



CHIP WAR

THE FIGHT FOR THE WORLD'S
MOST CRITICAL TECHNOLOGY

CHRIS MILLER

SCRIBNER

NEW YORK AMSTERDAM/ANTWERP LONDON
TORONTO SYDNEY/MELBOURNE NEW DELHI



Scribner
An Imprint of Simon & Schuster, LLC
1230 Avenue of the Americas
New York, NY 10020

For more than 100 years, Simon & Schuster has championed authors and the stories they create. By respecting the copyright of an author's intellectual property, you enable Simon & Schuster and the author to continue publishing exceptional books for years to come. We thank you for supporting the author's copyright by purchasing an authorized edition of this book.

No amount of this book may be reproduced or stored in any format, nor may it be uploaded to any website, database, language-learning model, or other repository, retrieval, or artificial intelligence system without express permission. All rights reserved. Inquiries may be directed to Simon & Schuster, 1230 Avenue of the Americas, New York, NY 10020 or permissions@simonandschuster.com.

Copyright © 2022 by Christopher Miller

All rights reserved, including the right to reproduce this book or portions thereof in any form whatsoever. For information, address
Scribner Subsidiary Rights Department,
1230 Avenue of the Americas, New York, NY 10020.

First Scribner trade paperback edition September 2025

SCRIBNER and design are trademarks of Simon & Schuster, LLC

Simon & Schuster strongly believes in freedom of expression and stands against censorship in all its forms. For more information, visit BooksBelong.com.

For information about special discounts for bulk purchases,
please contact Simon & Schuster Special Sales at 1-866-506-1949
or business@simonandschuster.com.

The Simon & Schuster Speakers Bureau can bring authors to
your live event. For more information or to book an event,
contact the Simon & Schuster Speakers Bureau at 1-866-248-3049
or visit our website at www.simonspeakers.com.

Interior design by Kyle Kabel

Manufactured in the United States of America

1 3 5 7 9 10 8 6 4 2

Library of Congress Cataloging-in-Publication Data has been applied for.

ISBN 978-1-9821-7200-8
ISBN 978-1-9821-7201-5 (pbk)
ISBN 978-1-9821-7202-2 (ebook)

To Liya

Contents

Cast of Characters	xiii
Glossary	xv
Introduction	xvii

PART I COLD WAR CHIPS

1 From Steel to Silicon	3
2 The Switch	9
3 Noyce, Kilby, and the Integrated Circuit	13
4 Liftoff	19
5 Mortars and Mass Production	23
6 “I . . . WANT . . . TO . . . GET . . . RICH”	29

PART II THE CIRCUITRY OF THE AMERICAN WORLD

7 Soviet Silicon Valley	35
8 “Copy It”	41
9 The Transistor Salesman	45

10	“Transistor Girls”	51
11	Precision Strike	57
12	Supply Chain Statecraft	63
13	Intel’s Revolutionaries	67
14	The Pentagon’s Offset Strategy	73

PART III LEADERSHIP LOST?

15	“That Competition Is Tough”	81
16	“At War with Japan”	85
17	“Shipping Junk”	91
18	The Crude Oil of the 1980s	97
19	Death Spiral	103
20	The Japan That Can Say No	109

PART IV AMERICA RESURGENT

21	The Potato Chip King	117
22	Disrupting Intel	123
23	“My Enemy’s Enemy”: The Rise of Korea	129
24	“This Is the Future”	135
25	The KGB’s Directorate T	141
26	“Weapons of Mass Destruction”: The Impact of the Offset	145
27	War Hero	151
28	“The Cold War Is Over and You Have Won”	155

PART V INTEGRATED CIRCUITS, INTEGRATED WORLD?

29	“We Want a Semiconductor Industry in Taiwan”	163
30	“All People Must Make Semiconductors”	171
31	“Sharing God’s Love with the Chinese”	177
32	Lithography Wars	183
33	The Innovator’s Dilemma	191
34	Running Faster?	197

PART VI OFFSHORING INNOVATION?

35	“Real Men Have Fabs”	205
36	The Fabless Revolution	209
37	Morris Chang’s Grand Alliance	215
38	Apple Silicon	221
39	EUV	225
40	“There Is No Plan B”	231
41	How Intel Forgot Innovation	235

PART VII CHINA’S CHALLENGE

42	Made in China	243
43	“Call Forth the Assault”	247
44	Technology Transfer	255
45	“Mergers Are Bound to Happen”	263
46	The Rise of Huawei	269
47	The 5G Future	277
48	The Next Offset	283

PART VIII THE CHIP CHOKE

49	“Everything We’re Competing On”	295
50	Fujian Jinhua	305
51	The Assault on Huawei	311
52	China’s Sputnik Moment?	319
53	Shortages and Supply Chains	327
54	The Taiwan Dilemma	335
	Conclusion	345
	Afterword	353
	<i>Acknowledgments</i>	365
	<i>Notes</i>	369
	<i>Index</i>	427

Cast of Characters

Morris Chang: Founder of Taiwan Semiconductor Manufacturing Company (TSMC), the world's most important chipmaker; previously, a senior executive at Texas Instruments.

Andy Grove: Former president and CEO of Intel during the 1980s and 1990s; notorious for his aggressive style and success in reviving Intel; author of *Only the Paranoid Survive*.

Pat Haggerty: Chairman of Texas Instruments; led the company as it specialized in building microelectronics, including for the U.S. military.

Jack Kilby: Co-inventor of the integrated circuit, in 1958; longtime Texas Instruments employee; winner of the Nobel Prize.

Jay Lathrop: Co-inventor of photolithography, the process of patterning transistors using specialized chemicals and light; formerly of Texas Instruments.

Carver Mead: Professor at the California Institute of Technology (Caltech); advisor to Fairchild Semiconductor and Intel; visionary thinker about the future of technology.

Gordon Moore: Cofounder of Fairchild Semiconductor and Intel; creator in 1965 of “Moore’s Law,” which predicted that the computing power on each chip would double every couple of years.

Akio Morita: Cofounder of Sony; coauthor of *The Japan That Can Say No*; represented Japanese business on the world stage during the 1970s and 1980s.

