elektor industry
innovate > launch > trade

Elektor Helps
Elektronik in Krisenzeiten

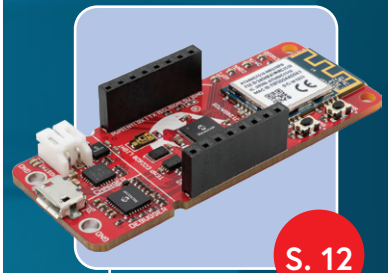
# EMBEDDED & TOOLS

embedded world 2021 Exhibition&Conference ...it's a smarter world DIGITAL

## embedded world DIGITAL Spezial

**MACHT TESTEN WENIGER SPASS ALS PROGRAMMIEREN?**
S. 37

**SECURE-BOOT-LÖSUNG FÜR RASPBERRY PI**
Viel Sicherheit zum vernünftigen Preis
S. 20


S. 12


S. 60


S. 27

**INTERVIEW**
Benedikt Weyerer, Leiter der embedded world, über digitale Messen
S. 10

**AVR-IoT- und PIC-IoT-Entwicklungsboards**
Sichere Kommunikation mit der AWS Cloud

**Aufstieg der kontaktlosen Manipulation**
Für empfindliche Komponenten

**Low-Power-Flaggschiff für Künstliche Intelligenz**
SMARC-Module mit NXP i.MX 8M Plus Prozessor

# Treten Sie jetzt der Elektor Community bei!

## Jetzt elektor GOLD membership Mitglied werden!


- Komplettes Webarchiv ab 1970
- 6x Elektor Doppelheft (Print)
- 9x Digital (PDF) inkl. Elektor Industry
- 10% Rabatt im Online-Shop und exklusive Angebote
- Elektor Jahrgangs-DVD
- Mit Tausenden von Mitgliedern des Online-Labors gemeinsam entwickeln mit Zugang zu über 1.000 Gerber-Dateien und direktem Kontakt zu unseren Experten!
- Veröffentlichen Sie Ihr eigenes Projekt oder verkaufen Sie direkt über unseren Shop!

---

## Auch erhältlich

Die digitale elektor GREEN membership Mitgliedschaft!

- Zugang zu unserem Webarchiv
- 10% Rabatt in unserem Online-Shop
- 6x Elektor Doppelheft (PDF)
- Exklusive Angebote
- Zugang zu über 1.000 Gerber-Dateien

**www.elektor.de/mitglied**

elektor design › share › sell

# Vorwort

Von **Jens Nickel**

*Chefredakteur*

## Digitale Chancen

Auch in diesem Vorwort komme ich nicht drum herum, die leidigen Viren zu erwähnen. Zu fest hat die Corona-Krise auch die Elektronikbranche noch im Griff. Und hier sind vor allem diejenigen betroffen, die in Unternehmen an vorderster Front stehen: Produktmanager, Marketing-Fachleute, Einkäufer. Und natürlich auch alle, die Plattformen für den Informationsaustauch bieten - zum Beispiel Messeveranstalter und auch wir, die Fachzeitschriften. Ein sichtbares Zeichen waren die geschlossenen Türen der Nürnberger Messe, auf denen normalerweise die *embedded world* stattgefunden hätte. Jedoch: Wie auch die Messe München bei der *electronica* hat auch die NürnbergMesse die Chance genutzt, die wichtige Veranstaltung ins Internet zu verlagern. Das Ganze ist nicht nur aus der Not geboren. Die virtuellen Events sind für die Messen auch ein Probeballon - die gesammelten Erfahrungen können helfen, neue Geschäftsfelder aufzubauen (siehe Seite 10). In absehbarer Zeit wird wohl nichts eine Veranstaltung ersetzen können, bei der man sich persönlich trifft. Doch muss sich beides ja auch nicht ausschließen. Eine digitale Messe zusätzlich zu einer „echten" Ausstellung hilft, schneller zu benötigten Infos zu kommen und das eigene Netzwerk mit den richtigen Ansprechpartnern zu ergänzen.

Beim Besuch der *embedded world DIGITAL* hat mein Kollege Mathias Claußen viele interessante Produkte entdeckt, Infos eingeholt und die Möglichkeit zum Video-Chat mit anderen Ingenieuren genutzt. Seinen (selbstverständlich ganz und gar nicht vollständigen) Überblick über neue Controller, Software-Tools und mehr finden Sie auf Seite 6. Der dort vorgestellte MAX78000 - ein SoC mit integriertem Arm Cortex-M4, einem RISC-V-Kern und einem KI-Beschleuniger - konnte den „embedded Award" der Messe gewinnen. Meine Kollegen aus dem Labor haben bereits mit dem MAX78000 gearbeitet; ein Design-Wettbewerb für interessierte Entwickler ist in Vorbereitung. Mehr darüber unter www.elektormagazine.de/ai-contest-max78000!

210204-02



WÜRTH ELEKTRONIK

**REDEXPERT.** Die einzigartige Online-Plattform von Würth Elektronik zur Auswahl elektronischer und elektromechanischer Bauelemente.
**www.we-online.de/redexpert**

- Weltweit genauestes AC-Verlustmodell für Speicherinduktivitäten
- Filtermöglichkeit für über 20 elektrische und mechanische Merkmale
- Simulation der Induktivität im DC/DC-Wandler
- Vergleichbarkeit anhand interaktiver Messkurven (Induktivität/Strom und Erwärmung/DC-Strom)
- Verfügbar in sieben Sprachen
- Messwertbasierte Online-Plattform
- Kein Login notwendig
- Integrierte kostenlose Musterbestellung
- Direkter Zugriff auf Produktdatenblatt

***WE* are here for you!**
Nehmen Sie teil an unseren kostenlosen Webinaren: **www.we-online.de/webinare**

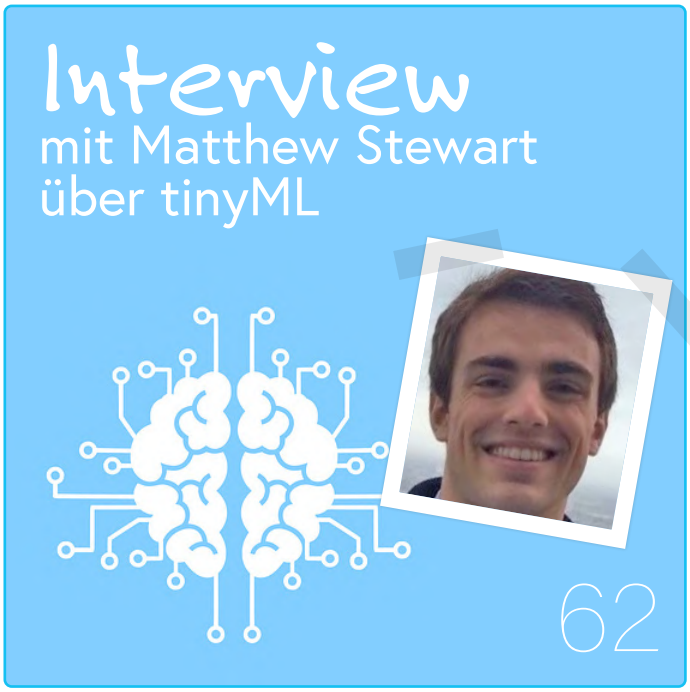
# Inhalt




## Rubriken



## Artikel





# Nächste Ausgabe

Die **Ausgabe 2/2021 des Magazins Elektor Industry** beschäftigt sich mit Sensoren sowie Equipment und Methoden für das Testen und Messen, mit Beiträgen von Unternehmen, Branchenspezialisten, Elektor-Redakteuren und freien Autoren.

Elektor Industry Ausgabe 2/2021 erscheint am 20. Mai 2021.

Änderungen vorbehalten.

# Impressum

**Elektor Industry**
Ausgabe 1/2021
Embedded & Tools
© 2021
www.elektormagazine.de

Elektor Industry, Deutsche Ausgabe, wird in 2021 drei Mal veröffentlicht.

**Hauptsitz des Verlages**
Elektor Verlag GmbH
Kackertstr. 10
52072 Aachen
Tel. +49 241 95509-190
Fax +49 241 95509-013

**Chefredakteur** (V.i.s.d.P.)
Jens Nickel
redaktion@elektor.de

**Redaktion**
C.J. Abate, Stuart Cording, Robert van der Zwan

**Anzeigen & Sponsoring**
Raoul Morreau
Tel. +31 (0)6 4403 9907
raoul.morreau@elektor.com

**Layout**
Harmen Heida, Giel Dols und Jack Jamar

**Herausgeber**
Don Akkermans

**Druck**
Pijper Media, Stettinweg 15, 9723 HD Groningen

**Aufgabe**
Elektor Industry bietet Elektronik-Ingenieuren, Innovatoren und Start-ups Informationen über und Einblicke in die neuesten Produkte, Techniken und Entwicklungen aus der Elektronik-Industrie.

**Distribution und Lieferung**
Elektor Industry erscheint zugleich auf Deutsch und Englisch. Die Ausgabe wird an Elektor Gold-Mitglieder als gedruckte Version und an Elektor Green-Mitglieder als pdf-Datei verschickt. Das Magazin wird auf Messen gratis verbreitet wie z.B. auf der Productronica (München), electronica (München), und embedded world (Nürnberg).

**Leserhinweis**
Elektor Industry enthält gesponserte Artikel.
Der Herausgeber bestätigt alle Trademarks in Verbindung mit Produkten, Diensten, Materialen und Firmennamen, die in dieser Veröffentlichung erscheinen. Die in der Elektor Industry vertretenen Ansichten stimmen nicht zwangsweise mit denjenigen des Herausgebers oder der Redaktion überein.

**Urheberrecht**
Die in dieser Zeitschrift veröffentlichten Beiträge, insbesondere alle Aufsätze und Artikel sowie alle Entwürfe, Pläne, Zeichnungen einschließlich Platinen sind urheberrechtlich geschützt. Ihre auch teilweise Vervielfältigung und Verbreitung ist grundsätzlich nur mit vorheriger schriftlicher Zustimmung des Herausgebers gestattet. Die veröffentlichten Schaltungen können unter Patent oder Gebrauchsmusterschutz stehen. Herstellen, Feilhalten, Inverkehrbringen und gewerblicher Gebrauch der Beiträge sind nur mit Zustimmung des Verlages und ggf. des Schutzrechtsinhabers zulässig. Nur der private Gebrauch ist frei. Bei den benutzten Warenbezeichnungen kann es sich um geschützte Warenzeichen handeln, die nur mit Zustimmung ihrer Inhaber warenzeichengemäß benutzt werden dürfen. Die geltenden gesetzlichen Bestimmungen hinsichtlich Bau, Erwerb und Betrieb von Sende- und Empfangseinrichtungen und der elektrischen Sicherheit sind unbedingt zu beachten. Eine Haftung des Herausgebers für die Richtigkeit und Brauchbarkeit der veröffentlichten Schaltungen und sonstigen Anordnungen sowie für die Richtigkeit des technischen Inhalts der veröffentlichten Aufsätze und sonstigen Beiträge ist ausgeschlossen.

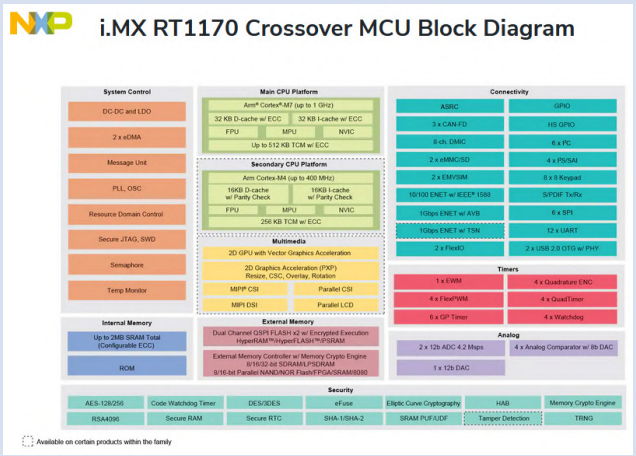
© Elektor International Media b.v. 2021
www.elektormagazine.de
Printed in the Netherlands

# embeddedworld2021 DIGITAL

## Ein kurzer Messeüberblick

Von **Mathias Claußen** (Elektor-Labor)

Letztes Jahr war die embedded world eine der letzten für internationale Aussteller und Besucher geöffneten Messen, bevor die COVID-Epidemie über den Globus fegte. In diesem Jahr wurde diese wichtige Messe komplett digital durchgeführt. Die Teilnehmer konnten über eine Online-Plattform teilnehmen, die für sie interessanten Aussteller aufsuchen und via Messages und Video mit Experten chatten. Kataloge, Broschüren und Produkt-Flyer standen per Download zur Verfügung. Selbstverständlich war das nicht das gleiche Erlebnis wie eine echte Messe, aber wer sich sehr für Mikrocontroller und Software begeistert (wie Mathias Claußen vom Elektor-Labor) konnte viele neue und interessante Produkte entdecken. Nachfolgend eine kleine Auswahl davon.




### MCUs mit 1 GHz

Mit dem i.MX RT1170 von **NXP Semiconductors** erreichen MCUs einen Takt im 1-GHz-Bereich. Die MCU bringt nicht nur einen Dual-Core Arm-Cortex-M (Cortex-M7 + Cortex-M4) und eine Fülle von Schnittstellen mit, sondern auch noch eine 2D-GPU mit OpenVG 1.1-Schnittstelle. Ein damit bestücktes Board Teensy 5.0 [1] wäre eine super Sache...

### AMD Ryzen V2000

**Congatec** hat seine auf Ryzen V2000-basierenden COM-Module angekündigt. Diese AMD-Prozessoren sind für Embedded-Systeme konzipiert und kombinieren eine hohe Rechenleistung mit einer Radeon-GPU in einem Chip. Durch die geringe Leistungsaufnahme eignen sich die Module gut für Systeme im Dauerbetrieb.

### MCUs mit ARM-Cortex-M33-Cores

Beim Cortex-M33 als Weiterentwicklung von Cortex-M4 handelt es sich um einen Armv8-M-Mainline-Core. Der STM32U5 von **STMicroelectronics** ist eine stromsparende, Cortex-M33-basierte MCU mit DSP-Befehlen und FPU. Die MCU ist in den Gehäusen LQFP48, QFN48, LQFP64, WLCSP90, LQFP100, UFBGA132, LQFP144 und UFBGA169 erhältlich.

Die Serie RA4M2 von **Renesas Electronics** umfasst die neueste Generation von 32-bit-MCUs mit Cortex-M33 und umfangreicher Peripherie wie USB 2.0 Full-Speed, SDHI, QSPI und erweiterten Analog-Funktionen.

**Key Features**

- 100MHz Arm® Cortex®-M33 with TrustZone®
- Secure element functionality
- 256kB - 512kB Flash memory and 64kB SRAM with Parity and 64kB SRAM with ECC
- 8kB Data Flash to store data as in EEPROM
- 1kB Stand-by SRAM
- Scalable from 48-pin to 100-pin packages
- Capacitive touch sensing unit
- USB 2.0 Full Speed
- CAN 2.0B
- QuadSPI
- SCI (UART, Simple SPI, Simple I²C)
- SPI/ I²C multimaster interface
- SDHI and MMC

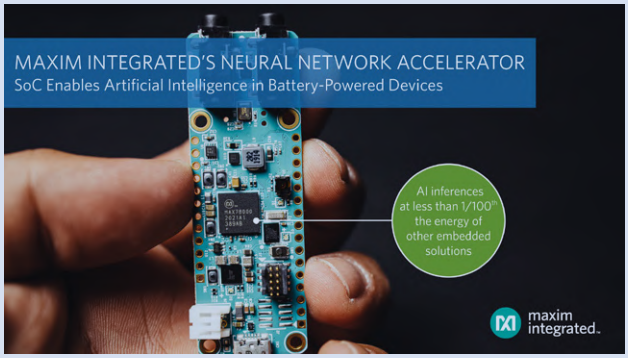

Das SoC nRF5340 von **Nordic Semiconductor** verfügt nicht nur über Bluetooth LE, Bluetooth Mesh, NFC sowie Thread- und ZigBee-Unterstützung, sondern auch über zwei Cortex-M33-Cores. Seine Peripherie macht es zur passenden MCU für IoT-Anwendungen etc. Die kleinere Variante nRF51822 steckt bereits im BBC micro:bit.

Das Modul NORA-B1 von **u-blox** basiert auf dem neuesten nRF5340-Chip von Nordic, der die neue Dual-Core-MCU ARM-Cortex-M33 enthält. Die NORA-B1-Serie eignet sich für Anwendungen mit hohem Leistungsbedarf.


Die Bluetooth-5.2-Low-Energy-NORA-B1-Module sind mit ARM-TrustZone und CryptoCell-312 für verbesserte Sicherheit ausgestattet und sind für industrielle Steuerungen, Asset-Tracking, Fernsteuerungen und Gateways, vernetzte Elektrowerkzeuge mit kontinuierlicher Motorsteuerung und medizinische Wearables gedacht.

## Low-Power-KI

Mit dem MAX78000 hat **Maxim** den „embedded award" der embedded world gewonnen. Der Chip enthält einen Beschleuniger für neuronale Netze und ist für mobile Anwendungen konzipiert. Dieses schönes Stück Hardware hat Clemens Valens in dem Artikel „KI-Peripherie: Erste Schritte mit dem MAX78000FTHR von Maxim Integrated" [2] vorstellt. Den MAX78000 kann man auf preiswerten Evaluation-Boards ausprobieren.

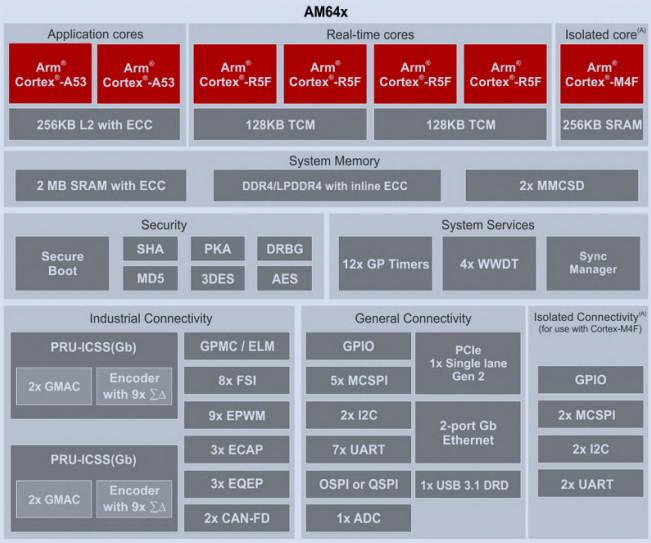



Der DS28E40 von **Maxim** ist ein 1-Wire-Authentifikator, der kryptographische Funktionen und Authentifizierung für 1-Wire-fähige Geräte bereitstellt. Der DS28E40 kann Geräte in vertrauenswürdige Hardware verwandeln.

## Energy Harvesting


**Renesas** hat seine Mikrocontroller-Familie um die RE-Familie erweitert. Die Ultra-Low-Power-MCU RE01 arbeitet bei niedrigen Spannungen bis zu 1,62 V mit 64 MHz Takt. Die hohe Leistung bei geringem Strombedarf ermöglicht längere Laufzeiten oder kleinere Akkus. Der On-Chip-Energy-Harvesting-Controller kann die Entwicklung eine Batterie etc. komplett überflüssig machen und so wartungsfreie Systeme ermöglichen. Die RE-Familie eignet sich für IoT-Anwendungen wie Hybrid-Uhren, intelligente Häuser/Gebäude, das Gesundheitswesen, intelligente Landwirtschaft, Strukturüberwachung und Tracker.

## Industrielle Anwendungen und Interfaces

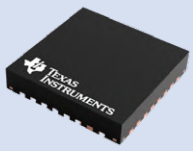



Der AM6442 von **Texas Instruments** ist für Motorantriebe und speicherprogrammierbare Steuerungen (SPS) konzipiert. Er bietet eine spezielle Kombination aus Echtzeitverarbeitung und Kommunikation. Der Chip kombiniert zwei Instanzen der Gigabit-TSN-fähigen PRU-ICSSG (programmierbare Echtzeiteinheit mit Gigabit-Industriekommunikations-Subsystem) von Sitara mit bis zu zwei ARM-Cortex-A53-Cores für Web-Services, bis zu vier Cortex-R5F-MCUs für Echtzeitberechnungen und einer Cortex-M4F-MCU.

## Long-Range-Kommunikation


Die STM32WL-MCU-Reihe von **STMicroelectronics** integriert LoRa, SigFox und Sub-1-GHz-Verbindungen mit einer Single- oder Dual-Core-Cortex-M-CPU bei niedrigsten Stromverbrauch.

Der 10BASE-T1L Single-Pair-Ethernet-PHY von **Texas Instruments** ermöglicht 10-Mbit/s-Ethernet-Verbindungen mit bis zu 2 km Leitungslänge. Der DP83TD510E erlaubt Kommunikationsnetzwerke vom Controller bis zum Edge Node, die Full-Duplex-Daten über ein einziges Paar verdrillter Adern übertragen. Durch den Wegfall von zusätzlichen Protokollen, Gateways und Kabeln lässt sich das



Netzwerkmanagement vereinfachen und gleichzeitig die Systemsteuerung und Interoperabilität bei großen Distanzen verbessern.

## Lower-Power-Positionsbestimmung


Das Modul MAX-M10S von **u-blox** basiert auf der Low-Power-GNSS-Plattform M10, die eine besondere Empfindlichkeit und Erfassungszeit für L1-GNSS-Systeme bietet. Der Stromverbrauch von weniger als 25 mW im kontinuierlichen Tracking-Modus ist ideal für batteriebetriebene Geräte wie Asset-Tracker.


Der intelligente Sensor BHI260AP von **Bosch** ist All-in-One-Lösung für den Dauerbetrieb, wie etwa Fitness-Tracking, Positionsbestimmung, maschinelle Analysen und Orientierungssysteme. Der Chip bietet viele Software-Funktionen, eine programmierbare 32-bit-MCU und eine 6-Achsen-IMU in einem Gehäuse.

## Flexible Hardware


**GOWIN** ist ein neuerer FPGA-Anbieter mit interessanten Chips. Im Angebot sind u.a. auch Chips mit fest verdrahteten Cortex-M-Cores.

## Runde Displays


**RockTouch** stellt sein erstes rundes Touch-LCD für die Outdoor- und Automobilanwendung vor. Mit einem 3 mm starken Deckglas kann es bei Temperaturen von -30° bis +85°C betrieben werden. Seine Auflösung von 1540×1540 Pixel bei 9,5" bietet neue Designoptionen.

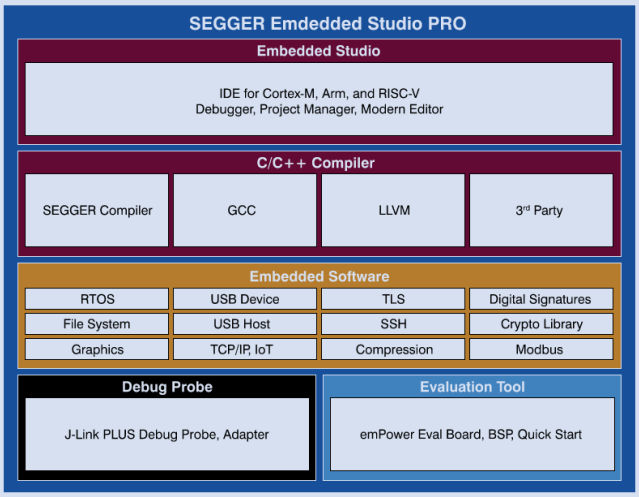
## Software- und Entwicklungs-Tools

**Green Hills** bietet Produkte rund um RISC-V an. Ihre Tools und Compiler unterstützen nun die 32- und 64-bit-Version von RISC-V. Besonders interessant an RISC-V ist die Unterstützung von kundenspezifischen Instruktionen. Die Tools von Green Hills berücksichtigen dies und bieten eine einfach zu erlernende Benutzeroberfläche zum Erstellen neuer Instruktionen für den Compiler.

Die Embedded Workbench von **IAR Systems** unterstützt nun ebenfalls RISC-V. Dabei werden die 32-bit-RISC-V-Cores und -Erweiterungen RV32 und RV32E einschließlich der RISC-V-P-Erweiterung für Packed-SIMD unterstützt. Benutzerdefinierte Erweiterungen helfen beim Entwurf benutzerdefinierter Cores mit exakten Definitionen für eine Anwendung oder ein Produkt.

Auch auf der Messe zu sehen waren die Debugging- und Trace-Pro-

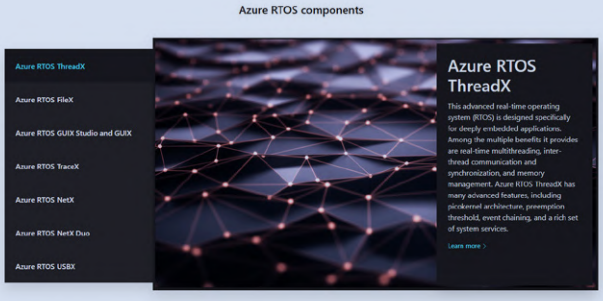
bes I-jet und I-jet-trace für Arm und RISC-V von **IAR Systems**, die leistungsstarke Funktionen in IAR Embedded Workbench ermöglichen.




Beim Embedded Studio von **SEGGER** handelt es sich um eine plattformübergreifende IDE für die Embedded-C-Programmierung, die u.a. einen C/C++-Compiler samt integriertem Debugger und direkter J-Link-Integration bietet. Alle diese Tools wurden speziell für Embedded-Systeme entwickelt und generieren extrem kompakten Code, um das Maximum aus den MCUs herauszuholen. Als All-in-One-Lösung bietet Embedded Studio alles, was für die professionelle Embedded-Entwicklung benötigt wird. Sie ermöglicht durchgängige Arbeitsabläufe von der Verwaltung über Build, Test bis hin zum Deployment von Embedded-Anwendungen. Der an Visual Studio angelehnte Stil bietet PC-Entwicklern eine sehr vertraute intuitive Bedienung. Für nicht-kommerziellen Einsatz und die Verwendung mit lizenzierten Geräten ist Embedded Studio kostenlos.

## Mehr RTOS-Auswahl


Das Projekt **Zephyr** dreht sich um ein skalierbares Echtzeitbetriebssystem (RTOS) mit Unterstützung mehrerer Hardware-Architekturen. Seine Open-Source-Lizenz nach Apache 2.0 erlaubt den Einsatz in kommerziellen und nicht-kommerziellen Anwendungen. Zephyr basiert auf einem Small-Footprint-Kernel, der für den Einsatz auf ressourcenbeschränkten Systemen entwickelt wurde. Es zielt auf einfache Embedded-Umweltsensoren und LED-Wearables bis hin zu anspruchsvollen Smartwatches und IoT-Funkgateways ab. Der Zephyr-Kernel unterstützt mit ARM-Cortex-M, Cortex-R, Cortex-A, Intel x86, ARC, MIPS Nios II, Tensilica Xtensa und RISC-V (32 und 64 Bit) viele Architekturen sowie etliche Boards.




Das RTOS Azure von **Microsoft** ist eine Embedded-Entwicklungssuite mit einem kleinen, leistungsstarken Betriebssystem, das schnelle und zuverlässige Lösungen für Geräte mit beschränkten Ressourcen erlaubt. Es wird bereits auf mehr als 6,2 Milliarden Geräten weltweit eingesetzt wird. Azure unterstützt die gängigsten 32-bit-MCUs und Embedded-Entwicklungstools.

Falls Sie sich fragen, warum Azure RTOS auch unter ThreadX bekannt ist: Vielleicht sagt Ihnen die Firma „Express Logic“ etwas, die einst ein RTOS namens ThreadX anbot. Um genau dieses RTOS handelt es sich. Express Logic wurde nämlich Ende 2019 von Microsoft übernommen und das Produkt dann in Azure RTOS ThreadX umbenannt. Der Code ist offen und kann auf GitHub [3] gefunden werden. Zum Testen kann das RTOS kostenlos genutzt werden. Bei der Nutzung lizenzierter MCUs (eine Liste findet sich unter [4]) fallen auch für die kommerzielle Nutzung keine Lizenzgebühren an.

210193-02


## WEBLINKS


[1] **Diskussion über Teensy und das Thema i.MX RT1170**: https://forum.pjrc.com/threads/57842-Future-Teensy-features-amp-pinout

[2] **C. Valens, „KI-Peripherie: Erste Schritte mit dem MAX78000FTHR von Maxim Integrated“, 5.3.2021, Elektor**: www.elektormagazine.de/articles/ki-peripherie-erste-schritte-mit-dem-max78000fthr-von-maxim-integrated

[3] **Microsoft Azure RTOS auf GitHub**: https://github.com/azure-rtos/threadx/

[4] **Microsoft Azure RTOS lizenzierte MCUs**: https://github.com/azure-rtos/threadx/blob/master/LICENSED-HARDWARE.txt

Q&A mit Benedikt Weyerer

(Director Exhibitions der embedded world, NürnbergMesse)

# Eine neue Welt, ein neuer Weg: **Die embedded world 2021 fand digital statt**

Fragen von **Shenja Panik** (Elektor)

Die embedded world, eine bekannte und beliebte Messe für Embedded-Profis, findet traditionell Anfang März statt. In diesem Jahr wurde die Veranstaltung komplett ins Internet verlegt. Elektor befragte Benedikt Weyerer, Leiter der embedded world, zu den Hintergründen, den Ergebnissen und der Zukunft der Veranstaltung.

**Elektor:** Bitte erzählen Sie uns etwas über die diesjährige embedded world.

**Weyerer:** Mit der *embedded world 2021 DIGITAL* erhielt die Embedded-Community ihre eigene Plattform zur Vernetzung. In diesem Jahr fand sie als rein digitale Veranstaltung statt und trat anstelle der Präsenzveranstaltung in Nürnberg. Die Veranstaltung dauerte statt der gewohnten drei Tage fünf Tage.

Gemeinsam mit der embedded world Conference DIGITAL und der electronic displays Conference DIGITAL bot die digitale Veranstaltung alle Themen mit Bezug auf die Entwicklung von eingebetteten Systemen an — von Bauelementen, Modulen und Komplettsystemen bis hin zu Betriebssystemen, Hardware, Software und Dienstleistungen. Über 300 Aussteller, sowohl führende Branchengrößen als auch innovative Neuunternehmen, waren auf der digitalen Plattform vertreten, wo sich die digitale Ausstellung und Konferenzen der Spitzenklasse perfekt ergänzten.

Teilnehmern wurden nicht nur Unternehmens- und Produktpräsentationen, sondern auch Vorträge und Anwendungsbeispiele auf dem Ausstellerforum geboten. Höhepunkte der Veranstaltung waren u.a. Fachgremien zu den Themen „Embedded Artificial Intelligence", „Safety and Security", „Embedded Vision", „Connectivity in IoT" und „Diversität als Erfolgsfaktor für Innovationen". Die zugehörigen Konferenzen zeichneten sich durch ein erstklassiges Vortragsprogramm, spannende Keynotes und namhafte Sprecher aus.

**Elektor:** Wie reagierte die Branche denn auf den Umstieg von live auf digital?

**Weyerer:** Das Feedback, das wir in zahllosen Gesprächen mit Kunden erhielten, ist, dass der Markt die embedded world braucht — vor allem zu Zeiten der Pandemie. Durch das digitale Format ergaben sich neue Möglichkeiten, von denen Aussteller und Teilnehmer profitierten. Für Teilnehmer aus den USA und China waren zum Beispiel verschobene digitale Termine vorgesehen, um die unterschiedlichen Zeitzonen zu überbrücken. Das wäre auf einer Präsenzveranstaltung bestimmt nicht möglich gewesen.

**Elektor:** Die Ausstrahlung des Personals am Messestand — und selbstverständlich die Präsentation des Standes — kann stark beeinflussen, wie viele Besucher sich an das Unternehmen wenden. Wie ließ sich das 2021 ausgleichen?

**Weyerer:** Unsere vier Ausstellerpakete waren maßgeschneiderte Lösungen, die den unterschiedlichen Anforderungen teilnehmender Unternehmen gerecht wurden. Wir haben außerdem einzeln zu buchende Dienstleistungen angeboten. Grundvoraussetzung für eine erfolgreiche Teilnahme ist jedoch auch bei einer digitalen Veranstaltung die aktive Mitwirkung der Aussteller vor, während und nach dem Event. Unsere Teilnahmeoptionen, wie Präsentationen auf dem Ausstellerforum, die Matchmaking-Funktion sowie mehrere Optionen zur Neukundengewinnung haben die besten Voraussetzungen dafür geschaffen.

# embedded world 2021

Exhibition&Conference

... it´s a smarter world

DIGITAL

Über 300 Aussteller, darunter führende Branchengrößen ebenso wie innovative Start-ups, präsentierten sich auf der digitalen Plattform, auf der sich die virtuelle Ausstellung und hochkarätige Konferenzen perfekt ergänzten.

**Elektor:** Nun ist die embedded world 2021 zu Ende gegangen. Wie ist es gelaufen?

**Weyerer:** 301 Aussteller und mehr als 13.400 Fachbesucher haben die Plattform genutzt, um sich über Trends, Produkte und Lösungen auszutauschen. Die embedded world DIGITAL hat damit die Bedeutung der Veranstaltung für die Embedded-Community unterstrichen - gerade in diesen besonderen Zeiten. Ich freue mich sehr, dass die embedded world DIGITAL so gut in der Branche angenommen wurde. Für uns ist das Bestätigung und Ansporn zugleich.

**Elektor:** Gibt es Lehren, die Sie aus der embedded world 2021 ziehen konnten?

**Weyerer:** Zunächst einmal, dass die Branche tatsächlich bereit war, das bisher unbekannte digitale Format für den Wissenstransfer zu nutzen. Natürlich haben wir auf unserer Seite alles dafür getan, trotzdem hätte das niemand mit Sicherheit vorhersagen können, weil es einfach keine Vorerfahrungen gab. Umso mehr freut es uns, dass das Programm und die angebotenen Formate so gut angenommen wurden. Wir haben aber auch erkannt, dass eine virtuelle Veranstaltung nicht so funktionieren kann wie eine Veranstaltung vor Ort. Die Vertriebs- und Marketingteams der Aussteller stehen bei der Online-Kommunikation vor anderen Herausforderungen und der einzelne Teilnehmer muss seinen „Messebesuch" anders planen. Die Teilnehmer müssen die Kommunikationstools aktiv nutzen, um das Beste aus dem Besuch herauszuholen.

**Elektor:** Konnten Sie auch Besucher aus Nicht-EU-Ländern mobilisieren, weil die Veranstaltung zu 100% im Netz stattfand?

**Weyerer:** Insgesamt haben wir uns über eine hohe Anzahl an internationalen Teilnehmern gefreut. Der Anteil war höher als bei der bereits sehr internationalen Veranstaltung in Nürnberg. Zur embedded world DIGITAL waren Teilnehmer aus mehr als 100 Ländern angemeldet, darunter ein großer Anteil aus den USA und Asien.

**Elektor:** Mehr als 96% der Messebesucher waren mit der embedded world Messe 2019 zufrieden und rund 97% der Fachbesucher gaben an, auch im nächsten Jahr die Messe besuchen zu wollen. Wie ist der Stand nach diesem sehr unsicheren Jahr?

**Weyerer:** Wir haben sehr viel positives Feedback für die embedded world 2021 DIGITAL erhalten, allerdings wurde uns auch häufig gesagt, dass die Branche einen Ort für den persönlichen Austausch braucht, wenn es die Rahmenbedingungen zulassen. Wie eine Face-to-Face-Veranstaltung angenommen wird, hängt natürlich sehr stark von den jeweiligen Rahmenbedingungen ab. Wir sind optimistisch, dass diese in einem Jahr eine Vor-Ort-Veranstaltung zulassen werden.

**Elektor:** Die embedded awards wurden zum 17. Mal verliehen. Wer waren die Gewinner in diesem Jahr?

**Weyerer:** In diesem Jahr gab es zwei Gewinner, die die Fachjury durch ihre herausragende Innovationsleistung überzeugt haben: Der embedded award ging an Maxim Integrated mit dem SoC MAX78000, während der embedded Start-up award an Kudan für die Software Kudan Visual SLAM ging.

**Elektor:** Vor der embedded world 2021 haben wir Sie gefragt, was Sie über die Zukunft der Messen denken. Wie ist Ihr Eindruck nach der digitalen Veranstaltung?

**Weyerer:** Die embedded world 2021 DIGITAL hat uns gezeigt, dass es auch mit einem digitalen Format möglich ist, Wissen zu vermitteln und Embedded-Experten aus aller Welt zu vernetzen. Die Vorbereitung einer solchen Veranstaltung erfordert jedoch, dass alle Beteiligten ihre Denkweise ändern. Am schwierigsten ist es meiner Meinung nach, charakteristische Elemente umzusetzen, die ein Messeerlebnis bietet. Persönliche Begegnungen, konstruktive Interaktion bei Fachgesprächen an Messeständen oder die Inspiration, die Besucher bei einem spontanen Stopp am Stand eines Ausstellers erfahren, können nur Live-Events bieten. Daher freuen wir uns auf die embedded world 2022 in Nürnberg!

210215-02

## WEBLINK

[1] **embedded world Website**: www.embedded-world.de

# AVR-IoT und PIC-IoT-Entwicklungsboards von Microchip

## Sichere Kommunikation mit der AWS-Cloud

Von **Mathias Claußen** (Elektor)

Die Entwicklungsboards AVR-IoT WA und PIC-IoT WA von Microchip für die WLAN-Anbindung zu AWS IoT Core (Amazon Web Services) bieten einen sicheren Zugang zu Cloud-Diensten und IoT-Anwendungen. Geringer Stromverbrauch, ein Onboard-Debugger und die integrierte Ladeelektronik für Lithium-Akkus machen sie zu einer eleganten Lösung für Entwickler.

Wie allseits bekannt hat Microchip 2016 Atmel übernommen und bietet jetzt nicht nur PIC-Mikrocontroller, sondern auch Atmel-Produkte einschließlich der AVR-Mikrocontroller an. Microchip wird die AVR-Serie nicht nur beibehalten, sondern sogar erweitern und neue Funktionen und Merkmale hinzuzufügen. Tools und Evaluierungsboards enthalten nun also das Beste von allen Produkten, die von Microchip angeboten werden. In diesem Kontext möchten wir Ihnen die Entwicklungsboards AVR-IoT und PIC-IoT WA von Microchip vorstellen, die der WLAN-Anbindung zu AWS IoT Core (Amazon Web Services) dienen.

Die ersten Schritte in die Welt von Cloud-Storage und -Computing sind nicht ganz trivial. Der Betrieb einer sicheren Datenübertragung zu einem Cloud-Service wie AWS bei geringem Energiebedarf kann für Entwickler ziemlich herausfordernd ausfallen. Microchip bietet hierfür zwei Entwicklungs-Boards an, die den Zugang zu AWS IoT Core ermöglichen und dabei eine sichere Kommunikation über WLAN bei geringem Energieverbrauch bieten und die nötige Entwicklungszeit verkürzen. Bevor wir die Hard- und Software detailliert beleuchten, werden wir zunächst den Lieferumfang zeigen und überlegen, was man von diesen Lösungen erwarten kann.

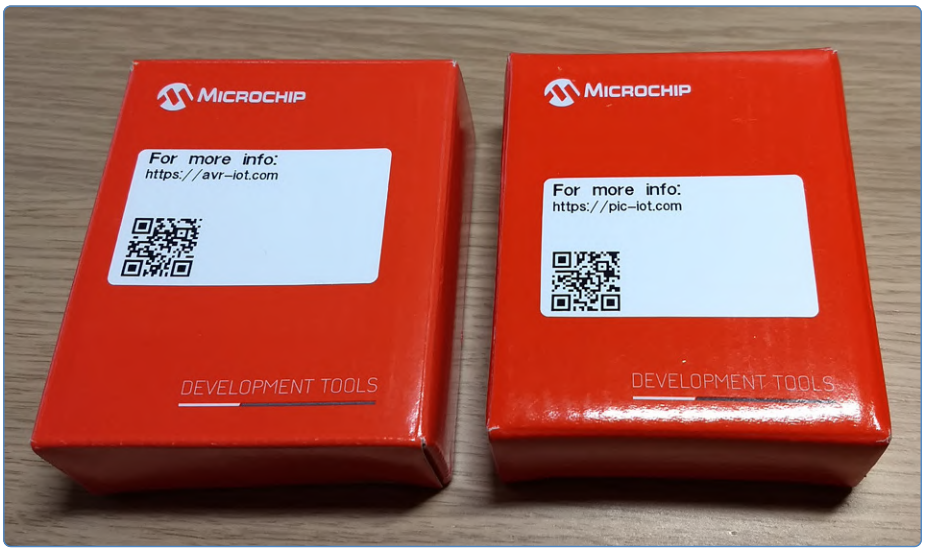



*Bild 1. Beide Kartons nebeneinander.*

### Was ist drin?

Auf den beiden fast identischen Kartons (**Bild 1**) ist ein Link und ein QR-Code aufgedruckt, die zu der zugehörigen Produkt-Webseite führen. Da fast jeder über mehr als genug USB-Kabel verfügt, sind diese aus Platz- und ökologischen Gründen nicht enthalten. Die beiden Kartons enthalten unterschiedliche Boards. Das Entwicklungsboard PIC-IoT WA (Produkt-Nummer: EV54Y39A) basiert auf der zentralen MCU PIC24FJ128GA705. Das Entwicklungsboard

AVR-IoT WA (Produkt-Nummer: EV15R70A) basiert auf einer ATmega4808-MCU aus der AVR-Reihe. Man hat also bei gleicher Funktion die Wahl zwischen einer PIC- oder einer AVR-Lösung.

In **Bild 2** sieht man, dass sich nicht nur die Kartons, sondern auch die Boards stark ähneln. Am rechten Rand der Platinen (**Bild 3**) sitzt eine Mikro-USB-Buchse, die das Board mit Strom versorgt und das System mit dem Debugger verbindet. Neben dem USB-Anschluss ist der Debugger zu sehen, und darüber befindet sich eine Ladeschaltung für Lithium-Akkus. Ist ein solcher angeschlossen, kann das Board als netzunabhängige Lösung verwendet werden.

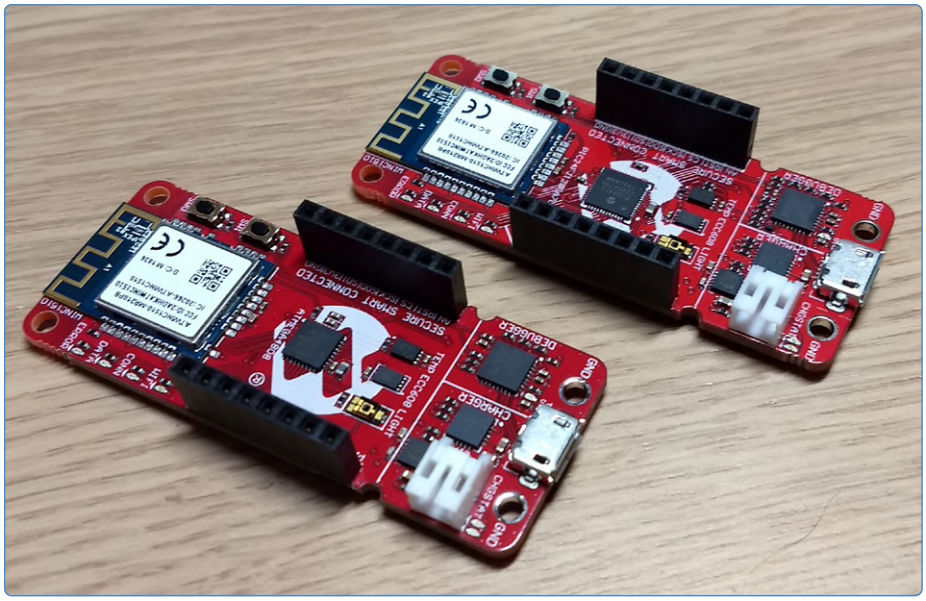


*Bild 2. Beide Entwicklungsboards ohne Verpackung.*

## Hardware der Entwicklungsboards

Die Stromversorgung beherbergt das Li-Ion/LiPo-Lade-IC MCP73871, das vom Abwärtswandler MIC33050 mit integrierter Induktivität unterstützt wird. Diese Kombination eignet sich für eine Lithium-Akkuzelle mit einer Ladeschlussspannung von 4,2 V (wichtig bei Auswahl der Zelle); der DC/DC-Wandler bietet einen Strom von bis zu 600 mA.

