

分类号 计算机软件与理论 密级 公开

UDC 004.72

学 位 论 文

基于小世界理论的

数据中心网络模型

(题名和副题名)

韦小宝

(作者姓名)

指导教师姓名、职务、职称、学位、单位名称及地址 陈近南 教授

南京大学计算机科学与技术系 南京市汉口路 22 号 210093

申请学位级别 博士 专业名称 计算机软件与理论

论文提交日期 2013 年 5 月 10 日 论文答辩日期 2013 年 6 月 1 日

学位授予单位和日期 _____

答辩委员会主席: 张三丰 教授

评阅人: 阳顶天 教授

张无忌 副教授

黄裳 教授

郭靖 研究员



南京大學

研究生畢業論文 (申請博士學位)

論 文 題 目 基于小世界理论的数据中心网络模型

作 者 姓 名 韦小宝

学科、专业方向 计算机软件与理论

指 导 教 师 陈近南 教授

研 究 方 向 计算机网络与信息安全

2013 年 5 月 1 日

学 号：**MG0033011**

论文答辩日期：**2013 年 6 月 1 日**

指 导 教 师：

(签字)

A Network Model of Data Centers Based on the Small World Theory

by

Xiaobao Wei

Directed by

Professor Jinnan Chen

Department of Computer Science and Technology
Nanjing University

May 1, 2013

*Submitted in partial fulfilment of the requirements
for the degree of Ph.D. in Computer Software and Theory*

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目： 基于小世界理论的数据中心网络模型

计算机软件与理论 专业 2010 级博士生姓名： 韦小宝

指导教师（姓名、职称）： 陈近南 教授

摘 要

复杂网络的研究可上溯到 20 世纪 60 年代对 ER 网络的研究。90 年后代随着 Internet 的发展，以及对人类社会、通信网络、生物网络、社交网络等各领域研究的深入，发现了小世界网络和无尺度现象等普适现象与方法。对复杂网络的定性定量的科学理解和分析，已成为如今网络时代科学研究的一个重点课题。

在此背景下，由于云计算时代的到来，本文针对面向云计算的数据中心网络基础设施设计中的若干问题，进行了几方面的研究。.....

关键词： 小世界理论；网络模型；数据中心

南京大学研究生毕业论文英文摘要首页用纸

THESIS: A Netork Model of Data Centers
Based on the Small World Theory
SPECIALIZATION: Computer Software and Theory
POSTGRADUATE: Xiaobao Wei
MENTOR: Professor Jinnan Chen

Abstract

This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract.

This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract. This is English abstract.

keywords: Small World, Network Model, Data Center

目 录

目 录	v
1 绪论	1
1.1 研究背景	1
1.2 研究目的与意义	1
1.2.1 现有解决方法	1
1.2.2 现有问题与不足	2
1.2.3 中心观点与思想	3
1.3 研究的应用背景	4
2 基于小世界模型的网络模型实现:SIDN	5
2.1 小世界现象	5
2.2 网络结构的重要指标	7
2.2.1 平均路径长度	7
2.2.2 聚类系数	7
3 结论	11
A 博士 (硕士) 学位论文编写格式规定 (试行)	13
A.1 适用范围	13
A.2 引用标准	13
A.3 印制要求	13
A.4 编写格式	13
A.5 前置部分	14
A.5.1 封面 (博士论文国图版用)	14

A.5.2	题名	14
A.5.3	前言	15
A.5.4	摘要	15
A.5.5	关键词	15
A.5.6	目次页	15
A.5.7	插图和附表清单	16
A.6	主体部分	16
A.6.1	格式	16
A.6.2	序号	16
A.6.3	绪论	16
A.6.4	正文	17
A.6.5	结论	19
A.6.6	致谢	19
A.6.7	参考文献表	20
A.7	附录	20
A.8	结尾部分 (必要时)	21
参考文献		23
简历与科研成果		25
致 谢		27
学位论文出版授权书		29

表格索引

1-1 测试表格	2
2-1 测试表格	6
2-2 测试表格	8
2-3 测试表格	9
3-1 测试表格	12

插图索引

1-1 测试插图	4
2-1 测试插图	5
2-2 测试插图	6
2-3 测试插图	8
2-4 测试插图	9
2-5 测试插图	10

第 1 章 绪论

1.1 研究背景

在分布式网络领域，沿着高性能集群、普世计算、网格计算的方向，现已走入云计算时代。

云计算对信息技术架构造成了越来越大的影响。例如，借助 Amazon EC2 云平台，用户借助其基础设施，可以十分方便的部署各类应用，以支持企业服务需求。用户可以按需购买计算资源，网络带宽，存储空间等各类资源以支持他们的业务需求，并在业务完成之后迅速的归还这些资源。通过云技术，用户可以集中在他们擅长的核心业务之中，而不会被诸如硬件购买、安装系统、网络设置、备份和安全等等问题干扰。

1.2 研究目的与意义

1.2.1 现有解决方法

在工程实践领域，目前在面向云计算的实际应用中，很少有在网络模型上有所突破的实际案例。以 Amazon EC2，Google 等云计算实现通常的解决方案为采购现有的成熟网络设备，以常规方法去组建网络，在底层网络、服务器节点、操作系统中都不做任何特殊处理。通过降低单个硬件成本，在单元资金投入的情况下采购更多的硬件，以数量获取性能，以冗余换取可靠性。

