

IU 'Twister' software improves Google's MapReduce for large-scale scientific data analysis

March 16 2010

(PhysOrg.com) -- "Twister," a new software tool released by Indiana University, supports faster execution of many data mining applications implemented as MapReduce programs. Developed by researchers from the Pervasive Technology Institute at IU, the tool extends the functionality of MapReduce, a distributed programming technique patented by Google for large-scale data processing in datacenter environments.

Twister allows MapReduce to achieve higher performance, perform faster data transfers, and reduce the time it takes to process vast sets of data for data mining and machine learning applications.

"MapReduce is an exceptionally valuable tool for finding meaning in very large scientific data sets," said Xiaohong "Judy" Qiu, Associate Director of the Community Grids Lab within the PTI Digital Science Center and lead on the project (Service Aggregated Linked Sequential Activities, or SALSA) that produced the Twister [software](#). "Twister makes MapReduce even more powerful for data-intensive disciplines such as physics, chemistry and the medical and life sciences."

Applications that currently use Twister include: K-means clustering, Google's page rank, Breadth first graph search, Matrix multiplication, and Multidimensional scaling. Twister also builds on the SALSA team's work related to commercial MapReduce runtimes, including Microsoft

Dryad software and open source Hadoop software. SALSA project work is funded in part by an award from Microsoft, Inc.

"Twister is especially effective for applications with iterative MapReduce Computations," said Jaliya Ekanayake, lead developer on the Twister project. "The architecture is based on pub/sub messaging that enables it to perform faster data transfers, minimizing the overhead of the runtime. Also, the support for long running processes improves the efficiency of the runtime for many iterative MapReduce computations."

To access these papers or to learn more about Twister, please visit www.iterativemapreduce.org .

To watch a video about Twister, please visit pti.iu.edu/video/twister.

Provided by Indiana University

Citation: IU 'Twister' software improves Google's MapReduce for large-scale scientific data analysis (2010, March 16) retrieved 3 May 2024 from <https://phys.org/news/2010-03-iu-twister-software-google-mapreduce.html>

<p>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.</p>
--