Distributed Cognition as a Solution to High Performance Computing Customer Support

--

Barbara Jennings, *Sandia National Laboratories*

ABSTRACT: *High Performance Computing users create new theories and modules based on the information that they have available to them. Within their community developers and users alike are originators. They are working with new problems, unique data sets and in most cases, forging the path for future work and new analysis. Redundancy exists in the public library sets, software modules, patches, and platforms that other previous developers have used. More likely than not, developers and analysts are reusing or repurposing some form of previously created information in at least a portion of their job. Consequently it is imperative that they have access to previous work and authors, in addition to access to software and hardware experts who may be able to provide assistance to achieving their goals.*

The technical support that this community requires is information from other experts, developers, and system administrators. What they need is access to shared knowledge. This paper documents a model for support that is based in actual experience and academic research. This design for shared knowledge is a wisdom community that provides a transitional learning environment based on the assimilated knowledge of experts. To enable such a community of information, first the knowledge must be gathered, then made accessible to others, and lastly, it must be managed to guarantee the reliability and delivery of the information.

KEYWORDS: HPC Support, Distributed Cognition, Online Collaboration, Wisdom Community

1. Shared Knowledge and Distributed Cognition

1.1 Distributed Cognition Defined

It is generally accepted that a person's cognition, their perception, understanding and reasoning exists exclusively in one's mind. Previously, little regard was given to the social, historical, physical, and artifactual ingredients that contribute to one's developed cognition. Researchers in this field have found that each of these variables contribute to one's cognition.

Gaveriel Salomon[1] advanced the idea that human cognition is not something that is solely possessed or residing in the mind of an individual. Rather it is an emerging quality created through the sharing of ideas, information, and resources. Cognition is the result of people thinking in conjunction or partnership with each other and through the use of culturally provided tools and implements. Thus suggesting the role of dynamic interaction by placing emphasis on sharing among all variables in a system. Salomon asserts that systems of distributed cognition grow and develop and that

[1] Gaveriel Salomon, PhD is a full professor at the University of Haifa in Israel. Among the many topics that he has researched is his presented work in intelligent computer tools and distributed cognition.

individuals interact with distributed cognitions independently yet with a reciprocal effect. Offering the effect of the balls on a pool table hitting each other, as illustration that distributed cognition is more than the sum of the whole and cannot be understood by examining individual parts. (Salomon, 2001)

1.2 Enabling Shared Knowledge

To enable shared knowledge, Salomon suggests partnerships, which evolve from guided simulation and qualitative scaffolding where one partner actively contributes to, or cultivates another partner's achievement. Other theorists building on this idea acknowledge that social and situational factors also have a strong impact on personal cognition, suggesting that social processes should be treated as cognitions. (Resnick[2], 1981 – as cited in Salomon, 2001).

The idea that distributed intelligence is "crafted", is offered by Roy Pea. He concludes that intelligence is manifested in activity that connects a means to ends achievement[3]. His assumption is that intelligence is the property of the mind of an individual. It is an aggregate of information distributed across other minds, persons, and environments. Asserting that distributed intelligence is socially constructed and distributed among thinking people in action, as Salomon, Pea considers the affect of tools as artifacts of distributed intelligence that may evoke new ways of contributing to activity. The community of members using the tools defines the activity.

Pea also investigates the relevance of computer technologies and the possibility of using this technology to bridge the gap between desires and actions. These technologies include the use of virtual reality, scientific visualization, and using technology to transfer data into display. He gives examples for collaborative activities and social construction processes; the linear mode becomes more of a cyclical process and the boundaries of each of the processes are blurred by cognition through technology. (Pea, Roy D., 200l Salomon, (Ed) 2001)

[2] Lauren Resnick, EdD is Director and Senior Scientist at the Research and Development Center at the University of Pittsburg.

[3] Roy Pea, PhD is a professor a professor at Stanford University and the Director of Stanford's Center for Innovation Learning. He is well known for his contribution to the scientific community including the CoVis, Project; collaborative visualization that mediates scientific knowledge through use of specific visualization tools in a collaborative learning context.

