Editor: Geoffrey Fox, gcf@indiana.edu

CLUSTERING SOFT-DEVICES IN THE SEMANTIC GRID

By Hai Zhuge

SOFT-DEVICES ARE A PROMISING NEXT-GENERATION WEB RESOURCE. THEY ARE SOFTWARE MECHANISMS THAT PROVIDE SERVICES TO EACH OTHER AND TO OTHER VIRTUAL ROLES ACCORDING TO THE CONTENT OF THEIR RESOURCES AND RELATED configuration information. They can contain various kinds of resources such as text, images, and other services. Configuring a resource in a soft-device is similar to installing software in a computer: both processes contain multistep human–computer interactions. Soft-devices combine the advantages of the active and intelligent features of intelligent agents, the semantics-based features of the Semantic Web, the configurable feature of hard devices, the benefits of abstraction, and the notion of the Semantic Grid. Although soft-devices are still in the research stage, they are easily implemented.

A soft-device "world"

A soft-device "world" is a high-level virtual community consisting of three roles: producer, consumer, and soft-device. Producers add content to and configure soft-devices. Consumers can post their requirements, select the services they need using a browser (a soft-device), and subscribe to long-term services offered by soft-devices. Soft-devices can play both the producer and the consumer role. They can accept content definition from multiple providers and actively provide services for multiple consumers by seeking requirements under a certain cost evaluation. People can play the role of either producer or consumer.

The soft-device world works on the Semantic Grid (www.semanticgrid.org), an Internet application platform that will be able to normally organize, effectively share, intelligently and dynamically cluster, and manage globally distributed resources in a uniform way.[1] Its major characteristics are semantic connectivity, dynamic clustering, and normal organization. *Normal organization* requires orthogonal semantic normal forms that can classify versatile resources appropriately.[2] The Semantic Grid will be a new computing and communication platform combining the advantages of client-server and peer-to-peer networking. Soft-devices can dynamically cluster resources by way of a single semantic image—that is, unifying various resources at the semantic level.

The soft-device world has no central controls; it is entirely peer-to-peer. It also includes no hardware: it regards any Internet-accessible hardware as a soft-device providing the same service. The operation of a hardware device is implemented by a corresponding soft-device.

Architecture

A soft-device consists of six major components:

- A container stores the content in a machine-understandable way, perhaps according to the markup languages and ontology mechanisms proposed by the Semantic Web community (www.semanticweb.org).[3]
- A *detector* is responsible for detecting the requirements in the requirement list published on the Semantic Grid.
- An *explainer* explains the resources' content.
- *Multiple built-in workflows* enable the soft-device to work according to different requirements. The workflow should be time sensitive[4] and adaptive.
- A *knowledge base* supports the explainer, the detector, and adaptive workflows.
- An *interface* supports the producer in defining resources' content. A browser is a special soft-device that conveys the output of the explainer to the consumer.

Sharing Web pages versus sharing soft-devices

Figure 1 compares Web-based information sharing and soft-device sharing. In the former, when Web pages are produced and browsed, semantic loss and distortion occur for two reasons:

- Current techniques do not let developers add machine-understandable semantic information to Web pages.
- Web browsers and search engines cannot understand the semantics of HTML-based pages.

Also, multipurpose resources would

have to be transformed into Web pages in order for them to become available via browsers, but this doesn't work. Moreover, sharing Web pages means that users and browsers must know the pages' locations.

Sharing soft-devices, on the other hand, has five major advantages:

- Producers can add semantic information to soft-devices using definition tools.
- The contents of a soft-device are organized in normal forms that guarantee the correctness of content-based operations.[5]
- Soft-devices are interconnected at the semantic level, so their semantics are not isolated and operations can be carried out only according to the semantics, regardless of the resources' form and location. That is, the soft-devices world is a single semantic image.
- Browsers are also soft-devices; they can manage resources under the privilege limitation in the single semantic image.
- Soft-devices can automatically cluster on demand in the best way possible and can then intelligently push the clustered services to consumers.

