

Comment ■

## In Response to: What Is a Grid?

THOMAS G. SAVEL, MD, LESLIE LENERT, MD, MS, JONATHAN C. SILVERSTEIN, MD,  
KENNETH E. HALL, MDIV

■ J Am Med Inform Assoc. 2008;15:705–706. DOI 10.1197/jamia.M2707.

In a recent letter to the editor,<sup>1</sup> Dr. Peter Szolovits called for “distinct names for distinct ideas” when discussing grid. His suggestion for better “precision of language” arose from listening to numerous talks at the 2006 American Medical Informatics Association (AMIA) Symposium “where speakers describe grids that have little in common.”<sup>1</sup>

At this early stage in the maturation of grid, as with any emerging technology, lack of clarity is natural and, in fact, stimulates an ever-improving “precision of thought.”<sup>1</sup> Specifically, this ambiguity has already begun to generate discussion, as demonstrated by Dr. Szolovits’ asking “What Is a Grid?” Most would agree that ambiguity is part of the normal lifecycle of adoption of new ideas, concepts, and technologies. Our reading and interpretation of the 2006 Gartner Hype Curve<sup>2</sup> illustrates this point quite graphically; it articulates the ambiguity with regard to “grid computing” by placing it midway between the “peak of inflated expectations” and the “trough of disillusionment,” on its way towards the “slope of enlightenment” and within 2–5 years of mainstream adoption.<sup>2</sup> What could be more ambiguous than living somewhere in-between “expectations” and “disillusionment?”

Another aspect of ambiguity is the role of perspective. The spectrum of ambiguity ranges from the least (for those actively working with grid) to the most (for those peripherally involved or newly learning about the field). To borrow from a well-known metaphor to help explain the ambiguity

around the grid-focused presentations at the 2006 AMIA Symposium, the speakers unintentionally actualized the American poet John Godfrey Saxe’s poem, “The Blind Men and the Elephant.”<sup>3</sup> In the poem, each blind man described the elephant differently from his own perspective—the elephant’s side felt like a wall, his tusk like a spear, his trunk like a snake, his knee like a tree, his ear like a fan and his tail like a rope. The AMIA speakers presented different and important views on the very large domain of grid, resulting in, unavoidably, a somewhat unclear overall picture. In time, with further discussion and education, this picture can, and should, become much clearer. In fact, at this year’s Symposium a tutorial was held on the topic “HealthGrid” for those interested in the area.

Grid computing has unambiguously functioned as a distributed and cost-effective way to boost computational power to solve large-scale mathematical and data-bound problems for the physics and the bioinformatics communities, and continues to excel in these areas today. However, grid was, from the beginning, a more mature and complex concept that addressed a specific problem: “The real and specific problem that underlies the grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations.”<sup>4</sup> As envisioned in the initial concept, grid has matured far beyond computational augmentation to include inter-organizational, distributed data management and exchange, industrial-strength security, distributed services and workload orchestration. To be more precise, grid represents a robust framework (based upon the principle of service-oriented architecture (SOA)) for performing distributed computing tasks on the scale of the Internet which can enable “Service-Oriented Science [and Medicine].”<sup>5</sup> To this end, grid may be defined as four distinct technology layers (resources, middleware, services, and applications) and one implicit (social) layer.<sup>4</sup> This infrastructure assumes a lower layer of grid-accessible resources (widely dispersed storage, computational power (CPU cycles), devices, sensors, data and relational databases amongst different enterprises), a middle layer of grid-specific middleware (as provided by an open-source project known as the Globus Alliance (<http://www.globus.org>), and an upper layer of grid-enabled services and applications distributed across the Internet (the Internet itself enabling connec-

---

Affiliations of the authors: Office of the Director, National Center for Public Health Informatics, Centers for Disease Control and Prevention (TGS, LL), Atlanta, GA; Computation Institute of The University of Chicago and Argonne National Laboratory (JCS), Chicago, IL; BearingPoint, Inc. (KEH), Atlanta, GA.

The findings and conclusions in this document are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Use of trademarked names and websites is for identification purposes only, and does not constitute endorsement by the authors and their respective institutions/organizations.

Correspondence: Ken Hall, BearingPoint, South Terraces Building, 115 Perimeter Center Place NE, Suite 380, Atlanta, GA 30346; e-mail: <[ken.hall@bearingpoint.com](mailto:ken.hall@bearingpoint.com)>.

