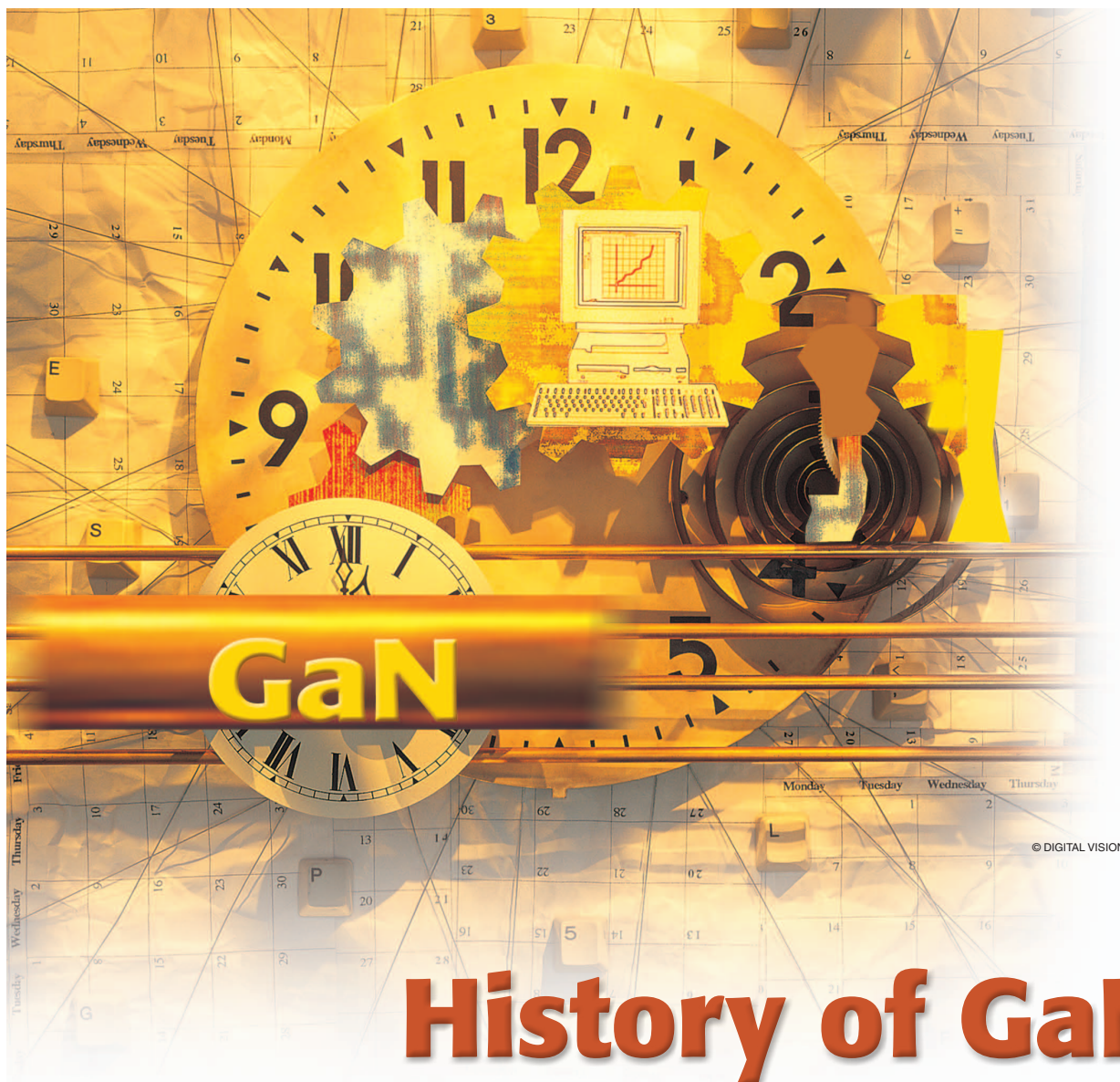




© DIGITAL VISION

History of GaN

*David W. Runton, Brian Trabert,
Jeffrey B. Shealy, and Ramakrishna Vetury*

In the early 1990s, gallium nitride (GaN) was deemed an excellent, next generation, semiconductor material for high power/high frequency transistors based on the material parameters of bandgap, electron mobility, and saturated electron velocity (Figure 1). The lack of bulk GaN source material led to the need for GaN growth on mismatched substrates such as Si, SiC and sapphire, but fundamen-

