

The end of sneakernet?

July 13 2017, by Cory Nealon



Credit: University at Buffalo

Not everyone marvels at the speed of the internet.

For researchers and companies sharing extremely large datasets, such as genome maps or satellite imagery, it can be quicker to send documents by truck or airplane. The slowdown leads to everything from lost

productivity to the inability to quickly warn people of natural disasters.

The University at Buffalo has received a \$584,469 National Science Foundation grant to address this problem. Researchers will create a tool, dubbed OneDataShare, designed to work with the existing computing infrastructure to boost data transfer speeds by more than 10 times.

"Most users fail to obtain even a fraction of the theoretical speeds promised by existing networks. The bandwidth is there. We just need new tools that take advantage of it," says Tefvik Kosar, PhD, associate professor in UB's Department of Computer Science and Engineering, and the grant's principal investigator.

Large businesses, [government agencies](#) and others can generate 1 petabyte (or much more) of data daily. Each petabyte is one million gigabytes, or roughly the equivalent of 20 million four-drawer filing cabinets filled with papers. Transferring this data online can take days, if not weeks, using standard high-speed networks.

This bottleneck is caused by several factors. Among them: substandard protocols, or rules, that govern the format of how data is sent over the internet; problems with the routes that data takes from its point of origin to its destination; how information is stored; and limitations of computers' processing power.

Rather than waiting to share data online, individuals and companies may opt to store the data on disks and simply deliver the information to its destination. This is sometimes called sneakernet—the idea that physically moving information is more efficient.

Managed file transfer service providers such as Globus and B2SHARE help alleviate data sharing problems, but Kosar says they still suffer from slow transfer speeds, inflexibility, restricted protocol support and other

shortcomings.

Government agencies, such as the NSF and the U.S. Department of Energy, want to address these limitations by developing high-performance and cost-efficient data access and sharing technology. The NSF, for example, said in a report that the cyberinfrastructure must "provide for reliable participation, access, analysis, interoperability, and data movement."

OneDataShare attempts to do that through a unique software and research platform. Its main goals are to:

- Reduce the time needed to deliver data. It will accomplish through application-level tuning, and optimization of Transmission Control Protocol-based data transfer protocols such as HTTP, SCP and more.
- Allow people to easily work with different datasets that traditionally haven't been compatible. In short, everyone's data is different, and it's often organized differently using different programs.
- Decrease the uncertainty in real-time decision-making processes and improve delivery time predictions.

OneDataShare's combination of tools—improving data sharing speeds, the interaction between different data programs, and prediction services—will lead to numerous benefits, Kosar says.

"Anything that requires high-volume data transfer, from real-time weather conditions and [natural disasters](#) to sharing genomic maps and real-time consumer behavior analysis, will benefit from OneDataShare," Kosar says.

Provided by University at Buffalo

Citation: The end of sneakernet? (2017, July 13) retrieved 7 July 2024 from <https://phys.org/news/2017-07-sneakernet.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.