根据 Google 早期披露的数据中心信息，Google 以集装箱为标准单位，一个集装箱内容纳 1160 台服务器，一个数据中心将容纳数万服务器节点。在运行的第一年中，有 1 千台机器故障；数千块硬盘损坏；一个 PDU 损坏导致 500-1000 台服务器下线 6 小时；一次网络重连导致 5% 的节点下线 2 天；20 个机柜故障

文档域类型	Java 类型	宽度 (字节)	说明
BOOLEAN	boolean	1	
CHAR	char	2	UTF-16 字符
BYTE	byte	1	有符号 8 位整数
SHORT	short	2	有符号 16 位整数
INT	int	4	有符号 32 位整数
LONG	long	8	有符号 64 位整数
STRING	String	字符串长度	以 UTF-8 编码存储
DATE	java.util.Date	8	距离 GMT 时间 1970 年 1 月 1 日 0 点 0 分 0 秒的毫秒数
BYTE_ARRAY	byte[]	数组长度	用于存储二进制值
BIG_INTEGER	java.math.BigInteger	和具体值有关	任意精度的长整数
BIG_DECIMAL	java.math.BigDecimal	和具体值有关	任意精度的十进制实数

表 1-1: 测试表格

导致 40-80 台节点下线；大量的小问题导致 30 秒左右的卡顿等等 [7]。

另一方面，在学术领域，在这方面有许多相关研究。大致上可分为两派，其一以改良为主，如针对扩展性问题，Fat-tree[4] 通过增加交换机的聚合端口，创造更多的上层交换节点，从而拓宽了聚合带宽，缓解了扩展的瓶颈问题；针对动态性问题，VL2[9] 提出了双层虚拟网络结构，允许虚拟机在保持网络配置不变的情况下，在物理节点中迁移；针对安全性问题，Amazon EC2[19] [12] 通过 Security Group 概念，以网络虚拟化的形式实现网络的物理隔离；针对广域网中面向海量用户所导致的流量激增与路由膨胀问题，通过 CDN[18]，以大量部署基础设施缓存数据来提高服务质量。

另一派则以革新为主要目的，以 OpenFlow 为代表，以重新设计整个网络协议为出发点。在实现上主要有基于 CPU 的软件仿真实现，如 Click[13]，Quagga[2]，VDE[3]，基于 FPGA 可编程虚拟网络设备，如 RouteBricks[8]，NetFPGA[14]，以及 GPU 等特殊硬件，如 PacketShader[11] 等。

1.2.2 现有问题与不足

从工程实现的角度去看，现有的实施方案依然以传统模式构建底层网络，存在难以满足廉价、高可靠性、高可用性、灵活性等需求的问题，只能以应用层冗余的形式进行缓解，难以彻底解决：

1. 网络结构中存在许多单点节点，对性能、扩展性造成瓶颈，同时单点故障会引起大面积的网络故障。
2. 网络结构的扩展性需要付出子网隔离的代价，造成管理复杂，并且与灵活性冲突。
3. 网络上层存在高聚合带宽，需要高端网络设备，因而随着网络规模的增大，组网成本也会指数上升。

虽然以大量设备去缓存复用等各种措施可以缓解上述问题。但解决问题的出发点并不是立足于问题本身，而是堆积资源强行绕开问题的一直规避手段。站在工程实践的角度而言，措施是有效的、风险和问题是可控的，但本文认为其核心问题存在改善的空间，针对其改进是更有意义的改善方式。

1.2.3 中心观点与思想

云计算在概念上通常被分为 IaaS、PaaS、SaaS 几个层面。但透过分类去理解其本质，可认为是上世纪 70 年代基于大型计算机的中心控制型瘦客户端终端模式，在如今技术水平上的一种新的表达，是在技术发展道路中，螺旋上升的结果。

与瘦客户端相比，云计算在设计结构上存在一定的相似性。

1. 中心控制的模式：通过中心的大规模硬件提供统一的计算，可大大降低管理成本，提高硬件资源利用率，同时降低客户端的硬件成本需求。例如 Nvidia 推出 GeForce GRID 平台 [1]，推出了 GaaS[†]概念。将 GPU 放置在云端，使得用户不需要再不断购买升级显卡，并可在更为广泛的终端（包括手机、平板、智能电视）和地点体验最新的游戏。
2. 数据集中：由于瘦客户端的关系，数据都集中存储在中心，可对数据提供可靠的保护，并且通过按需调用的实现方式，降低对网络带宽的需求。

在设计思路，两者都为了降低管理成本和硬件成本、以低能耗、高弹性等需求为设计目标。随着技术的进步，云计算在具体实现形态上与传统的大型机也有很大的不同：

一方面，云中心不再是传统的一台大型机，而是用大量廉价计算节点的互联来提供海量资源。云计算更强调资源规模的无缝、平滑扩展，以及高可靠性，无单点故障问题。另一方面，云计算时代的终端，也具备相当计算能力。

[†]Gaming as a Service

随着 web2.0 的整合，还有向胖客户端和智能终端发展的趋势。

总而言之，云计算在大框架中是传统的中心控制／终端的模式，但在中心与终端两方面，都引入分布式技术加以改良。核心的思路是在低成本的前提下做到高可靠性、高灵活性和高伸缩性。因此，云计算并不仅仅以数量换性能的表象，本质上为低成本高性能，追求高能效比，并在实现层面讲究可实现性和可操作性。