tal material development controlled the pace of maturation of GaN technology for both electronic and optoelectronic applications [1]. The development of GaN for RF electronics was significantly aided by the intense development that occurred in the race to first production of blue and, eventually, white light-emitting diodes (LEDs). Ultimately, advancements in the growth of device-grade aluminum gallium nitride (AlGaIn)/GaN

David W. Runton (drunton@ieee.org), Brian Trabert, Jeffrey B. Shealy, and Ramakrishna Vetury are with the RFMD, Chandler, Arizona.

Digital Object Identifier 10.1109/MMM.2013.2240853
Date of publication: 2 April 2013

Semiconductor (Typical Materials)		Silicon	Gallium Arsenide	Indium Phosphide	Silicon Carbide	Gallium Nitride
Characteristic	Unit					
Bandgap	eV	1.1	1.42	1.35	3.25	3.49
Electron Mobility at 300 °K	cm ² /Vs	1,500	8,500	5,400	700	1,000–2,000
Saturated Electron Velocity	× 10 ⁷ cm/s	1	1.3	1	2	2.5
Critical Breakdown Field	MV/cm	0.3	0.4	0.5	3	3.3
Thermal Conductivity	W/cm °K	1.5	0.5	0.7	4.5	>1.5
Relative Dielectric Constant	ϵ_r	11.8	12.8	12.5	10	9

Figure 1. Semiconductor material parameters.

heterostructures culminated in the demonstration of record power density RF amplifiers.

Several technical hurdles surrounding process development continued, but the promise to fulfill expectations was raised with initial demonstration of 10–11 W/mm devices, a 10× improvement compared to state-of-the-art 1 W/mm, from 28 V silicon (Si)-LDMOS. It was the eventual implementation of field plate topologies that offered the needed step change in performance [2]. By reducing the peak fields on the drain side of the gate, higher voltage operation was enabled and devices with 18.8 W/mm were demonstrated. Further optimization of the field plate topology and GaN material growth eventually led to devices capable of 32 W/mm at 4 GHz. Through field plate design, large signal performance was optimized for product release based on balancing, reliability, operational voltage range, and frequency performance [2].

Both defense and commercial applications for RF power transistors continue to push power and bandwidth limitations. Upper frequency performance limitations are inversely proportional to the capacitance of the device. The gate to source capacitance becomes the limiting parasitic, and this capacitance cascades as multiple devices are added in parallel to achieve higher powers. At some point, it is no longer possible to continue to add more gate periphery without impacting operational bandwidth [3]. A high power density commercialized disruptive technology, such as GaN on silicon carbide (SiC), offers the best potential for successfully filling these market needs.

GaN-on-SiC high-electron mobility transistor (HEMT) based power amplifiers have demonstrated high break-

down voltage, high current density, low thermal resistance, and low substrate parasitic capacitances. This combination of attributes makes this technology attractive for continual development of high power products for applications that include commercial and military radar, electronic warfare (EW), radio communications, and data backhaul (Figure 2).

Significant high voltage, high power, and wide-band capability has been demonstrated to date [4]–[6], however qualified, readily available GaN technologies from commercial sources were at first restricted to 28 V or 50 V drain bias operation. Furthermore, initial product offerings provided only one variant of the intrinsic device technology, which predetermined the trade-off between linearity and power capability.

