Estudo do uso de TLA+ para formalização de uma ferramenta de materialização de consultas com base em streaming[1]

Grasiela da Silva Ferreira[2], Cristiano Damiani Vasconcellos[3].

[1] Vinculado ao projeto "Inferência de Tipos e Efeitos para Análise Estática de Programas"
[2] Acadêmica do curso de Ciência da Computação – CCT – Bolsista PROIP
[3] Orientador, Departamento de Ciência da Computação – CCT – cristiano.vasconcellos@udesc.br

Entregar resultados de consultas em bases de dados é um problema de escala: quanto maior a base, mais lenta a consulta. As soluções mais comuns envolvem mecanismos de cache, onde geralmente são encontrados desafios com relação ao que deve ser mantido em cache e quando invalidar. A fim de alcançar escalabilidade é realizada uma distribuição desses processamentos em diversos computadores, caracterizando um Sistema Distribuído (SD), e por se tratar de um SD encontramos algumas dificuldades para contornar como, por exemplo, eventos chegando em ordens diferentes daquelas em que aconteceram, concorrência por recursos e distribuição de carga. Tais problemas de concorrência são intrinsecamente difíceis de resolver na sua totalidade porque o número de estados possíveis cresce rapidamente à medida que as alterações são executadas e a mente humana, sozinha, não consegue cobrir todas as possibilidades. A fim de garantir a consistência, dentre outras propriedades desejáveis, podem-se empregar métodos formais, que são técnicas matemáticas para especificar e raciocinar sobre propriedades de software, utilizando ferramentas para auxiliar na validação de tais propriedades. Uma vez que sistemas distribuídos estão sujeitos a muitas ordens de acontecimentos diferentes, navegar por todos os estados, seja mentalmente ou em testes automatizados, se torna inviável. Com uma especificação formal é possível validar propriedades e permitir otimizações mais agressivas, esses são apenas alguns exemplos dos benefícios no uso de certificações formais em sistemas computacionais, que podem ser cruciais para aplicações distribuídas ou críticas.

Este estudo tem como objetivo verificar a viabilidade e possíveis vantagens no uso de uma ferramenta de métodos formais (TLA+) na especificação de um sistema para materialização de consultas. TLA+ é uma linguagem de alto-nível para modelagem de programas e sistemas (em especial os sistemas concorrentes e sistemas distribuídos), suas especificações são escritas em uma linguagem formal que combina lógica temporal de ações e teoria de conjuntos de Zermelo-Fraenke (ZFC). TLA+ foi proposta por Lambort, 1994, baseada em transições de estados, sendo os estados permitidos modelados por uma ou mais ações. Modelar um sistema em TLA+ é definir uma sequência de estados que o sistema pode assumir. Como exemplo é apresentado na Figura 1 uma especificação que modela as possíveis transições para um semáforo. Nesse exemplo, a variável *color* representa a cor atual do semáforo, que pode assumir um de três estados: *green*, *yellow* ou *red*; e a variável *color'* representa a cor do próximo estado. *Init* define os possíveis estados iniciais do sistema e *Next* define as 3 ações permitidas. O semáforo é então definido por meio do estado inicial e das possíveis ações:

$$Spec == Init \wedge Next$$

Na verificação da especificação, que pode ser feita por meio da ferramenta TLC (Lamport. 2002), o sistema é tratado como uma grande máquina de estados, o que permite verificar se

determinadas propriedades são válidas em todos os estados possíveis do sistema. Por exemplo, em um sistema envolvendo vários semáforos poderíamos definir como propriedade que não é possível dois semáforos estarem simultaneamente com a cor verde (Mafra, 2019):

$$no\_collision == Cardinality(\{s \in SEMAPHORES : colors[s] = \text{“green”}\}) \leq 1$$

```
VARIABLE color

TurnRed  color == “yellow” ∧ color′ = “red”
TurnYellow == color = “green” ∧ color′ = “yellow”
TurnGreen == color = “red” ∧ color′ = “green”

Init == color ∈ {“red”, “yellow”, “green”}

Next == TurnRed ∨ TurnYellow ∨ TurnGreen
```

Figura 1: Semáforo em TLA+

Nesta etapa inicial o trabalho teve como objetivo o estudo de TLA+ e o uso do model checker TLC, não sobrando muito tempo para o estudo do problema que será modelado. O projeto foi proposto para tratar um problema real implementado por uma empresa da região, Uma das dificuldades encontradas foi a venda dessa empresa que teve como consequência a perda do interesse no projeto por parte da empresa, nos deixando sem uma referência para o problema real.

Referências

Gabriela Moreira Mafra. Tradução automática de especificação formal modelada em TLA+ para linguagem de programação. 2019. Trabalhos de Conclusão de Curso (Graduação). Universidade do Estado de Santa Catarina, Joinville, 2019.

Leslie Lamport. 1994. The Temporal Logic of Actions. ACM Trans. Program. Lang. Syst. 16, 3 (1994), 872–923.

Leslie Lamport. 2002. Specifying Systems: The TLA+ Language and Tools for Hardware and Software Engineers. Addison-Wesley, 2002. Disponível em: https://lamport.azurewebsites.net/tla/book.html>.

Palavras-chave: TLA+, Especificação formal, Streaming Database, Sistemas Distribuídos