Why data-over-sound is an integral part of any IoT engineer's toolbox: Chirp + Arm = frictionless low power connectivity

Kobus Marneweck, Senior Product Manager – Arm
James Nesfield, CEO - Chirp
Dr. Adib Mehrabi, Head of Research – Chirp
Dr. Daniel Jones, Chief Technology Officer – Chirp

White Paper

Contents

Data-over-sound presents a compelling solution for many device-to-device connectivity applications, particularly for use cases that require frictionless, low cost connectivity with nearby devices. This white paper introduces the fundamental concepts and benefits of data-over-sound connectivity, and explores key application areas within the Internet of Things, including provisioning smart devices and facilitating secure near-field communication in low-cost, low-power scenarios. It outlines how Chirp has capitalized on the digital signal processing (DSP) capabilities of the Arm Cortex-M4 and Cortex-M7 processors to engineer a software-defined data-over-sound solution that is robust and reliable without being resource intensive, enabling connectivity and application logic to reside on a single Arm core.

Contact

Arm www.arm.com/company/contact-us
Chirp info@chirp.io

1. Introduction

The rise of IoT devices in the home and workplace has created a world where data and connectivity are becoming increasingly complex. Global sales of smart home devices are set to reach almost 1 billion units by 2022[1]. In terms of revenue, we can expect to see over £10 billion spent on smart home devices in the UK alone in 2019[2]. These include smart speakers, appliances, lighting, locks, and controllers, to name but a few. As more and more devices are introduced to the home, we will see an increasing need for seamless interoperability and to device-device connectivity solutions. Early adoption of these devices in both domestic and industrial settings has demonstrated that there is no single device to rule them all. Instead, we are seeing smart functionalities and connectivity being added the plethora of devices already used for entertainment, utilities, and production.

"Global sales of smart home devices are set to reach almost 1 billion units by 2022."

Outside the home, the number of connected devices is set to grow significantly in industrial applications such as manufacturing, energy, and agriculture. The manufacturing sector is set to take the lead in terms of market size, with a predicted market spend of $430m by 2021[3]. The Industrial IoT (IIoT) presents an even more complex connectivity landscape, with tens, if not hundreds of devices, controllers, and sensors sitting side by side in the same factory or workplace. Looking at both the smart home and IIoT sectors combined, current forecasts predict a staggering 75 billion connected devices by 2025, up from 26 billion in 2019, as shown in Figure 1.

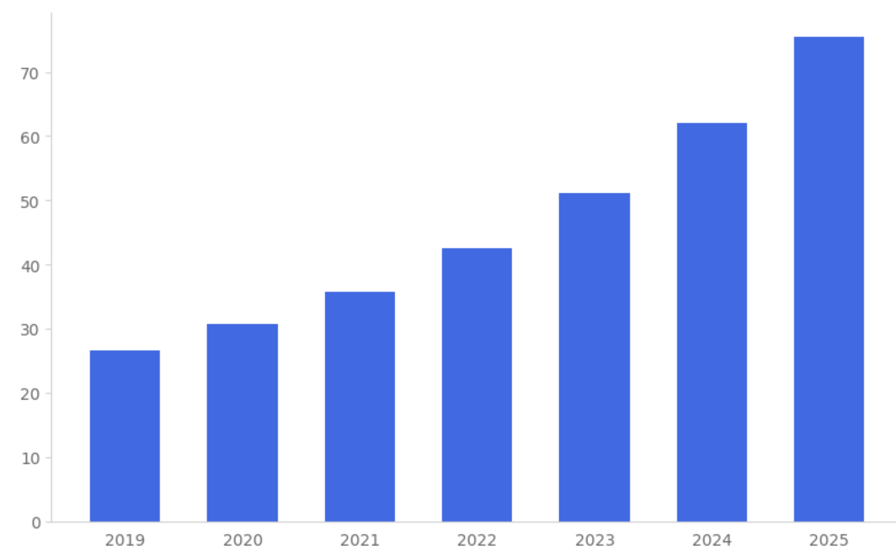


Figure 1: Internet of Things (IoT) connected devices installed base worldwide from 2019 to 2025 (in billions). Source: IHS[4].

