

# 哈尔滨工业大学

**2021年**

## 硕士研究生入学考试试题

考试科目代码: [854]

考试科目: 计算机基础

考生注意: 答案务必写在答题纸上, 并标明题号。答在试题上无效。

题号									总分
分数									150

### 第一部分 计算机系统 70 分

#### 一、单项选择题 (每小题 2 分, 共 20 分)

- 下列叙述错误的是( )。
  - 在一条指令执行周期中, CPU 至少访问一次内存
  - CPU 总是执行 CS::RIP 所指向的指令
  - call、ret 指令通过修改 CS: : RIP 的数值实现函数调用和跳转
  - "mov \$(123\*456+23/3), %eax" 不是合法的 X86-64 汇编指令
- 关于 C 语言中数据类型转换的叙述, 错误的是( )。
  - 从 short 向 int 转换时, 会进行符号扩展
  - 从 unsigned short 向 int 转换时, 会进行 0 扩展
  - 从 int 向 unsigned int 转换时, 会重新编码, 将最高符号位设置为 0
  - 从 int 到 unsigned short 转换时, 多出的位直接被截断丢弃
- C 语言程序中, x 是 int 类型变量, ux 是 unsigned int 类型变量, 下面叙述错误的是( )。
  - $(x*x) < 0$  可能为真
  - $(x + ux) \geq 0$  总为真
  - $ux \geq 0$  总为真
  - 若  $x < 0$ , 则  $-x > 0$  为真
- 条件跳转指令 JE 依据标志位( )来判断是否跳转。
  - ZF
  - OF
  - SF
  - CF

5. X86-64 汇编指令“add \$0x12ab,%ebx”的 16 进制指令编码的字节序列是“81 c3 \*\*\*”，其中源操作数“\*\*\*”应该是( )。
- A. ab 12    B. 12 ab    C. ab 12 00 00    D. 00 00 12 ab
6. 关于高速缓存的叙述，错误的是( )。
- A. 全相联高速缓存的结构最复杂，命中率更高，但命中时间长  
B. 组相联高速缓存能更好地在命中率、命中时间两方面折中  
C. 缓存命中率比不命中率能更直观地衡量系统的性能  
D. 提高程序的局部性，就能提高缓存命中率，从而优化性能
7. 链接时，两个目标文件中有同名的一个强符号和一个弱符号，且类型不同，链接器将选用( )做重定位。
- A. 与引用类型一致的符号    B. 强符号    C. 弱符号    D. 任意一个
8. CPU 中有专门的寄存器 CR3 存储一级页表的基地址 该基地址是( )。
- A. 虚拟地址    B. 物理地址    C. 虚拟页号    D. 物理页号
9. 关于虚拟页和页表的叙述，错误的是 ( )。
- A. 虚拟页可以从任意的虚拟地址开始划分  
B. 虚拟页和物理页大小相同  
C. 多级页表最能节省内存空间  
D. 多级页表中，访问最频繁的页表项在一级页表中
10. 下列关于信号的叙述，错误的是( )。
- A. 信号处理程序也能被其他信号处理程序中断  
B. 信号处理程序如果修改全局变量会有风险  
C. 在信号 k 的处理程序执行期间，新到达的信号 k 将被阻塞并计数  
D. 键盘输入 Ctrl+C，内核会给前台进程组中的每个进程发送一个 SIGINT 信号

## 二、填空题(每小题 1 分，共 20 分)

11. 在 X86-64 计算机中，若地址 0x1200~0x1207 的 8 个内存字节的数值依次是 0x50、0x51、0x52、0x53、0x54、0x55、0x56、0x57，寄存器%rbx=0x1020304050607080 则：  
执行汇编指令“mov \$0x1200,%ebx”后，寄存器 rbx 的数值是( )，  
继续执行指令“mov 2(%ebx),%eax”后，寄存器 ah 的数值是( )，  
继续执行指令“add (%ebx),%eax”后，寄存器 ah 的数值是( )。
12. C 语言函数 long proc(long p1, long p2, int \*p3)被调用时，它的三个参数依次通过( )、( )和( )传递，而该函数的返回值通过( )返回。若函数有大于 6 个整型参数，超出 6 个的部分要通过( )来传递。