1.3 研究的应用背景

云计算作为分布式技术的当前表现形式，通过将众多节点资源整合，以冗余、去中心化的分布式模式，实现传统技术中需要大型机才能解决的海量信息问题。一言而概之，“人多力量大”。

但随着节点数目的增多，问题的重点将逐步转换为如何对大量节点进行高效互联。图 1-1 所示为传统 IaaS 云中心网络结构的一部分。通过 BR 边界路由器，AR 接入路由器构建数据中心的主干；核心交换机和接入交换机 S，构成二层交换网络层，大量的服务器节点通过二层交换机被最终接入整个网络。



图 1-1: 测试插图

在上述传统网络中，当节点总数达到数千乃至万数量级时，上层链路的聚合带宽将不断提高，从而对核心交换机、接入路由器、边界路由器的指标提出了极高的要求。以至于少数核心网络设备，成为整个网络中的高价格、高性能单点。一方面与原本追求低成本、分布化、去中心化的云技术设计理念背道而驰，另一方面也降低了网络的健壮性。因而本文所做工作对数据中心的网络基础设施而言，具有较大应用价值。

第 2 章 基于小世界模型的网络模型实现:SIDN

2.1 小世界现象

生活中，常出现初次见面的陌生人却拥有双方都认识的共同熟人，于是大家时常会感叹：“这世界真小！”。这种现象被称为“小世界现象”，后又称为“六度分割理论”。



图 2-1: 测试插图

1909 年，现代无线电之父 Guglielmo Marconi 在其诺贝尔奖致辞中讨论了覆盖整个地球所需的无线电中继站数目，并根据他的实验结果计算出平均需要 5.83（近似为 6）个中继站 [16]。这个结论，被认为是“六度分割理论”中的常数 6 的最早出处 [5]。

1929 年，匈牙利作家 Frigyes Karinthy 发表了一部短篇小说集《Everything is Different》。其中一篇名为《Chain-Links》的小说以抽象的、概念性的和虚构的方式研究了网络理论领域的很多问题，而这些问题使得未来几代的数学家、社会学家和物理学家都为之着迷 [5, 17]。Karinthy 认为，随着通讯技术和交通技术的发展，人际关系网会变得越来越来大，扩张得越来越远，整个世界将因此而“缩小”。他认为，虽然人类个体之间的物理距离可能很远，但人类社交网络密度的增加使得人类个体之间的社会性距离变得非常小。根据这个假

文档域类型	Java 类型	宽度 (字节)	说明
BOOLEAN	boolean	1	
CHAR	char	2	UTF-16 字符
BYTE	byte	1	有符号 8 位整数
SHORT	short	2	有符号 16 位整数
INT	int	4	有符号 32 位整数
LONG	long	8	有符号 64 位整数
STRING	String	字符串长度	以 UTF-8 编码存储
DATE	java.util.Date	8	距离 GMT 时间 1970 年 1 月 1 日 0 点 0 分 0 秒的毫秒数
BYTE_ARRAY	byte[]	数组长度	用于存储二进制值
BIG_INTEGER	java.math.BigInteger	和具体值有关	任意精度的长整数
BIG_DECIMAL	java.math.BigDecimal	和具体值有关	任意精度的十进制实数

表 2-1: 测试表格

设，Karinthy 的小说的主角相信：“任何两个人之间可以通过不超过五个中间人相联系”。Karinthy 的想法直接或间接地影响了早期的社交网络理论的研究，他被认为是六度分隔 (six degrees of separation) 理论的最早提出者 [5]。

1961 年，Michael Gurevich 在社会学家 Ithiel de Sola Pool 的指导下完成了他的博士论文，对社交网络进行了实验性的研究。随后，数学家 Manfred Kochen 与 Sola Pool 一道在他们的手稿《Contacts and Influences》中对这些实验结果做了分析，发现在美国人口中，任意两个人之间通常只需不超过两个中间人即可互相联系 [6]。1973 年他们又利用计算机，基于 Gurevich 的数据，用 Monte Carlo 法做了模拟，并证实了该结论 [6]，从而为心理学家 Stanley Milgram 后来的发现打下了基础。



图 2-2: 测试插图

2.2 网络结构的重要指标

在刻画复杂网络结构的统计特性上有三个重要的指标：平均路径长度（average path length）、聚类系数（clustering coefficient）和度分布（degree distribution）。事实上，Watts 和 Strogatz 提出小世界网络模型的初衷，就是想建立一个既具有类似随机图的较小的平均路径长度，又具有类似规则网络的较大的聚类系数的网络模型。

2.2.1 平均路径长度

定义 2.1 (节点之间的距离)： 网络中两个节点 i 和 j 之间的距离 (distance) d_{ij} 定义为连接这两个节点的最短路径上的边数。

定义 2.2 (直径)： 网络中任意两个节点之间的距离的最大值称为该网络的直径，记为 D ，即

$$D = \max_{i,j} d_{ij} \quad (2-1)$$

定义 2.3 (平均路径长度)： 网络的平均路径长度 L 定义为任意两个节点之间的距离的平均值，即

$$L = \frac{2}{N(N+1)} \sum_{i \geq j} d_{ij} \quad (2-2)$$

其中 N 为网络节点数。网络的平均路径长度也称为网络的特征路径长度。

注意，为了便于数学处理，在公式(2-2)中包含了节点到其自身的距离（该距离为零）。如果不考虑节点到其自身的距离，那么公式 (2-2)的右端需要乘以因子 $(N+1)/(N-1)$ 。在实际应用中，该差别可以忽略不计。

