

高速缓存一致性和 量化分析



作者姓名： 高帅

邮箱： gshuai@mail.ustc.edu.cn

版本时间： 二〇二〇年七月十五日

写在最开始

缓存一致性 (Cache Coherent) 一直是芯片架构设计的重要问题，该文章配合仓库中的项目，来实现各种各样的 Cache 系统并比较 Cache 系统在特定场景下的性能。

本人毕业从事 AI 芯片架构设计岗位的工作，缓存一致性系统作为自己一个重要的技能栈，会广泛的学习各种发表或者开源的项目，加深自己对存储子系统的理解，服务于自己的设计工作。

目 录

第 1 章 基础知识	1
1.1 引言	1
1.2 存储器层次结构	1
1.3 缓存优化与方法	1
1.3.1 1	1
1.3.2 2	1
1.3.3 3	1
1.4 存储区优化与技术	1
1.4.1 二级节标题	1
1.5 脚注	1
第 2 章 浮动体	2
2.1 三线表	2
2.2 插图	2
2.3 算法环境	3
第 3 章 数学	4
3.1 数学符号和公式	4
3.2 量和单位	4
3.3 定理和证明	5
第 4 章 引用文献的标注	7
4.1 顺序编码制	7
4.1.1 角标数字标注法	7
4.1.2 数字标注法	7
4.2 著者-出版年制标注法	7
附录 A 补充材料	8
A.1 补充章节	8
致谢	9
在读期间发表的学术论文与取得的研究成果	10

符 号 说 明

a	The number of angels per unit area
N	The number of angels per needle point
A	The area of the needle point
σ	The total mass of angels per unit area
m	The mass of one angel
$\sum_{i=1}^n a_i$	The sum of a_i

第1章 基础知识

1.1 引言

1.2 存储器层次结构

1.3 缓存优化与方法

1.3.1 1

1.3.2 2

1.3.3 3

1.4 存储区优化与技术

1.4.1 二级节标题

1. 三级节标题

(1) 四级节标题

① 五级节标题

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.5 脚注

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. ^①

^①Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

第2章 浮 动 体

2.1 三线表

三线表是《撰写手册》推荐使用的格式，如表 2.1。

表 2.1 表号和表题在表的正上方

类型	描述
挂线表	挂线表也称系统表、组织表，用于表现系统结构
无线表	无线表一般用于设备配置单、技术参数列表等
卡线表	卡线表有完全表，不完全表和三线表三种

注：表注分两种，第一种是对全表的注释，用不加阿拉伯数字排在表的下边，前面加“注：”；第二种是和表内的某处文字或数字相呼应的注，在表里面用带圈的阿拉伯数字在右上角标出，然后在表下面用同样的圈码注出来

编制表格应简单明了，表达一致，明晰易懂，表文呼应、内容一致。排版时表格字号略小，或变换字体，尽量不分页，尽量不跨节。表格太大需要转页时，需要在续表上方注明“续表”，表头页应重复排出。

2.2 插图

有的同学可能听说“ \LaTeX 只能使用 eps 格式的图片”，甚至把 jpg 格式转为 eps。事实上，这种做法已经过时。而且每次编译时都要调用外部工具解析 eps，导致降低编译速度。所以我们推荐矢量图直接使用 pdf 格式，位图使用 jpeg 或 png 格式。



图 2.1 图号、图题置于图的下方

注：图注的内容不宜放到图题中。

关于图片的并排，推荐使用较新的 `subcaption` 宏包，不建议使用 `subfigure`

或 `subfig` 等宏包。

2.3 算法环境

模板中使用 `algorithm2e` 宏包实现算法环境。关于该宏包的具体用法，请阅读宏包的官方文档。

算法 2.1 算法示例 1

Data: this text

Result: how to write algorithm with L^AT_EX2_ε

```
1 initialization;
2 while not at end of this document do
3     read current;
4     if understand then
5         go to next section;
6         current section becomes this one;
7     else
8         go back to the beginning of current section;
9     end
10 end
```