Robert Noyce: Cofounder of Fairchild Semiconductor and Intel; co-inventor of the integrated circuit in 1959; known as the “Mayor of Silicon Valley”; first leader of Sematech.

William Perry: Pentagon official from 1977–1981 and later Secretary of Defense from 1994 to 1997 who advocated using chips to produce precision-strike weapons.

Jerry Sanders: Founder and CEO of AMD; Silicon Valley’s most flamboyant salesman; an aggressive critic of what he saw as unfair Japanese trade practices in the 1980s.

Charlie Sporck: Drove the offshoring of chip assembly while leading manufacturing operations at Fairchild Semiconductor; later CEO of National Semiconductor.

Ren Zhengfei: Founder of Huawei, China’s telecom and chip-design giant; his daughter Meng Wanzhou was arrested in Canada in 2018 on charges of violating U.S. law and trying to evade U.S. sanctions.

Glossary

Arm: a UK company that licenses to chip designers use of an instruction set architecture—a set of basic rules governing how a given chip operates. The Arm architecture is dominant in mobile devices and is slowly winning market share in PCs and data centers.

Chip (also “integrated circuit” or “semiconductor”): a small piece of semiconducting material, usually silicon, with millions or billions of microscopic transistors carved into it.

CPU: central processing unit; a type of “general-purpose” chip that is the workhorse of computing in PCs, phones, and data centers.

DRAM: dynamic random access memory; one of two main types of memory chip, which is used to store data temporarily.

EDA: electronic design automation; specialized software used to design how millions or billions of transistors will be arrayed on a chip and to simulate their operation.

FinFET: a 3D transistor structure first implemented in the early 2010s to better control transistor operation as transistors’ size shrank to nanometric-scale.

GPU: graphics processing unit; a chip that is capable of parallel processing, making it useful for graphics and for artificial intelligence applications.

Logic chip: a chip that processes data.

Memory chip: a chip that remembers data.

NAND: also called “flash,” the second major type of memory chip, used for longer-term data storage.

Photolithography: also known as “lithography”; the process of shining light or ultraviolet light through patterned masks: the light then interacts with photoresist chemicals to carve patterns on silicon wafers.

RISC-V: an open-source architecture growing in popularity because it is free to use, unlike Arm and x86. The development of RISC-V was partially funded by the U.S. government but now is popular in China because it is not subject to U.S. export controls.

Silicon wafer: a circular piece of ultra-pure silicon, usually eight or twelve inches in diameter, out of which chips are carved.

Transistor: a tiny electric “switch” that turns on (creating a 1) or off (0), producing the 1s and 0s that undergird all digital computing.

x86: an instruction set architecture that is dominant in PCs and data centers. Intel and AMD are the two main firms producing such chips.

CHIP WAR

Introduction

The destroyer USS *Mustin* slipped into the northern end of the Taiwan Strait on August 18, 2020, its five-inch gun pointed southward as it began a solo mission to sail through the Strait and reaffirm that these international waters were *not* controlled by China—at least not yet. A stiff southwestern breeze whipped across the deck as it steamed south. High clouds cast shadows on the water that seemed to stretch all the way to the great port cities of Fuzhou, Xiamen, Hong Kong, and the other harbors that dot the South China coast. To the east, the island of Taiwan rose in the distance, a broad, densely settled coastal plain giving way to tall peaks hidden in clouds. Aboard ship, a sailor wearing a navy baseball cap and a surgical mask lifted his binoculars and scanned the horizon. The waters were filled with commercial freighters shipping goods from Asia’s factories to consumers around the world.

On board the USS *Mustin*, a row of sailors sat in a dark room in front of an array of brightly colored screens on which were displayed data from planes, drones, ships, and satellites tracking movement across the Indo-Pacific. Atop the *Mustin*’s bridge, a radar array fed into the ship’s computers. On deck ninety-six launch cells stood ready, each capable of firing missiles that could precisely strike planes, ships, or submarines dozens or even hundreds of miles away. During the crises of

the Cold War, the U.S. military had used threats of brute nuclear force to defend Taiwan. Today, it relies on microelectronics and precision strikes.

As the USS *Mustin* sailed through the Strait, bristling with computerized weaponry, the People's Liberation Army announced a retaliatory series of live-fire exercises around Taiwan, practicing what one Beijing-controlled newspaper called a "reunification-by-force operation." But on this particular day, China's leaders worried less about the U.S. Navy and more about an obscure U.S. Commerce Department regulation called the Entity List, which limits the transfer of American technology abroad. Previously, the Entity List had primarily been used to prevent sales of military systems like missile parts or nuclear materials. Now, though, the U.S. government was dramatically tightening the rules governing computer chips, which had become ubiquitous in both military systems and consumer goods.

The target was Huawei, China's tech giant, which sells smartphones, telecom equipment, cloud computing services, and other advanced technologies. The U.S. feared that Huawei's products were now priced so attractively, partly owing to Chinese government subsidies, that they'd shortly form the backbone of next-generation telecom networks. America's dominance of the world's tech infrastructure would be undermined. China's geopolitical clout would grow. To counter this threat, the U.S. barred Huawei from buying advanced computer chips made with U.S. technology.

Soon, the company's global expansion ground to a halt. Entire product lines became impossible to produce. Revenue slumped. A corporate giant faced technological asphyxiation. Huawei discovered that, like all other Chinese companies, it was fatally dependent on foreigners to make the chips upon which all modern electronics depend.

The United States *still* has a stranglehold on the silicon chips that gave Silicon Valley its name, though its position has weakened dangerously. China now spends more money each year importing chips than it spends on oil. These semiconductors are plugged into all manner of devices, from smartphones to refrigerators, that China consumes

at home or exports worldwide. Armchair strategists theorize about China's "Malacca Dilemma"—a reference to the main shipping channel between the Pacific and Indian Oceans—and the country's ability to access supplies of oil and other commodities amid a crisis. Beijing, however, is more worried about a blockade measured in bytes rather than barrels. China is devoting its best minds and billions of dollars to developing its own semiconductor technology in a bid to free itself from America's chip choke.

