Analysis Paper

51/2021 09/12/2021

David Ramírez Morán

Squared millimeters of sovereignty

Visit the WEBSITE

Receive the E-NEWSLETTER

Squared millimeters of sovereignty

Abstract:

Microprocessors are at the heart of the information society and are the cornerstone of many of the processes that are part of its functioning. In a context where technology plays a central role, it is not uncommon for geopolitical clashes to occur over issues related to these small but highly capable devices. In the context of national security, a number of actions are being taken in this area. The vetoes on the sale of devices to certain countries and the confrontations that are taking place in the economic and political spheres require a certain degree of vigilance regarding the evolution of this technology, which is so fundamental to the day-to-day life of today's society. The recent COVID-19-driven disruption of the supply chain has raised concerns among companies, States and supranational organisations about effects that have transcended commercial boundaries.

Keywords:

Microprocessors, sovereignty, technology, globalization.

Milímetros cuadrados de soberanía

Resumen:

Los microprocesadores se encuentran detrás de la sociedad de la información y son piedra angular de muchos de los procesos que forman parte de su funcionamiento. En un contexto en el que la tecnología juega un papel primordial, no resulta raro que se produzcan enfrentamientos geopolíticos por cuestiones relacionadas con estos dispositivos de pequeño tamaño pero gran capacidad. Bajo el contexto de la seguridad nacional son diversas las acciones que se están llevando a cabo en este área. Los vetos a la venta de dispositivos a ciertos países y los enfrentamientos que se están produciendo en los ámbitos económico-políticos requieren mantener cierta alerta sobre la evolución de esta tecnología tan fundamental para el día a día de la sociedad actual. La reciente disrupción de la cadena de suministro motivada por la COVID-19 ha despertado las inquietudes de empresas, Estados y organizaciones supranacionales ante unos efectos que han trascendido los límites comerciales.

Palabras clave:

Microprocesadores, soberanía, tecnología, globalización.

How to cite this document:

RAMÍREZ MORÁN, David. *Square millimetres of sovereignty.* IEEE Analysis Paper 512021. https://www.ieee.es/Galerias/fichero/docs_analisis/2021/DIEEEA51_2021_DAVRAM_Milimetros_ENG.pdf and/or link bie[3] (accessed on day/month/year)

Introduction

15 November 2021 marked 50 years since the release of what is considered to be the first commercial microprocessor, the Intel 4004: 12 square millimetres, 2300 transistors, 1 core, 4 bits, 740 KHz, 10 µm technology and 92500 instructions per second.

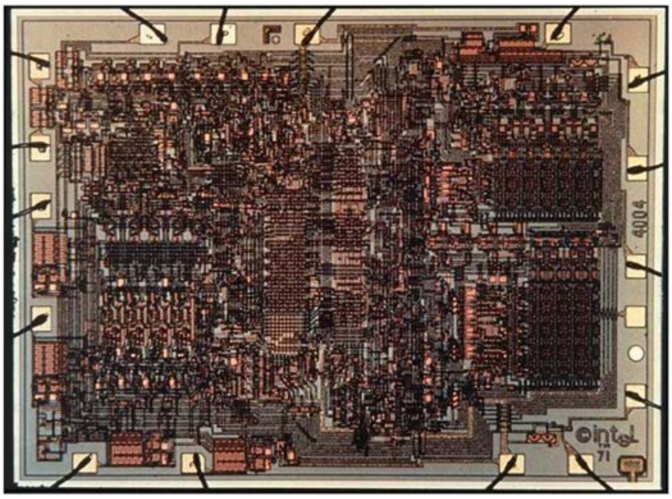


Figure 1 Picture of the 4004 chip (Source: intel.com)

This was the starting signal for a race in which Moore's Law, whereby every two years the number of transistors that could be manufactured in the same area would double, has served as a compass for the development of microprocessors up to the present day, far exceeding the 10-year deadline that the author had predicted. A line on a logarithmic scale that allows the exponential evolution of the complexity of devices to be represented in a comprehensible way.

