So Where's the Computation Superhighway?

Olivier Marin, Laboratoire d'Informatique de Paris

Let's reflect on the World Wide Web's history for a moment. The legend goes that it originated from the US Air Force's early notion of fault tolerance. They figured they wouldn't be happy if they lost a city to a nuclear missile, but they'd be even unhappier if that loss disabled their capacity to hit back with loads of their own missiles. The answer came swiftly: multiply the available resources, disseminate them throughout the country, and interconnect them and ensure the same computations would be workable on any location.

As the original Arpanet evolved, it opened to research communities and then to the general public. As the number of connected computers grew, so did the amount of resources. Nowadays this growth is nearly exponential, and many countries are still waiting to popularize Internet access. What we're arriving at is a planetary proliferation of networked equipment surrounding us that can run programs cooperatively. Starting from this conception, and putting aside a few technicalities for a jiffy, the WWW appears to be an unlimited pool of computing resources.

A huge pool, but not that much fishing

Where are all these resources going? This question automatically arises every time I launch a distributed program over the dozens of nodes my lab offers access to. Is someone using them overnight while nobody's looking? What's the point of all those NOPs (no operations) our computer processors keep performing? Do they have some utility we haven't figured out yet? Maybe it's some kind of message from a wider perspective: "Eventually what everyone does is a bit vain." Unfortunately, a more obvious answer to where the Internet's huge potential sum of
resources is going is down the drain. Then, the message from a wider perspective becomes, "Well, folks, just check out how much energy those little critters you created are wasting; now do you understand why I decided to take a break on that seventh day?"

There are several fair reasons for such a situation of "underexploitation":

- Alice just won't lend her PC to anyone.

- Bob tends to be more altruistic, yet wishes to remain selective and to choose whom he'll provide resources to.⁴

- Carla wreaks havoc as soon as she gets access somewhere—hence fully justifying both Alice's and Bob's prejudice.

- Dan would help, but he'd rather make money out of it at the same time.

- Ellen would also help, free of charge as far as she's concerned, but has no clue as to how she might share her PC resources. Ellen was enjoying a nice trip to the moon while the hype around SETI@home (http://setiathome.ssl.berkeley.edu) rose and then died, or more likely she's part of the majority of healthy people who don't spend so much time on Slashdot (http://slashdot.org) or on IEEE Distributed Systems Online.

Every one of these behaviors might partly explain why cooperative computations over open distributed networks have yet to emerge. There is one exception that can already be observed nowadays: Dan's stance is pretty much that of firms with expensive hardware willing to rent out their Web services. Such firms have invested a lot, and it seems all too normal that they should get restitution for the use of their resources. Provided the cost remains reasonable, such Web services do look very promising. However, you can't expect small firms or research facilities with low budgets to rent resources whenever the requirement arises. Yet this doesn't have to mean that short-funded teams should downsize their ambitions.

**Where the water's deep enough**

As it is, computer science research groups aren't the best endowed with computer resources. (For proof of this, visit TOP500 Supercomputer Sites, http://www.top500.org.) Most grid structures and supercomputers are owned by groups concerned with physics, biology, and other scientific areas involving massive data processing. Overall in computer science communities, real wide-scale experiments—as opposed to simulations and emulations—must wait until all
physicists, biologists, and such have decided to go for a nice, long walk in the countryside. Computer scientists are making some effort to share expenses and build common computer networks for their own experimentation needs. One such effort is Grid'5000 (http://www.grid5000.org), a French national project that aims to build an experimental test bed for grid-computing research. As far as this column is concerned, the most interesting aspect of Grid'5000 is the empirical study of the feasibility of large-scale desktop grids.

Taking this effort a step further could mean tapping into the ocean of resources that the WWW constitutes. The SETI@home example showed that individuals are ready to lend their resources, as long as a few conditions are respected: their security and privacy must be ensured, and they must feel involved in the project's long-term goal. Agreed, the latter will be tough to obtain: it's hard to imagine my neighbor getting thrilled with the idea of checking whether some new failure detection protocol scales well. Maybe it will take some kind of fishy incentive, an equivalent to file sharing for peer-to-peer networks. The exciting prospect is that this approach will motivate solutions for deploying and maintaining computations over large-scale open networks, while upholding strong guarantees as to dependability.

I firmly believe that the agent paradigm and multiagent systems represent a shrewd way of designing such solutions. Once the right environment has been downloaded and installed on a host, agents can access its resources—those the owner decided he or she could spare—and exchange data with other hosts and agents.

Of course, this raises sensitive issues. For example, hosts ought to be protected from malicious agents, and agents should be safeguarded from malevolent hosts. Similarly, given that hosts might crash or be rebooted at any time, some fault tolerance scheme is required. Agents certainly don't solve all the issues, but they definitely seem to present appropriate characteristics for large-scale, open environments.

**CONCLUSION**

Exploiting the huge amount of resources that sit idly on computers over the WWW sounds like the redistribution of goods without leveling everything to equality at the expense of comfort. Eventually the message from a wider perspective could become, "See? All you need is to get them all to cooperate, and there'll be nothing you can't achieve." Despite such a message's overwhelming cheesiness, don't you find it inconspicuously appealing?
References


Olivier Marin is an associate professor at the Laboratoire d'Informatique de Paris 6. Contact him at olivier.marin@lip6.fr.

Cite this article: Olivier Marin, "So Where's the Computation Superhighway?," IEEE Distributed Systems Online, vol. 6, no. 1, 2005.