

Review of Distributed Computing Technology

2017

一. 复习大纲

- 第一章：掌握分布式计算、集中式计算、JAVA 三种类型程序、JAVA 线程等基本概念；理解分布式计算与集中式计算的区别和它们的优缺点等；了解分布式计算与相关计算形式的区别；理解 CAP 理论和了解它在 Nosql 数据库的应用；了解 IPC 的概念、类型、接口原型（API），理解 IPC 的事件同步、死锁及其解决方法；掌握事件状态图和顺序状态图的使用方法。
- 第二章：掌握各种分布式计算范型的概念和原理，包括消息传递、客户-服务器范型、p2p 范型、消息系统、远程过程调用、分布式对象、网络服务、移动代理、云服务模型等，特别理解它们的区别。
- 第三章：理解数据包 SOCKET 和流式 SOCKET 的基本原理、及特别关注它们的区别和适用场景等；掌握两种 SOCKET 的程序设计方法和步骤等；理解客户服务器范型的概念和原理，掌握三层软件架构的开发方法，理解面向连接服务器与无连接服务器的区别、并发服务器与迭代服务器的区别、有状态服务器与无有状态服务器的区别；掌握使用三层软件架构开发 Echo 和 Daytime 服务器程序和客户端的方法（包括 UML 类图设计）。
- 第四章：理解分布式对象范型的体系结构、RMI 的体系结构；掌握一般 RMI 应用、RMI 客户回调应用、桩下载 RMI 应用的开发步骤和方法；了解 RMI 的 3 个高级特征：client callback, stub downloading and security manager。
- 第五章：理解 HTTP 协议的原理、HTTP 请求、HTTP 响应的格式；理解 CGI 的原理（结合图）和作用，掌握通过 FORM 请求 CGI 的两种方法（包括 GET 和 POST）及客户端传递参数给 CGI 的方式；掌握会话状态数据维护的技术（特别是 cookie 和 session 对象），了解 web 会话的概念；掌握 Applet, Servlet 的基本概念和原理；掌握 Applet 和 Servlet 程序设计的方法；理解 Servlet 与 Webservice 的区别；理解 rmi 和 webservice 的区别；理解一般 CGI 和 Servlet 的区别；了解 SSH 框架与应用开发；了解常用的 Web 开发技术。
- 第六章：理解 P2P 范型；理解 P2P 网络分类；了解 P2P 典型应用系统（如

DHT); 熟悉 P2P 编程实践的方法。

- 第七章：了解 Web Services 的体系结构，理解 Web Services 的技术原理和应用场景，了解 SOAP Web Services 的协议栈，理解 SOAP 协议的基本原理，了解 SOAP Web Services 的开发过程。
- 第八章：了解云计算的概念（三种类型的云服务）；理解云计算的关键技术；熟悉 Google 云计算原理和关键技术；了解亚马逊云服务。
- 第十章：熟悉网络存储、RAID 的基本概念和原理；熟悉云存储技术的分类（从分布式存储方面看）和应用场景；了解 NoSQL 数据库分类及主流产品；了解对象存储技术原理；了解存储技术趋势。
- 第十一章：了解大数据概念和特征；熟悉大数据存储平台和区别(Hadoop、Cassandra、Redis、MongoDB)，重点熟悉 Hadoop 的功能与组件；理解大数据计算模式，重点是 Mapreduce 和 Spark；了解大数据分析处理平台（Impala、HadoopDB）；了解云存储编程和大数据并行计算编程的基本方法。

二. 各章知识点总结

第一章：

- 1) 并行计算与分布式计算区别？
- 2) 分布式计算的核心技术是？
- 3) 解决进程间通信死锁的两种方法？
- 4) 分布式系统的 CAP 理论是什么？

第二章：

- 1) 网络应用中使用的最多的分布式计算范型是？
- 2) 消息传递范型与消息中间件范型异同？
- 3) 一个分布式应用能否使用多个分布式计算范型？
- 4) 抽象层次最低的分布式计算范型是？
- 5) Web Services 是基于网络服务范型吗？
- 6) 什么移动代理范型？与其他范型有何不同？
- 7) 分布式对象方向包括哪些子范型？
- 8) 云服务范型有哪三种服务类型？

第三章：

- 1)数据包 socket 与流式 socket 在传输层有何不同？
- 2)基于数据包 socket 的通信过程？
- 3)基于流式 socket 的通信过程？

- 4)数据包 socket 在应用层能否支持面向连接的通信?
- 5)流式 socket 在应用层能否支持面向连接的通信?
- 6)并发服务器与迭代服务器的区别?
- 7)会话 (session) 是什么?
- 8)网络应用(客户服务器应用)开发的三层软件架构是?
- 9) HTTP 服务器是有状态的还是无状态的?
- 10)迭代服务器必然是面向连接的服务器吗?