2.2.2 聚类系数

在图论中，聚类系数 (clustering coefficient) 是用来描述一个图中的顶点之间结集成团的程度的系数。具体来说，是一个点的邻接点之间相互连接的程度。许多大规模的实际网络都具有明显的聚类效应。例如生活社交网络中，你的朋友同时也是朋友的概率会随着网络规模的增加而趋向于某个非零常数。这意味着这些实际的复杂网络并不是完全随机的，而是在某种程度上具有类似于社会关系网络中“物以类聚，人以群分”的特性。

文档域类型	Java 类型	宽度 (字节)	说明
BOOLEAN	boolean	1	
CHAR	char	2	UTF-16 字符
BYTE	byte	1	有符号 8 位整数
SHORT	short	2	有符号 16 位整数
INT	int	4	有符号 32 位整数
LONG	long	8	有符号 64 位整数
STRING	String	字符串长度	以 UTF-8 编码存储
DATE	java.util.Date	8	距离 GMT 时间 1970 年 1 月 1 日 0 点 0 分 0 秒的毫秒数
BYTE_ARRAY	byte[]	数组长度	用于存储二进制值
BIG_INTEGER	java.math.BigInteger	和具体值有关	任意精度的长整数
BIG_DECIMAL	java.math.BigDecimal	和具体值有关	任意精度的十进制实数

表 2-2: 测试表格



图 2-3: 测试插图

集聚系数分为整体与局部两种。整体集聚系数可以给出一个图中整体的集聚程度的评估，而局部集聚系数则可以测量图中每一个结点附近的集聚程度。

定义 2.4 (整体聚类系数): 整体集聚系数的定义建立在闭三点组（邻近三点组）之上。假设网络中有一部分节点是两两相连的，那么可以找出很多个“三角形”，其对应的三点两两相连，称为闭三点组。除此以外还有开三点组，也就是之间连有两条边的三点（缺一条边的三角形）。这两种三点组构成了所有的连通三点组。整体集聚系数定义为一个网络中所有闭三点组的数量与所有连通三点组（无论开还是闭）的总量之比，即

$$C_{total} = \frac{3 \times G_{\Delta}}{3 \times G_{\Delta} + G_{\wedge}}$$

其中 C_{total} 表示网络的整体聚类系数， G_{Δ} 表示该网络中闭三点组的个数， G_{\wedge} 表示该网络中开三点组的个数 [15]。

文档域类型	Java 类型	宽度 (字节)	说明
BOOLEAN	boolean	1	
CHAR	char	2	UTF-16 字符
BYTE	byte	1	有符号 8 位整数
SHORT	short	2	有符号 16 位整数
INT	int	4	有符号 32 位整数
LONG	long	8	有符号 64 位整数
STRING	String	字符串长度	以 UTF-8 编码存储
DATE	java.util.Date	8	距离 GMT 时间 1970 年 1 月 1 日 0 点 0 分 0 秒的毫秒数
BYTE_ARRAY	byte[]	数组长度	用于存储二进制值
BIG_INTEGER	java.math.BigInteger	和具体值有关	任意精度的长整数
BIG_DECIMAL	java.math.BigDecimal	和具体值有关	任意精度的十进制实数

表 2-3: 测试表格

对图中具体的某一个点，它的局部集聚系数 C_i 表示与它相连的点抱成团（完全子图）的程度。Watts 与 Strogatz 在 1998 年的论文 [20] 中首次引入了这个概念，用以判别一个图是否是小世界网络。



图 2-4: 测试插图

定义 2.5 (局部聚类系数): 假设网络中的一个节点 i 有 k_i 条边与其他节点相连，这 k_i 个节点称为节点 i 的邻居。显然，在这 k_i 个节点之间最多可能有 $k_i(k_i - 1)/2$ 条边。而这 k_i 个节点之间实际存在的边数 E_i 和总的可能的边数 $k_i(k_i - 1)/2$ 之比就定义为节点 i 的聚类系数 (clustering coefficient) C_i ，即

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{2-3}$$

从几何特性上看，上式的一个等价定义为：

$$C_i = \frac{\text{与节点 } i \text{ 相连的三角形的数量}}{\text{与节点 } i \text{ 相连的三元组的数量}} \tag{2-4}$$

其中，与节点 i 相连的三元组是指由节点 i 和其两个邻居节点构成的组合。



图 2-5: 测试插图

知道了一个图里的每一个顶点的局部集聚系数后，可以计算整个图的平均集聚系数。这个概念也是 Watts 与 Strogatz 在 1998 年的论文 [20] 中引入的：

定义 2.6 (平均聚类系数)： 平均聚类系数定义为所有顶点的局部集聚系数的算术平均数，即

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i. \quad (2-5)$$

第 3 章 结论

本文在第一章中，通过考虑数据中心网络布局构建中的最大度限制问题，提出了符合数据中心网络基本要求的 DS 小世界模型，并分析了它的性质。随后提出 SIDN，将 DS 模型映射到具体的网络结构中，并分析了所构成网络的平均直径、网络总带宽、对故障的容错能力等各项网络性能。

分析与仿真实验证明，SIDN 网络具有很好的扩展能力，网络总带宽与网络规模成近似线性增长的关系；具有很强的容错能力，链路损坏与节点损坏几乎无法破坏网络的联通性，故障率对网络性能的影响与破坏节点/链路占总资源比率线性相关。

