DOI: 10.18727/0722-6691/5240

Upgrade Strategies for the ALMA Digital System

Benjamin Quertier[1]
Stéphane Gauffre[1]
Antsa Randriamantena[1]
Marina Studniarek[1]
Carlos De Breuck[2]
Tony Mroczkowski[2]
Gie Han Tan[2]
Ciska Kemper[2]
Neil Phillips[2]

[1] LAB, University of Bordeaux, France
[2] ESO

The Atacama Large Millimeter/submillimeter Array (ALMA) comprises 66 antennas working as a powerful interferometer. High-speed digitisation, signal transmission over several tens of kilometres from the receivers to the correlator, and complex data processing all require state-of-the-art technologies. The ALMA2030 Development Roadmap calls for an increase in the bandwidth by at least a factor of two, implying a major upgrade of the entire signal path. We present here the results of a detailed study looking at how to upgrade the ALMA digital system, including digitisation, data pre-processing, and data transmission to cope with bandwidths more than four times the current ones. At the same time, this system will contribute to increasing the nominal correlation efficiency from 88% to 99%, and prepare ALMA for longer baselines of up to 100 kilometres.

Introduction

ALMA is by far the most powerful (sub-)millimetre interferometer ever built, and has produced transformational science results throughout its first decade of operation (see, for example, Kemper, 2020). Bringing the signals from up to 66 antennas together in the correlators was one of the major technological challenges of ALMA. It involved laying more than 1000 kilometres of fibres between the 212 antenna pads and the correlator located in the Array Operations Site (AOS) technical building at an elevation of 5100 metres. During the construction phase, the University of Bordeaux (UB) was responsible for delivering two subsystems: the digitiser and the Tunable Filter Bank (TFB).

The digitiser module was built on two application-specific integrated circuits (ASICs), a 3-bit analogue-to-digital converter (ADC) and a demultiplexer, designed by UB with a bipolar complementary metal-oxide-semiconductor (BiCMOS) technology from ST-Microelectronics (see Baudry et al., 2006). This module performs the signal digitisation in the 2–4 GHz intermediate frequency (IF) bands[a] of two polarisations resulting from the first and second analogue frequency conversions, with the digitisers sampling in the second Nyquist zone. The receiver usually produces an IF of 8 GHz, so each antenna is equipped with four digitisation modules.

The TFB, part of the correlator, is a digital electronic system based on field-programmable gate array (FPGA) technology, clocked at 125 MHz, that performs massively parallel data processing. Each TFB card can process one of the 2-GHz basebands generated by the digitiser module to extract up to 32 subbands of 62.5 or 31.25 MHz, on which the correlation is performed. The TFB implements a spectral selection and zoom effect, and increases the total number of spectral channels produced by the correlator by a factor of 30 (7680 per baseband compared to 256 without the use of the TFBs).

Whilst ALMA inspired major progress in (sub-)millimetre technology during its construction phase, other (sub-)millimetre telescopes have also benefitted from these advances and have since overtaken ALMA, especially in terms of the instantaneous bandwidth that can be observed. To keep ALMA at the forefront of science and technology, the ALMA2030 Development Roadmap (Carpenter et al., 2019) both defined new key science drivers and prioritised increasing the overall instantaneous bandwidth of ALMA by a factor of at least two (but preferably four). This entails increasing the IF bandwidth of the receivers and the associated electronics and correlator. With these improvements, ALMA will be 2–4 times faster for spectral line surveys such as those required to study the chemical complexity in star-forming regions, but also when it is used as a redshift machine (see, for example, Reuter et al., 2020). All continuum projects will at the same time benefit from the improved sensitivity and imaging fidelity that comes with such a dramatic increase in bandwidth.

