

## 彭 成 章 个 人 简 历

电话: 13523181752/13379081097

email: pengcz.nwpu@gmail.com

出生年月: 1977年2月

就读学校: 西北工业大学

学历: 博士在读

籍贯: 湖南平江

家庭状况: 已婚, 育有一女

twitter/stackoverflow/github: abelard2008

### 教育经历

博士 2009.04 – 现在

方向: 大数据实时分析、存储

学校: 西北工业大学计算机学院

博士论文: 大数据管理系统基础架构的研究

硕士 2002.09 – 2005.03

方向: 嵌入式计算

学校: 东北大学计算机软件与理论

论文: 基于ARM Linux 的嵌入式系统及其应用研究

本科 96.09 – 02.06

专业: 财务会计

学校: 中国人民大学(北京自学考试)

### 项目经历 - 工程

类Stackoverflow — 2015.01 – 现在

项目概述: *www.yuelaw.com*, 一个主要基于*SocialStream*组件, 并参考Vish的*Rails*应用, 已实现的功能包括注册账号, 标签、提问/回答问题, 对问题/回答添加评论, 当有人回答/评论时, 给相关用户实时发送通知, 对问题/回答点赞/倒赞, 站内搜索。其中ORM使用*ActiveRecord*, 数据库使用*PostgreSQL*, 问题提交、评论和通知等都使用了*Ajax*, 组件主要用了*faye*, *ancestry*, *thinking-sphinx*, *strongex*, *private\_pub*, *ckeditor*, *whenever*等。另外, 实现过程中, 大量阅读了*SocialStream*的源代码。

开发语言及平台: *Ruby*, *Javascript*, *SCSS*, *Mac OS*;

个人负责: 全部

基于OpenFiler的NAS和SAN系统 — 2011.7-2012.5

项目功能概述: 在开源系统OpenFiler的基础上, 开发符合特定NAS和SAN要求的存储系统, 该系统主要为满足80路监控摄像的存储性能。

开发语言及平台: *PHP*, *C*, *Shell*, *Makefile*, *Linux OS*, *LVM*, *FTP*, *Samba*, *NFS*

个人负责: (1) 管理和研发 (2) 设计RAM文件系统; 由于以CentOS和多个开源组件基础的OpenFiler系统只提供执行系统, 并且是由*rPath*构建, 在针对特定设备如SAS HBA卡时, 无法从源代码加载驱动, 为此在没有RPM和yum的情况下搭建编译环境; (4) 阅读由PHP构建的web系统源码并指导组员进行针对性修改 (3) 前期的需求分析和客户沟通等

### 基于IOP348的SAN系统 — 2009.10-2011.3

项目概述: 在以intel的专用存储处理器*iop348*, *QLogic FC* 目标器芯片和*PMC SAS*扩展芯片为主要组件的硬件平台上, 开发硬件RAID后, 将其作为SCST (支持ISCSI或FC-SCSI协议) 的LUN块设备, 来构建一个高性能SAN系统, 用户通过ISCSI或FC-SCSI启动器(支持*windows*或类UNIX系统)使用SAN系统提供的存储。

开发语言及平台: *C, Shell, Makefile, Linux OS*

个人负责: (1) 管理和研发 (2) 在linux环境下, 搭建交叉编译环境, 基于*BusyBox*实现嵌入式Ext3文件系统; 分析SCSI协议; 针对特定的FC HBA卡, 实现RAID和SCST (支持ISCSI和FC-SCSI协议) 的系统集成; 基于*iop348*的DMA XOR和P+Q功能, 设计RAID 5/6的样例程序; 为了保护异常当机的内存数据, 修改系统启动的内存管理

### 机载设备数据采集系统 — 2009.7-2009.11

项目概述: 在Keil MDK-ARM环境下, 给基于Cortex的48路温度采集系统开发nand flash的无操作系统裸驱动和所有下位机程序。

开发语言及平台: *c, Keil MDK, Windows XP*

个人负责: 所有事务

### 种子基金(计算环境下系统智能迁移的研究) — 2010.12-2011.12

项目功能概述: 设计和实现基于系统智能迁移的云计算平台, 主要目的是解决Xen和KVM上虚拟机系统迁移的智能决策、虚拟机占用资源统计。

开发语言及平台: *Java, Shell, Eucalyptus* 云计算系统, *Linux OS*;