13. Linux 为每个进程维护一个内存区域结构链表, 当对一个内存地址的读/写操作发生缺页时, 系统检查该链表, 如果( ), 可判定发生非法内存访问/段错误(segmentation fault), 并终止该进程; 若多进程共享物理内存, 且希望一个进程对该内存的修改不影响其他进程, 即该修改对其他进程不可见, 需要将该内存区域设置为( )类型。
14. 若 8 路组相联 L1 d-cache 共 128KB、块大小 8 字节, 系统的物理地址是 52 位、虚拟地址是 48 位, 则 L1 d-cache 的块内偏移是( )位、组索引是( )位、tag 标记位是( )位。
15. 给定系统的物理地址 45 位、虚拟地址 42 位, 页面大小 2KB, 使用 4 级页表, 虚拟页号 VPN2、VPN3 和 VPN4 的长度都是 8 位, 则 VPN1 的长度是( )位, PPN 的长度是( )位。如所有页表项长度均为 8 字节, 每个三级页表项可以映射( )MB 内存区域, 每个一级页表项可以映射( )GB 内存区域。
16. 如果 C 语言程序在动态申请的内存使用完毕后没有释放, 会导致( ), 而具有( )机制的编程语言不存在这个问题。
17. 将源代码 hello.c 编译生成汇编语言程序的指令是: ( )。

### 三、分析题(每题 10 分, 共 20 分)

18. 下表中左侧是 C 语言函数, 右侧是其对应的 X86-64 汇编语言函数:

表 1 C 语言程序与对应的汇编语言程序

<pre>int myproc(int x[], int n, int k) {     int val = 0;     int m=0;     int i;     for(i=0; i ①, i ② )     {         val= ③ ;         m= ④ ;     }     if( ⑤ )         val /= m ;     return val; }</pre>	<pre>myproc.     movl    \$0, %ecx     movl    \$0, %r8d     movl    \$0, %eax     jmp    .L2 .L3:     movslq  %ecx, %r9     addl   (%rdi,%r9,4), %eax     addl   \$1, %r8d     addl   %edx, %ecx .L2:     cmpl   %esi, %ecx     jl    .L3     testl  %r8d, %r8d     jle   .L1     cld     idivl  %r8d .L1:     ret</pre>
--	---

(1) 为表 1 中汇编代码写出每条指令的注释。(5 分)

(2) 将表 1 中 C 语言函数补充完整。(5 分)

19. 顺序结构 Y86 CPU 的抽象结构如图 1 所示, 该 CPU 每个周期执行一条指令, CC 是三位二进制编码, 从高到低每位分别对应 ZF、SF、OF。CC 初值为 100, 按图 2 顺序执行指令序列。

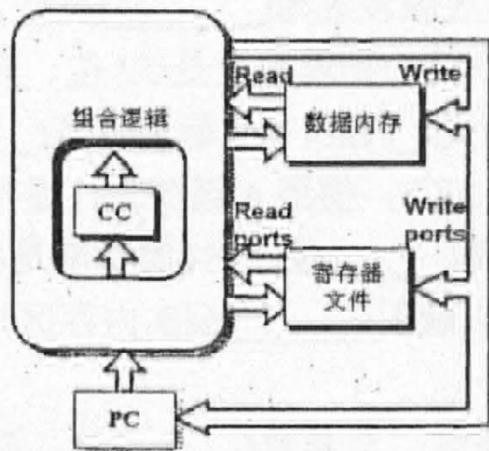


图 1 顺序结构 Y86 CPU 示意图



周期 1:	0x000: irmovq \$0x123,%rbx
周期 2:	0x00a: irmovq \$0x210,%rdx
周期 3:	0x014: addq %rdx,%rbx
周期 4:	0x016: je dest
周期 5:	0x01f: rmmovq %rbx,0(%rdx)

图 2 CPU 执行的指令序列

(1) 在图 1 中, 组合逻辑和状态单元(存储设备)分别包含哪些器件? 周期 3 结束前一刻即图 2 中时刻①时, CC、PC、rbx 的数值分别是多少, 组合逻辑是哪条指令的结果? 在周期 4 开始的后一刻, 即时刻②时, CC、PC、rbx 的数值是多少? (5 分)

(2) 流水线 CPU 的基本设计思路是什么, 划分的阶段数过多会有什么问题? (5 分)