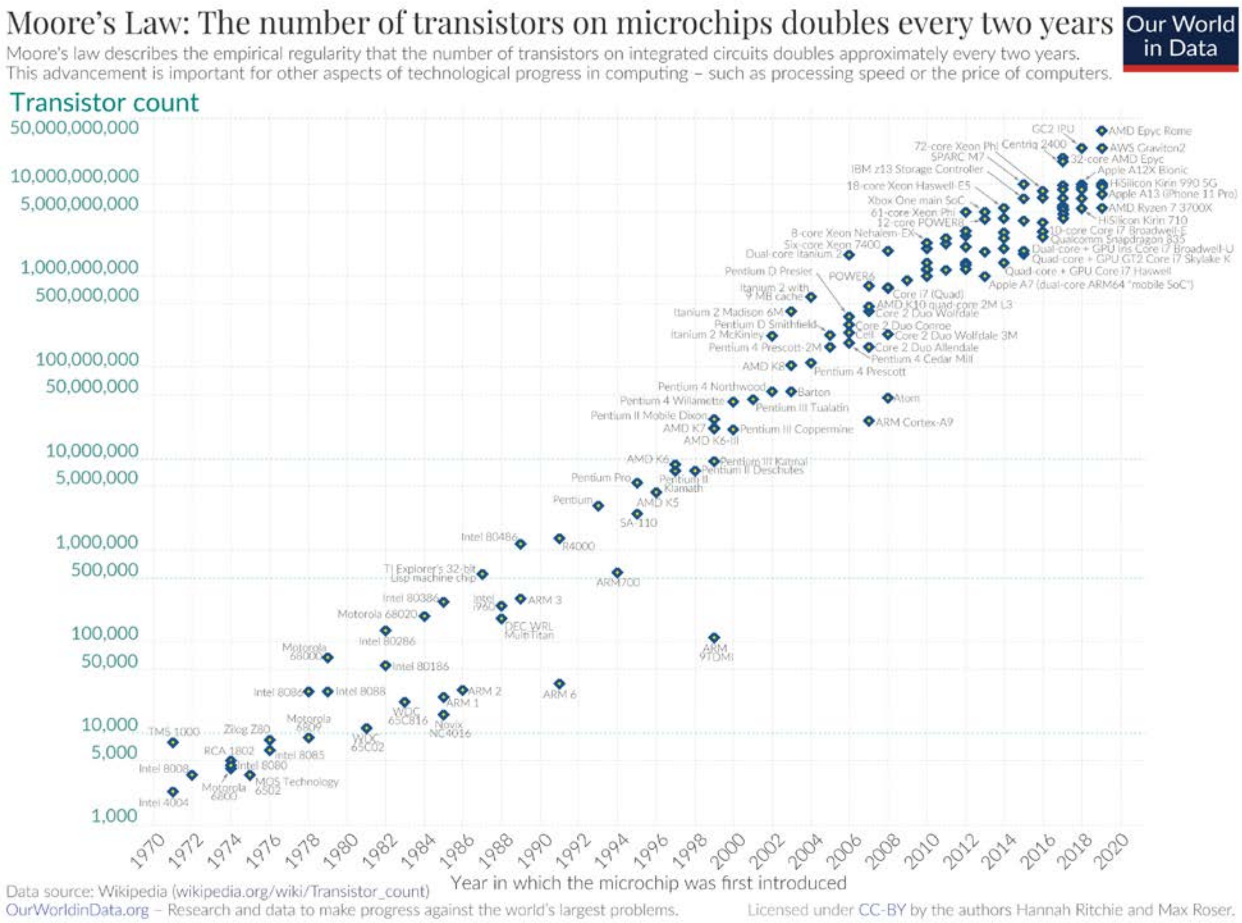


Figure 2 Evolution of microprocessors from 1971 to 2020

In 2020, each of the 8 modules of an AMD Epyc Rome features: 74 square millimetres, 3.8 billion transistors, 8 cores, 64 bits, 4 GHz, 7 nm technology and 1 billion instructions per second[1]. Almost 40 billion transistors in a single device that is installed in data centres where several thousand of these devices are interconnected and work in a coordinated way in parallel.