第四章:

- 1) 分布式对象范型有哪些具体子范型或协议?
- 2) RMI 应用的开发步骤?
- 3) 基于 RMI 与基于 SOCKET 开发分布式应用有什么区别?
- 4) RMI 对象能否被多个客户端并发访问?
- 5) RMI 客户回调有什么作用?
- 6) 桩下载(Stub downloading)可以解决 RMI 的什么问题? 桩下载 RMI 应用的执行流程?
- 8) RMI 客户回调应用开发与一般 RMI 应用开发有什么区别?

第五章:

- 1)HTTP 请求和响应命令的组成与格式?
- 2)HTTP 请求中常用的请求命令有哪些?
- 3)CGI 是什么作用? 可以编写 CGI 程序的语言有哪些?
- 4)在 web Form 中请求 CGI 有哪两种方法? 它们传递参数给 CGI 的方式分别是?
- 5)在客户端维护会话状态数据和服务器端维护会话状态数据各有什么优缺点?
- 6)JAVA 的三种类型程序的区别?
- 7)Servlet 程序与一般 CGI 的区别?
- 8)最常用的两种 WEB 会话跟踪技术 (维护) 是什么?
- 9)动态网页开发语言有哪些?
- 10) SSH 框架的作用是什么?
- 11)Web 开发技术有哪些?

第六章:

- 1) 什么是 P2P 范型?
- 2) P2P 网络的分类有哪些?
- 3) 常用的 DHT 网络系统有哪些?

第七章:

- 1) 开发 Web Services 的编程语言有哪些?
- 2) Web Services 的请求和响应是基于什么协议传输的?
- 3) 基于 SOAP 的 Web Services 的相关协议有哪些?
- 4) SOAP 消息是由哪些元素组成的 XML 模式?

第八章:

- 1) 云计算有哪三种服务类型?

- 2) 云计算有哪些关键技术?
- 3) Google 云计算平台有哪些关键技术?
- 4) 亚马逊提供哪些云计算服务?

第十章:

- 1) 网络存储有哪三种组网形态?
- 2) 磁盘阵列有哪些 RAID 级别?
- 3) 分布式存储有哪些分类?
- 4) 对象存储与传统存储有何不同?
- 5) 存储技术有哪些趋势?

第十一章:

- 1) 大数据存储平台有哪些? Hadoop 的功能与组件?
- 2) 大数据计算模式 Mapreduce 和 Spark 有何异同?
- 3) 大数据分析处理平台有哪些?
- 4) 大数据研究趋势有哪些?

三. 样题

A. Fill in the blank answers:

1. Please give five kinds of distributed computing paradigm: (1) _____ (2) _____ (3) _____ (4) _____ (5) _____.
2. The core technology of distributed computing is (1) _____.
3. When communication is from one process to a single other process, the IPC is said to be a (1) _____. When communication is from one process to a group of processes, the IPC is said to be a (2) _____.
4. Sockets that use UDP for transport are known as (1) _____, while sockets that use TCP are termed (2) _____.
5. A connection-oriented server can be threaded so that it can service multiple clients concurrently. Such a server is said to be a (1) _____. An unthreaded connection-oriented server is said to be an (2) _____.
6. Java RMI's three advanced features are: (1) _____ (2) _____ (3) _____.
7. A web form has two methods: (1) _____ and (2) _____ sending the query string to the server.
8. In the Java RMI architecture, the server-side proxy and the client-side proxy are (1) _____ and (2) _____.
9. (1) _____ is the most fundamental paradigm for distributed applications.
10. (1) _____ is the best known paradigm for network applications.
11. In (1) _____ paradigm, the participating processes play equal roles.
12. There are two subtypes of message system models: (1) _____ and (2) _____ message model.
13. (1) _____ is the object-oriented equivalent of RPC.
14. (1) _____ paradigm offers the abstraction for a transportable program or object.