Although intrinsic device reliability (from a wear-out perspective) has been shown by many manufacturers to approach and exceed one million hours,

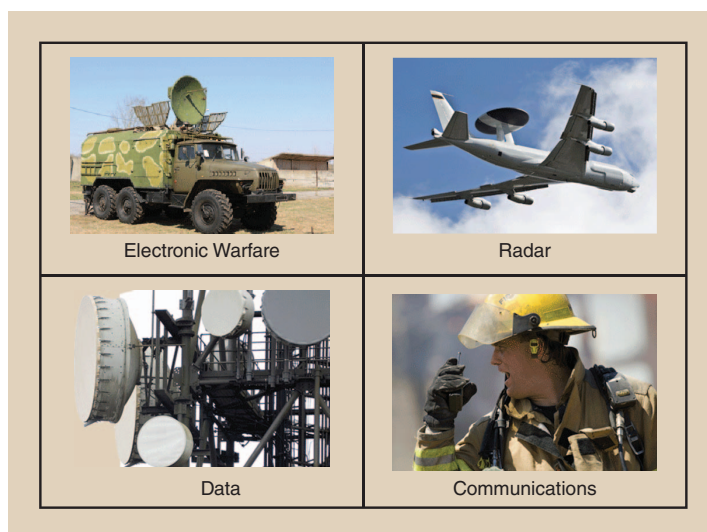


Figure 2. Applications for GaN-based RF power devices. (Photos courtesy of RFID.)

One straightforward way of providing more output power at a given gate periphery is to simply increase the drain voltage.

based on the median time to failure (MTTF) metric using an Arrhenius relationship, it is well known [7] that this metric is a necessary, but not sufficient, condition to ensure the manufacturability of a new technology. Field failure rates can arise from a variety of causes that need to be identified, contained, and eliminated. Such a process of continuous improvement, over time, results in decreasing failure rates—ultimately approaching failure rates as predicted from the purely wear-out-based Arrhenius relationship governing device MTTF.

Prior demonstrations of GaN FET reliability have been limited to 50 V or less by most manufacturers [8], [9], although RF high-temperature operating life (HTOL) data at 65 V drain bias voltage [10] has been reported. These reports, however, have been limited to three or fewer stress temperatures and are limited in the statistical description of confidence bounds on the projected lifetimes. In addition, there have been few reports [11] in the literature on the various failure modes and mechanisms that impact the failure rates in infancy and early life, which is a significant factor in determining the manufacturability of a device technology.

Of particular importance to high power PA design is the tradeoff between linearity and power density. This tradeoff arises fundamentally from the intrinsic GaN device technology, a topic insufficiently addressed thus far. Power handling capability at higher drain voltages and demonstration of

next generation device technology, offering flexibility between power and linearity performance, represent the next generation of GaN products currently being released.

GaN Technology Portfolio

First-Generation GaN Device Technologies and Incremental Improvements

Organometallic vapor phase epitaxy (OMVPE) of GaN materials on semi-insulating SiC substrates has developed to maturity. Starting with epitaxial materials with low defect density and utilizing a highly controlled fabrication process (Figure 3) with optimized SiN surface passivation, a device technology capable of high power and excellent reliability has been established.

Many initial device structures were based on a 0.5 μm optically defined gate. Although there are several options for field plate design topologies, one example uses the extension of the gate metallization into the gate-drain access region to form a gate connected field plate. A source connected field plate is additionally formed over top of this gate structure (Figure 4). This dual field plate structure achieves device breakdown voltages in excess of 300 V. SiC is the preferred substrate due to its superior thermal conductivity, low-temperature-related memory effects, and high-power-density compatibility. A baseline GaN device consists of a thin GaN cap with an AlGaN barrier on a GaN buffer.