If Beijing succeeds, it will remake the global economy and reset the balance of military power. World War II was decided by steel and aluminum, and followed shortly thereafter by the Cold War, which was defined by atomic weapons. The rivalry between the United States and China may well be determined by computing power. Strategists in Beijing and Washington now realize that all advanced tech—from machine learning to missile systems, from automated vehicles to armed drones—requires cutting-edge chips, known more formally as semiconductors or integrated circuits. A tiny number of companies control their production.

We rarely think about chips, yet they've created the modern world. The fate of nations has turned on their ability to harness computing power. Globalization as we know it wouldn't exist without the trade in semiconductors and the electronic products they make possible. America's military primacy stems largely from its ability to apply chips to military uses. Asia's tremendous rise over the past half century has been built on a foundation of silicon as its growing economies have come to specialize in fabricating chips and assembling the computers and smartphones that these integrated circuits make possible.

At the core of computing is the need for many millions of 1s and 0s. The entire digital universe consists of these two numbers. Every button on your iPhone, every email, photograph, and YouTube video—all of these are coded, ultimately, in vast strings of 1s and 0s. But these numbers don't actually exist. They're expressions of electrical currents, which are either on (1) or off (0). A chip is a grid of millions or billions of

transistors, tiny electrical switches that flip on and off to process these digits, to remember them, and to convert real world sensations like images, sound, and radio waves into millions and millions of 1s and 0s.

As the USS *Mustin* sailed southward, factories and assembly facilities on both sides of the Strait were churning out components for the iPhone 12, which was only two months away from its October 2020 launch. Around a quarter of the chip industry's revenue comes from phones; much of the price of a new phone pays for the semiconductors inside. For the past decade, each generation of iPhone has been powered by one of the world's most advanced processor chips. In total, it takes over a dozen semiconductors to make a smartphone work, with different chips managing the battery, Bluetooth, Wi-Fi, cellular network connections, audio, the camera, and more.

Apple makes precisely *none* of these chips. It buys most off-the-shelf: memory chips from Japan's Kioxia, radio frequency chips from California's Skyworks, audio chips from Cirrus Logic, based in Austin, Texas. Apple designs in-house the ultra-complex processors that run an iPhone's operating system. But the Cupertino, California, colossus can't manufacture these chips. Nor can any company in the United States, Europe, Japan, or China. Today, Apple's most advanced processors—which are arguably the world's most advanced semiconductors—can only be produced by a single company in a single building, the most expensive factory in human history, which on the morning of August 18, 2020, was only a couple dozen miles off the USS *Mustin*'s port bow.

Fabricating and miniaturizing semiconductors has been the greatest engineering challenge of our time. Today, no firm fabricates chips with more precision than the Taiwan Semiconductor Manufacturing Company, better known as TSMC. In 2020, as the world lurched between lockdowns driven by a virus whose diameter measured around one hundred nanometers—billionths of a meter—TSMC's most advanced facility, Fab 18, was carving microscopic mazes of tiny transistors, etching shapes smaller than half the size of a coronavirus, a hundredth the size of a mitochondria. TSMC replicated this process at a scale

previously unparalleled in human history. Apple sold over 100 million iPhone 12s, each powered by an A14 processor chip with 11.8 billion tiny transistors carved into its silicon. In a matter of months, in other words, for just one of the dozen chips in an iPhone, TSMC's Fab 18 fabricated well over 1 quintillion transistors—that is, a number with eighteen zeros behind it. Last year, the chip industry produced more transistors than the combined quantity of all goods produced by all other companies, in all other industries, in all human history. Nothing else comes close.

It was just over sixty years ago that the number of transistors on a cutting-edge chip wasn't 11.8 billion, but 4. In 1961, south of San Francisco, a small firm called Fairchild Semiconductor announced a new product called the Micrologic, a silicon chip with four transistors embedded in it. Soon the company devised ways to put a dozen transistors on a chip, then a hundred. Fairchild cofounder Gordon Moore noticed in 1965 that the number of components that could be fit on each chip was doubling annually as engineers learned to fabricate ever smaller transistors. This prediction—that the computing power of chips would grow exponentially—came to be called “Moore's Law” and led Moore to predict the invention of devices that in 1965 seemed impossibly futuristic, like an “electronic wristwatch,” “home computers,” and even “personal portable communications equipment.” Looking forward from 1965, Moore predicted a decade of exponential growth—but this staggering rate of progress has continued for over half a century. In 1970, the second company Moore founded, Intel, unveiled a memory chip that could remember 1,024 pieces of information (“bits”). It cost around \$20, roughly two cents per bit. Today, \$20 can buy a thumb drive that can remember many billions of bits.

When we think of Silicon Valley today, our minds conjure social networks and software companies rather than the material after which the valley was named. Yet the internet, the cloud, social media, and the entire digital world only exist because engineers have learned to control the most minute movement of electrons as they race across

slabs of silicon. “Big tech” wouldn’t exist if the cost of processing and remembering 1s and 0s hadn’t fallen by a billionfold in the past half century.

This incredible ascent is partly thanks to brilliant scientists and Nobel Prize–winning physicists. But not every invention creates a successful startup, and not every startup sparks a new industry that transforms the world. Semiconductors spread across society because companies devised new techniques to manufacture them by the millions, because hard-charging managers relentlessly drove down their cost, and because creative entrepreneurs imagined new ways to use them. The making of Moore’s Law is as much a story of manufacturing experts, supply chain specialists, and marketing managers as it is about physicists or electrical engineers.

The towns to the south of San Francisco—which weren’t called Silicon Valley until the 1970s—were the epicenter of this revolution because they combined scientific expertise, manufacturing know-how, and visionary business thinking. California had plenty of engineers trained in aviation or radio industries who’d graduated from Stanford or Berkeley, each of which was flush with defense dollars as the U.S. military sought to solidify its technological advantage. California’s culture mattered just as much as any economic structure, however. The people who left America’s East Coast, Europe, and Asia to build the chip industry often cited a sense of boundless opportunity in their decision to move to Silicon Valley. For the world’s smartest engineers and most creative entrepreneurs, there was simply no more exciting place to be.