随后在第二章中，分析了无尺度网络在数据中心网络构建应用中的理论方面问题。对 Scafida [10] 文中所述在最大度限制的情况下运用 BA 算法构造的网络并不会损失无尺度性质的观点，进行了深入的分析，并指出了该论点的局限性。

在给出了在引入节点最大度限制之后，利用分治和递归的思想，对无尺度网络进行多层构建，对所构造的网络进行度-度相关性，以及聚类性分析。

表3-1用于测试表格。随后分析了无尺度网络构造过程中，交换机节点与数据节点的角色区别，分析了两者在不同比率下形成的网络形态，以及对网络性能造成的影响。

通过理论分析和仿真实验，分析并找出比率因子 q 的最佳取值。此外，无尺度现象的引入提高了网络的聚类系数，从而在不失灵活性可靠性的基础上，进一步提升了网络的性能。

在第三章中，将关注点转移到交换机本身。由于图论难以描述数据中心网络中的交换设备，因此放弃基于图的抽象模型，转而基于多维簇划分的思想，提出并设计了 WarpNet 网络模型。

该网络模型突破了基于图描述的局限性，并对网络的带宽等指标进行理论

文档域类型	Java 类型	宽度 (字节)	说明
BOOLEAN	boolean	1	
CHAR	char	2	UTF-16 字符
BYTE	byte	1	有符号 8 位整数
SHORT	short	2	有符号 16 位整数
INT	int	4	有符号 32 位整数
LONG	long	8	有符号 64 位整数
STRING	String	字符串长度	以 UTF-8 编码存储
DATE	java.util.Date	8	距离 GMT 时间 1970 年 1 月 1 日 0 点 0 分 0 秒的毫秒数
BYTE_ARRAY	byte[]	数组长度	用于存储二进制值
BIG_INTEGER	java.math.BigInteger	和具体值有关	任意精度的长整数
BIG_DECIMAL	java.math.BigDecimal	和具体值有关	任意精度的十进制实数

表 3 – 1: 测试表格

分析并给出定量描述。最后对比了理论分析、仿真测试结果，并在实际物理环境中进系真实部署，通过 6 节点的小规模实验以及 1000 节点虚拟机的大规模实验，表明该模型的理论分析、仿真测试与实际实验吻合，并在网络性能、容错能力、伸缩性灵活性方面得到了进一步的提升。

在第四章中，针对网络模型研究这一类工作的共性，设计构造通用验证平台系统。以海量虚拟机和虚拟分布式交换机的形式，实现了基于少量物理节点，对大规模节点的模拟。其模拟运行的过程与真实运行在实现层面完全一致，运行的结果与真实环境线性相关。除为本文所涉若干网络模型提供验证外，可进一步推广到更为广泛的领域，为各种网络模型及路由算法的研究工作，提供分析、指导与验证。

附录 博士 (硕士) 学位论文编写格式规定 (试行)

A.1 适用范围

本规定适用于博士学位论文编写，硕士学位论文编写应参照执行。

A.2 引用标准

GB7713 科学技术报告、学位论文和学术论文的编写格式。

GB7714 文后参考文献著录规则。

A.3 印制要求

论文必须用白色纸印刷，并用 A4(210mm×297mm) 标准大小的白纸。纸的四周应留足空白边缘，上方和左侧应空边 25mm 以上，下方和右侧应空边 20mm 以上。除前置部分外，其它部分双面印刷。

论文装订不要用铁钉，以便长期存档和收藏。

论文封面与封底之间的中缝（书脊）必须有论文题目、作者和学校名。

A.4 编写格式

论文由前置部分、主体部分、附录部分 (必要时)、结尾部分 (必要时) 组成。

前置部分包括封面, 题名页, 声明及说明, 前言, 摘要 (中、英文), 关键词, 目次页, 插图和附表清单 (必要时), 符号、标志、缩略词、首字母缩写、单位、术语、名词解释表 (必要时)。

主体部分包括绪论 (作为正文第一章)、正文、结论、致谢、参考文献表。

附录部分包括必要的各种附录。

结尾部分包括索引和封底。

A.5 前置部分

A.5.1 封面 (博士论文国图版用)

1. 封面是论文的外表面, 提供应有的信息, 并起保护作用。
2. 封面上应包括下列内容:
 - (a) 分类号在左上角注明分类号, 便于信息交换和处理。一般应注明《中国图书资料分类法》的类号, 同时应注明《国际十进分类法 UDC》的类号;
 - (b) 密级在右上角注明密级;
 - (c) “博士学位论文”用大号字标明;
 - (d) 题名和副题名用大号字标明;
 - (e) 作者姓名;
 - (f) 学科专业名称;
 - (g) 研究方向;
 - (h) 导师姓名, 职称;
 - (i) 日期包括论文提交日期和答辩日期;
 - (j) 学位授予单位。

