A Public Health Grid (PHGrid): Architecture and value proposition for 21st century public health

T. Savel [a,*], K. Hall [a], B. Lee [b], V. McMullin [c], M. Miles [b], J. Stinn [b],
P. White [b], D. Washington [b], T. Boyd [a], L. Lenert [a]

[a] National Center for Public Health Informatics, CDC, Atlanta, GA, United States
[b] Deloitte, Consulting LLP, Atlanta, GA, United States
[c] Vanguard Consulting Group, Stone Mountain, GA, United States

ARTICLE INFO

ABSTRACT

Purpose: This manuscript describes the value of and proposal for a high-level architectural framework for a Public Health Grid (PHGrid), which the authors feel has the capability to afford the public health community a robust technology infrastructure for secure and timely data, information, and knowledge exchange, not only within the public health domain, but between public health and the overall health care system.

Methods: The CDC facilitated multiple Proof-of-Concept (PoC) projects, leveraging an open-source-based software development methodology, to test four hypotheses with regard to this high-level framework. The outcomes of the four PoCs in combination with the use of the Federal Enterprise Architecture Framework (FEAF) and the newly emerging Federal Segment Architecture Methodology (FSAM) was used to develop and refine a high-level architectural framework for a Public Health Grid infrastructure.

Results: The authors were successful in documenting a robust high-level architectural framework for a PHGrid. The documentation generated provided a level of granularity needed to validate the proposal, and included examples of both information standards and services to be implemented. Both the results of the PoCs as well as feedback from selected public health partners were used to develop the granular documentation.

Conclusions: A robust high-level cohesive architectural framework for a Public Health Grid (PHGrid) has been successfully articulated, with its feasibility demonstrated via multiple PoCs. In order to successfully implement this framework for a Public Health Grid, the authors recommend moving forward with a three-pronged approach focusing on interoperability and standards, streamlining the PHGrid infrastructure, and developing robust and high-impact public health services.

Published by Elsevier Ireland Ltd

1.   Introduction

1.1.   The state of public health

The infrastructure of "public health" consists of over 160,000 separate entities. Examples of these entities include county, city, and state health departments, public and private laboratories, volunteer organizations, private healthcare facilities, and federal agencies [1]. The Institute of Medicine in 1998 defined three core functions for all of public health: Assessment, Policy Development, and Assurance [2]. To achieve optimum assessment (i.e., monitoring) of the health of populations, well coordinated and timely access to a wide variety of information about a large number of people, across a broad range of health conditions, and over various geographic areas is required. Over time, these sometimes conflicting requirements have resulted in a public health surveillance ecosystem that is, unfortunately, essentially fragmented. Public health information systems play a crucial role in this public health surveillance ecosystem. Given the scale of public health, it has been estimated that, not surprisingly, there exist approximately 70,000 distinct information systems [3]. Given these issues, the Public Health community clearly faces significant challenges in achieving efficient and effective electronic exchange of data, information, and knowledge.

1.2.   A federal call to action

In April 2006, the House Committee on Government Reform released the report, "Strengthening Disease Surveillance" which focused on the fragmentation of public health surveillance systems in the United States. The Committee urged the CDC to address this issue of fragmentation [4]. In December 2006 Congress passed and the President signed the Pandemic and All-Hazards Preparedness Act (PAHPA), Public Law No. 109-417, which has broad implications for the Department of Health and Human Services' (HHS) preparedness and response activities [5]. Specifically, this Act states that the Secretary of HHS shall establish real-time electronic nationwide public health situational awareness capacity through a network of interoperable systems to share data and information to enhance early detection of rapid response to, and management of, potentially catastrophic infectious disease outbreaks and other public health emergencies. Innovative and emerging technology approaches to address this fragmentation challenge would clearly have significant public health impact.

1.3.   Grid computing

From the perspective of CDC and its partners, grid computing has significant potential as a technology approach to the public health challenges described earlier. Grid is a paradigm that proposes aggregating geographically distributed, heterogeneous computing, storage and network resources to provide secure and pervasive access to their combined capabilities [6]. Initially, grid technology development was driven by computing needs of the particle physics research community and enabled by the availability of high-performance networks. The term "grid" rapidly evolved toward a concept of ubiquitous and transparent computing to support a wide variety of applications, and builds on the well-known metaphor of the pervasive "electricity grid" [7]. As the grid industry continued to mature and advance, its scope began to widen as well. Specifically, its scope began to include not only massive computational (computer processing) challenges, but massively distributed data challenges as well. In other words, not only were computational grids being created, but "data grids" were beginning to be formed as well. It is the data grid aspect of grid computing that is most relevant to the focus of the author's research.

1.4.   Grid in action

Grid computing technologies are being developed in both the commercial space, and in the open-source community as well. Examples of open-source grid software/projects include: the Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) – an initiative in the pacific rim, and Enabling Grids for E-SciencE (EGEE) – an operational grid currently handling over 100,000 transactions per day, and growing [8]. The Open-Source initiative known as the Globus Alliance, produces open-source software that is central to science and engineering activities totaling nearly a half-billion dollars internationally and is the substrate for significant Grid products offered by leading IT companies [9].

1.5.   Grid and healthcare

Today, the grid-related activities in the healthcare space represent some of the most innovative drivers for progress in knowledge-based ubiquitous and transparent computing. Health-related grid activities are rapidly advancing in the US and abroad, and have resulted in the development of organizations, such as the HealthGrid.US (HG.US) alliance. HG.US is an affiliate of the international HealthGrid Association [10]. The European Union funded the international HealthGrid Association with €1 billion [11]. Nationally, the National Cancer Institute (NCI), supports the caBIG (cancer Biomedical Informatics Grid™) initiative, which was launched in 2004. caBIG leverages grid technologies to provide data and tools to scientists throughout the cancer community [12].

1.6.   Proposing a Public Health Grid (PHGrid)

Given the disparate nature of the public health community, and the maturing grid infrastructure, CDC and its partners have been investigating the development of a "Public Health Grid" which has the potential to afford public health a low-cost, and highly flexible and scalable environment (Fig. 1).

Please note that in this diagram, the smaller discs represent "nodes" which are connection (access) points to resources (i.e., services) which are maintained (i.e., controlled) by that local entity. Each node is essentially a technology connection point that is installed within an organization or partner site to share their resources and/or services with appropriate and authorized members of the Public Health Grid (PHGrid).

It is believed that such an information infrastructure could facilitate robust collaboration, seamless and timely exchange of health-related data, information and knowledge,