Once the chip industry took shape, it proved impossible to dislodge from Silicon Valley. Today’s semiconductor supply chain requires components from many cities and countries, but almost every chip made still has a Silicon Valley connection or is produced with tools designed and built in California. America’s vast reserve of scientific expertise, nurtured by government research funding and strengthened by the ability to poach the best scientists from other countries, has provided the core knowledge driving technological advances forward. The country’s

network of venture capital firms and its stock markets have provided the startup capital new firms need to grow—and have ruthlessly forced out failing companies. Meanwhile, the world's largest consumer market in the U.S. has driven the growth that's funded decades of R&D on new types of chips.

Other countries have found it impossible to keep up on their own but have succeeded when they've deeply integrated themselves into Silicon Valley's supply chains. Europe has isolated islands of semiconductor expertise, notably in producing the machine tools needed to make chips and in designing chip architectures. Asian governments, in Taiwan, South Korea, and Japan, have elbowed their way into the chip industry by subsidizing firms, funding training programs, keeping their exchange rates undervalued, and imposing tariffs on imported chips. This strategy has yielded certain capabilities that no other countries can replicate—but they've achieved what they have in partnership with Silicon Valley, continuing to rely fundamentally on U.S. tools, software, and customers. Meanwhile, America's most successful chip firms have built supply chains that stretch across the world, driving down costs and producing the expertise that has made Moore's Law possible.

Today, thanks to Moore's Law, semiconductors are embedded in every device that requires computing power—and in the age of the Internet of Things, this means pretty much *every* device. Even hundred-year-old products like automobiles now often include a thousand dollars worth of chips. Most of the world's GDP is produced with devices that rely on semiconductors. For a product that didn't exist seventy-five years ago, this is an extraordinary ascent.

As the USS *Mustin* steamed southward in August 2020, the world was just beginning to reckon with our reliance on semiconductors—and our dependence on Taiwan, which fabricates the chips that produce a third of the new computing power we use each year. Taiwan's TSMC builds almost all the world's most advanced processor chips. When COVID slammed into the world in 2020, it disrupted the chip industry, too. Some factories were temporarily shuttered. Purchases of chips for

autos slumped. Demand for PC and data center chips spiked higher, as much of the world prepared to work from home. Then, over 2021, a series of accidents—a fire in a Japanese semiconductor facility; ice storms in Texas, a center of U.S. chipmaking; and a new round of COVID lockdowns in Malaysia, where many chips are assembled and tested—intensified these disruptions. Suddenly, many industries far from Silicon Valley faced debilitating chip shortages. Big carmakers from Toyota to General Motors had to shut factories for weeks because they couldn't acquire the semiconductors they needed. Shortages of even the simplest chips caused factory closures on the opposite side of the world. It seemed like a perfect image of globalization gone wrong.

Political leaders in the U.S., Europe, and Japan hadn't thought much about semiconductors in decades. Like the rest of us, they thought "tech" meant search engines or social media, not silicon wafers. When Joe Biden and Angela Merkel asked why their country's car factories were shuttered, the answer was shrouded behind semiconductor supply chains of bewildering complexity. A typical chip might be designed with blueprints from the Japanese-owned, UK-based company called Arm, by a team of engineers in California and Israel, using design software from the United States. When a design is complete, it's sent to a facility in Taiwan, which buys ultra-pure silicon wafers and specialized gases from Japan. The design is carved into silicon using some of the world's most precise machinery, which can etch, deposit, and measure layers of materials a few atoms thick. These tools are produced primarily by five companies, one Dutch, one Japanese, and three Californian, without which advanced chips are basically impossible to make. Then the chip is packaged and tested, often in Southeast Asia, before being sent to China for assembly into a phone or computer.

If any one of the steps in the semiconductor production process is interrupted, the world's supply of new computing power is imperiled. In the age of AI, it's often said that data is the new oil. Yet the real limitation we face isn't the availability of data but of processing power. There's a finite number of semiconductors that can store and process

data. Producing them is mind-bogglingly complex and horrendously expensive. Unlike oil, which can be bought from many countries, our production of computing power depends fundamentally on a series of choke points: tools, chemicals, and software that often are produced by a handful of companies—and sometimes only by one. No other facet of the economy is so dependent on so few firms. Chips from Taiwan provide around a third of the world's new computing power each year. Two Korean companies produce 52 percent of the world's DRAM memory chips. The Dutch company ASML builds 100 percent of the world's extreme ultraviolet lithography machines, without which cutting-edge chips are simply impossible to make. OPEC's 33 percent market share of world oil production looks unimpressive by comparison.

The global network of companies that annually produces a trillion chips at nanometer scale is a triumph of efficiency. It's also a staggering vulnerability. The disruptions of the pandemic provide just a glimpse of what a single well-placed earthquake could do to the global economy. Taiwan sits atop a fault line that as recently as 1999 produced an earthquake measuring 7.3 on the Richter scale. Thankfully, this only knocked chip production offline for a couple of days. But it's only a matter of time before a stronger quake strikes Taiwan. A devastating quake could also hit Japan, an earthquake-prone country that produces 17 percent of the world's chips, or Silicon Valley, which today produces few chips but builds crucial chipmaking machinery in facilities sitting atop the San Andreas Fault.

Yet the seismic shift that most imperils semiconductor supply today isn't the crash of tectonic plates but the clash of great powers. As China and the United States struggle for supremacy, both Washington and Beijing are fixated on controlling the future of computing—and, to a frightening degree, that future is dependent on a small island that Beijing considers a renegade province and America has committed to defend by force.

The interconnections between the chip industries in the U.S., China, and Taiwan are dizzyingly complex. There's no better illustration of

this than the individual who founded TSMC, a company that until 2020 counted America's Apple and China's Huawei as its two biggest customers. Morris Chang was born in mainland China; grew up in World War II-era Hong Kong; was educated at Harvard, MIT, and Stanford; helped build America's early chip industry while working for Texas Instruments in Dallas; held a top secret U.S. security clearance to develop electronics for the American military; and made Taiwan the epicenter of world semiconductor manufacturing. Some foreign policy strategists in Beijing and Washington dream of decoupling the two countries' tech sectors, but the ultra-efficient international network of chip designers, chemical suppliers, and machine-tool makers that people like Chang helped build can't be easily unwound.