注意，我们可以在论文中插入算法，但是插入大段的代码是愚蠢的。然而这并不妨碍有的同学选择这么做，对于这些同学，建议用 `listings` 宏包。

第3章 数 学

3.1 数学符号和公式

《撰写手册》要求数学符号要根据 GB/T 3102.11-1993 《物理科学和技术中使用的数学符号》^① 使用，这与 L^AT_EX 默认的英美国家的数学符号习惯有所差异。本模板基于 `unicode-math` 宏包配置数学符号，以遵循国标的规定：

1. 大写希腊字母默认为斜体，如 `\Delta`： Δ 。
2. 有限增量符号 Δ (U+2206) 应使用 `\increment` 命令。
3. 向量、矩阵和张量要求粗斜体，应使用 `\sympbf` 命令，如 `\sympbf{A}`、`\sympbf{\alpha}`。
4. 数学常数和特殊函数使用正体，如圆周率 π 、 Γ 函数。应使用 `unicode-math` 宏包提供的 `\symup` 命令转为正体，如 `\symup{\pi}`。
5. 微分符号 d 使用正体，本模板提供了 `\dif` 命令。

注意，`unicode-math` 宏包与 `amsfonts`, `amssymb`, `bm`, `mathrsfs`, `upgreek` 等宏包不兼容。本模板作了处理，用户可以直接使用 `\bm`, `\mathscr`, `\upGamma`。关于数学符号更多的用法，参见 `unicode-math` 宏包的使用说明和符号列表 `unimath-symbols`。

在编辑数学公式时，最好避免直接使用字体命令，而应该定义一些语义命令取代字体命令，这样输入更简单，也让 L^AT_EX 代码更有可读性，而且还方便根据需要统一修改改格式。参考示例文档中的 `math-commands.tex`

更多的例子：

$$e^{i\pi} + 1 = 0 \quad (3.1)$$

$$\frac{d^2 u}{dt^2} = \int f(x) dx \quad (3.2)$$

$$\arg \min_x f(x) \quad (3.3)$$

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (3.4)$$

3.2 量和单位

宏包 `siunitx` 提供了更好的数字和单位支持：

- 12 345.678 90
- $1 \pm 2i$

^①原 GB 3102.11-1993，根据 2017 年第 7 号公告和强制性标准整合精简结论，自 2017 年 3 月 23 日起，该标准转化为推荐性标准。

- 0.3×10^{45}
- $1.654 \times 2.34 \times 3.430$
- $\text{kg} \cdot \text{m} \cdot \text{s}^{-1}$
- $\mu\text{m} \mu\text{m}$
- $\Omega \Omega$
- 10 和 20
- 10, 20 和 30
- 0.13 mm, 0.67 mm 和 0.80 mm
- $10 \sim 20$
- $10^\circ\text{C} \sim 20^\circ\text{C}$

3.3 定理和证明

示例文件中使用 **amsthm** 宏包配置了定理、引理和证明等环境。用户也可以使用 **ntheorem** 宏包。

定义 3.1 If the integral of function f is measurable and non-negative, we define its (extended) **Lebesgue integral** by

$$\int f = \sup_g \int g, \quad (3.5)$$

where the supremum is taken over all measurable functions g such that $0 \leq g \leq f$, and where g is bounded and supported on a set of finite measure.

假设 3.1 The communication graph is strongly connected.

例 3.1 Simple examples of functions on \mathbf{R}^d that are integrable (or non-integrable) are given by

$$f_a(x) = \begin{cases} |x|^{-a} & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 1. \end{cases} \quad (3.6)$$

$$F_a(x) = \frac{1}{1 + |x|^a}, \quad \text{all } x \in \mathbf{R}^d. \quad (3.7)$$

Then f_a is integrable exactly when $a < d$, while F_a is integrable exactly when $a > d$.

引理 3.1 (Fatou) Suppose $\{f_n\}$ is a sequence of measurable functions with $f_n \geq 0$. If $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ for a.e. x , then

$$\int f \leq \liminf_{n \rightarrow \infty} \int f_n. \quad (3.8)$$

注 We do not exclude the cases $\int f = \infty$, or $\liminf_{n \rightarrow \infty} \int f_n = \infty$.