Das Lade-IC fungiert als ideale Diode und verhindert einen Rückfluss des Stroms vom Akku zurück zum USB-Anschluss. Es gibt noch mehr auf der Platine, das eine genauere Betrachtung verdient: Weiter links sieht man mit dem Temperatursensor MCP9808 und dem Krypto-Co-Prozessor ATECC608A zwei ICs und einen Lichtsensor. Der Krypto-Chip kann zwar zur sicheren Kommunikation zwischen dem Board und den AWS IoT Core Services eingesetzt werden, doch ist er nicht darauf beschränkt, sondern ermöglicht sichere Kommunikation mit verschiedenen Cloud-Diensten.

In der Mitte befindet sich der zentrale Mikrocontroller. Die Variante mit dem PIC24FJ128GA705 bietet 128 KB ECC-Flash und 16 KB RAM. Bei 32 MHz Takt erreicht diese MCU 16 MIPS und bietet ein schönes Set an Peripherie sowie einen Low-Power-Modus. Bei der Variante ATmega4808, einer der neueren AVR-MCUs, stehen 48 KB Flash und 6 KB RAM zur Verfügung. Bezüglich Peripherie ähnelt er eher einer XMEGA- als einer typischen AVR-MCU. Der ATmega4808 erreicht mit einem Takt von bis zu 20 MHz maximal 20 MIPS und bietet erweiterte Peripherie sowie Energiespar-Optionen.

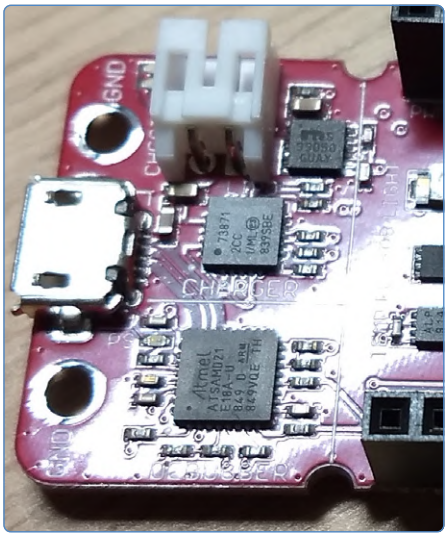



*Bild 3. Debugger- und Stromversorgung.*


*Bild 4. Das WLAN-Modul ATWINC1510.*

Am linken Platinenrand sitzt das WLAN-Modul ATWIN1510 (**Bild 4**). Es unterstützt als per SPI angebundenem 2,4-GHz-Single-Band-Controller den Netzwerkzugang nach IEEE 802.11 b/g/n. Das Modul bietet mehr als nur WLAN: Man kann es quasi als Netzwerk-Coprozessor betrachten, der anspruchsvolle Aufgaben wie die Verschlüsselung des Traffics und TCP/IP-Paketverarbeitung übernimmt.

Die mittigen Buchsenreihen an beiden Seiten (**Bild 5**) sind mikroBus-Click-Board-Steckverbinder, mit denen man zusätzliche Sensoren und Geräte aus dem Sortiment oder mikroBus-kompatible Geräte z.B. zur 3D-Bewegungserkennung aufstecken kann. Dreht man das Board um, dann entdeckt man, dass die verfügbaren Anschlüsse mit ihrem Namen und Pin-Nummer gekennzeichnet sind (**Bild 6**). Wer an weiteren Details der Platinen wie etwa Schaltplänen interessiert ist: Microchip stellt schön gestaltete Anleitungen für beide Varianten zur Verfügung, die nicht nur viele Informationen bieten, sondern auch einen Schnellstart für die AWS-Verbindung beinhalten.

## Vorinstallierte Software

Beide Boards kommen mit vorinstallierten Beispielanwendungen, die einen direkten Out-of-the-Box-Test für AWS IoT Core unter Verwendung eines Microchip-Accounts ermöglichen. An dieser Stelle eine Warnung, die auch in der Schnellstartanleitung zu finden ist: Die mit diesem Account verschickten Daten stehen allen Zugangsberechtigten zur Verfügung. Praktisch

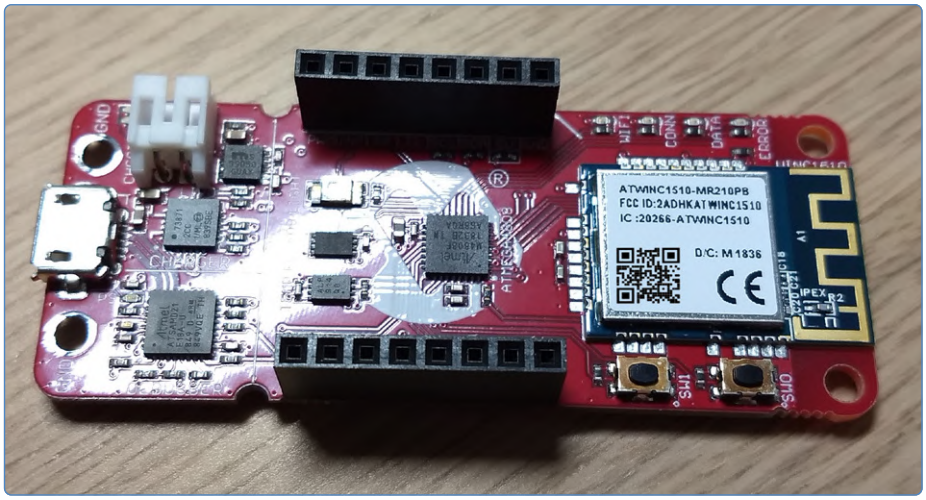



*Bild 5. Der mikroBus-Anschluss des Entwicklungsboards AVR IoT WA.*


*Bild 6. Die Rückseite des Entwicklungsboards mit Pin-Bezeichnungen.*

bedeutet das hier, dass sie quasi öffentlich zugänglich sind. Die Integration der Boards ins eigene Netzwerk ist unkompliziert: Beim Anschluss des Boards an einen PC erscheint ein neues Massenspeicherlaufwerk namens *CURIOSITY*. Die Datei *CLICK-ME.HTM* im Stammverzeichnis kann in einem Browser geöffnet werden. Sie führt durch die nötigen Konfigurationsschritte. Zuerst aktualisiert man die Firmware für die MCU auf die neueste Version (**Bild 7**), sonst gibt es Probleme bei der Konfiguration des WLANs. Da die Firmware des Sicherheits-Chips dadurch noch nicht aktualisiert ist, kann er noch veraltete Firmware samt veralteter oder Zertifikate enthalten. Sollten sich Probleme bei der AWS-Verbindung ergeben, werfen Sie einen Blick in den Textkasten „Was noch nicht im Manual steht“. Dort ist nachzulesen, wie man alle Komponenten des Boards auf die neueste Version aktualisiert. Für die Aktualisierung ziehen Sie die heruntergeladene Datei mit der Extension „hex“ auf das Laufwerk *CURIOSITY*. Anschließend wird die neue Firmware automatisch in den Flash-Speicher übertragen und die MCU neu gestartet. Das Board in das eigene WLAN zu bekommen ist eine Sache von wenigen Klicks. die Eingabe der Zugangsdaten zur Ihrem WLAN erfolgt direkt via *CLICK-ME.HTM* (**Bild 8**). Dabei wird eine WIFI.CFG-Datei erzeugt, die per Drag & Drop auf das Laufwerk *CURIOSITY* gezogen wird, um die WLAN-Zugangsdaten in der MCU zu speichern und damit eine Verbindung zum WLAN herzustellen. Ist diese Verbindung hergestellt, wird das Board versuchen, sich mit dem AWS-Service zu verbinden und dann mit der Veröffentlichung von Daten zu beginnen.

**Was noch nicht im Manual steht**

Es kann sein, dass beim gelieferten Board mit älterer Firmware die Daten für den ATECC608A veraltet sind. Dann kann sich das Board zwar problemlos mit Ihrem WLAN verbinden, aber keine Verbindung zum AWS IoT Core herstellen. Sie können dann nicht die Schritte der Anleitung nachvollziehen. Um die neueste Firmware auch für den ATECC608A und den Onboard-Debugger aufzuspielen, kann man das IoT Provisioning Tool aus [1] verwenden. Es läuft auf allen drei wichtigen Betriebssystemen als Kommandozeilen-Anwendung. Hierzu entpackt man die heruntergeladene Datei, und startet das Terminal (MacOS und Linux) bzw. eine Befehlszeile (cmd.exe bei Windows). Für Windows lautet der Befehl `iotprovision-bin.exe -c aws -m sandbox`. Anschließend ist das Board „up to date“ und alle erforderlichen Zertifikate sind vorhanden. Für die Verwendung von Google-Services sollte man sich das Video-Tutorial von Microchip zum Wechsel des Cloud-Providers [2] anschauen.

## Wie geht es weiter?

Der kleine Knopf unter den Sensorwerten führt durch die nächsten Schritte. Der erste ist die Einrichtung von *MPLABX* (**Bild 9**). Nach der Fusion von Microchip und Atmel wurde die MPLABX-Entwicklungsumgebung, die ursprünglich für die PIC-Entwicklung gedacht war, um AVR- und SAM-Chips ergänzt. Auch wenn das nicht mit dem Look & Feel von Atmel-Studio (basiert auf Visual Studio) übereinstimmt, ist es erfreulich, dass man so nicht mehr auf Windows festgelegt wird. Auch Benutzer von MacOS und Linux werden damit AVR-basierte Projekte entwickeln können – letztere gar auf AMD64-Maschinen. Wer nur eine 32-bit-CPU zur Verfügung hat oder aber noch ein 32-bit-Betriebssystem verwendet, wird mit der neuesten Version von MPLABX Pech haben. MPLABX basiert auf der Entwicklungsumgebung Netbeans, die einst als IDE für Java begann und im Laufe der Zeit auf andere Sprachen erweitert und später von Microchip zu MPLABX modifiziert wurde. Zum Entstehungszeitpunkt dieses Artikels war die Version 5.45 aktuell, die hier auf einem Windows-Rechner verwendet wurde. Die verschiedenen Architekturen benötigen entsprechende Compiler und die Support-Dateien. Der einfache Weg hierfür ist, sie bei genügend Platz auf dem Laufwerk gleich während der Installation mitzuinstallieren. Man

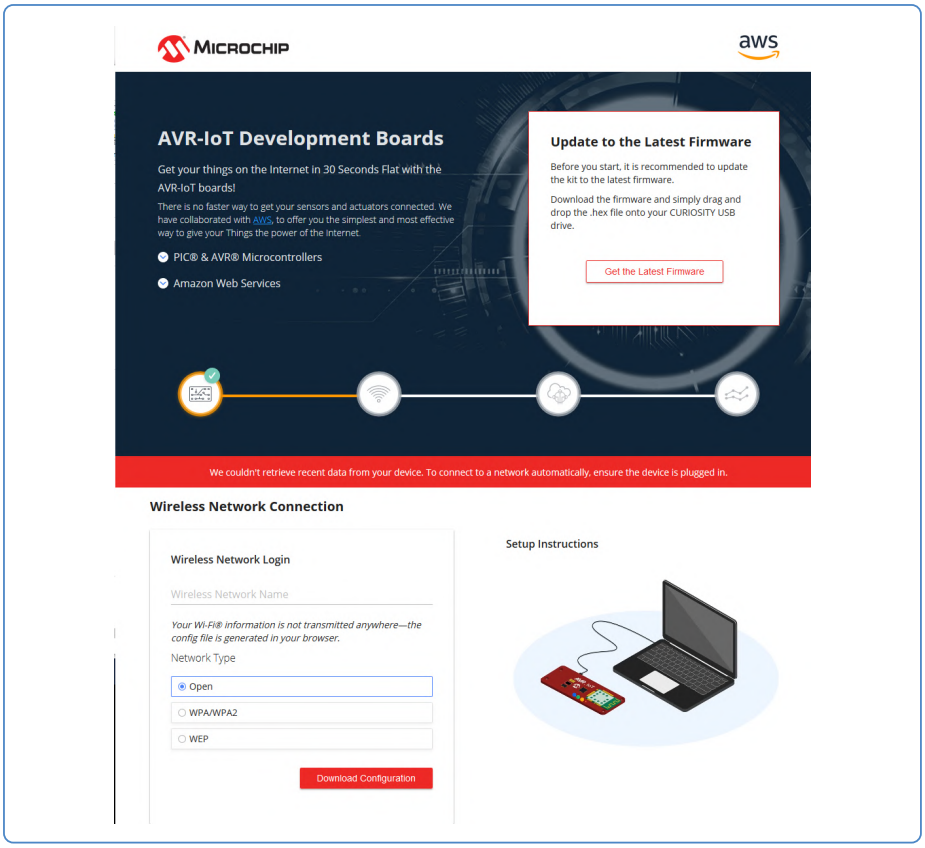



*Bild 7. Die Datei CLICK-ME.htm zur Konfiguration der Boards.*

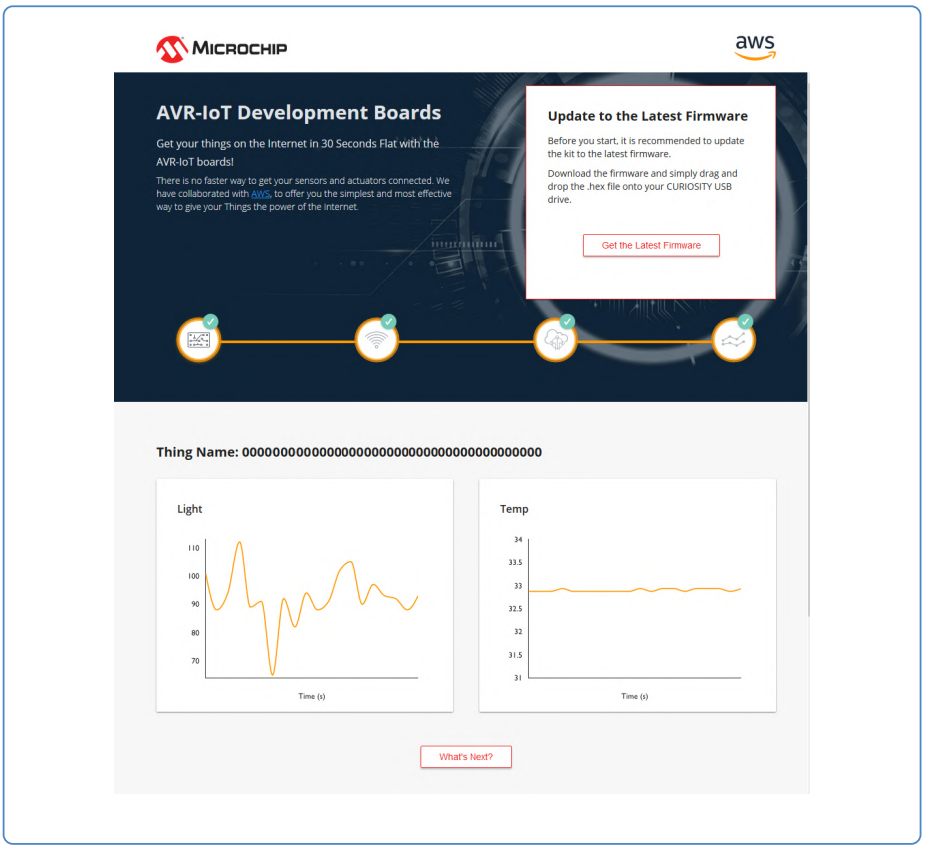



*Bild 8. Nach erfolgreicher WLAN-Verbindung kommen die ersten Daten an.*

muss lediglich die verwendete Architektur auswählen, um die Support-Dateien zu erhalten. Bezüglich Compiler wird man nach der Installation darüber informiert, dass man ihn von der Microchip-Website holen muss. Wer den von Atmel Studio 7 gewohnten AVR-GCC-Compiler verwenden will, muss diesen manuell von der Microchip-Seite herunterladen, die Zip-Datei extrahieren und den Compiler manuell zu MPLABX hinzufügen (**Bild 9**). Hier wird der Pfad zum Verzeichnis angegeben, in dem die Compiler-Dateien extrahiert wurden. Er muss zum Ordner *bin* führen – ansonsten wird die AVR8-Toolchain nicht erkannt. Da die Beispielprojekte standardmäßig mit der XC-Compiler-Serie kompiliert werden, ist dieser Schritt nicht erforderlich, aber hilfreich, wenn man an die Arbeitsweise des AVR-GCC gewöhnt ist.

## Entwicklungsboard AVR-IoT WA

Hier muss man die neueste AVR-IoT-WA-Anwendung herunterladen, die unter [3] zu finden ist. Nach dem Herunterladen des Codes kann dieser mit MPLABX geöffnet werden (**Bild 10**). Durch das angeschlossene Board bekommt man auch alle Ressourcen und Links zur Dokumentation präsentiert. Der Code wird für die nächsten Schritte des Tutorials benötigt und ermöglicht die Anpassung der Anwendung für eigene Zwecke oder aber er inspiriert zu eigenen Ideen und Entwicklungen. **Bild 11** zeigt das geöffnete und erweiterte Projekt. Der letzte Schritt des Tutorials für das AVR-Board zeigte zum Zeitpunkt meiner Experimente noch nicht die neueste Struktur des Codes. Es braucht einige Sekunden, um die wesentlichen Teile des Codes zu finden. Zur Erleichterung werden in **Bild 12** die erforderlichen Code-Änderungen gezeigt, damit die grüne LED als Aktuator arbeiten kann. Nachdem der Quellcode geändert wurde, kann er auf das AVR-Board hochgeladen werden und der LED-Status kann mit dem Schiebeschalter gesteuert werden (**Bild 13**). Abgesehen von diesem kleinen Glitch funktioniert das Board wunderbar. Wenn man das Tutorial durchgearbeitet hat, stehen einem noch die *Microchip IoT Developer Guides for AWS* [4] für weitere Informationen und zur Verwendung dieser Boards mit den Amazon Web-Services zur Verfügung.

## Entwicklungsboard PIC-IoT WA

Hier muss die PIC-IoT-WA-Anwendung [5]

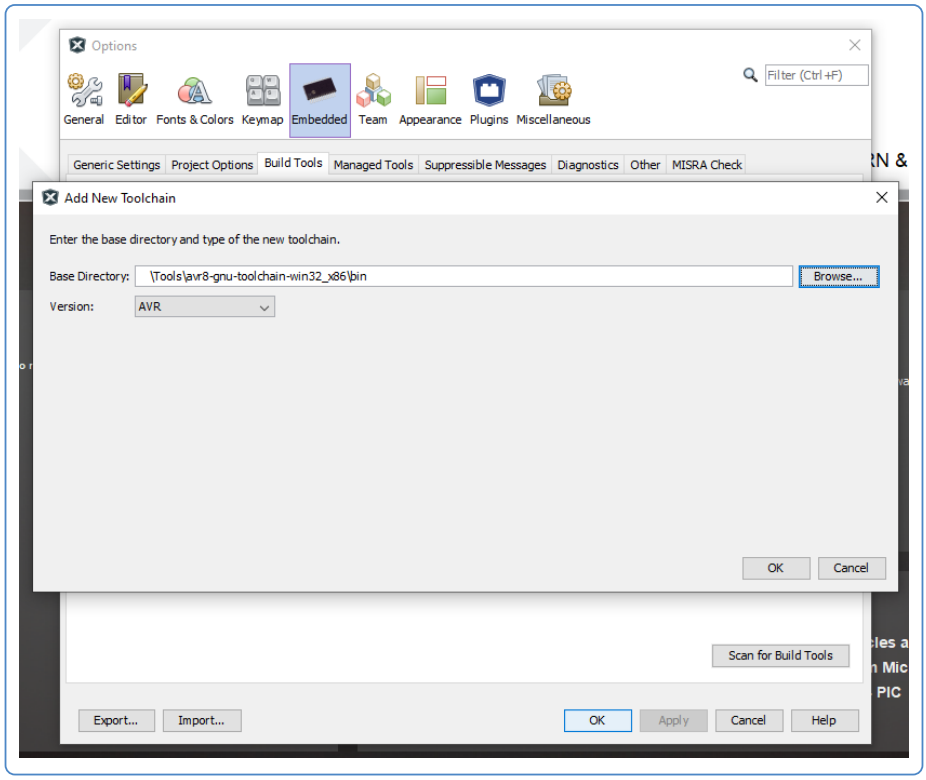



*Bild 9. Aufbau der AVR GNU Toolchain.*

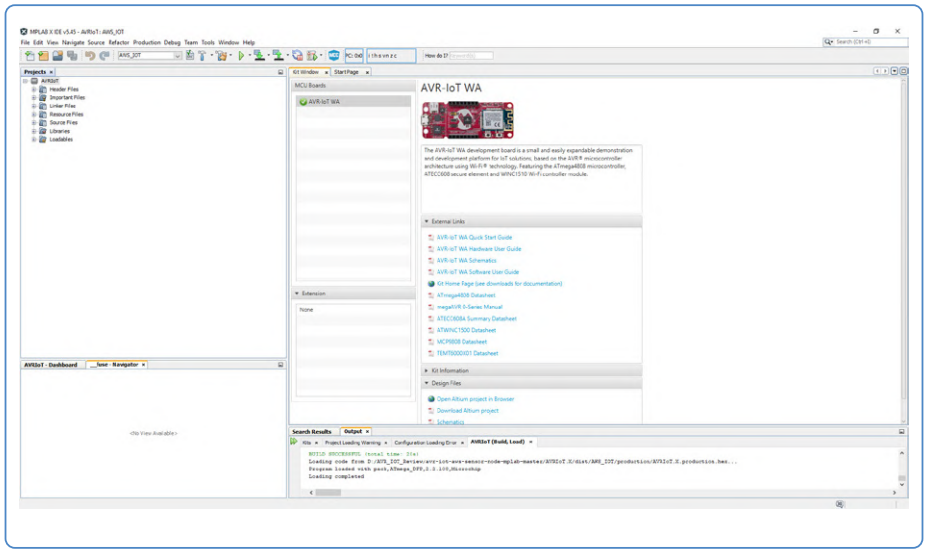



*Bild 10. Überblick über das MPLABX-Projekt und die Dokumentation.*

heruntergeladen werden. Der Code kann in MPLABX geöffnet werden und hat die gleiche Struktur wie der Code für den AVR. Das Tutorial funktioniert fast genauso wie das des AVR – nur dass ein PIC mehr Flash, zusätzliche Peripherie und auch etwas mehr RAM bietet (**Bild 14**).
Man beachte, dass ein PIC im Vergleich zu AVR mehr Breakpoints in der Hardware für das Software-Debugging und Daten-Watchpoints bietet, wobei die Instruktionen pro Takt weniger effizient sind als bei AVR-MCUs. Wer den bereitgestellten Code untersuchen möchte: Der Code von `main.c` ist wirklich kurz:

```
#include „mcc_generated_files/
application_manager.h“

int main(void)
{
    application_init();

    while (1)
    {
        runScheduler();
    }
    return 0;
}
```

Die Initialisierung aller Komponenten erfolgt in `application_init()`. Diese Funktion steckt in `application_manager.c`, wo die meisten höheren Funktionen zu finden sind. Der Code verwendet die Funktion `runScheduler()`, um alle anfallenden Aufgaben zu erledigen und zeitlich zu koordinieren. Hiervon wird die Funktion `timeout_callNextCallback()` aufgerufen, welche fällige Aufgaben ausführt:

```
inline void
timeout_callNextCallback(void)
{
    if (executeQueueHead == NULL)
        return;
    bool tempIE = (IEC0bits.T1IE ==
      1?1:0);
    IEC0bits.T1IE = 0;
    timerStruct_t *callBackTimer =
            executeQueueHead;
    // Done, remove from list
       executeQueueHead =
       executeQueueHead->next;
    // Mark the timer as not in use
    callBackTimer->next = NULL;
    if(tempIE)
    {
        IEC0bits.T1IE = 1;
    }
    uint32_t reschedule =
      callBackTimer->
        callbackPtr(callBackTimer->
          payload);
    // Do we have to reschedule it?
    //If yes then add delta to
    //absolute for reschedule
    if(reschedule)
    {
        timeout_
create(callBackTimer, reschedule);
    }
}
```

Der Code prüft, ob etwas zur Ausführung bereit ist, und wenn nicht, kehrt er sofort zurück. Im Folgenden deaktiviert der Code zunächst für eine kurze Zeit einen Timer-Interrupt, worauf wir später eingehen werden, greift das Element, das zur Ausführung bereit ist, und ruft die Funktion auf. Der Aufruf `uint32_t reschedule =`

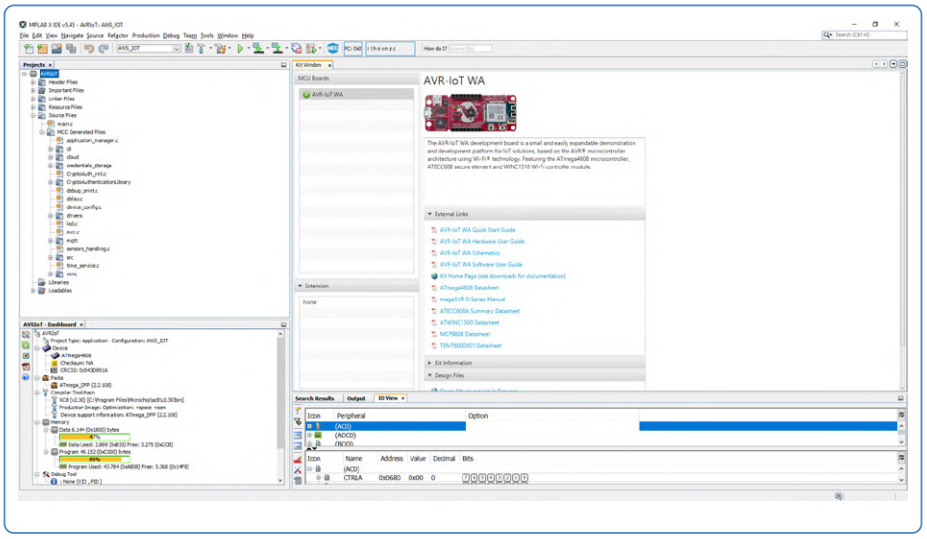



*Bild 11. Erweitertes Beispiel-Projekt für das AVR-Board.*

`callBackTimer->callbackPtr(callBackTimer->payload);` führt die Funktion aus, und die Funktion selbst gibt die Zeitspanne zurück, in der sie erneut aufgerufen werden muss. Wenn diese Zeitspanne größer als Null ist, wird sie mit der angegebenen Zeitspanne in eine Liste von wartenden Aufgaben zurückgestellt. Der erwähnte Timer und Timer-Interrupt ist der andere Teil, der das System die Tasks ausführen lässt. In einem definierten Intervall wird die Funktion `timeout_isr` aus dem Timer-Interrupt heraus aufgerufen.

```
void timeout_isr(void)
    {
        timerStruct_t *next =
           listHead->next;
        absoluteTimeofLastTimeout =
             listHead->absoluteTime;
        lastTimerLoad = 0;

    if (listHead != &dummy) {
    enqueueCallback(listHead);
    }

    listHead = next;
    startTimerAtHead();
}
```

Vereinfacht gesagt überprüft der Code eine Liste von Aufgaben. In der Liste können sich wirklich auszuführende Aufgaben befinden oder Dummy-Aufgaben, die nichts tun. Muss eine Aufgabe ausgeführt werden, wird sie in eine Liste mit fälligen Aufgaben eingetragen. Die nächste auf ihre Ausführung wartende Aufgabe der Liste wird geholt, und der Zeitgeber passt sein Intervall entsprechend an. Die Liste fälliger Aufgaben bearbeitet die schon erwähnte Funktion `timeout_callNextCallback`. Die ISR plant und bereitet die Aufgaben timer-gesteuert vor, die als Nächstes ausgeführt werden müssen. In Kombination mit der Funktion `timeout_callNextCallback` wird ein Zyklus erzeugt, in dem die Aufgaben verwaltet und zeitgesteuert ausgeführt werden. Wenn keine Tasks mehr ausgeführt werden müssen, könnte die MCU in den Ruhezustand versetzt werden und darauf warten, dass sie durch den Timer oder andere Events wieder geweckt wird. Dieses Konzept eignet sich nicht nur für diese Anwendung, sondern könnte auch für andere Projekte interessant sein, um Tasks mit reduziertem Stromverbrauch auszuführen.

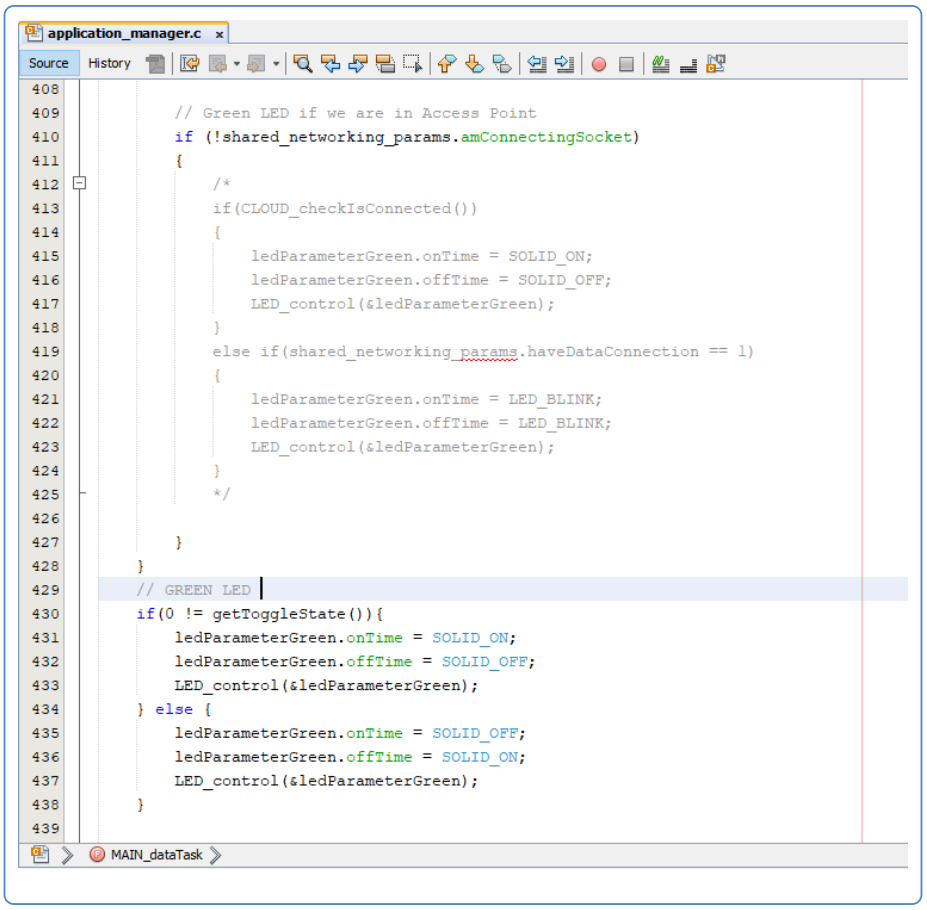



*Bild 12. Tutorial Schritt 3 – hier muss der Code modifiziert werden.*

## Passende Produkte

- **Microchip AVR-IoT WA Development Board**
  www.elektor.de/microchip-avr-iot-wa-development-board
- **Microchip PIC-IoT WA Development Board**
  www.elektor.de/microchip-pic-iot-wa-development-board

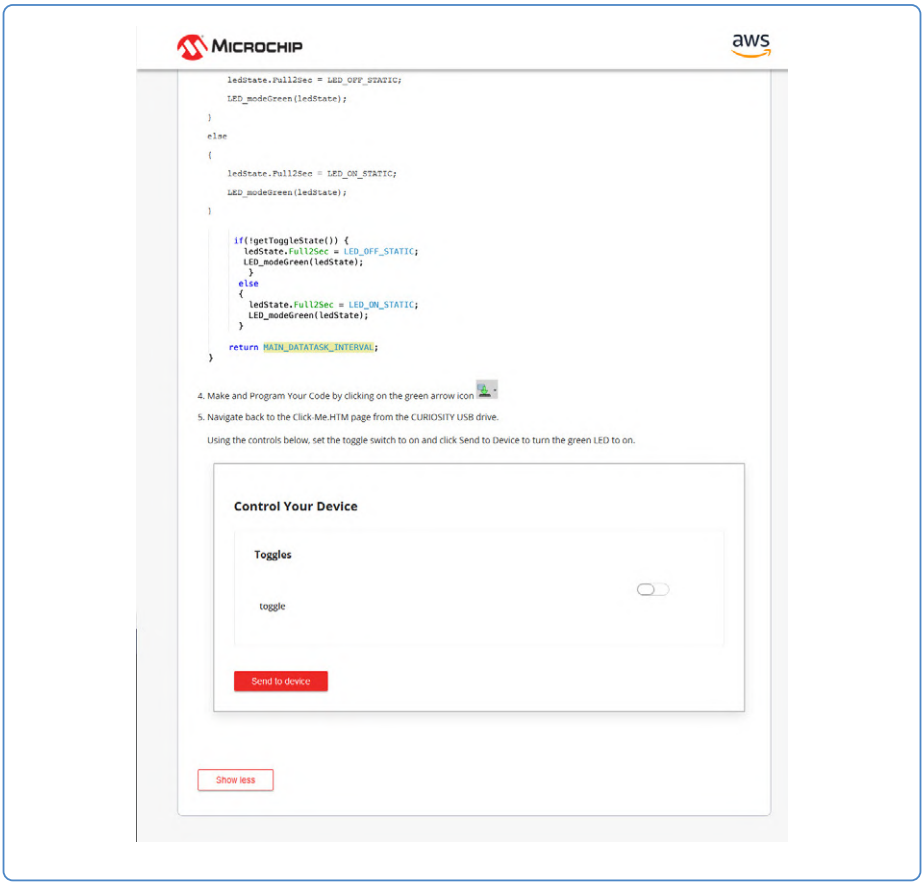



*Bild 13. Dieser Schalter steuert den Status der LED auf der Platine.*

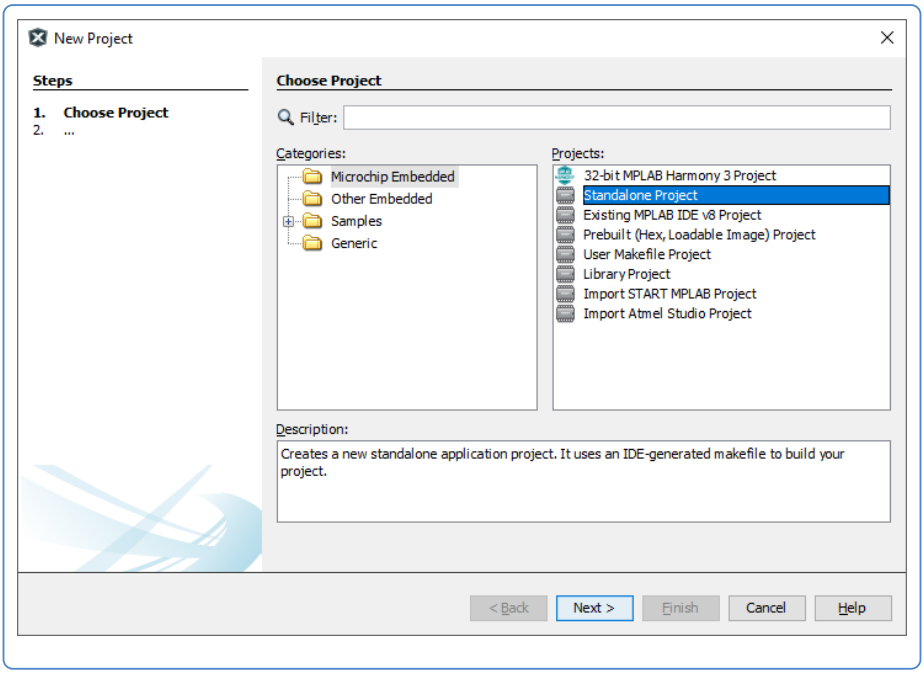



*Bild 15. MPLABX-Projektstart.*

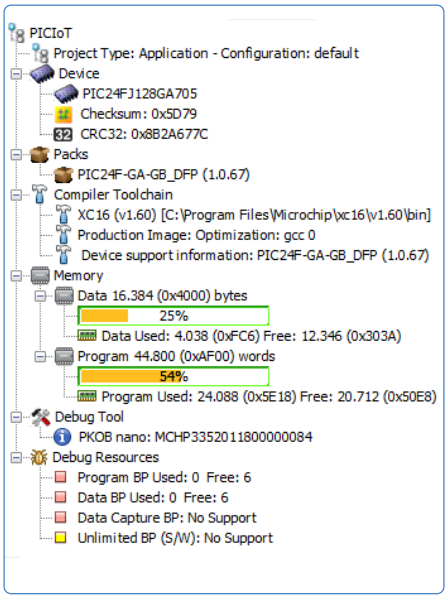



*Bild 14. Ressourcenverbrauch beim PIC-basierten Board.*

## Einen Schritt weiter

Für das AVR-IoT WA Board findet sich unter [6] eine Anleitung, um das Beispielprojekt mit Hilfe von MPLABX komplett neu zu erstellen. Dadurch wird man besser mit der Entwicklungsumgebung und der Programmierung vertraut. Doch Achtung: Auch wenn das alles gut funktioniert, werden einige Teile und Plug-Ins noch Verbesserungen ihrer Benutzerfreundlichkeit und Stabilität erfahren. Die **Bilder 15**, **16** und **17** zeigen einige Schritte auf dem Weg durch die Assistenten. Nach der Absolvierung der Schritte der Anleitung hat man ein Projekt, das auch mit dem AVR-GCC kompiliert werden kann, falls man für diesen speziellen Compiler geschriebenen Code einbinden muss.

Bild 17 zeigt auch die Komponenten, die im AWS-Beispiel-Projekt enthalten sind, darunter *CryptoAuthLibrary*, *WINC15XX* oder *Message Queue Telemetry Transport* (MQTT). Sie beschäftigen sich mit den Sicherheitsfunktionen der beiden Boards. AWS IoT Core verwendet MQTT für den Datenaustausch. Wenn keine zusätzlichen Messungen vorgenommen werden, werden die Daten im Klartext ausgetauscht, so dass die Daten auf dem Weg vom Gerät oder zurück verändert werden könnten. Um das zu verhindern nutzt man MQTT Transport Layer Security (TLS) zur Verschlüsselung der Kommunikation. Auf beiden Boards befindet sich ja ein ATECC608A, der die kryptographische Arbeit von der MCU abnimmt. Zwecks Sicherheit bietet der WINC1510 TLS-verschlüsselte Kommunikation und kann dabei durch den ATECC608A unterstützt werden, damit die übertragenen Daten nicht verändert werden können. Aufgrund der Verwendung

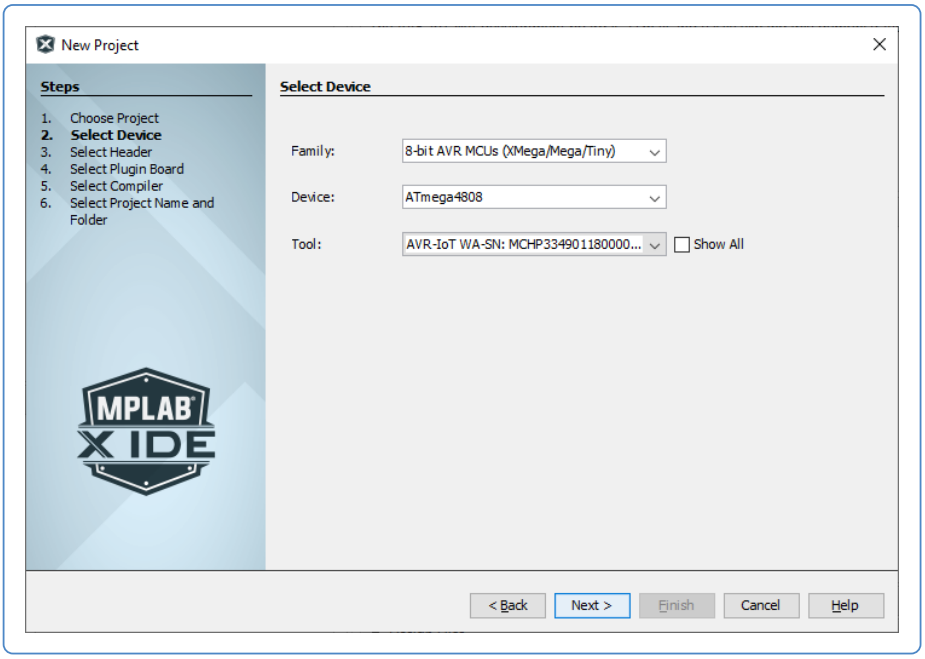



*Bild 16. Auswahl des Ziels.*

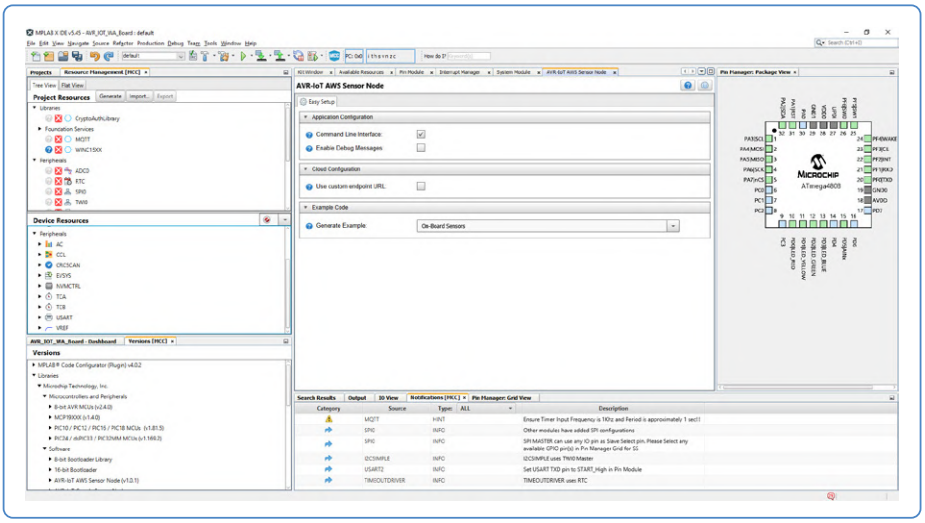


*Bild 17. Konfigurator für Microchip-Code.*

von MQTT ist man nicht auf den AWS IoT Core beschränkt. Man benötigt lediglich einen MQTT-Broker, der die Verbindungen annimmt und Daten verarbeitet.

## Einfacher Einstieg

Sowohl Preis der Boards als auch ihre Software-Unterstützung sind auf gutem Niveau. Kombiniert mit den gut ausgestatteten Dokumentations- und Debugging-Möglichkeiten bieten die Boards einen einfachen Einstieg in die Welt der vernetzten Geräte, die mit beschränkter Energieaufnahme laufen. Die integrierte Ladeschaltung für Lithium-Akkus erlaubt den direkten Einsatz im Feld. Der mikroBus-Steckverbinder ermöglicht Zugriff auf eine Menge zusätzlicher Hardware für eigene Projekte. Man kann mit diesen Boards bei der bevorzugten MCU-Architektur bleiben, oder aber man nutzt die Gelegenheit, die jeweils andere MCU-Gattung samt ihren Möglichkeiten kennenzulernen. In Kombination mit dem Onboard-Debugger und der plattformübergreifenden Lösung für alle drei wichtigen Betriebssysteme bieten die Boards eine mehr als vollständige Lösung.

200567-03

## Sie haben Fragen oder Kommentare?

Mit technischen Fragen können Sie sich gerne an die Elektor-Redaktion wenden unter der E-Mail-Adresse: redaktion@elektor.de.

