King's College London

This paper is part of an examination of the College counting toward the award of a degree. Examinations are governed by the College Regulations under the authority of the Academic Board.

Degree Programmes BSc, MSci

Module Code 6CCS3INS
Module Title Internet Systems

Examination Period January 2014 (Period 1)

Time Allowed *2 hours*

Rubric *ANSWER FOUR OF FIVE QUESTIONS*
If more than four questions are answered, the answers to the first four questions, in the order in the exam question paper, will count.

ANSWER EACH QUESTION IN A SEPARATE ANSWER BOOK AND WRITE ITS NUMBER ON THE COVER

Calculators Calculators may be used. The following models are permitted: Casio fx83 / Casio fx85

PLEASE DO NOT REMOVE THIS PAPER FROM THE EXAMINATION ROOM

Question One

(a) An HTTP message containing only the text below is sent over TCP/IP.

GET /index.html HTTP/1.1

Answer the following:

i. State what headers are added to this text for transmission over the internet *(2 marks)*

ii. What issues does each of these headers primarily handle in the communication of the message across the internet? *(4 marks)*

iii. Which internet header is sent and received first? *(1 mark)*

iv. How would the effect of the HTTP request above be different if the "GET" was replaced by "PUT"? *(2 marks)*

v. How does hierarchical addressing help make the internet manageable? *(4 marks)*

[13 marks]

Solution

The intended solution is the following, but see below for an alternative counting for some marks.

i. An IP header and a TCP header are added.

ii. The IP header provides the means to identify the receiver and sender using IP addresses, and handles the limits on network bandwidth using fragmentation. The TCP header handles reliability, through acknowledgements and checksums on data sent, and flow and congestion control, through window sizes.

iii. The IP header is the first to be sent as this is the lowest internet layer, so processed first.

iv. Changing the HTTP method to PUT means that the client is requesting to upload some data to /index.html rather than download the existing data from there. As there is no body to the HTTP message, the request would be to replace the existing file with an empty one.

v. Hierarchical addressing means that administrators are allocated network prefixes (or domain name roots) within which they can allocate addresses by appending suffixes. They know that no-one else in the world is using that prefix, so their addresses will be unique regardless of what suffix is chosen. This means they do

not need to coordinate over the allocation of addresses, which would be infeasible at the global scale.

An alternative solution, which does not really fit the questions as asked but may be considered a fair interpretation, is that the additional headers being referred to are parts of the HTTP header, rather than TCP/IP headers.

i. General, request and entity headers, or common HTTP header lines for GET such as Host, Accept, or User Agent.
ii. General headers are about the transmission, request headers are metadata about the request, entity headers are metadata about the entity. Host identifies the host the request is sent to, Accept identifies the MIME types of acceptable returned data, User Agent identifies the client software sending the request.
iii. General header, then request header, then entity header.
iv. As intended solution above.
v. As intended solution above.

[Delete as appropriate] **SEE NEXT PAGE / FINAL PAGE**

(b) An IPv4 network, N, has been divided into the following subnets.

- Subnet A: 101.55.65.128 /25
- Subnet B: 101.55.65.0 /25
- Subnet C: 101.55.64.128 /25
- Subnet D: 101.55.66.0 /23

Answer the following questions:

i. How many hosts can be addressed in each subnet? *(4 marks)*

ii. In making the above subnet allocation using the standard algorithm, which subnet address would have been allocated first, and why? *(2 marks)*

iii. Given that subnet C has a 3 bit subnet ID, what is the network address of the entire network N? *(2 marks)*

iv. What is the broadcast address of subnet B? *(2 marks)*

v. If the unallocated block of addresses in network N was allocated to a subnet, what would be the range of host IP addresses in that subnet? *(2 marks)*

[12 marks]

Solution

i. A, B and C can have 126 hosts each (32-25=7 host bit IDs, $2^7 - 2 = 126$), while D can have 510 (32-23=9, $2^9 - 2 = 510$).

ii. D would have been allocated first, because it is the largest.

iii. 101.55.64.0 /22 (because 7+3=10, and so we set the final 10 bits to 0 to get the network address of N)

iv. 101.55.65.127 /23 (set the 7 host ID bits to 1 to get the broadcast address)

v. 101.55.64.1 – 101.55.64.127 (the range of host addresses with subnet ID set to 000 is the block not yet allocated to a subnet)

Question two

(a) An IPv4 datagram with 1000 bytes of message data, Fragmentation Offset field set to 50, More Fragments flag set to 1, and Identification field set to 2001 (decimal) is divided into two fragments. The first fragment contains 800 bytes of message data. Neither fragment has header options. What are the values of the following fields in each of the fragments?

