

EP2

MAC0352 - Redes e Sistemas Distribuídos

Breno Helfstein Moura
Matheus Barcellos de Castro Cunha

IME-USP, outubro 2019

Outline

Arquitetura do Sistema

- Como é feita a divisão de trabalho

- Tipos de nós

- Eleição

Resultados dos testes

- Tempo de execução

- Uso de memória

- Carga da rede

Outline

Arquitetura do Sistema

- Como é feita a divisão de trabalho

- Tipos de nós

- Eleição

Resultados dos testes

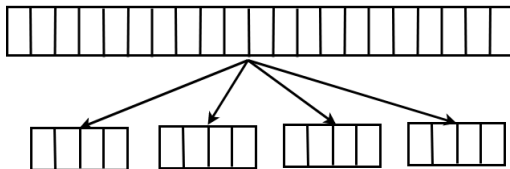
- Tempo de execução

- Uso de memória

- Carga da rede

Como é feita a divisão de trabalho

- ▶ O arquivo gigante é dividido em "*chunks*", o tamanho do "*chunk*" é 1% do tamanho da lista a ser ordenada. O tamanho máximo de um "*chunk*" é 10^6 e o tamanho mínimo é 10^1
- ▶ O arquivo é dividido entre as máquinas da rede, cada máquina usa o sort default do Go, e ao final é feito um intercala entre todos os "*chunks*" em $O(N \log K)$ onde N é o número de elementos da lista e K é o número de chunks.



Outline

Arquitetura do Sistema

Como é feita a divisão de trabalho

Tipos de nós

Eleição

Resultados dos testes

Tempo de execução

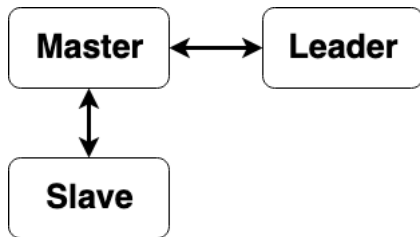
Uso de memória

Carga da rede

Tipos de nós

- ▶ O sistema possui 3 tipos de nós que podem entrar na rede, Mestre, Líder e Escravo
- ▶ O nó mestre possui o arquivo a ser ordenado. Ele recebe pedidos de trabalho do nó Líder para ordenar um pacote e de forma síncrona, pede para o nó indicado executar o trabalho. Ao fim retorna ao líder que o trabalho foi pedido com sucesso.
- ▶ O nó escravo ajuda o líder a ordenar o arquivo. Ele recebe pacotes a serem ordenados e devolve os pacotes ordenados. Um escravo pode ser promovido a líder.
- ▶ O nó líder divide os trabalhos. Ele pede para o mestre enviar trabalho para algum outro nó. Ele usa um "Round Robin Aleatorizado" em que ele pega uma permutação aleatória dos nós para pedir trabalhos.

Exemplo com 3 nós:



Detalhes

- ▶ O protocolo de rede foi feito em cima do TCP
- ▶ O Timeout para uma resposta no sistema é configurado como 45 segundos
- ▶ O HeartBeat é feito a cada 1 minuto
- ▶ Cada nó só pode ter 2 pedidos de ordenação por vez
- ▶ Se abre um novo socket para cada pedido de ordenação
- ▶ Ao fim da ordenação elegemos FINALSORT como líder, para focar os recursos do mestre no merge final

Outline

Arquitetura do Sistema

Como é feita a divisão de trabalho

Tipos de nós

Eleição

Resultados dos testes

Tempo de execução

Uso de memória

Carga da rede

Eleição

- ▶ A eleição é feita a cada 3 minutos de forma automática, ou é feita de forma forçada caso o líder morra
- ▶ O nó mestre começa pedindo votos aos nós da rede, e cada nó vota de forma aleatória dado seu conhecimento de que nós estão na rede.
- ▶ Ao final da eleição um novo líder é eleito e o nó mestre passa as informações necessárias para ele começar a divisão de trabalhos.
- ▶ Durante a eleição os nós colocam NULL como líder para pararem de aceitar a divisão de tarefas.