A.5.2 题名

1. 题名是最恰当、最简明的词语反映论文中最重要的特定内容的逻辑组合。

题名所用每一词语必须考虑到有助于选定关键词和编写题录、索引等二次文献可以提供检索的特定实用信息。

题名应避免使用不常见的缩略词、首字母缩写字、字符、代号和公式等。

题名一般不宜超过 20 字。

论文应有外文题名，外文题名一般不宜超过 10 个实词。

2. 可以有副题名。
3. 题名在整本论文中不同地方出现时，应完全相同。

A.5.3 前言

前言是作者对本论文基本特征的简介，如论文背景、主旨、目的、意义等并简述本论文的创新性成果。

A.5.4 摘要

1. 摘要是论文内容不加注释和评论的简单陈述。
2. 论文应有中、英文摘要，中、英文摘要内容应相同。
3. 摘要应具有独立性和自含性，即不阅读论文的全文，便能获得必要的信息，摘要中有数据、有结论，是一篇完整的短文，可以独立使用，可以引用，可以用于推广。摘要的内容应包括与论文同等量的主要信息，供读者确定有无必要阅读全文，也供文摘等二次文献引用。摘要的重点是成果和结论。
4. 中文摘要一般在 1500 字，英文摘要不宜超过 1500 实词。
5. 摘要中不用图、表、化学结构式、非公知公用的符号和术语。

A.5.5 关键词

关键词是为了文献标引工作从论文中选取出来用于表示全文主题内容信息款目的单词或术语。

每篇论文选取 3 – 8 个词作为关键词，以显著的字符另起一行，排在摘要的左下方。在英文摘要的左下方应标注与中文对应的英文关键词。

A.5.6 目次页

目次页由论文的章、节、附录等的序号、名称和页码组成，另页排在摘要的后面。

A.5.7 插图和附表清单

论文中如图表较多，可以分别列出清单并置于目次页之后。

图的清单应有序号、图题和页码。表的清单应有序号、表题和页码。

符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释表符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释说明汇集表，应置于图表清单之后。

A.6 主体部分

A.6.1 格式

主体部分由绪论开始，以结论结束。主体部分必须由另页右页开始。每一章必须另页开始。全部论文章、节、目的格式和版面安排要划一，层次清楚。

A.6.2 序号

1. 论文的章可以写成：第一章。节及节以下均用阿拉伯数字编排序号，如 1.1, 1.1.1 等。
2. 论文中的图、表、附注、参考文献、公式、算式等一律用阿拉伯数字分别分章依序连续编排序号。其标注形式应便于互相区别，一般用下例：图 1.2；表 2.3；附注 1)；文献 [4]；式 (6.3) 等。
3. 论文一律用阿拉伯数字连续编页码。页码由首页开始，作为第 1 页，并为右页另页。封页、封二、封三和封底不编入页码，应为题名页、前言、目次页等前置部分单独编排页码。页码必须标注在每页的相同位置，便于识别。
4. 附录依序用大写正体 A、B、C □ 编序号，如：附录 A。附录中的图、表、式、参考文献等另行编序号，与正文分开，也一律用阿拉伯数字编码，但在数码前题以附条序号，如图 A.1；表 B.2；式 (B.3)；文献 [A.5] 等。

A.6.3 绪论

绪论（综述）：简要说明研究工作的目的、范围、相关领域的前人工作和知识空白、理论基础和分析，研究设想、研究方法和实验设计、预期结果和意

义等。一般在教科书中有的知识，在绪论中不必赘述。

绪论的内容应包括论文研究方向相关领域的最新进展、对有关进展和问题的评价、本论文研究的命题和技术路线等；绪论应表明博士生对研究方向相关的学科领域有系统深入的了解，论文具有先进性和前沿性；

为了反映出作者确已掌握了坚实的基础理论和系统的专门知识，具有开阔的科学视野，对研究方案作了充分论证，绪论应单独成章，列为第一章，绪论的篇幅应达 1 ~ 2 万字，不得少于 1 万字；绪论引用的文献应在 100 篇以上，其中外文文献不少于 60 应按正文中引用的先后排列。

A.6.4 正文

论文的正文是核心部分，占主要篇幅。正文必须实事求是，客观真切，准确完备，合乎逻辑，层次分明，简便可读。

正文的每一章 (除绪论外) 应有小结，在小结中应明确阐明作者在本章中所做的工作，特别是创新性成果。凡本论文要用的基础性内容或他人的成果不应单独成章，也不应作过多的阐述，一般只引结论、使用条件等，不作推导；

A.6.4.1 图

图包括曲线图、构造图、示意图、图解、框图、流程图、记录图、布置图、地图、照片、图版等。

图应具有“自明性”，即只看图、图题和图例，不阅读正文，就可以理解图意。

图应编排序号。每一图应有简短确切的图题，连同图号置于图下。必要时，应将图上的符号、标记、代码，以及实验条件等，用最简练的文字，横排于图题下方，作为图例说明。

曲线图的纵、横坐标必须标注“量、标准规定符号、单位”。此三者只有在不必要标明 (如无量纲等) 的情况下方可省略。坐标上标注的量的符号和缩略词必须与正文一致。

照片图要求主题和主要显示部分的轮廓鲜明，便于制版。如用放大缩小的复制品，必须清晰，反差适中。照片上应该有表示目的物尺寸的标度。

A.6.4.2 表

表的编排，一般是内容和测试项目由左至右横读，数据依序竖排。表应有自明性。

表应编排序号。

每一表应有简短确切的表题，连同标号置于表上。必要时，应将表中的符号、标记、代码，以及需要说明事项，以最简练的文字，横排于表题下，作为表注，也可以附注于表下。表内附注的序号宜用小号阿拉伯数字并加圆括号置于被标注对象的右上角，如：x x x1)，不宜用“*”，以免与数学上共轭和物质转移的符号相混。

