Continued from page 23

Today most semiconductor firms have gone "fabless", meaning that they design the chips but contract out production to a foundry. Many prominent names, from US graphics card specialist Nvidia to communications giants Broadcom and Qualcomm do not make their own chips. While foundries are increasingly concentrated, design is competitive. AMD and Nvidia vie for graphics card supremacy, while Qualcomm and Taiwan's Mediatek square off for dominance of the mobile market.

The shift towards "fabless" production has been driven by steeply rising costs. Cramming ever more transistors onto a chip approaches the limits of physics. That doesn't come cheap. Capital expenditure in the industry more than doubled in the decade to 2020 at $109bn, according to data from Statista. TSMC's latest leading edge fab will cost more than $19bn to build. Unable to keep pace, most firms have dropped out of the foundry game.

A giant comes unstuck

Only a handful of firms still operate the traditional integrated model, designing and fabricating their own chips. The most prominent is Samsung, but it runs the foundry business at arm's length from its other operations. Others are niche players, such as analogue specialist Texas Instruments. Intel, the firm that created the microprocessor, has also stuck with the model, but that has proved a failing strategy. While TSMC and Samsung march ahead with 5nm chips, Intel, long the industry leader, has struggled to keep up. It delivered its 10nm chip five years late and won't release a 7nm chip until 2023. That has forced it to outsource some chip manufacturing to its Asian rivals.

The company was caught flat-footed by the rise of smartphones, which use different designs to the ones with which the company had long dominated the PC market. Now even Intel's position there is coming into question, as Apple ditches the firm's chips in favour of its own designs for Macs (these chips will be made by TSMC). The share price has fallen more than 16% over the past year, even as the wider chip industry soars. New CEO Pat Gelsinger has launched a turnaround, with plans to let the design and fabrication arms operate more independently. That is more in line with how Samsung does things. Yet rather than doing that as a first step to spinning off the foundries, Gelsinger is doubling down on them.

ASML's $150m extreme ultraviolet (EUV) machine

Intel has been splashing the cash, announcing new multi-billion-dollar developments in Arizona and Ohio. Flush with subsidies from Brussels, it plans to spend $95bn on building several European plants over the next decade. Gelsinger claims to "have more concrete trucks working for me today than any other human on the planet". There are plenty of sceptics: Intel will struggle to match TSMC's vast capital outlays. The company has a long track record of overpromising and underdelivering on new chips. But Gelsinger is nothing if not confident. He wants to regain "unquestioned leadership" in semiconductors by 2025. His secret weapon? A chip with transistors so minute that they measure just 20 angstroms across – the next unit down beneath the nanometre.

"As TSMC and Samsung march ahead, Intel has fallen behind"

Silicon wafers need to be handled in a dust-free environment

anyone – cannot respond immediately to higher demand. It takes about two years and several billion dollars to construct a new fabrication plant ("fab"). By the time new capacity comes online, it is often too late.

The latest crisis has triggered a rush of new investment. Up to 29 new fabs are expected to be built worldwide during the two years to the end of 2022, says a report from Deloitte. That could take global manufacturing capacity up by 36% compared to the start of 2020. TSMC raised capital expenditure 66% on the year in 2021 to $30bn. Industry analysts think global chip supplies will remain tight at least for the first half of this year, but come 2023 we could be heading for a chip glut. While supplies of leading 3nm, 5nm and 7nm should remain fairly tight, less advanced chipmakers could be heading for trouble.

Much of the new capacity is in the 22nm to 90nm range, says Nikkei Asia. There are signs that the chip cycle has already turned for memory chips, with prices falling during the four months to November last year. But others argue that the industry stands to benefit from major structural tailwinds. The rollout of 5G, electric vehicles and smart devices in homes and factories will see demand for all types of chips soar over the next decade. The rise of cloud computing and artificial intelligence are other big opportunities for semiconductors. "It took 50 years for the industry to grow to half a trillion dollars today, and now it is estimated that the industry will grow to a trillion in roughly eight years", said Tom Caulfield, CEO of GlobalFoundries, the world's fourth-biggest foundry, as work on its new Singapore fab began last year.