## Ein Beitrag von

Review und Text: **Mathias Claußen**
Redaktion: **Jens Nickel**
Übersetzung: **Dr. Thomas Scherer**
Layout: **Giel Dols**

## WEB LINKS

[1] **Microchip-IoT-Tool**: www.microchip.com/design-centers/internet-of-things/iot-dev-kits/iot-provision-tool

[2] **Video zum Wechsel von Amazon zur Google-Cloud und zurück**: www.youtube.com/watch?v=nwP8obSRaaE&feature=emb_logo

[3] **GitHub-Repository für das AVR-IoT WA Board**: ttps://github.com/microchip-pic-avr-solutions/avr-iot-aws-sensor-node-mplab

[4] **Microchip IoT Developer Guides for AWS** : https://github.com/microchip-pic-avr-solutions/microchip-iot-developer-guides-for-aws

[5] **GitHub-Repository für das PIC-IoT WA Board**: https://github.com/microchip-pic-avr-solutions/pic-iot-aws-sensor-node

[6] **Benutzerhandbuch des AVR-IoT-Boards**: http://ww1.microchip.com/downloads/en/DeviceDoc/AVR-IoT-WA-Development-Board-User-Guide-DS50002998B.pdf

# Secure-Boot-Lösung für Raspberry Pi

## Viel Sicherheit zum vernünftigen Preis

Vom **Elektor-Team**

Das Schöne an einer SD-Karte ist, dass sie einen einfachen Zugriff auf Daten ermöglicht. Der Inhalt einer Karte lässt sich auf Laptops, PCs und sogar Mikrocontrollern kopieren oder verändern. Dies stellt jedoch ein Sicherheits- und Wartungsrisiko für industrielle Maschinen und andere Anwendungen dar, wenn sie in Kombination mit einem Raspberry Pi verwendet werden. Die Schweizer Halbleiterfirma Swissbit hat eine Secure-Boot-Lösung für den Raspberry Pi entwickelt, wodurch die Daten auf Kartenebene mit Hilfe integrierter Verschlüsselungs- und Authentifizierungs-Hardware geschützt sind.

Die leichte Verfügbarkeit und Einfachheit des Raspberry Pi haben nicht nur Maker überzeugt. Im Laufe der Jahre und dank der Einführung des Raspberry-Pi-Compute-Moduls wurde er zunehmend auch in industriellen Anwendungen eingesetzt. Allerdings haben auf diesem Ansatz basierende Geräte eine große Achillesferse: Der primäre Massenspeicher in Form einer SD-Karte ist schlicht ungeschützt.

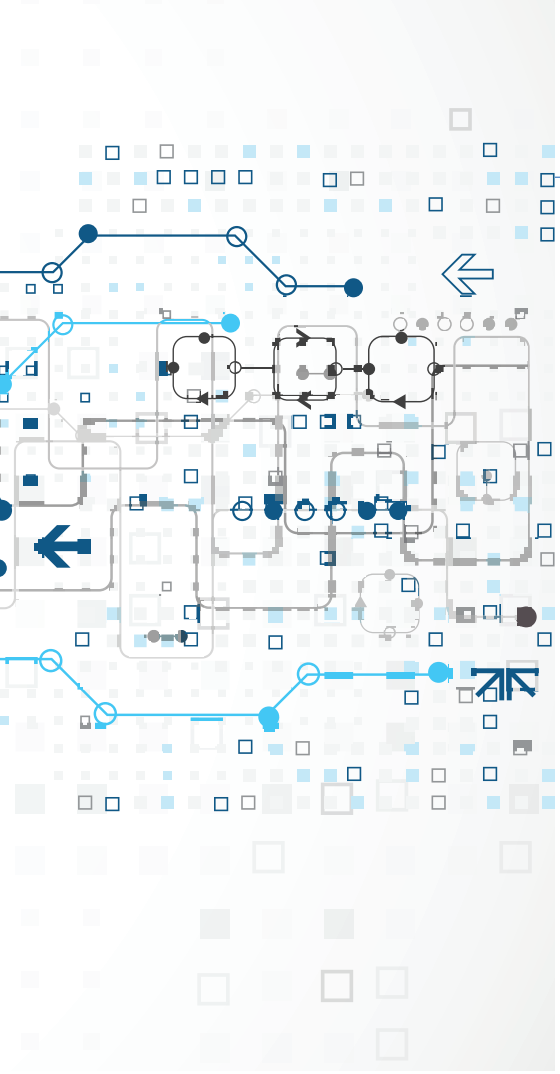
Ein Raspberry Pi selbst bietet keine Unterstützung für einen vertrauenswürdigen „Secure Boot“. Somit kann praktisch jeder Ihre SD-Karte entfernen, eine Kopie davon ziehen und die gespeicherten Dateien, die Firmware und die Benutzerdaten untersuchen. Um dies zu verhindern hat Swissbit eine Lösung entwickelt, welche die Daten auf der SD-Karte verschlüsselt und die für eine Secure-Boot-Lösung erforderliche Authentifizierungsinfrastruktur integriert.

### Swissbit AG

Die Schweizer Swissbit AG [1] war bis 2001 Teil der Siemens AG und hatte sich auf Speicherprodukte spezialisiert. Sie verfügt über eine Produktionsstätte in Berlin, um dort spezielle Speicherlösungen zu entwickeln und herzustellen. Ihre Secure Boot Solution für den Raspberry Pi ist ein solches Produkt.

### Raspberry Pi und die fehlende Sicherheit

Wer mit einem Raspberry Pi arbeitet, wird feststellen, dass er eine implizite Vertrauenskette verwendet. Das bedeutet, dass ein Raspberry Pi unabhängig vom verwendeten Boot-Medium versuchen wird, Code davon



auszuführen. Dies öffnet ihn für Angriffe, die das Betriebssystem des Geräts kompromittieren können. Da jeder Benutzer und selbst jeder Angreifer vollen Zugriff auf die Daten des Boot-Mediums hat, ist es schwer zu erkennen, ob ein System kopiert oder mit Malware bzw. Data-Mining-Tools kompromittiert wurde. Dies kann im schlimmsten Fall dazu genutzt werden, um die Netzwerke anzugreifen, mit denen der Raspberry Pi verbunden ist.

Eine Gegenmaßnahme ist die Erweiterung des Raspberry Pi um eine sichere Boot-Lösung. Da die Hardware selbst jedoch nicht verändert werden kann, braucht es eine intelligente Lösung, die auch in bestehenden Systemen nachgerüstet werden kann. Hier spielt die PS-45u DP „Raspberry Edition“ von Swissbit zu einem vernünftigen Preis ihre Stärken aus. Da es sich um eine SD-Karte handelt, kann sie einfach das bestehende Speichermedium ersetzen und so Ihre Hardware nahtlos um einen sicheren Bootvorgang samt Datenschutz ergänzen.

## Innenleben einer SD-Karte

Um zu verstehen, warum diese Lösung so besonders ist und warum eigene Speicher-Produktionsanlagen hier von Vorteil sind, muss man den Aufbau und die Funktion einer SD-Karte kennen. Die beiden entscheidenden Teile sind der eigentliche NAND-Flash-Speicher und der Speicher-Controller. Der Speicher-Controller wickelt die Schreib-, Lösch- und Lesesequenzen ab, die vom Host mit standardisierten Befehlen gesendet werden. NAND-Flash selbst ist aufgrund der hohen Zell-Dichte nicht fehlerfrei. Irgendwo im Speicher wird es einige Bits geben, die sich nicht wie vorgesehen verhalten. Solche Fehler müssen erkannt und wenn möglich korrigiert werden. Außerdem beeinträchtigt jedes Schreiben und Lesen auf einen NAND-Flash die Speicherzellen, was dazu führt, dass sie mit der Zeit ausfallen. Auch dies wird vom Speicher-Controller verwaltet.

Der Speicher-Controller verhält sich wie ein Mikrocontroller (siehe das Beispiel in **Bild 1**) mit einer CPU, spezieller Hardware für den Umgang mit Flash und der Host-Schnittstelle. Als Vermittler zwischen Host und rohem NAND-Speicher verbirgt er die eigentlichen Lese- und Schreibvorgänge, behandelt schlechte Bits und kümmert sich um die Reservezellen, so dass sich die Lebensdauer der SD-Karte verlängert. Je weniger selten die Datenbits beschrieben werden, desto geringer ist ihre Ausfallwahrscheinlichkeit.

## PS-45u DP „Raspberry Edition“

Da Swissbit SD-Karten selbst entwickelt und herstellt, kann ihr Speicher-Controller nach Bedarf optimiert oder erweitert werden. Während Consumer-SD-Karten dem Host lediglich den NAND-Speicher zur Verfügung stellen, hat Swissbit zusätzliche Sicherheitselemente hinzugefügt. Die PS-45u DP „Raspberry Edition“ [2] in **Bild 2** ist eine SD-Karte mit Sicherheits-Features, die unabhängig von der Host-CPU arbeiten. Weitere Details zu den technischen Spezifikationen der Karte finden sich in **Tabelle 1**.

Der Host (hier eine Raspberry Pi) sieht einfach eine SD-Karte und kann, falls entsperrt, auf die darauf befindlichen Daten zugreifen. Derzeit funktioniert diese „Raspberry Edition“ von PS-45u DP mit den Modellen Raspberry Pi 2 und Raspberry Pi 3B+, wobei die Authentifi-

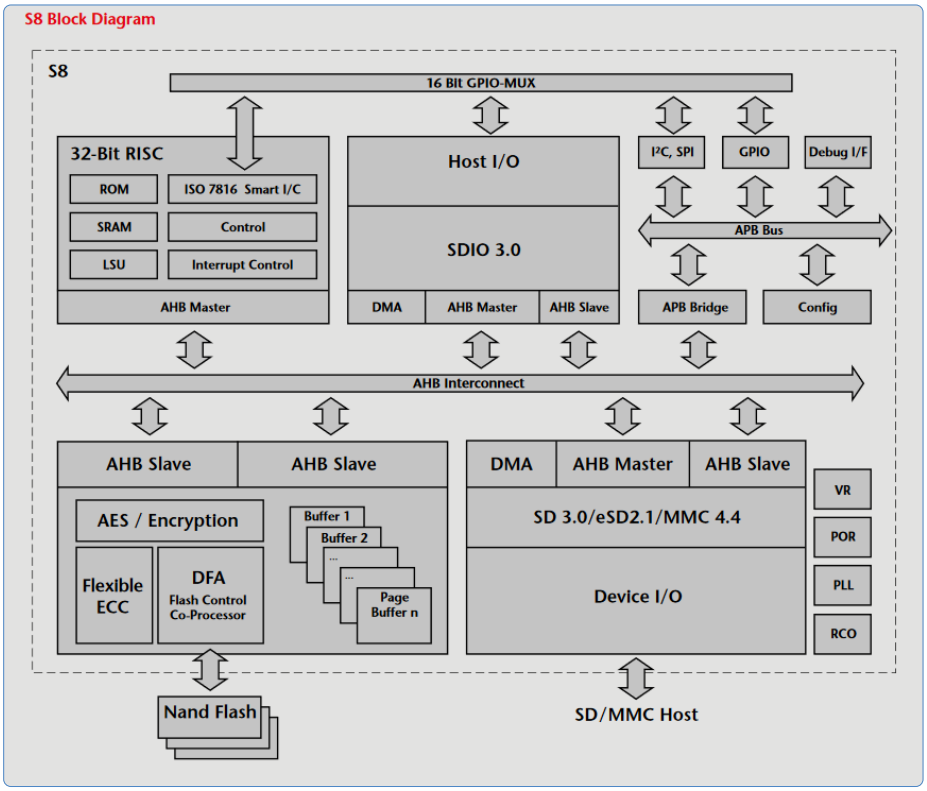



*Bild 1. Beispielhafte Implementierung eines SD-Karten-Controllers (Quelle: Hyperstone).*

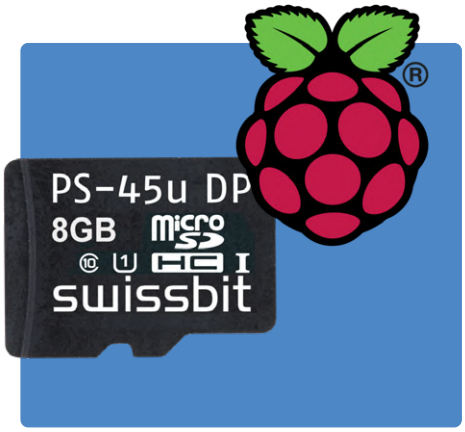



*Bild 2. PS-45u DP „Raspberry Edition" (Quelle: Swissbit).*


*Bild 3. USB-Stick PU-50n (Quelle: Swissbit).*

**Tabelle 1: Eigenschaften der SD-Karte PS-45u.**

- Größe 8 GB - 32 GB
- Sicherheitsfunktionen in Hardware
- Page-basiertes Flash-Management
- Speed Class 10
- MLC-Technologie
- Software-Tool für Authentifizierung
- Industrielle Temperaturbereiche -25°C bis 85°C und 0°C bis 70°C

**Tabelle 2: Vergleich der Authentifizierungsmethoden.**

| Security Function | PIN Policy | USB Policy | NET Policy |
|---|---|---|---|
| Know-How Protection | √ | √ | √ |
| License Protection | √ | √ | √ |
| Remote Attestation | √ | √ | √ |
| Data Protection | √ | √ | √ |
| Theft Protection | √ | | √ |
| Tamper Protection | | | √ |
| Lock Device | | | √ |
| Secure Unattended Boot | | √ | √ |

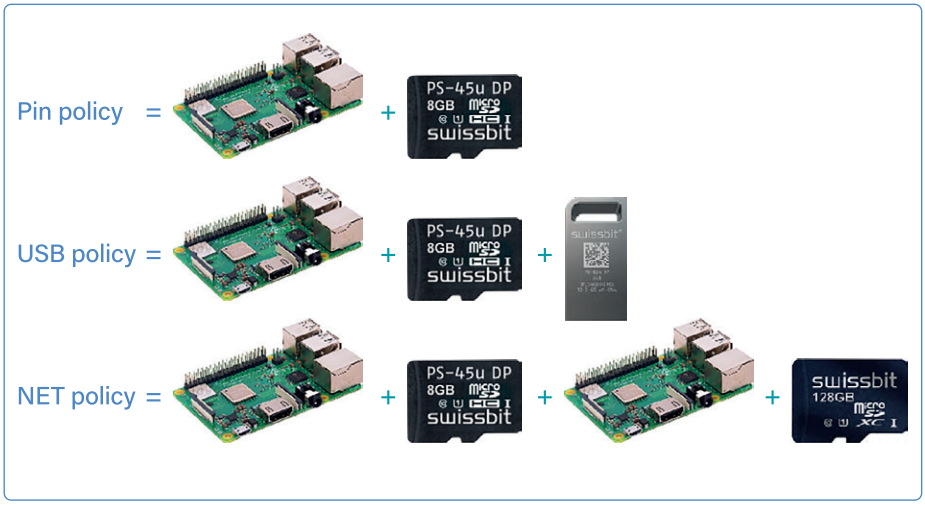



*Bild 4. Erforderliche Teile für die drei Policies.*

zierung in einer von Swissbit angepassten Pre-Boot-Phase implementiert ist, um den Zugriff auf die Daten der SD-Karte zu entsperren. Die Karte ist außerdem gehärtet, um die Datenintegrität bei Stromausfall zu gewährleisten. Wer einen Raspberry Pi mit SD-Karten betreibt, weiß wohl, dass die Karten einiger Hersteller selbst bei einem sauberen Herunterfahren aufgrund von Störungen auf den Strom- und Datenleitungen beschädigt werden können. Leider kann keine SD-Karte die Probleme verhindern, die ein unterbrochenes Schreiben des Betriebssystems auf die SD-Karte mit sich bringt. Im optimalen Fall hilft dann ein Wiederherstellungsversuch beim nächsten Hochfahren des Systems.

Die PS-45u DP „Raspberry Edition" ist mit Kapazitäten von 8 GB [3] und 32 GB [4] direkt bei Mouser Electronics erhältlich. Neben dem Plus an Robustheit bieten diese Karten auch eine integrierte Verschlüsselung und Authentifizierung. Zwecks Authentifizierung werden derzeit drei Methoden unterstützt, die in **Tabelle 2** aufgeführt sind. Dabei werden die Daten nur mit dem richtigen Zugangs-Token – einer PIN, einem USB-Token oder einem Authentifizierungsserver – freigeschaltet und zugänglich gemacht.

PIN-Policy bedeutet, dass eine vordefinierte PIN beim Hochfahren an der Kommandozeile eingegeben werden muss. Die USB-Policy arbeitet mit dem USB-Stick PU-50n [5] (**Bild 3**) als Authentifizierungs-Token [6], der ebenfalls von Swissbit erhältlich ist. Seine technische Spezifikation ist in **Tabelle 3** aufgeführt. Der USB-Stick

| Tabelle 3: Technische Daten des PU-50n. | |
|---|---|
| Standard | USB 3.1 |
| Geschwindigkeit | USB 3.1 SuperSpeed |
| Temperaturbereich | -25°C bis 85°C |
| Kompatibel mit | U-50n |
| Flash-Technologie | MLC |
| Density DP Data Protection | 8 GB - 64 GB |

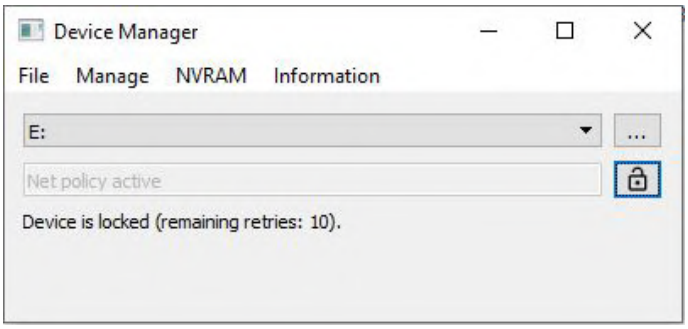



*Bild 5. SD-Karten-Konfigurationsprogramm.*

entspricht einem Lizenz-Dongle, wie er für hochpreisige CAD-Tools erforderlich ist. Die NET-Policy erfordert die Implementierung eines Authentifizierungsservers, der irgendwo im Netzwerk läuft und die Verwendung der SD-Karte authentifiziert oder verweigert.
Um ein System mit PIN-Policy zu evaluieren, benötigen Sie einen Raspberry Pi (empfohlen: Modell 3B+) und eine PS-45u DP „Raspberry Edition". Für die Evaluierung der NET-Policy benötigen Sie zusätzlich zu den Voraussetzungen für die PIN-Policy einen zweiten Raspberry Pi und eine reguläre microSD-Karte (empfohlen: S-50u) [7].
Für die Evaluierung der USB-Policy benötigen Sie zusätzlich zu den Anforderungen der PIN-Policy den USB-Stick PU-50n DP „Raspberry Edition". **Bild 4** zeigt das benötigte Zubehör zur Auswertung einer der drei Policys.

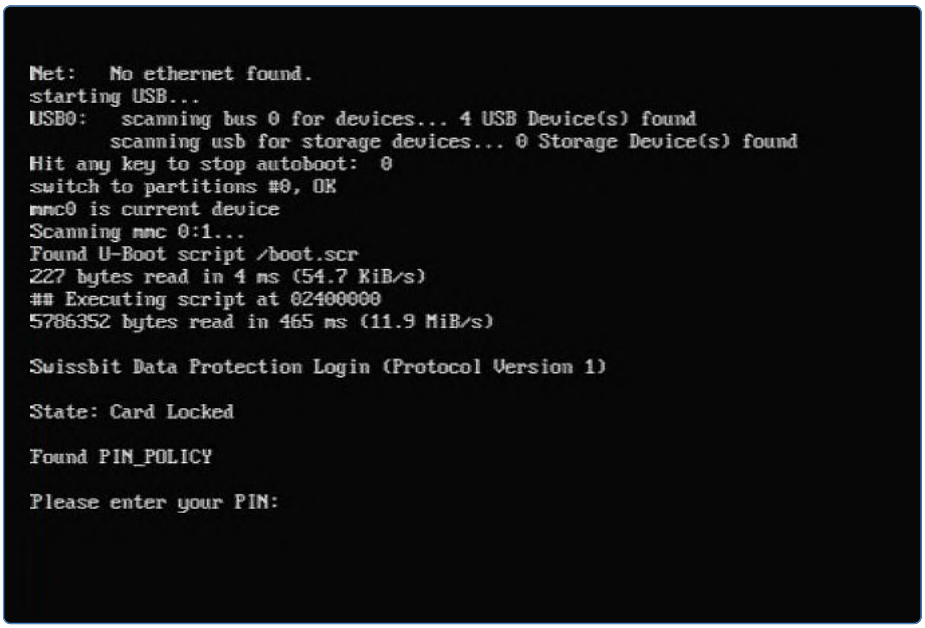



*Bild 6. Abfrage der PIN.*

## Sicherheitsnachrüstung bestehender Anwendungen

Da der Raspberry Pi beim Booten Zugriff auf einige Core-Firmware-Dateien benötigt, wurde die PS-45u DP „Raspberry Edition" mit einer besonderen Sicherheitsfunktion ausgestattet, die eine nahtlose Integration in bestehende Projekte ermöglicht. Wenn die SD-Karte gesichert ist, befindet sich die erste Partition in einem Nur-Lese-Zustand. Dies erlaubt dem Pi, die Basis-Firmware und den ersten Teil der Software zu laden, die prüft, ob ein gültiger Token vorhanden ist. Während diese Firmware der ersten Stufe von jedem gelesen werden kann, ist sie vor Änderungen geschützt und verhindert, dass ein Angreifer Zugriff auf den Rest der SD-Karte erhält. Dies schützt die Software vor unbefugten Zugriffen mit einer der folgenden drei möglichen Authentifizierungsmethoden.

### PIN-Policy

Die Verwendung der SD-Karte mit der PIN-Policy ist der einfachste Ansatz. Die SD-Karte muss anhand des SecureBoot SDK RPi Handbuchs von Swissbit konfiguriert werden. Derzeit benötigt man zur Einrichtung einer SD-Karte einen PC unter Windows (siehe **Bild 5**). Mit den mitgelieferten Tools kann eine PIN für die Karte gesetzt werden – die Verwendung des Standard-ASCII-Zeichensatzes wird dringend empfohlen. Unter Windows kann man die Karte zwar auch mit Sonderzeichen sperren und entsperren, jedoch erzwingt der Raspberry Pi beim Booten die Verwendung eines UK-Tastaturlayouts (qwerty). Dies kann bei Verwendung von Nicht-UK Tastaturen und alphanumerischen PINs zu Irritationen führen. Wenn etwa im Windows-Tool eine PIN „qwertz" erstellt wurde und der Raspberry Pi mit einer angeschlossenen deutschen Tastatur bootet, muss man zur erfolgreichen Authentifizierung „qwerty" eingeben, da „y" und „z" im Vergleich zum UK-Layout vertauscht sind.

Neben den Problemen mit dem Tastaturlayout erfordert dieser Ansatz, dass die PIN jedes Mal beim Booten eingegeben werden muss, um die SD-Karte zu entsperren und mit dem restlichen Bootvorgang fortzufahren. **Bild 6** zeigt eine PIN-Eingabeaufforderung. Brute-Force-Angriffe, bei denen viele Kombinationen ausprobiert werden, sind ein generelles Risiko bei einem solchen Authentifizierungsansatz. Dies kann die SD-Karte per Hardware verhindern. Bei der

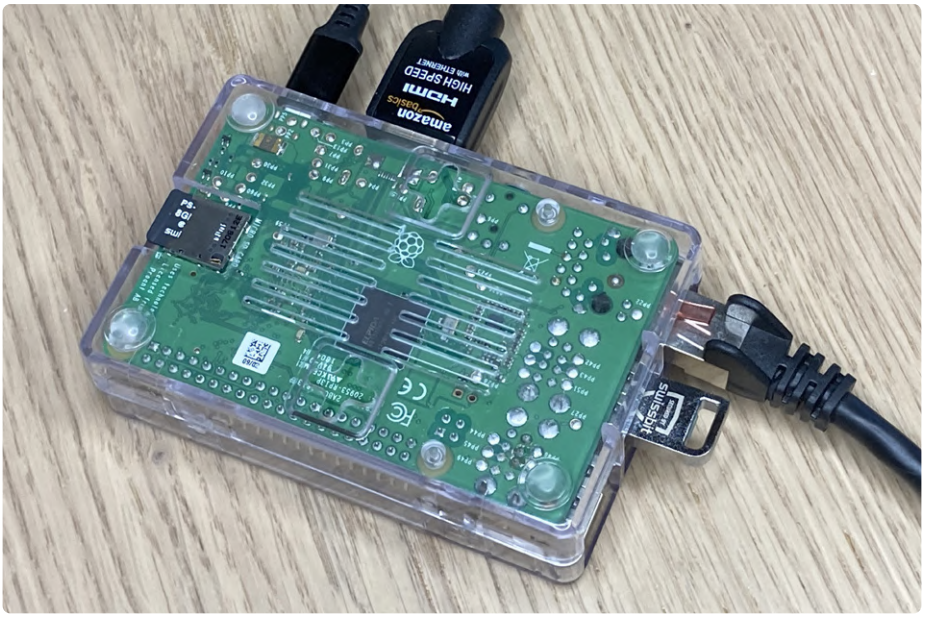


*Bild 7. Secure Boot mit einem USB-Token.*

```
Net:   No ethernet found.
starting USB...
USB0:   scanning bus 0 for devices... 5 USB Device(s) found
        scanning usb for storage devices... 1 Storage Device(s) found
Hit any key to stop autoboot:  0
switch to partitions #0, OK
mmc0 is current device
Scanning mmc 0:1...
Found U-Boot script /boot.scr
227 bytes read in 10 ms (21.5 KiB/s)
## Executing script at 02400000
5786352 bytes read in 984 ms (5.6 MiB/s)

Swissbit Data Protection Login (Protocol Version 1)

State: Card Locked

Found USB_POLICY VendorId <0x0> ProductId <0x0>

USB0:   scanning bus 0 for devices... 5 USB Device(s) found
1 Storage Device(s) found
USB0:   scanning bus 0 for devices... 5 USB Device(s) found
1 Storage Device(s) found
Kernel image @ 0x080000 [ 0x000000 - 0x584af0 ]
## Flattened Device Tree blob at 2eff8b00
   Booting using the fdt blob at 0x2eff8b00
   reserving fdt memory region: addr=0 size=1000
   Using Device Tree in place at 2eff8b00, end 2f002f0a

Starting kernel ...
```

*Bild 8. USB-Authentifizierung mit dem PU-50n.*


*Bild 9. Benutzeroberfläche des NetPolicyServers.*

Konfiguration der Karte kann die zulässige Anzahl falsch eingegebener PINs festgelegt werden, bevor die Karte weitere Versuche verweigert. Ist das Limit erreicht, kann die Karte nur noch mit einer (hoffentlich eingestellten und notierten) Supervisor-PIN entsperrt bzw. damit die PIN der SD-Karte geändert werden. Andernfalls bleibt die Karte dauerhaft gesperrt.

## USB-Policy

Die USB-Policy ermöglicht einen unbeaufsichtigten sicheren Bootvorgang. Dies funktioniert durch Hinzufügen des USB-Sticks PU-50n DP von Swissbit als Sicherheits-Token (Bild 3). Das System wird anhand des Handbuchs eingerichtet [8], wobei SD-Karte und USB-Stick mit der gleichen PIN vorbereitet werden. Auf dem PU-50n DP ist die gleiche sichere Firmware installiert wie auf der PS-45u DP. Nachdem der Raspberry Pi mit der SD-Karte bootet, findet eine Challenge-Response-Authentifizierung statt. Wenn dabei die PINs von PU-50n DP und PS-45u DP übereinstimmen, wird der Bootvorgang fortgesetzt. Ein Raspberry Pi mit diesem Setup ist in Bild 7 zu sehen. Bild 8 zeigt einen erfolgreichen Bootvorgang mit angeschlossenem USB-Stick PU-50n DP.

## NET-Policy

Während PIN- und USB-Policy eine einfache Sicherheitsnachrüstung für Standalone-Geräte ohne Netzwerkzugang ermöglichen, ist die NET-Policy für an ein Netzwerk angeschlossene Geräte gedacht. Hierzu ist ein Lizenzserver innerhalb dieses Netzwerks erforderlich, den ein Raspberry Pi mit einer PS-45u-DP-Karte zum Entsperren verwenden kann. Für den Lizenzserver wird ein vorgefertigtes Raspberry-Pi-Image bereitgestellt, mit dem eine handelsübliche 16-GB-SD-Karte beschrieben werden kann. Der Lizenzserver läuft unter Raspberry OS und verfügt über ein Webinterface zur Gerätekonfiguration. Der Setup-Prozess ist im Handbuch zum NetPolicyServer von Swissbit beschrieben [9].

Nach dem Hochfahren startet der Desktop und es kann eine Verbindung über einen Browser zu *https://localhost/NetPolicyServer/* hergestellt werden. Daraufhin wird die Konfigurationswebseite aufgerufen (Bild 9). Um ein Gerät zum Lizenzserver hinzuzufügen, wird die UID der zu aktivierenden SD-Karte zusammen mit dem zugehöri-

gen Kennwort benötigt. Diese Informationen stellt das Windows-Tool bereit, das auch für die Vorbereitung der SD-Karte verwendet wird. Laut Anleitung [8] zur SD-Karten-Vorbereitung muss eine IP-Adresse (Bild 10) eingegeben werden. Derzeit ist das System auf IPv4-Netzwerke beschränkt. Wenn Sie keine IP-Adresse für Ihren Server eingeben möchten, gibt es stattdessen eine inoffizielle Unterstützung für einen DNS-Namen, die jedoch nicht im Handbuch beschrieben ist und auf eigene Gefahr verwendet werden sollte.

Nachdem der Client konfiguriert wurde, wird bei jedem Startversuch der NetPolicyServer kontaktiert. Raspberry Pi und Server verwenden ein Challenge-Response-Schema, um die SD-Karte zu entsperren. Wenn der Server nicht die richtige Antwort liefert oder nicht verfügbar ist, wird der Boot-Vorgang abgebrochen und die SD-Karte bleibt gesperrt. Diese zentralisierte Implementierung ermöglicht die Verwaltung der Geräte vom Server aus, worüber sie deaktiviert werden können, falls sie nicht mehr zum Booten berechtigt sind. Bild 11 zeigt einen Boot-Prozess mit serverseitig falsch gesetzten Anmeldeinformationen und fehlgeschlagenem Boot-Vorgang, Bild 12 einen erfolgreichen Bootvorgang.

## Vorhandene Hardware nachrüsten

Der Vorteil beim Nachrüsten bestehender Hardware ist die Einfachheit des Vorgangs beim Austausch einer SD-Karte durch die PS-45u DP „Raspberry Edition“. Die mitgelieferten Beispiele und Dateien gehen davon aus, dass eine Linux-Version betrieben wird. Das bedeutet, dass auch eine bestehende Installation verwendet werden kann, sobald die Dateien zur Aktivierung der gewählten PS-45u-DP-Sicherheitsfunktionen gemäß Anleitung hinzugefügt wurden.
Man könnte argumentieren, dass der Raspberry Pi in der Lage sein sollte, das auch mit systemeigener Festplattenverschlüsselung zu tun, aber leider müssen auf dieser Plattform alle kryptografischen Operationen in Software ausgeführt werden. Dies würde auf einem Raspberry Pi 2 oder 3 bei Festplattenzugriffen zu einer hohen Prozessorlast führen, was Leistungsengpässe bei der Ausführung von Standardanwendungen zur Folge hat. Da die PS-45u DP „Raspberry Edition“ die

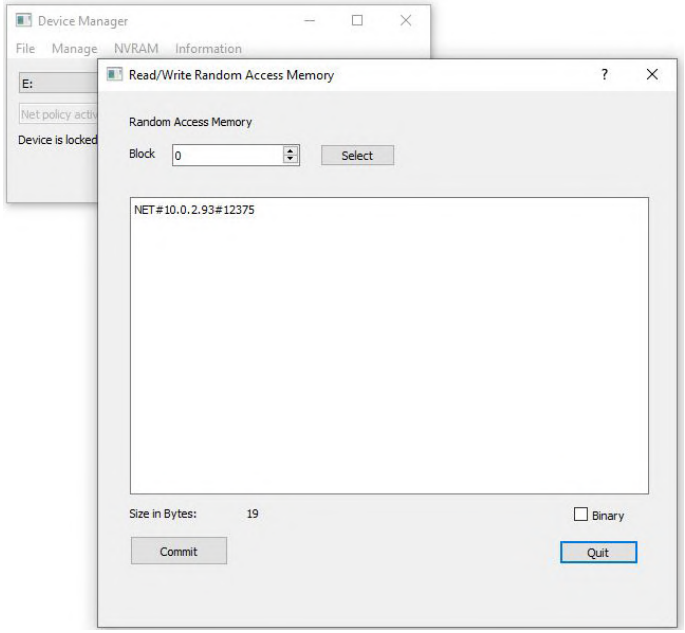



*Bild 10. Einrichtung der IP-Adresse des NetPolicyServers.*

```
Net:   No ethernet found.
starting USB...
USB0:   scanning bus 0 for devices... 4 USB Device(s) found
        scanning usb for storage devices... 0 Storage Device(s) found
Hit any key to stop autoboot:  0
switch to partitions #0, OK
mmc0 is current device
Scanning mmc 0:1...
Found U-Boot script /boot.scr
227 bytes read in 5 ms (43.9 KiB/s)
## Executing script at 02400000
5786352 bytes read in 463 ms (11.9 MiB/s)

Swissbit Data Protection Login (Protocol Version 1)

State: Card Locked


Found NET policy [redacted]:12375

lan78xx_eth Waiting for PHY auto negotiation to complete........ done
DHCP broadcast 1
net_dns_server  : [redacted]
net_netmask     : [redacted]
net_gateway     : [redacted]
Our IP address is [redacted]

IP addr for [redacted] is [redacted]

lan78xx_eth Waiting for PHY auto negotiation to complete....... done
Using lan78xx_eth device

DP login server [redacted] is alive

lan78xx_eth Waiting for PHY auto negotiation to complete....... done


Send Packet : Wait for reply from server

.Response from server recieved:

Failed to do hashed based login to card in SD card slot
unexpected return code: 0x6F02
 Please check PIN in Server
```

*Bild 11. Fehlgeschlagener Bootvorgang aufgrund einer falschen PIN.*

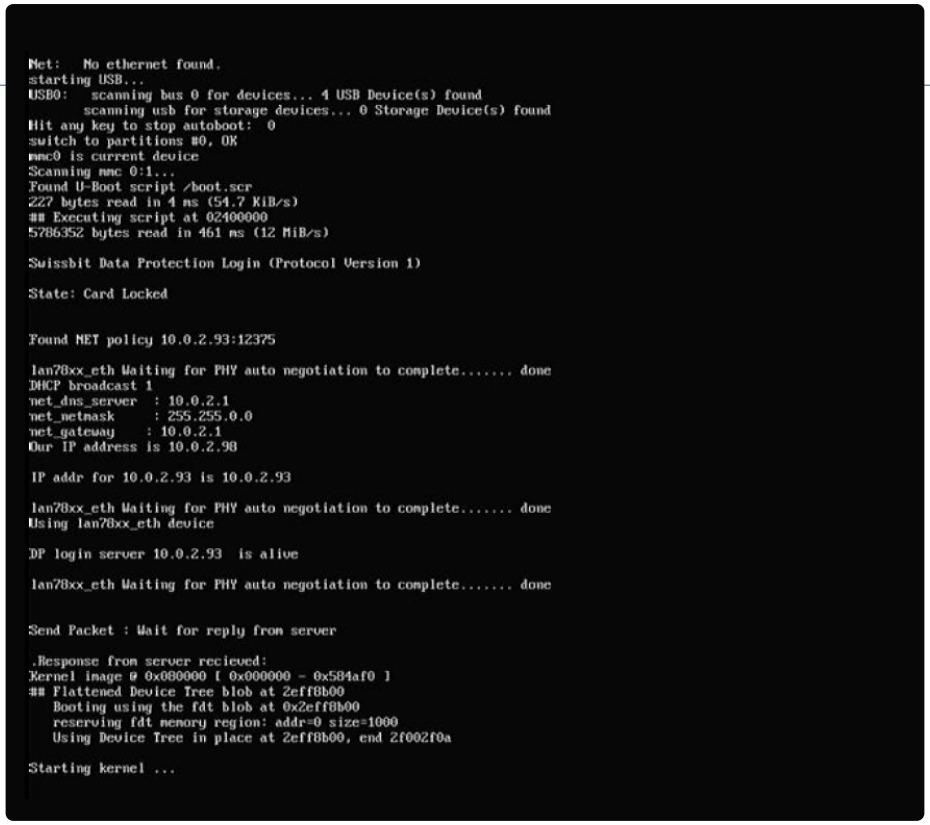



*Bild 12: Booten mit korrekter PIN.*

Verschlüsselung mit Hilfe des integrierten Controllers der SD-Karte implementiert, wird der Verschlüsselungsaufwand ausgelagert. Das hat den zusätzlichen Vorteil, dass die kryptographischen Schlüssel nicht offengelegt werden. Deshalb funktionieren die installierten Anwendungen nach dem Entsperren einer PS-45u DP ohne Leistungseinbußen, als kämen sie von einer normalen SD-Karte. Je nach Anwendung und gewählter Implementierungspolitik benötigt eine Sicherheitsumrüstung nicht sonderlich viel Zeit.

Laut Swissbit wurde die „Raspberry Edition" als praxisnahes Lern- und Evaluierungs-Kit konzipiert. Swissbit will zeigen, wie Secure-Boot-Funktionen auf Embedded-Plattformen nachgerüstet werden können, die keine native Unterstützung dafür bieten. Für produktive Anwendungen auf Basis des Raspberry Pi empfiehlt Swissbit die Verwendung des robusten Raspberry Compute Modules CM3+.

## Wünsche und Fazit

Die PS-45u DP „Raspberry Edition" unterstützt derzeit die Modelle Raspberry Pi 2 und 3, da dies die am häufigsten verwendeten SoC-Versionen sind (sowohl für die Standard- als auch die Compute-Module).

Wer Unterstützung für den Raspberry Pi 4 sucht, muss warten, bis die Softwarevoraussetzungen erfüllt sind, denn der Raspberry Pi 4 weist eine geänderte Boot-Architektur auf. Das Gleiche gilt auch für den NetPolicyServer, der eher als funktionierende Tech-Demo denn als vollwertiges, ausgefeiltes Produkt betrachtet werden kann. Es wäre schön, wenn es eine Dokumentation nebst Hintergrundinformationen gäbe, wie man einen eigenen NetPolicyServer zusammen mit zu den verwendeten Protokollen aufbaut, anstatt sich durch den Code wühlen zu müssen. Außerdem würden nicht mit Windows arbeitende Zeitgenossen PS-45u DP bestimmt gerne konfigurieren, ohne dass eine virtuelle Maschine oder ein Dual-Boot-Setup erforderlich sind.

Trotz dieser kleinen Wünsche für die nächste Generation bietet der Gesamtansatz dieser sicheren SD-Karte eine einfache Möglichkeit, Daten vor unerwünschtem Zugriff zu schützen. Der zentrale Netzwerkserver für die Authentifizierung ist ein weiterer Bonus. All das macht die Lösung von Swissbit zu einer guten Wahl für die Nachrüstung bestehender Geräte mit mehr Sicherheit. PS-45u DP schützt Ihre Daten nicht nur vor Diebstahl, sondern schränkt auch den physischen Zugriff auf das Dateisystem des Raspberry Pi ein, der es etwaigen Angreifern sonst erlauben würde, das System zu kompromittieren.

210008-02

## WEBLINKS


[1] **Swissbit**: www.swissbit.com/en/

[2] **PS-45u DP „Raspberry Edition"**: www.swissbit.com/en/products/security-technology/security-products/secure-boot/

[3] **PS-45u DP „Raspberry Edition" 8 GB bei Mouser.com**: https://bit.ly/359jFpO

[4] **PS-45u DP „Raspberry Edition" 32 GB bei Mouser.com**: https://bit.ly/3hNyAeq

[5] **PU-50n USB-Stick bei Mouser.com**: https://bit.ly/3nht3xu

[6] **PU-50n USB-Stick**: www.swissbit.com/en/products/security-technology/security-products/pu-50n/

[7] **Swissbit S-50u bei Mouser.com**: https://bit.ly/2L4xbnx

[8] **Swissbit SecureBoot SDK RPi User Manual**: www.swissbit.com/files/public/data/security/raspberry-edition/Swissbit_SecureBoot_SDK_RPi_User_Manual_v2.0.pdf

[9] **Swissbit NetPolicyServer User Manual**: www.swissbit.com/files/public/data/security/raspberry-edition/Swissbit_NetPolicyServer_User_Manual_v2.6.pdf

# Low-Power Flaggschiff für Künstliche Intelligenz

## SMARC-Module mit NXP i.MX 8M Plus Prozessor

Von **Martin Danzer** (Leiter Produkt-Management, congatec AG)

Congatec erweitert seine SMARC-Familie um ein neues Modul mit NXP i.MX 8M Plus Prozessor, der sich hervorragend für Embedded AI-Applikationen eignet. Dank des umfassenden Ökosystems mit applikationsfertigen 3,5-Zoll-Carrierboards, Basler Kameras und AI-Software-Stack sind schnelle Proofs-of-Concepts möglich.

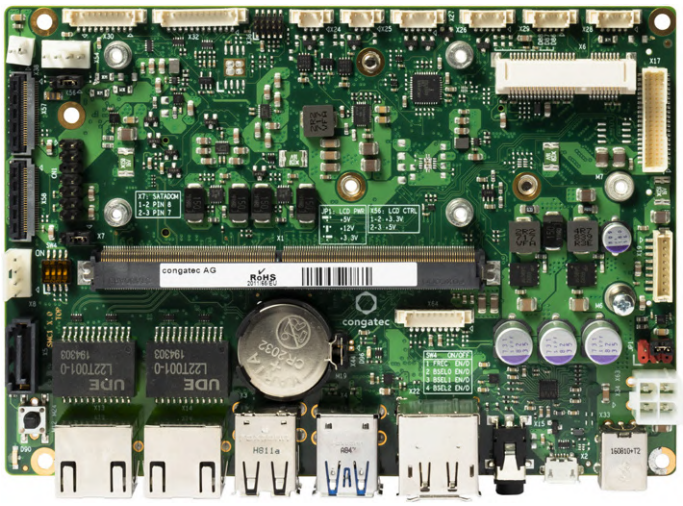


*Bild 1. Congatec Carrierboard für SMARC-Module.*

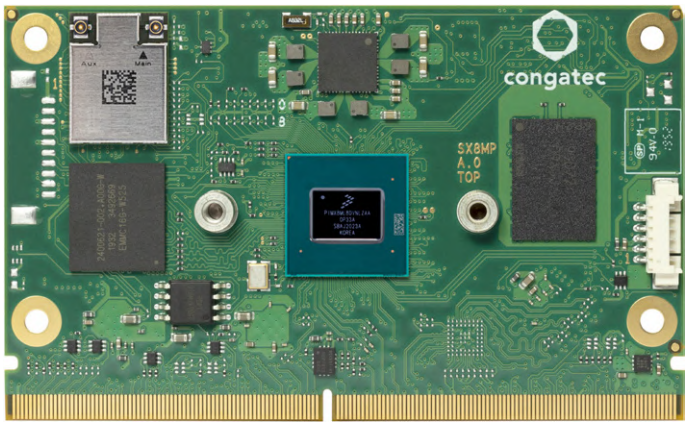



*Bild 2. Congatec SMARC-Modul mit NXP i.MX 8M Plus Prozessor.*

Das ging nicht immer so einfach. Wollte man früher nämlich neueste Prozessortechnologie aus dem ARM-Umfeld als fertiges System einsetzen, dann war dies in der Regel deutlich schwieriger als im x86-Umfeld. Dies weil das Ökosystem an System-Plattformen nicht so umfassend ist, da historisch vielfach individuell zugeschnittene Custom-Designs in großen Stückzahlen eine Rolle spielten. Mit dem modularen Ansatz mittels der SMARC Computer-on-Modules Spezifikation ist es nun jedoch möglich, Standardformfaktoren aus dem x86 Box-PC-Bereich auch mit ARM-Prozessorbestückung zu erhalten. So bietet der Embedded Computing Spezialist congatec beispielsweise ein 3,5-Zoll-Board an, das bedarfsgerecht mit Modulen seiner SMARC-2.1-Familie bestückt werden kann (**Bild 1**). Mit jedem Launch eines neuen SMARC-Moduls erweitert sich damit auch automatisch mit leichtem Zeitversatz das Portfolio an möglichen Konfigurationen. Neuester Flaggschiff-Prozessor ist der neue NXP i.MX 8M Plus Prozessor (**Bild 2**), den congatec auf SMARC offiziell zur virtuellen Embedded World 2021 launchen wird und für den der offizielle Support des 3,5-Zoll-Carriers voraussichtlich im Sommer fix und fertig abgeschlossen ist.

