FBAR—From Technology Development to Production--

Rich Ruby

Director of Technology

Agilent Technologies Inc.

Abstract — IN 2001, FBAR technology—in the form of PCS duplexer—showed up in cell phones. By the year 2002, Agilent began high volume FBAR manufacturing. The value proposition for FBAR duplexers was compelling. Up to 2001, PCS duplexers were made out of ceramic and were quite large. SAW technology was perceived as unable to meet the very stringent power requirement, not to mention the exacting electrical specifications. Agilent FBAR was not created overnight. Prior to 2001, Agilent invested eight years of R&D on FBAR. Major R&D investment was done on further improvements to the FBAR device and development of the microcap. Along the way, Agilent closed one FBAR fabrication facility and transferred the process to a 6" CMOS facility. In 2004, Agilent has received its first orders for microcap’d FBAR filters and duplexers. This paper attempts to highlight some of the issues going to high volume manufacturing.

I. INTRODUCTION

FBAR devices first appeared in the literature in the early to mid 1980’s [1,2]. The first reference concerning FBAR’s true value proposition was discussed in Lakin’s paper published in 1982[3]. However, few heeded the vision described in Reference 3.

To be fair, there were a few, rather severe, obstacles FBAR would need to overcome. Zinc Oxide (ZnO), the material of choice for many researchers in the early days, was difficult to deposit and was, by nature, a semiconductor. Aluminum Nitride (AlN), another choice for the piezoelectric material, was even more unknown and thought to be particularly difficult to deposit.

Even if AlN or ZnO films could be made reliably, it was thought unlikely that physical sputtering would give a consistent film that had a reproducible coupling coefficient, $k_t^2$. But, if either ZnO or AlN could be put down in some reproducible manner such that $k_t^2$ was consistent, it was generally believed that the Q of the material and the electrodes would degrade the electrical performance (at that time, Au or Al were the choices of electrodes). It would seem unlikely that a sputtered piezoelectric film could have the same Q’s as the crystal grown Lithium Tantalate or Lithium Niobate (substrates for SAW filters and resonators). Even if the films could reliably be put down with consistent $k_t^2$ and the Q made tolerable, there was no guarantee that control of the transmission poles and zeros of any filter design could be had. Assuming all the above could be made, there was no way to control the final frequency of the filter. And, of course, any reasonable uniformity on 4” wafers would be unthinkable, much less uniformity for 6” and ultimately 8” wafers. [For any cell phone application, frequency accuracy and uniformity (i.e. total thickness) would have to be at least 0.2%].

One could also be forgiven for doubting the sanity of anyone who believed they could do all of the above and control the stress of the acoustic stack in the free standing membrane and not have the membrane rip itself apart due to large residual stresses.

Lastly, assuming all of the above could be demonstrated in high volume (i.e. in volumes of thousands of wafers per month), one wonders whether such effort would be worth it given the relative manufacturing ease and low cost of ceramic and SAW technologies.

As an example of the perceived challenge FBAR faced, a year after cell phones with Agilent FBAR duplexers became available, this Author was repeatedly informed that FBAR would not work and that Agilent was “not really making duplexers, and, if they were, Agilent could never make them cheaply”. Others went further, “even if Agilent could demonstrate small volumes of FBAR filters in some boutique fab, there would be no way any CMOS fab would touch FBAR”. Given the number of challenges and difficulties making FBAR, skeptics had good reason to be, well, skeptical.

II. Overcoming the Challenges

One of the first choices made at Agilent (then HP) was to use AlN for the piezoelectric and Molybdenum for the electrodes. It was clear that even if ZnO could be made reliably, its incompatibility with most semiconductor facilities made it unlikely to transfer into a high volume fab. The choice of Molybdenum was more fortuitous. Early work on superconducting circuits[4] and their need for high quality resistors had given extensive experience on how to deposit “good” Mo.
Furthermore, the fundamental chemistries for AlN (a material that etches well in chlorine gas) and Mo (a material that etches well in fluorine gases) were complimentary. Molybdenum is extremely stiff and thus a low loss acoustic material with relatively high electrical conductivity. This gave strong motive to make Mo successful.

Proof that one could sputter deposit AlN with consistent $k^2_\text{T}$ was published in 2000[5] and is shown in Fig. 1. Later, it became clear that there were several “knobs” that could be turned to increase or decrease $k^2_\text{T}$ as necessary. Good control over the mass loading[5], repeatable AlN films with good $k^2_\text{T}$ and the high Q associated with Mo electrodes went a long way to establishing that high quality resonators could be made and filters could be built – albeit with a range of frequencies across the wafer.

