

Instructions: There are **five** questions which count 100 points in total. Each question may have several sub-questions. Please read the questions carefully before answering.

1. (24 pts) A virtual memory system can be structured on fixed-size pages and variable-size segments.
 - A. (4 pts) Compare the two approaches; describe one scenario the paging approach work better than the segment approach.
 - B. (4 pts) Following the previous question, describe one scenario the segment-based approach works better than the paging approach.
 - C. (4 pts) Please explain the different between internal fragmentation and external fragmentation.
 - D. (4 pts) Which one of the above two fragmentations occurs in paging systems? Please explain your answer.
 - E. (8 pts) Describe two tradeoffs involved in having a design of smaller pages versus larger pages?

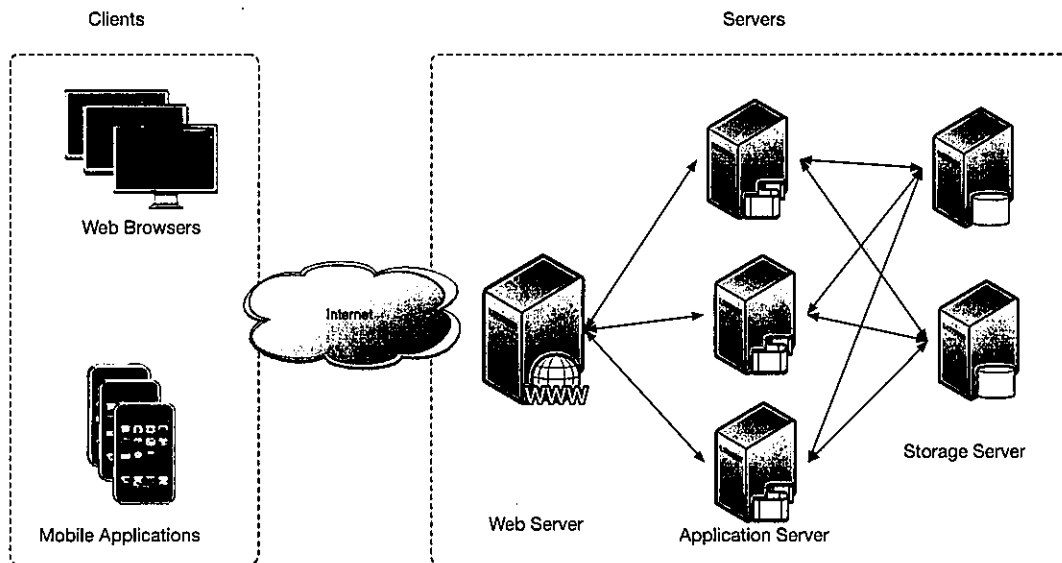
2. (26 pts) Two common approaches to structure an operating system kernel: Monolithic kernel vs. microkernel. A monolith approach puts all operating system into one executable while a microkernel approach puts only allows bare essential functions to run in privileged mode.
 - A. (8 pts) which approach provides better flexibility and why?
 - B. (8 pts) which approach provides better scalability and why? (Scalability refers to the ability of supporting large number of processors/cores in the system.)
 - C. (10 pts) which approach provides better security and why?

3. (10 pts) In 2014, Taiwan Highway Commission is officially discontinuing the highway toll ticket system in favor of an electronic toll collection system, called ETC for short. The service company, Far Eastern Electronic Toll Collection (FETC) systems, develops several services to allow the users to query historical toll charge, the balance of their account if there is any, and the estimated charge via mobile applications and web services. In this question, we will go through the software design exercise to design a highly available and responsive service via proper system services and operator system configuration.

The client server model shown below is the general application model to build such services. The users use mobile applications or web browser to use the services provided by FETC. To serve the queries from mobile applications and web sites, a three-tier server model is adopted. In the three-tier server model, there are web server, application servers, and storage servers. The web servers accepts the requests from mobile applications or web browsers, formats the query results in proper ways, and returns the formatted results. The applications conduct the computation logic, query the database server for required data, and compute the results. The storage servers store the user data, vehicle information, and charge to the vehicle.

On January 1st and 2nd, the web services were fail to respond to user's requests, either from mobile applications or web site.

- A. (5 pts)The service company's first reaction was to increase the network bandwidth of the systems. Under what condition the increased bandwidth may resolve the problem.
- B. (5 pts)The service remained stalled after the first reaction. What's the possible cause of the systems?



4. (20 pts)(Continued from Question 4.) On the client side, the (mobile or web) application issues the requests, receives the requests from the server and presents the results to the user. Please answer the following questions.
 - A. (10 pts) The application may use either blocking or nonblocking calls to send the requests. Please describe these two types of function calls including process behaviour, data transmission, and storage access of both calling and callee processes.
 - B. (10 pts) Suppose blocking functions are used to issue requests. When the servers are too busy to process the requests, the application will be stalled. Signals are usually used to stop the request. Please describe how to stop the request by signals.

5. (20 pts) Continued from Question 4.) All these servers are supposed to accept simultaneous requests, handled by either multiple processes or multiple threads server processes. Please answer the following questions.
 - A. (10 pts) When the services are implemented as multiple process applications. Please describe the difference between `fork()` and `vfork()` defined in POSIX standard (or IEEE Std 1003) to create child processes and handle incoming requests. You should discuss their memory layout, scheduling policy, and data sharing between parent and child processes.
 - B. (10 pts) To handle the simultaneous requests and avoid blocking the request, should you use `fork` or `vfork` to create child processes? Please discuss your design.