Big Data Analytics Research in
ASTRI-HP Information Technology Research Centre

Dr. YAN Shi Xing
Principal Investigator, ASTRI-HP Information Technology Research Centre
A new era of Internet of Things (IoT)

2013

Every 60 seconds

- 98,000+ tweets
- 695,000 status updates
- 11 million instant messages
- 698,445 Google searches
- 168 million+ emails sent
- 1,820TB of data created
- 217 new mobile web users

Growing Internet of Things (IoT)

Pervasive Connectivity | Smart Device Expansion | Explosion of Information

By 2020

- 30 Billion (1)
- 40 Trillion GB (2)
- 10 Million (3)
- ... for 8 Billion (4)

A new style of IT required for IoT solutions


© Copyright 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
IoT solutions demand the new style of IT

Moore's law

Energy Consumption by IoT

Comprehensive solutions

Hardware  Software  Services
Big Data Analytics need

End Users
- Domain Expertise
- Defines the requirements of big data analytics

Data Scientists
- Hard to find
- Limited domain knowledge
- Specialty in Data Mining, Machine Learning, Statistics

Time and Money needed for new Data Analytics requirement

End Users Domain Experts

Software Developers Easy to Find

Software Developers

End Users Domain Experts

Data Scientists
### Big Data Professionals

**Hadoop:** Open-Source software framework that supports data-intensive distributed applications,

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup Cluster</td>
<td>Manually install OS and set up connectivity in each node.</td>
</tr>
<tr>
<td>Deploy Applications</td>
<td>Manually deploy and configure applications in each node.</td>
</tr>
<tr>
<td>Create applications</td>
<td>Write programs using APIs or scripts</td>
</tr>
</tbody>
</table>

[Diagram of Hadoop workflow]

© Copyright 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Hadoop Practitioner

public class MapReduceFunction {
    public void MapFunction(...) {
        // MapReduce code here...
    }
    public void ReduceFunction(...) {
        // MapReduce code here...
    }
}

Hadoop MapReduce code in Java

... AND MORE ...

© Copyright 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Bamboo Eco-system for Big Data Analytics

- Simplify the data analytics process
- Shortening development lifecycle
- Shared Data Scientist resources
- Less programming

Bamboo

App Store for Data Analytics

Workflow Designer

Data Scientists

End Users
Domain Experts

Data Visualizer

Software Developers
ASTRI-HP Information Technology Research Centre

Officially Launched on 16 Dec 2013
Located in Hong Kong Science Park
Big Data Analytics R&D
Thank you

syan@hp.com