2. Technical Cooperation

2.1 Economies of Cooperation

Exploiting online technology to meet unique needs is presented in the research of Peter Kollock[4] who refers to "economies of cooperation" that are created by online communities where online interactions and exchange of information are practiced. Comparing reciprocal interaction to gifting, he distinguishes the roles that develop within a community for the giver(s) and the recipient(s) of the information. He references examples[5] as far back as the WELL and traditional Usenet demonstrating that a professional who could be paid at an hourly fee for advice to an individual off line is willing to provide the same information to an unlimited group of individuals on the Internet. This "gifting" is motivated by individual and social benefits as it improves technology. Contributors are remembered within the community by their contributions as well as their identity in the form of user name or "foot-print" that they may leave within the code. The choice of complete anonymity is also theirs. When a known contributor requests assistance from fellow developers online, fellow participants are very eager to come to their aide. This evidences the economies of cooperation.

Perl code developers who provide information to online sites such as FREEPerlCode.com evidence demonstration of cooperative sharing.[6] This site exhibits a common form of distribution for Open Source software. Illustrating a form of information and code sharing, this case is similar to public goods in that 1) it is nonrival, one person's consumption of the good does not reduce another's and 2) it is nonexcludable as it is difficult to exclude who will be able to use the good. (Kollock, 2003) Two areas of challenge that are presented with this type of online cooperation are 1) the need to motivate contributors versus attracting only consumers of the information, and 2) the challenge of coordinating the information. (Discussed further in this research.)

[4] Peter Kollock is an Associate Professor and Vice Chair of the Department of Sociology at the University of California, Los Angeles.

[5] The WELL (Whole Earth L'Lectronic Link) is a San Francisco based, online community using threaded message board discussions and collaborative site started in 1985 and Usenet groups; an informal virtual network of asynchronous online message boards originally implemented in 1979-80 by Steve Bellovin, Jim Ellis, Tom Truscott, and Steve Daniel at Duke University.

[6] www.freeperlcode.com is a site for Perl code and CGI module distribution. Contributors must identify themselves and they may contribute new modules of modify existing ones.

2.2 Tools of the Trade

Linux is the dominant development platform in the HPC world and has the foothold. Technical experts and designers do not want to leave their development platform environment to collaborate with fellow professionals. The shift here is from technologies that drive collaboration to cooperative communities that drive collaboration. The HPC environment itself offers collaborative situations through scientific visualization, simulations, and virtual environments that may be designed to enhance not impede progress. Users do not want to grapple with duelling operating system (OS) environments in order to communicate with each other. Admittedly, often times it is the case that a PC desktop or laptop running Macintosh or Windows OS sits adjacent to one running Linux order to facilitate corporate and administrative interactions.

One tool that has been used widely and adopted by the HPC community is Access Grid® (AG). This collection of resources includes multimedia large-format displays, presentation and interactive environments, and interfaces to Grid middleware and to visualization environments. AG is commonly used to provide distributed services such as training, large-scale meetings and to facilitate collaborative sessions. First developed by the Futures Laboratory at Argonne National Laboratory; widely regarded for their research in advanced communications, collaboration and visualization technologies such as telimmersion.[7] Each "site" requires specific technology service to connect to AG. Over 1500 members worldwide utilize high-end audio and visual capability to create the research environment required for the development, distribution, and collaboration of data and visualization. The focus of this "space" is on many-to-many communication rather than the individual communication afforded by desktop-to-desktop capability.

Access Grid is a good tool for many-to-many communication, but it is not a reasonable solution to address the one-to-many communication. One-to-many is more readily satisfied by tools such as email and Web based collaboration sites.

[7] Deployed by the NCSA PACI Alliance, Access Grid is still a focus of the Futures Lab at Argonne National Laboratory where research is continued to find new ways to improve the Access Grid. Telimmersion describes people meeting together in a 3D virtual environment that enables users in different geographic locations to come together in a simulated environment to interact. This technology enables users will feel like they are actually looking, talking, and meeting with each other face-to-face.