A summary of the process capability of a first generation GaN RF power process (RFMD's GaN-1 in this case) is shown in Figure 5. The baseline device exhibits a mean threshold voltage of -3.5 V with a mean maximum current density of 1000 mA/mm. Under continuous wave (CW) operation at 2.14 GHz

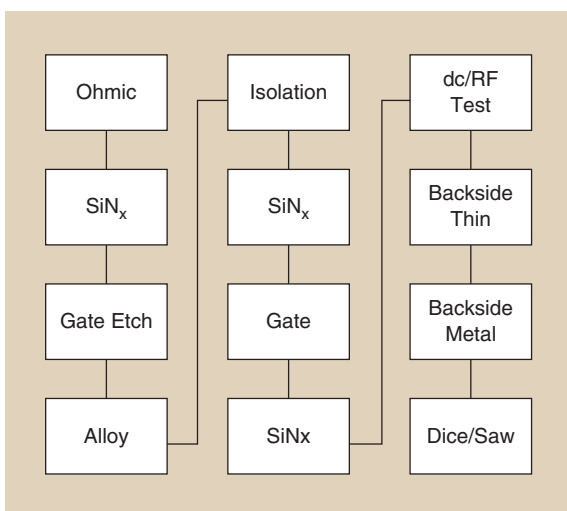


Figure 3. Simplified process flow used to fabricate GaN RF power technology [15].

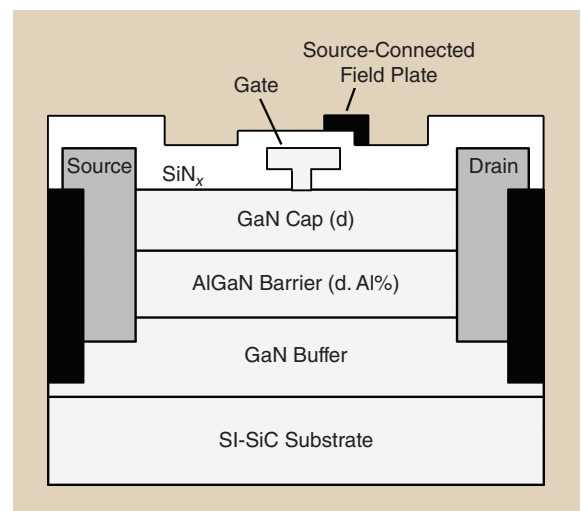


Figure 4. AlGaN/GaN device layer structure on semi-insulating SiC substrate including dual field plates [15].

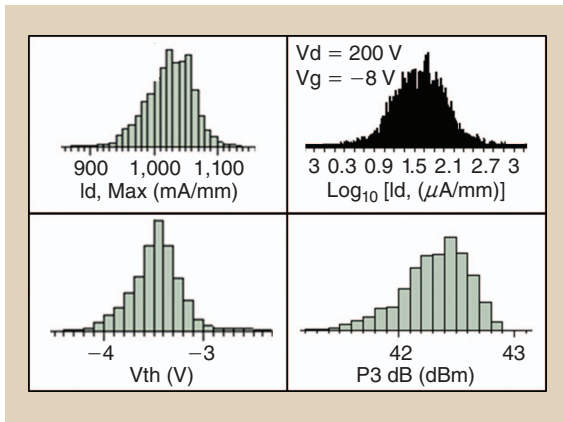


Figure 5. First generation GaN (RFMD's GaN-1) RF power process capability (threshold voltage, output power at 3-dB compression, drain current leakage at 200-V drain bias, and maximum drain current) [15].

and class-AB bias, a typical device at 48 V obtains 8.0 W/mm peak output power and over 65% peak drain efficiency. The drain leakage current (measured at $V_{ds} = 200$ V and $V_{gs} = -8$ V) shows a mean of 1.5 ($\log_{10}(I_d)$, with I_d in $\mu A/mm$).

The current and power gain cut-off frequencies (f_T and f_{max}) measured on 2.2 mm periphery devices at 44 mA of drain bias current are 12 GHz and 16 GHz respectively, at 28 V drain voltage and 11 GHz and 18 GHz, at 48 V drain voltage.