Received for review: 01/02/08; accepted for publication: 03/24/08.

tivity) that are useful to combine into arbitrary virtual combinations to achieve some common goal. The first two technology layers (resources and middleware), provide the core, domain-independent infrastructure. The *grid-enabled* upper layers represent the algorithmic stack which encapsulates the business logic required to support the shared processes (services and applications) within a specific virtual organization or "community of practice," (e.g., the health-care domain), and defines the domain-dependent grid-based services (data, computational, visualization, etc.). It is notable that these services harmonize with Web Service standards. The social layer, then, encompasses the governance, coordination and policy activities required to assure an efficient and effective community of practice while the technology layers provide a robust framework to execute such shared policies and goals. In other words, the community of practice could be thought of as the social manifestation of the grid. Fundamentally, grid computing interconnects people, organizations, processes, application silos and data silos in completely innovative ways, and has the potential to significantly augment the efficiency and effectiveness of virtual organizations (flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources).

In order to lessen the ambiguity one degree more, a brief comparison of grid computing with another distributed architecture should add both historical and technical depth to this conversation. The most similar architectural pattern relative to grid, at least on the surface, is the Common Object Request Broker Architecture, also known as CORBA. Both approaches attempt to integrate data, resource sharing, communications, conventions and organizational exchanges; nevertheless, CORBA contrasts more readily with web services than grid computing per se. First and foremost, grid provides a secure overarching framework in which web services may interoperate; in other words, web services are an integral part of the grid fabric. Secondly, grid supplies a rich set of additional computational, federated data and collaborative services above and beyond existing distributed systems. The CORBA may interoperate and coexist with web services but by no means are web services endemic to its architecture. Furthermore, while CORBA is distinctly tightly coupled, object-oriented and stateful, web services are loosely coupled, utilize a message exchange model and are stateless. These differences give web services an inherent flexibility and simplicity unavailable in CORBA implementations. Grid, therefore, is unique from other methods of distributed computing by its ability to coalesce the technology layers with the social layer through its commercial-strength security model (X.509 certificates), thus enabling the dynamic creation of multiple virtual organizations and providing unprecedented technical and social agility. Grid and CORBA represent different models for building distributed systems and are complementary rather than competing technologies and architectures.

So, while it may appear that the 2006 AMIA Symposium presenters spoke of different definitions of the same grid, demonstrating a lack of precision of thought, we would argue that they are, in fact, accurate definitions of different

aspects and layers of the same grid framework, as described above. Thus, grid (as a framework), in fact:

- provides a technical infrastructure that enables "a community of common interests,"<sup>1</sup>
- requires a governance model that supports "a social and funding infrastructure to encourage data sharing,"<sup>1</sup>
- delivers a comprehensive set of tools and consequently "a technical approach to what we used to call federated databases,"<sup>1</sup>
- provides a flexible, scaleable and secure technical framework that facilitates "standardization and ontology construction for specific fields, and of course,
- offers distributed computing far beyond its "original meaning."<sup>1</sup>

In this broader context, grid computing has the potential to offer a vast set of production-tested software tools designed specifically for building virtual organizations beyond the limits and silos of institutional boundaries. To further extend the metaphor, one could ask, "what is an interconnected national healthcare system if not a virtual organization?"<sup>4</sup> At the Centers for Disease Control & Prevention's (CDC) National Center for Public Health Informatics (NCPHI), we are attempting to answer this and other questions by engaging in active preliminary research to discover the potential role that grid-based technologies may provide in interconnecting and augmenting the virtual organizations known as public health and clinical care. Clearly the task is challenging. Fortunately, there are many worldwide initiatives from which to learn. In contrast to the US, Europe and other countries around the world have made significant advancements in the development and utilization of grid computing technologies across health and other domain boundaries. In addition, the five-year-old International Health-Grid organization (<http://community.healthgrid.org>) and the newly formed HealthGrid.US Alliance (<http://www.healthgrid.us>) are leading the way towards exploring and discovering the capability of grid computing for healthcare.

Altogether, grid is clearly a very broad domain, and the ambiguity Dr. Szolovits addressed is a challenge to be expected and addressed. To minimize ambiguity and accelerate its introduction into the healthcare domain, we propose that grid be conceived of as an overall framework with layers of core technologies, processes and sociology. Education and collaboration are critical to quickly mature its place in the US biomedical informatics field. We look forward to continued research and discussion on this topic.

#### References ■

1. Szolovits P. What Is a Grid? JAMIA 2007;14:386.
2. Pettey C. Gartner's 2006 Emerging Technologies Hype Cycle Highlights Key Technology Themes 2006. Available at: <http://www.gartner.com/it/page.jsp?id=495475>; accessed July 14, 2008.
3. Saxe JG. Blind Men and the Elephant. Available at: [http://www.wordinfo.info/words/index/info/view\\_unit/1/?letter=B&page=3](http://www.wordinfo.info/words/index/info/view_unit/1/?letter=B&page=3); accessed July 14, 2008.
4. Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. Int J. Supercomp App 2001;15(3).
5. Foster I. Service-Oriented Science. Science, 2005;308(5723): 814-17.