As IoT technology advances and the demand for efficient ways of communicating data between these devices grows, the world has witnessed a rise in emerging new data transmission technologies which are looking to provide secure and effective solutions for sharing information. One solution rising to meet these new demands is "data-over-sound", which harnesses devices' existing speakers and microphones to send and receive data over an acoustic channel. Because it doesn't require any additional networking hardware, data-over-sound has captured the interest of companies interested in adding wireless connectivity functionality to existing devices. Moreover, the status of data-over-sound as a pairing-free, one-to-many medium means it addresses some of the pain points of longstanding alternatives such as Bluetooth and Wi-Fi, presenting an attractive general-purpose solution for frictionless data transmission.

One of the main advantages of data-over-sound is that the physical infrastructure needed to facilitate sonic data transfer is already largely in place. The voice is gaining momentum as the primary control mechanism for many IoT devices, and as such, microphones are increasingly being incorporated into more and more IoT devices. Beyond mobile devices and home assistants such as Alexa and Google Assistant, we are seeing voice control being added to smart TVs, fridges, door bells, vacuum cleaners, light bulbs, locks, and thermostats. As humans continue to communicate with IoT devices using sound, we are seeing millions of devices of all form factors already equipped with the required processor, speaker, and microphone for data-over-sound functionality - without requiring any physical upgrades to existing hardware or additions to a BOM. With companies always looking to innovate and future-proof their services, many are now realizing the potential of data-over-sound to provide seamless device-to-device connectivity either to nearby devices or remotely (e.g. down a phone line), using nothing but sound.

"...the world has witnessed a rise in emerging new data transmission technologies which are looking to provide secure and effective solutions for sharing information. One solution rising to meet these new demands is 'data-over-sound'."

This shift is not without its challenges. One of the major threats to the mass adoption of IoT technology is the age-old conflict between performance and price. We are entering an age where people expect even the cheapest smart devices to be able to process real-time audio and perform intelligent digital signal processing (DSP) on the edge device. The Cortex-M series of processors addresses this need by providing an extremely cost-effective solution that is capable of high-performance real-time audio DSP.

This white paper explores the growing role of seamless, low-power and low-cost device-to-device communication in the emerging IoT landscape. We will look at the solutions offered by data-over-sound technology for a number of IoT use cases, explore how data-over-sound can be realised in embedded scenarios by the optimized DSP of the Arm Cortex-M4 and Cortex-M7 processors, and outline how this technology can be implemented by third-party engineers in their IoT products.

2. The missing piece of the connectivity puzzle

Table 1: Feature comparison for Chirp vs. alternative connectivity technologies.

There already exists a plethora of connectivity solutions, including extremely short range (NFC and QR); short range, high bandwidth (Bluetooth and Wi-Fi); long range, low power (Sigfox and LoRa); and short range, low power (Zigbee). Each technology has certain

	Chirp	QR	NFC	Blue-tooth	Wi-Fi	Li-Fi	ZigBee 802.15.4	LoRa	Sigfox
Two-way communication	•			•	•	•	•	•	•
One-to-many broadcast	•					•	•		
Non-line-of-sight transmission	•			•	•		•	•	•
Works in RF-restricted environments	•	•	•			•			
Zero setup / pairing / configuration	•	•	•						
Available to application by default	•	•							
Low power operation	•	•	•	•			•	•	•
Can transmit with sub-$2 electronics	•	•	•	•			•		
Can receive with sub-$2 electronics	•			•			•		
Wireless broadcasts confined to room boundaries	•	•	•						
Transmit over ranges > 10m	•	•		•	•	•	•	•	•
Can limit the transmission range to < 1m	•	•	•	•					
Supported by dumb media channels and P.A. systems	•	•							
Supported by typical mobile devices	•	•	•	•	•				
Supported by low-end mobile devices	•	•		•	•				
Typical usable max data rate	1 kb/s	3kb	424 kb/s	25 mb/s	70 mb/s	1 gb/s	250 kb/s	50 kb/s	100 b/s
Typical max range	100m	10:1	20cm	100m	50m	10m	100m	10km	40km

advantages that make it more or less suitable for specific applications. As any technology provider in the IoT arena will appreciate, the application-specific needs are not always as simple as choosing a technology based on cost, data rate, and range. Things can get a lot more complicated when one considers the finer details of device support, backwards compatibility, user experience, and security. Further detail is introduced by the technical specifications of the protocol, such as the nature of the transfer medium (RF, optical, acoustic), number of required channels, security concerns, or whether two way and one-to-many communication is needed.

