

L^AT_EX Author Guidelines for 8.5 × 11-Inch Proceedings Manuscripts

Bernardo Simões
Afonso Oliveira
Rui Francisco
Instituto Superior Técnico
Plataformas para Aplicações Distribuídas da Internet

Abstract

No projecto de PADI (Plataformas para Aplicações Distribuídas na Internet), foi-nos pedido para projectar e implementar uma PADITable, um sistema distribuído que gere em memória volátil conjuntos de chave-valor que suportam as duas operações fundamentais get e put de forma atómica. Este documento contém a descrição dos protocolos implementados para a nossa solução deste sistema, onde discutimos formas de processamento de transacções atómicas, topologia da rede e gestão dos servidores, clientes e uma directoria central. Seguidas de uma avaliação das mesmas referindo vantagens e desvantagens de cada implementação e são apresentados resultados.

1. Introdução

Este artigo permite descrever a solução para a implementação da PADITable, um sistema distribuído para gerir o armazenamento volátil de conjuntos chave-valor. Este sistema é constituído por 4 tipos de nós: uma directoria central (referida em frente como o nó CD (*Central Directory*)), um conjunto de servidores e um conjunto de clientes que são controlados por um *puppet master*. Os conjuntos chave-valor estão armazenados nos servidores que são operados pelos clientes usando operações de *put* e *get*. A directoria central tem a informação dos clientes e servidores ligados à rede e o *puppet master* é responsável por controlar os clientes a fim de testar e fazer *debug* do sistema.

Como todos os sistemas distribuídos é necessário saber onde colocar determinada informação. Como tal é necessário um algoritmo para dividir diferentes chaves pelos servidores existentes de maneira a no futuro se saber melhor onde se encontra cada chave. A este problema segue-se o problema de que os servidores do sistema podem iniciar-se em alturas diferentes, o que faz com que a distribuição anterior de chaves fique desactualizada. Será então necessário mover a localização de algumas chaves a

fim de tornar o sistema mais e melhor distribuído.

Quando um cliente precisar de executar operações de *put* e *get* irá necessitar de saber a localização dos servidores com as chaves a que quer aceder ou então será necessário um sistema para reencaminhar os pedidos para o servidor certo. A solução para este problema deverá usar o menor número de comunicações possível a fim de ser uma solução otimizada.

Uma vez que o cliente saiba a localização dos servidores que necessita aceder, ele terá de assegurar que uma transacção completa se execute de maneira sequencial de maneira a evitar estados inconsistentes.

A fim de se aumentar a disponibilidade do sistema e aumentar a capacidade de acesso a uma chave o sistema irá ser replicado. Onde e como replicar a informação é um factor que deve ser tomado em conta e que terá impacto no desempenho do sistema. Esta replicação deverá tornar o sistema acessível em caso de falha de qualquer servidor. A informação replicada deverá estar sempre consistente de maneira a evitar que sejam lidos valores desactualizados.

O sistema deverá estar preparado para que em cada chave sejam guardados vários valores, cada um associado a uma marca temporal. O sistema guarda assim um historial de valores antigos. Esta situação pode ser usada para diminuir o número de operações apenas de leitura que falham.

Na secção ?? irão ser identificadas e descritas as soluções para os requisitos do sistema e os problemas que estes levantam. De seguida na secção ?? irão ser descritas as vantagens das soluções optadas. Nesta secção serão também apresentados resultados de *benchmarks* ao sistema implementado, descrevendo o impacto das decisões tomadas nos valores obtidos.

2. Solução

Please read the following carefully.

2.1. Algoritmos de Colocação de Dados

2.2. Encaminhamento

2.3. Protocolo de Transacções Consistencia das Réplicas

2.4. Falhas do Servidor

2.5. Multi-Versão

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times that you have access to.

MAIN TITLE. Center the title 1-3/8 inches (3.49 cm) from the top edge of the first page. The title should be in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave two blank lines after the title.

AUTHOR NAME(s) and **AFFILIATION(s)** are to be centered beneath the title and printed in Times 12-point, non-boldface type. This information is to be followed by two blank lines.

The **ABSTRACT** and **MAIN TEXT** are to be in a two-column format.

MAIN TEXT. Type main text in 10-point Times, single-spaced. Do NOT use double-spacing. All paragraphs should be indented 1 pica (approx. 1/6 inch or 0.422 cm). Make sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs. Figure and table captions should be 10-point Helvetica boldface type as in

Figure 1. Example of caption.

Long captions should be set as in

Figure 2. Example of long caption requiring more than one line. It is not typed centered but aligned on both sides and indented with an additional margin on both sides of 1 pica.

Callouts should be 9-point Helvetica, non-boldface type. Initially capitalize only the first word of section titles and first-, second-, and third-order headings.

FIRST-ORDER HEADINGS. (For example, **1. Introduction**) should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one blank line after.

SECOND-ORDER HEADINGS. (For example, **1.1. Database elements**) should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after. If you require a third-order heading (we discourage it), use 10-point Times, boldface, initially capitalized, flush left, preceded by one blank line, followed by a period and your text on the same line.

2.6. Footnotes

Please use footnotes sparingly¹

2.7. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [?]. Where appropriate, include the name(s) of editors of referenced books.

2.8. Illustrations, graphs, and photographs

All graphics should be centered. Your artwork must be in place in the article (preferably printed as part of the text rather than pasted up). If you are using photographs and are able to have halftones made at a print shop, use a 100- or 110-line screen. If you must use plain photos, they must be pasted onto your manuscript. Use rubber cement to affix the images in place. Black and white, clear, glossy-finish photos are preferable to color. Supply the best quality photographs and illustrations possible. Penciled lines and very fine lines do not reproduce well. Remember, the quality of the book cannot be better than the originals provided. Do NOT use tape on your pages!

2.9. Color

The use of color on interior pages (that is, pages other than the cover) is prohibitively expensive. We publish interior pages in color only when it is specifically requested and budgeted for by the conference organizers. DO NOT SUBMIT COLOR IMAGES IN YOUR PAPERS UNLESS SPECIFICALLY INSTRUCTED TO DO SO.

2.10. Symbols

If your word processor or typewriter cannot produce Greek letters, mathematical symbols, or other graphical elements, please use pressure-sensitive (self-adhesive) rub-on symbols or letters (available in most stationery stores, art stores, or graphics shops).

¹Or, better still, try to avoid footnotes altogether. To help your readers, avoid using footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence).

2.11. Copyright forms

You must include your signed IEEE copyright release form when you submit your finished paper. We **MUST** have this form before your paper can be published in the proceedings.

2.12. Conclusions

Please direct any questions to the production editor in charge of these proceedings at the IEEE Computer Society Press: Phone (714) 821-8380, or Fax (714) 761-1784.