Such a major upgrade is of course more complex for an interferometer with as many as 66 antennas. Whilst most of the development efforts have thus far concentrated on completing the original receiver complement, this effort will come to an end in a few years now that the final receiver (Band 2) is on track to be built by an international consortium led by ESO (Yagoubov et al., 2020). This allows ALMA to now take on the ALMA2030 bandwidth upgrade, involving receivers, digitisers and the correlator. During the last decade, the Laboratoire d'Astrophysique de Bordeaux (LAB) has completed a set of

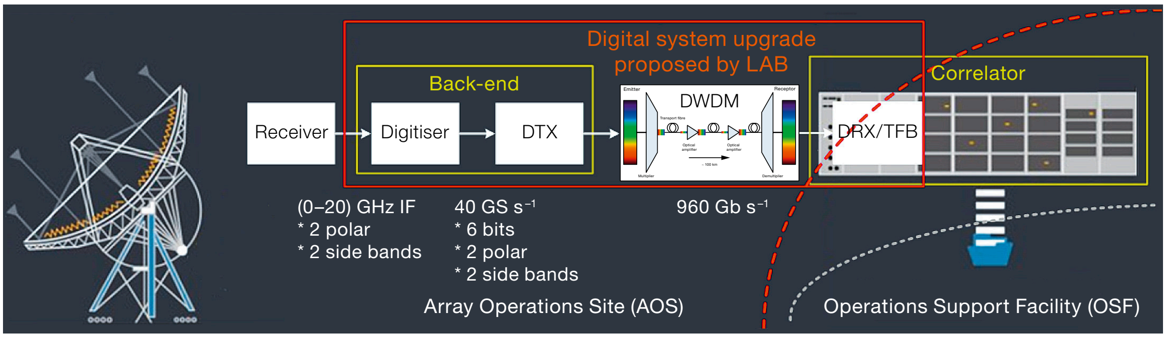


Figure 1. Upgrade of the digital system as proposed by the LAB, University of Bordeaux (both AOS and OSF locations are considered for the new correlator).

development studies supported by ESO to identify appropriate solutions for the digitiser, digital signal processing and data transport that enable meeting these ambitious new goals. In addition to the bandwidth extension, the more modern ADCs also allow one to increase the nominal correlation efficiency from the current 88% to almost 99%, which in terms of sensitivity is equivalent to adding 6–8 new antennas to the array, even for single-line observations. The advanced digitiser solutions studied by the LAB match well the aims of the future ALMA system architecture using only a single, very wide-band IF stage instead of the dual heterodyne conversion approach currently in use.

This single frequency conversion architecture has several major advantages for the user:

- There is no loss of instantaneous bandwidth arising from the use of multiple, staggered, non-ideal IF pass bands. In the current system 4 adjacent bands of 2 GHz each are used; however, the effective total bandwidth is only 7.5 GHz.
- The lack of discontinuities in the bandpass curves of adjacent filters, since the whole IF range will be covered by one passband, makes calibration for this anomaly redundant.
- There is no longer a need for oscillators to convert the astronomical signal from the IF_1 range to the IF_2 range. In practice these strong oscillator signals produce unwanted spurious, ghost signals in the astronomical observations.

The main objective of the first LAB study in 2013 was to evaluate technologies which could be used within a ten-year timeframe to upgrade the sub-systems originally delivered by the LAB. The second study performed by the LAB was focused on the identification and the evaluation of the critical devices required to upgrade digitisation, data transmission, and digital signal processing, in accordance with the general ALMA2030 roadmap, which had been progressively defined in the meanwhile. This second study also included system architecture considerations and comparisons. The most recent LAB study, which is ending in June 2021, aims to confirm these critical choices regarding the ALMA2030 specifications which are now available. All three studies focus on upgrading the ALMA digital system, from digitisation to correlation (see Figure 1).

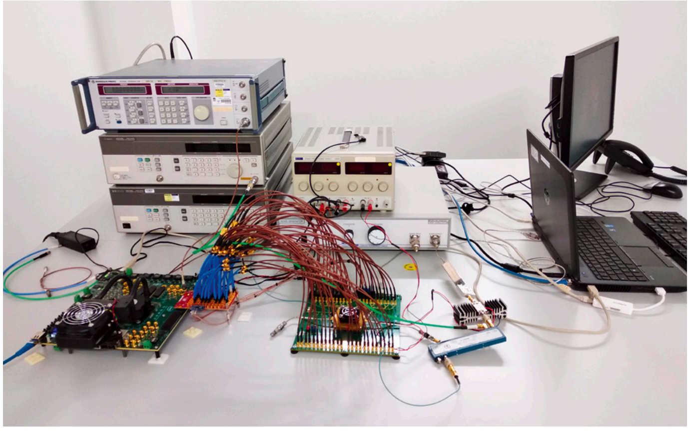


Figure 2. Test bench of the Micram digitiser module in the technical facility at the LAB, University of Bordeaux. The digitiser performance is measured at 40 GS s^{-1} using the Intel Stratix 10 GX FPGA development kit.

