**Application Requirements**

The author works primarily in the world of smart and connected community applications and services and their infrastructure. If selected to present this white paper at the BDEC2 conference, most of the time would be spent on smart community requirements for cyberinfrastructure, including volume, time constraints, reliability, resistance to attack, economics, veracity, privacy, and value to society. US Ignite has more than a hundred applications and services developed or under development in its 26 Smart Gigabit Communities, and the most demanding of these in terms of data and computation will be discussed along with their requirements. Examples will come from transportation, energy management, healthcare, and education, depending on the time available. To prototype these applications, we've had to handcraft some edge clouds and national network paths. For a few, we have videos in which the PIs describe them.




**Related Sources of Application Requirements**

Additional sources of smart and connected community application requirements come from the Global City Teams Challenge and their Action Cluster blueprints, and from the EU effort for FIWARE Smart Cities. In Asia, smart city requirements are present in the NSF-funded CENTRA effort. The EU, Asia, and Brazil have committed to joint smart communities applications and infrastructure research via GEFI. US Ignite, together with Northeastern University, is responsible for the PAWR advanced wireless research which addresses the high bandwidth and low latency wireless infrastructure needed by these applications and services.

US Ignite is working with the Alliance for Telecommunications Industry Solutions (ATIS) to develop a data exchange for smart community information. This effort is intended to help make smart community data more comparable and actionable, improving the deployability of smart community applications.

Glenn Ricart is co-founder of the San Diego Supercomputer Center and co-author of the Federal HPC Bluebooks in the 1990s. He was mentor to the West Big Data Hub during its formation. His current effort is US Ignite, a nonprofit where he's helping solve the chicken-and-egg problems in Smart and Connected community applications and their edge clouds. He holds an adjunct appointment at the University of Utah School of Computing.