WANG, XU

| CONTACT INFORMATION | 104, Building East 2, Jimenli, Haidian District, Beijing, China 100088 | *Mobile:* +86 13911635527 *Email:* gnawux@gmail.com *Homepage:* http://wangxu.me/blog/ |

SUMMARY

Having nearly 4 years experience on Hadoop HDFS development; Core developer of BigCloud team of China Mobile Research Institute; Focusing on metadata management of NameNode and volume management of DataNode in recent projects; Actively participating Hadoop community activities.

Having been using and maintaining Linux system for more than 10 years; Developed with C/C++, Java, Python, Perl, and Skillful in Unix utilities such as sed, find.

OBJECTIVE

Jobs on *Apache Hadoop* related development and mantaince, especially HDFS related; Or R&D Job about distributed storage systems in Beijing.

CURRENT OCCUPATION

Hadoop Related Work

Researcher and *Project Manager* (2007 – present) on Cloud Computing in *China Mobile Research Insititute*, being responsible of *Hadoop HDFS* related development and maintenance, and communicate with upstream community, includes evaluation and testing, bug fix, and feature development (Apache JIRA ID gnawux):

- 2010–2011, Feature of DataNode volume refreshment `HDFS-1362`;
- 2008, Fixed a critical bug of Hadoop 0.18, which may lead to replica mistaken deleting: `HADOOP-4742`;
- Aprial 2008–present, Maintaining the CMRI distribution of Hadoop, including NameNodes synchronization patches:
  `http://github.com/gnawux/hadoop-cmri`
- 2008–2009, Maintaining the CMRI Hadoop stress testing tool (Python/Java):
  `http://code.google.com/p/hadoop-test/`
- September 2010, Present in Hadoop in China (HIC 2010), introduced Hadoop related developments in China Mobile Research Institute:
  `http://wangxu.me/blog/p/445`
- November 2009, Present on Apache Roadshow Asia 2009 (Nov, 2009), introduced Hadoop HDFS development in CMRI.
- Update the Unofficial CMRI Hadoop Team Blog: http://gnawux.info/hadoop/
- Weblog on Hadoop related developments: `http://gnawux.info/hadoop/`

Other Work

- *Linux Cluster Management*, Since 2010, Having been responsible for system management of the cloud cluster of China Mobile Research Institute, including the configuration maintenance, optimization of virtualization system and Hadoop platform;
- *Cloud Storage*, Since 2010, as Project Manager and Architect, Having been responsible of Cloud Storage Project of China Mobile Research Institute;
- *Book Writing*, Being responsible for Book of CMRI on Cloud related Technologies;
- *Translation*, Translating *Cassandra: The Definitive Guide* for Turing Book (expect to finish in April, 2011) in spare time;

PAST PROJECTS
AND
EXPERIENCE

Linux & Open Source related

- /Linux/ *Life in Debian*, ISBN: 978-7-115-19466-4, Published in September 2009.
- /Linux/ Organizer and Topic Contributor of LinuxFB: Monthly Salon of Linuxer in BUPT.
- /Linux,VT,NoSQL/ Techonology Author and Translater on Linux kernel, Filesystem, NoSQL, etc., Articles posted to Personal Web Site (http://wanxu.me/blog/) and Magazines.
- /LaTeX/ LaTeX Template for Thesis of BUPT, including code, examples, and manual.

Training Lecturer   **2005 – 2007**
Staff training or certification training on Linux and UMTS/TD-SCDMA Core Networks, Services, and Air interfaces, for operators and vendors, in Beijing, Guangdong, Wuhan, etc.

Software Developer   **2002 - 2003**
Having been responsible for development of wireless field test system, etc.

EDUCATION

Beijing University of Posts and Telecommunications (BUPT)

Ph.D., Circuits and Systems, 2002.09 — 2007.07

Beijing University of Technology

B.S., Electronic Engineering, 1998.09 — 2002.07

REFERENCES

Part of Articles

Hadoop related technical writing (`http://wangxu.me/blog/p/tag/hadoop`)

- DataNode Volume Refreshment in HDFS-1326 (English):
      `http://gnawux.info/hadoop/2010/12/datanode-volume-refreshment-in-hdf`
- How to Decomission DataNodes in HDFS: `http://labs.chinamobile.com/mblog/225_1837`

Articles on system management and version control, some of them are records of problem resolving and scripts:

- Synchronizing SVN with SVM and the lock issue: `http://wangxu.me/blog/p/523`
- 1 line scripts for log analysis: `http:////wangxu.me/blog/p/145`

Part of Translations

NoSQL, Hadoop, and Java related translations:

- Cassandra by Example: `http://wangxu.me/blog/p/383`
- Observers: Making ZooKeeper Scale Even Further: `http://wangxu.me/blog/p/320`
- Avro: a New Format for Data Interchange: `http://wangxu.me/blog/p/279`
- Java SE 6 Hotspot ^{TM} Virtual Machine Garbage Collection Tuning :
      `http://wangxu.me/blog/p/209`

Linux kernel, graphics and other software related translations:

- LWN: The Fun of Null Pointers (1): `http://wangxu.me/blog/p/39`
- KernelTrap: Linux: The Journaling Block Device: `http://wangxu.me/blog/p/151`
- LWN: What is RCU, Fundamentally?: `http://wangxu.me/blog/p/122`

Note: All the original articles are from Internet, and announced the info about original articles; and most articles, including those from LWN, cloudera has asked the authorization of original publishers