Through continued optimization of this and similar process topologies, further generations of GaN technology are being released by multiple GaN suppliers. Next-generation products continue to push performance boundaries through refinement of process parameters and foundry capabilities. These advances are advertising gate dimensions of 0.4 μm , and, with adjustments of field plate parasitics, enable 8 W/mm devices to operate through 6 GHz [12]. Even further refinements, demonstrated by Cree, TriQuint Semiconductor, and NXP Semiconductors N.V., are pushing frequency capability, with recent announcements demonstrating 0.25 μm gate topologies for use with drain voltages up to 40 V, power densities up to 6 W/mm, and operating up to 18 GHz [12].

Expanded and Optimized GaN Process Technology Portfolio

As mentioned previously, both defense and commercial applications for RF power transistors continue to push power and bandwidth limitations of existing solutions. Attempting to meet the demands of high peak power and saturated efficiency applications, like radar systems, will lead to a different technology optimization from those needed for linear, high peak-to-average ratio applications, such as the wireless infrastructure market. The lines, however, are being blurred as system requirements for

The development of GaN for RF electronics was significantly aided by the intense development that occurred in the race to first production of blue and, eventually, white LEDs.

next generation military communication systems start to overlap those of wireless infrastructure systems and vice versa. For example, next-generation military communication systems are also adopting high peak-to-average ratio communications approaches, while still demanding extreme bandwidth capability, in order to provide system agility. Equally, wireless infrastructure systems are exploring saturation mode amplifiers such as higher class (class-D/E) or envelope tracking systems to push system efficiency.

In order to address the two differing product requirements, a shift in design of the GaN technology portfolio is warranted. It is possible to optimize the transistor for linear applications by studying the causes of nonlinearities. Intentionally adjusting the shape and nonlinearity points of the three terminal device capacitances versus voltage is a prime candidate. Through optimizations of the GaN crystal structure during epitaxial growth, the gate to drain capacitance can be more linearized for the I-V plane of operation. When making these adjustment, where linearity is desired over peak power, further optimizations become possible as modifications of $I_d(max)$ and breakdown voltage provide further degrees of freedom. This allows amplifier designers other options to explore tradeoffs between power capability and linearity, while still maintaining key advantages of GaN, such as adequately high breakdown voltage and associated bandwidth enhancements.

65-V GaN Reliability

Since both defense and commercial applications continue to push power and bandwidth limitations, qualification of devices that disruptively enable higher power for a given periphery, becomes a market driver. As mentioned before, the gate to source capacitance becomes the limiting parasitic and simple device scaling eventually reaches a limit where gate periphery starts to impact bandwidth [3]. One straightforward way of providing more output power at a given gate periphery is to simply increase the drain voltage.

At this time, GaN qualification for high power RF at drain voltages of 65-V has been advertised extensively by RFMD with significant published information. This

Learning curve theory is a useful tool for projecting the cost of the n th wafer produced in a wafer factory as well as providing forward pricing projections.

section, therefore, focuses on RFMD's qualification activities for 65-V drain voltage operation. Consistent with conventions followed for III-V field-effect transistor (FET) technologies, to assess the lifetime limitation imposed by long term, thermally activated, wear-out failure mechanisms, a series of multitemperature stress tests were conducted. RFMD's power optimized, GaN-1 and linearity optimized, GaN-2 technologies were qualified at 48-V drain bias with GaN1 additionally qualified for 65-V operation.

To obtain a greater degree of confidence in the prediction of reliable life at extrapolated temperatures relevant to device use conditions, temperature tests were conducted at 48 V and 65 V at five temperatures, rather than the industry standard of three temperatures. The measured failure times were fit to an Arrhenius model, showing 95% confidence bounds by delta method (Figure 6), to extract a projected MTTF for 65 V operation.

