DI/FCT/UNL
Mestrado Integrado em Engenharia Informática

Confiabilidade de Sistemas Distribuídos
2º Semestre, 2015/2016

Prova de Exame (1/Julho/2016)

PARTE I (Sem Consulta, 1h15 m)

Questão 1

Considerando o problema do consenso distribuído, enuncie rigorosamente em que consiste o princípio conhecido pelo resultado de impossibilidade FLP (*Fisher, Lynch* and *Paterson*).

Questão 2

Num modelo de replicação baseado em máquina de estado (*SMR Replication Model*) cada servidor (réplica) do sistema distribuído gere uma cópia (estado local do serviço), exportando numa API um conjunto de operações *Oi* (tipicamente de leitura e escrita de valores em variáveis de estado). Cada operação: *(i)* fornece um resultado de saída com base em argumentos de entrada e *(ii)* a sua execução opera uma mudança ou transição do estado interno da réplica (*state-transition*) acordada com as outras réplicas.

a) Em que consiste a noção de determinismo de uma operação *Oi* ?
b) Quando se considera que o sistema replicado (na sua globalidade) é determinista ?
c) Considere que o sistema é determinístico e todas as réplicas do sistema funcionam corretamente. Para obtenção da noção de replicação de máquina de estados que garantias de ordenação das operações devem existir em qualquer réplica quando um conjunto arbitrário de operações é executado no sistema.
d) Considere que tem ao seu dispor um algoritmo de consenso distribuído. Explique brevemente como poderia implementar um sistema de replicação de máquina de estado com base no algoritmo de consenso.

Questão 3

Considere a definição de um protocolo do tipo consenso determinístico, enunciada da seguinte forma:

- Cada réplica que executa corretamente o protocolo de consenso tem uma variável Vi (*proposal*) que pretende propor como novo valor global do sistema;
- Cada réplica tem uma variável de decisão (Di) que inicialmente tem o valor "*not-decided*".
- O objetivo é o sistema chegar a um estado em que todas as réplicas acordem e decidam no mesmo valor para *Di*, face ás propostas das diferentes réplicas.
- A decisão é tomada salvaguardando as seguintes quatro propriedades garantidas pelas réplicas que operam corretamente:

  (P1) *validity*, (P2) *agreement*, (P3) *termination*, (P4) *integrity*.

a) Defina cada uma destas propriedades.
b) Dadas as anteriores propriedades e a forma como as enunciou, qual a diferença entre um protocolo de consenso determinístico e um protocolo do tipo consenso probabilístico.

Questão 4

O protocolo PAXOS na sua versão base permite implementar um consenso distribuído garantindo as propriedades de *liveness* e *safety*, sob certas condições.

a) Qual das seguintes condições não é verdadeira como pressuposto para o correto funcionamento do protocolo tendo em vista garantir as propriedades enunciadas? Justifique a sua resposta, argumentando adequadamente a sua justificação.

   A1: O sistema distribuído é assíncrono;

A2: Todas as operações realizadas corretamente pelas réplicas são determinísticas
A3: As réplicas podem exibir um falha do tipo *failt-stop* provocada por um falha acidental ou por um ataque ativo que provoque essa falha;
A4: As réplicas não podem exibir falhas arbitrárias que ocorram acidentalmente ou que resultem de ataques bizantinos, nem podem exibir um comportamento de falha ou de incorreção resultante de conluio (*collusion*).

b) Os sistemas que usam o PAXOS tipicamente executam o papel de *proposer*, *acceptor* e *learner* (papéis desempenhados em diferentes *threads*). Explique brevemente qual o objetivo de cada um destes componentes do PAXOS.

Questão 5

O sistema Depsky implementa um sistema de armazenamento confiável, com base numa arquitetura multi-cloud, utilizando mecanismos de replicação, fragmentação bem como mecanismos criptográficos específicos, incluindo o mecanismo designado por "*secret-sharing*". De acordo com o seu estudo para que é usado e qual o interesse ou vantagem da utilização do mecanismo criptográfico *secret-sharing* na arquitetura e operação do sistema? Justifique, contextualizando a utilização desse mecanismo nas operações de escrita e leitura tal como são suportadas no sistema.

Questão 6

Considere o suporte de TPMs (*Trusted Platform Modules*) em soluções HW que implementam os serviços e funções normalizadas para TPM. Os serviços básicos suportados e disponibilizados na normalização desse tipo de soluções são os seguintes: *Authenticated Boot Service*, *Certification ou Attestation Service* e *Encryption Service*. Refira de forma completa em que consiste o suporte de cada um desses serviços e como asseguram as respetivas garantias de segurança

Questão 7

Considere o contexto do sistema COCA, apresentado como caso representativo de uma proposta que utiliza técnicas de replicação combinadas com recuperação pró-ativa. Este sistema utiliza ainda um mecanismo criptográfico conhecido por "assinaturas de limiar" (ou *threshold signatures*). Qual o contexto de uso deste mecanismo no objetivo do sistema e que vantagens encontra na utilização desse mecanismo de segurança, tendo em mente a arquitetura do sistema? Justifique.