Digitisation

Digitisation is one of the main technological challenges in respect of increasing the ALMA instantaneous bandwidth because the market for ADCs with analogue bandwidths and sampling frequencies above 10 GHz is extremely small and unsteady. For the past ten years the electronics group at the LAB has been investigating homemade fully custom ASIC solutions, based on CMOS or BiCMOS technologies, continuing the development work it undertook for the construction phase, and has in parallel carried out an ongoing survey of the solutions announced or under development by small microelectronics startups and global electronics companies. Several of the most promising solutions have been evaluated in the laboratory to measure their actual performance and assess the complexity of their implementation and potential use for ALMA. Various digitisation topologies (for example, multi-rate, interleaved) have also been considered to achieve digitisation of a bandwidth larger than half the sampling frequency, the limit given by the Shannon sampling theorem (see, for example, Tan & Jiang, 2019).

Our primary objective has always been to design a digital back-end including the digitisation, data transmission, and possibly some digital signal processing functions, which would be able to digitise the full IF bandwidth delivered by the new generation of receivers and format the digital samples for transmission over optical fibre. Removing most of the analogue parts, currently used for the second frequency conversion prior to digitisation, we would enhance back-end versatility, reproducibility and reliability, and ease calibration and failure analysis.

Today the baseline solution for the ALMA2030 digitisation is an ADC module from Micram Microelectronic GmbH. This commercial device was initially specified for 34 GS s^{-1}, but has been extensively evaluated by the LAB group up to 40 GS s^{-1}, and subsequently also by Micram. The Micram digitiser chip is built on two internally interleaved ADC cores. Complex calibration is required to optimise the linearity of each core and minimise the mismatch between cores. With this solution we can achieve direct digitisation of, for example, 2–19.5 GHz IF at 40 GS s^{-1}, ten times better than the current digitiser, and we also significantly increase the quantisation efficiency since the Micram device is a 6-bit ADC (as against 3 bits for the current digitiser).

Digitiser requirements

The requirements that drive the selection of the digitiser are the sampling frequency, the bandwidth and the quantisation efficiency. Indeed, the maximal sampling frequency (combined with the digitisation topology adopted) and the digitiser bandwidth together define the effective IF band which can be digitised, and thus the instrument's instantaneous bandwidth. Moreover, the quantisation efficiency is a direct contributor to the overall instrument sensitivity. If the question of the bandwidth and sampling frequency is relatively straightforward, the question