### Attraktive Features

Technische Highlights des neuen SMARC Moduls mit i.MX 8M Plus Prozessor sind die vier leistungsstarken ARM Cortex-A53 Prozessor-Cores mit der zusätzlichen Neural Processing Unit (NPU), die bis zu 2,3 TOPS an KI-Rechenleistung addiert. Sie sind dank integriertem Image Signal Processor (ISP) speziell für KI-Inferenzen und Machine-Learning-Anwendungen entwickelt und können sehr effizient die Daten der zwei ebenfalls integrierten MIPI-CSI Schnittstellen verarbeiten. Ein großes Anwendungsfeld sind hier gestenbasierte Steuerungen, beispielsweise um sich in medizinischen Anwendungen kontaktlos an Intensivpflege-Computern anmelden zu können, um die Gabe von Infusionen einzustellen. Die Multimedia-Funktionen umfassen zudem 3D/2D-Grafikbeschleunigung sowie Video-Decodierung und Encodierung einschließlich H.265, sodass man Kamerastreams direkt auf das Netzwerk geben kann. Hohe Auflösungen werden sowohl in Inspek-

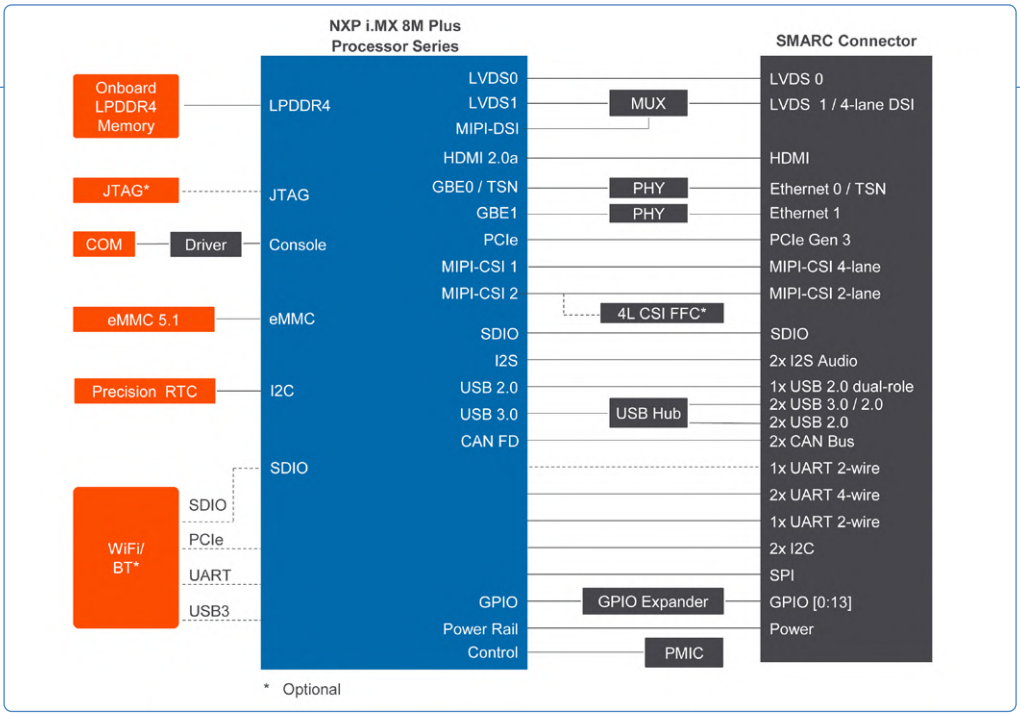



*Bild 3. Blockdiagramm des NXP i.MX 8M Plus Prozessormoduls.*

tions- als auch Überwachungssystemen immer wichtiger, um Details besser auswerten zu können. Die NPU kann hier bei der Voranalyse helfen, damit nicht die Rohdaten das Netzwerk überlasten. Interessant ist auch der hochwertige DSP für Audio- und Sprachfunktionen. Über die Kombination mit den Rechenwerken ist zum Beispiel eine nutzerspezifische Spracherkennung mit rund 40.000 unterschiedlichen Wörtern möglich. Und das, anders als bei kommerziellen Lösungen wie Alexa, Google oder Siri, rein lokal und ganz ohne Cloudanbindung, was für kontaktlose Sprachbefehle sehr nützlich sein kann.

Zudem ist auch eine Steuerung in Echtzeit über den Cortex-M7 möglich. Dies auch über Gigabit-Ethernet mit Time Sensitive Networking (TSN). Für hohe Datensicherheit und Zuverlässigkeit sorgt erstmals die in-line ECC-Funktion des Prozessors, welche auf Basis von LPDDR4-Speicher für bis zu 6 GB Kapazität Error Correction Code umsetzen kann. Der Cortec-M7 kann zudem als Fail-Safe Unit eingesetzt werden, was für fehlertolerant auszulegende Kundenapplikationen gemäß IEC 61508/IEC61511 von entscheidender Bedeutung ist. Betrachtet man dieses runde Leistungspaket und setzt weitere applikationsspezifische Funktionen über PCIe Gen 3 sowie 2x USB 3.0 und 2x SDIO um, hat man eine höchst zuverlässige und robuste Linux-, Android oder FreeRTOS Plattform, die man je nach Variante sogar im industriellen Temperaturbereich von -40 °C bis 85 °C einsetzen kann. Das von congatec primär unterstützte BSP basiert übrigens auf dem Yocto Framework. Attraktiv ist zudem vor allem auch der geringe Energieverbrauch von 2 bis 6 Watt sowie der 128 GB große onboard eMMC-Speicher, der auch im robusten pSLC-Mode betrieben werden kann, was die Lebensdauer des ohnehin schon robusten Speichers nochmals verlängert (**Bild 3**).

## Vielfältige Einsatzbereiche

Die neuen SMARC Module mit i.MX 8M Plus-Familie fokussieren damit hoch zuverlässige industrielle Applikationen in Kombination mit Embedded Vision, Machine Learning und AI sowie fortschrittliche Multimedia-Applikationen. Sie zielen damit auf Industrie 4.0 /IIoT, visuelle Inspektions- und Überwachungssysteme sowie smarte Infrastrukturen und Smart Cities ab. Weitere Anwendungsfälle finden sich aber auch in der Landwirtschaft und Bauwirtschaft sowie der Gebäudeautomation und der Messtechnik. Die adressierten Märkte sind also immens breit. Eines der größten Anwendungsfelder findet sich im breiten Spektrum der HMIs, denn nahezu jedes Gerät erhält heute eine grafische Bedienoberfläche und erweiterte Multimedia-Funktionalitäten für zunehmend immersive Erlebnisse, auch im Low-Power und Mobile Devices Bereich. Bis zu drei unabhängige Displays werden zudem unterstützt, was das neue Prozessormodul auch für Multi-Domain-Anwendungen prädestiniert. Die NXP i.MX 8M Plus Prozessorfamilie ist damit eine perfekter Ersatz für alle bestehenden i.MX6 Applikationen, die bereits in die Jahre gekommen sind. Sie skaliert zudem auch das SMARC-basierte AMD und Intel Portfolio von congatec in Richtung noch energieeffizienterer, kompakterer und mobilerer Lösungsplattformen. Hinzu kommen mobile Handheld-Devices und In-Vehikel-Applikationen. Dank der Kamera- und KI-Integration sind die Module zudem auch für eine Vielzahl unterschiedlicher Vision-Applikationen geeignet, für die congatec zusammen mit seinem Partner Basler erweitere Services bietet.

## Industrie-4.0-Applikationen mit Echtzeitvernetzung

Für die Echtzeit-Steuerung von sicherheitskritischen Applikationen und vernetzte Industrie-4.0-Systeme, wie beispielsweise Roboter oder autonome Logistikfahrzeuge, bringen die neuen SMARC-Module mit Prozessoren der i.MX8 M Plus Familie viele attraktive Features. Das beginnt beim intergierten Watchdog, den man für die Ausfallerkennung in redundant ausgelegten Maschinensteuerungen einsetzen kann. Über die Real Time Clock und den TSN (Time Sensitive Networking) fähigen Gbit-Ethernet-Port können zudem echtzeitsynchronisierte verteilte Edge-Systeme etabliert werden, beispielsweise in vernetzten Fertigungszellen oder der Prozessindustrie. Zusätzliche Sicherheit gegen Cyberangriffe, deren Risiko durch die Vernetzung steigt, bietet die NXP Trustzone eine sichere Ausführungsumgebung. Sie basiert auf einem Verschlüsselungsmodul (CAAM) für eine hardwarebeschleunigte ECC- und RSA-Verschlüsselung, dem Ressource Domain Controller (RDC) für die isolierte Ausführung kritischer Software sowie dem abgesicherten High-Assurance-Bootmodus, der die Ausführung unautorisierter Software während des Bootvorgangs verhindert.

*Bild 4. Congatec wird sein applikationsfertiges Embedded Vision Set nun auch mit dem neuen SMARC 2.1 conga-SMX8-Plus Modul mit i.MX 8M Plus Prozessoren ausrüsten.*

## Was ist mit Qseven?

Neben SMARC bietet congatec auch Qseven-Module an. Die Verfügbarkeit des NXP i.MX 8M Plus Prozessors auf diesem Formfaktor ist für das zweite Halbjahr 2021 geplant. Beide Formfaktoren unterscheiden sich nicht wesentlich. Grob kann man jedoch sagen, dass SMARC eher auf multimedialere Applikationen abzielt, während Qseven eher in tiefer eingebettenen Applikationen zum Einsatz kommt.

## Embedded-Visionsysteme sind ein Fokusmarkt

Die Verfügbarkeit der SMARC Module mit i.MX 8M Plus ist damit auch ein weiterer Meilenstein der Lösungsplattformen des Unternehmens für Embedded-Visionsysteme. In 2019 hatte congatec bereits mit Basler und NXP Semiconductors ein Proof-of-Concept für eine Retail Deep Learning Applikation erstellt. Auf dieser Plattform läuft eine integrierte KI-Software, mit dem automatisierte Kassensysteme im Einzelhandel realisiert werden können. Im Zuge der pandemiebedingt aufkommenden Kontaktlos-Technologien ist dieses Set heute aktueller denn je. Der NXP i.MX 8M Plus Prozessor bietet mit der neuen NPU und leistungsfähigem Image Signal Processor für eine noch leistungsfähigere KI-basierte Bildauswertung in Echtzeit für solche Applikationen ein noch attraktiveres Featureset (**Bild 4**). So kann der ISP Videostreams von 375 Megapixel pro Sekunde (MP/s) aufbereiten. Das entspricht 180 Full-HD Bildern pro Sekunde, bei denen beispielsweise der Bildkontrast optimiert oder Objektivverzeichnungen korrigiert werden müssen. Zusammen mit dem intergierten Neural-Prozessor sowie dem Support von OpenGL ES 3.1, Vulcan VX-Erweiterungen, OpenCL 1.2 FP und OpenVG 1.1 erhalten Vision-Applikationen eine einzigartig hohe Verarbeitungsleistung.

## Mit COM & Carrier zum kundenspezifischen Design

SMARC-Module sind – so wie alle Computer-on-Modules – für Entwickler aber auch ohne 3,5-Zoll-Board attraktiv, weil applikationsfertige Carrierboards oftmals schon vorhanden sind und das SMARC-Modul direkt auf diesen Kundensystemen eingesetzt werden kann. So entfällt praktisch die Hardwareentwicklung, was auch die Softwareentwicklung wesentlich beschleunigt und schlussendlich zu einer kürzeren Time-to-Market-Phase führt. Im weiteren Verlauf des Produktlebenszyklus kann man zudem Performancevarianten bilden. Durch den einfachen Austausch des Moduls und für extreme Langzeitverfügbarkeit kann man bei Abkündigung eines Prozessors mit geringstem NRE-Aufwand auf die nächste Generationen wechseln. So lässt sich der ROI letztendlich deutlich verlängern. Aus diesem Grund haben COMs im Markt der Embedded Computer Boards und Systeme auch eine führende Stellung eingenommen. Attraktiv bei SMARC ist zudem die Tatsache, dass dieser Standard sowohl ARM als auch x86 unterstützt. Infolgedessen werden potenzielle Migrationspfade noch breiter. Aktuell skalieren kann man SMARC-Module von congatec über 24 Modulvarianten hinweg (**Tabelle 1**). Wenn die Performance des Systems aber ausreicht, wird man die Module in professionellen Applikationen noch lange nachordern können. Congatec gibt für das neue SMARC-Modul mit NXP i.MX 8M Plus Prozessor eine erweiterte Verfügbarkeit von 15 Jahren an.

210062-02

**Tabelle 1. Alle SMARC-Module von congatec.**

| SMARC Module | ARM | x86 |
|---|---|---|
| conga-SMX8M-Plus  | i.MX 8M Plus Quad<br>i.MX 8M Plus Quad mit NPU | – |
| conga-SMX8  | NXP i.MX8 Quad Max<br>NXP i.MX8 QuadPlus | – |
| conga-SMX8X  | NXP i.MX8 QuadXPlus<br>NXP i.MX8 DualXPlus | – |
| conga-SMX8-Mini  | NXP i.MX 8M Mini Quad (industrial)<br>NXP i.MX 8M Mini Dual (industrial)<br>NXP i.MX 8M Mini Solo (industrial) | – |
| conga-SA7  | – | Intel® Atom® x6425E<br>Intel® Atom® x6413E<br>Intel® Atom® x6211E<br>Intel® Atom® x6212RE<br>Intel® Atom® x6425RE<br>Intel® Atom® x6414RE<br>Intel® Pentium® J6425<br>Intel® Celeron® J6413 |
| conga-SA5  | – | Intel® Atom™ x7-E3950<br>Intel® Atom™ x5-E3940<br>Intel® Atom™ x5-E3930<br>Intel® Celeron® N3350<br>Intel® Celeron® J3455<br>Intel® Pentium® N4200 |

# Ein **roter Würfel** und seine Kollegen

## REDCUBE-Terminals für den Anschluss hoher Ströme

Ein Beitrag von **Würth Elektronik**

Würth Elektronik bietet mit den REDCUBE-Terminals eine Produktgruppe von Hochstromkontakt-Lösungen in verschiedenen Ausführungen an. Ihren Namen verdanken sie einer innovativen Variante mit Schnellverriegelungssystem, die tatsächlich als roter Würfel die Firmenfarbe des Herstellers auf Leiterplatten bringt. Bei näherer Betrachtung stehen auch die weniger farbigen Modelle der Reihe ihrem roten Kollegen in puncto Funktionalität in nichts nach.

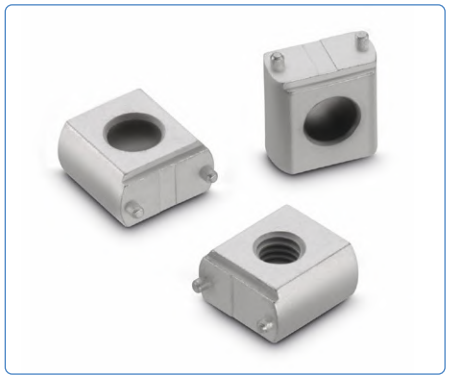


*Bild 1. Die rechtwinkligen SMD-Terminals sind für Nennströme bis maximal 50 A ausgelegt und benötigen in der Bauform M3 nur eine geringe Lötfläche.*

Wenn es um die Verbindung von Leistungssteckverbindern mit der Leiterplatte geht, haben viele zwei Möglichkeiten vor Augen: die klassische, automatisierte SMT-Bestückung und die manuelle THT-Bestückung. Warum gerade die Oberflächenmontage (Surface Mount Technology) und die Durchsteckmontage (Through-Hole-Technology) so bekannt und beliebt sind, hat gute Gründe. Bei Anschlüssen im Kleinspannungs- und mA-Bereich ist und bleibt die SMT-Bestückung die effektivste und kostengünstige Methode (**Bild 1**). Die Bauteile werden im vollautomatischen Bestückungsprozess per Pick & Place auf die bereits mit Lötpaste versehende Leiterplatte gebracht. Im anschließenden Reflow-Prozess wird das Lot verflüssigt – daher auch der Name Reflow – um die Komponenten mit der Leiterplatte zu verbinden. Das Besondere an diesem Verfahren ist, dass sich beide Seiten der Leiterplatte in ein und demselben Prozess verarbeiten lassen.

War die SMT-Bestückung lange Zeit dem Bereich niedriger Ströme vorbehalten, hat sie sich inzwischen auch im Mittelfeld etabliert. So sind Wire-to-Board-Steckverbinder mit Strömen um 5 A mittlerweile Standard. Auch bei Terminal Blocks bis 10 A sind SMT-Anschlüsse auf dem Vormarsch. Und was ist kabelseitig bei Stromstärken über 10 A zu beachten? Eine höhere Stromstärke bedeutet eine größere Wärmeentwicklung, der man mit einem größeren Leitungsquerschnitt begegnen muss. Das wiederum heißt, dass die mechanische Belastung am Leiterkartenanschluss steigt. Und dieser kann die SMT-Verbindung nicht standhalten. Deshalb steigen hier viele auf die Durchsteckmontage um.

### Das Eckige muss in das Runde

THT-Steckverbinder sind bedrahtete Bauteile, haben also kleine Anschlussbeinchen, die später per Handbestückung durch eine Durchgangsbohrung in der Leiterplatte gesteckt und mittels Wellenlöten mit der Platine verbunden werden. Die THT-Bestückung erfolgt nach der SMT-Fertigung. Somit sind meist beide Seiten der Platine bereits bestückt und die Gefahr, bereits aufgelötete Bauteile bei der Wellenlötung zu beschädigen, ist einfach zu groß. Daher muss hier sehr oft auf Handlötung bzw. Selektivlötung zurückgegriffen werden. Diese zusätzlichen Fertigungsschritte verursachen entsprechend zusätzliche Kosten in der Produktion. Doch muss man diesen Zeit- und Kostenfaktor tatsächlich immer in Kauf nehmen, wenn die Stromstärken steigen? Nein!

Wer bis zu 85 A (bei 20° C) auf die Leiterplatte bringen möchte, der kann mit der Produktfamilie REDCUBE SMT bei der klassischen SMT-Bestückung bleiben, um Einzelleitungen an die Leiterkarte anzuschließen. Und auch diejenigen, die für kleine und mittlere Ströme nicht auf die mechanische Belastbarkeit von THT-Bauteilen verzichten möchten, müssen nicht vom Reflow-Prozess abweichen. Leistungssteckverbinder mit THR-(Through Hole Reflow-)Anschlüssen lassen sich zusammen mit den SMT-Komponenten im selben Reflow-Prozess verarbeiten. Beim Reflow-Löten sind die Bauteile im Vergleich zum Wellenlöten jedoch höheren Temperaturen ausgesetzt. Deshalb müssen sowohl Steckverbinder als auch Stiftkontakte hinsichtlich Konstruktion und Material an die THR-Bestückung angepasst werden. Die konstruktiven Anpassungen sind auf der Gehäuseunterseite zu finden: Abstandsstifte schaffen Platz für die Lötpaste und fördern den Luftfluss unter dem Gehäuse, damit sich die Löttemperatur gleichmäßig in der Leiterplatte verteilen kann. Und auch die Kontaktstifte müssen für die THR-Technik die richtige Länge aufweisen, damit weder die Stabilität des Kontaktstifts in der Durchstecköffnung beeinträchtigt noch Lötpaste durch die Durchgangsbohrung gedrückt wird und Lötperlen verursacht. All diese Anpassungen und auch die Zusatzkosten für die Tape & Reel-Verpackung führen unweigerlich zu einem höheren Preis der THR-Elemente. Doch deshalb auf die herkömmliche THT-Bestückung umsteigen? Vor allem wenn es sich nur vereinzelt um THT-Komponenten handelt, ist der THR-Prozess in der Gesamtkalkulation merklich günsti-

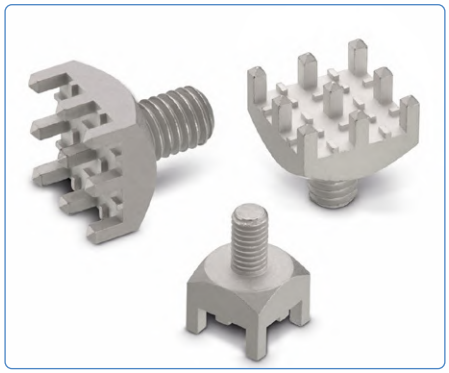


*Bild 2. Statt der am Markt bekannten Bügelelemente verfügen die REDCUBE THR Terminals über ein spezielles und massives Pin-Design, welches ein perfektes Lötergebnis. eine hohe mechanische Stabilität und deutlich bessere Drehmomente bietet.*

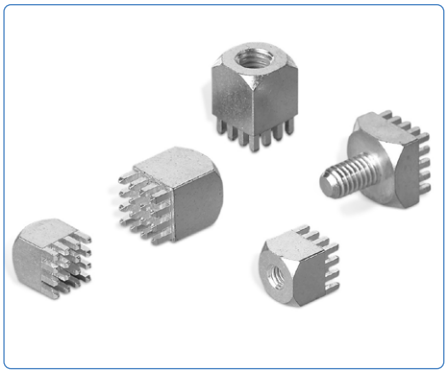


*Bild 3. PRESS-FIT Terminals vereinen sehr hohe Stromtragfähigkeit mit einer extrem hohen Langzeitzuverlässigkeit und lassen sich dennoch nahtlos und somit kostengünstig in den Bestückungsprozess einreihen.*

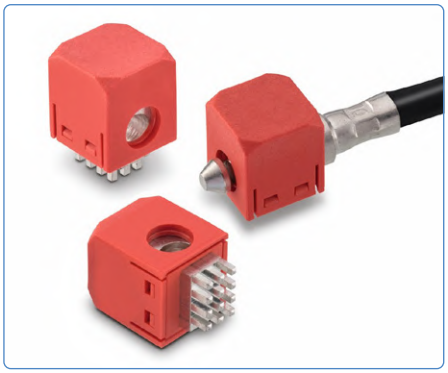


*Bild 4. Vier unterschiedliche Steckanschlüsse von 50 bis 120 A bietet die Reihe REDCUBE PLUG für die einfache und schnelle Wire-to-Board-Hochstromverbindung.*

ger. Darüber hinaus lassen sich mit diesem Verfahren deutliche Platzeinsparungen auf der Platine erreichen.

Ob Stapler, Aufzug oder einfach nur der Ladeanschluss einer Batterie – eine hohe Vibrationsbeständigkeit ist bei Leistungssteckverbindern in Anwendungen, bei denen sie regelmäßig Bewegungen oder vielen Steckzyklen ausgesetzt sind, ein Muss. Die Leistungsanschlüsse REDCUBE THR eignen sich deshalb so besonders für solche Anwendungen, weil ihre Kontaktstifte im Vergleich zu denen anderer Leistungsteckverbinder nicht flexibel sondern starr ausgeführt sind (**Bild 2**). Übrigens spielt auch die Stiftgeometrie beim THR-Vorgang eine wichtige Rolle, denn diese muss bei den Berechnungen des Lötpastenvolumens berücksichtigt werden. Zur Ermittlung des benötigten Lötpastenvolumens gibt es eine einfache Formel:

(Volumen der Durchgangsbohrung – Pin-Volumen im Via) x 2

Der Faktor 2 ist nötig, da lediglich 50% der Lötpaste metallisch ist und Zusätze wie z. B. Flussmittel im Prozess verschwinden. Optimal ist ein Füllgrad von 100% - nach IPC-A-610D ist eine mehr als 75% gefüllte Lötstelle akzeptabel.

## Einpressen statt Löten

Doch wie bringt man Anschlüsse auf eine Leiterplatte, von denen eine noch höhere Stromtragfähigkeit gefordert wird und deren Verbindung mechanisch viel stabiler und über lange Zeit zuverlässig sein muss, als es eine Lötverbindung jemals leisten kann, zum Beispiel in Windkraftanlagen? Hier gesellt sich mit der Einpresstechnik ein weiteres Verfahren dazu.

Eine leistungsfähige Einpressverbindung wird hergestellt, indem ein Einpresspin in das durchkontaktierte Loch einer Leiterplatte eingepresst wird (**Bild 3**). Dieser Vorgang erzeugt eine hohe Reibung, die zu einer homogenen Kaltverschweißung an den Anbindungsstellen zwischen Pin und Durchkontaktierung führt. Das Ergebnis ist eine gasdichte, starke mechanische Verbindung mit einem Übergangswiderstand von weniger als 200 µΩ.

Dabei werden die Durchkontaktierungen grundsätzlich auf dieselbe Weise hergestellt wie die Löcher zur Aufnahme von THT-Bauelementen. Der entscheidende Unterschied zur Löttechnik ist, der Einpresspin kontaktiert direkt das Kupfer der Durchkontaktierung. Somit erreicht eine Einpressverbindung einen bis zu vierfach niedrigeren Übergangswiderstand gegenüber einer Lötstelle. Das Lötzinn selbst hat in der Regel einen bis zu 20-fach schlechteren elektrischen Leitwert als Kupfer. Im Vergleich zu anderen Komponenten zur Stromübertragung hat REDCUBE PRESS-FIT somit bei gleichen Stromwerten die niedrigste Wärmeentwicklung.

Des Weiteren ist beim Löten nicht immer gewährleistet, dass das in die Löcher eingebrachte Lot auch die komplette Hülse hochsteigt. Dadurch entstehen noch höhere Übergangswiderstände und eine geringe mechanische Festigkeit vor allem bei Belastungen mit Drehmomenten oder Vibration. Bezüglich der Langzeitzuverlässigkeit überzeugt REDCUBE PRESS-FIT mit dem niedrigsten Failure-in-Time-Wert des Gesamtsystems. Dieser ist bis dreißig Mal besser als der einer SMT-Lötstelle.

Ein weiterer Vorteil gegenüber der Löttechnik ist, dass sich Leiterplatten mit einer Dicke von bis zu 3,2 mm und hoher Kupferbelegung einfach verarbeiten lassen. Um die Vorzüge der Verbindung durch Einpresstechnik einer Vielzahl von Anwendungen zugänglich zu machen, bietet Würth Elektronik die REDCUBE PRESS-FIT Terminals aus Messing mit verzinnter Oberfläche in verschiedenen Varianten an: Innengewinde (zweireihig, umlaufend oder vollflächig), Außengewinde vollflächig, 90° zweireihig oder vollflächig sowie zweiteilig als Grund- und Support-Element für IGBT- und BTB-Anbindung.

Für alle, die schnell und einfach eine Wire-to-Board-Hochstromverbindung herstellen möchten, bietet Würth Elektronik mit dem REDCUBE PLUG den jüngsten Zugang der Serie (**Bild 4**). Auch dieser Anschluss kommt mit allen Vorteilen der Einpresstechnik und völlig ohne Unterlegscheiben und Muttern/Schrauben aus. Der kleine rote Würfel aus glasfaserverstärktem Kunststoff ist in der Lage, vier unterschiedliche Steckanschlüsse von 50 bis 120 A aufzunehmen. Die schraublose Verbindung wird hergestellt, indem das Gehäuse von oben manuell betätigt und der zugehörige Stecker eingeführt wird. Mittels Federkraft wird der Stecker automatisch im Gehäuse verriegelt. Auf diese Weise lässt sich die Verbindung auch wieder lösen. Das Einpressen der REDCUBE PLUG erfolgt nach demselben Prinzip wie bei REDCUBE PRESS-FIT. Kabelseitig werden Cable Connector und Kabel, mit Querschnitten von 4, 6, 10 oder 16 mm², über einen Hexagonal Crimp miteinander verbunden. Diese lösbare, mehrfach steckbare Wire-to-Board- und Board-to-Board-Verbindung eignet sich besonders bei der Montage auf kleinem Raum oder an schwer zugänglichen Stellen, da kein Werkzeug benötigt wird. Für den Einsatz in Ladegeräten können die REDCUBE PLUG in anderen Farben zur leichteren Unterscheidbarkeit geliefert werden.

210067-02

# 0xDEADBEEF: Pufferüberläufen auf der Spur

## Software-Bugs mit Daten-Köder jagen

Von **Stuart Cording** (Elektor)

Software-Entwicklung auf Mikrocontrollern kann eine echte Herausforderung sein. Manchmal funktioniert der Code einfach nicht richtig, und Sie finden einfach den Fehler nicht. Dann ist es angebracht, die Herausforderung wie Columbo, Jessica Fletcher oder Ihr Lieblingsdetektiv anzugehen, investigative Nachforschungen anzustellen und den Übeltäter festzunageln. Und dafür müssen Sie vielleicht totes Rindfleisch als Köder auslegen: 0xDEADBEEF.

Manchmal funktioniert Code scheinbar grundlos nicht so wie er soll. Vielleicht ändern sich unerwartet einige Daten und das Programm friert einfach ein oder der Mikrocontroller setzt sich unerwartet zurück. Zunächst sollte man sich selbst eingestehen, dass es sich mit höchster Wahrscheinlichkeit um ein Problem in der Software handelt. Mikrocontroller tun schließlich nur das, was ihnen befohlen wird. Folglich sind wohl wir schuld an ihrem widerspenstigen Verhalten. Solche Probleme können auch entstehen, wenn man nicht ganz versteht, wie Compiler und Linker funktionieren.

### Den Fehler erneut provozieren

Falls der Fehler nur sporadisch auftritt, sollte man als ersten Schritt herausfinden, wie sich der Fehler provozieren lässt. Nur wiederholtes Auftreten des Fehlers ermöglicht die Durchführung weiterer Tests. Nur so lässt sich sicher feststellen, ob man den Fehler sicher zum Verschwinden bringen kann. Nur wenn die Ursache klar ist, können Sie sicher sein, dass die Änderung den Fehler tatsächlich behoben hat – und nicht lediglich das Problem an eine andere Stelle verlagert hat! Es ist verlockend, diesen Prozess mit „Testcode" und Variablen zu unterstützen. Wenn dies aber den Fehler „verschwinden" lässt, könnte es sein, dass ein Speicherproblem etwa als Folge eines Buffer- oder Stack-Überlaufs den Fehler verursacht.

Puffer werden bei Mikrocontrollern in C typischerweise als Arrays mit fester Länge implementiert. Seltener aber ebenfalls möglich ist ein zur Laufzeit mit „alloc" (via malloc oder `calloc`) alloziertes, dynamisch zugewiesenes Array. So oder so ist es mit C sehr einfach „machbar", dass Code auf Speicher außerhalb der definierten Grenzen eines Arrays zugreift. Dann aber können beim Schreiben in ein Array versehentlich benachbarte Variablen überschrieben werden.

### Mit einem Köder in die Falle locken

Um zu überprüfen, ob Arrays über ihre Grenzen hinaus beschrieben werden, kann man sie vor Verwendung mit einem bekannten Muster füllen. Da das Hexadezimalsystem die Buchstaben A bis F enthält, lassen sich damit einfache und leicht erkennbare Wörter bilden. Ein Beispiel wäre `0xDEADBEEF`.

Das Vorhaben lässt sich auch direkt mit der IDE und dem verwendeten Debugger realisieren (**Bild 1**). Zunächst wird die Start- und Endadresse des untersuchten Arrays ermittelt. Dann ist es normalerweise möglich, mit Hilfe einer Speicheransicht ein auffälliges Muster von der Start- bis zur Endadresse in das RAM zu schreiben.

Der beste Ansatz für diese Debugging-Methode ist, den Code auf den Mikrocontroller zu laden und ihn laufen zu lassen, bis `main()` erreicht ist. Hierzu kann man einen Breakpoint bei `main()` setzen oder eine „run to cursor"-Funktion verwenden. Nach der Bestimmung von Start- und Endadresse des Arrays wird das Muster an diese Stelle geschrieben und dann der weitere Code eine Weile ausgeführt. Nach dem Anhalten der Code-Exekution erfolgt die Prüfung der mit dem `0xDEADBEEF`-Muster gefüllten Speicherstelle. Wenn das Muster nicht mehr auftaucht, wurden wohl irgendwann Daten über das Ende des Arrays hinaus geschrieben. Jetzt muss nur noch festgestellt werden, wann und warum dies geschah.

Hinweis: Die verwendete IDE erlaubt möglicherweise nicht das Schreiben des vollständigen `0xDEADBEEF`-Musters. Mit MPLAB X konnten in den Beispielen hier lediglich 16-Bit-Muster in den PIC24F-Speicher geschrieben werden (da es ein 16-Bit-Mikrocontroller ist). Aber es geht auch kürzer: ein Muster wie `0xBEEF` ist ebenfalls gut zu erkennen. Alternativ lässt sich das Muster auch in eine ASCII-Datei schreiben und von dort in den Speicher laden.

### Hat sich der Speicher verändert?

Eine weitere hilfreiche Funktion vieler Debugger ist die Hervorhebung geänderter Speicherstellen in einer bestimmten Farbe (z.B. rot), wenn die Code-Ausführung angehalten wird. Ein Stopp des Mikrocontrollers passend zum untersuchten Speicherbereich zeigt dann die geänderten Bytes in der Highlight-Farbe an (**Bild 2**). Dies ist besonders dann hilfreich, wenn Funktionen oder Bibliotheken von Drittanbietern aufgerufen wurden, die möglicherweise einen Buffer Overflow verursacht haben.

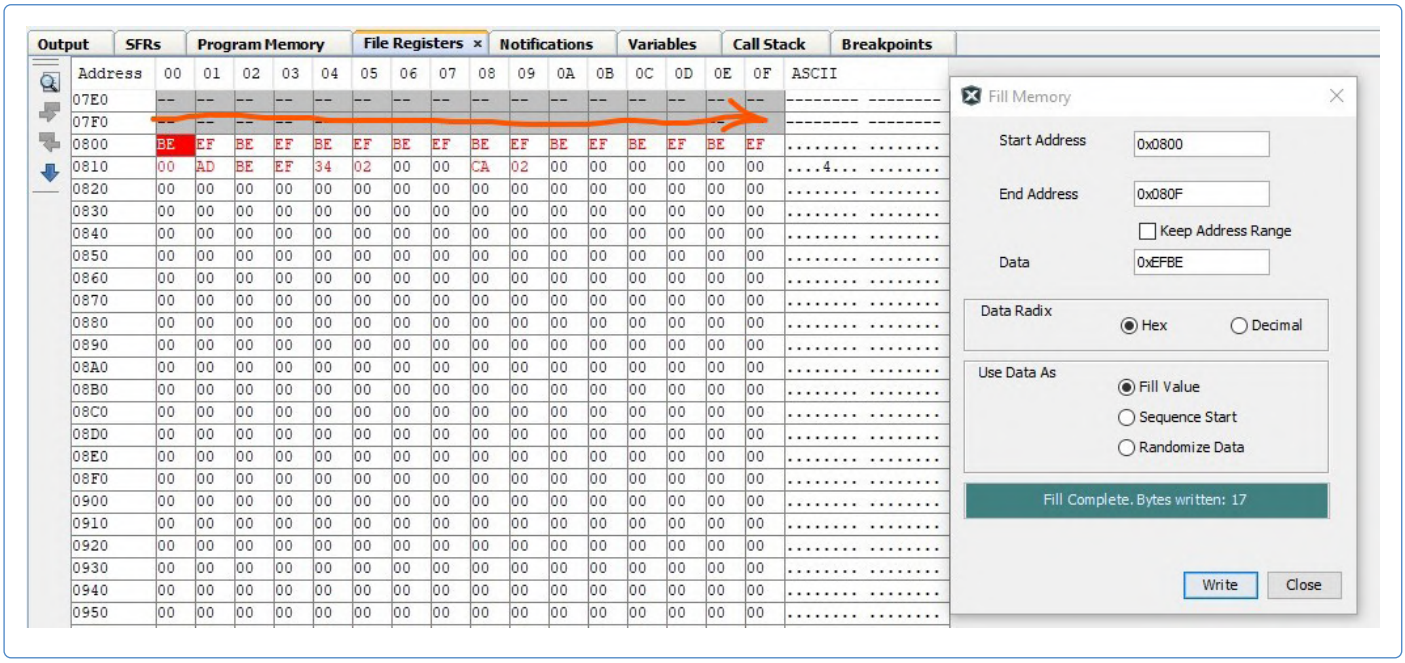



*Bild 1. Mit der Entwicklungsumgebung (hier MPLAB X IDE) kann ein RAM-Bereich mit einem Muster gefüllt werden.*

Wenn das `0xDEADBEEF`-Muster komplett verschwunden ist, wird es schwierig zu beurteilen, wie weit über die Grenzen des Arrays hinaus geschrieben wurde. In solchen Fällen generiert man zusätzliche Arrays als „Puffer" vor und nach dem untersuchten Array (**Bild 3**). Dies ist in C/C++ etwas knifflig, da sich ohne zusätzlichen Code nicht die Adresse definieren lässt, an dem Variablen alloziert werden. Man kann aber mit einiger Zuversicht darauf vertrauen, dass direkt hintereinander definierte Variablen auch im Speicher direkt hintereinander abgelegt werden. Wenn man also je ein „Extra-Puffer-Array" vor und nach dem zu untersuchenden Array definiert, werden diese wahrscheinlich auch direkt davor und danach im Speicher alloziert. Anschließend füllt man einfach alle Arrays mit `0xDEADBEEF`, führt den Code aus und prüft dann, ob das Muster in den Arrays davor und danach noch intakt ist.

## Debuggen von Echtzeit-Code

Viele Echtzeitanwendungen etwa für Motorsteuerungen oder USB- und Ethernet-Schnittstellen können nicht angehalten und zeilenweise

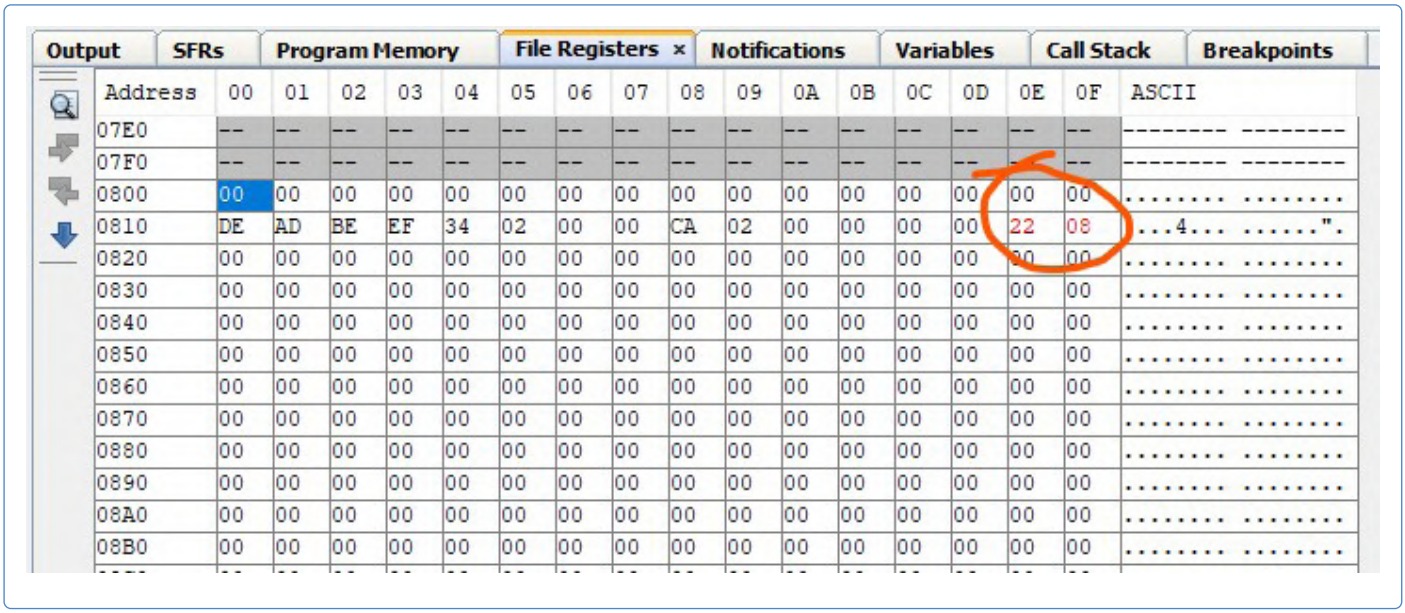



*Bild 2. Seit dem letzten Stopp der MCU hat sich die orange markierte, rote Speicherstelle verändert. Schwarze Werte sind unverändert. Wenn sich diese Speicherstelle unerwartet verändert haben sollte, sind Sie dem Softwarefehler wohl schon auf die Spur gekommen.*

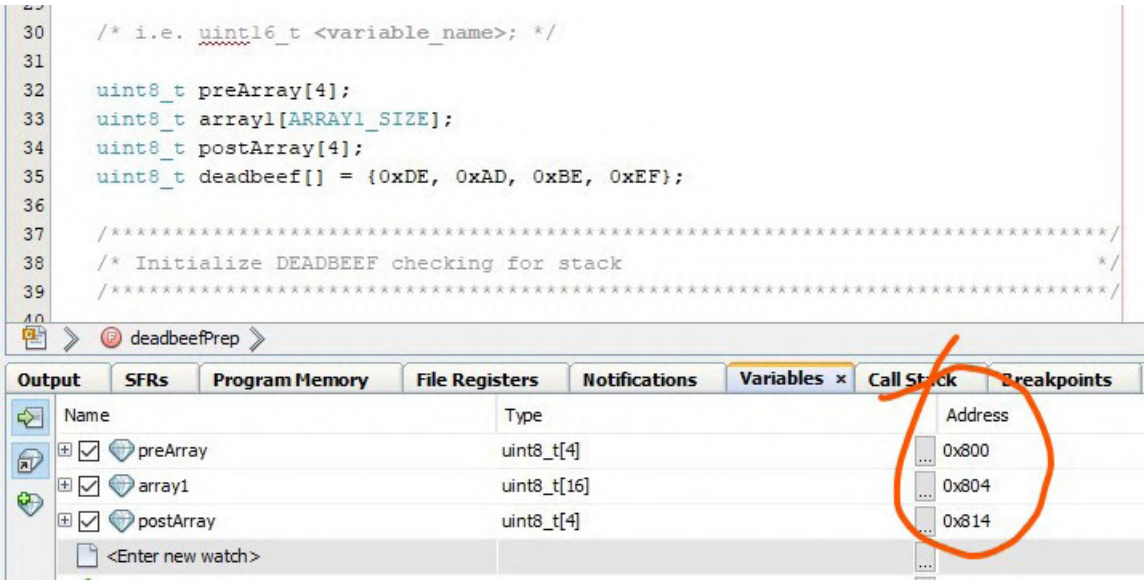



*Bild 3. Je ein zur Unterstützung der Fehlersuche definiertes extra Array befindet sich vor und nach dem untersuchten Array. Der Linker kann auch gezwungen werden, bestimmte Adressen zu verwenden, falls das nicht automatisch passiert.*

debuggt werden, da sie sonst nicht mehr funktionieren. Hier lässt sich dann auch nicht einfach per IDE die Programmausführung an der richtigen Stelle anhalten, um das des `0xDEADBEEF`-Musters in die Puffer zu schreiben. Stattdessen muss das in Software erfolgen. Das ist aber nicht weiter schwierig: Man verwende schlicht ein bereits mit `0xDEADBEEF` gefülltes Array und etwas Code, der diese Bytes in das/die zu untersuchenden Array(s) kopiert (**Bild 4**).

```
/* Example code that writes 0xDEADBEEF pattern into an
array on a Microchip PIC24F microcontroller */

#define ARRAY1_SIZE 16
 uint8_t deadbeef[] = {0xDE, 0xAD, 0xBE, 0xEF};
 uint8_t array1[ARRAY1_SIZE];

 int16_t main() {
    uint8_t x = 0;
    uint8_t y = 0;

    while(1)
    {
        y = 0;
        for (x = 0; x < ARRAY1_SIZE; x++) {
            array1[x] = deadbeef[y];
            y++;
            if (y >= 4) {
                y = 0;
            }
        }
    }
 }
```

Nun kann der Code laufen, bis er die Ausführung aufgrund eines Fehlers stoppt oder bis man den Mikrocontroller selbst anhält. Anschließend lässt sich der Speicher mit Hilfe der IDE analysieren. Selbst wenn der Fehler einen Reset der MCU verursacht hat, wird der Speicher bei weiter eingeschalteter MCU nicht gelöscht – es sei denn, ein Start-Code löscht diese Speicherstellen. Man kann also auch nach einem Reset noch per Debugger auf das RAM in dem Zustand zugreifen, in dem es sich vor dem Reset befand. Es lässt sich also auch dann noch nach Pufferüberläufen suchen, welche `0xDEADBEEF`-Muster überschrieben haben.