Outline

Arquitetura do Sistema

Como é feita a divisão de trabalho

Tipos de nós

Eleição

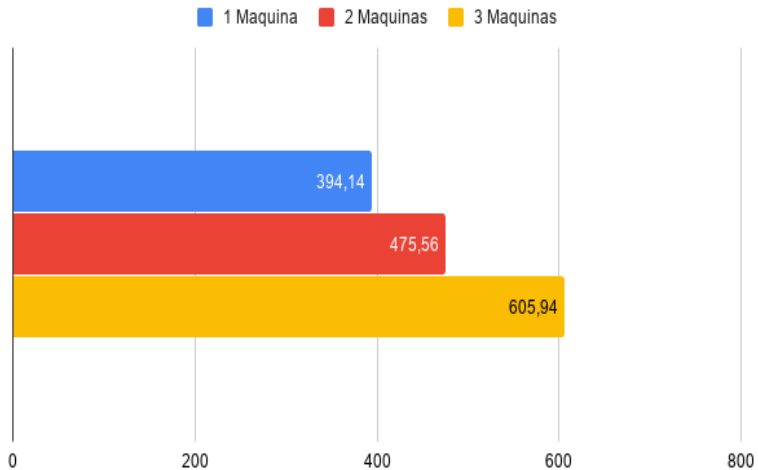
Resultados dos testes

Tempo de execução

Uso de memória

Carga da rede

Análise de Desempenho (Tempo (segundos) x Máquinas)



Tempo de execução

Devido a necessidade de fazer envios a diferentes máquinas utilizando o protocolo TCP, o tempo de ordenação do arquivo aumenta conforme o número de computadores conectados ao sistema aumenta. Isso pode parecer contra-intuitivo dado que com mais máquinas teríamos mais poder de processamento, contudo dada a implementação da coordenação com líderes e eleições, mais máquinas acabam levando a um tempo de resolução maior.

Outline

Arquitetura do Sistema

Como é feita a divisão de trabalho

Tipos de nós

Eleição

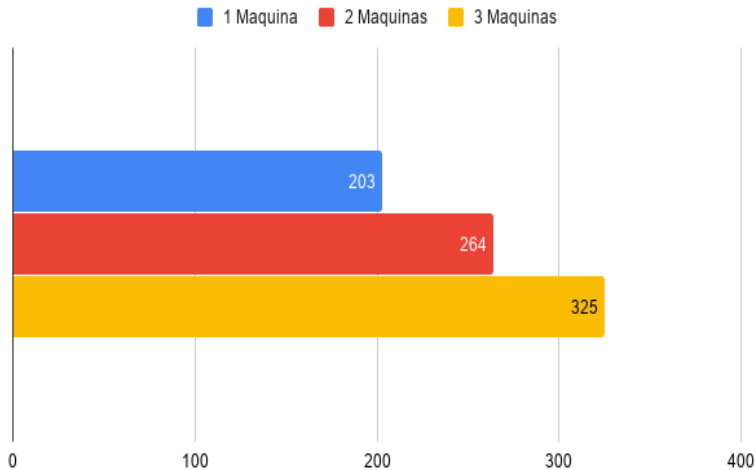
Resultados dos testes

Tempo de execução

Uso de memória

Carga da rede

Análise de Desempenho (Uso da CPU (%) x Máquinas)



Uso de memória

Similar a ideia do aumento do tempo de execução, há um custo de processamento para enviar pacotes para outras máquinas, e maior o número de máquinas conectadas mais processamento será gasto com operações de envio e recebimento de dados.

Outline

Arquitetura do Sistema

Como é feita a divisão de trabalho

Tipos de nós

Eleição

Resultados dos testes

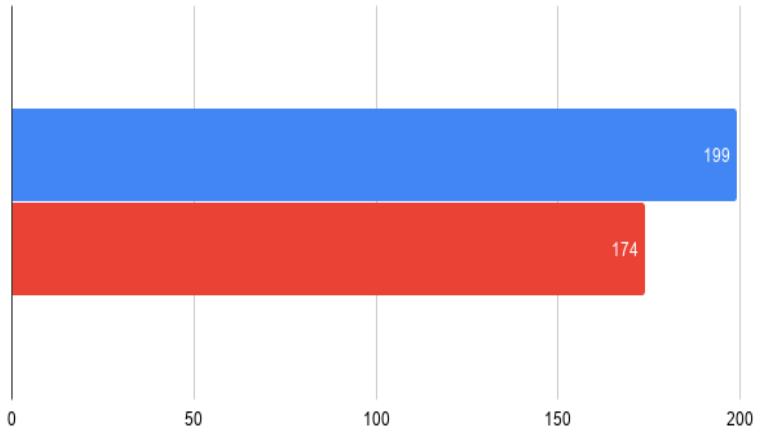
Tempo de execução

Uso de memória

Carga da rede

Análise de Desempenho (KB/Máquinas x Máquinas)

■ 2 Maquinas ■ 3 Maquinas



Carga da rede

- ▶ Tendo uma divisão de trabalho baseada no número de computadores conectados ao sistema, a diminuição da média de Kilobytes enviados para cada computador também será menor.
- ▶ Para medir o tráfego da rede sem interferência de outras aplicações, foi utilizado o Wireshark para filtrar a análise de tráfego de IPs específicos para IPs específicos.