15. In the stream socket API, the (1) _____ class is for the establishment of connections, while the (2) _____ class is for the transfer of data.
16. In the basic socket API, the send operation is (1) _____ and the receive operation is (2) _____. (noblocking or blocking)
17. In the client-server model, we use the term (1) _____ to refer to the interaction between the server and one client.
18. The functionalities of a distributed application can be classified in three layers: (1) _____ (2) _____ (3) _____
19. In the development RMI application, the stub and skeleton are generated from the implementation of a (1) _____ interface using a tool: (2) _____.

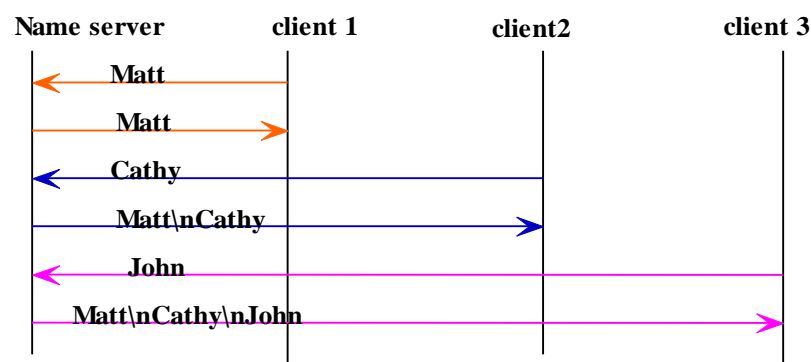
B. True or False.

- () 1. Stream sockets can support both connectionless and connection-oriented communication at the application layer.
- () 2. Datagram sockets can support both connectionless and connection-oriented communication at the application layer.
- () 3. Iterative server can service multiple clients concurrently.
- () 4. Concurrent server can service multiple clients concurrently.
- () 5. The distributed objects paradigm is action-oriented.
- () 6. The message-passing paradigm is data-oriented.
- () 7. Client-server is the most fundamental paradigm for distributed applications.
- () 8. HTTP is a connection-oriented, stateless protocol.
- () 9. HTTP is a text-based request-response protocol.
- () 10. The CGI script can dynamically generate web contents.
- () 11. Sending the query string to the CGI program with the GET method, the CGI program will receive the encoded form input from the standard input.
- () 12. Sending the query string to the CGI program with the POST method, the CGI program will receive the encoded form input from environment variable.
- () 14. Callback allows the server initiates a remote method invocation to the client.
- () 15. In the stateless server, the session state information is maintained by the client, so it is more complex to design and implement than the stateful server.
- () 16. A CGI program can be written in Perl.
- () 17. A CGI program can be written in SQL.
- () 19. Well known Internet services include HTTP, FTP, DNS, are client-server applications.
- () 20. Well known Internet services include HTTP, FTP, are peer-to-peer applications.
- () 21. The Accept() method of ServerSocket class is a blocking operation
- () 22. The Accept() method of ServerSocket class is a nonblocking operation

C. Question

1. Draw a sequence diagram for the daytime protocol.

2. Is it possible for a daytime client to be blocked indefinitely? Explain.
3. Process 1 sends three messages successively to process 2. What is the possible order in which the messages may arrive at process 2 if
 - a. Connectionless socket is used to send each message?
 - b. connection-oriented socket is used to send each message?
4. Write a Java code fragment that may appear in a main method to open a datagram socket for receiving a datagram of up to 100 bytes, timing out in 5seconds. If a timeout does occur, a message “timed out on receive” should be displayed on screen.
5. Using the three-tier software architecture presented in this chapter, design and implement a client-server suite for the following protocol (it is not a well-known service): Each client sends to the server a name. The server accumulates the names received from successive clients (by appending each, with a linefeed ('\n'), to a static string). Upon receiving a name, the server sends the names that it has collected to the client. The figure below illustrates the sequence diagram of a session of the protocol.



- a. Is this server a stateful server? If so, what kind of state information (global or session) does it maintain?
- b. Create one or more of the following suites for the protocol:
 - i. Connectionless server;
 - ii. Connection-oriented, iterative server,
 - iii. Connection-oriented, concurrent server.
6. Please describe the Java RMI architecture. What is the role of the RMI Registry?
7. Please describe the steps for building an RMI application.
8. Please describe two schemes that make use of client-side facilities to maintain session data.
9. Please describe three types of Java programs.

10. Please give two ways to create a Java thread.
11. Please give two examples of distributed computing application.
12. Please describe the concept of distributed computing
13. Please give an example of a deadlock during interprocess communications(involving send/receive operations).