推论 3.2 Suppose f is a non-negative measurable function, and $\{f_n\}$ a sequence of non-negative measurable functions with $f_n(x) \leq f(x)$ and $f_n(x) \rightarrow f(x)$ for almost every x . Then

$$\lim_{n \rightarrow \infty} \int f_n = \int f. \quad (3.9)$$

命题 3.3 Suppose f is integrable on \mathbf{R}^d . Then for every $\epsilon > 0$:

i. There exists a set of finite measure B (a ball, for example) such that

$$\int_{B^c} |f| < \epsilon. \quad (3.10)$$

ii. There is a $\delta > 0$ such that

$$\int_E |f| < \epsilon \quad \text{whenever } m(E) < \delta. \quad (3.11)$$

定理 3.4 Suppose $\{f_n\}$ is a sequence of measurable functions such that $f_n(x) \rightarrow f(x)$ a.e. x , as n tends to infinity. If $|f_n(x)| \leq g(x)$, where g is integrable, then

$$\int |f_n - f| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (3.12)$$

and consequently

$$\int f_n \rightarrow \int f \quad \text{as } n \rightarrow \infty. \quad (3.13)$$

证明 Trivial. □

Axiom of choice Suppose E is a set and E_α is a collection of non-empty subsets of E . Then there is a function $\alpha \mapsto x_\alpha$ (a “choice function”) such that

$$x_\alpha \in E_\alpha, \quad \text{for all } \alpha. \quad (3.14)$$

Observation 1 Suppose a partially ordered set P has the property that every chain has an upper bound in P . Then the set P contains at least one maximal element.

A concise proof Obvious. □

第4章 引用文献的标注

模板使用 **natbib** 宏包来设置参考文献引用的格式，更多引用方法可以参考该宏包的使用说明。

4.1 顺序编码制

4.1.1 角标数字标注法

<code>\cite{knuth86a}</code>	\Rightarrow	[?]
<code>\citet{knuth86a}</code>	\Rightarrow	?]
<code>\cite[42]{knuth86a}</code>	\Rightarrow	[?] ⁴²
<code>\cite{knuth86a,tlc2}</code>	\Rightarrow	[? ?]
<code>\cite{knuth86a, knuth84}</code>	\Rightarrow	[? ?]

4.1.2 数字标注法

<code>\cite{knuth86a}</code>	\Rightarrow	[?]
<code>\citet{knuth86a}</code>	\Rightarrow	?]
<code>\cite[42]{knuth86a}</code>	\Rightarrow	[?] ⁴²
<code>\cite{knuth86a,tlc2}</code>	\Rightarrow	[? ?]
<code>\cite{knuth86a, knuth84}</code>	\Rightarrow	[? ?]

4.2 著者-出版年制标注法

<code>\cite{knuth86a}</code>	\Rightarrow	?
<code>\citep{knuth86a}</code>	\Rightarrow	(?)
<code>\citet[42]{knuth86a}</code>	\Rightarrow	? ⁴²
<code>\citep[42]{knuth86a}</code>	\Rightarrow	(?) ⁴²
<code>\cite{knuth86a,tlc2}</code>	\Rightarrow	??
<code>\cite{knuth86a, knuth84}</code>	\Rightarrow	??

注意，参考文献列表中的每条文献在正文中都要被引用[????????????????]。

附录 A 补充材料

A.1 补充章节

补充内容。

致 谢

在研究学习期间，我有幸得到了三位老师的教导，他们是：我的导师，中国科大 XXX 研究员，中科院 X 昆明动物所马老师以及美国犹他大学的 XXX 老师。三位深厚的学术功底，严谨的工作态度和敏锐的科学洞察力使我受益良多。衷心感谢他们多年来给予我的悉心教导和热情帮助。

感谢 XXX 老师在实验方面的指导以及教授的帮助。科大的 XXX 同学和 XXX 同学参与了部分试验工作，在此深表谢意。

在读期间发表的学术论文与取得的研究成果

已发表论文

1. A A A A A A A A A
2. A A A A A A A A A
3. A A A A A A A A A

待发表论文

1. A A A A A A A A A
2. A A A A A A A A A
3. A A A A A A A A A

研究报告

1. A A A A A A A A A
2. A A A A A A A A A
3. A A A A A A A A A