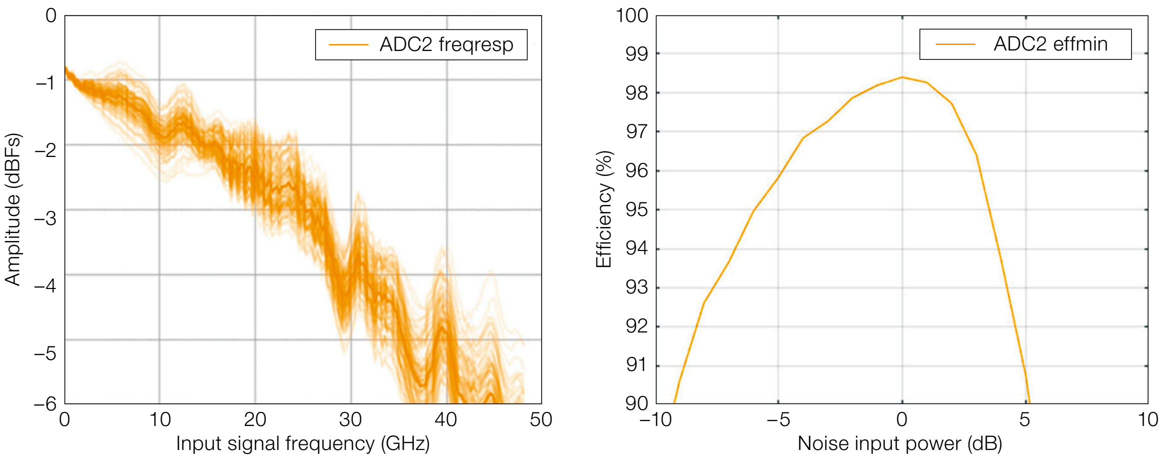


Figure 3. The Micram digitiser's bandwidth (measured using a sine wave input; note that only frequencies below 20 GHz would be used when sampling the first Nyquist zone at 40 GS s^{-1}) and quantisation efficiency (measured using a Gaussian noise input that includes a passband gain variation of 5.4 dB over 16 GHz).

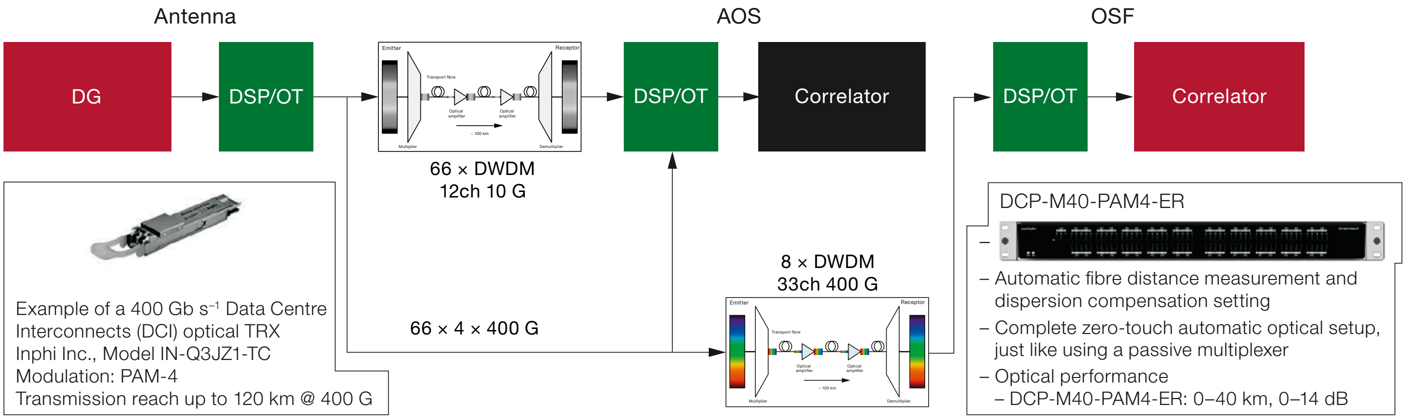
of the quantisation efficiency, and how to define and measure it, is much more complex.

Quantisation efficiency is the relative loss in signal-to-noise ratio (SNR) resulting from the quantisation process; it depends on the number of quantisation levels and the statistical property of the input signal. Thompson, Emerson & Schwab (2007) provide a method for calculating the quantisation efficiency for any number of uniformly spaced levels as a function of the level spacing, using mathematical formulas with radio-astronomy-like signal (Gaussian for a perfect ADC). The deviation from the ideal quantisation step width, which necessarily exists with a real ADC because of manufacturing imperfections, is called differential nonlinearity. Appropriate calibration is needed to minimise this nonlinearity, but some degradation of the quantisation efficiency is unavoidable. In our latest study we describe a novel approach which consists of estimating the effective quantisation efficiency of a real ADC from SNR measurements instead of over-constraining the digitiser specification using the effective number of bits (ENOB) as is often suggested. Indeed, it would be particularly unfortunate to be mistaken in the digitiser evaluation and selection process, considering that there are very few candidates that are capable of achieving the ALMA2030 objectives.