#### 四、综合设计题(10 分)

20. 动态内存分配器使用带边界标记的块格式、隐式空闲链表, 其 C 语言代码的部分片段如下:

```
#define WSIZE      4 /*字长是 4 字节*/
#define DSIZE      8 /*双字长是 8 字节*/
#define OVERHEAD  8 /*头部和脚部总长 8 字节*/
#define MAX(x, y) ((x) > (y) ? (x) : (y))
#define PACK(size, alloc) ((size) | (alloc))
#define GET(p)      (*(size_t*)(p))
#define PUT(p, val) (*(size_t*)(p) = (val))

#define GET_SIZE(p) (GET(p) & ~0x7)
#define GET_ALLOC(p) (GET(p) & 0x1)
#define HDRP(bp)    ((char*)(bp) - WSIZE)
#define FTRP(bp)    ((char*)(bp) + GET_SIZE(HDRP(bp)) - DSIZE)
#define NEXT_BLKP(bp) ((char*)(bp) + GET_SIZE(((char*)(bp) - WSIZE)))
#define PREV_BLKP(bp) ((char*)(bp) - GET_SIZE(((char*)(bp) - DSIZE)))

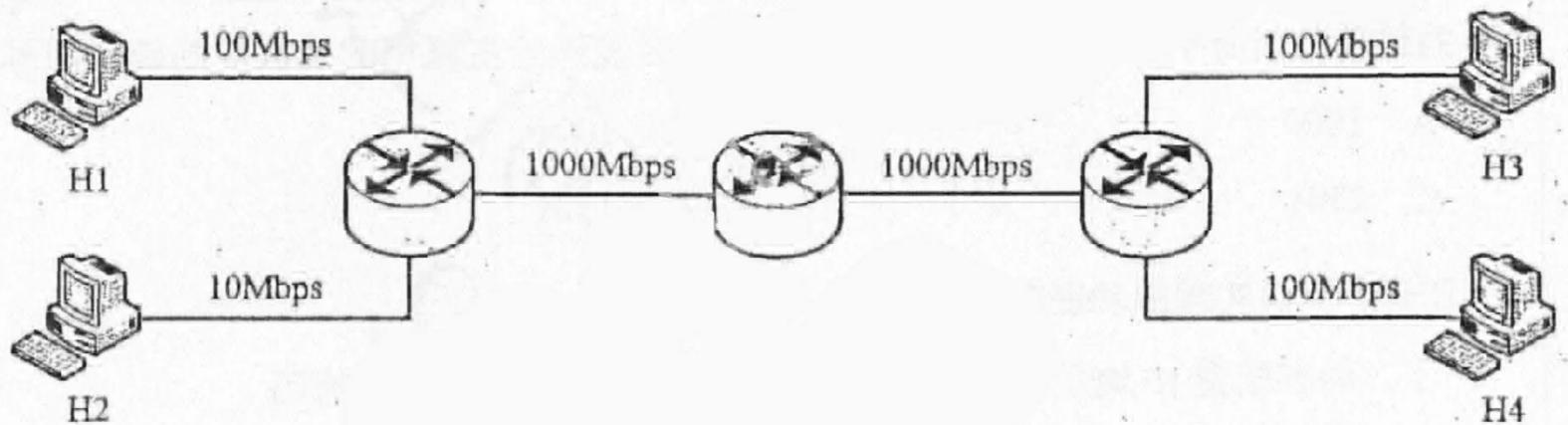
void mm_free_coalesce(void *bp){//释放内存块 bp, 并进行空闲块合并
.....}
```

请简述在内存块释放并合并空闲块的过程需要完成的处理? 写出内存块释放及空闲块合并的 C 语言函数 mm\_free\_coalesce() 并添加关键注释(可使用上述宏和宏函数)。