3. Motivation to Develop Community

3.1 Social Dynamics

The social community that is developed within an online site and the capability that members have to join and contribute to the community is brought to light by researcher Amy Jo Kim[8]. She asserts that members will come to a site for information, but they will continue to return to the site for community. Comparing online communities to physical communities, she emphasizes social dynamics. Three principles underlie her strategy for community requirements for a site: 1) design for change, 2) create and maintain "if" feedback loops and, 3) empower the members to have a role in the building and maintenance of the site over time. Her theory investigates the social dynamics that occur in physical communities to maintain connections, deepen relationships, and meet others who share the same interests. (Kim, 2000)

The Linux development story is an example of successful online community. In 1991 Linus Torvalds[9], released a new operating system, the Linux kernel, on the Internet. The next stage of development for this "open source" software is one of the earliest and most successful examples of online collaboration, specifically code development. The code, released as General Public release software, free of charge. Torvalds then led an ad-hoc international team of computer professionals to refine the technology. Program modules were released to the Internet along with the understood invitation to other developers to make contributions and to hone the code.

The motivation that captivated these developers was driven by both their code needs and the idea of contributing to a pursuit to take control of the operating instructions from the hands of companies that make billions of dollars by selling software preinstalled on new computers.

The Linux community 'roles' consists of: Torvalds - the Lead Developer, Andrew Morton, the "Colonel of the Kernel"[10] who performs as a conduit to maintaining the

[8]Ms. Kim is known in the field on online community as a social architect. Specializing in Visual Psychophysics, she combines programming skills with psychology to create sites and software, which encourages interactivity.

[9] Linus Torvalds released the Linux kernel as a student project while attending Helsinky University in Finland.

[10] Colonel Kernel is the role of primary software management, version control and distribution for the LINUX Kernel software. The Colonel is currently Andrew Morton, a long-time system software developer who as made many significant contributions to the LINUX Kernel since 2000.

code, a primary mailing list for code developers used to distribute messages, mirror sites - locations maintained gratis by individuals for the sole purpose of code distribution, and the unlimited number of contributors who create, modify and validate the Linux code modules. There are also a gazillion periphery sites for purchasing support, packaged releases, obtaining training, etc.

Linux kernel development illustrates a community of online knowledge sharing, controlled by respected roles and expected rules for collaboration. It demonstrates distributed cognition and the remarkable advancements that can come out of cooperative technologies.

Linux further exemplifies collaboration, motivation, and community in the technical arena. Collaboration, primarily via asynchronous communication, makes use of documentation and one-to-many exchanges all in one simplified package. Experts and new contributors alike are motivated to contribute to Linux. This effort is largely about experimentation, research, trial and error, and participation in a community. Roles that assimilate out of such collaboration are well defined. They are performed without payment, generally resulting in the respect among peers.

4. Coordination of Information

4.1 Knowledge Management

Knowledge management of the information that is provided online includes access to the site. The site must contain a reasonable level of security for the data Additionally the information must have a high rate of reliability if it is expected to be sought out by other experts. As with the physical community, the site must provide maps or directions, as required by the users and the information must be easily navigable. Most importantly, the information must be directly applicable to the site's reason for presence at all.

In the case of HPC Support the information contained within should be directly related to the users task at hand. It should not contain any superfluous information and should not distract the user with such errands as finding and loading a browser plugin. Information should be structured in a way that is familiar to the user; categorically, sharing a known taxonomy, and it may even mimic the physical community that is serves. Use of the tool such as searching for information and general interaction with the site must be intuitive.

Gathering the information requires careful consideration. It must be easy and inviting for a user to add information to the community. When possible such information gathering should be designed to occur in conjunction with the user's work. Quality control of the submitted information must be included in the design. Volunteer experts may perform review of content if the site is not for profit.

Certifying the validity of information after it has been approved for release poses an interesting challenge for knowledge management. One solution is to date the information. This provides a caveat declaring that the data was valid at a particular date. A schema for organization may also include computer platform, operating systems or application information to provide greater level of granularity.

4.2 Knowledge Management Tools

Using tools that are familiar to the users is another important matter for consideration. Natural language searches with extended Boolean capability for example are well known in this computer community. Most HPC users would not welcome required sql or some other form of database query code to perform searches.