Unless, of course, something explodes. Beijing has pointedly refused to rule out the prospect that it might invade Taiwan to "reunify" it with the mainland. But it wouldn't take anything as dramatic as an amphibious assault to send semiconductor-induced shock waves careening through the global economy. Even a partial blockade by Chinese forces would trigger devastating disruptions. A single missile strike on TSMC's most advanced chip fabrication facility could easily cause hundreds of billions of dollars of damage once delays to the production of phones, data centers, autos, telecom networks, and other technology are added up.

Holding the global economy hostage to one of the world's most dangerous political disputes might seem like an error of historic proportions. However, the concentration of advanced chip manufacturing in Taiwan, South Korea, and elsewhere in East Asia isn't an accident. A series of deliberate decisions by government officials and corporate executives created the far-flung supply chains we rely on today. Asia's vast pool of cheap labor attracted chipmakers looking for low-cost factory workers. The region's governments and corporations used offshored chip assembly facilities to learn about, and eventually domesticate, more advanced technologies. Washington's foreign policy strategists embraced complex semiconductor supply chains as a tool to bind Asia

to an American-led world. Capitalism's inexorable demand for economic efficiency drove a constant push for cost cuts and corporate consolidation. The steady tempo of technological innovation that underwrote Moore's Law required ever more complex materials, machinery, and processes that could only be supplied or funded via global markets. And our gargantuan demand for computing power only continues to grow.

Drawing on research in historical archives on three continents, from Taipei to Moscow, and over a hundred interviews with scientists, engineers, CEOs, and government officials, this book contends that semiconductors have defined the world we live in, determining the shape of international politics, the structure of the world economy, and the balance of military power. Yet this most modern of devices has a complex and contested history. Its development has been shaped not only by corporations and consumers but also by ambitious governments and the imperatives of war. To understand how our world came to be defined by quintillions of transistors and a tiny number of irreplaceable companies, we must begin by looking back to the origins of the silicon age.

PART I

COLD WAR CHIPS

CHAPTER 1

From Steel to Silicon

Japanese soldiers described World War II as a “typhoon of steel.” It certainly felt that way to Akio Morita, a studious young engineer from a family of prosperous sake merchants. Morita only barely avoided the front lines by getting assigned to a Japanese navy engineering lab. But the typhoon of steel crashed through Morita’s homeland, too, as American B-29 Superfortress bombers pummeled Japan’s cities, destroying much of Tokyo and other urban centers. Adding to the devastation, an American blockade created widespread hunger and drove the country toward desperate measures. Morita’s brothers were being trained as kamikaze pilots when the war ended.

Across the East China Sea, Morris Chang’s childhood was punctuated by the sound of gunfire and air-raid sirens warning of imminent attack. Chang spent his teenage years fleeing the Japanese armies that swept across China, moving to Guangzhou; the British colony of Hong Kong; China’s wartime capital of Chongqing; and then back to Shanghai after the Japanese were defeated. Even then, the war didn’t really end, because Communist guerillas relaunched their struggle against the Chinese government. Soon Mao Zedong’s forces were marching on Shanghai. Morris Chang was once again a refugee, forced to flee to Hong Kong for the second time.

Budapest was on the opposite side of the world, but Andy Grove lived through the same typhoon of steel that swept across Asia. Andy (or Andras Grof, as he was then known) survived multiple invasions of Budapest. Hungary's far-right government treated Jews like the Groves as second-class citizens, but when war broke out in Europe, his father was nevertheless drafted and sent to fight alongside Hungary's Nazi allies against the Soviet Union, where he was reported missing in action at Stalingrad. Then, in 1944, the Nazis invaded Hungary, their ostensible ally, sending tank columns rolling through Budapest and announcing plans to ship Jews like Grove to industrial-scale death camps. Still a child, Grove heard the thud of artillery again months later as Red Army troops marched into Hungary's capital, "liberating" the country, raping Grove's mother, and installing a brutal puppet regime in the Nazis' place.

Endless tank columns; waves of airplanes; thousands of tons of bombs dropped from the skies; convoys of ships delivering trucks, combat vehicles, petroleum products, locomotives, rail cars, artillery, ammunition, coal, and steel—World War II was a conflict of industrial attrition. The United States wanted it that way: an industrial war was a struggle America would win. In Washington, the economists at the War Production Board measured success in terms of copper and iron, rubber and oil, aluminum and tin as America converted manufacturing might into military power.

The United States built more tanks than all the Axis powers combined, more ships, more planes, and twice the Axis production of artillery and machine guns. Convoys of industrial goods streamed from American ports across the Atlantic and Pacific Oceans, supplying Britain, the Soviet Union, China, and other allies with key materiel. The war was waged by soldiers at Stalingrad and sailors at Midway. But the fighting power was produced by America's Kaiser shipyards and the assembly lines at River Rouge.

In 1945, radio broadcasts across the world announced that the war was finally over. Outside of Tokyo, Akio Morita, the young engineer,

donned his full uniform to hear Emperor Hirohito's surrender address, though he listened to the speech alone rather than in the company of other naval officers, so he wouldn't be pressured to commit ritual suicide. Across the East China Sea, Morris Chang celebrated the war's end and Japan's defeat with a prompt return to a leisurely teenaged life of tennis, movies, and card games with friends. In Hungary, Andy Grove and his mother slowly crept out of their bomb shelter, though they suffered as much during the Soviet occupation as during the war itself.

World War II's outcome was determined by industrial output, but it was clear already that new technologies were transforming military power. The great powers had manufactured planes and tanks by the thousands, but they'd also built research labs that developed new devices like rockets and radars. The two atomic bombs that destroyed Hiroshima and Nagasaki brought forth much speculation that a nascent Atomic Age might replace an era defined by coal and steel.

Morris Chang and Andy Grove were schoolboys in 1945, too young to have thought seriously about technology or politics. Akio Morita, however, was in his early twenties and had spent the final months of the war developing heat-seeking missiles. Japan was far from fielding workable guided missiles, but the project gave Morita a glimpse of the future. It was becoming possible to envision wars won not by riveters on assembly lines but by weapons that could identify targets and maneuver themselves automatically. The idea seemed like science fiction, but Morita was vaguely aware of new developments in electronic computation that might make it possible for machines to "think" by solving math problems like adding, multiplying, or finding a square root.