This is the infrastructure that supports the unstoppable advance of today's information and communications technologies. An infrastructure on which new paradigms such as artificial intelligence, cloud, edge computing, big data, high-performance computing and high-capacity ubiquitous communications are built.

A wide range of services run on this infrastructure with ever-increasing requirements and complexity, shaping the fourth industrial revolution: the information society. These are essential technologies for the day-to-day life of today's developed societies.

1 https://wccftech.com/amd-2nd-gen-epyc-rome-iod-ccd-chipshots-39-billion-transistors/

The digitisation of information is inherent to this new relationship model in which information needs to be collected and used at all scales, from the control of a household appliance to the integration of large volumes of data. For this reason, it is also necessary to equip "things" with the electronics they require to carry out these functions.

For example, in addition to its mechanics, a car requires numerous processors that handle everything from engine control and monitoring to music playback. Features such as anti-lock braking systems (ABS) or the multiple cameras and sensors in today's vehicles contribute to passenger safety and are being incorporated into regulated equipment for the marketing of a product. The digital devices needed to provide these functions are in most cases undemanding in terms of performance, but in many cases essential for the correct functioning of the complete system. Even semiconductors of modest complexity become critical elements in the supply chain, as demonstrated by the problems that automobile manufacturers are having on a near-global scale to maintain production.

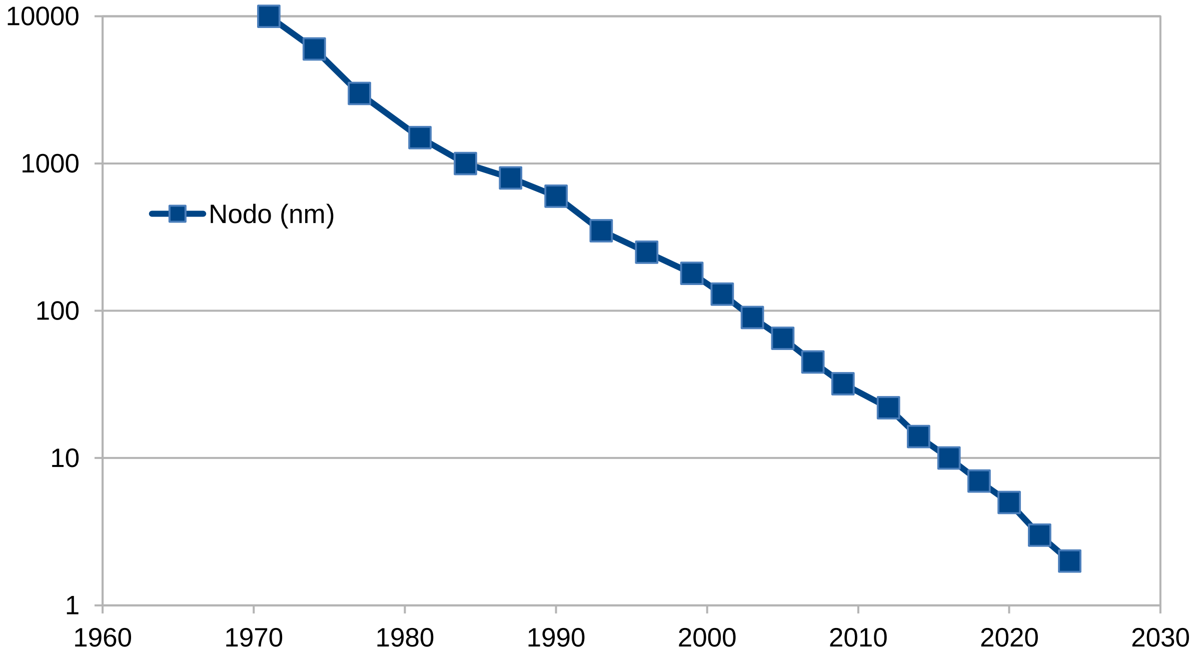
In the FMCG sector, mobile phone and PC manufacturers are also struggling to respond to business models and a level of demand that has changed in unforeseen ways.