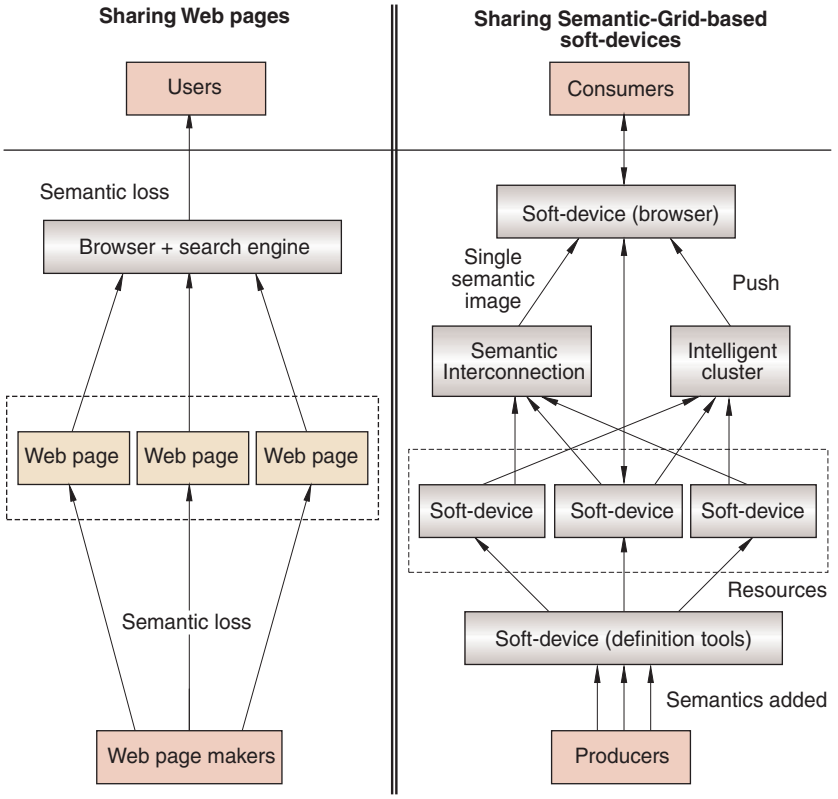

Figure 1. Web page sharing versus soft-device sharing.

Dynamic clustering

Figure 2 shows an approach to the dynamic clustering of soft-devices. Workflows between soft-devices are predefined according to consumers' business process requirements. Consumers can post their requirements (activities in the workflows) on the published requirements list. Soft-devices actively and intelligently search the list and inform the broker (another soft-device) if they find requirements that match their capabilities. The broker selects the best soft-devices among several candidates, clusters particular soft-devices to provide an integrated service, or adapts existing clusters to respond to changes in consumers' requirements. The knowledge in soft-devices should also support the adaptation of their services to best suit the work of the cluster. Knowledge flows can be formed in a soft-device cluster to realize effective knowledge sharing and management.[6,7]

Figure 2 also shows the differences among Web-based searching, soft-device clustering, and the normal-organization approach. Web-based searching uses search engines to explore the Web's limitless and still-expanding resource space (1-to-*m*), so it cannot be considered efficient or effective in terms of accuracy. There are also no mechanisms to remove useless and out-of-date Web pages. In contrast, soft-device clustering (*m*-to-*m*) is, in a way, the reverse of Web searching—that is, resources find people rather than people find resources. By removing the matched and out-of-date requirements in the list, this approach can limit the search space, thus satisfying search efficiency and efficacy. People, represented by the person in Figure 2 lying on the globe, can enjoy the services provided by soft-devices. Finally, the normal-organization approach guarantees the correctness and accuracy of content operations by means of normal forms,[2,5,8] just as the normal forms of relational databases guarantee the

Producers can add semantic information to soft-devices using definition tools.