## Füllen des Stacks mit 0xDEADBEEF vor main()

Ein weiteres mögliches Problem ist ein Stack-Überlauf. Viele Mikrocontroller haben eine Hardware-Erkennung hierfür. Fehlt solch eine Funktion oder wurde die RAM-Verwendung anders definiert, kann der Stack wie bereits erwähnt via IDE gefüllt werden. Es kann jedoch sein, dass dieser Platz schon bevor der Code ausgeführt bzw. `main()` erreicht wird mit dem `0xDEADBEEF`-Muster gefüllt werden muss. Mikrocontroller haben vor dem Erreichen von `main()` mehr zu tun, als hier behandelt werden könnte. Über Hooks lässt sich aber Anwendercode ausführen, bevor die CPU zu `main()` springt. Dies kann dazu genutzt werden, den Stack mit `0xDEADBEEF` zu füllen.
Das konkrete Vorgehen hängt zwar vom verwendeten Mikrocontroller ab, basiert aber auf dem folgenden generellen Ablauf:

- Bestimmung des Stack-Anfangs.
- Bestimmung des Stack-Endes.
- Füllung dieses Bereichs mit `0xDEADBEEF`.

Das in C gehaltene Beispiel in Listing 1 für den 16-Bit-Mikrocontroller PIC24F implementiert die Funktion `user_init` (in der Compiler-Dokumentation [2] finden sich weitere Details). Diese Funktion wird noch vor `main()` aufgerufen, sobald der Chip initialisiert ist.

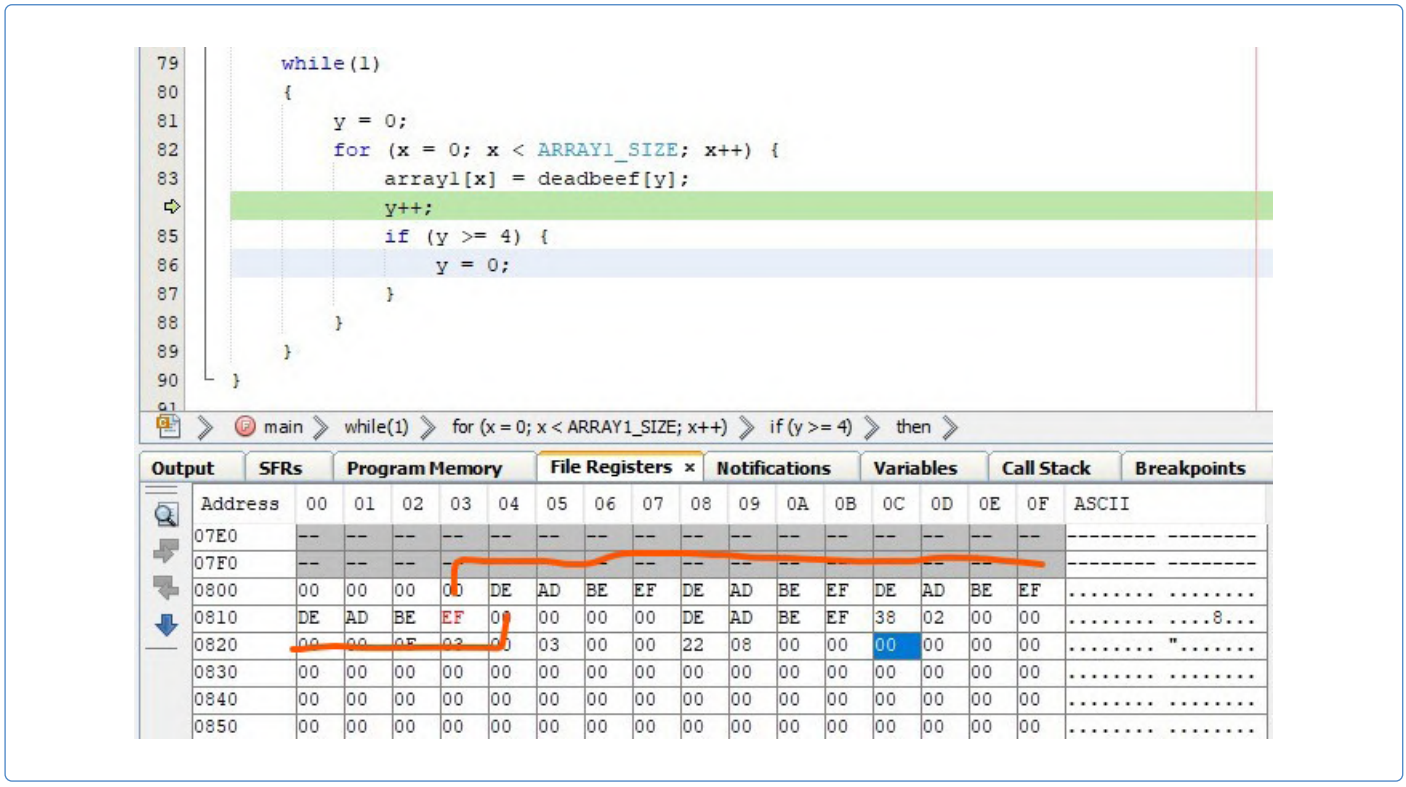
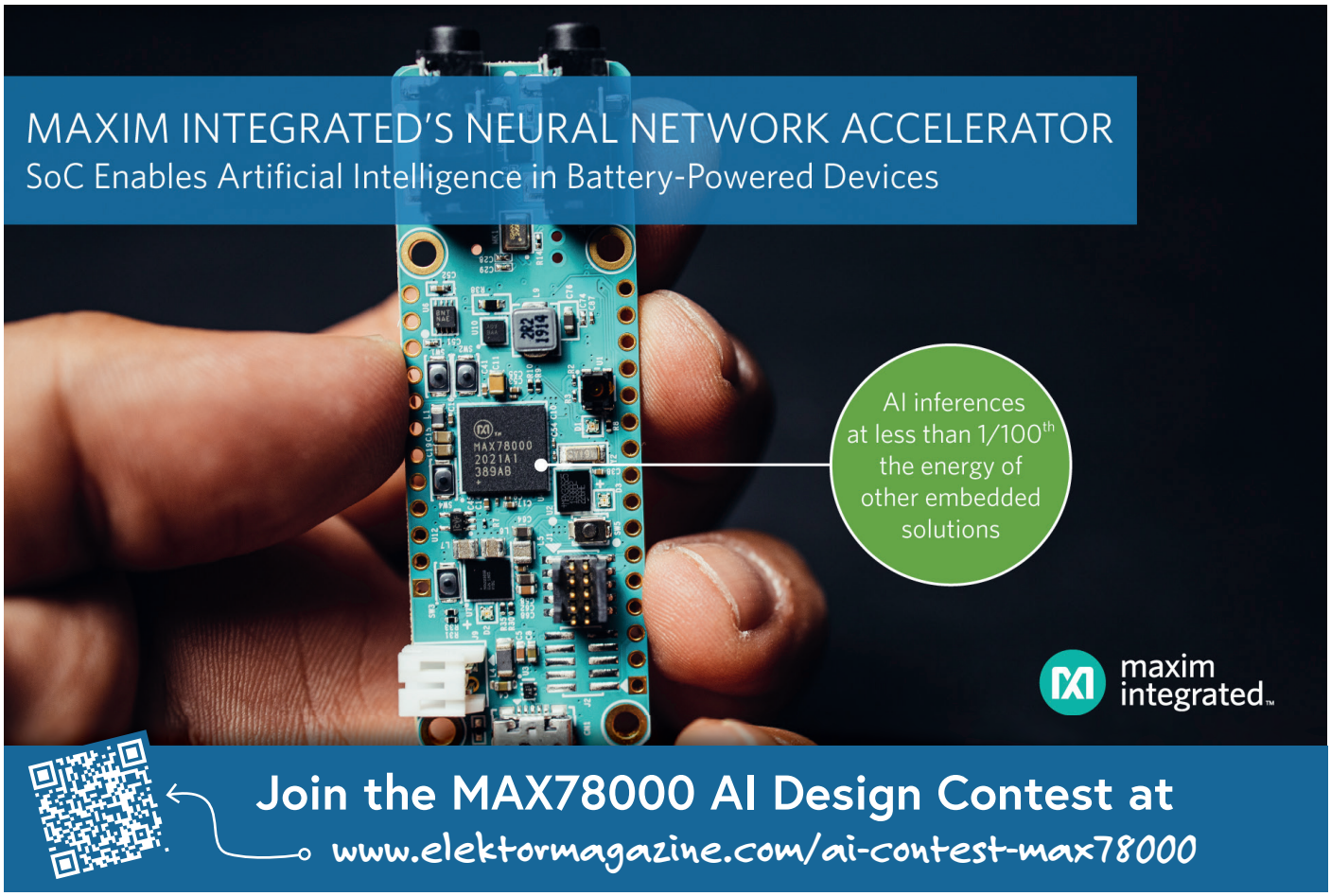



*Bild 4. Der Beispiel-Code hat das Muster `0xDEADBEEF` in das untersuchte Array geschrieben (orange markiert).*



**Listing 1. Die Funktion user_init für eine PIC24-MCU.**

```
/*****************************************************************************/
/* Initialize DEADBEEF checking for stack                                    */
/*****************************************************************************/
void __attribute__((user_init)) deadbeefPrep(void) {
    uint8_t *stackPointer = (uint8_t*)0;
    uint16_t y = 0;

    // Find out where stack is pointing now from SP (WREG15)
    stackPointer = (uint8_t*) WREG15;

    // Increment to next memory position after current stack pointer location
    // (we have a byte pointer into word-addressable memory)
    stackPointer += 2;

    // Loop through stack area until we reach stack limit address (in SPLIM register)
    for (y = 0; stackPointer < (uint8_t*) SPLIM; ++stackPointer) {
        // Write part of 'DE AD BE EF' pattern into stack
        *stackPointer = deadbeef[y];

        // Cycle through four bytes of DE AD BE EF array
        ++y;
        if (y >= 4) {
            y = 0;
        }
    }
}
```

Da der Stack-Pointer hier im Register `WREG15` gespeichert ist, wird die Adresse, auf die er zeigt, in einer Pointer-Variable gespeichert. In dieser Architektur zeigt dieser Zeiger auf die aktuelle Rücksprungadresse der gerade ausgeführten Funktion, so dass ein Überschreiben vermieden werden muss. Daher wird der Pointer um zwei Bytes inkrementiert (dies ist eine 16-Bit-MCU), damit das Schreiben des `0xDEADBEEF`-Musters an der nächsten wortadressierbaren Stelle im Speicher beginnen kann. Der Rest des Codes schreibt das Muster und hält an, wenn das Ende des Stacks erreicht ist. Diese Grenze wird anhand des Inhalts des Stack-Limit-Registers `SPLIM` bestimmt. Dieses Register gibt es ebenfalls nur in dieser Prozessorarchitektur.

Dieser Ansatz kann recht einfach an die meisten anderen Mikrocontroller angepasst werden, erfordert aber eine eingehende Lektüre der Dokumentation von MCU und Compiler-Toolchain.

## Optimaler Ansatz zur Entdeckung von Buffer- und Stack-Überläufen

Wir fassen zusammen: Unerwartetes Verhalten eines Mikrocontrollers ist wahrscheinlich auf einen Programmierfehler zurückzuführen. Wenn Sie herausfinden, was das Problem auslöst und somit replizierbar macht, kann es leicht untersucht werden. Dann ist es einfacher festzustellen, welche der von Ihnen vorgenommenen Code-Änderungen das Problem wirklich beheben und was die eigentliche Ursache sein könnte. Wenn Buffer- und Stack-Überläufe vermutet werden, können Sie dies anhand der hier besprochenen Ansätze mit `0xDEADBEEF` testen.

200722-02


**Ein Beitrag von**

Text: **Stuart Cording**
Redaktion: **Jens Nickel**
Übersetzung: **Dr. Thomas Scherer**
Layout: **Giel Dols**


**Sie haben Fragen oder Kommentare?**

Haben Sie technische Fragen oder Anmerkungen zu diesem Artikel? Schicken Sie eine E-Mail an den Autor über stuart.cording@elektor.com oder kontaktieren Sie Elektor unter editor@elektor.com.

## WEBLINKS


[1] **Elektor-Artikel zu Mikrocontrollern**: www.elektormagazine.de/tags/mikrocontroller

[2] **Microchip Technology, „MPLAB C Compiler for PIC24 MCUs and dsPIC DSCs User's Guide"**: https://bit.ly/3btCHec

# Macht Testen weniger Spaß als Programmieren?

Von **Stuart Cording**

Der Autor geht der Frage nach, warum die Testphase im Bereich der Embedded-Software immer noch ein Schattendasein führt - und stellt Lösungsansätze für dieses Problem vor.

Es sieht so aus, als würden die Entwickler von eingebetteter Software den Diskussionen über Softwarearchitektur, Abstraktion, Modellierung und Simulation gern aus dem Weg gehen, weil sie sich zu sehr auf das Programmieren konzentrieren. Bei diesem Software-Typ ist es systembedingt schwieriger, einen ein Mal freigegebenen Code zu aktualisieren; deshalb sind Fehler hier besonders problematisch. Angesichts der verfügbaren Modellierungs- und Simulationswerkzeuge stellt sich die Frage, warum der Software-Test so vernachlässigt wird.

## Vorsicht vor dem Specht

Bei der Recherche von Hintergrundmaterial für die Entwicklung eines Trainingskurses fragte ich mich, warum Entwickler von Embedded-Software die Methoden, die in anderen Zweigen der Softwareentwicklung alltäglich sind, nur zögerlich anwenden. Ich kenne viele Ingenieure, die keine Coding-Standards verwenden, obwohl mehrere Personen an einem Projekt arbeiten. Niemand hatte irgendwelche Regeln definiert, die einen portablen, wiederverwendbaren Code ermöglichen würden: Objektorientierte Entwicklung, Modeling und Simulation waren nicht vorgesehen. Das folgende Zitat erscheint mir dazu recht passend:

„Wenn Zimmerleute ihre Gebäude so errichten würden, wie Programmierer ihre Programme erstellen, dann würde der erste Specht, der vorbeikommt, den gesamten Dachstuhl zerstören."

Diese leicht abgewandelte Version von Weinbergs Gesetz [1] gab mir zu denken. Warum nutzen die Entwickler von eingebetteter Software nicht die Forschungsergebnisse der Informatik und erzeugen ihren Code mit einer geeigneten Architektur oder nutzen Modellierung und Simulation?

## Ein Generationsproblem?

Vielleicht liegt es an der Generation der betroffenen Ingenieure? Sie sind mit den in ihren Ressourcen eingeschränkten 8-Bit-Mikrocontrollern aufgewachsen und haben gelernt, mit Assembler jedes noch so kleine Quäntchen Leistung aus ihnen regelrecht heraus zu „kitzeln". Selbst der Wechsel zu C war für einige eine Herausforderung, trotz der Nähe zur verwendeten Hardware.

Ein weiterer Aspekt ist die Art und Weise, wie C normalerweise gelehrt wird: Der Schwerpunkt liegt auf der „Grammatik", und nicht darauf, wie man diese am besten anwendet. Doch genau wie bei einer gesprochenen Sprache machen gute Grammatik-Kenntnisse den Studenten noch lange nicht zu einem eloquenten Redner.

Bis zum Aufkommen von Internet-of-Things-Anwendungen (IoT) konnte die Firmware eines Mikrocontrollers nicht aktualisiert werden, ohne die Installation der Umgebung zu berücksichtigen. Trotz dieses systembedingten Mangels an Spielraum für Fehler blieb die gewohnte Herangehensweise an die Code-Entwicklung bei einigen Programmierern von Embedded-Software unverändert.

## Neue Lösungen

Konnte der Code, der auf Mikrocontrollern ausgeführt wird, früher oft auf einfache Weise implementiert werden, so sind die heutigen Herausforderungen wirklich enorm. So ist es heute nicht unüblich, dass ein Mikrocontroller bei der Steuerung eines Elektromotors eine ganze Reihe komplizierter Aufgaben übernimmt. Allein die Bestimmung des Rotorwinkels und damit des nächsten Kommutierungspunktes in Echtzeit ist trotz existierender Software ein erheblicher Aufwand.

Bei Waschmaschinen muss dabei auch noch die Norm *IEC 60730* [2] berücksichtigt werden (Automatische elektrische Steuerungen - Teil 1: Allgemeine Anforderungen). Die meisten Mikrocontroller-Hersteller bieten eine „Class B"-Bibliothek [3] an, die eine Reihe von Tests ausführt, die CPU-Register, Programmzähler und Speicherintegrität abdecken und auch die Uhr überprüfen. Dieser Code muss „parallel" zur Echtzeit-Motorsteuerung ausgeführt werden. Zusätzlich sind oft noch Kommunikationsschnittstellen im Einsatz, die ebenfalls zeitnah reagieren müssen, ebenso wie eine Mensch-Maschine-Schnittstelle. Auch wenn ein Entwickler von Embedded-Software sicher sein kann, dass all diese Funktionen isoliert funktionieren: Was passiert, wenn sie alle integriert werden?

Eine Möglichkeit besteht in der Verwendung der *Unified Modeling Language* (UML), um ein Modell der Anwendung zu simulieren. Doch hier gibt es oft erheblichen Widerstand seitens der Embedded-Software-Entwickler. Eines der Hauptbedenken ist, dass dies eine Zeitverschwendung sei, denn wenn man sich schon die Mühe macht, ein Modell der Anwendung zu erstellen, kann man es auch gleich kodieren. Außerdem hat die UML einen schlechten Ruf bekommen, da man ungünsti-

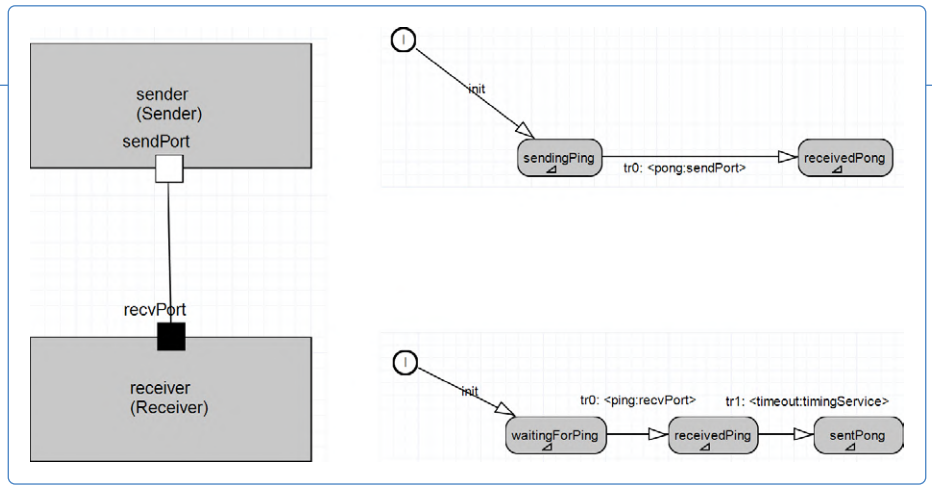



*Bild 1. Ein beispielhaftes „Ping-Pong"-Projekt in eTrice, das die Akteure „ping" und „pong" (links) zusammen mit ihrem Verhalten (rechts) zeigt.*

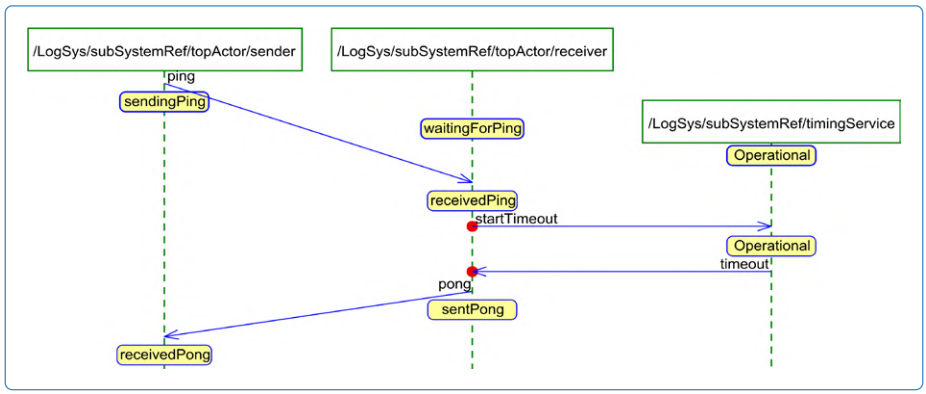



*Bild 2. Die Message Sequence Chart, die sich aus der Simulation des „Ping-Pong"-Modells ergibt.*

gerweise versucht hat, sie auf jeden Teil eines Systems anzuwenden [4]. Diese abstrahierte Arbeitsweise erfordert auch eine andere Entwicklungsmethodik, und der Versuch, ein Team von Entwicklern davon zu überzeugen, neue Methoden zu übernehmen, ist oft eine Herausforderung. Alles zweimal zu kodieren, einmal als Modell und dann wieder als Anwendungscode, wird als Phasendiskontinuität bezeichnet [5]. Idealerweise sollte jeder Modellierungsansatz dieses Problem lösen, was bei der domänenspezifischen Sprache (DSL) *Real-Time Object-Oriented Modeling* (ROOM) [6] auch der Fall ist. Sie wurde speziell für die Bedürfnisse von Entwicklern ereignisgesteuerter Echtzeitsysteme entwickelt und unterstützt die Modellierung und Simulation, während sie gleichzeitig in der Lage ist, Code zu generieren, der auf der Ziel-Hardware verwendet werden kann.

ROOM beschreibt Softwaresysteme entlang der drei Dimensionen Struktur, Verhalten und Vererbung. Strukturen können durch Diagramme dargestellt werden, die wiederum Akteure enthalten, die über Ports miteinander kommunizieren. Nachrichten, die zwischen ihnen ausgetauscht werden, werden von Laufzeitbibliotheken verarbeitet, die die dafür erforderlichen Dienste bereitstellen. Das Verhalten von Akteuren wird durch hierarchische Finite-State-Maschinen beschrieben, die ebenfalls als Diagramme erzeugt werden können. Für die Zustände kann in der Zielsprache des verwendeten Codegenerators (C-Compiler für einen Mikrocontroller) ein Ein- und Ausstiegscode definiert werden. Zustandsübergänge erfolgen beim Empfang von Nachrichten über Ports, und eine daraus resultierende Antwort kommt über die Ports zurück.

Die Vererbung stellt den objektorientierten Aspekt von ROOM dar, indem Akteure als Klassen behandelt werden, die mehrfach instanziiert werden können. Jede Instanziierung erbt auch den Zustandsautomaten, der natürlich bei Bedarf erweitert werden kann. Ein weiteres Konzept ist schließlich das der Schichtung. Dies ermöglicht zum Beispiel die Kommunikation zwischen Anwendungen oder den Zugriff auf Dienste wie beispielsweise Timer über Service Access Points (SAP).

Unterstützt wird dieser Softwareentwicklungsansatz durch eTrice [7], ein Paket, das der Eclipse IDE hinzugefügt werden kann. Modelle können grafisch oder als Text erstellt werden, die Code in C, C++ oder Java generieren. Die Simulation des Modells erzeugt sogenannte *Message-Sequence-Charts* (MSC) zur Visualisierung im externen Tool Trace2UML, das Teil des UML-Tools Astade [8] ist. Ein Tutorial-Projekt zeigt, wie eine einfache Ping-Pong-Anwendung implementiert werden kann (**Bild 1**), wobei die Verzögerung der Antwort durch einen Timer-Service-Access-Point bereitgestellt wird. Die Struktur und das Verhalten können grafisch oder in der ROOM-DSL untersucht werden. Das Ausführen des Modells liefert die Ergebnisse in einer Kommandozeile. Schließlich kann das Ergebnis der Simulation als *Message-Sequence-Chart* untersucht werden (**Bild 2**).

## Die Überwindung der Zurückhaltung

Nachdem sie überstürzt mit der Codierung begonnen haben, scheinen viele Embedded-Entwicklungsteams beim Test sehr zurückhaltend zu sein - jedenfalls bis ein guter Teil der Anwendungsentwicklung abgeschlossen ist. Die Gründe dafür sind vielfältig. Das beginnt schon bei der irrigen Ansicht, dass Testen teuer ist und von einem Testteam oder einer Testabteilung durchgeführt werden muss. Hinzu kommt der Mangel an klaren Anforderungen für einzelne Softwaremodule. Und vielleicht gibt uns die Verfügbarkeit leistungsfähiger Debugger für die Entwicklungs-IDE das trügerische Gefühl, dass wir beim Schreiben des Codes schon alles überprüft haben. Die Folgen von Programmierfehlern werden in Lisa Simones Buch „*If I Only Changed the Software, Why is the Phone on Fire?*" [9] humorvoll geschildert. In einem Beispiel ist ein Fehler in einem Algorithmus zur Temperaturregelung das Ergebnis eines falsch gewählten Datentyps. Simone beschreibt zwar hervorragend den Debugging-Prozess, um solche Probleme nachträglich zu finden, aber diese hätten durch Unit-Tests bereits viel früher entdeckt werden müssen.

Unit-Tests werden mit Kenntnis des Quellcodes durchgeführt, zumeist von der Person, die ihn entwickelt hat. Bei einem Mikrocontroller erfordert dies das Schreiben von Anwendungen, die den Code ausführen, die Antworten auswerten und die Ergebnisse ausgeben können, zum Beispiel über eine serielle Schnittstelle

an eine Konsole. Das bedeutet, dass eine Hardwareplattform benötigt wird, auf der die Tests ausgeführt werden können. Ein Arduino und viele seiner Klone, die einen UART-zu-USB-Wandler enthalten, machen dies viel einfacher, da die Testergebnisse auf einer Konsole ausgegeben werden können. Aber es ist immer noch ein hoher Arbeitsaufwand, eine Testanwendung für jedes entwickelte Softwaremodul zu pflegen.

Die Software testIDEA von iSYSTEM, ein Hersteller von Debugging-Tools, bietet eine hervorragende und einfache Möglichkeit, den Testaufwand zu reduzieren [10]. Bei den betreffenden Tools hat die Entwicklungsumgebung bereits über die Debugging-Schnittstelle Zugriff auf den Quellcode und das Innenleben des Mikrocontrollers. Sie benötigt lediglich eine kompilierte Applikation (ELF-Format), die in den Flash-Speicher des Mikrocontrollers geladen werden kann.

In testIDEA werden Tests definiert, die einzelne Funktionen mit definierten Testwerten für die Funktionsparameter ausführen (**Bild 3**). Bei vollem Zugriff auf das RAM können alle Datentypen und Zeiger problemlos eingefügt werden. Das Testergebnis hängt von dem Rückgabewert ab, den die aufgerufene Funktion liefert. Zusätzlich kann während des Testens die Ablaufverfolgung (Tracing) aktiviert werden, die zusätzlich zu den Testergebnissen auch eine Codeabdeckung liefert. C++ wird ebenfalls unterstützt, erfordert aber den zusätzlichen Schritt des Aufrufs des Konstruktors, bevor eine der Methoden getestet werden kann. Die Tests können auch in Python exportiert werden, wodurch die Testausführung automatisiert werden kann.

Ein beeindruckendes Feature besteht darin, dass alle internen Variablen des Typs *private* mit Werten injiziert oder als Teil des Testablaufs ausgewertet werden können. Der mit dem Debugging beauftragten Person steht jede Möglichkeit offen. Dies ist hilfreich, wenn man einen Code testet, der mit Hilfe von Optimierungsmethoden kompiliert wurde. Es ist bekannt, dass Software, die auf diese Weise erstellt wurde, nicht genau so gut funktioniert wie der nicht optimierte Code, zum Beispiel wegen des Entfernens von ungenutztem Code und des Umschichtens von Codefragmenten. Die Fehlersuche in solch einem Code ist besonders schwierig, da ein Großteil des Bezugs zwischen Quell- und Binärcode verloren geht. Durch das Ausführen derselben Suite von Tests mit testIDEA ist es möglich, festzustellen, ob der grundlegende Code noch wie erwartet funktioniert, bevor man dazu übergeht, auf der Integrations- oder Systemebene zu testen.

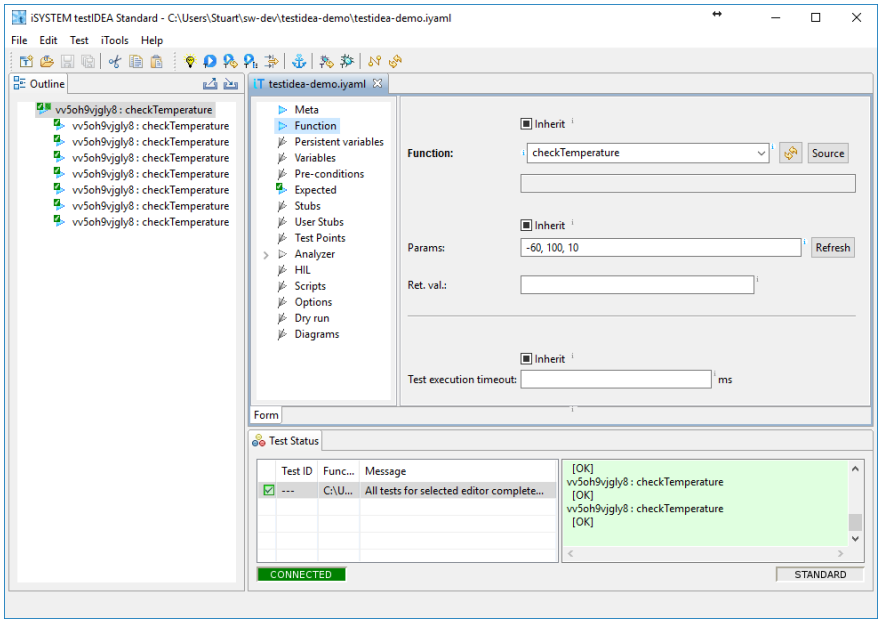



*Bild 3. Testen einzelner Softwarefunktionen auf der Mikrocontroller-Hardware in testIDEA (Quelle: iSYSTEM AG).*

## Wie viele Tests sind genug?

Für unerfahrene Tester kann es schwierig sein, die optimale Strategie zu bestimmen. Testen erinnert schnell an das Kaninchenloch von Alice im Wunderland, denn je mehr man sich ansieht, was zu testen ist, desto mehr scheint es zu testen zu geben. Auch hier wird es aufgrund der mehrdimensionalen Aspekte und Abhängigkeiten schnell unklar, wie viele Tests genug sind. Jeder, der schon einmal an einem sicherheitskritischen Projekt gearbeitet hat, weiß auch, dass die geltenden Normen offen für Interpretationen sind, zum Beispiel die ISO 26262 für die Automobilindustrie und die ISO 14971 und IEC 60601 für die Medizintechnik. So wird beispielsweise. in der ISO 26262 die Zweigabdeckung für strukturelle Abdeckungsmetriken „sehr empfohlen", während die Anweisungsabdeckung und MC/DC (modifizierte Bedingungs-/Entscheidungsabdeckung) nur „empfohlen" wird [11]. Nur durch die Zusammenarbeit mit Testexperten können Sie bestimmen, was dies im Kontext Ihrer Anwendung bedeutet und wie Sie es am besten implementieren.

Um sicherzustellen, dass alle möglichen Zweige des Codes abgedeckt sind und möglichst alle Permutationen des Codes getestet wurden, gibt es Hilfe in Form der MBTsuite [12] der *sepp.med GmbH*. Dieses modellbasierte Testframework eignet sich hervorragend für Integrations- und Systemtests, da es in diesem Stadium des Entwicklungsprozesses schwierig ist, alle möglichen Testpermutationen zu ermitteln.

Der Prozess beginnt mit der Untersuchung der Anforderungen an das System und deren Eingabe in ein Werkzeug wie *Enterprise Architect* [13]. Von hier aus kann ein Modell des Systems erstellt werden, wobei das Modell mit verwandten Anforderungen verknüpft wird, um zu zeigen, wie diese getestet werden. Komplexe Aspekte der Entwicklung können in Untermodelle aufgeteilt werden, wobei die oberste Ebene clean und das Modell gut abstrahiert bleibt (**Bild 4 und Bild 5**).

Wenn das Modell vollständig ist, wird es in die MBTsuite importiert. Das betreffende Werkzeug ermöglicht die Erstellung eines Testplans nach verschiedenen Strategien. *Full Path Coverage* würde versuchen, einen Testplan zu entwickeln, der das gesamte Modell abdeckt, während *Shortest Path*, wie der Name schon sagt, versucht, die kürzesten Pfade durch das Modell zu finden (**Bild 6**). Die Zufallsstrategie kann dabei helfen, Probleme zu entdecken, die mit den eher formalen Ansätzen nicht gefunden werden können. Sie ist auch ideal für die Durchführung eines sogenannten *Smoke-Tests*, um sicherzustellen, dass alles korrekt funktioniert.

Wie im Modell der elektrischen Parkbremse (EPB) zu sehen ist (Bild 4), ist es möglich, sich zyklisch durch viele Abschnitte des Modells zu bewegen, was zu einer Endlosschleife führen kann. Um solche rekursiven Schleifen einzuschränken, kann jede Prüfstrategie mit Parametern (wie zum Beispiel der

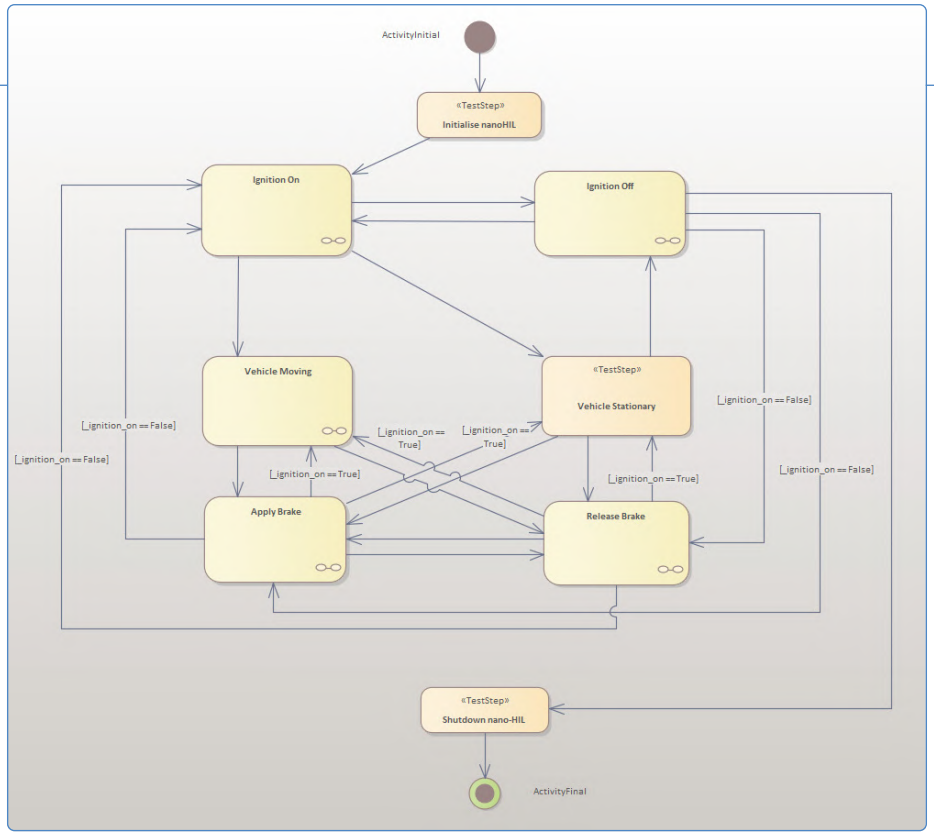



*Bild 4. Das Modell einer elektrischen Parkbremse in Enterprise Architect.*

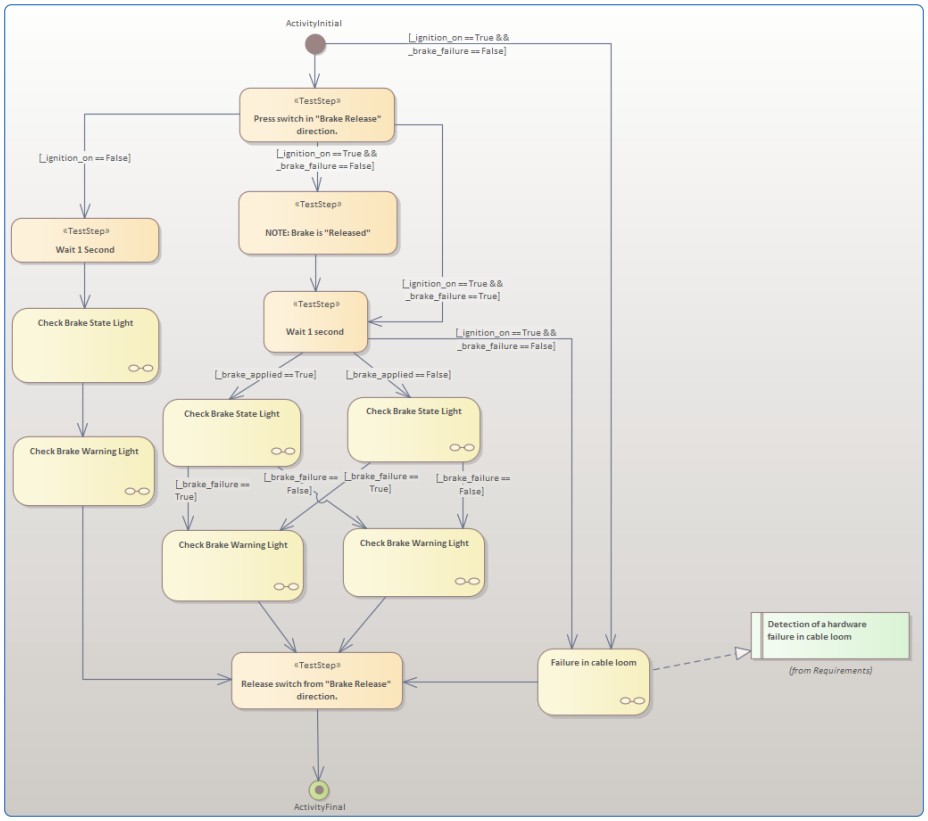



*Bild 5. Das Teilmodell der „Feststellbremse" in Enterprise Architect mit Verknüpfung zu den Anforderungen.*

Anzahl der Schleifendurchläufe und der maximalen Pfadlänge) im Umfang begrenzt werden. Auch bestimmte Abschnitte des Modells können untersucht werden, so dass Tests für eine bestimmte Funktion oder Fähigkeit erstellt werden können.

Im einfachsten Fall generiert MBTsuite die Tests als Word- oder Excel-Dokument, das Testschritte und Verifikationspunkte auflistet, wobei für jeden Schritt ein Ergebnis angegeben wird. Manuelle Tests können dann ein *Pass-Fail*-Ergebnis liefern, indem für jeden Schritt ein Kästchen angekreuzt wird. Wenn jedoch ein automatisiertes Testen gewünscht wird, kann der Test in einem geeigneten Format (z. B. Python) in Kombination mit einer geeigneten Vorlage exportiert werden.

## Schneller Hardware-Test

Idealerweise wird ein unter Umständen teurer Hardware-in-the-Loop-Test (HIL) benötigt, um eine Anwendung auf Herz und Nieren zu prüfen. Dies ist besonders hilfreich für Anwendungen, die mit gefährlichen Spannungen oder beweglichen mechanischen Teilen arbeiten. Oft steht auch nur eine Hardware-Anordnung zur Verfügung, um mehrere Varianten der entwickelten Produkte zu testen. Infolgedessen kann der HIL-Test zu einem ernsthaften Engpass werden, so dass das Testen daher erst sehr spät im Entwicklungsprozess erfolgt. Unter den aktuellen Corona-Einschränkungen sind HIL-Tests möglicherweise sogar noch nicht einmal verfügbar.

Um diese Probleme zu umgehen, hat die PROTOS Software GmbH ihre miniHIL-Plattform entwickelt. Sie besteht aus einer Hardwareplatine von etwa der Größe eines A4-Blattes. Auf der rechten Seite ist Platz für ein STM32 Nucleo Entwicklungsboard, während auf der linken Seite ein leistungsfähiger STM32H743 sitzt. In der Mitte befindet sich eine Anordnung von Pins, die als Verbindungsmatrix zwischen den beiden Seiten dienen. Der STM32H743 dient dazu, die Anlage zu simulieren, die der STM32 Nucleo steuert, und alle Signale bereitzustellen, die die Anlage erzeugt. Damit ist er ideal zum Testen von Motorsteuerungs-Software, ohne dass der Code auf einer echten Wasch- oder Bohrmaschine laufen muss.

Die Umgebung wird von eTrice und der Sprache CaGe [14] (Case Generator) zur Entwicklung der Tests unterstützt. Damit kann der Entwickler nicht nur schnell überprüfen, ob die Codeänderungen funktionieren: Die gesamte Plattform kann auch mit einer *Continuous-Integration-Plattform* (CI) wie Jenkins [15] kombiniert werden. Automatisierte Tests können so zum Beispiel auch über Nacht regelmäßig auf der Hardware ausgeführt werden und liefern die Ergebnisse auf einem Dashboard, das am nächsten Tag dann gleich überprüft werden kann (**Bild 7**).