"...data-over-sound has certain advantages that make it particularly attractive for a number of use cases that are suited to the Cortex-M series of processors..."

An overview of these considerations is given in Table 1. This is by no means exhaustive, and each use-case will bring its own specific requirements. Nonetheless, data-over-sound has certain advantages that make it particularly attractive for a number of use cases that are suited to the Cortex-M series of processors, including:

- **Device interoperability:** The simple hardware requirement of a speaker and/or microphone make data-over-sound arguably the most wide-reaching wireless communication technology in terms of device compatibility. Mobile phones, voice-controlled devices, and any device with an alarm speaker are able to communicate using data-over-sound. This includes many legacy devices. Data can also be transmitted over media channels such as radio and TV, and over existing telephone lines.

- **Frictionless UX:** Data-over-sound requires no pairing or configuration, making data transfer as simple as pressing a button.

- **Physically bounded:** Because sound waves respect room boundaries, particularly in the near-ultrasonic range commonly used in data-over-sound, transmissions do not pass between neighbouring rooms. This means that it can be used for detecting the presence of a device with room-level granularity.

- **Low or no bill-of-materials (BOM):** Data-over-sound adds nothing to a device BOM if the specification already includes audio I/O, which is becoming increasingly commonplace. If audio I/O is only required for data-over-sound functionality, the BOM is less than $2 for sender and receiver components.

- **Zero power:** The advent of 'wake-on-sound' MEMs microphones such as the Vesper VM1010 enables devices to communicate using data-over-sound, whilst draining virtually no battery power in between communications (< 10 µA).

Table 2: Key use cases of Chirp

USE CASE	CONNECTIVITY REQUIREMENTS
IoT device provisioning	✣ Seamless UX and rapid setup ✣ No pairing/config ✣ One-to-many ✣ >10m range ✣ Non-line of sight ✣ No additional hardware requirements on router ✣ No additional infrastructure requirements
Proximity detection	✣ Respects room boundaries ✣ One-to-many ✣ Universal device support
Two-way acoustic NFC	✣ Two-way ✣ Low-cost, low-power ✣ Universal device support ✣ Supports industry-standard cryptography
Telemetry in RF-restricted environments	✣ No use of EM spectrum ✣ Up to 100m range ✣ Non-line of sight ✣ Two-way communication ✣ Low data rate

"By providing Chirp SDKs for the Arm Cortex-M4 and Cortex-M7 processors, we have substantially lowered the barrier to incorporating data-over-sound into these IoT applications."

Based on these affordances, four key use cases for Chirp are emerging, which are identified in the following section and Table 2 (above). Data-over-sound is the only data-communication solution that meets all of the connectivity requirements for each use case. In addition, the specific hardware requirements of affordability, price, and power, can be comfortably provided by the Cortex-M series of processors. By providing Chirp SDKs for the Arm Cortex-M4 and Cortex-M7 processors, we have substantially lowered the barrier to incorporating data-over-sound into these IoT applications.

Application Areas

Some key application areas have emerged as the drivers of data-over-sound in IoT contexts, primarily in scenarios in which it significantly reduces friction or human effort.

Provisioning

Configuring a new smart device with network credentials and functional configuration remains a disproportionately complex task, particularly for headless devices. Although technologies such as WPS have emerged to address this challenge, these typically require support on the wireless access point or wider infrastructure, or, in some cases, physical access to the AP itself.

Chirp offers an approach to provisioning which is offline, locally-bounded, and does not require any infrastructure modifications. Credentials are encoded as audio, optionally layered with cryptography for secure scenarios, and broadcast over-the-air to nearby smart devices.