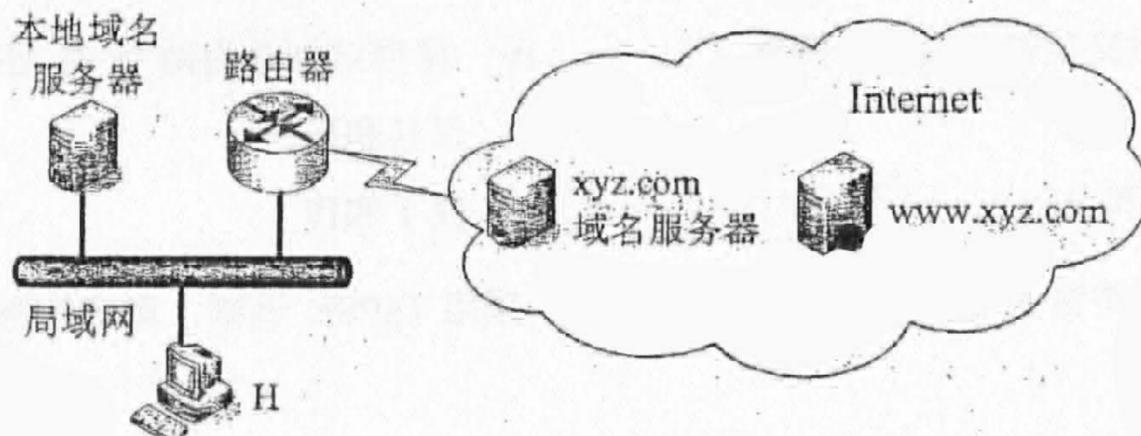
## 第二部分 计算机网络 40 分

### 五、单项选择题 (每小题 2 分, 共 20 分)

1. 在 OSI 参考模型和 TCP/IP 参考模型中, 实现数据表示转换功能的层分别是  
A. 应用层, 应用层  
B. 表示层, 应用层  
C. 表示层, 传输层  
D. 传输层, 传输层
2. 如下图所示分组交换网络, 分组长度为 1000B。若 H1 向 H3、H2 向 H4 同时分别发送 1 个大小为 10 MB 的文件, 则从开始发送时刻起, 到 H3 和 H4 收到文件为止, 所用时间至少分别约为



- A. 0.1s, 1s      B. 1s, 1s      C. 0.8s, 8s      D. 8s, 8s
3. 假设下图所示网络中的本地域名服务器无缓存, 局域网内主机访问 Internet 上各服务器的往返时间  $RTT=10ms$ , 忽略其他各种时延。若主机 H 通过超链接 <http://www.xyz.com/index.html>, 请求浏览纯文本 Web 页 `index.html`, 则从点击超链接开始到浏览器接收到 `index.html` 页面为止, 所需的时间至少是