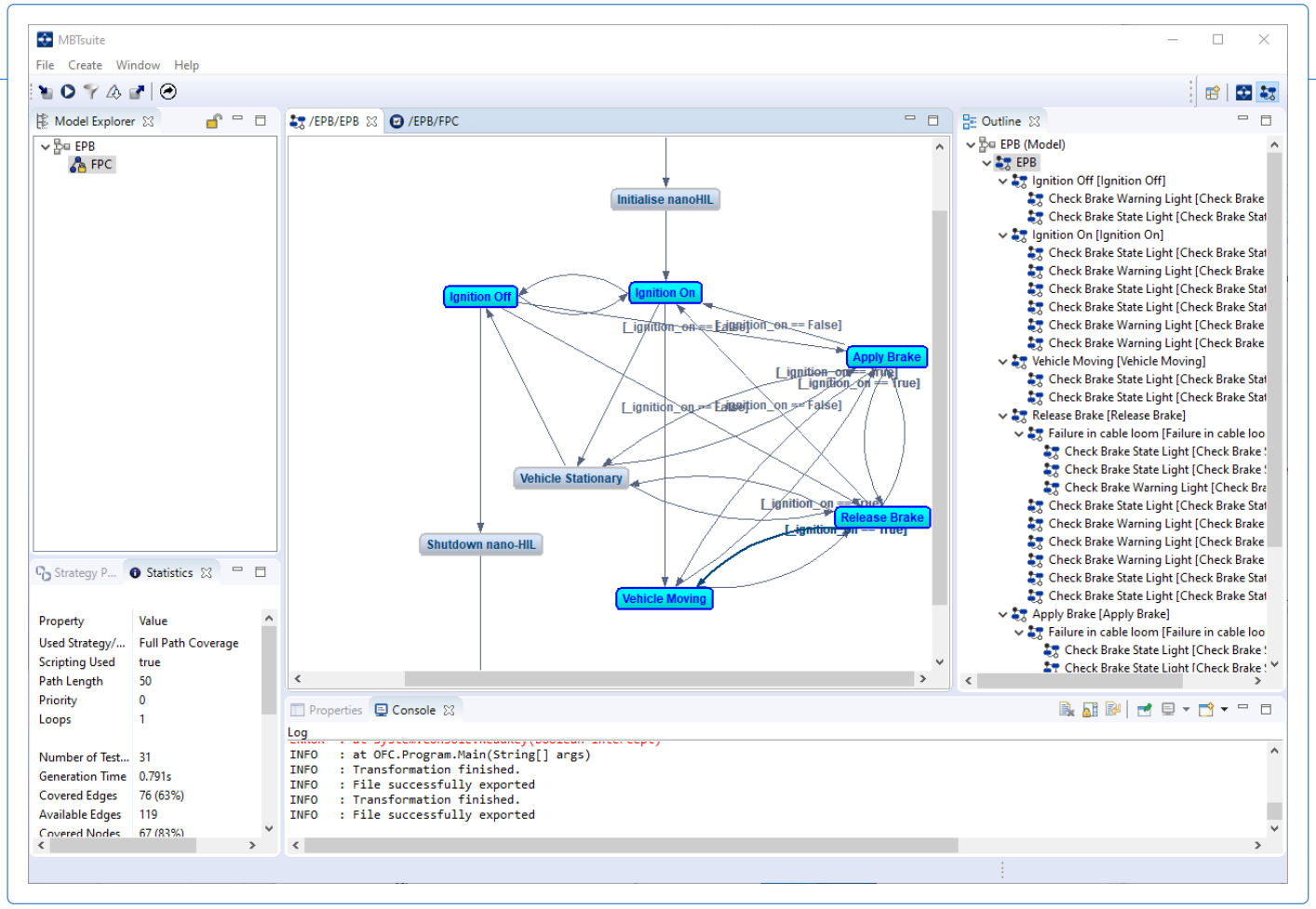



*Bild 6. Das Modell der „Electric Park Brake" in der MBTsuite mit Test nach einer Full Path Coverage Strategie.*

## Identifizierung sporadischer Ausfälle

Heutige Automobile sind außerordentlich komplex, wobei der Umfang des Codes in einem Fahrzeug in etwa dem von Windows NT entsprechen kann [16]. Mit der zunehmenden Verteilung der Code-Entwicklung an viele Zulieferer und in einigen Fällen sogar an den Kunden selbst wird nicht die Funktionalität, sondern das Timing zunehmend zur Ursache für sporadische, schwer zu findende Fehler, was durch den Einsatz von Multicore-Prozessoren und virtuellen Maschinen noch verstärkt wird.

Trotz umfangreicher Tests passieren immer noch merkwürdige Dinge, wenn die Hardware erst einmal ausgeliefert ist oder, schlimmer noch, wenn sie bereits im Einsatz ist. Ein Beispiel kommt aus dem Automobilbereich, wo zwei über einen CAN-Bus verbundene, elektronische Steuergeräte (ECU) intermittierende Ausfälle zeigten. Steuergerät A zeigte nur Ausfälle, wenn es mit Steuergerät B verbunden war. Steuergerät B und Steuergerät C arbeiteten dagegen problemlos zusammen. Das Ganze war auf leichte Unterschiede in der Quarzfrequenz in jedem Steuergerät zurückzuführen. Mit der Zeit wich das Timing der Mikrocontroller so weit ab, dass gelegentlich eine CAN-Nachricht verloren ging.

Werkzeuge wie chronSIM und chronVIEW der Inchron AG [17] und die TA Tool Suite der *Vector Informatik GmbH* [18] geben dem Timing den gleichen Stellenwert wie der Funktionalität. Sie erlauben es, Modelle des Zielsystems zu erstellen, bevor Code geschrieben wird, und ermöglichen es, die optimale Architektur für den Code auf einer Vielzahl von Zielhardware zu bestimmen. Auch die Auswirkungen der Zuweisung von

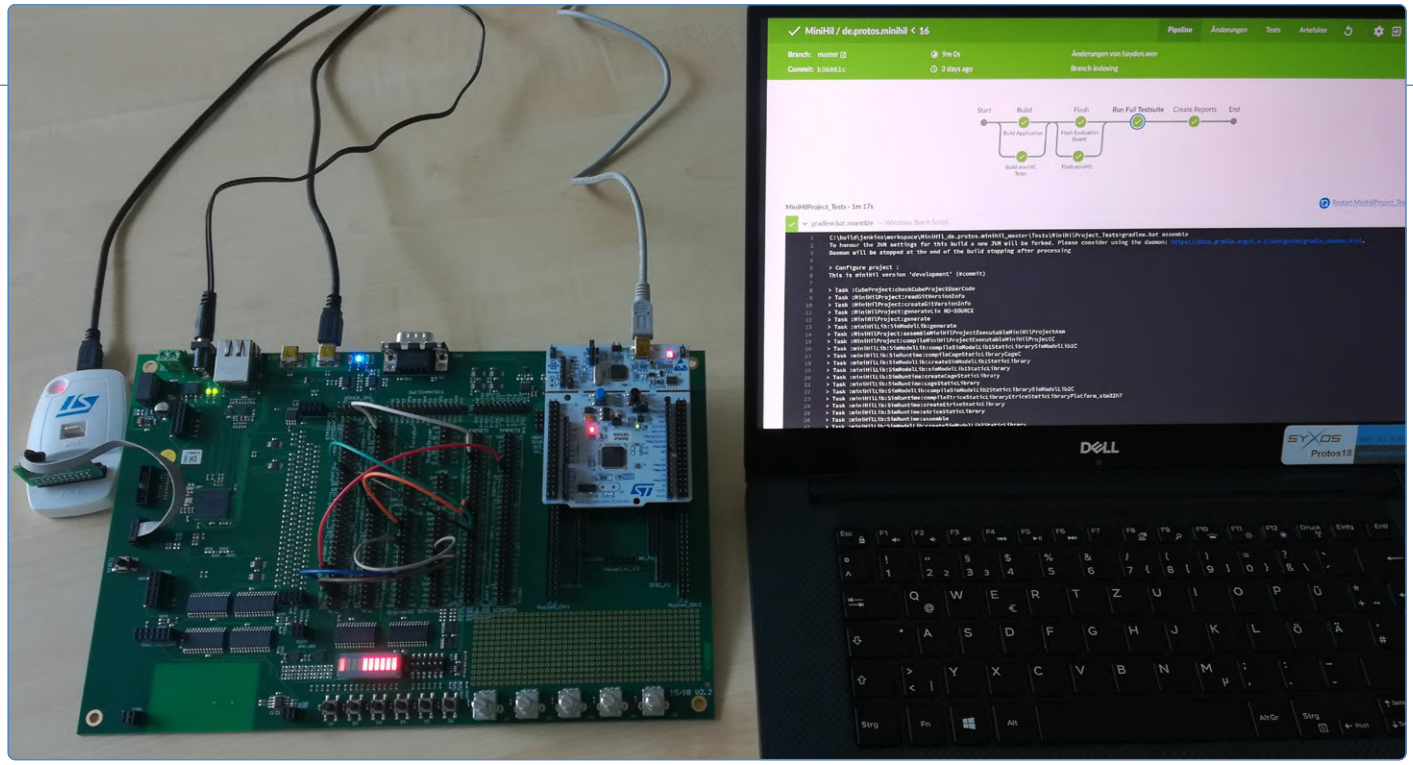




*Bild 7. MiniHIL-Setup integriert auf einem PC, auf dem ein Jenkins Continuous Integration Server läuft (Quelle: PROTOS Software GmbH)*

Code zu unterschiedlichen Kernen können bewertet werden, was bei der Verwendung heterogener Multicore-Prozessoren mit unterschiedlichen Frequenzen unerlässlich ist. Außerdem lassen sich Ereignisketten (der zeitliche Ablauf vom Sensor über das System bis zum Aktor) auswerten (**Bild 8**), was für das autonome Fahren sehr wichtig ist, denn hier werden die Eingaben von mehreren Sensoren mit verschiedenen Datenraten kombiniert.

Durch die Fokussierung auf das Timing, die Funktionalität und die Definition der zulässigen Ausführungszeit von Code-Abschnitten kann das bei CAN-Geräten geschilderte

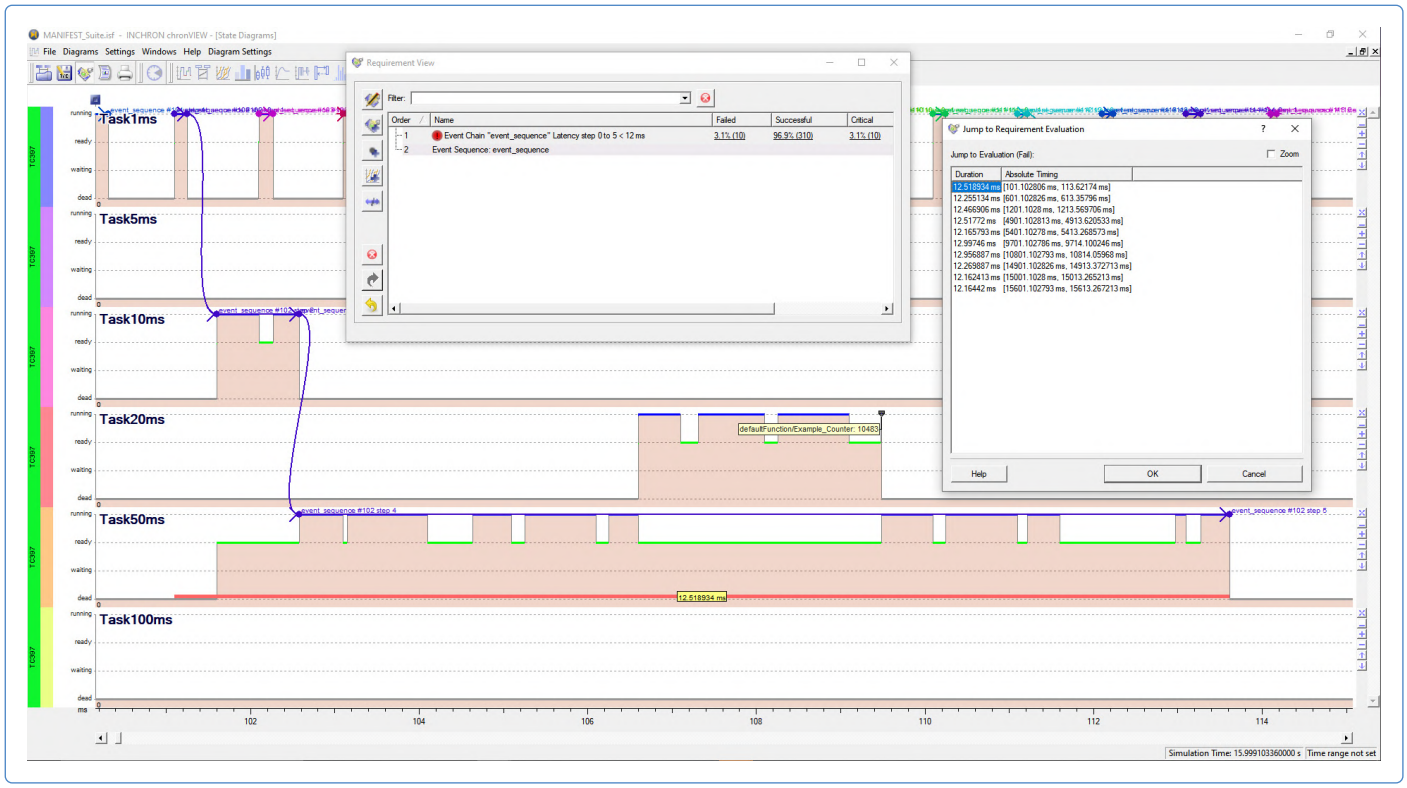



*Bild 8. Die Inchron Tool-Suite stellt durch Modellierung, Simulation und Analyse das Timing und nicht die Funktionalität in den Vordergrund. Hier wird eine Ereigniskette analysiert (Quelle: INCHRON AG).*

Zeitproblem vermieden werden. Modelle können auch simuliert werden, um die Auswirkungen von Taktfrequenzvariationen, Core-Zuweisung, Code-Ausführungszeit oder Änderungen der Taktfrequenz besser zu verstehen. Um sicherzustellen, dass die Timing-Anforderungen erfüllt werden, können diese Werkzeuge auch die Trace-Daten analysieren, die vom Mikrocontroller im Betrieb auf realer Hardware erfasst wurden, und prüfen, ob die Ausführungszeit von Code-Abschnitten unter unterschiedlichen Testbedingungen immer noch den Anforderungen entspricht.

## Zeit, das Modellieren ernster zu nehmen?

Die Entwicklung eingebetteter Software ist sehr traditionell geblieben. Die Fixierung auf den Code wird durch die große Auswahl an Entwicklungs-Toolchains von Mikrocontroller-Lieferanten verstärkt, die sich auf die Codierung und das *Debugging on target* konzentrieren. Da sie größtenteils kostenlos sind, schmälern sie den Wert alternativer kostenpflichtiger Toolchains.

Werkzeuge, die die Modellierung und Simulation von eingebetteten Systemen unterstützen, bieten eine Möglichkeit, einen Schritt zurück zu treten und zu abstrahieren, bevor man sich an die harte Arbeit der Codierung macht. Aufgrund der inhärenten, sich wiederholenden Tests und Analysen, die sich aus diesem Ansatz ergeben, können Probleme in der Architektur ermittelt werden, bevor die Anwendungsentwicklung ein Stadium erreicht, in dem es zu teuer wird, den Kurs zu ändern. Bench-top HIL mit Simulationslösungen erlauben es, selbst kleine Codeänderungen in kompakten Zeitzyklen zu bewerten, anstatt auf einen großen Test des vollständig integrierten Codes zu warten. Schließlich müssen die komplexesten Echtzeitanwendungen, wie sie für die Automobilindustrie entwickelt werden (wo Modellierung und Test bereits weit verbreitet sind) über die funktionalen Anforderungen hinausgehen und auch Timing-Anforderungen einschließen.

Die vielleicht größte Hemmschwelle bei der Einführung dieser modellbasierten Werkzeuge ist die Angst vor Veränderungen. Sie drängt traditionelle Embedded-Entwickler in ein Reich des Unbekannten, weit weg von der Hardware, die sie kennen und lieben. Änderungen in den Entwicklungsprozessen führen zu Widerstand. So lange sie nicht stattfinden, bleiben ihre Vorteile, obwohl sie auf dem Papier klar und sinnvoll sind, unerreichbar. Viele der hier besprochenen Tools gibt es jedoch schon seit mehr als einem Jahrzehnt, was darauf hindeutet, dass sie ihren Platz und ihre Rolle gefunden haben. Es bleibt den einzelnen Entwicklern überlassen, jede Option im Kontext ihrer Prozesse und der von ihnen entwickelten Produkte auf Brauchbarkeit zu untersuchen. Es lohnt sich, denn dies kann zu Einsparungen bei der Entwicklungszeit und den Entwicklungskosten sowie zu Qualitätsverbesserungen führen.

200677-02

**Sie haben Fragen oder Kommentare?**

Wenn Sie technische Fragen zu diesem Artikel haben, können Sie sich gerne direkt per E-Mail an den Autor wenden: stuart.cording@elektor.com.

## WEBLINKS


[1] **Weinbergs Gesetz**: http://bit.ly/2WBOeiU

[2] **IEC 60730-1:2013**: http://bit.ly/37Bq8Lw

[3] **Hitex Class B Library**: https://bit.ly/2KorSz5

[4] **A. E. Bell, „Death by UML Fever", ACM QUEUE, April 2004**: https://bit.ly/3auSCsv

[5] **B. Selic, G. Gullekson, P. Ward, Real-Time Object-Oriented Modeling, Wiley, 1994**: http://bit.ly/3mJdJtC

[6] **Real-Time Object-Oriented Modeling - Wikipedia**: http://bit.ly/34xxLkq

[7] **eTrice**: http://bit.ly/3mGWk4p

[8] **Astade**: http://bit.ly/34vLxnG

[9] **L. Simone, „If I Only Changed the Software, Why is the Phone on Fire?", Newnes, 2007**: http://bit.ly/3my2nIx

[10] **testIDEA - iSYSTEM**: http://bit.ly/3rh2BHS

[11] **Feabhas Ltd, „A Quick Guide to ISO 26262", 2016**: https://bit.ly/3nHMeSk

[12] **MBTsuite**: https://mbtsuite.com/

[13] **Enterprise Architect**: https://sparxsystems.com/

[14] **CaGe User Manual**: http://bit.ly/3h8ht6M

[15] **Jenkins**: https://www.jenkins.io/

[16] **Infografik**: How Many Millions of Lines of Code Does It Take?: http://bit.ly/37BS4iH

[17] **INCHRON Test-Suite**: https://www.inchron.com/

[18] **Vector TA Tool Suite**: http://bit.ly/3mBMocF

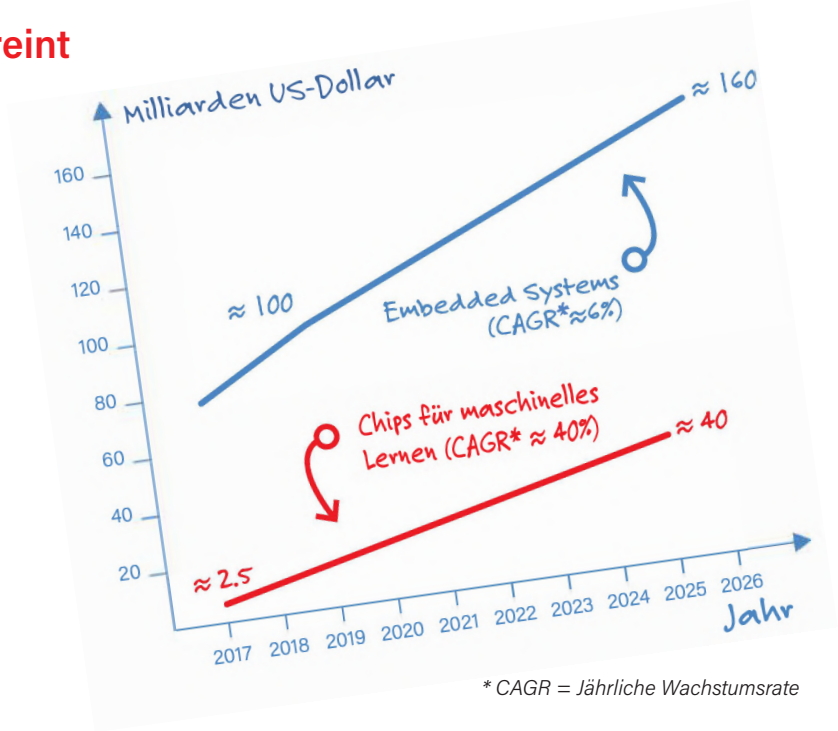
# Elektor infographics

**Von Robert van der Zwan**

## Embedded und KI: Endlich vereint

Auf den ersten Blick scheinen künstliche Intelligenz (KI) und *Embedded Computing* unvereinbar. *Embedded* bedeutet schließlich „eingebettet", also unantastbar und unflexibel. Bei KI geht es um Begriffe wie *Anpassungsfähigkeit, maschinelles Lernen* und *Deep Learning*. Können die beiden zusammenkommen? Ja: KI wird zunehmend in eingebettete Systeme implementiert. Das Wachstum von *Embedded Computing* wird zu einem großen Teil von ML, dem maschinellen Lernen, abhängen. Schauen Sie sich doch einmal das ähnliche Wachstum in absoluten Zahlen von Embedded Computing und ML an.

*(Quellen: Allied Market Research, Global Market Insights, Market Resarch Future, Research and Markets)*

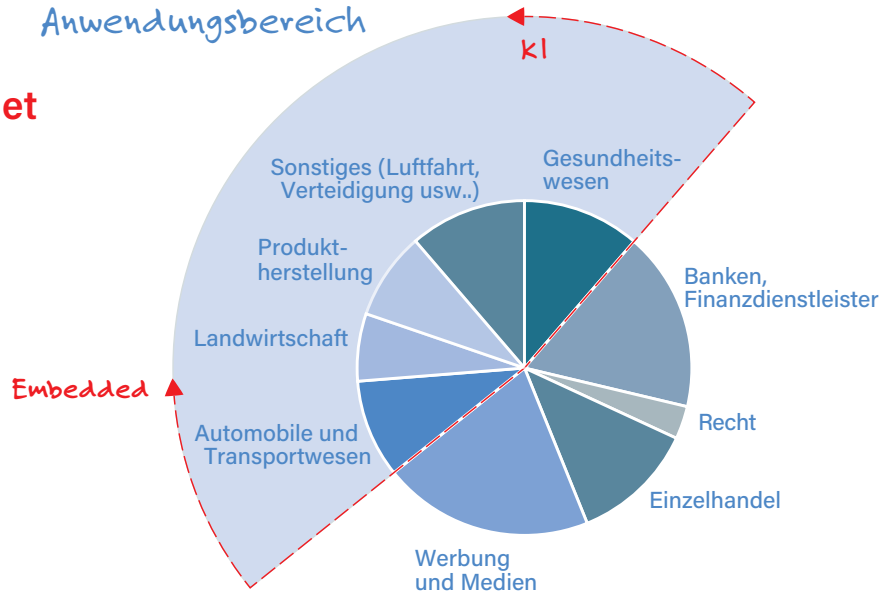



** CAGR = Jährliche Wachstumsrate*

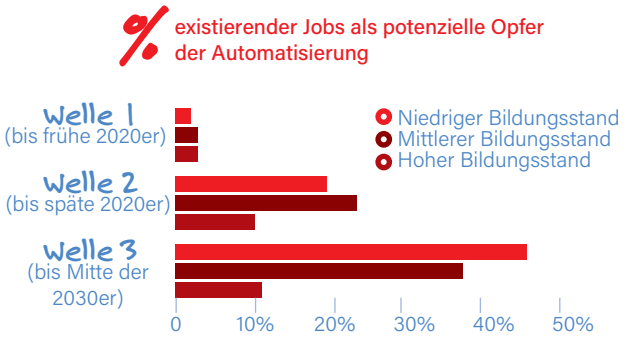
## Ein breites Anwendungsgebiet

Die Hälfte aller KI-Anwendungen wird in Embedded- und nicht in regulärer Software implementiert werden. Nach einem Tortendiagramm von *Grand View Research* zu urteilen, werden Sektoren wie zum Beispiel das Gesundheitswesen, die Automobil- und Transportbranche, die Landwirtschaft, die Fertigungsindustrie, das Militär sowie die Luft- und Raumfahrt das maschinelle Lernen in eigener Software einsetzen. Andere Sektoren wie Banken oder Einzelhandel sind dagegen eher auf KI angewiesen, die aus externer Software stammt.

*(Quelle: Grand View Research)*



## KI - Jobkiller oder Jobwunder?



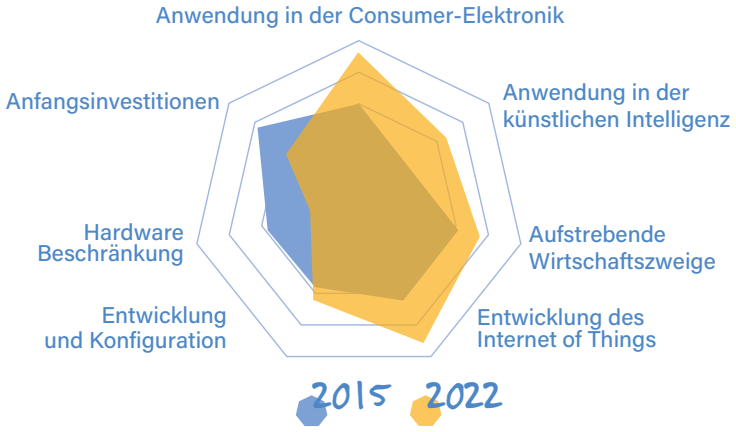

Das Wirtschaftsprüfungs- und Beratungsunternehmen *PwC* hat eine umfangreiche Untersuchung darüber durchgeführt, wie sich Automatisierung und KI auf die globale Erwerbstätigkeit auswirken. Es ist schwer zu sagen, ob eingebettete Systeme eher Jobkiller sind oder neue Jobs schaffen. Eines ist jedoch sicher: Wenn Embedded Computing und KI aufeinandertreffen, ist das Bildungsniveau entscheidend. Die erste Welle in den frühen 2020er Jahren lässt noch Chancen für jede Art von Job. Ab den 2030er Jahren setzt sich eingebettete KI dann wirklich durch, vor allem in den Bereichen Transport und Fertigung.

*(Quelle: PwC, basierend auf OECD/PIAAC-Daten)*

## Faktoren, die den globalen Embedded-Computing-Markt beeinflussen

Was macht das Embedded-Computing zu einem so solide wachsenden Markt? Es gibt viele Gründe. Zwei negative Faktoren verlieren an Bedeutung: Erstens, Anfangsinvestitionen sind nicht mehr so oft notwendig, und zweitens, Hardware-Beschränkungen werden durch eine erfolgreiche Entwicklungsarbeit reduziert. Die positiven Faktoren sind in der Mehrheit und auf dem Vormarsch: Die Entwicklung wird von Tag zu Tag intelligenter, die Unterhaltungselektronik wird immer beliebter und die Intelligenz der KI nimmt zu. Auch die aufstrebenden Volkswirtschaften und das Wachstum des IoT gehören zu den positiven Faktoren..

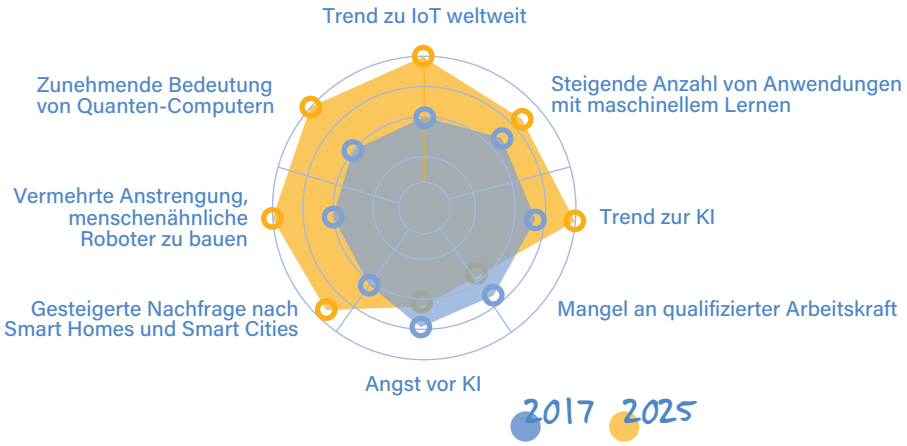
*(Quelle: Allied Market Research)*




## Faktoren, die den globalen Markt für Chips im ML-Bereich beeinflussen

Welche Faktoren bestimmen den Erfolg oder Misserfolg des Chip-Marktes für maschinelles Lernen? *Allied Market Research* prognostiziert, dass die Einbettung von ML in Chips immer beliebter wird. Das liegt nicht nur daran, dass das IoT die KI vorantreibt, oder dass menschenähnliche Roboter in Zukunft sehr gefragt sein werden, sondern auch daran, dass der Fachkräftemangel langsam zurückgeht, ebenso wie das Misstrauen der Öffentlichkeit gegenüber KI. Die Chancen wachsen, die Skepsis aber nicht.

*(Quelle: Allied Market Research)*




210004-02

# Fertigungsgerechtes Design leicht gemacht

Ein Beitrag von **Eurocircuits**

Fertigungsgerechtes Design (DFM) hat entscheidende Vorteile: Es senkt die Kosten bei der Prototypen- und Serienfertigung, erhöht die Qualität und Zuverlässigkeit der Baugruppen und schont wertvolle Ressourcen. Eurocircuits frei verfügbarer Visualizer macht DFM möglich: schnell, einfach und auf Anhieb richtig.

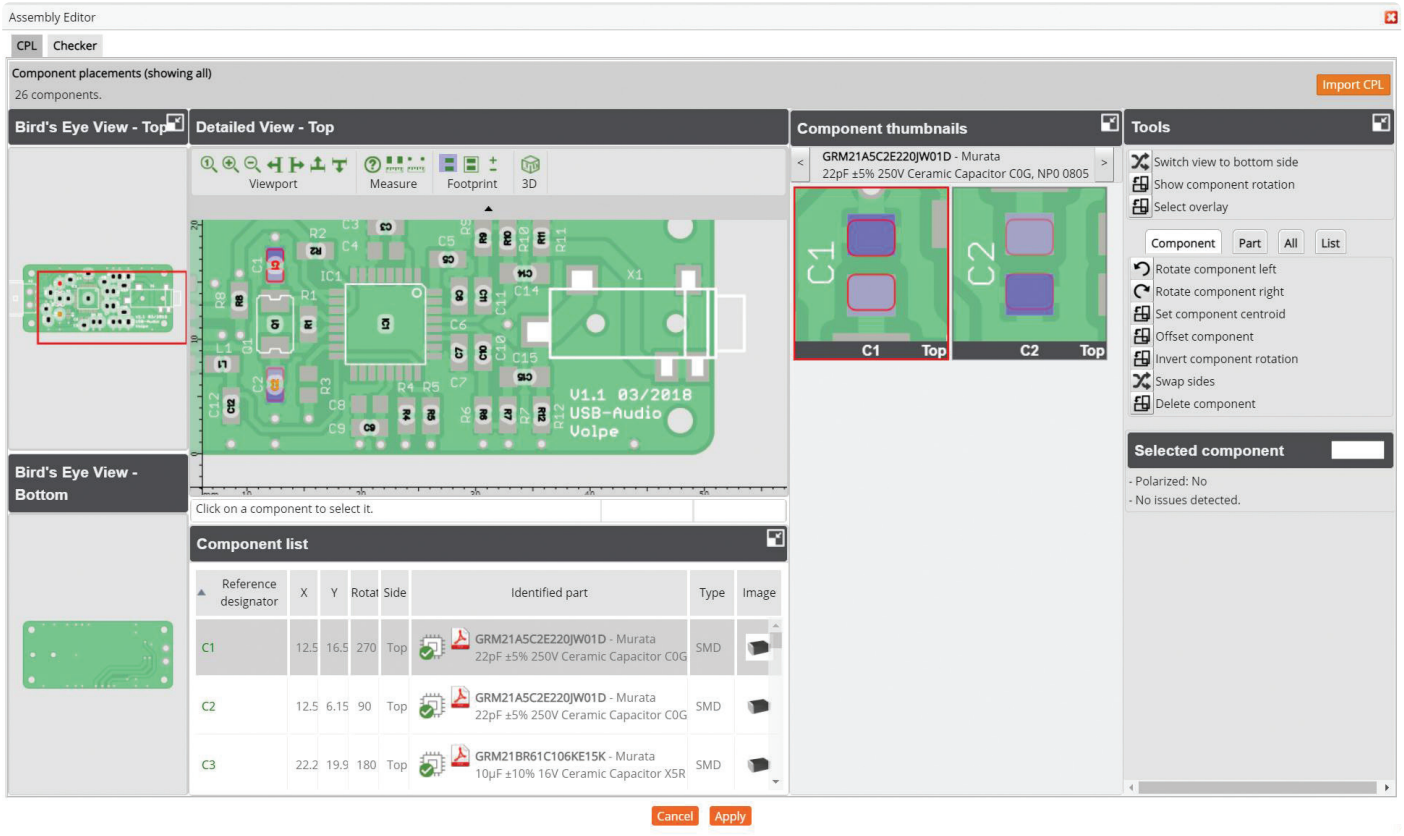



*Bild 1. Mit dem Eurocircuits Visualizer lassen sich Layout-Parameter validieren oder kostengünstigere Optionen finden..*

Beim Leiterplattendesign richtet sich das Augenmerk natürlich zuerst auf die elektrische Funktion. Doch die Leiterplatte muss produziert und mit Bauteilen bestückt werden. Diese Überlegung sollte so früh wie möglich in den Designprozess fließen. Denn der überwiegende Teil der späteren Fertigungskosten wird in der Prototypenphase festgelegt.

Mit dem Layout der Leiterplatte und der Auswahl und Art der Bauteile entscheiden Leiterplattendesigner über die Fertigungskosten und die Zuverlässigkeit der Hardware. Hinzu kommt: Ein kostenoptimierter Prototyp lässt sich zügig in die

Serienfertigung überführen und ermöglicht einen schnellen Markteintritt. Außerdem lässt sich die Fertigung für ein gut industrialisiertes Design leicht aus- oder verlagern. Designer tun gut daran, die Realität bei der Leiterplattenfertigung und Bestückung zu beachten. Eine Leiterplatte ist eine sehr komplexe und technisch anspruchsvolle Konstruktion aus verschiedenen Materialien und Prozessschritten mit unvermeidbaren Fertigungstoleranzen. Der erfahrene Leiterplattenhersteller weiß, wie man die Einflussfaktoren handhabt und den Prozess optimiert. Trotzdem sind Industrietoleranzen die Realität und unvermeidbar.

## Auf Anhieb richtiges Design in Industriequalität

Beim Konstruieren fehlerfreier und zuverlässiger Leiterplatten hilft das Wissen über die Fertigungsprozesse und Prozesstoleranzen. Eurocircuits hat seine Fertigungserfahrung in der Online-Engineering-Plattform Visualizer gebündelt (**Bild 1**). Der Visualizer ist ein leistungsstarkes Werkzeug zur Designprüfung und Verifikation. Der Visualizer steht kostenlos, in sechs Sprachen und rund um die Uhr zur Verfügung und ist zugleich ein nutzerfreundliches Online-Bestell- und Kalkulationssystem. Die Plattform wurde 2020 komplett überarbeitet und weiterentwickelt.

Im Visualizer integrierte interaktive DFM-Werkzeuge für die Leiterplattenfertigung und das Bestücken der Leiterplatte prüfen die Designdaten auf Vollständigkeit, zeigen kritische Stellen im Design und geben konkrete Vorschläge, um Designfehler zu beheben. Das Ziel ist immer, Lösungen, Alternativteile oder Konstruktionsänderungen vorzuschlagen, um ein fertigungsgerechtes Design zu erreichen (**Bild 2**).

Der Visualizer arbeitet mit über 700 Regeln zur Validierung der Parameterauswahl und verfügt über 900 vordefinierte Leiterplattenaufbauten für unterschiedliche Leiterplattentypen. PCB-Designer prüfen mit dem Visualizer auch die Daten für die Bestückung der Leiterplatte. Hierfür werden die BOM und CPL hochgeladen und die Daten validiert. Zentrales Element ist eine eigene Datenbank mit über 250.000 elektronischen Bauteilen.

Jede Leiterplatte bzw. elektronische Baugruppe wird vorab virtuell gefertigt. Damit erhalten Entwickler und Leiterplattendesigner ein auf Anhieb richtiges Design in Industriequalität – eine validierte Stückliste und validierte Bestückungs- und 3D-Daten der Baugruppe. Die validierten Daten verkürzen die Entwicklungszeiten und ermöglichen, Prototypen optimal in die Serienproduktion bei jedem beliebigen Fertiger zu führen.

210166-02

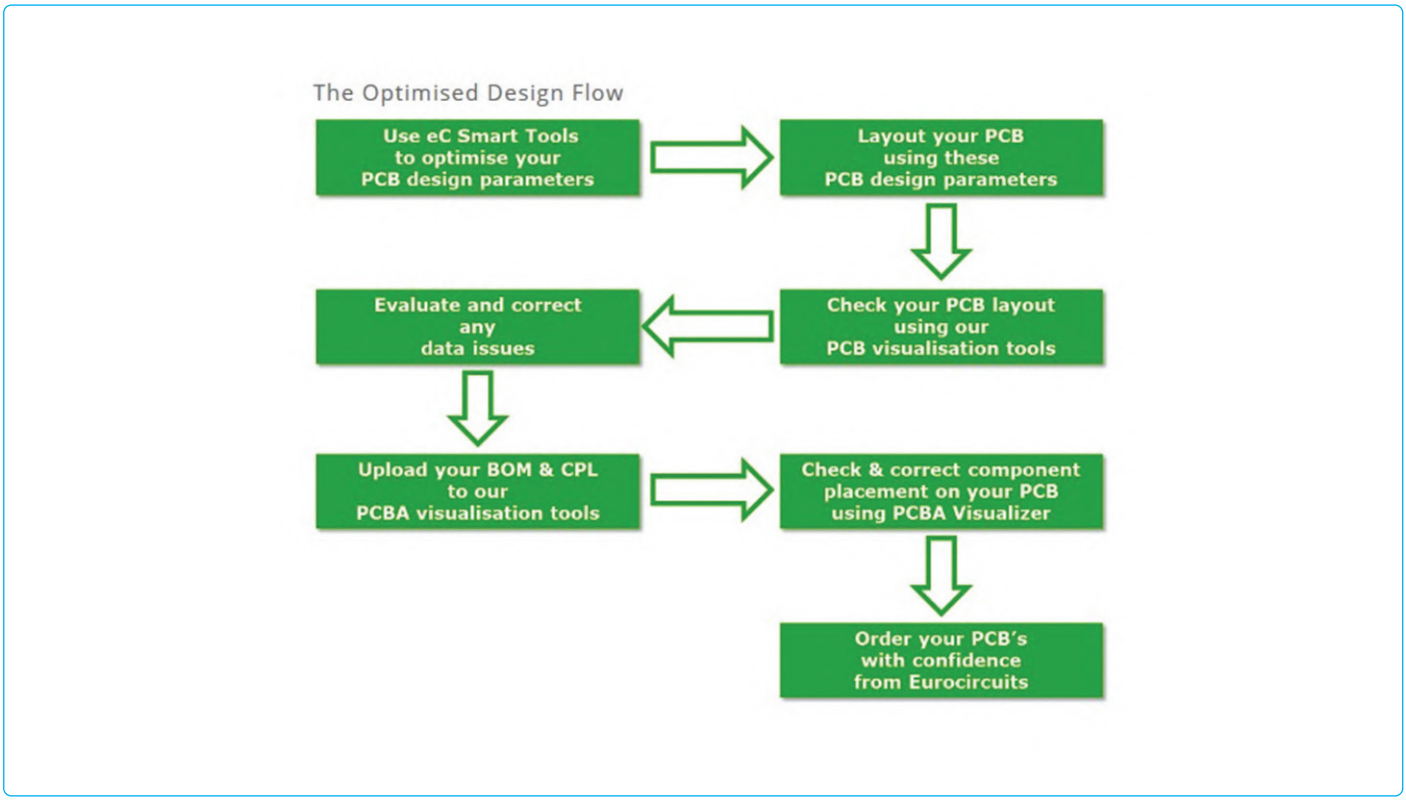



*Bild 2. Die Online-Werkzeuge von Eurocircuits unterstützen den optimalen Ablauf mit dem Ziel eines auf Anhieb richtigen Designs (Right First Time for Manufacture).*

## WEBLINK

[1] **Website von Eurocircuits**: www.eurocircuits.de

# Effektivere Analyse von Umweltsensordaten

Von **Mark Patrick** (Mouser Electronics)

## Durch automatisierte Cloud-gestützte Analyse

Für das Internet der Dinge (Internet of Things, IoT) werden in den nächsten Jahren voraussichtlich mehrere Milliarden Knotenpunkte in Betrieb genommen. Die Überwachung verschiedener wichtiger Umweltparameter wird eine der wichtigsten Anwendungen sein. Es ist nicht so einfach, die riesigen Datenmengen, die kontinuierlich erfasst werden, effizient zu analysieren und dann eine entsprechende Maßnahme zu ergreifen. Für maximale Effizienz und möglichst kurze Reaktionszeiten ist ein automatisierter Entscheidungsfindungsprozess notwendig.

Für das Internet der Dinge werden in den nächsten Jahren voraussichtlich mehrere Milliarden Knotenpunkte in Betrieb genommen. Mit dem Beginn der flächendeckenden Einführung dieser Technologie wird die Überwachung verschiedener wichtiger Umweltparameter wie beispielsweise Temperatur, Luftfeuchtigkeit, Kohlenmonoxidwerte usw. eine der wichtigsten Anwendungen sein. Auf der Grundlage dieser Messdaten können dann Entscheidungen getroffen und entsprechende Maßnahmen eingeleitet werden. Nehmen wir zum Beispiel eine Anlage zur chemischen Verarbeitung: Wenn der aus einem Schornstein austretende Gasgehalt einen vorgegebenen Schwellenwert überschreitet, kann es erforderlich sein, die Produktion zu diesem Zeitpunkt zu reduzieren. Es könnte sich aber auch um ein Anzeichen für einen längerfristigen Trend handeln, der untersucht werden muss. Möglicherweise muss die betroffene Verarbeitungsanlage gewartet werden oder hat das Ende ihrer Lebensdauer erreicht und sollte daher komplett erneuert werden.

Es ist nicht so einfach, die riesigen Datenmengen, die kontinuierlich von den Sensoren in einem weit verzweigten Netzwerk von IoT-Knotenpunkten erfasst werden, effizient zu analysieren und dann eine entsprechende Maßnahme zu ergreifen. Wenn es sich nur um Daten handelt, die von einer kleinen Anzahl von Knoten eingehen, mag das keine besondere Herausforderung sein, aber die Verarbeitung der riesigen Datenmengen in einem industriellen Steuerungs- oder Umweltüberwachungssystem ist etwas ganz anderes. Dies gilt ebenso für ein breites Spektrum potenzieller Anwendungsbereiche, wie z. B. Smart Homes/Intelligente Gebäude, Fabrikautomatisierung, Landwirtschaft, Intelligente Städte und mehr.

Für maximale Effizienz und möglichst kurze Reaktionszeiten ist ein automatisierter Entscheidungsfindungsprozess notwendig – und dafür ist der Einsatz fortschrittlicher Algorithmen erforderlich. Angesichts der schieren Anzahl von IoT-Knoten, die voraussichtlich beteiligt sein werden, ist es nicht effektiv, einfach eine Reihe von Regeln in einer Datenbank zu speichern und diese mit den empfangenen Daten abzugleichen. Eine solche Struktur kann sehr schnell überlastet werden. Es ist daher ein intelligenterer, aber dennoch schlanker Ansatz erforderlich.

### Der Rete-Algorithmus

Forschungen haben ergeben, dass der Rete-Algorithmus der Schlüssel dazu ist. Dieser Algorithmus wurde bereits in den späten 1970er Jahren entwickelt und wird in vielen verschiedenen Bereichen eingesetzt. Der Rete-Algorithmus stellt im Kern einen Mustervergleichsmechanismus bereit, mit dem die riesigen Mengen an Musterdaten (wie sie jetzt von IoT-Netzwerken erzeugt werden) schnell mit einer Datenbank verglichen werden können, die viele verschiedene Objekte enthält. Eine Iteration der Daten ist nicht erforderlich, da der Status eines bestimmten Musters während des gesamten Prozesses in einem Zwischenspeicher gehalten wird. Dadurch entfällt der Rückgriff auf zuvor angewandte Regeln und der gesamte Vorgang wird beschleunigt und ist somit wesentlich zeitsparender als alternative Verfahren.