The bill of materials needed to receive and decode credentials can be as minimal as an Arm Cortex-M processor plus a digital MEMS microphone.

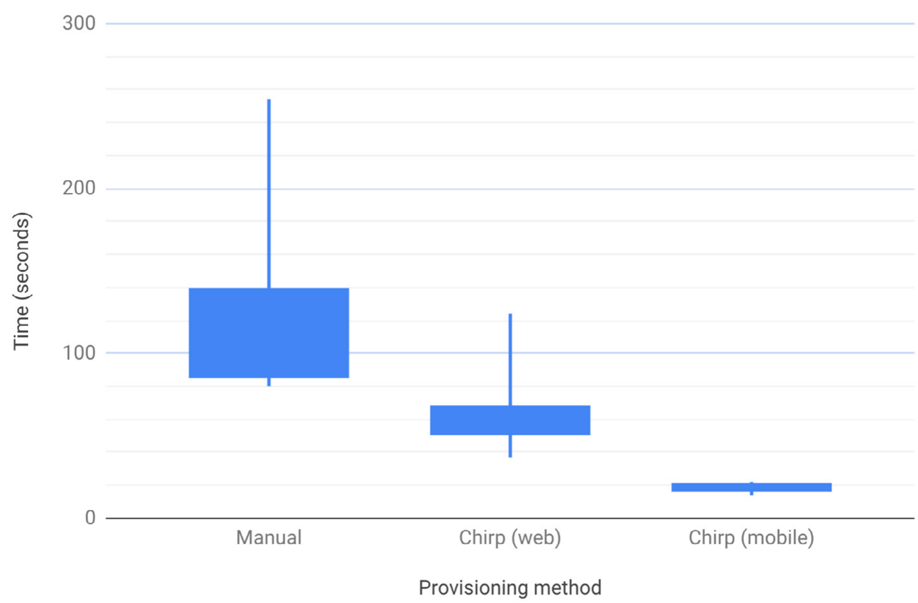


Figure 2: Comparison of provisioning methods in terms of the time taken to provision a headless smart device.

"The Chirp-enabled solutions reduced the provisioning time by 48% with the web-based scenario, and 86% with the mobile scenario."

Figure 2 shows the results of a study in which users with some experience of IoT were asked to provision a headless smart device using three methods:
1 - Manually, by creating a temporary Wi-Fi hotspot on the device, connecting to the hotspot with a web browser, and entering credentials;
2 - Using a Chirp-enabled web interface which asks the user to enter credentials and transmits them to the device via data-over-sound;
3 - Using a Chirp-enabled mobile app which was able to use the device's existing knowledge of the credentials.

The Chirp-enabled solutions reduced the provisioning time by 48% with the web-based scenario, and 86% with the mobile scenario. This could be the difference between a happy and frustrated user of a smart home device, and has the potential to offer significant time and cost savings in industrial scenarios, where many devices may need to be regularly provisioned.

Proximity and Presence Detection

A key property of acoustic signals is that their propagation respects room boundaries, particularly in the near-ultrasonic range. As a consequence, there is a growing uptake for near-ultrasonic acoustic beacons for presence sensing at room-level granularity. Devices can send out sonar-like beacon packets at regular intervals, enabling them to be discovered by nearby peers. Chirp's implementation uses multi-channel encoding to support multiple devices within hearing range, with carrier sensing to detect free channels.

Two-way acoustic NFC

Data-over-sound also supports near-field communication scenarios whilst offering a two-way full-duplex mode of communication, thus addressing a critical limitation of NFC. Enabling two-way data exchange means that devices can perform challenge-and-response dialogues - for example, Diffie-Hellman key exchange for secure financial transactions, or securely sending un-spoofable receipts to a merchant or buyer. The low power requirements and low cost of audio components, combined with fact that data-over-sound is software-defined thus requires no additional hardware, means that acoustic NFC is a powerful emerging approach to peer-to-peer payments and minimal-cost POS devices.

"...data-over-sound is no more complex than a traditional telephone modem. Data is encoded into an acoustic signal, which is then played through a medium ."