It remains to be seen whether Caulfield is right. Alarmed about their dependence on distant suppliers, governments have been lavishing subsidies on the industry. America is spending $52bn to encourage more domestic semiconductor manufacturing, with Intel and TSMC planning to open new US plants. The EU is pushing for "digital sovereignty", with a plan to double the continent's share of the business by the end of the decade. China wants to be 70% self-sufficient in chips by 2025. That will require massive investment. The country only sourced 16% of the chips it needed domestically in 2020, according to IC Insights. Surprisingly few chips are "Made in China".

Growing subsidies and protectionism are likely to distort the market. Today's semiconductor supply chain is highly specialised, says Ma Tieying of DBS in a recent note. "If many countries attempt to establish their domestic semiconductor supply chain encompassing all segments, it could result in lower efficiency and higher costs in the long run."

Absolutely fabless

Semiconductor manufacturing is highly concentrated. Roughly 80% of chips are made in northeast Asia. TSMC alone accounts for more than half of chip manufacturing, followed by Samsung's foundry division with 17%. The likes of Taiwan's UMC, GlobalFoundries (an American firm with Singaporean roots) and China's SMIC specialise in producing less advanced "trailing-edge" chips.

Continued on page 24

"Supplies of advanced chips should remain tight, but other chipmakers could be heading for trouble"

The firms cashing in on the semiconductor shortage

Last year's chip crunch brought home how dependent the world is on these tiny pieces of silicon. Chipmakers are rushing to build new factories. Will they turn a shortfall into a glut? Alex Rankine reports

The world doesn't have enough semiconductors. Over the past year shortages of these tiny pieces of silicon have had big consequences. Car assembly lines have been halted for want of chips. Video gamers have found it almost impossible to acquire the latest consoles. Politicians have been fretting about national security: it hasn't escaped their notice that 92% of the world's most advanced chips are made by just one firm: Taiwan Semiconductor Manufacturing (TSM).

Meanwhile, investors have been in the money. The PHLX Semiconductor Index, which tracks the industry's leading lights, gained 43% last year. However, as firms rush to build new factories, some question whether the industry is about to go from shortages to an impending glut.

A technological marvel

Semiconductors play a central role in powering today's economy. Modern life would certainly be impossible without them. Most consumers think about chips, if at all, when acquiring a new laptop or a phone. A newer chip means better graphics and a snappier performance. But the sophisticated "system on a chip" that powers a smartphone represents only a fraction of the global market. Vast numbers of microcontrollers and sensors are needed to keep everything from 5G masts to fridges to electric toothbrushes running. Pets are microchipped in case they go missing. The global car industry has been hit hard by shortages because modern cars need upwards of 100 chips to power everything from the steering to the parking sensors. Such sensors play a growing role in "smart manufacturing" too – more data helps a firm to optimise its production lines.

Semiconductors are made from silicon, the second-most abundant element on earth. After being mined from silica-rich quartz, it is refined into a wafer, a flat, shiny disk that is about the size of a dinner plate. This is then shipped off to a fabrication plant. A speck of dust is enough to disrupt the delicate task of etching the circuits, so processing happens in a room that is kept far cleaner than an operating theatre.

An intricate, microscopic web of transistors is etched into the wafer by machines that seem to come straight out of science fiction. A semiconductor crams billions of transistors onto a slice of silicon about the size of a ten-pence piece. Apple's M1 Max chip, which is used in high-end laptops, has 57 billion of them (the human brain, for reference, hosts about 86 billion neurones). Each transistor is just nanometres in size, about 10,000 times thinner than a human hair.

The highest end chips are made using a lithography machine that "generates a specific wavelength of high-energy ultraviolet light by blasting molten drops of tin with a laser 50,000 times a second", says Clive Thompson in the MIT Technology Review. Such extreme ultraviolet (EUV) is "devilishly hard to manipulate". Wavelengths are so short that they are absorbed by air, meaning the whole thing must happen in a vacuum. Only one company in the world, the Dutch firm ASML, produces the machines. The "size of a small bus", each can cost up to $150m.