More than 170 devices from three lots were stress tested for 1,000 hours at five temperatures chosen between 300 and 335 °C. The devices were biased with the source grounded and drain at 65 V, while the gate voltage was dynamically adjusted to maintain the desired drain current of 150 mA/mm. This methodology ensures consistent power dissipation and channel temperatures during the accelerated stress test. A similar test methodology was followed for the 48 V qualification test, the results of which were previously published [13].

The results of both the 48 V and 65 V stress tests indicated that the MTTF at the lower 95% confidence bound

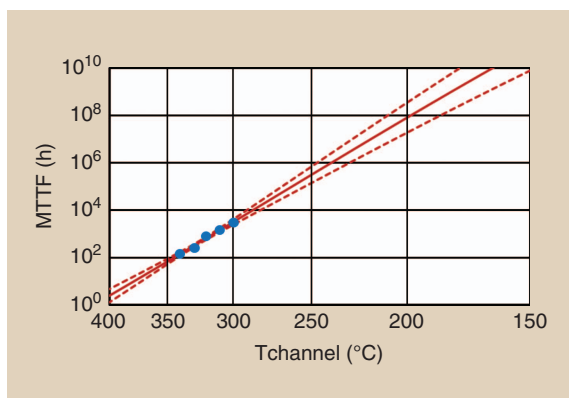


Figure 6. Result of multitemperature stress tests conducted at 65 V showing an extrapolated lifetime of greater than one million hours at 200 °C.

was greater than 1 million hours at 200 °C, an industry-leading result, and the first statistically demonstrated solid state reliability result at 65 V for a III-V FET technology.

Manufacturing

Wafer Factory and Learning Curve

Learning curve theory is a useful tool for projecting the cost of the n th wafer produced in a wafer factory as well as providing forward pricing projections. The relationship between cost and cumulative units produced was first quantified by the aircraft industry in the early 20th century. Unit cost reduction versus cumulative unit production was later generalized as the “experience curve,” which encompasses all aspects of a factory or manufacturing system. For the semiconductor industry, the learning curve is found to be between 68% and 80% [14].

As an example, Figure 7 shows the learning curve for RFMD's III-V wafer factory over the last 13 years [15]. The vertical axis shows wafer cost as a percent of the base unit cost. The horizontal axis represents the cumulative volume of GaAs shipped from the factory. As of March 2011, the company had shipped in excess of 1.5 billion square millimeters of GaAs. In this case, the RFMD factory operates on a 79% learning curve—meaning every time cumulative volume increases by two, the wafer cost is reduced by 21%. Since GaN is produced in the same factory, on the same equipment, and with the process same staff as GaAs, GaN manufacturing cost will benefit from the same scale and manufacturing learning cycles already demonstrated by GaAs as wafer volume continues to grow.

Wafer Acceptance Testing

Although MTTF is a necessary technology metric, field failure rates can arise from a variety of causes that need to be identified, contained, and ultimately eliminated. The process of continuous improvement,

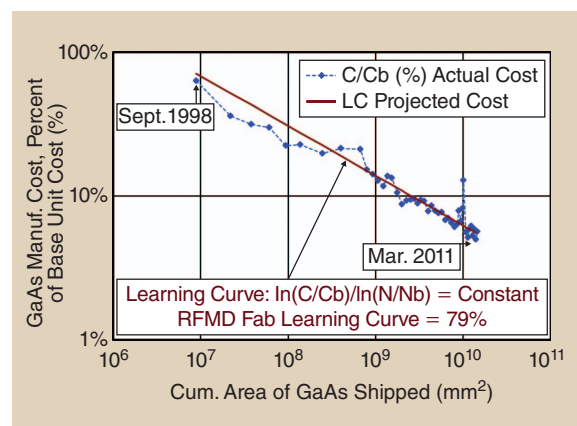


Figure 7. 13-year factory learning curve, GaAs manufacturing costs versus cumulative area shipped.

over time, results in decreasing device failure rates that approach the Arrhenius based technology limits, but to fully understand these improvements, it is necessary to capture device performance and baseline reliability in an ongoing fashion.