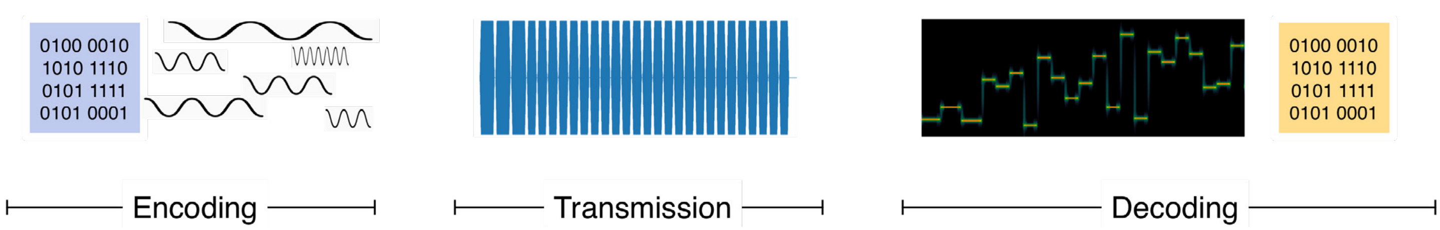
Telemetry in RF-restricted environments

In many sensitive environments, RF-based communications are prohibited due to the risk of sparks or interference with equipment that pre-dates RF regulation. Chirp overcomes these limitations, allowing the industrial IoT to harness the benefits of wireless communication without limitation.

3. Data-over-sound 101

The basic idea of data-over-sound is no more complex than a traditional telephone modem. Data is encoded into an acoustic signal, which is then played through a medium (typically the air, although it could equally be a wired telephone line or VoIP stream), and received and demodulated by a 'listening' device. The listening device then decodes the acoustic signal and returns the original data. This process is illustrated in Figure 3.

Figure 3: The lifecycle of a Chirp.



As with other data transmission protocols, the entire message can include specific sequences to indicate the start and expected length of a message, as well as additional bytes for error detection and correction, as shown in Figure 4.

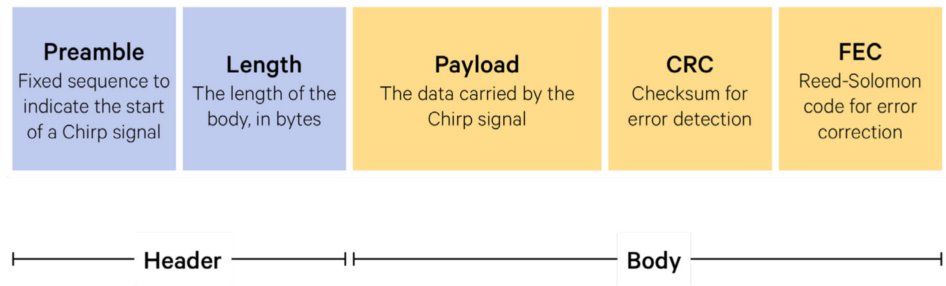


Figure 4: The makeup of a typical data-over-sound message.

Messages are typically encoded using frequency-shift keying (FSK) modulation, which is robust to background noise and distortion from acoustic effects such as reverberation. Any part of the sonic spectrum can be used, although unless data-over-sound-specific hardware is being designed for the application, the bandwidth is limited by the transducers and analogue-digital, digital-analogue converters found on many target devices. Following extensive testing of a large range of consumer devices, Chirp considers the 'safe' range to be between 1kHz-20kHz. This enables the production of audible messages, or near-ultrasonic when limited to ~17kHz-20kHz, which is imperceptible to most adult humans.

Data rate, range, and reliability

As with any wireless communication protocol, there is a natural trade-off between rate, range and reliability in data-over-sound signalling. An increase in range will result in a decrease in reliability or data rate (see Figure 5), and an increase in reliability at a constant range will invariably result in a lower data rate. In theory, data-over-sound can be used with data rates considerably higher than 1kbps. However, in practice the typical data rates are in the order of hundreds of bits per second.

