

Desempenho em Sistemas Distribuídos

Seminário 1 de SDI

Miguel A. Nunes
miguel.nunes@edu.udesc.br

22/08/2022

O que significa desempenho no contexto de sistemas distribuídos?

- Quão rapidamente o sistema consegue realizar uma tarefa (tempo de execução);

O que significa desempenho no contexto de sistemas distribuídos?

- Quão rapidamente o sistema consegue realizar uma tarefa (tempo de execução);
- Quão rapidamente os componentes do sistema se comunicam (latência);

O que significa desempenho no contexto de sistemas distribuídos?

- Quão rapidamente o sistema consegue realizar uma tarefa (tempo de execução);
- Quão rapidamente os componentes do sistema se comunicam (latência);
- Quão rapidamente o sistema consegue responder a uma requisição (tempo de resposta);

O que significa desempenho no contexto de sistemas distribuídos?

- Quão rapidamente o sistema consegue realizar uma tarefa (tempo de execução);
- Quão rapidamente os componentes do sistema se comunicam (latência);
- Quão rapidamente o sistema consegue responder a uma requisição (tempo de resposta);
- Qual a fração do tempo que o sistema fica disponível (disponibilidade);

O que significa desempenho no contexto de sistemas distribuídos?

- Quão rapidamente o sistema consegue realizar uma tarefa (tempo de execução);
- Quão rapidamente os componentes do sistema se comunicam (latência);
- Quão rapidamente o sistema consegue responder a uma requisição (tempo de resposta);
- Qual a fração do tempo que o sistema fica disponível (disponibilidade);
- Quão resistente à falhas/erros é o sistema (resiliência).

- Quanto tempo leva para o sistema computar algo;

- Quanto tempo leva para o sistema computar algo;
- Em específico, quanto tempo de CPU é gasto para realizar alguma computação, desconsiderando tempo de comunicação;

- Quanto tempo leva para o sistema computar algo;
- Em específico, quanto tempo de CPU é gasto para realizar alguma computação, desconsiderando tempo de comunicação;
- Balanceamento de Carga: Todos os componentes do sistema devem realizar a mesma quantidade de trabalho;

- Quanto tempo leva para o sistema computar algo;
- Em específico, quanto tempo de CPU é gasto para realizar alguma computação, desconsiderando tempo de comunicação;
- Balanceamento de Carga: Todos os componentes do sistema devem realizar a mesma quantidade de trabalho;
- Concorrência: O máximo possível de trabalho deve ser executado em paralelo.

- Quanto tempo leva para componentes do sistema se comunicarem;



- Quanto tempo leva para componentes do sistema se comunicarem;
- Como os componentes do sistema estão conectados está intimamente relacionado à latência;

- Quanto tempo leva para componentes do sistema se comunicarem;
- Como os componentes do sistema estão conectados está intimamente relacionado à latência;
- Topologias de rede que não são escaláveis/tornam elementos muito distantes são mais lentas;

- Quanto tempo leva para componentes do sistema se comunicarem;
- Como os componentes do sistema estão conectados está intimamente relacionado à latência;
- Topologias de rede que não são escaláveis/tornam elementos muito distantes são mais lentas;
- Problemas de rede: Sobrecarga em enlaces, topologia da rede, protocolos que não balanceiam bem a carga entre si.