Die cloudbasierte Analysesoftware von UrsaLeo läuft mit dem Thunderboard 2 Sensormodul von Silicon Labs sowie dem Raspberry Pi 3B+ (auf dem Yocto Linux vorinstalliert ist) und arbeitet nach diesem Prinzip. Die Software ermöglicht die Verarbeitung der erfassten Daten durch Anwendung einer baumbasierten Regelstruktur

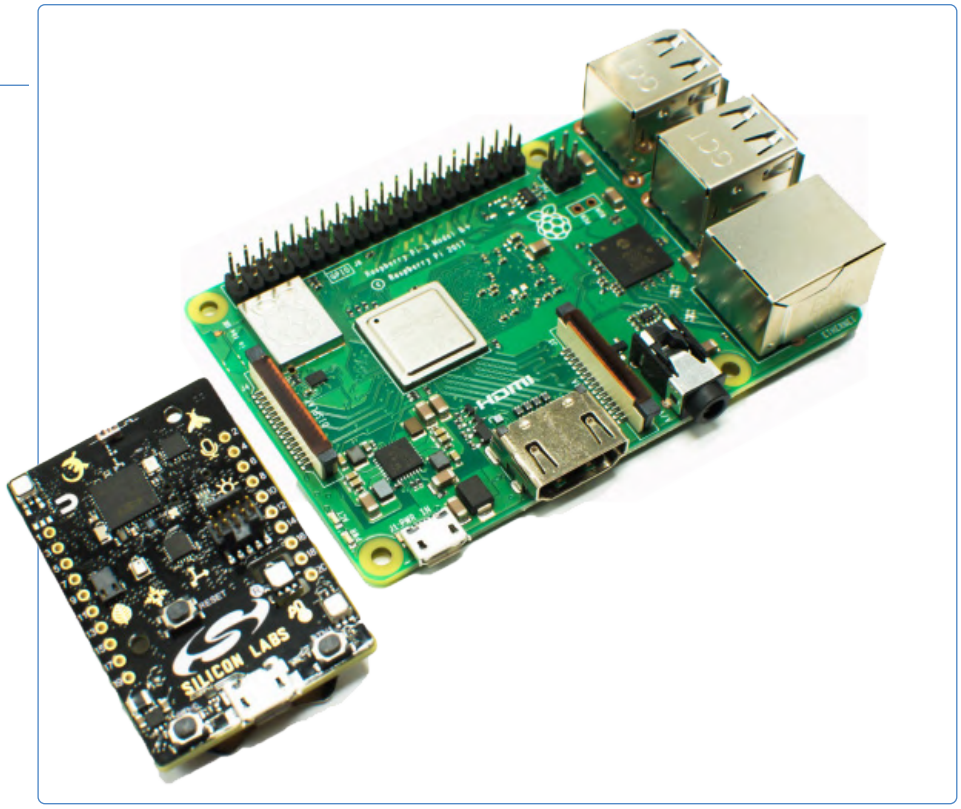



*Bild 1. UrsaLeo IoT-Hardware – bestehend aus einem Thunderboard 2-Modul und einem Raspberry Pi 3B+.*

**Über den Autor**

Als Technical Marketing Manager für EMEA bei Mouser Electronics ist Mark Patrick für die Erstellung und Verbreitung von technischen Inhalten in der Region verantwortlich - Inhalte, die für Mousers Strategie zur Unterstützung, Information und Inspiration der Elektronik-Branche von zentraler Bedeutung sind. Bevor er das Technische Marketing Team leitete, war Patrick Teil des EMEA-Lieferanten-Marketing-Teams und spielte eine wichtige Rolle beim Aufbau und der Entwicklung von Beziehungen zu wichtigen Produktionspartnern. Zusätzlich zu einer Vielzahl von technischen und Marketing-Positionen war Patrick acht Jahre lang bei Texas Instruments in den Bereichen Anwendungsunterstützung und technischer Vertrieb tätig.

für den Musterabgleich, anstatt immer wieder auf eine Lookup-Tabelle zurückgreifen zu müssen. Dadurch lassen sich fundiertere Entscheidungen treffen, ohne dass das System mit unerwünschten Latenzen belastet wird. Benutzerdefinierte Dashboards mit Anzeige der kompilierten Datensätze können aufgerufen und Trigger eingerichtet werden, um Mitarbeiter zu informieren, wenn bestimmte Ereignisse auftreten, bei denen ein menschliches Eingreifen erforderlich ist.

## Das System kann selbst mit einem mittelgroßen Server 500.000 Meldungen pro Sekunde verarbeiten.

Mit dem intuitiv zu bedienenden, visuellen Editor lassen sich Regeln erstellen, die genau auf die Anforderungen der jeweiligen Anwendung abgestimmt sind. So kann beispielsweise ein Trigger ausgelöst werden, wenn eine Nachricht eintrifft, oder umgekehrt, wenn eine Nachricht innerhalb eines bestimmten Zeitraums nicht empfangen wurde. Alternativ kann ein Trigger in regelmäßigen Abständen (nach Ablauf einer Stunde, eines Tages oder einer Woche) ausgelöst werden, um eine Langzeitüberwachung zu ermöglichen. Außerdem lassen sich Geo-Fencing-Regeln definieren, die ausgelöst werden, wenn sich ein mobiler Knoten außerhalb / innerhalb eines bestimmten Bereichs bewegt. Diese Funktion kann z.B. verwendet werden, um Gabelstapler in der Fabrikhalle innerhalb bestimmter Grenzen zu halten, oder in Anwendungen für das Flottenmanagement oder die Verfolgung von Viehbeständen. Wenn ein erfasster Datenwert ungewöhnlich ist (z.B. wenn er konstant bleibt oder dauerhaft auf Null steht), kann dies darauf hindeuten, dass der Sensor nicht richtig funktioniert. Ein Techniker kann dann gezielt vor Ort geschickt werden, um die notwendigen Reparaturen vorzunehmen.

### Vielfalt vernetzter Sensoren

Das Thunderboard 2-Modul kann Umgebungsdaten von zahlreichen vernetzten Sensoren erfassen, wie z. B. Umgebungslicht, Luftqualität, Luftdruck, relative Luftfeuchtigkeit und Temperatur. Außerdem bietet es Funktionen zur Gaserkennung. Ein 6-Achsen-Trägheitssensor (zur Raumorientierung) und ein Hall-Effekt-Sensor (zur geografischen Orientierung) sind ebenfalls enthalten. Das zugehörige Raspberry-Pi-Board dient als IoT-Gateway (mit direkter Anbindung an die Google-Cloud-Plattform). Mit diesem Board lassen sich alle gesammelten Daten entweder über die drahtlose (WLAN) oder die drahtgebundene (Ethernet) Verbindung in die Cloud übertragen. Je nach Anwendungsszenario kann die Hardware über USB oder über eine Li-Ionen-Knopfzelle mit Strom versorgt werden.

Durch die baumbasierte Struktur ist das System vollständig skalierbar. Es kann eine beliebige Anzahl von verbundenen IoT-Knoten innerhalb eines Netzwerks verwalten und ist nicht auf eine bestimmte Anzahl begrenzt. Somit ist es möglich, Hunderttausende von Ereignissen/Alarmen pro Sekunde zu verarbeiten. Das System kann selbst mit einem mittelgroßen Server bereits 500.000 Meldungen pro Sekunde verarbeiten. Mit einem Hochleistungsserver lässt sich dies auf 1 bis 2 Millionen steigern. Die Kombination aus hochoptimierter, cloudfähiger Hardware und leistungsstarken Algorithmen sorgt für eine wesentlich zeiteffizientere Verarbeitung von IoT-Daten als bislang. Dies ermöglicht den Aufbau von IoT-Anwendungen mit Hunderttausenden von vernetzten Knoten und somit auch die Ausschöpfung der großen Vorteile dieser Technologie in den verschiedensten Industriezweigen.

210036-02

# Embedded-Systeme mit RISC-V-Prozessoren

## Einführung und Marktübersicht

Von **Allan He** (Embedded System Association, China)

Dieser Artikel dreht sich um die Ursprünge und die Entwicklung der ISA von RISC-V. Nach einer kurzen Beschreibung der technischen Merkmale geht es um die Auswahlkriterien verschiedener RISC-V-CPU-Cores und SoC-Design-Plattformen. Weiter werden die Eigenschaften der RISC-V-Befehlserweiterungen sowie die Open-Source- und kommerziellen Software-Entwicklungs-Tools vorgestellt.

In den vergangenen zwei Jahrzehnten hatte die Firma ARM einige Meilensteine bei mobilen und Embedded-Systemen gesetzt. Heute sieht man, wie ARM seine Marktführerschaft im IoT etabliert und dass kommerzielle Prozessorarchitekturen wie etwa MIPS im Niedergang begriffen sind. Nachdem ARM auch die zuvor von x86 dominierten Märkte für PCs und Server betreten hat, steht Intel unter Druck. Genau in diesem Moment zieht der Aufstieg von RISC-V die Aufmerksamkeit der Branche auf sich. IT-Giganten waren auf der Suche nach Alternativen zu ARM. Die Offenheit der ISA (Instruction Set Architecture) von RISC-V macht diese Plattform zu einer brauchbare Option. Im akademischen Bereich verdrängt RISC-V nun in rasantem Tempo MIPS und x86 als Architektur der Wahl in Lehrbüchern. Da Regierungen und Unternehmen auf RISC-V als Standard setzen, entsteht eine Flut neuer CPU- und SoC-Designs, ein wachsendes Ökosystem und eine immer engagiertere Entwickler-Community.

### Was ist RISC-V?

Eine ISA umfasst die von einer CPU unterstützten Befehle und deren Byte-Kodierung. Sie ist die Brücke zwischen Computer-Software und Hardware. CPU-Familien wie x86, PowerPC oder ARM, haben ihre eigenen individuellen ISAs, aber RISC-V ist derzeit die einzige offene ISA. Bei der ISA von RISC-V handelt es sich nicht um einen konkreten CPU-Chip, sondern um Spezifikationen und Standards für Prozessoranweisungen - tatsächlich ist es nicht einmal ein vollständiger Befehlssatz. RISC-V hat seinen Ursprung an der University of California, Berkeley. Im Sommer 2010 startete dort Prof. Krste Asanovic mit seinen Studenten Andrew Waterman und Yunsup Lee ein dreimonatiges Forschungsprojekt. Das Ziel war es, eine offene ISA zu entwickeln, um die technische Komplexität und die IP-Probleme von x86 und ARM zu umgehen.

2015 wurde dann die RISC-V Foundation als gemeinnützige Organisation gegründet. Der Vorstand der Stiftung besteht aus Vertretern von Bluespec, Google, Microsemi, NVIDIA, NXP, der UC Berkeley und Western Digital. Sie wird von Prof. Asanovic geleitet. Die Stiftung legt die Standards fest und unterstützt das Ökosystem von RISC-V-Cores. Die Standards sind öffentlich und stehen zum Download bereit. Es gibt mehr als 300 zahlende Mitglieder der Stiftung, darunter Qualcomm, NXP, Alibaba und Huawei. Die Mitglieder sind berechtigt, RISC-V-Markenzeichen zu verwenden. RISC-V ist unter der Open-Source BSD-Lizenz lizenziert. Jede Firma, jede Institution und jede Einzelperson kann eine CPU mit RISC-V-Architektur entwickeln. Im vergangenen Jahr wurde die RISC-V Foundation in „RISC-V International Association" umbenannt und verlegte ihren Sitz von den USA in die Schweiz, wo sie im März 2020 registriert wurde [1].

Nach einem Jahrzehnt der Entwicklung zeichnen sich bedeutende Errungenschaften bei CPU-IP-Core, Plattform, SoC und Anwendungen ab. Einige Anwendungsbeispiele für RISC-V-Cores: Die SSD- und HDD-Controller von Western Digital sind mit SweRV-Kernen ausgestattet, das Bluetooth-Headset TWS von Shenzhen Bluetrum, die AIoT-SoC-Chips K210 von China Canaan Creative und die 32-bittigen Hochgeschwindigkeits-MCUs von Nanjing Qinheng Microelectronics.

### Der Befehlssatz von RISC-V

Der Befehlssatz von RISC-V verwendet ein modulares Design, um Befehle zu organisieren, wobei ein Buchstabe ein Modul darstellt. Der grundlegendste und am meisten benötigte Befehlssatz in RISC-V ist das Integer-Set mit der Bezeichnung „I". Selbst mit einem einzigen Set lässt sich ein kompletter Software-Compiler implementieren. Andere Befehlssätze wie etwa *M, A, F, D* oder *C* sind optional. Bei der RISC-V-Variante RV32IMAC sind also Befehlssätze *I, M, A* und *C* implementiert. Die *32I*- und *64I*-Sets in RISC-V sind bereits „eingefroren", ebenso die *MAFDQC*-Erweiterungen. *32E, 128I, JBJTPV* und *ZAM*-Zugriffserweiterungen sind noch in der Entwicklung. Befehlssatzerweiterungen sind einzigartige technische Merkmale von RISC-V.

Dank der Kooperation mit Mitgliedsinstitutionen und der Industrie legt RISC-V hier ein stabiles Wachstum an den Tag.

## RISC-V-Core, -SoC und ChipRISC-V-Core

Vor dem Eintauchen in die Entwicklung von Embedded-Systemen mit RISC-V sollen zunächst einige Konzepte geklärt sein: RISC-V Prozessor Core (als *Core* bezeichnet), SoC-Plattform und SoC-Chip. Seit den Anfängen von RISC-V hat es Dutzende RISC-V-Cores und SoC-Chips gegeben. Einige davon sind Open-Source, andere interne Projekte und wieder andere sind Cores und Plattformen für Firmen. Bei SweRV (RV32IMC) von Western Digital handelt es sich um eine 32-Bit-Architektur mit sequentieller Ausführung im Zwei-Wege-Superskalar-Design mit neunstufiger Pipeline in 28-nm-Technologie. Beim Maximaltakt von 1,8-GHz erreicht die SweRV-CPU eine Leistung von 4,9 CoreMark/MHz und liegt damit leicht über dem ARM Cortex A15. SweRV ist ein Open-Source-Projekt und wurde bereits auf SSD/HDD-Controllern von Western Digital eingesetzt.

Es gibt eine ganze Reihe von Beispielen für Open-Source-RISC-V-Cores: Rocket Core von UC Berkeley ist ein klassisches RV64-Design, während BOOM Core derselben Universität auf höhere Leistung abzielt. Der Zero-riscy der ETH Zürich ist ein typisches RV32-Design, während ihr R15CY Core für Ultra-Low-Power-Lösungen gedacht ist und als RV32E konfiguriert werden kann. Der Core PicoRV32 von Clifford Wolf konzentriert sich auf einen kleinen Footprint und die Optimierung des Takts.

Obwohl die Open-Source-Cores für Forschung und Lehre geeignet sind, gibt es noch mehr zu berücksichtigen, wenn man ein kommerzielles SoC entwerfen will. Yunsup Lee, ein Mitentwickler von RISC-V, gründete SiFive. Im Jahr 2017 kündigte SiFive seine erste RISC-V-Kern- und SoC-Familie zusammen mit Software und Entwicklungsboards an. Zu diesen Chips gehören die in 28 nm gefertigte und Linux unterstützende 64-Bit-Multicore-CPU U500 und der in 180-nm gefertigte preiswerte IoT-Core E300. Viele weitere Hersteller wie etwa Codasip, Syntacore, T-Head, Andes sowie Startups wie Nuclei System Technology beschäftigen sich mit der Entwicklung von RISC-V-Prozessorkernen.

## RISC-V als SoC-Plattform

PULPino der ETH Zürich, das Open-Source-Projekt LowRISC und der Open-Source-SoC-Generator Rocket Chip der UC Berkeley (basierend auf Chisel) sind einige der bekanntesten RISC-V-SoC-Plattformen. In China ist Zhenbo Hu als Gründer von Nuclei ein Initiator der populären Open-Source-Softcore-SoC-Plattform Hummingbird E200 [3].

## RISC-V-CPUs und MCUs

In den letzten Jahren erlebten RISC-V-SoCs einen rasanten Aufstieg. Ein Beispiel für bekannte SoCs ist der MCU-Chip von GigaDevice (GD32VF103), der auf dem Bumblebee-Kern von Nuclei (RV32IMAC) basiert. Die GD32VF103-Serie bietet einen Takt von 108 MHz, 16...128 KB On-Chip-Flash, 6...32 KB SRAM, vier 16-Bit-AP-Timer, zwei 16-Bit-Basis-Timer und zwei Mehrkanal-DMA-Controller. Die GD32VF103-MCU verfügt über einen neu entwickelten Enhanced Core Local Interrupt Controller (ECLIC), der bis zu 68 Peripherie-Interrupts und 16 Interrupt-Prioritätsstufen unterstützt, was die MCU für Echtzeitanwendungen hoher Leistung empfiehlt.

Es gibt mehrere GD32VF103 MCU-Entwicklungsboards: GD32VF103V-EVAL ist ein Exemplar mit vollem Funktionsumfang, und GD32VF103-START ist für den Einstieg bzw. die Ausbildung gedacht. Zusätzlich gibt es die Boards RV-Start und Sipeed Longan Nano von Nuclei [4] sowie das neueste RISC-V GD32 EVAL Kit von IAR [5] (siehe **Bild 1**).


**Über den Autor**

Allan He ist Spezialist für Embedded-Systeme. Er war stellvertretender Leiter der Redaktion des „Microcontrollers and Embedded Systems Journal“. Allan ist der Gründer einer Firma im Bereich Embedded-Software. Er hält häufig Vorträge auf Konferenzen und lehrt zu den Themen Embedded-Betriebssysteme und IoT an Universitäten und Firmen.


K210 von China Canaan Creative ist ein RISC-V-AIoT-SoC. Sein Prozessor besteht aus zwei RV64GC-Cores mit MAFD-Befehlserweiterungen. Der K210 enthält eine KPU-CPU für neuronale Netzwerke, der menschliche Gesichter und Objekte in Echtzeit erkennt. Der FFT-Beschleuniger des K210 ist in Hardware implementiert [6].

NXP RV32M1 integriert vier Kerne: RISC-V RI5CY, RISC-V Zero-riscy, ARM Cortex M4 und ARM Cortex M0+ (**Bild 2**). Aus professioneller Sicht ist der RV32M1 eher ein Prototyp, der von Entwicklern evaluiert werden kann. NXP hat die Open ISA-Community (open-isa.org) gegründet und leistet so einen Beitrag zur Wartung der Toolchain und zum Aufbau des Software-Ökosystems. NXP spielte eine wichtige Rolle bei der frühen Verbreitung von RISC-V, indem Entwicklungsboards verschickt [7] und Wettbewerbe/Hackathons in China veranstaltet wurden.

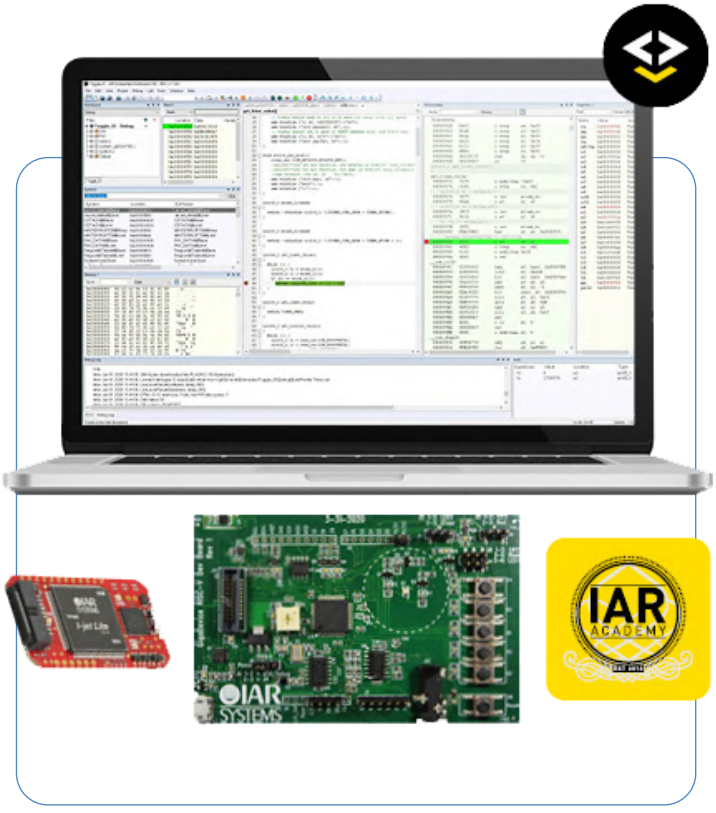



*Bild 1. IAR RISC-V GD32V Evaluation Kit (Quelle: iar.com).*

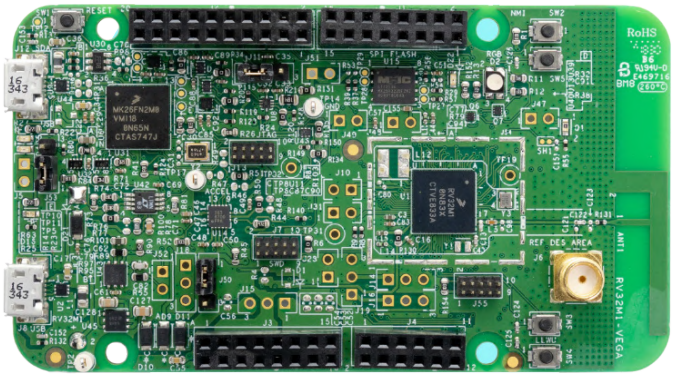


*Bild 2. NXP RISC-V Vega Development Board (Quelle: open-isa.org).*

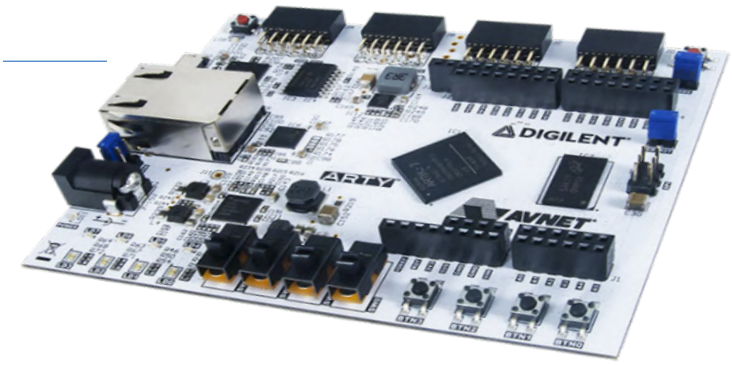


*Bild 3. Arty FPGA-Entwicklungsboard für RISC-V Softcore (Quelle: digilentinc.com).*

Microchips PolarFire SoC ist ein kostengünstiges Multi-Core-SoC-FPGA aus vier 64-Bit RV64GC RISC-V-Cores und einem RV64IMAC-Core für Echtzeit- und Überwachungsaufgaben. PolarFire unterstützt Linux und eignet sich besonders für die Entwicklung industrieller Steuerungs- und IoT-Anwendungen.

## Auswahl zwischen RISC-V Core, SoC und Chip

Wer sich an der Forschung, Entwicklung und Ausbildung rund um RISC-V beteiligen möchte, kann sich an folgenden Aspekten orientieren:

- Chip-Designer können sich dafür entscheiden, eigene Chips mit RISC-V-Cores und -SoCs zu konstruieren. Beispielsweise eignet sich die PULPino-Plattform nebst den RI5CY- und Zero-Riscy-Cores für die Entwicklung von SoCs. Einige Firmen und Universitäten in China verwenden diese in ihren Projekten.
- Auch Embedded-System- und IoT-Entwickler können von RISC-V-SoC-Chips profitieren. Für die MCUs der GD32VF103-Serie gibt es zahlreiche Entwicklungsboards, die mehrere Toolchains unterstützen. Für AIoT-Anwendungen bietet sich der K210 an, der von einem ausgereiften SDK begleitet wird und sowohl FreeRTOS als auch Bare Metal unterstützt. Kürzlich wurde mit Linux 5.8 die Unterstützung von K210 RISC-V zum Main-Release hinzugefügt. K210 wurde bereits erfolgreich beim maschinellen Sehen und Hören für die Gesichtserkennung und das intelligente Ablesen von Energiezählern eingesetzt.
- Universitäten und Forschungseinrichtungen können den Open-Source-RISC-V-Core auf FPGAs für die Bereichen Computerarchitektur, Betriebssysteme, Compiler und Embedded Systems in Forschung und Lehre einsetzen. Das Arty-FPGA-Entwicklungsboard hat die quelloffene Freedom E310-MCU von SiFive implementiert (siehe **Bild 3**) und ist eine ausgereifte Plattform mit gutem Software-Toolchain-Support [8].

## Vorteile von RISC-V in Embedded-Systemen

Offensichtlich sind Embedded-Systeme, IoT und AIoT die aktivsten Anwendungen für RISC-V. Für embedded Systeme ergeben sich eine ganze Reihe von Vorteilen. Die Wichtigsten:

- Open-Source. Man kann auch mit offenen Systemen erfolgreich sein. Außerdem ist dies die beste Lerngrundlage für Studenten und Ingenieure. Open-Source-ISA bedeutet, dass Sie ihre eigene, auf spezifische Anwendungsszenarien ausgerichtete Chip-Architektur entwerfen können. Niedrigste Kosten senken die Hürden für den Einstieg in das Chip-Design.
- Einfachheit und Flexibilität. Es gibt 50 grundlegende RISC-V-Befehle in vier zentralen Befehlssätzen. Dank modularen Design

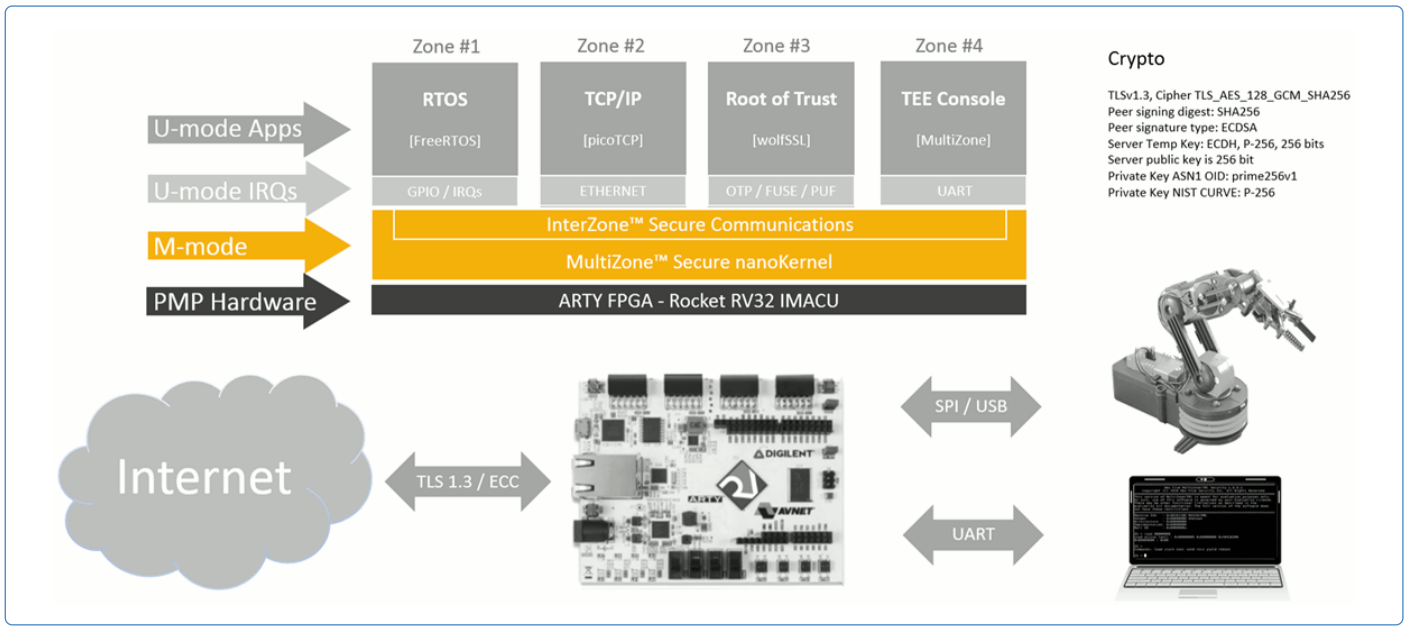



*Bild 4. Referenzanwendung des MultiZone Security IoT Stack (Quelle: riscv.org).*

kann man eine RISC-V-CPU mit einem vereinfachten Befehlssatz konzipieren, was zu niedriger Codedichte und geringem Energiebedarf führt. Dank der Flexibilität von RISC-V werden eine breite Palette von Prozessoren vom 8051 bis zu ARMs A-Serie unterstützt.

- Effizienz und Sicherheit. RISC-V mit seinem reservierten Code-Raum kann leicht mit benutzerdefinierten Befehlen erweitert werden. Befehlserweiterungen können Berechnungen beschleunigen und die IoT-Sicherheit verbessern. Ein allgemeiner Ansatz bei der IoT-Sicherheit besteht darin, die Trusted Execution Environment (TEE) von der nicht vertrauenswürdigen Umgebung abzugrenzen. Hardware-TEE ist Teil der ISA-Spezifikation von RISC-V und kann auf jedem RISC-V-Chip implementiert werden. Die Spezifikationen beinhalten auch Physical Memory Protection (PMP).

Hex-Five hat eine Demonstration von RV32 Core mit Hardware-basierter IoT- Sicherheits-Domänentrennung [9] (siehe **Bild 4**) vorgelegt. Der Code ist auf GitHub [10] verfügbar. Zum Design gehören ein X300 Bitstream SoC mit Rocket RV32IMAC Core und ein Digilent Arty A7 Entwicklungsboard mit Ethernet-Unterstützung. Die Software MultiZone definiert vier Zonen des TEE:

- Zone 1: FreeRTOS, führt CLI-, PWM-LED- und Roboterarm-Aufgaben aus.
- Zone 2: TCP/IP mit PicoTCP, Kanal verschlüsselt mit TLS.
- Zone 3: WolfSSL TLS 1.3, vertrauenswürdige Wurzel, Verschlüsselungsschlüssel/Passwörter, geschützte Dateien.
- Zone 4: UART, TEE-Konsole.

## RISC-V embedded Software Ökosystem

Das Software-Ökosystem von RISC-V besteht größtenteils aus Open-Source-Software mit einigen kommerziellen Programmen. Mehrere GNU-Open-Source-Toolchains wie z.B. der C-Compiler riscv-gcc, die riscv-binutils (Tools/Linker/Assembler), der Debugger riscv-gdb und die Open-Source Debugging-Software OpenOCD (für PC) sowie die JTAG-Hardware J-link unterstützen RISC-V. OpenOCD mit GDB-Server und Unterstützung von GDB-Befehlen ist eingebaut.

### GNU Open-Source Toolchains

Zu den gängigen Open-Source-IDEs gehören SiFive Freedom Studio, AndesSight und die Nuclei Studio IDE. Alle basieren auf Eclipse und sind für die CPU-Kerne der jeweiligen Hersteller optimiert. Ein Entwickler kann das JDK, die Eclipse IDE für C/C++, die GNU MCU Eclipse Build-Tools und die Toolchain riscv32-unknown-elf-gcc herunterladen, um seine eigene RISC-V-IDE zu erstellen. Der QEUM-Prozessoremulator unterstützt bereits den RV32/RV64-Befehlssatz„ Entwickler können Software und Betriebssysteme auswählen, die auf dem Emulator laufen sollen, darunter FreeRTOS, Zephyr und Linux [11].

Firmen im Bereich Embedded-Software haben bereits begonnen, RISC-V RV32 und einige Cores/Chips (z.B. GD32VF103 und SiFive E310) zu unterstützen. Die Embedded Workbench for RISC-V der schwedischen Firma IAR und das Embedded Studio RISC-V der deutschen Firma SEGGER gehören zu den beliebtesten Produkten auf dem Markt.

### IAR Embedded Workbench

Embedded Workbench von IAR ist ein Entwicklungs-Tool, das ähnliche Bedeutung wie Keil MDK für Embedded-Systeme und MCUs hat. Es bietet eine exzellente Optimierung von Code-Größe und -Effizienz. Entwickler können es zum Kompilieren, Analysieren und Debuggen von Code einsetzen. Die Features der neuesten Version 1.30 der Embedded Workbench for RISC-V sind:

- Weitere Optimierung des Compilers und der Laufzeitbibliothek.
- Unterstützung von DSP-Befehlen in der P-Erweiterung und der Packed-SIMD-Spezifikation.
- Unterstützung von Debugging auf Basis des Nexus IEEE-ISTO-5001-Protokolls und der SiFive Insight Lösung.
- Unterstützung für mehr als ein Dutzend MCUs (z.B. GigaDevice GD32V).

### SEGGER Embedded Studio

SEGGER ist bekannt für seinen Hardware-JLINK-Debugger für Embedded-Systeme. Das Embedded Studio von SEGGER (IAR) unterstützt seit 2017 auch RISC-V. Zusätzlich zu JLINK ist Embedded Studio kompatibel zu anderen Debugging-Schnittstellen wie OpenOCD und GD-Link. Embedded-Studio bietet umfassende Unterstützung für den Single-Core-RV32-Befehlssatz, einschließlich RV32I, RV32IMA, RV32IMAC, RV32IMAF, RV32IMAFC, RV32G und RV32GC; die Unterstützung für RV64 und Multi-Core wird angeblich derzeit getestet. Es unterstützt außerdem den RISC-V-Core von Nuclei Technology, Real-Time Transfer (RTT) zwischen Host und Debugger sowie das Software-Analysetool SEGGER SystemView.


SIGLENT®

*Das Beste Preis-Leistungs-Verhältnis in der Messtechnik*

Schneller zur Marktreife dank innovativer Messtechnik

- Oszilloskope 50MHz –> 1GHz
- Signal Generatoren 10MHz –> 500MHz
- Labornetzteile –> 200W
- Elektronische Lasten 200 –> 300W
- Multimeter 4½ –> 6½ Stellen

***Siglent HF-Messgeräte: Helfer bei der Entwicklung fehlerfreier und zuverlässiger Kommunikation***

**Spektrum- und Vektornetzwerkanalysator**

- EMI-Pre-Compliance
- Modulationsanalyse
- Distance-to-Fault Analyse

**HF-Signalquellen**

- mit/ohne IQ-Modulation

**Siglent Technologies Germany GmbH**
www.siglenteu.com
Info-eu@siglent.com

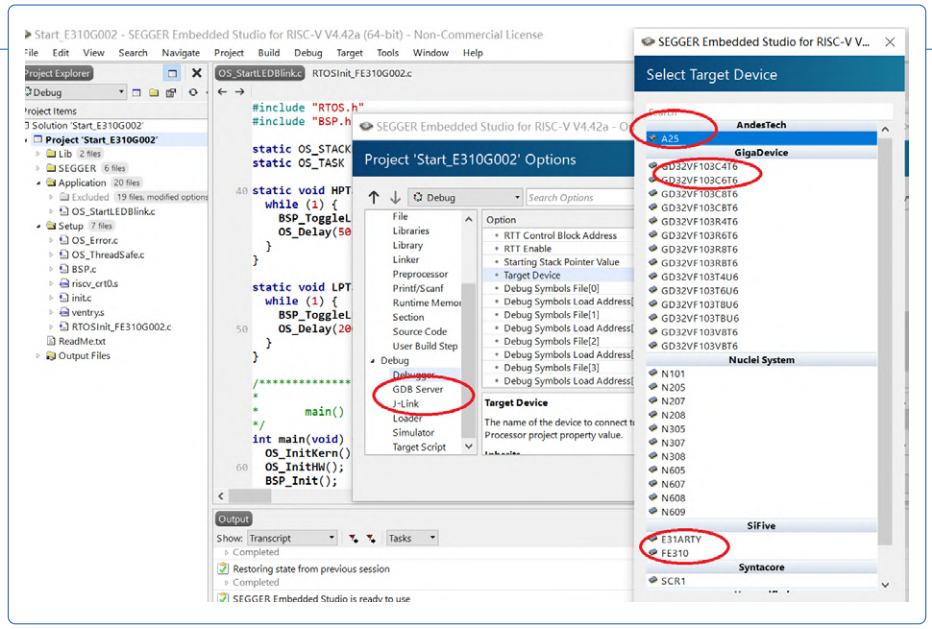



*Bild 5. SEGGER Embedded Studio.,*

Embedded Studio ist ein integriertes Entwicklungs-Tool. Es nutzt laut SEGGER die C/C++-Compiler Clang/LLVM und GCC. Außerdem unterstützt es auch externe Toolchains. Ein Beispiel: Ein Embedded-Studio-Projekt von Nuclei verwendet standardmäßig die von Nuclei optimierte GCC-Toolchain. Weitere Informationen zu Embedded Studio sind in Rolf Seggers Blog *The SEGGER Compiler* [12] zu finden.
Embedded Studio ist für Studenten/Dozenten und nicht-kommerzielle Nutzung kostenlos. Man kann es sofort nach Zustimmung zur Lizenzvereinbarung einsetzen.

## Embedded-Betriebssysteme

Ein RISC-V-Prozessor muss logischerweise von einem Embedded-Betriebssystem unterstützt werden. Es gibt bereits RISC-V-Portierungen in FreeRTOS, Zephyr OS, ThreadX und µC/OS. Für FreeRTOS V. 10.3 gibt es bereits Portierungen auf NXP Vega und SiFive Freedom HiFive1-revB Boards. Dabei werden GCC und IAR unterstützt. SEGGER embOS hat eine Demo mit emWin auf dem Board GD32VF103-EVAL. SiFive hat die Amazon FreeRTOS-Portierung auf Github veröffentlicht. Derzeit können das Learn Inventor Education Board von SiFive und die Andes Corvette-1 N25-Plattformen davon profitieren. In China hat RT-Thread eine Demo auf HiFive1-revB, Huawei LiteOS auf dem Board GD32VF103-EVAL und TencentOS Tiny die Überwachungsanwendung PM2.5 auf dem Board GD32VF103 vorgestellt.
Da die RISC-V-ISA den High-End-Computing-Markt erreichen will, ist es für RISC-V entscheidend, Unterstützung durch Linux zu erhalten. Dank der Ingenieure von Andes, Western Digital und SiFive hat es schon einige Entwicklungen gegeben. Es gibt Demos, in denen Fedora und Debian auf SiFive HiFive Unleashed laufen. An der neuen Prozes-

## WEBLINKS UND LITERATUR


[1] **Geschichte von RISC-V**: https://riscv.org/about/history/

[2] **D. Patterson und A. Waterman, „The RISC-V Reader: An Open Architecture Atlas“**, Strawberry Canyon LLC, 2018

[3] **Hu, Zhenbo, „RISC-V Architecture and Introduction to Embeddded System Development“,** China Posts & Telecom Press, 2019, S. 34-39

[4] **Chinesische RISC-V-Community**: https://www.riscv-mcu.com/

[5] **Evaluierungskit für RISC-V**: www.iar.com/iar-embedded-workbench/tools-for-risc-v/evaluation-kit-for-risc-v/

[6] **Kendryte K210**: https://canaan-creative.com/developer

[7] **V32M1-VEGA Development Board User Guide, aus**: https://github.com/open-isa-org/open-isa.org/releases/download/1.0.0/Documentation.zip

[8] **Arty FPGA Evaluation Kit Hands-On + Verifikation von RISC-V-Befehlen mit dem SiFve Chip (chinesischer Blog)**: https://www.cnblogs.com/zjutlitao/p/9745365.html

[9] **Das erste TEE für RISC-V**: https://hex-five.com/multizone-security-sdk/

[10] **Source-Code auf GitHub**: https://github.com/hex-five/multizone-iot-sdk

[11] **RISC-V Foundation, „RISC-V - Getting Started Guide“**, Mai 2019, S. 5-14

[12] **Der Segger-Compiler**: https://blog.segger.com/the-segger-compiler/

[13] **Linux für RISC-V**: https://kernelnewbies.org/Linux_5.8#RISCV

[14] **RISC-V Computer Architecture Course (enthält Lehrmaterial und praktische Übungen für Studenten)**: https://bit.ly/3tXvr1h

[15] **rCore OS (Chinesische Website)**: https://rcore.gitbook.io

[16] **Embedded System und IoT Development Online Seminar (3)**: http://hexiaoqing.net/courses/

[17] **S. Greengard, „Will RISC-V Revolutionize Computing?“**, Communication of the ACM, Band 63 Nr. 5, Mai 2020

sorarchitektur RISC-V haben Open-Source-Linux-Entwickler offensichtlich großes Interesse. Da es bisher leider noch nicht viele preiswerte RISC-V-Entwicklungs-Boards gab, war es für Entwickler schwierig mit RISC-V zu arbeiten. Vor kurzem wurde mit Linux 5.8 die offizielle Unterstützung von K210 [13] hinzugefügt, was bedeutet, dass K210 nicht mehr nur von no-MMU Linux unterstützt wird. Entwickler können jetzt leicht ein K210-Board (RV64GC) bekommen, um damit zu arbeiten.

## Ausbildung und Entwicklung

Im September 2020 kündigte Imagination Technologies einen Kurs für Studenten über die RISC-V-basierte Computer-Architektur mit dem Titel „RVfpga: Understanding Computer Architecture" [14] an. Der Kurs wurde von Imagination und Sarah Harris Harvey vom Mudd College entwickelt. Harris ist Mitautorin des bekannten Lehrbuchs „Digital Design & Computer Architecture". Sie meint: *„RISC-V übetrifft frühere Prozessor-Generationen in jeder erdenklichen Hinsicht, vom Stromverbrauch über die Leistung bis hin zu erhöhter Sicherheit. Als ein weiterer großer Schritt vorwärts in der Computerarchitektur ist es für Studenten wichtig, RISC-V auf einer fundamentalen Ebene zu verstehen."*

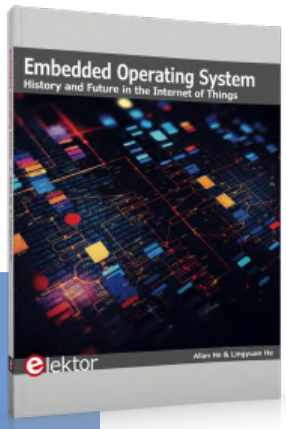
Yu Chen von der Tsinghua University in China leitet ein Team, das an der Implementierung eines RISC-V-Betriebssystems in RUST gearbeitet hat. Die Ziele ihres rCore OS, einer Portierung von uCore in RUST, sind die Verbesserung der OS-Entwicklung und -Qualität unter Verwendung einer modernen Programmiersprache sowie die Erforschung des zukünftigen OS-Designs und der Implementierung [15]. Ich selbst habe am 2. und 3. Online-Kurs zur Entwicklung von Embedded-Systemen und IoT teilgenommen, die gemeinsam von BMR, IAR, SiFive, GigaDevice und Nuclei entwickelt wurden. Diese Kurse zielen darauf ab, Entwickler, Dozenten und Studenten über RISC-V, Embedded-System-Entwicklung und Entwicklungs-Tools zu informieren. Lab-Code-Projekte und zugehörige Materialien zu den Kursen wurden zur Verfügung gestellt [16].

RISC-V steht bei Anwendungen in Embedded-Systemen vor vielen Herausforderungen. Erstens erfordert die Entwicklung von Embedded-Systemen standardisierte Entwicklungsplattformen, von denen es nur sehr wenige gibt. Zweitens ist Open-Source-Software in kommerzieller Qualität wichtig. Zwar gibt es nun die Unterstützung von Linux, aber Android wird wohl in unmittelbarer Zukunft nicht auf RISC-V laufen. Drittens hat RISC-V zwar ein aktives Ökosystem, aber es hat sich noch nicht genug in der Community verbreitet.

Zweifelsohne eignet sich RISC-V besonders für die Forschung und Lehre in den Bereichen Elektronik und IT. Ein brandneues Open-Source-Hardwaremodell begünstigt eine Umgebung, die Innovation und Zusammenarbeit fördert. Zum Schluss ein Zitat von Michael Taylor von der School of Computer Science and Engineering der University of Washington in Seattle:

*„Es gibt keine ernsthaften technischen oder praktischen Probleme mit RISC-V. Es wird schließlich x86 und ARM als primären Befehlssatz für Mikroprozessoren verdrängen. Es wird die Computerwelt grundlegend verändern."*

200659-02



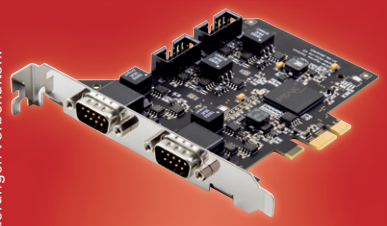

**PASSENDE PRODUKTE**

- **A. He and L. He, Embedded Operating System: History and Future in the Internet of Things, Elektor 2020.**

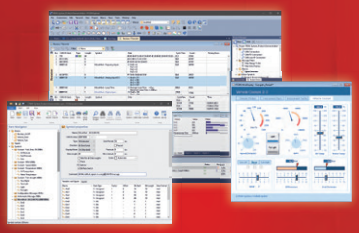
www.elektor.de/embedded-operating-system

You CAN get it...
Hardware und Software
für CAN-Bus-Anwendungen...