The COVID-19 pandemic has affected supply chains and led to bottlenecks that have significantly affected markets and countries' economies.

These problems have triggered an avalanche of actions to address the current problems and to try to take the necessary measures to reduce the impact of similar scenarios in the future.

The nodes

Evolutions in digital device fabrication technologies are referred to as nodes, which describe the size in nanometres of the gate that controls the current flow of a FET transistor. The logarithmic scale of the graph indicates the exponential complexity of the devices and shows no signs of exhaustion. However, proven 3 nm and 2 nm technologies are starting to present problems due to the small size of the components, which are already atomic in scale and have sub-atomic effects that limit their performance. The most advanced node commercially available today is the 5 nm node from TSMC (Taiwan Semiconductor Manufacturing Company).



The smallest nodes are those that achieve the best power-to-capacity ratios and are therefore necessary for the development of compute-intensive systems or high-powered handheld devices. However, the remaining nodes are still used for more modestly performing devices where less integration is required or where power consumption is not such a critical factor.

In 2020, the European Commission noted that only two companies had the capacity to produce 7 nm devices: Samsung in South Korea and TSMC in Taiwan. It used this information and other data to justify the need for manufacturing capabilities for semiconductor devices, including microprocessors, in European territory.

Most of the high-performance devices marketed by Intel, the industry's largest company, are 14 nm and some 10 nm devices are beginning to be released.

GlobalFoundries had traditionally manufactured AMD devices but stayed with 14 nm technology. AMD is currently selling devices made with TSMC's 7nm technology.

One of the main suppliers of technology for the manufacture of integrated circuits for the most advanced nodes is ASML, a Dutch supplier to both TSMC and Samsung. In fact, it is almost the only supplier of EUV Extreme UltraViolet technology with 13.3-13.7 nm laser light.

The international high-performance microprocessor market

Microprocessors can be classified according to what is known as ISA, Instruction Set Architecture. A device is capable of executing a set of instructions defined by the ISA, and the implementation of devices to execute programs designed for that ISA is subject to intellectual property rights.

Intel architecture, both in its almost extinct 32-bit x32 version and the current 64-bit amd64 or x64 version, is one of the most widely used to date. There are two main manufacturers of devices with this architecture: Intel and AMD, companies that have their headquarters in the United States.

AMD and Intel have a slightly different business model in that, while Intel designs, produces and sells its product directly, AMD designs its devices and sends them to companies such as Global Foundries or TSMC for manufacturing. It therefore has no manufacturing capacity and relies on external companies to carry it out.

Intel devices are found in a large majority of data centres and also in a large proportion of desktop and notebook computers.

Intel incorporates integrated graphics cards (GPU Graphics Processing Unit) in its devices, which reduces the number of components required to build a workstation, although performance is limited and they are generally not used for intensive calculations beyond those required for the visual representation of screen content.

AMD, on the other hand, acquired the graphics card manufacturer ATI and, although it has produced some mixed microprocessor and graphics card devices in a single device, it generally uses external graphics cards. These graphics cards are being harnessed for artificial intelligence processing and other computationally intensive operations. Its architecture has many independent processing nodes and is much more efficient for performing certain operations than a general-purpose microprocessor.

ARM is an architecture that has slowly been gaining a foothold in the market, especially in the notebook market, due to its low power consumption thanks to its RISC architecture, compared to the CISC architecture of Intel and AMD. ARM devices are also starting to be used in data centres to take advantage of their lower power consumption and thus lower heat generation.