"Each transistor is 10,000 times smaller than a human hair"

A wafer goes through hundreds of processing steps as it is repeatedly heated, blasted with a laser and measured until the design in achieved. The wafers are tested (despite the precautions, a certain percentage are always defective and must be thrown away). Finally, the individual chips are cut out of the wafer and packaged ready for use.

The cutting edge of design

Greater numbers of transistors require ever more miniaturisation – smaller chips are more efficient and can pack more circuits onto a wafer, which reduces production cost. Chips are categorised by the nominal size of their transistors (although today this is more of a marketing label than an exact measurement). Higher numbers mean less advanced technology: 28nm chips were at the cutting edge in 2013, powering the latest Xbox. Today they are more likely to be found keeping your Wi-Fi running. Today's leading "process" is the 5nm chip. Just two companies – TSMC and Samsung in South Korea – are able to produce them.

Leading-edge chips attract all the attention, but they only account for about 11% of the output of contract chip makers like TSMC. Simpler memory chips, sensors and microcontrollers make up the bulk of the market by volume. These widgets have been running especially short of late, says Wired. A simple power-controlling integrated circuit, usually priced at $1, "can now sell for as much as $150", according to Josh Pucci of Sourceability. Lead times for simple components have gone from four to eight weeks up "to 24-52 weeks", according to IC Insights.

Even humble analogue chips are hard to find, says Yang Jie in The Wall Street Journal. Rather than using binary code, such chips "treat incoming information about temperature, sound and electrical current more like a human would, on a scale with many gradations." Such chips might be used to stop a battery overheating or power a phone display. Last year, laptop maker Asustek's assembly line suffered stoppages, apparently because of "a lack of analogue chips costing a few dollars apiece".

The chips are down

Before the pandemic, semiconductors were going through a rough patch. The last chip cycle had peaked in 2018, with sales falling back 12% in 2019, according to data from the Semiconductor Industry Association. The pandemic looked like another blow: carmakers cancelled their orders, anticipating a fall in vehicle demand. Instead, lockdowns drove an unprecedented surge in sales of consumer electronics and work-from-home office equipment. Sales were up an estimated 25.6% year-on-year in 2021, making semiconductors into a half-a-trillion-dollar industry.

Semiconductors are highly cyclical. On the demand side, when supply is tight many customers double orders to be sure they have enough supplies. That makes it difficult for chipmakers to gauge the true level of demand. "Phantom orders" stalk the nightmares of the industry's executives. On the supply side, contract foundries – firms that make chips for

Santa Clara

Nvidia's Arm deal in doubt: Nvidia, the most valuable chip company in the US, is quietly preparing to abandon its takeover of Cambridge-based Arm Holdings, says Tim Culpan on Bloomberg. Arm created the key architecture of many of the semiconductors used in almost every facet of the world economy – its technology forms the "very core" of processors made by Apple, Qualcomm and Samsung. So it's no surprise that when the $40bn cash and shares deal – the biggest chip deal in history – was announced in September 2020, it drew a backlash from regulators everywhere. Their belief that the merger would create a "virtually unstoppable powerhouse" in a world already starved of chips (see page 22) is not unreasonable. But it creates a headache for Arm's owner, Japanese conglomerate SoftBank. More than half of the proceeds were to come in Nvidia's stock, which, since the deal was struck, has soared 92% in value. The share portion, which makes up around half of the deal, is alone now worth $40bn, with the $17bn cash portion the icing on the cake. "SoftBank really needs the money." It spent $31bn taking Arm private in 2016 and it currently values the firm at $24bn. Even that "looks rich", while Softbank's balance sheet is "looking quite burdened". The ratio of current assets to current liabilities has fallen to 0.78, implying that "more money is due in the coming year than it has available". Softbank may pursue an initial public offering of Arm instead. "But it won't get nearly as much."