Of course, the idea of using devices to compute wasn't new. People have flipped their fingers up and down since *Homo sapiens* first learned to count. The ancient Babylonians invented the abacus to manipulate large numbers, and for centuries people multiplied and divided by moving wooden beads back and forth across these wooden grids. During

the late 1800s and early 1900s, the growth of big bureaucracies in government and business required armies of human “computers,” office workers armed with pen, paper, and occasionally simple mechanical calculators—gearboxes that could add, subtract, multiply, divide, and calculate basic square roots.

These living, breathing computers could tabulate payrolls, track sales, collect census results, and sift through the data on fires and droughts that were needed to price insurance policies. During the Great Depression, America’s Works Progress Administration, looking to employ jobless office workers, set up the Mathematical Tables Project. Several hundred human “computers” sat at rows of desks in a Manhattan office building and tabulated logarithms and exponential functions. The project published twenty-eight volumes of the results of complex functions, with titles such as *Tables of Reciprocals of the Integers from 100,000 Through 200,009*, presenting 201 pages covered in tables of numbers.

Organized groups of human calculators showed the promise of computation, but also the limits of using brains to compute. Even when brains were enhanced by using mechanical calculators, humans worked slowly. A person looking to use the results of the Mathematical Tables Project had to flip through the pages of one of the twenty-eight volumes to find the result of a specific logarithm or exponent. The more calculations that were needed, the more pages had to be flipped through.

Meanwhile, the demand for calculations kept growing. Even before World War II, money was flowing into projects to produce more capable mechanical computers, but the war accelerated the hunt for computing power. Several countries’ air forces developed mechanical bombsights to help aviators hit their targets. Bomber crews entered the wind speed and altitude by turning knobs, which moved metal levers that adjusted glass mirrors. These knobs and levers “computed” altitudes and angles more exactly than any pilot could, focusing the sight as the plane homed in on its target. However, the limitations were obvious. Such bombsights

only considered a few inputs and provided a single output: when to drop the bomb. In perfect test conditions, America's bombsights were more accurate than pilots' guesswork. When deployed in the skies above Germany, though, only 20 percent of American bombs fell within one thousand feet of their target. The war was decided by the quantity of bombs dropped and artillery shells fired, not by the knobs on the mechanical computers that tried and usually failed to guide them.

More accuracy required more calculations. Engineers eventually began replacing mechanical gears in early computers with electrical charges. Early electric computers used the vacuum tube, a lightbulb-like metal filament enclosed in glass. The electric current running through the tube could be switched on and off, performing a function not unlike an abacus bead moving back and forth across a wooden rod. A tube turned on was coded as a 1 while a vacuum tube turned off was a 0. These two digits could produce any number using a system of binary counting—and therefore could theoretically execute many types of computation.

Moreover, vacuum tubes made it possible for these digital computers to be reprogrammed. Mechanical gears such as those in a bombsight could only perform a single type of calculation because each knob was physically attached to levers and gears. The beads on an abacus were constrained by the rods on which they moved back and forth. However, the connections between vacuum tubes could be reorganized, enabling the computer to run different calculations.

This was a leap forward in computing—or it would have been, if not for the moths. Because vacuum tubes glowed like lightbulbs, they attracted insects, requiring regular “debugging” by their engineers. Also like lightbulbs, vacuum tubes often burned out. A state-of-the-art computer called ENIAC, built for the U.S. Army at the University of Pennsylvania in 1945 to calculate artillery trajectories, had eighteen thousand vacuum tubes. On average, one tube malfunctioned every two days, bringing the entire machine to a halt and sending technicians scrambling to find and replace the broken part. ENIAC could multiply

hundreds of numbers per second, faster than any mathematician. Yet it took up an entire room because each of its eighteen thousand tubes was the size of a fist. Clearly, vacuum tube technology was too cumbersome, too slow, and too unreliable. So long as computers were moth-ridden monstrosities, they'd only be useful for niche applications like code breaking, unless scientists could find a smaller, faster, cheaper switch.

The Switch

William Shockley had long assumed that if a better “switch” was to be found, it would be with the help of a type of material called semiconductors. Shockley, who’d been born in London to a globe-trotting mining engineer, had grown up amid the fruit trees of the sleepy California town of Palo Alto. An only child, he was utterly convinced of his superiority over anyone around him—and he let everyone know it. He went to college at Caltech, in Southern California, before completing a PhD in physics at MIT and starting work at Bell Labs in New Jersey, which at the time was one of the world’s leading centers of science and engineering. All his colleagues found Shockley obnoxious, but they also admitted he was a brilliant theoretical physicist. His intuition was so accurate that one of Shockley’s coworkers said it was as if he could actually *see* electrons as they zipped across metals or bonded atoms together.

Semiconductors, Shockley’s area of specialization, are a unique class of materials. Most materials either let electric current flow freely (like copper wires) or block current (like glass). Semiconductors are different. On their own, semiconductor materials like silicon and germanium are like glass, conducting hardly any electricity at all. But when certain materials are added and an electric field is applied, current can begin

to flow. Adding phosphorous or antimony to semiconducting materials like silicon or germanium, for example, lets a negative current flow.

“Doping” semiconductor materials with other elements presented an opportunity for new types of devices that could create and control electric currents. However, mastering the flow of electrons across semiconductor materials like silicon or germanium was a distant dream so long as their electrical properties remained mysterious and unexplained. Until the late 1940s, despite all the physics brainpower accumulated at Bell Labs, no one could explain why slabs of semiconductor materials acted in such puzzling ways.

In 1945, Shockley first theorized what he called a “solid state valve,” sketching in his notebook a piece of silicon attached to a ninety-volt battery. He hypothesized that placing a piece of semiconductor material like silicon in the presence of an electric field could attract “free electrons” stored inside to cluster near the edge of the semiconductor. If enough electrons were attracted by the electric field, the edge of the semiconductor would be transformed into a conductive material, like a metal, which always has large numbers of free electrons. If so, an electric current could begin flowing through a material that previously conducted no electricity at all. Shockley soon built such a device, expecting that applying and removing an electric field on top of the piece of silicon could make it function like a valve, opening and closing the flow of electrons across the silicon. When he ran this experiment, however, he was unable to detect a result. “Nothing measurable,” he explained. “Quite mysterious.” In fact, the simple instruments of the 1940s were too imprecise to measure the tiny current that was flowing.

