Marcos Paulo Tomás Ferreira – NUSP 13747950

O que são Distributed Hash Tables (DHTs)?
As DHTs são tabelas usadas na busca de dados em redes P2P. As DHTs associam um texto de busca a um valor, que está armazenado em um determinado *peer* da rede. As buscas em DHTs somente suportam *matches* exatos. DHTs foram inspiradas por outros sistemas P2P, como Freenet, Gnutella e Napster. As DHTs surgem para combinar a eficiência do Napster com a descentralidade do Freenet e Gnutella.

Como as DHTs funcionam?
As DHTs combinam Autonomia e descentralização, tolerância a falhas e escalabilidade. Cada par na rede apenas se comunica com outros poucos pares, mantendo o trabalho a ser feito com um limite superior da ordem de O(log n). A estrutura de uma DHT pode ser decomposta em conjuntos de strings, chamados *keyspace*. Um *keyspace* pode ter diversos tamanhos; suponhamos que o nosso *keyspace* seja um conjunto de strings, com 20 Bytes para cada string. Cada *keyspace* tem sua propriedade partilhada entre os nós participantes. E, então, uma *overlay network* (rede sobreposta) conecta todos os nós.

O processo de salvamento de um arquivo é o seguinte:

1. O nome do arquivo passa pela função de hash SHA-1, e, então, é gerado uma chave *k*;
2. Uma mensagem *put(k, data)*, onde *data* são os dados dos arquivos, é enviada para qualquer nó participante na DHT; a mensagem é passada de nó em nó através da *overlay network* até atingir um único nó que seja responsável pela chave *k*, como foi identificado no particionamento do *keyspace;*
3. Ao chegar no nó responsável, este armazena a chave *k* e os dados *data*.

O processo de recuperação de um arquivo é semelhante. Vale ressaltar que qualquer outro componente da rede pode recuperar esse arquivo, veja como:

1. O nome do arquivo novamente passa pela função SHA-1, para recuperar a chave *k*;
2. Uma mensagem de recuperação de arquivo, *get(k)*, é enviada a outro nó da rede; a mensagem é roteada até o nó que contém o arquivo;
3. Chegando no nó que armazena o arquivo com chave *k*, este responde à mensagem com os dados *data* do arquivo.

Como conhecimento adicional, SHA-1 é uma função de criptografia, que já foi quebrada no passado com ataques de colisão provados. O uso de SHA-1 já não é mais recomendado e está sendo substituídos por SHA-2 ou SHA-3. Outros algoritmos de criptografia, como a Cifra de Rinjdael (AES-256), são recomendados. Não tenho informações se SHA-1 foi substituído no caso das DHTs.

Como as DHTs ajudam na arquitetura P2P?
As DHTs, por usarem uma tabela de hash, tem tempo constante O(1) para cada nó, o que torna a busca de dados e informações rápidas. Além disso, são descentralizadas – diminuindo a chance de ataques em um ponto focal -, escaláveis e tolerantes a falhas.

Fonte: https://en.wikipedia.org/wiki/Distributed_hash_table