Performance of the Micram solution

We ran four test campaigns with the Micram device over the past 3 years. The first two were held at the Micram facility and allowed us to measure the general performance, the stability over time and the temperature of the digitiser module operating at 34 GS s^{-1}, on noise signals. Two additional test campaigns were then undertaken in the LAB technical facility (see Figure 2), based on our own digital signal processing and calibration procedures, when we were able to demonstrate the performance up to 40 GS s^{-1}. This was later confirmed by Micram on a sample of 100 devices. It was also demonstrated that the bandwidth at 40 GS s^{-1} is beyond 20 GHz (see Figure 3), that the digital sample capture and synchronisation with an FPGA was suitable for long-term applications, that the mismatch calibration process of the two cores appears efficient at rejecting the ghost image arising from interleaving below the quantisation noise floor, and that the threshold response estimated from Gaussian noise histograms makes it possible to optimise the ADC linearity for radio astronomy. Finally, we have been able to estimate the quantisation efficiency from SNR measurements at around 98% for standard astronomical observations (passband gain variations < 5.4 dB over 16 GHz and signal level changes ≤ 4 dB).

Figure 4. Possible data transmission scheme from antennas to correlator (both AOS and OSF locations are considered).



Example of a 400 Gb s^{-1} Data Centre Interconnects (DCI) optical TRX
Inphi Inc., Model IN-Q3JZ1-TC
Modulation: PAM-4
Transmission reach up to 120 km @ 400 G

DCP-M40-PAM4-ER

- Automatic fibre distance measurement and dispersion compensation setting
- Complete zero-touch automatic optical setup, just like using a passive multiplexer
- Optical performance
  - DCP-M40-PAM4-ER: 0–40 km, 0–14 dB

Digital data pre-processing and data transmission

The antennas are connected to the AOS technical building by a network of optical fibre cables. Each antenna is connected by eight single-mode optical fibres, allocated as follows: one fibre for the Data Transmission System, one fibre for the Photonic Local Oscillator and low-frequency timing references, two fibres for the monitoring and control communication, and four spare fibres. The data flow generated by each antenna is transmitted over a single optical fibre using wavelength division multiplexing, a technology which multiplexes a number of optical carrier signals onto a single optical fibre by using different wavelengths (i.e., colours) of laser light. The current system is based on 20-channel dense wavelength division multiplexing (DWDM) with10 G per channel, where only 12 channels are populated. The current throughput of 120 Gb s^{-1} per antenna could then be increased to 200 Gb s^{-1}, but this would not in any case be sufficient to support the IF bandwidth doubling that is a minimum objective for ALMA 2030.

The location of the new-generation correlator will almost certainly be at the Operation Support Facility (OSF) and no longer at the AOS building. This preference has been expressed by several key stakeholders over the past years, for example at the ALMA 2030 workshops organised by the US National Radio Astronomy Observatory (NRAO) and the National Astronomical Observatory of Japan (NAOJ) in 2020. For the upgrade design there is a near consensus within the ALMA developer community that the FFX architecture (where the first F indicates a frequency division of the signal, the second F stands for a Fourier transform stage and the X represents the cross-correlation stage of the signal processing) would be preferred, considering the possibilities offered by the wide multipliers in digital signal processing (DSP) hardware such as FPGA and the finest spectral resolution target of 1 kHz. Baudry et al. (2017) give several examples of high-spectral-resolution architecture where a frequency division of the input baseband is followed by a pure FX correlator. This pre-processing, often just called "the 1st F" in the DSP architecture, has been extensively investigated by the LAB group over the past 5 years, in order to propose an upgrade to our initial contribution with the TFB. Note that the current baseline correlator architecture is a digital hybrid XF design (or FXF). However, when frequency division of the input baseband is bypassed, then the correlator behaves as a pure XF system; both operating modes (frequency-division mode and time-division mode) are offered to users. This brings greater flexibility and makes it possible to have an independent zooming factor within different spectral windows (a tradeoff between effective bandwidth and spectral resolution).