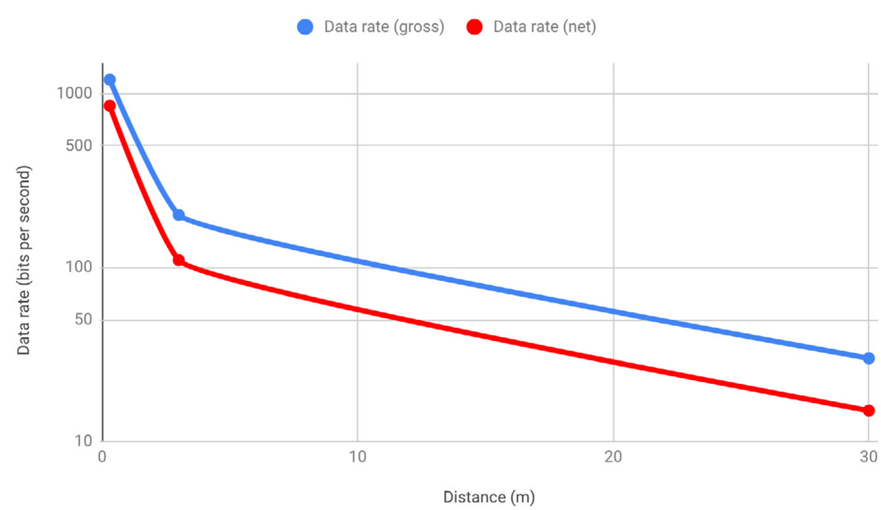


Figure 5: The effect of distance on data rate.

"All of Chirp's standard profiles are tested to perform at >99% reliability under normal operating conditions."

In terms of range, one can expect reliable transmission from 1cm to 100m, depending on the data rate requirements. All of Chirp's standard profiles are tested to perform at >99% reliability under normal operating conditions. Of the existing data-over-sound solutions on the market, the main differentiator for Chirp is this reliability and robustness of the internal decoding technology. This has been demonstrated and stress-tested in acoustically extreme environments, including nuclear power stations with upwards of 100dB(A) of background noise, across ranges exceeding 100m.

Encoding a Chirp signal is a lightweight process in terms of CPU, only requiring the generation of error correction symbols and sinusoidal oscillator synthesis. The bulk of Chirp's processing requirements are found in the decoder. This is where high-performance DSP power is required, and where the DSP features of the Cortex-M series processors become critically important.

4. Cortex-M for DSP applications

"The Cortex-M processor series is designed to enable developers to create cost-sensitive and power-constrained solutions for a broad range of devices."

The Cortex-M processor series is designed to enable developers to create cost-sensitive and power-constrained solutions for a broad range of devices. The addition of DSP extensions to the Thumb instruction set and the optional floating-point unit (FPU), are designed to improve the performance of numerical algorithms, addressing the need for high-performance generic code processing, as well as digital signal processing applications. The signal processing operations are performed directly on the Cortex-M4, Cortex-M7, Cortex-M33 and Cortex-M35P processors, while maintaining the ease of use of the Cortex-M programmer's model.

The Cortex-M7 processor, used in the Chirp solution, is the highest-performance member of the Cortex-M family. It is highly energy efficient and designed for mixed-signal devices. Its DSP capability and flexible system interfaces makes it suitable for a wide variety of applications - from industrial IoT solutions to smart monitors.

The Cortex-M4 features include table branch instructions and conditional execution (using IT instruction), hardware divide instructions, multiply-accumulate (MAC), and various bit field operations. In addition, it supports DSP-specific instructions such as SIMD, saturation arithmetic instructions, a wide range of MAC instructions which can execute in single cycles, and an optional floating unit that support single precision floating point operations.

The instruction set support in the Cortex-M7 processor is similar to Cortex-M4, with additional floating-point instruction, optional double precision floating instructions, and support preload data (PLD) instruction for caching data in advance.

5. The Chirp Arm SDK: Chirp on your Cortex-M processor

The suite of Chirp C SDKs are compact but fully featured, allowing developers to use Chirp's technology even on low-power, memory-constrained devices. The Chirp Arm SDK is optimized for the Cortex-M processors, with support for both the Cortex-M4 and Cortex-M7. It exists as a static library (approximately 740KiB in size) with a straightforward API for sending and receiving data via Chirp signals.

