

# Pervasive DataRush™: A New Weapon to Help Win the Cyber Security War

## Cyber Security: The Key to Winning is in the Data

Use your data to vastly improve your ability to affect outcomes

The threat of cyber attack (warfare, terrorism, or criminal) has been identified by the United States government as one of the most dangerous issues facing our nation and its economic infrastructure. Every year, the number of attacks on US information systems increases over the prior year's average. The U.S. Department of Defense, U.S. Department of Commerce, Homeland Security and NASA are some of the organizations that have been attacked. Pervasive Software is committed to provide innovative solutions to support those fighting the cyber security battles. Pervasive DataRush is the latest in a line of next generation data processing tools that provide the speed, scalability and agility to meet these ever growing "Big Data" challenges.

## Pervasive DataRush: Solving Cyber Security Problems

The Bigger the Data – The Bigger the Advantage

Pervasive DataRush is a proven, next generation software platform that provides **game changing speeds** and **dynamic scalability** for applications that process big data or perform deep analytics. It's a true **force multiplier** in the cyber security war.

**Game Changing Speeds** – Applications running Pervasive DataRush typically perform at **30X** or greater over existing implementations by fully utilizing the power of single or clustered multicore systems.

Dynamic Scalability – Applications written on Pervasive DataRush will dynamically scale within a multicore system with no code changes. For example, an application written on an 8-core system will automatically scale to fully consume all resources when installed on a 24-core system. This future proofing of applications means that organizations can simply add more cores to keep up with growing data volumes.

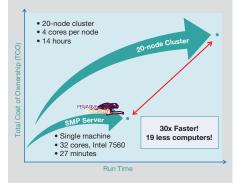
**Extensibility** – Pervasive DataRush is a single, extensible platform that allows organizations to ingest, enhance and analyze **ALL** of their data. This single platform approach greatly simplifies solution architectures, reduces the number of discrete applications and lowers overall total cost of ownership. Programmers can easily add extensions to the Pervasive DataRush libraries to help meet unique organizational challenges.





#### **MalStone B Benchmark**

The MalStone benchmarks have been developed by the Open Cloud Consortium as a means of assessing data intensive application performance when run on hosted hardware in a cloud. A common framework often cited for use with the MalStone benchmarks is Apache Hadoop, an open-source software project for reliable, scalable, distributed computing and storage. Performance results reported by Grossman on a cluster of 20 nodes with 4 cores each show Hadoop MapReduce was able to process 100 billion 100-byte records in a little over 840 minutes.



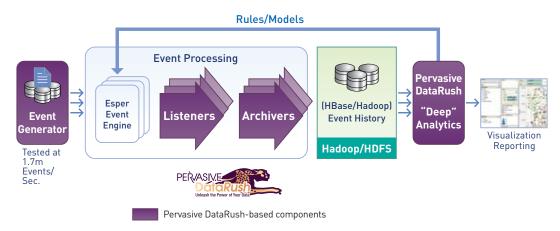
Pervasive ran the same MalStone
B benchmark using Pervasive
DataRush on a single machine
(node) with 24 cores (a 4 socket,
2.6 GHz Intel 7560 system) in about
27 minutes – more than 2 Tb per
hour processing speed. Pervasive
DataRush yielded a >30x speedup
using 19 fewer computers (nodes).
With additional cores, the observed
speedup with Pervasive DataRush
would be even more dramatic.



## Capturing ALL the Data - Historical Event Processing

One of the biggest challenges facing organizations in the cyber war is to capture and extract actionable knowledge from *all* the data that's being created. In this use case, 100 Million Syslog events were processed and loaded in only 58 seconds – a rate of **1.7 million events per second – orders of magnitude faster than any solutions on the market today.** Pervasive DataRush was used to capture *ALL* events that passed through Esper engines and then loaded them into HBase, the Hadoop database, at speed. A single 48 core machine was used in this case. The processing and loading rates will dynamically scale up as more cores or machines are added.

Once the data is loaded, Pervasive DataRush is used to perform deep analytics to discover previously unknown correlations, patterns, etc. The extensibility of Pervasive DataRush allows application specific algorithms to be applied to the data. This overall capability now give organizations the ability to use *all* of their data to vastly improve their ability to affect outcomes.



### The Pervasive DataRush Difference

- · Ability to use thousands of variables in models
- Eliminate need to sample data
  - Full utilization of multicore system provides game changing throughputs for fine grain parallelism
- No observed memory limitations
  - Data flow architecture overcomes memory heap sizes
- Dynamic Scalability more cores = more speed
- Extensibility
  - Reduce the number of discreet applications
  - Easily add to libraries to meet unique organizational needs
- Use Industry Standard JVM Languages
  - Java, Python, JRuby, etc.
  - No specialized parallel programming skills required
    - Pervasive DataRush Platform handles multi-threading, deadlocks, etc.
- Reduced processing cost per watt

#### **About Pervasive Software**

For over two decades, Pervasive has helped thousands of customers around the world in every industry to manage, integrate and analyze their data. Pervasive DataRush is an embeddable software platform solution for data-intensive processing applications such as claims processing, risk analysis, fraud detection, bioinformatics and life sciences.

#### Contact Us

www.pervasivedatarush.com 1-866-980-RUSH (7874) 1-512-231-6818 (Main Office)

info@pervasivedatarush.com

©2011 Pervasive Software Inc. All rights reserved. All Pervasive brand and product names are trademarks or registered trademarks of Pervasive Software Inc. in the United States and other countries All other marks are the property of their respective owners.