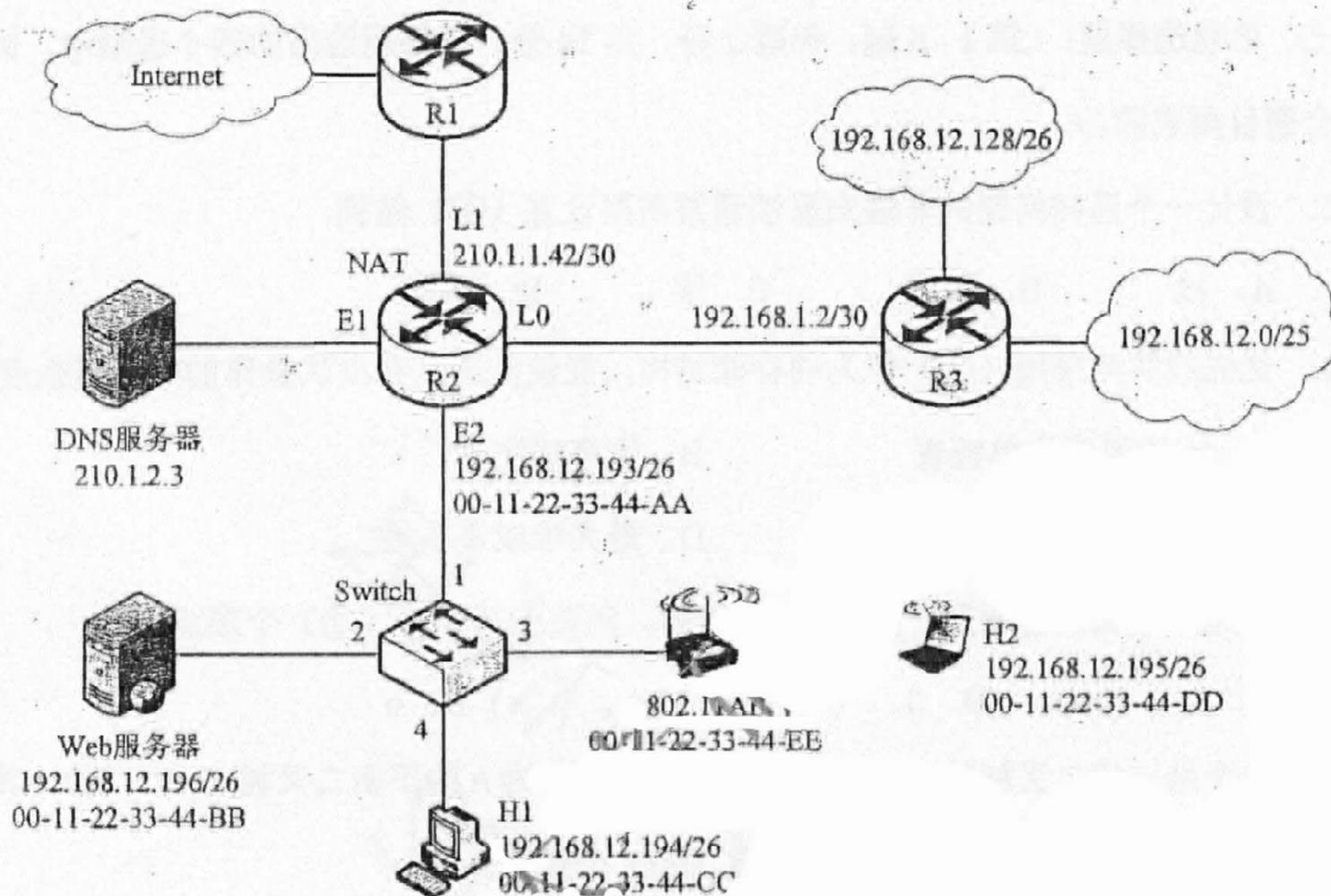


- A. 10 ms      B. 20 ms      C. 30 ms      D. 40 ms
4. 在 Internet 邮件系统中, 邮件服务器之间发送和接收邮件的应用层协议是  
A. HTTP      B. IMAP      C. POP3      D. SMTP

5. 主机甲与主机乙通过一条链路带宽为 10 Mbps ( $M=10^6$ ) 的链路互连, 单向传播时延为 7.6 ms。若甲采用滑动窗口协议向乙发送数据, 数据帧长为 1000B, 忽略确认帧帧长, 乙的接收窗口为 3, 数据帧序号为 3 比特编号, 则甲的最大信道利用率约为
- A. 15%                      B. 25%                      C. 29%                      D. 48%
6. 若主机甲与主机乙已建立一条 TCP 连接, 最大段长 (MSS) 为 1 KB, 往返时间 RTT=5ms, 则拥塞窗口从 4KB 增长到 32KB 所需的时间至少是
- A. 15 ms                      B. 20 ms  
C. 140 ms                      D. 145 ms
7. 主机甲与主机乙之间已建立一个 TCP 连接, 在主机乙向主机甲发送了确认序号 ack\_seq=1000 的 TCP 段后, 依次收到 “seq=1001, data=300” 和 “seq=1301, data=200” (seq 为序号, data 为封装的载荷字节数) 2 个段, 则主机乙此时发送给主机甲的确认段的确认序号是
- A. 1000                      B. 1001  
C. 1301                      D. 1501
8. DHCP 协议实现的功能是
- I. 动态配置 IP 地址                      II. 动态配置子网掩码  
III. 动态配置 MAC 地址                      IV. 动态配置默认网关 IP 地址
- A. 仅 I、II                      B. 仅 I、II、III  
C. 仅 I、II、IV                      D. I、II、III、IV
9. 如果期望将同属于一个广播域的两个网络分割为两个独立的较小广播域网络, 可以选择的方案是
- I. 利用路由器互联两个网络                      II. 利用网桥互联两个网络  
III. 利用交换机互联两个网络                      IV. 将网络划分为两个 VLAN
- A. 仅 I 和 II                      B. 仅 II 和 III  
C. 仅 III 和 IV                      D. 仅 I 和 IV
10. 若某信道带宽为 2MHz, 信噪比为 30dB, 采用 16PSK 调制, 则该信道的极限码元速率 (波特率) 是
- A. 5M 波特                      B. 10M 波特  
C. 20M 波特                      D. 80M 波特