The SDK implements the physical and data link layers of the data-over-sound OSI stack. Chirp is software-defined, meaning that any additional layers can be easily implemented on top of the SDK if features such as encryption, application-specific protocols, or end-to-end transmission/QoS are required.

Chirp can be incorporated into the software stack for existing projects by integrating the library's object (.a) and header (.h) files. API calls are available for the following functions:

- **Initialisation and termination**, to allocate and free the main structures of the SDK
- **Audio processing interaction**, in which arrays of audio samples are exchanged between the audio I/O drivers and the Chirp SDK
- **Payload handling**, including creation and transmission of data
- **Parameter getters and setters** for properties such as software volume, SDK states, and internal parameters such as audio sample rate
- **Configuration of user profile and audio transmission properties**, for application-specific communication protocols
- **Debugging helpers**, for converting payloads and error codes into human-readable strings

The protocol definitions are handled via a cloud service and associated with specific users/clients. Protocols are often updated and in some cases users will require bespoke protocols for specific applications. This provides a simple way to update the protocol without having to change any of the code implementations.

The protocol definitions are handled via a cloud service and associated with specific users/clients. Protocols are often updated and in some cases users will require bespoke protocols for specific applications. This provides a simple way to update the protocol without having to change any of the code implementations.

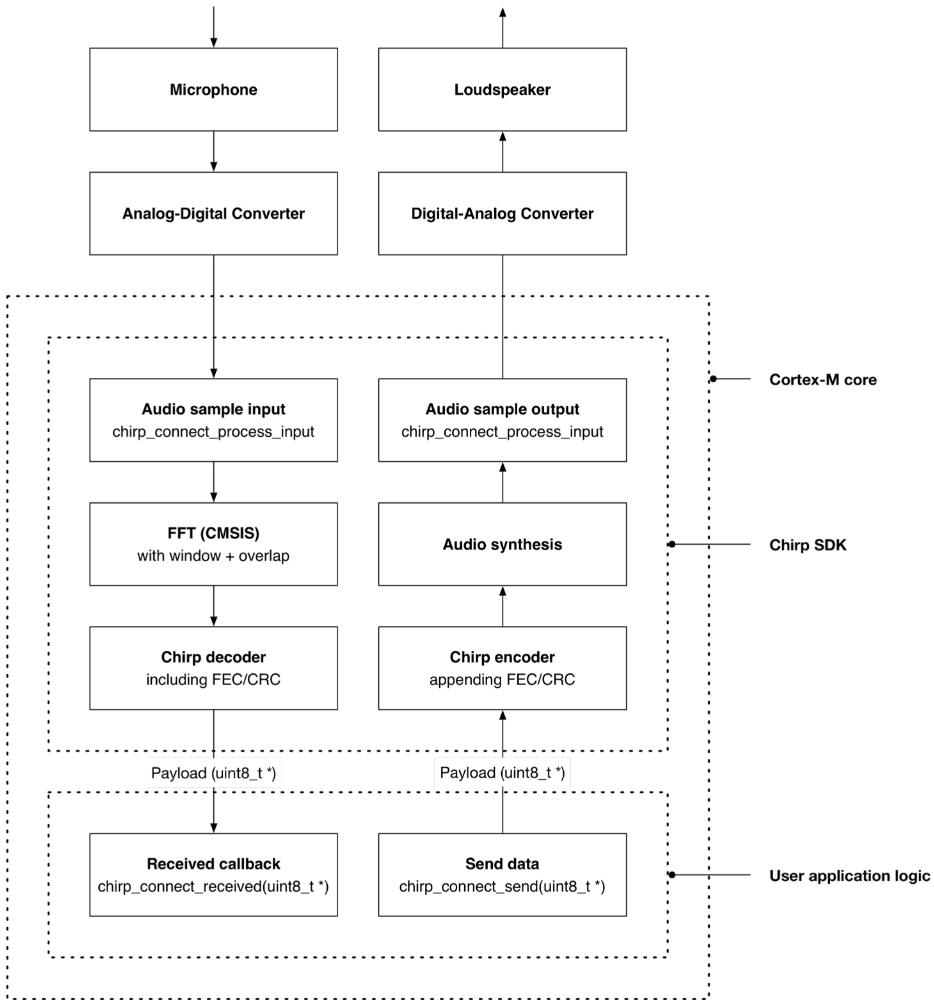


Figure 6: System architecture incorporating the Chirp SDK.