Hard work went into the modification and ultimately the re-design of the AlN magnetron sputtering tools[6]. All of which helped enormously with repeatable AlN thickness, stress and coupling coefficient, $k^2_\text{T}$

II. FBAR Filter vs. Ceramic and SAW Filters

When comparing FBAR filters to SAW and ceramic filters, FBAR devices have superior Q and Agilent FBAR coupling coefficients, $k^2_\text{T}$, are as high as 7.1% (comparable to LiTaO3 SAW material. The FBAR temperature response is superior to SAWs (~25ppm/C vs. ~42 ppm/C) but not as good as ceramic filters. Likewise, power handling is superior to SAW devices but inferior to ceramic filters. The biggest drawback to ceramic filters and resonators is their size and the tradeoff between size and Q. In comparison, the loaded Q’s of an FBAR routinely sit at about 1000 and occasionally (depending on product and design) hit 2000.

An important Figure of Merit (FOM) is the $k^2_\text{T}$*Q product. High $k^2_\text{T}$*Q for FBAR filters directly impacts the corners of the passband filter response, the higher this product, the sharper the corners. This is particularly important when the part is operating in a phone using one of the high frequency channels at high temperature and power. The insertion loss of a filter with high FOM will stay low at the corner even under adverse conditions. Filters made with low FOMs will have increasing insertion loss at the corners. This increasing loss causes the active gain elements to work harder, draw more power, and get even hotter. A thermal run-away begins to take occur.

Two advantages of FBAR being manufactured on a silicon substrate is that the process can be implemented in a CMOS fab, thus eliminating much of the wafer cost and establishing relatively large wafer size. The second advantage of using silicon as a substrate is the development of an “all-silicon” package. This eliminates much of the back end costs (ceramic LCC or LTCC package and lid plus assemblies). An all-silicon package is formed by bonded-wafer chip scale package technique and takes full advantage of batch processing (analogous to the batch processing of transistors in an IC).

The size of these packaged filters are quite small. Figure 2 shows four packaged microcap’d filters with bumps for attachment on a single grain of rice.

III. MICROCAP: WAFER-LEVEL, CHIP SCALE PACKAGING

Microcap started in Agilent labs sometime in late 1997/98 resulting in various IP [7,8,9,10] The first published reference for microcap’ing or wafer-bonded chip scale package for FBAR was given in Feb. 2002 at the ISSCC conference[11]. At that time, we focused on a very simple process technique that utilized deep silicon etched holes wide enough to allow wire bonds onto the FBAR chip. Later as our processing became more sophisticated we “shrunk” the holes and back filled these vias with gold. In both cases hermeticity was proven.

One inherent advantage to wafer-bonding schemes was the ability to develop an understanding of hermeticity and the various factors affecting hermeticity. On a given bonded wafer there are 20 to 30 thousand testable die. This allows one to build up meaningful statistics and to understand the root-cause failure mechanisms. These statistics are also meaningful when studying lot splits. Lastly, knowledge of the special location of failing die helped (in the early days) to diagnose a myriad of problems associated with making the microcap package hermetic. Figure 3 is an early “snapshot” of approximately 20,000 die after soaking the wafer for 24 hours of 95C/98%R.H. The die are measured before and after the “soak”. As one can see the failed die (due to this harsh strife test) were clustered at the edge and near the flat of the wafer. In this case, “failure” is defined as: those die that change frequency by more than 0.5MHz. [Die that either could not be successfully tested before or after strife are ignored. In Figure 3, an example of a false “negative” due to test can be seen when there is a streak of “failed” devices along one row.]

Besides wet high temperature storage (WHTS), there are a multitude of other reliability criteria customers have come to expect. These include Wet high temperature operating life (WHOTOL), High temperature storage (HTS), high temperature operating life (HTOL), 5X IR (to mimic multiple refloors in assembly), ESD, power handling, low temperature operating life (LTOL), temperature shock...
(TS) and temperature cycling (TMCL), and mechanical shock.

Mirocap FBAR products have been evaluated on each of these criteria and have been fully qualified all of the major handset manufacturers.

IV. FBAR FABRICATION FACILITIES

One of the major stumbling blocks FBAR faces is the fact that manufacturing the device in a small fab (suitable for small volumes) results in high costs. This, in turn, limits the market size to only a few high-end applications.

To break this conundrum, one must find a large diversity fab that can help defray some of the costs inherent to fabrication (cost of land, shared equipment, etc…). Agilent made the decision two years ago to transfer FBAR from a small facility in Newark to Agilent’s 6” CMOS facility in Fort Collins. This transfer allowed for continued capacity expansion without the large investment needed to build a new facility.

What is especially nice about this arrangement is that as the CMOS products become mature, they begin to move off-shore to low cost manufacturing sites. Thus, more capacity becomes available to FBAR, creating a low cost path to high volumes. The more unit processes FBAR shares with CMOS or silicon IC processes, the less the capital outlays for new tools.