六、综合应用题（共 20 分）

某网络拓扑如下图所示，图中 R1、R2、R3 为路由器；Switch 为 100Base-T 以太网交换机。



R2 的路由表结构为：

目的网络	子网掩码	下一跳 IP 地址	接口
------	------	-----------	----

请回答下列问题。

- 请给出 R2 的路由表，要求路由表项尽可能少。
- 若交换机 Switch 的交换表为空时，主机 H2 向 H1 发送一个封装 IP 数据报 P 的 802.11 帧，则该 802.11 帧的地址 1、地址 2 和地址 3 分别是什么？AP 向交换机 Switch 转发的封装 P 的以太网帧的目的 MAC 地址和源 MAC 地址分别是什么？当 H1 接收到 P 时，交换机 Switch 的交换表中包含哪些交换表项？（交换表结构为：<MAC 地址，接口编号>）
- 假设 NAT 转换表结构为：

公网 IP 地址	公网端口号	私网 IP 地址	私网端口号
----------	-------	----------	-------

如果期望外网可以通过默认端口号访问 Web 服务器，请给出一个可行的 NAT 转换表配置。

- 假设 R1 与 R2 之间链路的 MTU=600B，R2 在向 Internet 转发一个总长度为 1500B、头部长度为 20B 的 IP 分组时，进行了分片。若分片时尽可能分为最大片，则至少需要分为几个分片？每个分片的总长度字段和片偏移量字段的值分别是多少？

### 第三部分 数据结构 40 分

七、单项选择题：(第 1-5 题，每题 2 分，共 10 分。在每题给出的四个选项中，请选出一项最符合题目要求的。)

1. 设计一个递归问题的非递归算法通常需要设置 (①) 结构。  
A. 栈      B. 数组      C. 堆      D. 队列
2. 优先级队列采用 (②) 作为的存储结构，能使其进队和出队操作的时间复杂度一样。  
A. 无序或有序单链表      B. 无序顺序表  
C. 循环有序顺序表      D. 最大堆或最小堆
3.  $G$  是一个无向连通图，共有 22 条边的，则该图至少有 (③) 个顶点。  
A. 9      B. 8      C. 7      D. 6
4. 一个结点的二叉树的高度为 1。假设一棵高度为  $h$  的平衡二叉树 (AVL 树)，其每个非终端结点的平衡因子均为 0，则该树共有 (④) 个结点。  
A.  $2^{(h-1)}-1$       B.  $2^{(h-1)}$       C.  $2^h-1$       D.  $2^h$
5. 在对长度为  $n$  ( $n>2$ ) 的顺序存储的有序表进行折半查找，查找每个元素的比较次数均 (⑤) 对应的折半查找判定树的高度。  
A. 小于      B. 小于等于      C. 等于      D. 大于等于

八、简答题：(第 6 题，共 15 分。)

6. (15 分) 在一个长度为  $n$  的数组里，所有元素都是  $0 \sim n-1$  范围内的整数。某些元素在数组中可能重复出现，但不知道哪些是重复出现的，也不知道重复出现多少次。现要尽可能快地找出数组中所有重复出现的元素。请回答下列问题：
  - 1) 设计相关的数据结构。
  - 2) 描述求解问题的方法步骤，并说明时间和空间效率。

九、算法设计题：(第 7 题，共 15 分。)

按以下要求设计算法：

- (1) 给出算法的设计思想。
- (2) 使用 C 或 C++ 或 Java 语言，给出相关数据类型定义。
- (3) 根据设计思想，采用 C 或 C++ 或 Java 语言描述算法，栈和队列的操作可以直接使用。

7. (15分) 假设一棵BST(二叉查找树) T 采用二叉链表表示, 其结点包括三个域: data—数据域、left 域—指向其左子树的指针和 right 域—指向其右子树的指针。请设计一个算法 Search(T, p), 对于给定结点 p, 在 T 上查找比 p 的 data 值小且具有最大 data 值的结点。若找到, 则返回该结点的指针; 否则, 返回 NULL。