个人负责: 实现KVM虚拟机迁移

## 项目经历 - 研究

### 云计算和云存储 — 2010.09-2011.08

简介: 在分析和研究业界各种云计算和云存储后, 搭建了开源Eucalyptus云计算平台, 并在尝试各种云计算应用后, 最后确定研究Eucalyptus中的云存储walrus, 经过对Eucalyptus架构等各方面的分析后, 考虑将使用本地存储空间的walrus扩展为分布式的walrus, 基本思路是将Cassandra作为核心存储组件, 根据Walrus的存储内容和读写方式, 设计基于Cassandra的存储控制组件。

主要成绩: (1) 构建了基于SOAP和Cassandra云存储系统, 用户通过客户端可以存储和读取文件, 客户端和存储中心通过SOAP方式进行交流。(2) 通过阅读Eucalyptus的源代码, 理清其启动流程; 明白了框架的核心组件是由一个自己实现的Servlet容器; 第一次见识了一个由Java, Groovy, Python, c, shell等组成的应用系统。(3) 阅读Cassandra的源代码, 因其使用了经典的SEDA用于高并发, 进而阅读了该论文和部分源代码, 并就源代码与作者Matt Welsh沟通。

大数据实时分析 — 2011.09 – 2014.12

**早期引擎 (SPE):** (1) 为了理解SPE方向经典的理论和设计方法, 搭建了与早期论文(2001年流处理 ~2007年)匹配的系统并设法重现论文中的实验, 这些系统有Brown, Brandeis和MIT三所大学合作的Aurora, Borealis, 以及斯坦福大学的 STREAM和流引擎 (SPE): 处理引擎基准测试系统Linear Road, 和Berkeley的TelegraphCQ。(2) 将Borealis移植到Fedora 20 x86\_64上, 成功地运行了其中一个应用 Borealis的射击游戏Cube, 并找到一个Borealis本身的Bug。(3) 在运行 Magdalena Balazinska (MIT) 的博士论文中的容错(Fault Tolerance), 多点故障(multi-nodes Fault)实验时, 通过阅读Borealis和她的例程代码, 确定现有版本与论文描述有误后, 与她联系并获得她8年前写的源码, 找到了问题所在。(4) 阅读了Borealis的几个核心模块: 查询处理器, Aurora节点, DataPath和使用Antlr工具实现的解释性语言(主要用于应用程序的配置) (5) 阅读Borealis使用的, 基于SEDA的网络和消息组件NMSTL源码, 并自己尝试了重写部分代码。(6) 为更好地掌握Borealis系统的数据库知识, 阅读了Raghu Ramakrishnan编写的经典教材 “Database Management Systems” 和minibase源代码中的Buffer Manager和B+ Trees部分, 并参考PostgreSQL的源码进行修改, 使其可以在Fedora 20 x86\_64上运行。

**近期的流处理引擎:** (1) 主要关注了Twitter的storm和(瑞典)乌普萨拉大学的SCSQ, 并重现论文相关的实验, (2) 阅读了storm的源代码, 主要包括启动、集群管理和数据处理, 以及从函数式语言的角度分析几个重要的宏。

**基于Hadoop的大数据实时处理:** (1) 主要研究Berkeley的HOP论文和实现, 在基于源代码详细分析Hadoop的MapReduce架构和编程模式后, 研究HOP相对于原始Hadoop所做的大数据修改和修改后所带来的优缺点及其适应场合 (2) 修改Hadoop的输入方式, 将原来的HDFS提取数据, 修改为通过基于ZeroMQ构建的支持TCP的EndPoint实时收集数据来进行处理 (3) 将Mapper和Reducer之间pull方式修改适应实时分析的push方式。(4) 重点阅读了Hadoop-0.20.203\*中MapReduce的RPC系统, JobTracker, TaskTracker, Mapper和Reducer以及启动和数据处理过程等的代码 (5) 在HOP中, Tyson Condie使用了自己团队开发的, 用于网络协议的申明式语言Overlog, 为了理解的完整性, 干脆用他们最新的Bloom完整实现了P2P中经典算法Chord。

## 教学实践

教学实践: — 2010.09-2013.7

**简介:** 每个学期作为 “嵌入式系统” 课程的助教, 负责基于ARM S3C2440开发版实验指导, 并讲解包括由自己实现的xworks驱动, 搭建Linux和Windows CE的开发环境和开发简单的应用程序。

## 工作经历

焦作大学机电工程系任教师 — 2006.09-2009.03

任教课程: C 语言程序设计, 单片机原理(8051系列), Protel电路设计, VHDL电子设计

主要课外项目: (1) 为带毕业实习, 自己动手设计8051电路板原理图和PCB布线(两层PCB板), 移植原始的ucosII到8051板; (2) 远程监控系统: 在基于 ARM7的ucosII上, 添加开源的tcp/ip协议栈, 其数据链路层使用ppp协议, 配合GPRS模块开发远程监控系统, 在window xp上使用Windows 32 API开发一个可以发送和接收短信的微型客户端

