

Sixth Edition

John L. Hennessy | David A. Patterson

COMPUTER ARCHITECTURE

A Quantitative Approach

MK
MORGAN KAUFMANN

Computer Architecture Formulas

1. $CPU\ time = \text{Instruction count} \times \text{Clock cycles per instruction} \times \text{Clock cycle time}$
2. $X\text{ is }n\text{ times faster than }Y: n = \text{Execution time}_Y / \text{Execution time}_X = \text{Performance}_X / \text{Performance}_Y$
3. $Amdahl's\ Law: \text{Speedup}_{\text{overall}} = \frac{\text{Execution time}_{\text{old}}}{\text{Execution time}_{\text{new}}} = \frac{1}{(1 - \text{Fraction}_{\text{enhanced}}) + \frac{\text{Fraction}_{\text{enhanced}}}{\text{Speedup}_{\text{enhanced}}}}$
4. $Energy_{\text{dynamic}} \propto 1/2 \times \text{Capacitive load} \times \text{Voltage}^2$
5. $Power_{\text{dynamic}} \propto 1/2 \times \text{Capacitive load} \times \text{Voltage}^2 \times \text{Frequency switched}$
6. $Power_{\text{static}} \propto \text{Current}_{\text{static}} \times \text{Voltage}$
7. $Availability = \text{Mean time to fail} / (\text{Mean time to fail} + \text{Mean time to repair})$
8. $Die\ yield = \text{Wafer yield} \times 1 / (1 + \text{Defects per unit area} \times \text{Die area})^N$

where Wafer yield accounts for wafers that are so bad they need not be tested and N is a parameter called the process-complexity factor, a measure of manufacturing difficulty. N ranges from 11.5 to 15.5 in 2011.

9. *Means—arithmetic (AM), weighted arithmetic (WAM), and geometric (GM):*

$$AM = \frac{1}{n} \sum_{i=1}^n \text{Time}_i \quad WAM = \sum_{i=1}^n \text{Weight}_i \times \text{Time}_i \quad GM = \sqrt[n]{\prod_{i=1}^n \text{Time}_i}$$

where Time_i is the execution time for the i th program of a total of n in the workload, Weight_i is the weighting of the i th program in the workload.

10. $\text{Average memory-access time} = \text{Hit time} + \text{Miss rate} \times \text{Miss penalty}$
11. $\text{Misses per instruction} = \text{Miss rate} \times \text{Memory access per instruction}$
12. $\text{Cache index size: } 2^{\text{index}} = \text{Cache size} / (\text{Block size} \times \text{Set associativity})$
13. $\text{Power Utilization Effectiveness (PUE) of a Warehouse Scale Computer} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$

Rules of Thumb

1. *Amdahl/Case Rule:* A balanced computer system needs about 1 MB of main memory capacity and 1 megabit per second of I/O bandwidth per MIPS of CPU performance.
2. *90/10 Locality Rule:* A program executes about 90% of its instructions in 10% of its code.
3. *Bandwidth Rule:* Bandwidth grows by at least the square of the improvement in latency.
4. *2:1 Cache Rule:* The miss rate of a direct-mapped cache of size N is about the same as a two-way set-associative cache of size $N/2$.
5. *Dependability Rule:* Design with no single point of failure.
6. *Watt-Year Rule:* The fully burdened cost of a Watt per year in a Warehouse Scale Computer in North America in 2011, including the cost of amortizing the power and cooling infrastructure, is about \$2.

In Praise of *Computer Architecture: A Quantitative Approach* Sixth Edition

“Although important concepts of architecture are timeless, this edition has been thoroughly updated with the latest technology developments, costs, examples, and references. Keeping pace with recent developments in open-sourced architecture, the instruction set architecture used in the book has been updated to use the RISC-V ISA.”

—from the foreword by Norman P. Jouppi, Google

“*Computer Architecture: A Quantitative Approach* is a classic that, like fine wine, just keeps getting better. I bought my first copy as I finished up my undergraduate degree and it remains one of my most frequently referenced texts today.”

—James Hamilton, Amazon Web Service

“Hennessy and Patterson wrote the first edition of this book when graduate students built computers with 50,000 transistors. Today, warehouse-size computers contain that many servers, each consisting of dozens of independent processors and billions of transistors. The evolution of computer architecture has been rapid and relentless, but *Computer Architecture: A Quantitative Approach* has kept pace, with each edition accurately explaining and analyzing the important emerging ideas that make this field so exciting.”

—James Larus, Microsoft Research

“Another timely and relevant update to a classic, once again also serving as a window into the relentless and exciting evolution of computer architecture! The new discussions in this edition on the slowing of Moore's law and implications for future systems are must-reads for both computer architects and practitioners working on broader systems.”

—Parthasarathy (Partha) Ranganathan, Google

“I love the ‘Quantitative Approach’ books because they are written by engineers, for engineers. John Hennessy and Dave Patterson show the limits imposed by mathematics and the possibilities enabled by materials science. Then they teach through real-world examples how architects analyze, measure, and compromise to build working systems. This sixth edition comes at a critical time: Moore's Law is fading just as deep learning demands unprecedented compute cycles. The new chapter on domain-specific architectures documents a number of promising approaches and prophesies a rebirth in computer architecture. Like the scholars of the European Renaissance, computer architects must understand our own history, and then combine the lessons of that history with new techniques to remake the world.”