Data transport from antenna to OSF via AOS

One of the ALMA2030 priorities gets implemented by having extended baselines. The cable length to such remote new antenna pads may be significantly longer than the baseline length to the AOS, so lengths of 50 or even 60 kilometres cannot be excluded at this point. The 400G ZR data centre interconnects (DCI) technology allows for transmission reach up to 120 kilometres. The Inphi Corporation, a leader in high-speed data movement interconnects, announced in March 2021 the commercial availability and production ramp-up of COLORZ II 400ZR, the industry's first quad small-form pluggable double-density coherent transceivers for cloud DCIs. It would be the perfect solution for the antenna-to-AOS links. The distance between the AOS and the OSF is less than 40 kilometres. 48 single-mode optical fibres are available for this connection. Installing additional fibres might not be possible so we have identified a fully commercial solution for a complete 40-channel multiplex/demultiplex DWDM module compatible with 1/10/40/100/200/400G Ethernet. Eight of these modules would increase the ALMA optical transport capacity to 2 Tb s^{-1} per antenna, while using only eight single-mode fibres. This solution, combined with 400G ZR transceivers, allows for antenna-to-AOS distances of up to 80 kilometres.

Generic board DSP/optical transmission for the "1st F" and data transmission

For ALMA2030, state-of-the-art FPGAs are required at both ends of the optical fibre, because at each stage we need many high-speed FPGA transceivers to capture, transmit, and receive the extended data flow. Since this kind of FPGA will necessarily come with a very large number of computational resources, we can consider various partitioning configurations of the overall signal processing between the antenna, the AOS and the OSF buildings in order to minimise the cost, ease the deployment, and increase the versatility of the system. For example, the pre-processing and data transmission functions of the current system could be merged in a single FPGA. Moreover, having a first DSP stage like a frequency division at the antenna would be the most straightforward strategy to transmit the overall data stream over the optical system because of considerations concerning the detailed electrical interface of the transceivers and impacts on frame synchronisation. Figure 5 gives an overview of one of the possible architectures we have identified to perform the frequency division prior to the correlation. Here, oversampled polyphase filter banks are used to provide the ability to switch between time-division mode and frequency-division mode while keeping the same usable bandwidth and throughput. One important conclusion is that the transceiver requirements and DSP partitioning are compatible with the idea of a generic FPGA board (DSP/OT, as seen in Figure 4) that could be used at both ends of the optical data transmission system, to implement various functions required along the data path. The only differences would be the firmware to be implemented in the generic FPGA board. This would be a tremendous advantage in terms of validation, manufacturing, fault analysis and maintenance. Using these FPGAs to perform the 1st F, at either end of the DTS link, rather than having additional FPGAs in the correlator for this task, could considerably improve the cost and power efficiency of ALMA2030.

Outlook: prototyping and on-site demonstration

Our current plan for this upgrade is described by the global block diagram in Figure 6, with on-site demonstration in potentially three steps to ensure compatibility with the existing infrastructure (physical locations, power supplies, reference

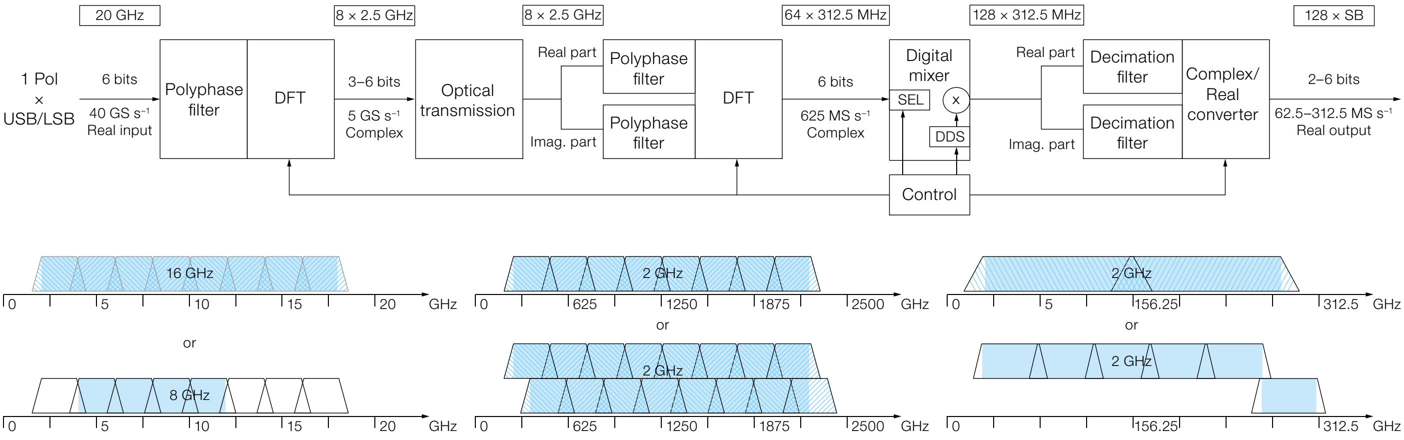


Figure 5. (Above) Example of DSP architecture.