北京德信无线电研发工程师 — 2005.3 – 2006.7

个人负责: (1) 基于高通BREW平台开发手机中间件, 向应用组提供菜单接口, (2) 维护日本ACCESS公司的, 使用VC6开发的手机模拟器

## 发表论文

(1) Peng Chengzhang; Jiang Zejun\*, Building a Cloud Storage Service System, Proceedings of the 3rd International Conference on Environmental Science and Information Application Technology, ESIAT 2011, Vol 10 PP 691-696, 西安, 2011.8. 20-21

(2) Peng ChengZhang; Ze-Jun Jiang\*; Xiao-Bin Cai; Zhi-Ke Zhang, Implementing chord with declarative networking language, Proceedings of the 11 International Conference on Machine Learning and Cybernetics, ICMLC 2012, Vol 2 PP 586-590, 西安, 2012.7.15-17

(3) Peng ChengZhang; Ze-Jun Jiang\*; Xiao-Bin Cai; Zhi-Ke Zhang, Real-time analytics processing with MapReduce, Proceedings of the 11 International Conference on Machine Learning and Cybernetics, ICMLC 2012, Vol 4 PP 1308-1311, 西安, 2012.7.15-17

(4) 彭成章, 蒋泽军, 蔡小斌, 张志珂. 基于分布式编程语言的Chord协议和算法. 计算机应用, 33(7):1885-1889 July 2013

(5) (软件著作权) 王丽芳, 蒋泽军, 刘志强, 彭成章, 张志珂, 韦涌泉. 高性能磁盘阵列系统, 登记号: 2011SR043424, 中国, 2011.

(6) (软件著作权) 王丽芳, 蒋泽军, 刘志强, 彭成章, 张志珂, 韦涌泉. 基于 XML的磁盘阵列管理系统, 登记号: 2011SP563428, 中国, 2011.

## 专业技能

使用过的语言: C/C++, Java, Ruby, JavaScript, Shell, Makefile, Scheme, Haskell, Scala, X86, ARM汇编等;

操作系统: 从研究生开始间断地使用RedHat Linux, 到2006年开始将所有事情都在Fedora上完成, 偶尔会使用CentOS和Ubuntu, 对Linux2.4和2.6进行过低级别的裁剪、驱动设计和编译, 另外, 是一个忠实的Emacs和vi编辑器使用者;

算法：除了应付考试，没有完整读过一本严格的算法书，但是在遇到问题和别人代码中的算法时，一定会去查阅理论和分析、写一个自己的算法，如B+树，红-黑树，外部排序，Hash、回溯和H-99: Ninety-Nine Haskell Problems中的前70个问题等

## 英语水平

大学英语四级，有较强的英语阅读写作能力，经常会与计算机领域的欧美科技工作者进行在线和离线英语交流。

## 这些年阅读和参考的部分书

计算机：通灵芯片(丹尼尔·希利斯)，逻辑的引擎(马丁·戴维斯)，Programming in Scala(Martin Odersky)，A little Java/A few Patterns/The Little Schemer(Daniel P. Friedman)，Programming Languages: Application and Interpretation(Shriram Krishnamurthi)，LINUX DEVICE DRIVERS(Jonathan Corbet)，Algorithms 4th Edition(Robert Sedgewick)，计算程序设计艺术(Knuth)，Real World Haskell(Bryan O' Sullivan)，Learn You a Haskell for Great Good(Miran Lipovaca)，MetaProgramming Ruby(Paolo Perrotta)

其它：活法(稻盛和夫)，易经与人生(傅佩荣)，柏拉图对话集(柏拉图)，苏菲的世界(乔斯坦·贾德)，1984(乔治·奥威尔)，怪诞行为学(丹·艾瑞里)，自控力(凯利·麦格尼格尔)，万历十五年(黄仁宇)，平面国(艾勃特)，返璞归真(C.S.路易斯)，沉思录(马可·奥勒留)，上帝的语言(弗兰西斯·柯林斯)，我读天下无字书(丁学良)，罗马帝国衰亡史第一卷(爱德华·吉本)

## 自我评价

有一定的上进心；乐观、专注、执着、自我激励、热爱生活；喜欢倾听内心的声音。