表的各栏均应标明“量或测试项目、标准规定符号、单位”。只有在无必要标注的情况下方可省略。表中的缩略词和符号，必须与正文中一致。

表内同一栏的数字必须上下对齐。表内不宜用“同上”，“同左”和类似词，一律填入具体数字或文字。表内“空白”代表未测或无此项，“—”或“□”(因“—”可能与代表阴性反应相混)代表未发现，“0”代表实测结果确为零。

如数据已绘成曲线图，可不再列表。

A.6.4.3 数学、物理和化学式

正文中的公式、算式或方程式等应编排序号，序号标注于该式所在行(当有续行时，应标注于最后一行)的最右边。

较长的式，另行居中横排。如式必须转行时，只能在 $?$ ， $?$ ， $?$ ， $?$ ， $<$ ， $>$ 处转行。上下式尽可能在等号“=”处对齐。

小数点用“.”表示。大于 999 的整数和多于三位数的小数，一律用半个阿拉伯数字的小间隔分开，不用千位撇。对于纯小数应将 0 列于小数点之前。

示例：应该写成 94 652.023 567；0.314 325，不应写成 94, 652.023, 567；.314, 325

应注意区别各种字符，如：拉丁文、希腊文、俄文、德文花体、草体；罗马数字和阿拉伯数字；字符的正斜体、黑白体、大小写、上下脚标(特别是多层次，如“三踏步”)、上下偏差等。

A.6.4.4 计量单位

报告、论文必须采用国务院发布的《中华人民共和国法定计量单位》，并遵照《中华人民共和国法定计量单位使用方法》执行。使用各种量、单位和符号，必须遵循附录 B 所列国家标准的规定执行。单位名称和符号的书写方式一律采用国际通用符号。

A.6.4.5 符号和缩略词

符号和缩略词应遵照国家标准的有关规定执行。如无标准可循，可采纳本学科或本专业的权威性机构或学术团体所公布的规定；也可以采用全国自然科学名词审定委员会编印的各学科词汇的用词。如不得不引用某些不是公知公用的、且又不易为同行读者所理解的、或系作者自定的符号、记号、缩略词、首字母缩写字等时，均应在第一次出现时一一加以说明，给以明确的定义。

A.6.5 结论

报告、论文的结论是最终的、总体的结论，不是正文中各段的小结的简单重复。结论应该准确、完整、明确、精炼。在结论中要清楚地阐明论文中有那些自己完成的成果，特别是创新性成果；

如果不可能导出应有的结论，也可以没有结论而进行必要的讨论。可以在结论或讨论中提出建议、研究设想、仪器设备改进意见、尚待解决的问题等。

A.6.6 致谢

可以在正文后对下列方面致谢：

国家科学基金、资助研究工作的奖学金基金、合作单位、资助或支持的企业、组织或个人；

协助完成研究工作和提供便利条件的组织或个人；

在研究工作中提出建议和提供帮助的人；

给予转载和引用权的资料、图片、文献、研究思想和设想的所有者；

其他应感谢的组织或个人。

A.6.7 参考文献表

A.6.7.1 专著著录格式

主要责任者, 其他责任者, 书名, 版本, 出版地: 出版者, 出版年

例: 1. 刘少奇, 论共产党员的修养, 修订 2 版, 北京: 人民出版社, 1962

A.6.7.2 连续出版物中析出的文献著录格式

析出文献责任者, 析出文献其他责任者, 析出题名, 原文献题名, 版本: 文献中的位置。

例: 2. 李四光, 地壳构造与地壳运动, 中国科学, 1973 (4): 400 — 429

参考文献采用顺序编码制, 按论文正文所引用文献出现的先后顺序连续编码。

A.7 附录

附录是作为报告、论文主体的补充项目, 并不是必需的。

1. 下列内容可以作为附录编于报告、论文后, 也可以另编成册;
 - (a) 为了整篇论文材料的完整, 但编入正文又有损于编排的条理和逻辑性, 这一材料包括比正文更为详尽的信息、研究方法和技术更深入的叙述, 建议可以阅读的参考文献题录, 对了解正文内容有用的补充信息等;
 - (b) 由于篇幅过大或取材于复制品而不便于编入正文的材料;
 - (c) 不便于编入正文的罕见珍贵资料;
 - (d) 对一般读者并非必要阅读, 但对本专业同行有参考价值的资料;
 - (e) 某些重要的原始数据、数学推导、计算程序、框图、结构图、注释、统计表、计算机打印输出件等。
2. 附录与正文连续编页码。
3. 每一附录均另页起。

A.8 结尾部分 (必要时)

为了将论文迅速存储入电子计算机, 可以提供有关的输入数据。可以编排分类索引、著者索引、关键词索引等。

参考文献

- [1] , . Nvidia grid. <http://www.nvidia.com/object/cloud-gaming.html>. Accessed: 03/10/2013.
- [2] , . Quagga routing suite. <http://www.nongnu.org/quagga/>. Accessed: 03/09/2013.
- [3] , . Vde - virtual distributed ethernet. <https://vde.sourceforge.net>. Accessed: 03/09/2013.
- [4] Al-Fares, M., Loukissas, A., Vahdat, A., 2008. A scalable, commodity data center network architecture, in: ACM SIGCOMM Computer Communication Review, ACM. pp. 63–74.
- [5] Barabási, A.L., 2003. Linked: How Everything is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. Plume, New York.
- [6] de Sola Pool, I., Kochen, M., 1978. Contacts and influence. Social Networks 1, 5–51.
- [7] Dean, J., 2009. Designs, lessons, and advice from building large distributed systems. Keynote from LADIS 2009 .
- [8] Dobrescu, M., Egi, N., Argyraki, K., Chun, B.G., Fall, K., Iannaccone, G., Knies, A., Manesh, M., Ratnasamy, S., 2009. Routebricks: exploiting parallelism to scale software routers, in: Proceedings of the ACM SIGOPS 22nd symposium on Operating systems principles, ACM. pp. 15–28.
- [9] Greenberg, A., Hamilton, J.R., Jain, N., Kandula, S., Kim, C., Lahiri, P., Maltz, D.A., Patel, P., Sengupta, S., 2009. V12: a scalable and flexible data center net-