Two years later, two of Shockley’s colleagues at Bell Labs devised a similar experiment on a different type of device. Where Shockley was proud and obnoxious, his colleagues Walter Brattain, a brilliant experimental physicist from a cattle ranch in rural Washington, and John Bardeen, a Princeton-trained scientist who’d later become the only person to win two Nobel Prizes in physics, were modest and mild-mannered. Inspired by Shockley’s theorizing, Brattain and Bardeen

built a device that applied two gold filaments, each attached by wires to a power source and to a piece of metal, to a block of germanium, with each filament touching the germanium less than a millimeter apart from the other. On the afternoon of December 16, 1947, at Bell Labs' headquarters, Bardeen and Brattain switched on the power and were able to control the current surging across the germanium. Shockley's theories about semiconductor materials had been proven correct.

AT&T, which owned Bell Labs, was in the business of telephones, not computers, and saw this device—soon christened a “transistor”—as useful primarily for its ability to amplify signals that transmitted phone calls across its vast network. Because transistors could amplify currents, it was soon realized, they would be useful in devices such as hearing aids and radios, replacing less reliable vacuum tubes, which were also used for signal amplification. Bell Labs soon began arranging patent applications for this new device.

Shockley was furious that his colleagues had discovered an experiment to prove his theories, and he was committed to outdoing them. He locked himself in a Chicago hotel room for two weeks over Christmas and began imagining different transistor structures, based on his unparalleled understanding of semiconductor physics. By January 1948, he'd conceptualized a new type of transistor, made up of three chunks of semiconductor material. The outer two chunks would have a surplus of electrons; the piece sandwiched between them would have a deficit. If a tiny current was applied to the middle layer in the sandwich, it set a much larger current flowing across the entire device. This conversion of a small current into a large one was the same amplification process that Brattain and Bardeen's transistor had demonstrated. But Shockley began to perceive other uses, along the lines of the “solid state valve” he'd previously theorized. He could turn the larger current on and off by manipulating the small current applied to the middle of this transistor sandwich. On, off. On, off. Shockley had designed a switch.

When Bell Labs held a press conference in June 1948 to announce that its scientists had invented the transistor, it wasn't easy to understand

why these wired blocks of germanium merited a special announcement. The *New York Times* buried the story on page 46. *Time* magazine did better, reporting the invention under the headline “Little Brain Cell.” Yet even Shockley, who never underestimated his own importance, couldn’t have imagined that soon thousands, millions, and billions of these transistors would be employed at microscopic scale to replace human brains in the task of computing.

Noyce, Kilby, and the Integrated Circuit

The transistor could only replace vacuum tubes if it could be simplified and sold at scale. Theorizing and inventing transistors was simply the first step; now, the challenge was to manufacture them by the thousands. Brattain and Bardeen had little interest in business or mass production. They were researchers at heart, and after winning the Nobel, they continued their careers teaching and experimenting. Shockley's ambitions, by contrast, only grew. He wanted not only to be famous but also to be rich. He told friends he dreamed of seeing his name not only in academic publications like the *Physical Review* but in the *Wall Street Journal*, too. In 1955, he established Shockley Semiconductor in the San Francisco suburb of Mountain View, California, just down the street from Palo Alto, where his aging mother still lived.

Shockley planned to build the world's best transistors, which was possible because AT&T, the owner of Bell Labs and of the transistor patent, offered to license the device to other companies for \$25,000, a bargain for the most cutting-edge electronics technology. Shockley assumed that there'd be a market for transistors, at least for replacing vacuum tubes in existing electronics. The potential size of the transistor

market, though, was unclear. Everyone agreed transistors were a clever piece of technology based on the most advanced physics, but transistors would take off only if they did something better than vacuum tubes or could be produced more cheaply. Shockley would soon win the Nobel Prize for his theorizing about semiconductors, but the question of how to make transistors practical and useful was an engineering dilemma, not a matter of theoretical physics.

Transistors soon began to be used in place of vacuum tubes in computers, but the wiring between thousands of transistors created a jungle of complexity. Jack Kilby, an engineer at Texas Instruments, spent the summer of 1958 in his Texas lab fixated on finding a way to simplify the complexity created by all the wires that systems with transistors required. Kilby was soft-spoken, collegial, curious, and quietly brilliant. “He was never demanding,” one colleague remembered. “You knew what he wanted to have happen and you tried your darndest to make it happen.” Another colleague, who relished regular barbecue lunches with Kilby, said he was “as sweet a guy as you’d ever want to meet.”

Kilby was one of the first people outside Bell Labs to use a transistor, after his first employer, Milwaukee-based Centralab, licensed the technology from AT&T. In 1958, Kilby left Centralab to work in the transistor unit of Texas Instruments. Based in Dallas, TI had been founded to produce equipment using seismic waves to help oilmen decide where to drill. During World War II, the company had been drafted by the U.S. Navy to build sonar devices to track enemy submarines. After the war, TI executives realized this electronics expertise could be useful in other military systems, too, so they hired engineers like Kilby to build them.

Kilby arrived in Dallas around the company’s July holiday period, yet he’d accumulated no vacation time so he was left alone in the lab for a couple of weeks. With time to tinker, he wondered how to reduce the number of wires that were needed to string different transistors together. Rather than use a separate piece of silicon or germanium to build each transistor, he thought of assembling multiple components

on the same piece of semiconductor material. When his colleagues returned from summer vacation, they realized that Kilby's idea was revolutionary. Multiple transistors could be built into a single slab of silicon or germanium. Kilby called his invention an "integrated circuit," but it became known colloquially as a "chip," because each integrated circuit was made from a piece of silicon "chipped" off a circular silicon wafer.

About a year earlier, in Palo Alto, California, a group of eight engineers employed by William Shockley's semiconductor lab had told their Nobel Prize-winning boss that they were quitting. Shockley had a knack for spotting talent, but he was an awful manager. He thrived on controversy and created a toxic atmosphere that alienated the bright young engineers he'd assembled. So these eight engineers left Shockley Semiconductor and decided to found their own company, Fairchild Semiconductor, with seed funding from an East Coast millionaire.