ARM's business model is based on royalties. The company designs the microprocessors and licenses these designs to the customer for production. The client company uses companies to implement the design on a chip and then it is sent to one of the available foundries to be manufactured. In addition to the licence fee, ARM's customer also pays royalties based on the number of units produced.

ARM has its headquarters in Cambridge in the United Kingdom. In September 2020, US graphics card maker Nvidia acquired a majority stake in ARM from Japanese conglomerate Softbank, which had acquired it in 2016 with funding from Saudi Arabia. This created a third player with the ability to create microprocessors and graphics cards. This acquisition still needs to be authorised by countries such as the United States, the European Union and, unsurprisingly, the United Kingdom[2]. This authorisation is necessary as the purchase could affect competition.

Against this backdrop of proprietary architectures, an initiative has emerged in which the ISA is open and does not require licensing fees or the purchase of devices from specific manufacturers. RISC-V is an ISA that users can implement freely. In a model similar to ARM, but without economic return, the design of the device is made available to the user. The user will have to realise this design by implementing it so that it can be manufactured in a foundry.

Moreover, as the information is published, US law does not allow export bans, so countries subject to these limitations may prefer this model of architecture[3]. The architecture is maintained by a non-profit foundation that has recently moved its headquarters to Switzerland to avoid US legislation.

When it was created and for many subsequent years, Apple initially used Motorola microprocessors and later PowerPCs; both had a different architecture to Intel. In 2006 it abandoned these architectures and switched to Intel microprocessors. In 2021, the first Apple device using an Apple-designed microprocessor with ARM architecture tailored to its needs will be commercially available. This brings it closer to its goal of standardising its mobile, laptop and desktop devices to reduce the variants of the software it develops.

2 A 'national security' issue: UK.gov blocks Nvidia's Arm deal for now, inserts deeper probe https://www.theregister.com/2021/11/16/nvidia/
3 https://www.economist.com/science-and-technology/2019/10/03/a-new-blueprint-for-microprocessors-challenges-the-industrys-giants

Supply chain issues

With the COVID-19 pandemic, there have been global semiconductor supply chain problems. There are several reasons behind these problems, and while some can be attributed to direct effects of the pandemic, others are the result of uncertainty and business decisions that have led to a domino effect.

Due to the pandemic, many countries confined their populations, resulting in entire production chains ceasing to function. They stopped foundries that generate the wafers, the high-purity semiconductor wafers on which integrated circuits are fabricated.

Integrated circuit production plants were also shut down. Device packaging was another activity that was also slowed down by the reduction of active staff in the sector. Recently, some stoppages have continued to occur due to rising raw material costs because this industry, and especially wafer foundries, is very energy intensive.

With the recovery of economies, problems have now also occurred in the logistical supply chains of goods, and the time it takes to move goods from one place to another is lengthening. Shipments of electronic components are taking longer and longer, further delaying the normalisation of supply chains.

The pandemic period also saw a change in consumer habits that have now become logistical problems. Faced with falling demand for new vehicles, manufacturers cancelled their orders for components and, instead of being stockpiled, they were simply discontinued due to lack of demand.

Demand shifted to electronic equipment, such as computers, tablets, telephones and connection equipment, in response to working from home. The shortage of materials also meant certain types of devices were prioritised for manufacture and certain components were discontinued in favour of others.

With the recovery, consumption habits have trended towards normality, but a new problem has emerged. Logistical operations during the pandemic have disrupted existing orchestration in the movement of ships, containers and goods and bottlenecks are occurring. This is leading to a considerable increase in transport costs, as they are being prioritised due to less usable capacity, and an increase in delivery times in the face of unforeseen events that may occur in a shipment.

Supply issues could have been limited if there had been greater geographical diversification of production sites. However, in the current situation it is not possible to provide a solution of this type, as the creation of new production infrastructures requires heavy investment and, more importantly, involves several years of construction time.

The automotive sector is one of the sectors where the effects of lower initial demand and the cancellation of orders for materials has had the greatest impact. Many plants around the world have had to stop production due to a lack of materials. The inability to produce has also led to shortages of finished products and a major market contraction despite latent demand. In Spain, a significant fraction of GDP comes from the sector and therefore an economic impact is to be expected.