- work, in: ACM SIGCOMM Computer Communication Review, ACM. pp. 51–62.
- [10] Gyarmati, L., Trinh, T., 2010. Scafida: A scale-free network inspired data center architecture. ACM SIGCOMM Computer Communication Review 40, 4–12.
- [11] Han, S., Jang, K., Park, K., Moon, S., 2010. Packetshader: a gpu-accelerated software router. ACM SIGCOMM Computer Communication Review 40, 195–206.
- [12] Hu, F., Qiu, M., Li, J., Grant, T., Taylor, D., McCaleb, S., Butler, L., Hamner, R., 2011. A review on cloud computing: Design challenges in architecture and security. Journal of Computing and Information Technology 19, 25–55.
- [13] Kohler, E., Morris, R., Chen, B., Jannotti, J., Kaashoek, M.F., 2000. The click modular router. ACM Transactions on Computer Systems (TOCS) 18, 263–297.
- [14] Lockwood, J.W., McKeown, N., Watson, G., Gibb, G., Hartke, P., Naous, J., Raghuraman, R., Luo, J., 2007. Netfpga—an open platform for gigabit-rate network switching and routing, in: Microelectronic Systems Education, 2007. MSE’07. IEEE International Conference on, IEEE. pp. 160–161.
- [15] Luce, R.D., Perry, A.D., 1949. A method of matrix analysis of group structure. Psychometrika 14, 95–116.
- [16] Marconi, G., 1909. Wireless telegraphic communication. Nobel Prize address.
- [17] Newman, M., Barabási, A.L., Watts, D., 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton, NJ.
- [18] Pallis, G., Vakali, A., 2006. Insight and perspectives for content delivery networks. Communications of the ACM 49, 101–106.
- [19] Varia, J., 2010. Architecting for the cloud: Best practices. Amazon Web Services .
- [20] Watts, D., Strogatz, S., 1998. Collective dynamics of “small-world” networks. Nature 393, 440–442.

简历与科研成果

基本信息

韦小宝，男，汉族，1985 年 11 月出生，江苏省扬州人。

教育背景

2007.9 — 2010.6	南京大学计算机科学与技术系	硕士
2003.9 — 2007.6	南京大学计算机科学与技术系	本科

攻读博士学位期间完成的学术成果

1. Xiaobao Wei, Jinnan Chen, “Voting-on-Grid Clustering for Secure Localization in Wireless Sensor Networks,” in Proc. IEEE International Conference on Communications (ICC) 2010, May. 2010.
2. Xiaobao Wei, Shiba Mao, Jinnan Chen, “Protecting Source Location Privacy in Wireless Sensor Networks with Data Aggregation,” in Proc. 6th International Conference on Ubiquitous Intelligence and Computing (UIC) 2009, Oct. 2009.

攻读博士学位期间参与的科研课题

1. 国家自然科学基金面上项目“无线传感器网络在知识获取过程中的若干安全问题研究”（课题年限 2010.1 — 2012.12），负责位置相关安全问题的研究。
2. 江苏省知识创新工程重要方向项目下属课题“下一代移动通信安全机制研究”（课题年限 2010.1 — 2010.12），负责 LTE/SAE 认证相关的安全问题研究。

致 谢

首先感谢我的母亲韦春花对我的支持。其次感谢我的导师陈近南对我的精心指导和热心帮助。接下来，感谢我的师兄茅十八和风际中，他们阅读了我的论文草稿并提出了很有价值的修改建议。

最后，感谢我亲爱的老婆们：双儿、苏荃、阿珂、沐剑屏、曾柔、建宁公主、方怡，感谢你们在生活上对我无微不至的关怀和照顾。我爱你们！

学位论文出版授权书

本人完全同意《中国优秀博硕士学位论文全文数据库出版章程》（以下简称“章程”），愿意将本人的学位论文提交“中国学术期刊（光盘版）电子杂志社”在《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》中全文发表。《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》可以以电子、网络及其他数字媒体形式公开出版，并同意编入《中国知识资源总库》，在《中国博硕士学位论文评价数据库》中使用和在互联网上传播，同意按“章程”规定享受相关权益。

作者签名：_____

_____年____月____日

论文题名	基于小世界理论的数据中心网络模型				
研究生学号	MG0033011	所在院系	计算机科学与技术系	学位年度	2010
论文级别	<input type="checkbox"/> 硕士 <input type="checkbox"/> 博士 <input type="checkbox"/> 硕士专业学位 <input type="checkbox"/> 博士专业学位 (请在方框内画勾)				
作者电话	13671413272		作者 Email	xiaobao.wei@gmail.com	
第一导师姓名	陈近南 教授		导师电话	13671607471	

论文涉密情况：

☐ 不保密

☐ 保密，保密期：_____年____月____日至_____年____月____日

注：请将该授权书填写后装订在学位论文最后一页（南大封面）。