Optimizing Chirp for Cortex-M

"Using the CMSIS open-source DSP library, the Chirp Arm SDK maximises the performance of the Cortex-M processors by using fully optimized instruction sets."

Decoding messages requires the processing of relatively large amounts of audio data in a short space of time. Using the CMSIS open-source DSP library, the Chirp Arm SDK maximises the performance of the Cortex-M processors by using fully optimized instruction sets. In particular, the SIMD instructions allow the Chirp decoder to work on large audio buffers in real time without incurring CPU overruns.

In order to monitor and manage the memory usage, we have developed the Chirp Memory Manager, an open-source library for tracking memory usage in C. This provides a single entry point for allocating and freeing dynamic memory. Using the memory manager, it is possible to track the dynamic allocations and query the current allocated memory at any time in the life cycle of the program. If there is sufficient headroom in the memory, additional statistics can be calculated and stored during the lifecycle of the program, logging information regarding every allocation and free (relative time, file, function, type allocated, amount of the allocation).

Using the CMSIS library and Chirp Memory Manager, the memory and CPU requirements of the Chirp Arm SDK have been reduced to levels that not only perform within the hardware limitations, but leave ample headroom for additional processes and sophisticated application logic to run alongside Chirp.

Table 3: System requirements for the Chirp SDK on Arm Cortex-M processors.

	Cortex-M4 (STM32F437VIT7 @ 168 MHz)		**Cortex-M7 (STM32F746NGH6U @ 200 MHz)**	
	Audible	Ultrasonic	Audible	Ultrasonic
Audio sample rate	16kHz	48kHz	16kHz	48kHz
Heap usage	125.9KiB	126.8KiB	127.5KiB	126.8KiB
Stack usage	5.54KiB	5.54KiB	5.46KiB	5.46KiB
%CPU usage	15.8%	37.7%	10.7%	34.6%
Data rate	Up to ~200bps (dependent on range and reliability requirements)			
Typical application areas	Thermostats, locks, sensors		Smart speakers, hubs	

6. Summary

In this paper we introduced data-over-sound technology for low-power devices. Data-over-sound has unique affordances that provide the potential for it to become a key device-to-device connectivity solution, particularly in the smart home and industrial IoT sectors. It provides a means to set up and configure devices with minimal effort and low BOM hardware requirements. The rise of microphones in IoT devices means that the hardware requirements are already met for billions of existing devices. Combined with the advent of zero-power microphones, data-over-sound provides an extremely low power, 'always connected', device-to-device data transfer solution.

We have developed the Chirp data-over-sound SDK for the Arm Cortex-M4 and Cortex-M7 based MCUs. The high-performance DSP features on these chips allow for high sample rate audio processing without the need for hardware acceleration, enabling the encoding and decoding of ultrasonic chirps. This places the technology in the hands of device manufacturers and developers who wish to take advantage of a frictionless, low-cost solution for device provisioning, zero-power monitoring, or RF-free telemetry in general.

Get the Chirp data-over-sound SDK
Explore Arm DSP solutions for Cortex-M

7. References

[1] https://www.idc.com/getdoc.jsp?containerId=prUS43701518

[2] https://www.pwc.co.uk/industries/power-utilities/insights/energy2020/connected-home.html

[3] http://files.shareholder.com/downloads/PMTC/0x0x907546/309A7969-7F29-4110-9763-012ED05CAF0C/IoT_Platform_Market_Report_2015-2021.pdf

[4] Lucero, S. (2016). IoT platforms: enabling the Internet of Things (white paper). IHS.