—Cliff Young, Google

This page intentionally left blank



Computer Architecture

A Quantitative Approach

Sixth Edition

John L. Hennessy is a Professor of Electrical Engineering and Computer Science at Stanford University, where he has been a member of the faculty since 1977 and was, from 2000 to 2016, its 10th President. He currently serves as the Director of the Knight-Hennessy Fellowship, which provides graduate fellowships to potential future leaders. Hennessy is a Fellow of the IEEE and ACM, a member of the National Academy of Engineering, the National Academy of Science, and the American Philosophical Society, and a Fellow of the American Academy of Arts and Sciences. Among his many awards are the 2001 Eckert-Mauchly Award for his contributions to RISC technology, the 2001 Seymour Cray Computer Engineering Award, and the 2000 John von Neumann Award, which he shared with David Patterson. He has also received 10 honorary doctorates.

In 1981, he started the MIPS project at Stanford with a handful of graduate students. After completing the project in 1984, he took a leave from the university to cofound MIPS Computer Systems, which developed one of the first commercial RISC microprocessors. As of 2017, over 5 billion MIPS microprocessors have been shipped in devices ranging from video games and palmtop computers to laser printers and network switches. Hennessy subsequently led the DASH (Director Architecture for Shared Memory) project, which prototyped the first scalable cache coherent multiprocessor; many of the key ideas have been adopted in modern multiprocessors. In addition to his technical activities and university responsibilities, he has continued to work with numerous start-ups, both as an early-stage advisor and an investor.

David A. Patterson became a Distinguished Engineer at Google in 2016 after 40 years as a UC Berkeley professor. He joined UC Berkeley immediately after graduating from UCLA. He still spends a day a week in Berkeley as an Emeritus Professor of Computer Science. His teaching has been honored by the Distinguished Teaching Award from the University of California, the Karlstrom Award from ACM, and the Mulligan Education Medal and Undergraduate Teaching Award from IEEE. Patterson received the IEEE Technical Achievement Award and the ACM Eckert-Mauchly Award for contributions to RISC, and he shared the IEEE Johnson Information Storage Award for contributions to RAID. He also shared the IEEE John von Neumann Medal and the C & C Prize with John Hennessy. Like his co-author, Patterson is a Fellow of the American Academy of Arts and Sciences, the Computer History Museum, ACM, and IEEE, and he was elected to the National Academy of Engineering, the National Academy of Sciences, and the Silicon Valley Engineering Hall of Fame. He served on the Information Technology Advisory Committee to the President of the United States, as chair of the CS division in the Berkeley EECS department, as chair of the Computing Research Association, and as President of ACM. This record led to Distinguished Service Awards from ACM, CRA, and SIGARCH. He is currently Vice-Chair of the Board of Directors of the RISC-V Foundation.

At Berkeley, Patterson led the design and implementation of RISC I, likely the first VLSI reduced instruction set computer, and the foundation of the commercial SPARC architecture. He was a leader of the Redundant Arrays of Inexpensive Disks (RAID) project, which led to dependable storage systems from many companies. He was also involved in the Network of Workstations (NOW) project, which led to cluster technology used by Internet companies and later to cloud computing. His current interests are in designing domain-specific architectures for machine learning, spreading the word on the open RISC-V instruction set architecture, and in helping the UC Berkeley RISELab (Real-time Intelligent Secure Execution).

Computer Architecture

A Quantitative Approach

Sixth Edition

John L. Hennessy

Stanford University

David A. Patterson

University of California, Berkeley

With Contributions by

Krste Asanović

University of California, Berkeley

Jason D. Bakos

University of South Carolina

Robert P. Colwell

R&E Colwell & Assoc. Inc.

Abhishek Bhattacharjee

Rutgers University

Thomas M. Conte

Georgia Tech

José Duato

Proemisa

Diana Franklin

University of Chicago

David Goldberg

eBay

Norman P. Jouppi

Google

Sheng Li

Intel Labs

Naveen Muralimanohar

HP Labs

Gregory D. Peterson

University of Tennessee

Timothy M. Pinkston

University of Southern California

Parthasarathy Ranganathan

Google

David A. Wood

University of Wisconsin–Madison

Cliff Young

Google

Amr Zaky

University of Santa Clara



MK

MORGAN KAUFMANN PUBLISHERS

AN IMPRINT OF ELSEVIER

Morgan Kaufmann is an imprint of Elsevier
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States

© 2019 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-12-811905-1

For information on all Morgan Kaufmann publications
visit our website at <https://www.elsevier.com/books-and-journals>



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Katey Birtcher

Acquisition Editor: Stephen Merken

Developmental Editor: Nate McFadden

Production Project Manager: Stalin Viswanathan

Cover Designer: Christian J. Bilbow

Typeset by SPi Global, India

To Andrea, Linda, and our four sons

This page intentionally left blank