5. Collaborative Learning, Information, and Knowledge: CLIK

Sandia National Laboratories (SNL) and Cray, Inc. the system developer are on schedule to release Red Storm, the next massively parallel supercomputer. The support required for the High Performance Computing (HPC) environment must address several concerns: 1) the dynamic landscape that is constantly changing and actually being created on a daily basis, 2) customers are often the experts, 3) problem solving requires research into areas that may or may not have previous record, 4) users are local as well as remote and 5) support must be intraoperative among the support teams from local sites as well as the remote host locations.

5.1 Specific Needs

The users of this system require access to timely information gathered in a dynamic environment from a wide-range of information resources. To meet the needs of this unique environment, a model for HPC customer support has been designed by the Scientific Computing Department at SNL. This design incorporates an online collaborative support environment and is supported by advanced academic research using the FOCAL[11] model that was developed for online collaborative learning as a basis. (Gunawardena, et al. 2004) The SNL model is referred to as Collaborative Learning, Information, and Knowledge; CLIK.[12] The software interface design

[11] Focused Outcome Centered Around the Learner. Developed at UNM in the Organizational Learning and Instructional Technologies Department.
[12] CLIK Project Lead/Software Designer, Barbara Jennings, Member of Technical Staff, Sandia National Laboratories.

follows two methodologies, Minimalist[13] and Flow[14] Minimalist is a design specifically for experience computer users. Flow was developed to create an environment of concentration, one that does not divert the user's attention off of the task at hand.

In brief, this environment is a shared wisdom community, making available, the experience and knowledge of the users, designers, application developers, system administrators, and vendors for the HPC environment among three national laboratories and their extended user base. The customer scope requirement mandates secure access by local and remote customers to multiple previously established information resources and the creation of a resource to capture ongoing knowledge as it is developed.

5.2 A SocioTechno Community

The goal of CLIK is to create a culture of shared information gathering and exchange, for users of a community of HPC systems. In effect we are facilitating the ability to construct technical knowledge through social collaboration. We provide common access and usability across a wide range of skill and experience levels, and a mechanism for the HPC community to actively learn and share their knowledge. In the capability platform realm today, there are many more questions than answers. Tomorrow, the knowledge integration and dissemination acquired through CLIK will provide user access to answers these questions. As shown in Figure 1, this model provides access to a shared knowledge community. This community is an aggregate of share online resources, and access to expert knowledge.

This design, based on information gathering, dissemination, and management, results in a collaborative knowledge resource that doubles as a tool to be utilized by users who need to research information and consultants who need to research and compile available problem solving techniques. The CLIK model provides interfaces to: experts (via telephone, email, and web access), online training, and asynchronous collaboration areas. Thus enabling user transformation from novice to expert.

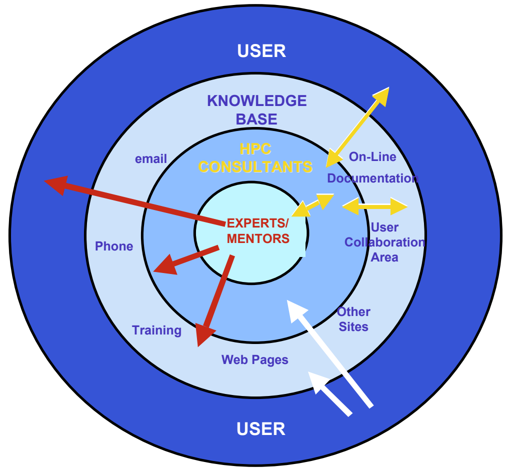

Figure 1. CLIK - Collaborative Learning Information and Knowledge model for High Performance Computing support at Sandia National Laboratories

5.3 Motivation

CLIK enables a user the ability to search existing knowledge using a standard web browser. It provides a tool for documentation to groups of individuals who need a shared location to access knowledge gathered. Email messages to consultants are tracked automatically to monitor the status of user requests for assistance while providing the standard problem tracking techniques of vendor supplied applications. Requests for assistance by means of the web based application are monitored and managed through a sql database and custom code. CLIK also makes use of the corporate capabilities such as: authentication, secured access, user identity, and account information customized to the SNL environment.

A novice will find information on how to get started on a system, along with access to expert knowledge via a familiar search appliance. Access to experts for additional technical support is provided as well. The goal is to provide the path for a novice to learn and the opportunity for them to eventually add back to the community through their contributions.