Questão 8

Considere o contexto do estudo dos sistemas híbridos e distribuídos de detecção de intrusões. Estes sistemas combinam diferentes subsistemas de detecção de intrusões, incluindo NIDS, HIDS e sistemas do tipo *Honeypot*.

a) Qual o papel dos sistemas *Honeypot* nestas arquiteturas tendo em conta o tipo de detecção de intrusões já suportada nas soluções NIDS e HIDS ?
b) Na caracterização das tipologias de sistemas *Honeypo*t consideram-se duas classes principais que foram identificadas como "*Low-Interaction Honeypots*" e "*High-Interation Honeypots*". Qual a diferença nesta classificação?

PARTE II (Com Consulta): até 1h15 m

Questão 1

Considere o protocolo ABD e a possibilidade do mesmo tolerar Falhas Bizantinas. Neste protocolo são utilizados *nonces* nas mensagens, sendo os mesmos importante no processamento do protocolo. Suponha que não se usam *nonces* e em vez das mensagens incluírem um *nonce,* passam a incluir um código de integridade de cada mensagem (por exemplo do tipo HMAC), com base numa chave (HMAC) partilhada em segurança por todos os participantes (e que foi supostamente obtida com base num protocolo seguro de autenticação e de distribuição de chaves que distribuiu essa chave a todos os participantes que vão depois processar o protocolo).

A utilização do HMAC (independentemente de exigir um processamento de síntese no envio e recepção/verificação das mensagens) permite substituir o uso dos *nonces* ? Justifique.

Questão 2

Considere o seguinte algoritmo:

```
preference ← input (0 ou 1)
round ← 1
while true do
   send (round, preference) to all processes
   wait to receive n – f (round, *) messages
   if received more than n / 2 (round, v) messages then
     output ← v
     preference ← v
   else if received less than n / 2 – f (round, preference) messages then
     preference ← coin_flip
   end
end
```

Explique porque é que é seguro mudar de voto nas condições indicadas no protocolo.

Sugestão: considere o que pode acontecer nos outros processos quando a condição que leva à mudança de *preference* num dado processo se verifica.

Questão 3

a) Qual a diferença entre a utilização de IPSec em modo em modo túnel e em modo transporte.
b) Discuta as diferenças bem como vantagens e inconvenientes de proteger as comunicações num sistema distribuído com IPSec (usado em modo transporte) em vez de usar SSL (ou TLS). Se quiser complementar a sua discussão com um exemplo, utilize como exemplo o sistema que implementou na sua avaliação prática, tendo em vista que pretende proteger as comunicações entre o cliente e o serviço SIFTBox e as comunicações entre diferentes réplicas SIFTBox.

   Sugestão: considere na sua discussão pelo menos os seguintes critérios (para além de outros que considere relevantes): propriedades de segurança garantidas, eficiência, escalabilidade para o serviço funcionar na Internet e potenciais limitações operacionais no deployment real da solução para a Internet.

Questão 4 (parte II)

Considere o estudo do sistema Depsky

a) No sistema Depsky a utilização de *erasure codes* é particularmente importante uma vez que o suporte do repositório de dados é baseado em *key-value-stores* suportados em diferentes Clouds (diferentes *cloud-providers*). Que principais vantagens argumentaria para o uso dessa solução face a um estratégia de replicação nas diversas *clouds* utilizadas? Justifique.
b) Do seu estudo do sistema Depsky, deve ter reparado que não é possível suportar operações concorrentes de escrita em algum dos protocolos variantes (Depsky-A ou Depsky-CA)? Que suporte ou solução é preconizada pelos autores para melhorar este aspeto? Da sua visão crítica, que overheads podem verificar-se pela adoção dessas soluções?

Questão 5

Considere o seu estudo do sistema PBFT utilizando o mecanismo de Pro-Active Recovery proposto pelos autores. Indique alguma possibilidade ou uma tipologia de um ataque por intrusão que consiga subverter o sistema, no pressuposto de que: (1) as condições de estabelecimento do consenso e suporte para BFT (Byzantine Fault-Tolerance) são salvaguardadas, (2) o protocolo PBFT está bem implementado e (3) o mecanismo de Pro-Active Recovery também foi corretamente implementado.

Questão 6

Considere o estudo do sistema CryptDB.

a) "Um dos problemas em aberto que pode ser levantado face à solução é que à medida que se vão suportando *queries*, a confidencialidade dos dados no servidor vai ficando mais vulnerável". Concorda ou não com esta afirmação? Justifique.
b) Qual a vantagem que encontra na utilização da abordagem do tipo Onion-Encryption por coluna? Justifique a sua resposta.