The eight defectors from Shockley's lab are widely credited with founding Silicon Valley. One of the eight, Eugene Kleiner, would go on to found Kleiner Perkins, one of the world's most powerful venture capital firms. Gordon Moore, who went on to run Fairchild's R&D process, would later coin the concept of Moore's Law to describe the exponential growth in computing power. Most important was Bob Noyce, the leader of the "traitorous eight," who had a charismatic, visionary enthusiasm for microelectronics and an intuitive sense of which technical advances were needed to make transistors tiny, cheap, and reliable. Matching new inventions with commercial opportunities was exactly what a startup like Fairchild needed to succeed—and what the chip industry needed to take off.

By the time Fairchild was founded, the science of transistors was broadly clear, but manufacturing them reliably was an extraordinary challenge. The first commercialized transistors were made of a block of germanium with different materials layered on top in the shape of a mesa from the Arizona desert. These layers were fabricated by covering a portion of the germanium with a drop of black wax, using

a chemical to etch off the germanium that wasn't covered with wax, and then removing the wax, creating mesa shapes atop the germanium.

A downside of the mesa structure was that it allowed impurities like dust or other particles to become lodged on the transistor, reacting with the materials on the its surface. Noyce's colleague Jean Hoerni, a Swiss physicist and avid mountaineer, realized the mesas weren't necessary if the entire transistor could be built into, rather than on top of, the germanium. He devised a method of fabricating all the parts of a transistor by depositing a layer of protective silicon dioxide on top of a slab of silicon, then etching holes where needed and depositing additional materials. This method of depositing protective layers avoided exposing materials to air and impurities that could cause defects. It was a major advance in reliability.

Several months later, Noyce realized Hoerni's "planar method" could be used to produce multiple transistors on the same piece of silicon. Where Kilby, unbeknownst to Noyce, had produced a mesa transistor on a germanium base and then connected it with wires, Noyce used Hoerni's planar process to build multiple transistors on the same chip. Because the planar process covered the transistor with an insulating layer of silicon dioxide, Noyce could put "wires" directly on the chip by depositing lines of metal on top of it, conducting electricity between the chip's transistors. Like Kilby, Noyce had produced an integrated circuit: multiple electric components on a single piece of semiconductor material. However, Noyce's version had no freestanding wires at all. The transistors were built into a single block of material. Soon, the "integrated circuits" that Kilby and Noyce had developed would become known as "semiconductors" or, more simply, "chips."

Noyce, Moore, and their colleagues at Fairchild Semiconductor knew their integrated circuits would be vastly more reliable than the maze of wires that other electronic devices relied on. It seemed far easier to miniaturize Fairchild's "planar" design than standard mesa transistors. Smaller circuits, meanwhile, would require less electricity to work. Noyce and Moore began to realize that miniaturization and

electric efficiency were a powerful combination: smaller transistors and reduced power consumption would create new use cases for their integrated circuits. At the outset, however, Noyce's integrated circuit cost fifty times as much to make as a simpler device with separate components wired together. Everyone agreed Noyce's invention was clever, even brilliant. All it needed was a market.

CHAPTER 4

Liftoff

Three days after Noyce and Moore founded Fairchild Semiconductor, at 8:55 p.m., the answer to the question of who would pay for integrated circuits hurtled over their heads through California's nighttime sky. Sputnik, the world's first satellite, launched by the Soviet Union, orbited the earth from west to east at a speed of eighteen thousand miles per hour. "Russ 'Moon' Circling Globe," declared the headline in the *San Francisco Chronicle*, reflecting Americans' fears that this satellite gave the Russians a strategic advantage. Four years later, the Soviet Union followed Sputnik with another shock when cosmonaut Yuri Gagarin became the first person in space.

Across America, the Soviet space program caused a crisis of confidence. Control of the cosmos would have serious military ramifications. The U.S. thought it was the world's science superpower, but now it seemed to have fallen behind. Washington launched a crash program to catch up with the Soviets' rocket and missile programs, and President John F. Kennedy declared the U.S. would send a man to the moon. Bob Noyce suddenly had a market for his integrated circuits: rockets.

The first big order for Noyce's chips came from NASA, which in the 1960s had a vast budget to send astronauts to the moon. As America set its sights on a lunar landing, engineers at the MIT Instrumentation Lab were tasked by NASA to design the guidance computer for

the Apollo spacecraft, a device that was certain to be one of the most complicated computers ever made. Everyone agreed transistor-based computers were far better than the vacuum-tube equivalents that had cracked codes and calculated artillery trajectories during World War II. But could any of these devices really guide a spacecraft to the moon? One MIT engineer calculated that to meet the needs of the Apollo mission, a computer would need to be the size of a refrigerator and would consume more electricity than the entire Apollo spacecraft was expected to produce.

MIT's Instrumentation Lab had received its first integrated circuit, produced by Texas Instruments, in 1959, just a year after Jack Kilby had invented it, buying sixty-four of these chips for a price of \$1,000 to test them as part of a U.S. Navy missile program. The MIT team ended up not using chips in that missile but found the idea of integrated circuits intriguing. Around the same time, Fairchild began marketing its own "Micrologic" chips. "Go out and buy large quantities of those things," one MIT engineer ordered a colleague in January 1962, "to see if they are real."

Fairchild was a brand-new company, run by a group of thirty-year-old engineers with no track record, but their chips were reliable and arrived on time. By November 1962, Charles Stark Draper, the famed engineer who ran the MIT lab, had decided to bet on Fairchild chips for the Apollo program, calculating that a computer using Noyce's integrated circuits would be one-third smaller and lighter than a computer based on discrete transistors. It would use less electricity, too. The computer that eventually took Apollo 11 to the moon weighed seventy pounds and took up about one cubic foot of space, a thousand times less than the University of Pennsylvania's ENIAC computer that had calculated artillery trajectories during World War II.

MIT considered the Apollo guidance computer one of its proudest accomplishments, but Bob Noyce knew that it was his chips that made the Apollo computer tick. By 1964, Noyce bragged, the integrated circuits in Apollo computers had run for 19 million hours with only two