The effects on companies have been diverse as illustrated in the examples below.

Ahead of the launch of the new iPhone 13 handset, Apple has encountered supply problems for some of the components and has suffered a considerable delay in terms of its original plans.

Google has also identified supply chain vulnerabilities and has proposed an initiative to standardise certain industry products to help mitigate risks[4]. In this case, it is focusing on the production chain for integrated circuits and is working to standardise intermediate products in order to reduce dependence on specific producers as much as possible. In fact, it proposes defining common, non-mandatory standards to leave open the option of being able to develop its own technology.

In the United States, vehicle manufacturer Ford is considering a new supply chain for its electronic component needs. In talks with Global Foundries to "develop a *collaborative model to accelerate the new wave in automotive chip design*".[5]

In India, vehicle manufacturer Tata has also been affected by the chip shortage. It is currently taking steps towards electronic component production and foresees investment of $300 million[6] to set up a device assembly and test factory, which could serve both the conglomerate itself and other companies such as Intel, AMD or ST Microelectronics.

4 https://www.theregister.com/2021/11/15/usa_100_day_semiconductor_supply_chain_review/
5 https://www.cnet.com/roadshow/news/ford-semiconductor-chips-shortage/
6 https://indianexpress.com/article/explained/tata-group-semiconductor-manufacturing-computer-chip-shortage-tejas-networks-7450091/

Supply chains are also being disrupted in the communications field and, for example, companies such as Cisco, which provides professional networking equipment to companies in general, and telecommunications companies in particular, have extended delivery times for their equipment due to a lack of materials to complete their products.

This observed vulnerability has led to institutional measures being taken to deal with future scenarios. The United States has introduced the CHIPS for America Act, while the European Union is also talking about an equivalent act[7].

United States

Semiconductor production in the US has progressed from 37% in 1990 to 12% in 2020. In November 2020, the H.R. proposal was presented in the House of Representatives. 7178

> "To restore American leadership in semiconductor production by increasing federal incentives to enable advanced research and development, secure the supply chain, and ensure long-term national security and economic competitiveness."[8]

Called the CHIPS for America Act (Creating Helpful Incentives to Produce Semiconductors). The FABS Act accompanied this law and was intended to provide tax credits to encourage private investment.

Investments planned as a result of this law:

- Establish a national semiconductor technology centre with a contribution of three billion euros from 2021 to 2030.
- Make two billion dollars available to DARPA in 2021 until 2025.
- Two billion dollars for basic semiconductor research programmes at Department of Energy
- and five billion dollars to set up the national institute for advanced packaging manufacturing.

7 BRETON, Thierry. *How a European Chips Act will put Europe back in the tech race* https://ec.europa.eu/commission/commissioners/2019-2024/breton/blog/how-european-chips-act-will-put-europe-back-tech-race_en
8 https://www.congress.gov/bill/116th-congress/house-bill/7178

There is a highly significant, direct reference to the Department of Defence: "The Department of Defense shall prioritize the use of specified available amounts for programs, projects, and activities in connection with semiconductor and related technologies" with a minimum annual amount of 50 million dollars.

European Union

On 15 September 2021, the European Commission published a working document to accompany the Proposal for a Decision of the European Parliament and the Council document setting out the 2030 policy agenda "Road to the Digital Decade"[9] as an action of the "2030 Digital Compass: the European way for the Digital Decade" communication previously published on 9 March.

> "*The European Union's ambition is to be digitally sovereign in an open and interconnected world, and to pursue digital policies that empower people and businesses to achieve a digital future that is people-centred, inclusive, sustainable and more prosperous.*"

The document sets out objectives to be achieved to this end in the following areas:

- A digitally empowered population and highly skilled digital professionals.
- Secure and powerful sustainable digital infrastructures
- Digital transformation of companies
- Digitisation of public services

Within the infrastructure area, it includes a specific point for semiconductors in which a target is set whereby "the production of sustainable, state-of-the-art semiconductors in Europe including processors is at least 20% of world production by value". A baseline value of only 10% was identified in 2020.