i. Internet Header Length *(2 marks)*
ii. Total Length *(2 marks)*
iii. Identification *(2 marks)*
iv. More Fragments *(2 marks)*
v. Fragment Offset *(2 marks)*

[10 marks]

Solution

i. 5 (measured in 4-byte blocks) in both fragments, because an IPv4 header with no options contains 20 bytes
ii. 820 in the first fragments (800 bytes of message data + 20 bytes of header), 220 in the second (1000-800=200 + 20 header bytes)
iii. 2001 in both fragments (all fragments of one message have the same Identification field)
iv. 1 in the both fragments (both have more fragments to follow, as the original datagram had more fragments to follow)
v. 50 in the first fragment (original FO of 50 and no further offset), 150 in the second fragment (original FO of 50 plus 800/8=100 offset of 8-byte chunks from first fragment)

If the "message data" was interpreted as including the header- not conventional but plausible- the answers would be adjusted as follows.

i. As above
ii. 800 in the first fragment (780 bytes of message data + 20 bytes header), and 220 (200 bytes of message data + 20 bytes header)
iii. As above
iv. As above
v. 50 in the first fragment, 147/148 in the second (doesn't work in practice as 780 is not a multiple of 8)

[Delete as appropriate] **SEE NEXT PAGE / FINAL PAGE**

(b) In a TCP session, the client sends the server a series of segments, each containing 10 bytes of message data. The server responds with five messages in which the Acknowledgement field is set to 111, 111, 111, 111, and 151 respectively.

Explain what plausibly has occurred between the client and server to cause these acknowledgements to be sent *(10 marks)*. This explanation should include stating the values of the Sequence Number field in each of the segments sent by the client *(5 marks)*.

[15 marks]

Solution

The following is a plausible chain of events:

- The client sends the first segment with Sequence Number 101. The server returns an acknowledgement that it has bytes up to but not including sequence number 111.
- The client sends the second segment with Sequence Number 111, but it is lost in transmission.
- The client sends three further segments with Sequence Numbers 121, 131, and 141, and each time the server acknowledges that it has bytes up to but not including sequence number 111.
- The client determines from the repeated acknowledgements that the second segment was lost, and so re-sends the segment with Sequence Number 111.
- The server receives the re-sent segment and acknowledges that it now has all bytes up to but not including sequence number 151.

Question three

(a) Decode the base64 string “ey1D” back into 8-bit bytes, giving the decimal values of those bytes.

(5 marks for the method, 2 marks for the correct decimal values)

base64 encoding table

0	A	8	I	16	Q	24	Y	32	g	40	o	48	w	56	4
1	B	9	J	17	R	25	Z	33	h	41	p	49	x	57	5
2	C	10	K	18	S	26	a	34	i	42	q	50	y	58	6
3	D	11	L	19	T	27	b	35	j	43	r	51	z	59	7
4	E	12	M	20	U	28	c	36	k	44	s	52	0	60	8
5	F	13	N	21	V	29	d	37	l	45	t	53	1	61	9
6	G	14	O	22	W	30	e	38	m	46	u	54	2	62	+
7	H	15	P	23	X	31	f	39	n	47	v	55	3	63	/

[7 marks]

Solution

- Mapping from each symbol using the table above: 30, 50, 53, 3
- Write these as 6-bit binary: 011110, 110010, 110101, 000011
- Rearrange these as 8-bit binary: 01111011, 00101101, 01000011
- In decimal: 123, 45, 67

[Delete as appropriate] **SEE NEXT PAGE / FINAL PAGE**

(b) A file is downloaded from the following URL:

http://www.example.org:8080/myfiles/page.html

The web server retrieves the file from its local hard disk at path

/local/web/site/myfiles/page.html

Answer the following questions:

i. What is the document root used by the server? *(2 marks)*

ii. On what TCP port is the server running? *(2 marks)*

iii. What is the URI scheme used by the URL? *(2 marks)*

iv. What relative URL would be included in page.html to provide a link to the file stored at /local/web/site/images/photo.jpg ? *(2 marks)*

[8 marks]

Solution

i. /local/web/site

ii. 8080

iii. http

iv. ../images/photo.jpg

(c) For each of the following elements found in a WSDL interface document, state what its content specifies: Description, Schema, Message, Operation, and Port Type *(2 marks each)*.