As an example, a reasonable, fully-depreciated 6” fab built in the late 80’s would have between 10,000 to 20,000 CMOS wafer starts per month capacity. Assuming a reasonable number of masks for a CMOS process of 20, then there would be 400,000 lithography steps available (ultimately) for FBAR (this assumes a one month cycle time). Assuming, (in round numbers) a 10 level FBAR mask set, then the potential lithographic capacity for FBAR could be 40,000 6” FBAR wafers per month – plenty of capacity at no extra cost.

It is very important to create an FBAR process flow that utilizes standard silicon processes. Any new or exotic materials or manufacturing processes increase the cost of FBAR significantly. This is the reason Zinc Oxide as a choice of piezoelectric was never an option for Agilent FBAR.

IV. SAW VS FBAR COST

Only time will tell which technology is truly the lowest cost method of manufacturing. SAW devices have the undeniable advantage of fully depreciated fabs and over 20 years of manufacturing experience. In contrast, Agilent announced it sold one million FBAR duplexers in early 2002. Using this as the starting date for the manufacturing ramp, FBAR production is now only two years old!

Compared to SAW die, FBAR has many more masks (although far fewer than CMOS wafers). This means longer cycle time and more tools (capital investment). However, SAW processes have become more complicated in order to address the increasingly tough specifications. The metal transducer electrodes are now multilayer composites to handle the power. To reduce the TCE (temperature coefficient of expansion), SAW manufacturers are resorting to sputtered quartz on top of the electrodes and even of bonding the SAW substrate to another substrate[12].

Mitigating factors affecting FBAR die cost (besides the making FBAR in a CMOS diversity fab), is the relatively larger wafer size and smaller die.

The biggest cost to SAW devices is the package. Today, SAW companies have various packaging strategies involving flip chipping into an LTCC package and sealed with a lid or onto a LTCC substrate with an overcoating of plastic and metal. In comparison, microcap’ing is part of the FBAR process flow. This means, the 10,000 to 30,000 FBAR die are packaged in a few steps using batch processing.

V. THE FUTURE

Beyond FBAR stand-alone products in filters and duplexers, there arises the intriguing possibility of co-designing FBAR devices with active elements. In the first iteration, Agilent is investigating and sampling Front-End Modules (FEM) that incorporate the point filter, the power amplifier (using Agilent EPHEMT technology) and duplexer.

What is exciting about this collaboration is that it quickly becomes apparent that specifications (for filter or duplexer) from the handset designer is very incomplete. Specifications from the handset manufacturers are given typically in units of dB and demand certain conditions on the magnitude of the filter response (low insertion loss in band, high rejection out of band etc…). But, what is left out is any requirement on phase. However, phase plays an important role in both the ACPR (Adjacent channel power ratio) and the PAE (power added efficiency) of the PA.

As an example of this, multiple designs of cell band duplexers were made. Each layout was designed to meet the isolation specifications given by handset manufacturers. However, one design demonstrated excellent isolation in the FEM above what the duplexer itself could do. Figure 4 shows the isolation at the Rx port....
over frequency at maximum power out at the antenna port. First it is interesting to compare the isolation measured in an FEM compared to the cellband component. The difference is believed to be due to subtle interactions in the phasing between duplexer local matching and PA.

Another specification used when comparing PA technologies (EPHEMT, LDMOS or HBT) that could be redefined is PAE. Typically, PAE is defined as

\[ \text{PAE} = \frac{P_{\text{out}} - P_{\text{in}}}{P_s} \]

where \( P_s \) is the average dc power used by the module. However, it is battery life that really counts and is what is “sensed” by the end user. A more relevant PAE specification should include the losses of the duplexer as well.

For the FEM, \( P_{\text{out}} \) is defined at the antenna port and includes the insertion loss of the duplexer. Here is where tenths of dB count. For every tenth of a dB improvement in the duplexer insertion loss, the PAE adds another 2%!

ACKNOWLEDGEMENT

The Authors wish to express their gratitude to the Fort Collins Fab engineers and management for bringing these products into commercial reality and to our management.

REFERENCES


Figure 1. Plot of median \( k_t^2 \) vs. lot run (2000)
Figure 2. Four microcap’d and bump’d PCS filters for GSM phones on a grain of rice (California short grain)

Fig. 4. A plot of the power at the Rx port of an Agilent FEM consisting of Agilent EPHEMT Power Amplifier, and Agilent Cell Band Duplexer. Isolation at the Rx port at maximum input power (~26dBm) ranges from –63 to –76 dB.

Figure 3. 2-D Plot of good die after 24 hour “soak” in 95C/95R.H. The red die are die that did not get correctly measured prior to soak. The orange die are die that either did not measure after “soak” or failing test before soak became “good” after “soak”. Yellow indicates die that shifte 0.5 MHz or more due to soak.