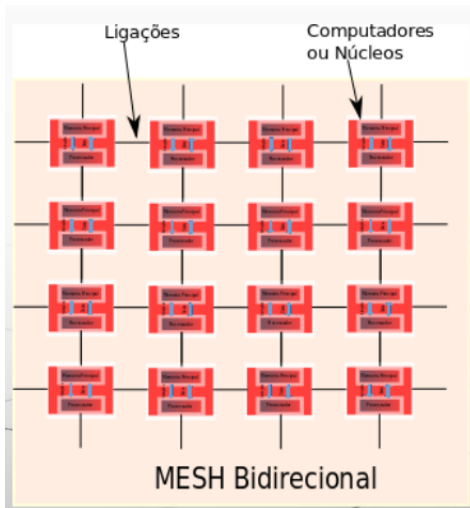


Figura: Exemplo de topologia de rede – componentes ligados diretamente

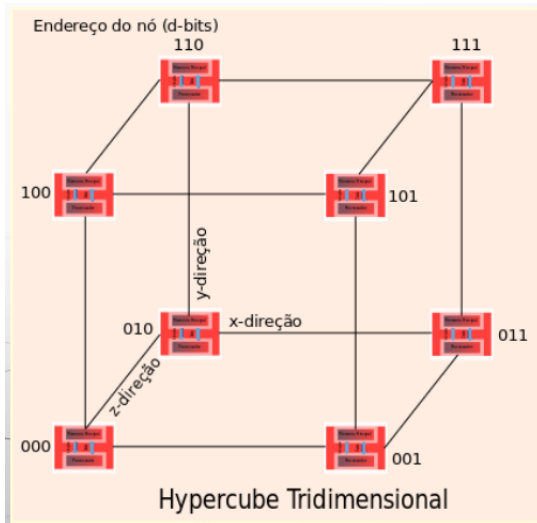


Figura: Exemplo de topologia de rede – conexões formam um cubo n-dimensional

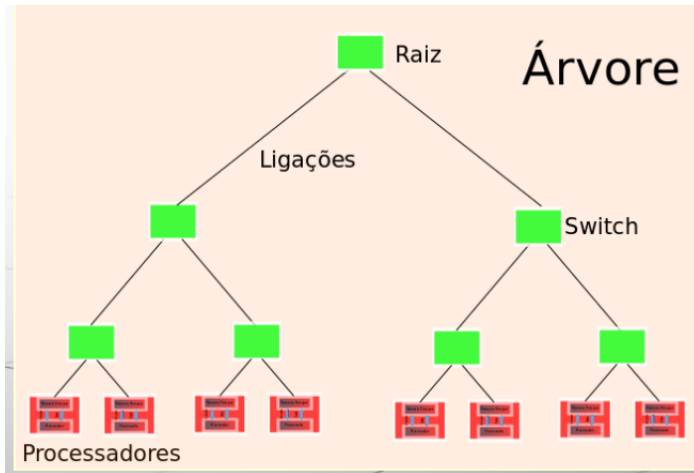


Figura: Exemplo de topologia de rede – componentes ligados em árvore

- Quanto tempo leva para o sistema responder a uma requisição;

- Quanto tempo leva para o sistema responder a uma requisição;
- Tempo de Execução + Latência + Tempo para a informação chegar até a origem da requisição;

- Quanto tempo leva para o sistema responder a uma requisição;
- Tempo de Execução + Latência + Tempo para a informação chegar até a origem da requisição;
- Largura de Banda: Quantos dados podem ser transmitidos em um dado instante de tempo

- Quanto tempo leva para o sistema responder a uma requisição;
- Tempo de Execução + Latência + Tempo para a informação chegar até a origem da requisição;
- Largura de Banda: Quantos dados podem ser transmitidos em um dado instante de tempo
- *Throughput*: Quantidade de tarefas realizadas em um dado tempo.

- A fração de um dado intervalo de tempo onde o sistema está operando corretamente;
- “*O SIGA tem disponibilidade de 40%*” – o SIGA passa 40% de todo o tempo operando corretamente

- A fração de um dado intervalo de tempo onde o sistema está operando corretamente;
- “*O SIGA tem disponibilidade de 40%*” – o SIGA passa 40% de todo o tempo operando corretamente
- Uma fração do sistema que está continuamente operando;
- “*Os servidores da Amazon tem disponibilidade de 95%*” – 95% de todos os servidores da Amazon estão operando corretamente a todo instante.

- A fração de um dado intervalo de tempo onde o sistema está operando corretamente;
- “*O SIGA tem disponibilidade de 40%*” – o SIGA passa 40% de todo o tempo operando corretamente
- Uma fração do sistema que está continuamente operando;
- “*Os servidores da Amazon tem disponibilidade de 95%*” – 95% de todos os servidores da Amazon estão operando corretamente a todo instante.
- Um fator “alto nível”, vai além dos próprios componentes do sistema, manutenção do sistema, manutenção da rede interna e externa.

- Se sistema consegue lidar com falhas/erros;
- Se sim, quais tipos de falhas;



- Se sistema consegue lidar com falhas/erros;
- Se sim, quais tipos de falhas;
- Outro fator de alto nível, resiliência a certos tipos de falhas pode ir além dos componentes do sistema, bugs no hardware ou OS, ataques externos

Early Christmas morning (Pacific Standard Time), the Steam Store was the target of a DoS attack which prevented the serving of store pages to users. [...] During the second wave of this attack, a second caching configuration was deployed that incorrectly cached web traffic for authenticated users. This configuration error resulted in some users seeing Steam Store responses which were generated for other users.

Como medir desempenho?

- Modelos teóricos ou “em escala” do sistema;

Como medir desempenho?

- Modelos teóricos ou “em escala” do sistema;
- Simular cargas de trabalho no sistema em execução;

Como medir desempenho?

- Modelos teóricos ou “em escala” do sistema;
- Simular cargas de trabalho no sistema em execução;
- Medir o sistema em execução.