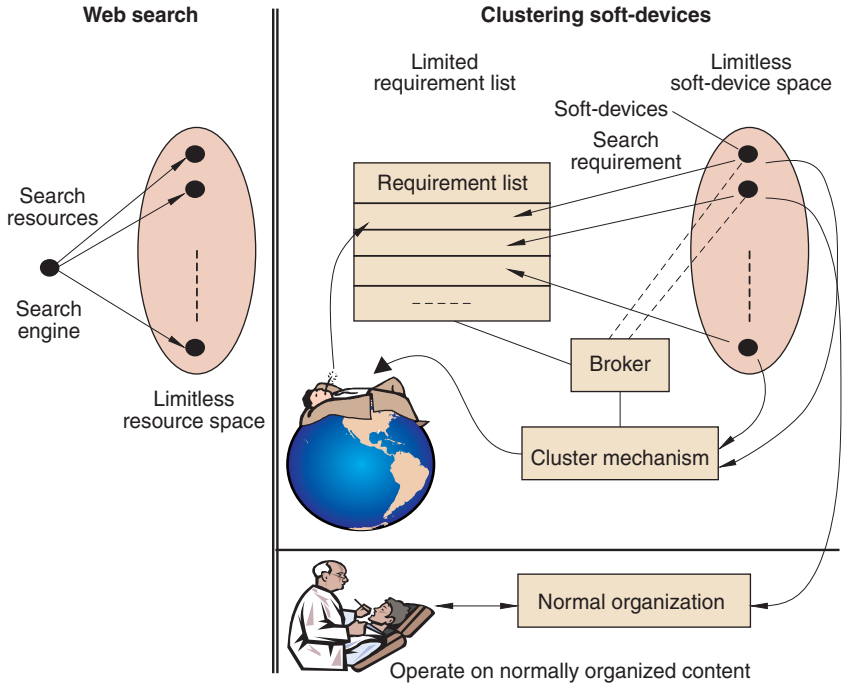

Figure 2. Web-based searching, soft-device clustering, and normal organization.

correctness of data operations. Normal organization lets consumers access and use content through soft-device–like database management—as easily as a dentist operating on the organized content (the teeth) in a patient's mouth (see Figure 2). An intelligent Semantic Grid should adopt one or both of the latter two approaches. CiSE

References

1. H. Zhuge, "Semantic Grid Research Action in China," to be published in *Comm. ACM*, 2003.
2. H. Zhuge, "VEGA-KG: A Way to the Knowledge Web," *11th Int'l World Wide Web Conf.*, poster session, ACM, New York, 2002, www.2002.org.
3. J. Hendler, "Agents and the Semantic Web," *IEEE Intelligent Systems*, vol. 16, no. 2, Mar./Apr. 2001, pp. 30–37.
4. H. Zhuge, T.Y. Cheung, and H.K. Pung, "A Timed Workflow Process Model," *J. Systems and Software*, vol. 57, no. 3, July 2001, pp. 231–243.
5. H. Zhuge, "Resource Space Model, Its Design Method and Applications," to be published in *J. Systems and Software*, 2003.
6. H. Zhuge, "A Knowledge Flow Model for Peer-to-Peer Team Knowledge Sharing and Management," *Expert Systems with Applications*, vol. 23, no. 1, July 2002, pp. 23–30.
7. H. Zhuge, "Distributed Team Knowledge Management by Incorporating Knowledge Flow with Knowledge Grid," *Proc. 2nd Int'l Conf. Knowledge Management*, Springer-Verlag, New York, 2002, pp. 218–223.
8. H. Zhuge, "A Knowledge Grid Model and Platform for Global Knowledge Sharing," *Expert Systems with Applications*, vol. 22, no. 4, May 2002, pp. 313–320.

Hai Zhuge is a professor of computer engineering at the Institute of Computing Technology in the Chinese Academy of Sciences and the leader of the China Knowledge Grid Research Group (http://kg.ict.ac.cn). His research involves the concepts and modeling of the Knowledge Grid and the Intelligent Semantic Grid. Contact him at Zhuge@ict.ac.cn.