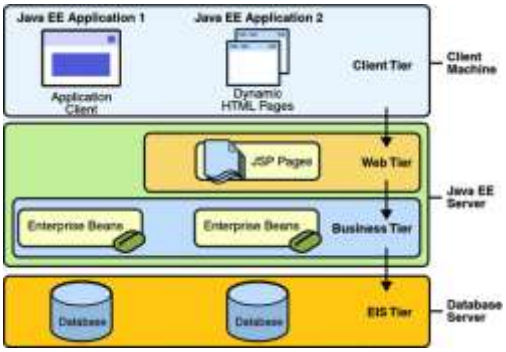
Gestão de Centros de Dados

Data: 18/Novembro/2011 Duração: 1h45m

Nota: Algumas questões podem não estar completamente especificadas. Isto é intencional, e deverá assumir hipóteses razoáveis para depois formular as suas respostas.

Q-1 Num centro de dados pretendemos instalar uma arquitectura de suporte a uma aplicação distribuída em 3 níveis:



(iii)

(ii)

(i)

i) uma base de dados, suportada por um SGBD relacional (e.g., MS SQLserver);
ii) um servidor aplicacional suportado por uma tecnologia distribuída (e.g., Java Beans);
iii) uma camada de apresentação suportada numa tecnologia integrada com Web servers.

Numa primeira fase, as três camadas estão implementadas sobre um único servidor: um sistema com 4 sockets para CPUs, totalmente preenchidos, com 32 GB de RAM (expansíveis até 128GB), e com 4 discos internos SATA de 500 GB – um usado para o SO (**`DiscoSO`**), outro para o software "aplicacional" (SGBD, Web server, Application server, etc. - **`DiscoSfw`**), outro usado para os conteúdos estáticos (páginas estáticas, imagens, filmes promocionais, etc. - **`Disco1`**) e outro para a BD (**`DiscoBD`**). Contudo, cedo nos apercebemos que a solução não é aceitável tanto do ponto de vista de fiabilidade, como de desempenho.

1) Caracterização da situação ao nível do subsistema de discos. **`DiscoSO`**: o padrão de acessos corresponde a uns 50% de leituras (praticamente todas acontecendo quando o sistema arranca) e 50% de escritas – destas últimas, 95% são referentes a **`paging`** e acontecem sempre que o número de utilizadores concorrentes excede a centena. **`DiscoSfw`**: o padrão de acessos corresponde a uns 99% de leituras (praticamente todas acontecendo quando o sistema arranca) **`Disco1`**: o padrão de acessos corresponde a uns 99% de leituras, que são consequência das visitas dos utilizadores às páginas. **`DiscoBD`**: o padrão de acessos corresponde a uns 80% de leituras e 20% de escritas resultando da actividade dos utilizadores (compras on-line).

2) Caracterização da situação ao nível do subsistema de memória. Nota-se um esgotamento da RAM, com intensa actividade de *paging*, quando número de utilizadores concorrentes excede a centena.

3) Caracterização da situação ao nível dos CPUs. Nota-se crescendo de consumo com o aumento do número de utilizadores, chegando a atingir 80% quando esse número atinge a centena; contudo, depois, a taxa de utilização do CPU decresce!

Deve apresentar as soluções que propõe em duas alíneas separadas:

a) [4 val.] **Mantendo** o sistema actual, tentando resolver os problemas de fiabilidade e desempenho de cada um dos subsistemas dentro do possível ☺

b) [4 val.] Tem "carta branca" para efectuar as alterações necessárias para promover o aumento de fiabilidade / alta disponibilidade e desempenho.

Q-2 [6 val.] Considere estes dois "componentes" de uma sistema computacional: o servidor (CPU, memória, periféricos vários) e os periféricos de armazenamento ("discos"), sendo que estes podem ser instalados "dentro" ou "fora" do servidor.

a) Discuta (i) as diferentes opções tecnológicas que conhece para os "discos" e suas vantagens e inconvenientes - com referência ao contexto/exemplo em que se verifica a referida (in)conveniência; (ii) as diferentes arquitecturas/tecnologias de interligação dos discos/ /subsistemas de armazenamento aos servidores; e (iii) os protocolos (pilhas protocolares, se for o caso) utilizados na interligação.

b) Defina, no contexto de um subsistema de armazenamento (do tipo *disk array*) (i) o que são as funcionalidades designadas respectivamente *thin provisioning* e *snapshots*; (ii) descreva um cenário em que seja vantajoso que o subsistema de armazenamento ofereça tais funcionalidades.

Q-3 [6 val.] Um grande centro de dados (como o visitado em Évora) dispõe de uma série de infra-estruturas que albergam os equipamentos e mantêm um ambiente adequado ao seu funcionamento. Nomeie cada uma ("tudo menos os servidores e equipamentos de armazenamento e de rede") e descreva o seu papel. Indique, se possível, quais as grandezas físicas que devem ser monitorizadas e, para cada uma, que intervalos acha que constituem "um ambiente adequado".