**PCAN-Router Pro FD**
Frei programmierbarer 6-Kanal-Router für CAN und **CAN FD** mit I/O und Datenlogger. Auslieferung inkl Entwicklungspaket mit Beispielen.
ab 980 €



**PCAN-PCI Express FD**
**CAN-FD**-Interface für PCI Express-Steckplätze mit Datenübertragungsraten bis 12 Mbit/s. Lieferung inklusive Monitor-Software und APIs.
ab 240 €



**PCAN-Explorer 6**
Professionelle Windows-Software zur Steuerung und Überwachung von **CAN-FD**- und CAN-Bussen.
ab 510 €

Alle Preise verstehen sich zzgl. MwSt., Porto und Verpackung. Irrtümer und technische Änderungen vorbehalten.

www.peak-system.com

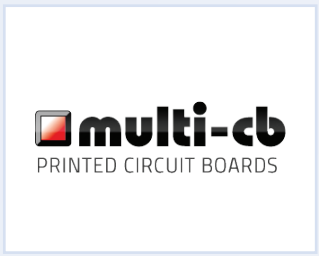
PEAK System
Otto-Röhm-Str. 69
64293 Darmstadt
Germany
Tel.: +49 6151 8173-20
Fax: +49 6151 8173-29
info@peak-system.com

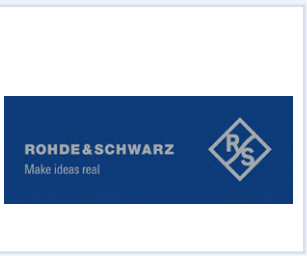
# Wall of Fame

März 2021

Elektor International Media ist stolz darauf, Ihnen die Wall of Fame 2021 präsentieren zu können! Wir arbeiten mit führenden Elektronikunternehmen zusammen - von globalen Zulieferern bis hin zu Geräteherstellern - um über innovative Produkte und Dienstleistungen im High-Tech-Bereich zu berichten.

Die in der diesjährigen Wall of Fame gelisteten Unternehmen bieten Elektroingenieuren, Makern und Studenten ein breites Spektrum an unverzichtbaren Werkzeugen und Lösungen für die Entwicklung von spannenden Elektronikprojekten und -produkten!






**Hier** könnte
Ihr Logo stehen.
Informationen unter
+49 (0)241 955 09 178

# Lösungen für PCAP-Touch-Anwendungen mit hohen EMV-Anforderungen


Ein Beitrag von SCHURTER

Touchpanel-Produkte für den europäischen Markt müssen den europäischen Richtlinien entsprechen und CE-zertifiziert sein. Eine wesentliche Voraussetzung dafür ist, dass diese Produkte die Anforderungen an die elektromagnetische Verträglichkeit (EMV) erfüllen. Doch welche sind dies?

Andere elektrische Geräte in unmittelbarer Nähe eines Touchpanels dürfen den Betrieb dieses Gerätes nicht beeinflussen. Das Touchpanel muss also eine hohe Störfestigkeit aufweisen. Umgekehrt darf das Panel aber auch andere Produkte nicht stören, seine eigene Emission muss also gering sein. Die EMV-Anforderungsklassen sind je nach Anwendungsbereich der Applikation unterschiedlich. Beispielsweise erfordern Medizingeräte aus Sicherheitsgründen eine höhere Störfestigkeit.

## EMV-Spezialist

SCHURTER verfügt über jahrzehntelange Erfahrungen im Bereich EMV. Das zeigt sich darin, dass die EMV-Eigenschaften sowohl der Touchscreens als auch des Endproduktes bereits im Designstadium berücksichtigt werden. Um ein gut durchdachtes Design auf dem Gebiet der EMV zu erreichen, sind die Ingenieure von SCHURTER oft schon in der Anlaufphase in das Projekt eingebunden. Vorzugsweise wird ein gemeinsames Entwicklungsteam zusammengestellt. Dies hat den Vorteil, dass alle spezifischen Wünsche bereits in einem frühen Stadium erfasst und in den Entwurf miteinbezogen werden.

## Cleveres Design

Störaussendung (Emission) und Störfestigkeit (Immission) werden nicht nur durch die elektronischen Schaltungen und deren Komponenten bestimmt, sondern auch durch den mechanischen Aufbau, die Positionierung der Elektronik-Baugruppen, die Wahl der Materialien und die Kabelführung. Da SCHURTER diese Einflüsse bereits in der frühen Designphase berücksichtigt, werden EMV-Probleme und nachträgliche, kostspielige Änderungen nach Fertigstellung des Produkts vermieden.

## Niedrige Emissionen

Viele elektronische Geräte arbeiten heutzutage mit hochintegrierten, hochgetakteten Schaltungen, die häufig die Ursache von ausgesandten Störsignalen sind. Außerdem weisen diese Signale oft eine sehr hohe Flankensteilheit auf, was zu weiteren Oberwellen führt. Durch die Wahl der richtigen Komponenten und Filterelemente werden diese Emissionen an der Quelle wirksam bekämpft.

## Emissionsquelle Controller

Der von SCHURTER eingesetzte Touchscreen-Controller-Chip ist bei einer unsachgemäßen Beschaltung eine potentielle Emissionsquelle. Durch den Entwurf eines optimierten, kundenspezifischen PCB-Layouts in Zusammenarbeit mit dem Kunden und dem Chiphersteller wird die Emission so gering wie möglich gehalten. Berücksichtigt werden eine korrekte Leiterbahnführung, die Filterung der Signalleitungen, die Abgrenzung der Versorgungsleitungen und die Platzierung von Masseflächen in der Leiterplatte. Ein Multilayer-Design bietet eine zusätzliche Abschirmung gegen hochfrequente Störsignale. Da sich der kapazitive Touchscreen an der Außenseite des Endprodukts befindet, kann dieser wie eine Sendeantenne für höhere Frequenzen wirken. Für diese Frequenzen und deren Oberwellen ist es manchmal notwendig, eine interne Abschirmung der Elektronik in Form von Metallfolie oder ein komplettes Metallgehäuse zu integrieren.

## Hohe Störfestigkeit

Ein PCAP-Touchscreen darf niemals eine Aktion auslösen wenn er nicht berührt wird (ein so genannter Ghosttouch). Ein Touchscreen kann durch hochfrequente Energie von außen gestört werden. Dies kann eine leitungsgebundene oder eingestrahlte

Störenergie sein. Um die erforderliche hohe Störfestigkeit zu erreichen, werden zusätzliche Vorkehrungen getroffen. Diese können etwa aus Hardware-Lösungen wie Signal- und Netzleitungsfilterung, der Abschirmung der Leiterplatte sowie der Verdrahtung bestehen. Andererseits verfügt der Touchscreen Controller-Chip auch über eine Reihe von eingebauten Softwarefunktionen. Zu nennen wären hier etwa ein Algorithmus zur Rauschunterdrückung oder das Frequenzsprungverfahren (Frequency Hopping Methode). Unter Berücksichtigung aller Aspekte wählt SCHURTER sorgfältig die bestmögliche Lösung aus.

## Leitungsgeführte Störfestigkeit

Die Entwickler bei SCHURTER verfügen über professionelles Prüfequipment zur Ermittlung der leitungsgeführten Störfestigkeit. Bereits bei den ersten Prototypen wird die Störfestigkeit auf dem Niveau von Pre-Compliance-Tests offizieller Prüflabors verifiziert. Werden bei einem solchen Test Störungen festgestellt, dann wird der Ingenieur „on-the-fly" die richtigen Parameter in der Firmware des Controllers justieren und die optimale Funktionalität feinabstimmen.

## Montage des Controllers

Für optimale EMV-Eigenschaften des Endproduktes, sowohl in Bezug auf die Emission als auch auf die Störfestigkeit, ist es entscheidend, dass der Touchscreen-Controller korrekt montiert ist. Insbesondere die Masseverbindungen, die Verdrahtung und die elektrische Verbindung mit weiteren Komponenten des Gerätes sind in dieser Hinsicht von wesentlicher Bedeutung. Erdungsverbindungen, ob getrennt oder nicht, sind ebenfalls wichtig, da sie die endgültige Leistung eines Touchscreens mitbestimmen. Für Kunden hält SCHURTER eine Montageanleitung für die korrekte Montage des Touchpanels mit dem entsprechenden Controller im Detail bereit.

## Touchscreen-Sensor und kapazitive Tastpunkte

Eine bedeutende Komponente in der Konstruktion ist der Sensor. Der Sensor kann aus Polyester oder ITO-Glas gefertigt werden. Tastpunkte können auch direkt auf einer Leiterplatte integriert sein. Ein optimales Leiterbahnmuster des Sensors ist für eine gute Funktionalität unerlässlich. Leitfähiges ITO-Glas ermöglicht transparente Schaltbereiche zur einfachen Hinterleuchtung. Der Sensor selbst wird hinter das gewählte Coverlens-Material laminiert oder gebondet. Für jeden Sensor ist eine Anschlussfahne aus Polyester oder Kapton erforderlich. Da die Sensoren in einem Ätz- oder Siebdruckverfahren hergestellt werden, sind sie in jeder gewünschten Größe und Form erhältlich. SCHURTER fertigt das Sensordesign selbst, passend zur jeweiligen Anwendung. Zusätzlich zu den Umgebungsinterferenzen kann ein Berührungssensor auch Störsignale von dem darunter liegenden LC-Display aufnehmen. Dies kann über die Elektronik des Controllers oder über eine zusätzliche transparente Abschirmfolie zwischen Display und Sensor gelöst werden.


*Bild 1. EMV-Prüfaufbau für die leitungsgebundene Störfestigkeit von Touchpanels.*


*Bild 2. EMV-Prüfung der leitungsgebundenen Störfestigkeit eines Touchpanels.*

210042-02

### WEBLINK

[1] **SCHURTER Eingabesysteme**: www.schurter.com/de/Eingabesysteme

Touchless Automation

# Der Aufstieg der kontaktlosen Manipulation

Das Bewegen von Objekten ohne sie dabei zu berühren war schon immer einer der beliebtesten Zaubertricks. Leider können diese Tricks in industriellen Umgebungen nicht zur Herstellung tatsächlicher Produkte angewendet werden. Die Fähigkeit Komponenten zu manipulieren, ohne sie zu berühren, wäre besonders nützlich für sehr empfindliche oder zerbrechliche Teile wie Mikrostrukturen, dünne Chips oder Photonik. Die letztgenannte Anwendung ist besonders wichtig, da fast jedes Bauteil, welches sich mit der Wechselwirkung zwischen Licht und Materie befasst, eine nahezu perfekte Oberfläche haben muss.

Die Fertigung dieser speziellen Teile ist umständlich und viele Produzenten melden Fehlerraten im zweistelligen Bereich. Gleichzeitig hat die Implementierung solcher Komponenten in den letzten Jahren in allen Bereichen stetig zugenommen. Beispielsweise hatten im Jahr 2010 nur Smartphones der oberen Preisklasse zwei Kameras. Und die eingebauten Sensoren konnten einzig und allein feststellen, ob sich das Telefon in der Nähe des Ohrs befindet. Jetzt verfügen Top-Geräte über mehr als sechs Kameras und eine ganze Reihe von Sensoren, die eine Vielzahl von Aufgaben wie etwa Gesichtserkennung oder 3D-Bildgebung ausführen können. Aber nicht nur die kommerzielle Elektronik-Branche hat diese Entwicklung durchgemacht. Der Fokus auf optisch ausgerichtete Bauteile nimmt täglich zu. Auch die bevorstehende Welle autonom fahrender Autos wird stark von Photonik-Komponenten abhängig sein.

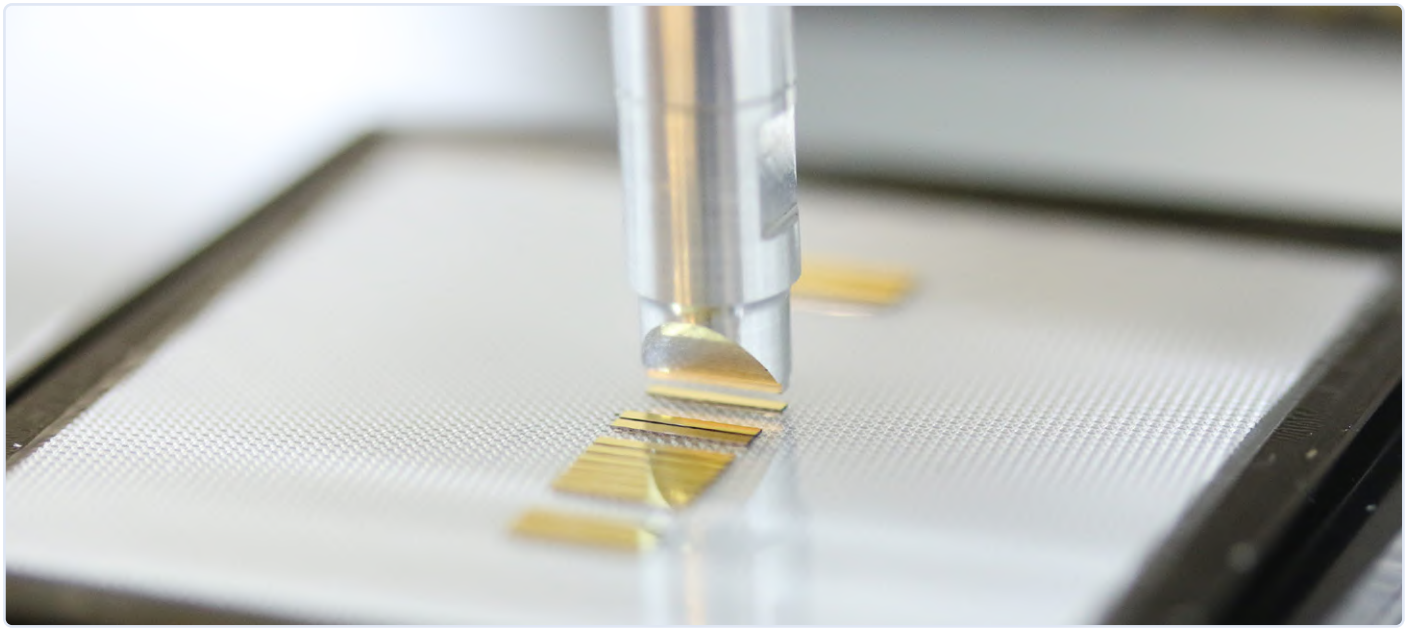
Die Photonik hat somit an Bedeutung gewonnen. Aber im Gegensatz zu Standardhalbleitern liegt das Herz ihrer Forschung, Entwicklung und Produktion immer noch in den westlichen Ländern. In einem am Anfang dieses Jahres von der Europäischen Kommission herausgegebenen Dokument („Eine neue Industriestrategie für Europa", 10.03.2020) wird Photonik eindeutig als eine der Schlüsseltechnologien für die Zukunft eingestuft und verspricht weitere Investitionen in diesem Bereich. Dieser Fortschritt ist auch dank mehrerer Initiativen auf europäischer Ebene, wie dem European Photonics Industry Consortium, EOS und vielen weiteren lokalen Initiativen wie SwissPhotonics oder Photonics France, besonders effektiv.

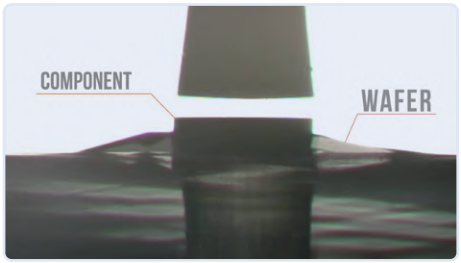
Trotz der großflächigen Verbreitung der optischen Technologie ist die Fertigung immer noch mit zahlreichen Herausforderungen konfrontiert. Das Aufnehmen aus dem Wafer Frame, das Bearbeiten oder Verkleben der Bauteile führt vielfach zu Oberflächenverunreinigungen. Diese müssen danach gereinigt werden oder führen zu Defekten, welche die Funktionalität der Komponenten endgültig beeinträchtigen. Hier kommt die kontaktlose Manipulation ins Spiel.

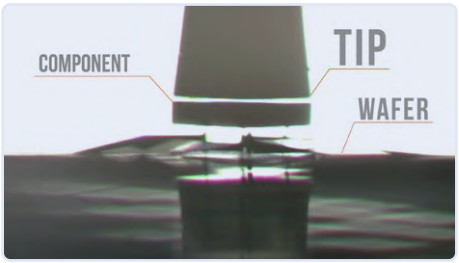
## Kontaktloses Handling ist Realität

*Touchless Automation* konzentriert sich seit seiner Gründung auf die Bereitstellung kontaktloser Manipulationslösungen. Das Unternehmen befindet sich in Biel, einer kleinen Stadt in der Schweiz, die als Zentrum des Uhrentals bekannt ist. Ausgehend von der Notwendigkeit der Uhrmacher, perfekte Oberflächen für ihre High-End Teile zu haben, perfektionierte *Touchless Automation* dieses Handling und brachte eine Reihe von Industriemaschinen auf den Markt, die sich mit Photonik-Komponenten befassen.

Das neueste Produkt „Levio" wurde 2019 auf der SPIE Photonics West in San Francisco vorgestellt und arbeitet bereits in Kundenanwendungen. Diese vollautomatische Maschine kann Komponenten aus einem Wafer kontaktlos aufnehmen, visuell prüfen und nach den Prüfergebnissen in verschiedene Behältnisse sortieren.




Die meisten Kundenanfragen kommen von Herstellern von Laserdioden, Mikrolinsen und optischen Filtern. Die Technologie hat sich jedoch bereits in einer Vielzahl anderer Bereiche als nützlich erwiesen. *Touchless Automation* hat Projekte für MedTech-Komponenten, wie Herzschrittmacherteile oder Standard-Halbleiterkomponenten, erfolgreich abgeschlossen.

Die Flexibilität und Vielfältigkeit der kontaktlosen Anwendung erregte bereits großes Interesse. Unter anderem bei der Europäischen Kommission, die einen Teil der Entwicklung mitfinanzierte und bei der Europäischen Weltraumorganisation, welche *Touchless Automation* in ihr ESA BIC Schweiz-Programm aufnahm.

## Die Technologie

Die Innovation von *Touchless Automation* liegt in seiner Technologie. Das kontaktlose Handling ist eine Kombination zweier entgegengesetzter Kräfte. Ein Vakuum zieht das Bauteil nach oben und ermöglicht das Anheben - gleichzeitig vibriert die Spitze des Greifers mit hoher Frequenz. Diese Vibration erzeugt ein dünnes Luftkissen zwischen Greifer und Bauteil, welches verhindert, das sie in Kontakt kommen.

Diese Manipulationstechnologie ermöglicht einen sehr stabilen Griff, ohne dass es dabei zu einem physischen Kontakt kommt. Auf diese Weise bleiben alle Komponentenoberflächen unversehrt und können fehlerfrei weiterverarbeitet werden.

Unter Berücksichtigung von Aero- und Fluidphänomenen ist es möglich, Komponenten aus jedem Material und mit einer Vielzahl von Formen zu hantieren. Sogar Teile mit konvexen oder konkaven Oberflächen können bewegt werden, vorausgesetzt solche Formen sind bei der Entwicklung des Systems bekannt.

Die Hauptstärken von *Touchless Automation* liegen jedoch nicht nur in der Manipulation von Komponenten. Mit jahrelanger Erfahrung im Maschinenbau und deren Inbetriebnahme konzentriert sich das Unternehmen auf die Entwicklung von Systemen. Ausgehend von der Levio-Plattform kann diese problemlos an viele verschiedene Prozesse angepasst werden, unter anderem auch ein Inspektionssystem. Beispielsweise erfordern Laserdiodenfacetten normalerweise die Erfassung von Submikrondefekten, während für optische Komponenten solche Defekte um ein Vielfaches größer sein können, um toleriert zu werden. Diese Flexibilität ermöglicht Kosteneffizienz. Das Ganze ist eine Lösung nach Maß, die individuell auf den Kunden zugeschnitten ist.

## Eine bessere Zukunft

Die kontaktlose Manipulation wird die Prozesse der Technik verbessern und die Qualität der Komponenten erhöhen. Gleichzeitig werden die Fehlerraten gesenkt und Reinigungsschritte überflüssig gemacht, was die Reduzierung von Industrieabfällen ermöglicht. Dies sind jedoch nur die offensichtlichsten Effekte. Die Möglichkeit Bauteile auf solche Weise zu handhaben eröffnet völlig neue Entwurfsmöglichkeiten für Mikrokomponenten, die bis jetzt noch als unmöglich galten. Zum Beispiel lassen sich Materialien einsetzen, die zu spröde oder weich sind, um konventionell gehalten zu werden. Viele Branchenexperten sehen diese Technologie als revolutionär und möglicherweise als neues Paradigma der zukünftigen Fertigung.

Mit Kunden auf drei Kontinenten ist *Touchless Automation* in kurzer Zeit schnell gewachsen und strebt an, in den kommenden Jahren Branchenführer zu werden. Viele verschiedene Projekte konnten bereits umgesetzt werden. Das Unternehmen sucht jedoch immer wieder nach neuen Herausforderungen.

200210-02

Q&A mit Matthew Stewart

# Was ist tinyML?

Fragen von **C.J. Abate** (Elektor)

Maschinelles Lernen (ML) ist eine Unterkategorie der künstlichen Intelligenz und wird in vielen Bereichen wie Atmosphärenforschung oder Computer Vision angewendet. Laut Doktorand Matthew Stewart eignet sich tinyML zur Entwicklung „schneller, ressourcenschonender und energieeffizienter Algorithmen für maschinelles Lernen, die auf ressourcenbeschränkten Mikrocontrollern eingesetzt werden können".

**Abate:** Zu Ihrem Hintergrund: Wann haben Sie sich das erste Mal für maschinelles Lernen [1] interessiert? Sind Sie eher durch Programmierung oder Hardware-Design auf dieses Thema gekommen?

**Stewart:** Für meinen Bachelor in Maschinenbau habe ich Erfahrungen in Programmierung und Mechatronik sammeln können. Mit maschinellem Lernen wurde ich jedoch erst zu Beginn meines Studiums an der Harvard University konfrontiert. Mein Interesse daran wurde geweckt, als ich im ersten Jahr meiner Doktorandenzeit einen Einführungskurs in Data Science belegte und mir das enorme Potenzial maschinellen Lernens sowohl allgemein als auch speziell für die Atmosphärenforschung bewusst wurde.

**Abate:** Was hat Sie an die Harvard University geführt?

**Stewart:** Harvard ist eine der weltweit führenden Forschungseinrichtungen. Ein Studium an dieser Universität ist das Ziel vieler engagierter und strebsamer Studenten. Auch die Forschungsinteressen meines Betreuers spielten eine Rolle, da er den tropischen Regenwald im Amazonasgebiet mit Hilfe von Drohnen untersuchte. Mein Interesse an Umweltforschung entstand im Laufe meines Maschinenbaustudiums, als mir immer klarer wurde, dass die meisten wichtigen technischen Probleme der Moderne Umweltprobleme sein werden - also Klimawandel, Energieversorgung und Nachhaltigkeit. Die Arbeit mit Drohnen im Amazonas-Regenwald erschien mir aufgrund meiner Interessen und meines ingenieurwissenschaftlichen Hintergrunds ideal. Dies war das zentrale Motiv zum Gang nach Harvard.

**Abate:** Wie halten Sie sich als Umweltwissenschaftler über Embedded-Systeme und Programmierung auf dem neuesten Stand? Bei all den neuen Entwicklungen im KI-Bereich, in der Sensorik, bei eingebetteten Systemen etc. ist es ja schwer, immer auf dem Laufenden zu bleiben. Wie informieren Sie sich?

**Stewart:** Das ist aufgrund des intensiven und schnellen Fortschritts in diesen Bereichen ein echtes Problem für viele Studenten und Akademiker. Ich persönlich nutze mehrere Ressourcen. Z.B. hilft Twitter dabei, neue Forschungsergebnisse zu entdecken, die von anderen Akademikern auf dem Gebiet gepostet werden. Ich nehme auch an mehre-

ren Slack-Kanälen teil, in denen Kollegen regelmäßig Neuigkeiten und Forschungsartikel zu verwandten Themen austauschen. Darüber hinaus schaue ich mir regelmäßig neue Arbeiten an, die in einschlägigen Zeitschriften veröffentlicht wurden. Dort suche ich nach Auffälligem, das mich zum ausführlichen Lesen motiviert. Glücklicherweise sind die meisten veröffentlichten Arbeiten von geringer Relevanz für meine eigene Forschung, und breitere Trends sind oft Gegenstand von Seminarvorträgen, die von verschiedenen Abteilungen und Interessengruppen innerhalb der Universität gehalten werden.

**Abate:** Obwohl ich tinyML in einem Interview mit Daniel Situnayake [2] vor ein paar Monaten angesprochen habe, ist es für viele Ingenieure in der weltweiten Elektor-Community noch ein neues Thema. Was würden Sie sagen? Ist tinyML ein Ansatz für die Implementierung von Machine-Learning auf Mikrocontrollern?

**Stewart:** Ja, das ist der Kern. tinyML ist keine spezifische Technologie. Genau genommen ist es eher eine Proto-Plattform mit Synergien aus den Bereichen Computerarchitektur, Performance Engineering und maschinellem Lernen. tinyML zielt auf die schnelle, ressourcenschonende und energieeffiziente Implementierung von Algorithmen für maschinelles Lernen, die sich für Mikrocontroller mit beschränkten Ressourcen eignen. Dies umfasst auch die Entwicklung von spezieller Hardware für bestimmte Aufgaben, neuen Algorithmen für ressourcenbeschränkte Anwendungen, neuen Tools zur Portierung von Algorithmen oder die Optimierung ihrer Leistung über eine breite Palette von Hardware-Architekturen. Es wird auch davon gesprochen, dass tinyML für das maschinellem Lernen auf Mikrocontrollern mit weniger als 1 MB RAM und einer Leistungsaufnahme von <1 mW gedacht ist, aber dies ist keine strenge oder erschöpfende Definition.

„tinyML ist keine spezifische Technologie oder ein Satz von Prinzipien, es ist eher eine Proto-Plattform mit Synergien aus den Bereichen Computerarchitektur, Performance Engineering und maschinellem Lernen."

**Abate:** Es geht also nicht um Plattformen für den NVIDIA Jetson und Raspberry Pi? Der Fokus liegt also auf viel schwächeren Controllern mit <1 mW und RAM, das eher in Kilobyte als Megabyte berechnet wird?

**Stewart:** Korrekt. Raspberry Pi und NVIDIAs Jetson stehen nicht im Fokus von tinyML - ebenso wenig wie Technologien im Zusammenhang mit selbstfahrenden Autos, die erhebliche Rechenleistung benötigen. Das Schlüsselwort ist „ressourcenbeschränkt". In tinyML muss man Entscheidungen darüber treffen, wie die Leistung eines Algorithmus bezüglich anwendungs- und hardwarespezifischer Beschränkungen optimiert werden kann.
Z.B. kann es in einigen Anwendungen zwingend erforderlich sein, sowohl eine schnelle Inferenz als auch eine hohe Genauigkeit zu erreichen. Um die Inferenzgeschwindigkeit zu verbessern, könnte man 8-bit-Integerzahlen anstelle von Fließkommazahlen verwenden, aber dies würde sich auf die Genauigkeit des Algorithmus auswirken sowie Auswirkungen auf die erforderlichen Speicher- und Rechenressourcen haben. Dies verdeutlicht, warum ich tinyML als Proto-Plattform betrachte. Man muss hier mehr über funktionale Anforderungen nachdenken, die oft in direkter Konkurrenz zueinander stehen, was Kompromisse erfordert.

**Abate:** Können Sie ein paar Beispiele aus der Praxis anführen?

**Stewart:** Es gibt bereits recht verbreitete Beispiele für tinyML in Smartphones. Eines davon ist das „Keyword-Spotting", bei dem es um die Erkennung von Wörtern wie „Hey Siri" oder „Hey Google" geht. Würden Smartphones das Mikrofon kontinuierlich mit Hilfe der normalen CPU überwachen, um diese Wörter zu erkennen, würde der Akku des Handys nur ein paar Stunden durchhalten. Stattdessen erfolgt die kontinuierliche Überwachung durch einen schlanken Signalprozessor. Dieser weckt dann die CPU auf, wenn das Schlüsselwort von einem bekannten Sprecher gesprochen wurde. Die CPU verarbeitet dann weitere Spracheingaben.
Ein weiteres Beispiel sind Smartphones, die erkennen, wann es in die Hand genommen wird. Die Daten der eingebauten Beschleunigungssensoren und des Gyroskops werden kontinuierlich durch einen energiesparenden Controller überwacht. Sobald das Smartphone aufgenommen wird, weckt dieser die CPU auf.
Noch ein Beispiel ist die Detektion der Anwesenheit einer Person durch einen Mikrocontroller via Kamera. Es kann z.B. erkannt werden, ob der Benutzer eine Maske trägt, was besonders während der aktuellen Pandemie nützlich ist. Die Erkennung von Anomalien wird zukünftig in der Industrie eine größere Rolle spielen, und zwar dort, wo Signale von großen Maschinen kontinuierlich überwacht werden, um eine vorausschauende Wartung zu realisieren.

**Abate:** 2019 haben Sie einen faszinierenden Artikel mit dem Titel „The Machine Learning Crisis in Scientific Research" [3] veröffentlicht, der sich mit der Frage beschäftigt, ob maschinelles Lernen zu einer „Reproduzierbarkeitskrise" in der Wissenschaft beiträgt. Wenn zum Beispiel ein Wissenschaftler einen „schlecht verstandenen" ML-Algorithmus in einem Experiment verwendet, kann das bedeuten, dass andere Wissenschaftler die originalen Forschungsergebnisse nicht reproduzieren können. Selbst Nicht-Wissenschaftler können das Problem verstehen. Vermutlich hat sich die Debatte „maschinelles Lernen vs. traditionelle Statistik" im letzten Jahr verschärft. Wie denken Sie darüber?

**Stewart:** Das ist ein wichtiges Thema in der akademischen Welt. Mein Artikel war eine Reaktion auf die Reproduzierbarkeitskrise, die zuerst durch die Kontroverse um einige Arbeiten zum Thema „Power Posen" von Amy Cuddy, einer ehemaligen Professorin der Harvard Business School, aufkam. Andrew Gelman schrieb ein wichtiges Paper, in dem er schlechte Forschungspraktiken im Bereich der Psychologie anprangerte, die unaufrichtige Datenanalysen mit Techniken wie p-hacking, post-hoc Rationalisierung und Rosinenpickerei von Daten beinhalteten, um statistisch signifikante Ergebnisse zu produzieren. Dies führte zu einer Reihe von Experimenten mit dem Ziel, einige wichtige Ergebnisse der psychologischen Literatur zu reproduzieren, von denen

sich viele als nicht reproduzierbar erwiesen. Dies beleuchtet problematische Fehler in der Forschung, da Studien zur Reproduzierbarkeit oft nicht finanziert wurden, weil sie als unnötig und Verschwendung von Ressourcen angesehen wurden. Mittlerweile zeigte sich, dass die Reproduzierbarkeitskrise auch andere Bereiche wie Literatur und Wirtschaft betrifft.
Natürlich führt diese Korrumpierung der Integrität des Forschungsprozesses zu Bedenken hinsichtlich der Verwendung großer Datensätze und maschinellem Lernen. Bei einer ausreichend großen Anzahl von Variablen in einem Datensatz ist es letztlich unvermeidlich, dass einige statistisch signifikante Ergebnisse vorhanden sind. Dies lässt vermuten,

„Jede Branche, die mit vielen IoT-Geräten arbeitet, wird von tinyML aufgrund des geringeren Stromverbrauchs und der geringeren Netzwerklast stark profitieren."

dass Scheinkorrelationen leichter zu finden sein werden, wenn Experimente gerade solche Hypothesen testen und eben nicht eine Vielzahl von Hypothesen gleichzeitig auf korrekte Weise. Big Data macht es einfacher, mit Daten zu „schummeln". Und Maschinelles Lernen macht es darüber hinaus noch einfacher, den Betrug zu „verstecken". Eingeschränkte Interpretierbarkeit, nuanciertes Verhalten vieler Algorithmen des maschinellen Lernens und mangelnde Ausbildung im Bereich des maschinellen Lernens werden es schwieriger machen, diese Probleme in veröffentlichten Forschungsarbeiten aufzudecken. Glücklicherweise ist die Lösung für dieses Problem recht einfach: Die Finanzierung von Reproduzierbarkeitsstudien und die Aufklärung von Forschern über das richtige Design von Experimenten und die korrekte Verwendung von maschinellem Lernen für Forschungszwecke.

**Abate:** Sie haben in Ihrem Artikel einen interessanten Punkt angesprochen: „Eines der anderen Probleme von Algorithmen des maschinellen Lernens ist, dass der Algorithmus eine Vorhersage machen muss. Der Algorithmus kann nicht sagen: ‚Ich habe nichts gefunden'." Das klingt, als gäbe es Fälle, für die maschinelles Lernen nicht taugt.

**Stewart:** Ich würde zwar zustimmen, dass maschinelles Lernen für einige Fälle nicht passt. Dennoch glaube ich nicht, dass dies der Grund dafür ist. Eines der Probleme bei binären Klassifizierungsproblemen ist z.B. die unpassende Zusammenfassung, was dann zu einer falschen Dichotomie führt. Unter bestimmten Umständen kann es sinnvoller sein, die nahe an einer Entscheidungsgrenze liegenden Daten von einem Menschen genauer zu beurteilen, anstatt den Algorithmus eine Entscheidung treffen zu lassen. Diese Art der Entscheidungsfindung wird manchmal als „Human-in-the-Loop"-Entscheidungsfindung bezeichnet und wäre vor allem da sinnvoll, wo die zu treffende Entscheidung wichtige Auswirkungen hat – etwa bei der Kreditvergabe oder bei Krebs-Diagnosen.

**Abate:** In welchen Branchen sehen Sie die größten Chancen für Innovationen durch tinyML?

**Stewart:** Vermutlich erwarten viele in diesem Bereich arbeitende Leute, dass tinyML oder Vergleichbares eine neue industrielle Revolution in Gang setzen wird. Aus diesem Grund bezeichnen einige diese Entwicklung als „Industrie 4.0". Jede Branche, die mit vielen IoT-Geräten arbeitet, wird von tinyML aufgrund des geringeren Stromverbrauchs und der geringeren Netzwerklast stark profitieren.
Genauer gesagt: Es gibt Branchen, die vom richtigen Einsatz der neuen Möglichkeiten von tinyML massiv profitieren werden. Die Landwirtschaft wäre ein solcher Bereich. tinyML könnte in der Landwirtschaft intelligente Sensorfunktionen ermöglichen, ohne dass eine Verbindung zu einem Stromnetz erforderlich ist, tinyML würde hier helfen, die Zeitpunkte für das Düngen, Wässern und die Ernte bestimmter Pflanzen zu bestimmen.
Ein weiteres gutes Beispiel ist die Schwerindustrie, in der vorausschauende Wartung mittels Anomalie-Erkennung zu Kosteneinsparungen und Effizienzsteigerungen führen könnte. Die Vorbeugung von Problemen bei großen Maschinen ist wahrscheinlich preiswerter und führt zu geringerem Produktivitätsverlust als die Folgen eines Ausfalls.

**Abate:** Was ist mit Unternehmen, die an der Entwicklung energieeffizienter Computerlösungen interessiert sind?

**Stewart:** Apple und ARM sind wahrscheinlich die größten Unternehmen, die sich im Moment auf energieeffizientes Computing konzentrieren. Die Entwicklung leistungsstarker und energieeffizienter Architekturen war bei Smartphones entscheidend für die Verbesserung der Akkulaufzeit bei gleichzeitiger Erhöhung der Funktionen und der Leistung. In den letzten Jahren haben sich die mobilen Architekturen bezüglich Leistung und Energieeffizienz erheblich verbessert, während die traditionellere Architekturen wie etwa von Intel eher stagniert haben. Folglich konkurrieren mobile Architekturen jetzt mit traditionelleren Architekturen, haben aber einige zusätzliche Vorteile wie die bessere Energieeffizienz. Dies hatte erste größere Auswirkungen, als Apple seinen neuen ARM-basierten M1-Chip vorstellte und damit warb, dass er die „längste Akkulaufzeit aller Zeiten in einem Mac" bieten würde. Dieser Schritt von Apple weg von Intel wird als ein Wendepunkt in der Computerindustrie angesehen, der in den kommenden Jahren auch Auswirkungen auf die gesamte Industrie haben wird.

**Abate:** Welche Rolle spielt tinyML bei Ihrer Forschung mit Drohnen und Systemen zur Überwachung von Chemikalien?

**Stewart:** Es wurden bereits Arbeiten veröffentlicht, tinyML für verschiedene Mikrodrohnen-Anwendungen nutzen. Dabei geht es darum, leichtgewichtige Drohnen zu entwickeln, die mit Hilfe von Reinforcement-Learning-Methoden intelligent in einer Umgebung navigieren. Dies könnte zukünftig z.B. bei der Erkennung von Gaslecks oder der Lokalisierung von Schadstoffemissionen, sowohl im Innenbereich als auch im Freien sehr nützlich werden.
Für chemische Überwachungssysteme im weiteren Sinne könnte tinyML helfen, weit entfernte, vom Stromnetz unabhängige Sensornetzwerke zu schaffen. Außerdem ist eine intelligentere Nutzung der Sensorinformationen machbar. Anstatt kontinuierlich Daten an einen Cloud-Server zu übertragen, könnte sich das System z.B. nur auf anomale Daten konzentrieren. Dies würde die Belastung des Kommunikationsnetzes und auch den mit kontinuierlicher Überwachung verbundenen Stromverbrauch reduzieren. Diese Aspekte werden in den kommenden Jahren immer wichtiger werden, da die Anzahl der eingesetzten IoT-Geräte exponentiell ansteigt.

### Mehr über Maschinelles Lernen

Möchten Sie mehr über die Arbeiten von Matthew Stewarts oder maschinelles Lernen im Allgemeinen erfahren? Nachfolgend finden Sie weitere Informationen:

- Matthew Stewart: https://mpstewart.net/
- Towards Data Science: „Matthew Stewart - Tiny ML and the future of on-device AI", 11 2020: http://bit.ly/video-tinyML
- W. Trojan, „KI für Anfänger", Elektor 9-10/2020: www.elektormagazine.de/articles/ki-fr-einsteiger-1
- C. Valens, „KI und das AIY-Vision-Kit", ElektorMagazine.de: www.elektormagazine.de/news/ki-und-das-aiy-vision-kit
- C. Valens, „Intelligenter Assistent im Selbstbau mit dem Voice Kit V2 von AIY", ElektorMagazine.de: www.elektormagazine.de/news/intelligenter-assistent-im-selbstbau-mit-dem-voice-kit-v2-von-aiy

**Abate:** Ihre Artikel und Forschungen werden wahrscheinlich viele Mitglieder unserer Community dazu inspirieren, sich näher mit tinyML zu befassen. Haben Sie neben dem Buch über tinyML von Pete Warden und Daniel Situnayake irgendwelche Empfehlungen für Ingenieure und sonstige interessierte Elektroniker, die mehr über das Thema erfahren wollen?

**Stewart:** Leider ist eine der Kehrseiten von neuen Technologie, dass es oft nur wenig Informationsquellen gibt. Allerdings gibt es immer mehr Peer-Review- Literatur zum Thema tinyML (wenn auch oft unter anderem Namen). Ein beträchtlicher Teil dieser Literatur wird auf dem Preprint-Server arXiv in der Kategorie Hardware-Architektur [4] veröffentlicht. Ich vermute, dass es bald mehrere Zeitschriften geben wird, die sich auf das Thema konzentrieren. Eine weitere Ressource ist der von der tinyML Foundation veranstaltete tinyML Summit (www.tinyml.org/event/summit-2021), der im März 2021 digital stattfindet, und auf dem wahrscheinlich einige neue und spannende Entwicklungen vorgestellt werden.

210014-02

## WEBLINKS


[1] **Maschinelles Lernen (Elektor)**: www.elektormagazine.de/search?query=maschinelles+Lernen

[2] **C. Abate, „Die Zukunft des maschinellen Lernens", Elektor 11-12/2020**: www.elektormagazine.de/magazine/elektor-158/59072

[3] **M. Stewart, „The Machine Learning Crisis in Scientific Research", TowardsDataScience.com, 11 2019**: http://bit.ly/ml-crisis-stewart

[4] **Cornell University, „Hardware Architecture", arXiv.org**: https://arxiv.org/list/cs.AR/recent



Ganz einfach Preis und Verfügbarkeit aller Bauteile prüfen, die Sie benötigen

Price Availability Assistant

STOCK • PRICE • BUY


MOUSER ELECTRONICS

mouser.de/price-availability-assistant

elektorstore

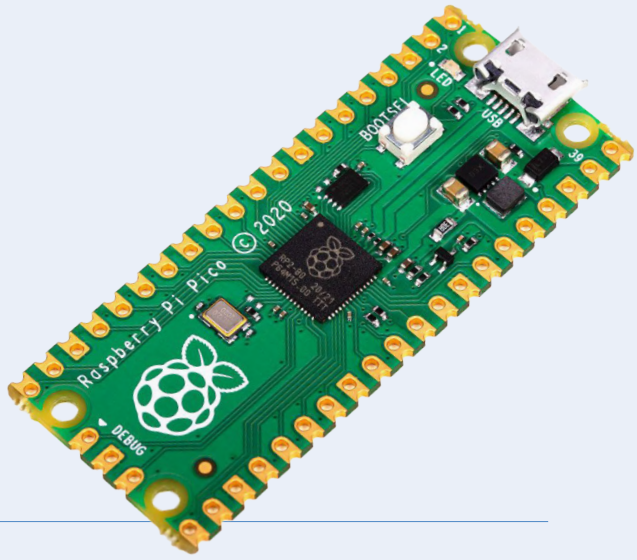
www.elektor.de

# Der Elektor Store

## Nie teuer, immer überraschend!

Der Elektor Store hat sich vom Community-Store für Elektor-eigene Produkte wie Bücher, Zeitschriften, Bausätze und Module zu einem umfassenden Webshop entwickelt, der einen großen Wert auf überraschende Elektronik legt. Wir bieten die Produkte an, von denen wir selbst begeistert sind oder die wir einfach ausprobieren wollen. Wenn Sie einen Produktvorschlag haben, sind wir hier erreichbar (sale@elektor.com). Unsere Bedingungen:
**Nie teuer, immer überraschend!**

### Raspberry Pi Pico Mikrocontroller-Board




**Preis: 4,95 €**

**www.elektor.de/19562**

### JOY-iT JT-RD6006 Labornetzteil (360 W)


Preis: 89,95 €
**Mitgliederpreis: 80,96 €**

**www.elektor.de/19564**

# elektor e-zine

Your dose of electronics


Jede Woche, in der Sie den Elektor e-zine Newsletter nicht abonnieren, ist eine Woche mit großartigen Artikeln und Projekten zum Thema Elektronik, die Sie verpassen!

Also, worauf warten Sie noch? Melden Sie sich heute für unseren Elektor e-zine Newsletter unter www.elektor.de/ezine an und erhalten Sie zusätzlich ein kostenloses Raspberry Pi Projektbuch!

## Was können Sie erwarten?

**Redaktioneller Elektor-Newsletter**
Jeden Freitag erhalten Sie die wichtigsten Artikel und Projekte der Woche. Wir zeigen MCU-basierte Projekte, IoT, Programmierung, KI und mehr!

**Elektor-Newsletter mit exklusiven Angeboten**
Verpassen Sie nicht unsere Shop-Angebote, jeden Dienstag und Donnerstag haben wir eine besondere Aktion für Sie.

**Mailing von externen Partnern**
Sie wollen über die laufenden Aktivitäten in der Branche informiert bleiben? Dann gibt Ihnen diese E-Mail die besten Einblicke. Unregelmäßig, aber immer mittwochs.

elektor
design > share > sell

# Jede Bewertung spiegelt ein persönliches Erlebnis wider


Wir lieben Elektronik und Projekte, und wir setzen alles daran, die Bedürfnisse unserer Kunden zu erfüllen

Der Elektor-Store: '**Never expensive, always surprising**'

Elektor Deutschland is rated **Excellent**

Based on 30 reviews

Trustpilot

Sehen Sie sich weitere Bewertungen auf unserer Trustpilot-Seite an: www.elektor.de/TP

Oder bilden Sie sich selbst eine Meinung und besuchen Sie unseren Elektor Store, www.elektor.de

elektor
design > share > sell