Semiconductor geopolitics

In 2015, the United States vetoed the sale of what were then the most advanced processors for high-performance computing. China's intention to create the most powerful

9 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0574

supercomputer was being thwarted because the US believed the devices would be used for the development of nuclear weapons.[10]

In 2020, President Trump added twelve Chinese companies, including SMIC, to the Entity List[11], making the export of technology with the ability to manufacture semiconductors from 10nm nodes and below subject to licensing by the US government. This closed the door to China being able to develop this type of technology at home using tools provided by other countries. This would have allowed it to take the technological edge far more quickly, given that the infrastructures it relies on (those of SMIC) are several nodes behind the technologies available on the market. They were added to the list on the grounds that the technology developed could be used to support Chinese military activity, posing a risk to US national security.

Conclusions

Semiconductors present a very diverse vertical market in which very few players wield enormous power. Just a few players provide each service at each step in the market. There are, in practice, three commercial manufacturers in architecture design with the capacity to meet today's needs, and all are based in the same country. An open initiative is also emerging, although certain applications, such as defence applications where specific certifications are required, aren't expected to be able to make use of these alternatives. On the production side, there is also a high concentration of suppliers, especially for higher performance devices needed to cope with the growing demand for processing in a context where energy efficiency is crucial to meet emission targets.

Both the US and Europe have identified weaknesses in microprocessor supply chains and are taking specific measures in an area of growing dependence. However, this isn't something new: in 2018 a similar initiative was taken, in this case within the preparatory action on another component where the supply chain was of concern, FPGA programmable logic devices, which presented a model quite similar to that of microprocessors[12] for the military field:

10 Ramírez Morán, David. "Is computing a geopolitical tool?" https://www.ieee.es/Galerias/fichero/docs_analisis/2015/DIEEEA43-2015_Supercomputacion_DRM.pdf
11 https://www.washingtonpost.com/technology/2020/12/18/china-smic-entity-list-ban/
12 https://defaiya.com/news/International%20News/North%20America/2020/11/25/mbda-to-contribute-to-the-foundation-of-a-european-source-of-field-programmable-gate-arrays

Early in November, the European Defence Agency (EDA) announced the implementation of EXCEED, the last of the three defence research projects selected under the 2018 call for proposals for the EU Preparatory Action on Defence Research (PADR).

EXCEED, which stands for 'trustEd and fleXible system-on-Chip for EuropEan Defence applications', aims to create a European supply chain for a reconfigurable, flexible and trustable programmable system-on-a-chip family.

As of today, most of the Field Programmable Gate Array (FPGA) used by European industries are American products manufactured on Asian industrial lines. The lack of a sovereign European solution on the market entails a dependency to ITAR or ITAR-like policies and represents a risk in terms of security and restrictions of use. Considering the strategic nature of FPGA, last week the EDA signed a grant agreement worth €12 million ...

Trends of recent decades, whereby production is becoming geographically concentrated, have been one of the main causes of the shortages of certain materials that are affecting the recovery of markets after the pandemic.

In today's knowledge society, where data is the primary raw material, sovereignty over the tools that make it possible to operate with this raw material can easily be undermined because there are many points at which control over the production, availability and security of goods can be lost.

Several actors have seen their sovereignty affected because they have not been able to make decisions on a globalised supply chain that is located, to a large extent, beyond their borders. For these reasons, important and controversial decisions are being taken, such as the financing of private activities with public funds, in order to avoid the recurrence of such problems in future situations.

The effects that the COVID-19 pandemic has had on supply chains can be identified as a national security risk. Events such as the need to stop production lines due to a lack of supplies have a high economic impact and directly affect the state's financial accounts.

David Ramírez Morán
IEEE Senior Analyst
@darammor