Project teams are motivated to participate as information may be shared using need-to-know access security that may be applied to information as the information owner decides. Developers and analysts with developed sites or databases resources may also share their information via a well know "Google" search applicance.

This methodology is necessary to create, grow, and control an information-based site. Following this design principle ensures controlled community formation within a site where the member can have an expected outcome from interacting with the community.

---

[13] John M. Carroll's framework of learning for experienced computer users, completed while at the Department of Computer Science at Virginia Tech University, 1998.
[14] Mihaly Csikszentmihaly, Flow Theory for which he received Thinker of the Year Award in 2000.

6. Where do we go from here?

"Build it and they will come", the field of dreams marketing strategy needs an addendum if it is to apply to online sites that promote collaboration. The statements needs to include the message that "allowing the members to 'add-on' will keep them coming back." Contribution to the community gives members a sense of identity and belonging. Encouragement for a member to define their identity via footprints, tags or a personal web site within the CLIK community, contributes to the development of the community. Allowing members to provide feedback and add to the information on the site enables joint responsibility. This further ensures that rules will be followed, as members understand that the "rules" are necessary to ensure cooperation.

Successful online communities promote shared knowledge. The resulting distributed cognition, which initially defines the individual member's intelligence gain from the information, also defines the benefit to the community at large. This form of community provides support to the HPC user through enabling information sharing and providing access to such information.

The tools that are provided need to facilitate collaboration within the development environment if they are intended to enable technical cooperation. Tools should be focused on the effort at hand if they are to foster the desired collaboration. There is much to be gained from the economies of collaboration by the individual as well as the community as a whole. Use of a good is not diminished if it is online information rather its value is enhanced. Individuals are motivated to contribute knowledge for personal satisfaction and cultural gain.

The area of online collaboration is wide open to research and development. Understanding how to make the most use of tools and information to facilitate individual's collaboration and sharing of knowledge online is an ever-moving target. Each industry has unique needs and requirements, which are always changing and continue to challenge and online community to develop. Communities must be designed with flexibility to enable the required growth. Tools that are not useful or well understood by the members will not be used and may cause a member to retreat from the community altogether. This difficult task is eased by the fact that intelligence is not gained in isolation and that individuals will search out information or create it as the need arises.

7. Acknowledgments

The author would like to thank the Sandia Management Team of 9300 and 9900 for the opportunity to develop this innovative work, the professors in the OLIT department and her committee for their thoughtful critique and suggestions. Additionally the Sandia Scientific Computing Support team who continue to contribute to the design, development, and running of CLIK, our developer, the users, and system administrators all for being flexible to new solutions and open to change.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

8. About the Author

Barbara Jennings is a Member of Technical Staff at Sandia National Laboratories in the Scientific Computing Systems department. She is completing her doctoral studies at the University of New Mexico where she has her directed study to Online Collaboration for Technical Users. Ms. Jennings has an active 13-year history in the field of Internet applications and High Performance Computing at Sandia National Laboratories and contributions to the Internet Engineering Task Force.

References

Argonne National Laboratories. (2005) The Access Grid Project, *Access Grid Newsletter March 2005*. Retrieved on April 25, 2005, from http://www.accessgrid.org/

Gunawardena, C. N., Jennings, B., Ortegano-Layne, L., Frechette, C., Carabajal, K., Lindemann, K., & Mummert, J. (2004). Building an online wisdom community: a transformational design model. Journal of Computing in Higher Education, 15(2), 40-62.

Kim, Amy Jo (2000) Community Building on the Web. Peachpit Press, Berkeley, CA.

Rheingold, H. (2003) Smart Mobs, The Next Social Revolution. Perseus Publishing, Cambridge, MA.

Smith, Marc A. and Kollock, Peter (Eds.) (2003). Communities in Cyberspace. Routledge Publishing, New York, NY.

Salomon, Gavriel (Ed.), (2001). "Distributed Cognitions Psychological and Educational Considerations". Cambridge University Press, Cambridge, UK.

Streicher, Martin (2003, September 15) The Colonel of the Kernel: Feature Story. LINUX Magazine. Retrieved from http://www.linux-mag.com/content/view/1437/2211/ on April 24, 2005.