[10 marks]

Solution

- Description: The whole service interface definition
- Schema: This defines the structure used by messages to and from the service
- Message: Defines a message to be sent to/from a service
- Operation: Links messages, stating that when the service receives one, it will send the other in response
- Port Type: Combines a set of operations of a similar kind, which will be invoked via the same URI

Question four

(a) The following data is sent by HTTP when an HTML form is completed by a user and submitted:

a=b+c&d=e&f=g-h

Write the HTML snippets for input text boxes in a form which would produce such encoded data *(6 marks)*. Give the values that the user would have inputted into your form to produce the above encoded data *(3 marks)*. Explain the differences in the HTTP request sent when the HTML form uses GET or uses POST as its submission method *(4 marks)*.

[13 marks]

Solution

<input name="a" type="text"/>
<input name="d" type="text"/>
<input name="f" type="text"/>

Field a would contain "b c" (the space is converted to +), field d would contain "e", and field f would contain "g-h".

When the form method is GET, the form data is added to the submission URL, as the query part (after the ?). In the HTTP request, this means that the form data is on the resource path on the request line (first). There is no body to the message. When the form method is POST, the form data is put into the body of the HTTP request, with the resource path (URL) unaltered.

(b) The following XML Schema element definition is given, with XXXX replacing a missing element name.

```
<element name = "a">
  <complexType>
    <XXXX>
      <element name = "b" type = "xs:string"/>
      <element name = "c" type = "xs:string"/>
    </XXXX>
  </complexType>
</element>
```

For each of the following XML snippets, say what name XXXX must be in order for that snippet to be valid according to the schema.

i. <a><c>x</c>y *(2 marks)*

ii. <a><c>x</c> *(2 marks)*

Next, write an absolute XPath to return the text content of the `c` element in a document whose root element is an `a` element *(2 marks)*.

Finally, alter the element definition so that

<a>yyyyy<c>123</c>

is valid, but

<a><c>123</c>y and

<a>y<c>hello</c>

are not valid, according the schema *(6 marks)*.

[12 marks]

Solution

i. XXXX = all

ii. XXXX = choice

XPath: /a/c/text()

A solution to the latter part is to make changes marked in bold below:

```
<element name = "a">
  <complexType>
    <sequence>
      <element name = "b" type = "xs:string"
```

```
        maxoccurs = "unbounded"/> (or could be "5"+)
      <element name = "c" type = "xs:integer"/>
    </sequence>
  </complexType>
</element>
```


[Delete as appropriate] **SEE NEXT PAGE / FINAL PAGE**

Question five

(a) Alice wishes to send a message to Bob. Alice acquires a digital certificate from a website, where the certificate allegedly gives the credentials, including public key, of Bob. Explain the steps conducted by Alice (or her security software) to verify that the credentials are definitely those of Bob, and then send the message to Bob such that only Bob can read it.

[8 marks]

Solution

Alice reads the certificate to determine which certificate authority (CA) signed the certificate. Alice obtains the public key of the CA, and decrypts the certificate's signature. Alice then creates a digest of the certificate credentials, and compares this to the decrypted signature. If they match, then this indicates that the credentials are genuine, as certified by the CA. To send the message to Bob, Alice encrypts it with Bob's public key, obtained from the certificate. Only Bob can decrypt this, because only Bob has the corresponding private key.

(b) Generate a matching RSA public key and private key from the primes 3 and 11. Show your working.

[7 marks]

Solution

- Product: n = 3 x 11 = 33
- Totient: m = 2 x 10 = 20
- Co-prime to m: e = 3
- Integers d such that d=(1+im)/e for some i: d = 7 (with i=1)
- Key pairs: (33, 3) and (33, 7)

Other values of e and d are possible.

(c) Write some Turtle RDF statements for which, if the SPARQL query below is used to search them, a value of ex:Simon for ?name and a value of ex:London for ?city would be one of the results returned *(4 marks)*.

Next, write two RDF statements to state that the property ex:inCity relates an address (class ex:Address) to a city (class ex:City). *(4 marks)*

Finally, write a statement that ex:London is of the class ex:City. *(2 marks)*

```
PREFIX ex: <http://www.example.com/>
SELECT ?name ?city
WHERE
      { ?name  ex:livesIn  ?address .
        ?address  ex:inCity  ?city  }
```

[10 marks]

Solution

ex:Simon ex:livesIn ex:simonsAddress .
ex:simonsAddress ex:inCity ex:London .

(ex:simonsAddress can be any URI or a blank node)

ex:inCity rdfs:domain ex:Address .
ex:inCity rdfs:range ex:City .

ex:London rdf:type ex:City . (or ex:London a ex:City .)