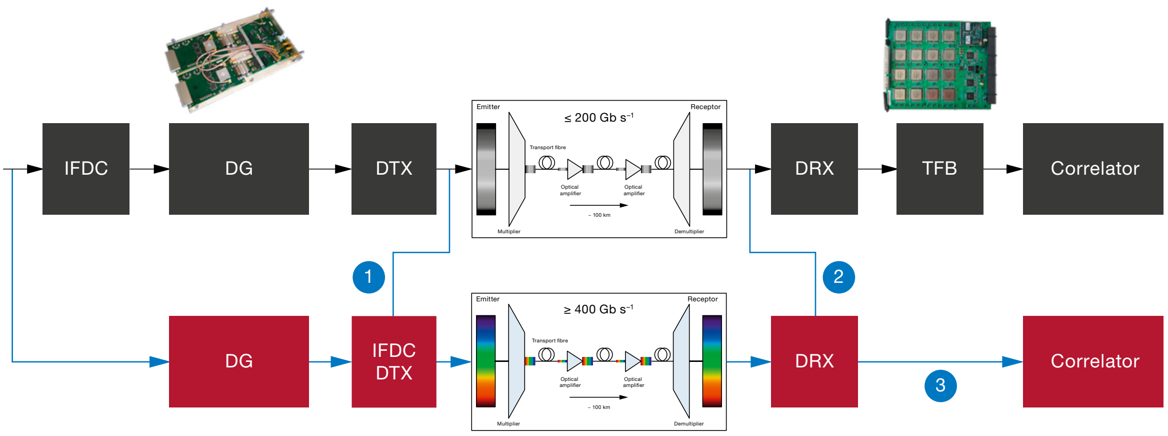


Figure 6. (Left) In black: the current system showing the digitiszer module and the Tunable Filter Bank, two subsystems of the ALMA instrument designed and delivered by the LAB, University of Bordeaux.; In colour: the digital system upgrade plans proposed by the LAB, Université de Bordeaux, with on-site demonstration in three potential steps before a full production and integration.

Step	Quantisation	Efficiency	Bandwidth	Spectral coverage
#1	3 bits × 2 bits	0.84	2 × 4 GHz/pol	128 × 62.5 MHz/pol
#12	6 bits × 2 bits	0.88	2× 16 GHz/pol	128 × 62.5 MHz/pol
#13	6 bits × 6 bits	0.99	2× 16 GHz/pol	256 × 125 MHz/pol

signals, control and monitoring interfaces and optical fibres) before a full production and integration. The recommended strategy is to deploy the new ALMA 2030 digital system in parallel with existing hardware, with all new components installed at different physical locations, so that the current system can continue science operations until the new system is commissioned. The bulk of commissioning could then be accomplished during periods of unfavourable or less favourable observing conditions (i.e., in daytime), while science operations continue with the current system during optimal conditions.

Acknowledgements

We would like to acknowledge the Micram and Intel companies for their technical support and various hardware loan agreements which made this work possible.

References

Kemper, C. 2020, The Messenger, 180, 42
Baudry, A. et al. 2006, The Messenger, 125, 37
Baudry, A. et al. 2017, ALMA Memo, 607
Carpenter, J. et al. 2019, The ALMA Development Roadmap, arXiv:1902.02856
Reuter, C. et al. 2020, ApJ, 902, 78
Yagoubov, P. et al. 2020, A&A, 634, A46
Tan, L. & Jiang, J. 2019, *Digital Signal Processing*, (Cambridge, Massachusetts: Academic Press), 13 et seq.
Thompson, A. R., Emerson, D. T. & Schwab, F. R. 2007, Radio Science, 42

Notes

[a] The term “IF band” is used to distinguish the on-sky, radio frequency (RF) band (in the 35–950 GHz range for ALMA) from that accessible to the digitisation electronics (typically < 20 GHz).