Optimizing manufacturing processes while simultaneously maintaining long-term reliability can be competing goals. One way to capture baseline information and help guarantee reliability through sampling is to implement an ongoing wafer acceptance test (WAT). A primary goal of a WAT is early detection and screening of wafers that pose reliability or performance risks. It is worth noting that this process is not exclusive to high-power RF GaN; it was also implemented in the early days of RF power device production of both Si-LDMOS (at Sirenza Microdevices) and Si-Bipolar (at TRW).

A typical WAT process begins by spatially sampling die across a wafer such that samples are representative of the overall wafer (Figure 8). The die are assembled into product equivalent packages and subjected to an application specific suite of dc and RF tests. The wafers may also be sampled through a high temperature life test to simulate long term reliability.

The data generated by this process may also be combined with epitaxial, process control monitor (PCM), single die dc test, and any other available on-wafer based measurements. Through this process, the wafer can be analyzed for acceptance against specifications that have been statistically

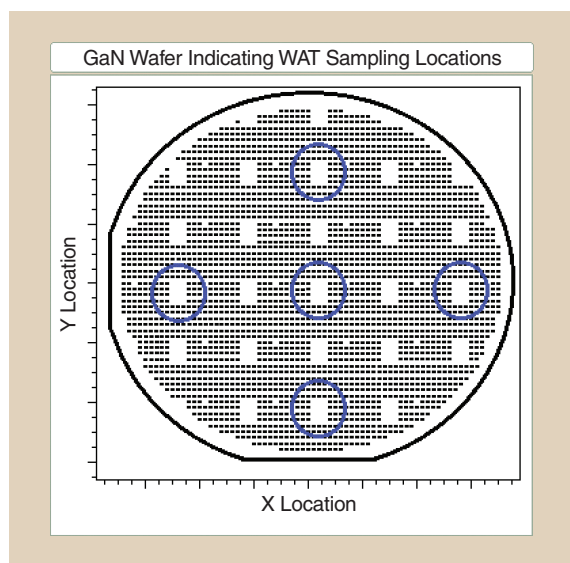


Figure 8. An illustration of how dice are picked from top, middle, bottom, left, and right regions in order to insure the WAT builds appropriately sample the within-wafer variation. Maintaining traceability of the spatial relationship of the dice is key to defining higher order correlations and interactions of the starting material on the final RF performance [15].

shown to demonstrate product reliability. Furthermore, failing wafers can be sent back for failure analysis to determine root cause and drive continuous improvement. An illustration of a typical WAT flow is shown (Figure 9).

A significant benefit of a WAT process is the vast amount of useful information contained in the measurements generated. This data set can be continually mined for insight into design sensitivities, parametric correlations, and process trends, as well as rates of specific failure modes. As the data set grows with every additional wafer, the quality of feedback to the wafer fabrication facility, assembly, and design teams becomes increasingly precise.

As a compelling example, GaN production ramp WAT data set, in conjunction with advanced failure analysis techniques, was demonstrated to resolve several different physical defects in a GaN production process [15]. A clear example is demonstrated by the identification of a type of defect in the gate metal of the GaN die with long-term reliability implications. Review of the WAT data revealed a unique electrical signature that was indicative of the defect (Figure 10). Pinpointing the electrical signature made the development of targeted screens at both die probe and final production test steps straightforward.

Although initial WAT data analysis can focus on correlating specific physical defects to easily detectable electrical signatures and then devising screens at one or more of the various production test points, the end goal is to identify fundamental root-causes for any given defect and improve the manufacturing process such that electrical screening becomes redundant. Timely, data-driven feedback, through implementation of a WAT process, results in die probe and final electrical test solution improvements, which can lead to development of surgical screens for suspect or out-of-family product earlier in the manufacturing flow.

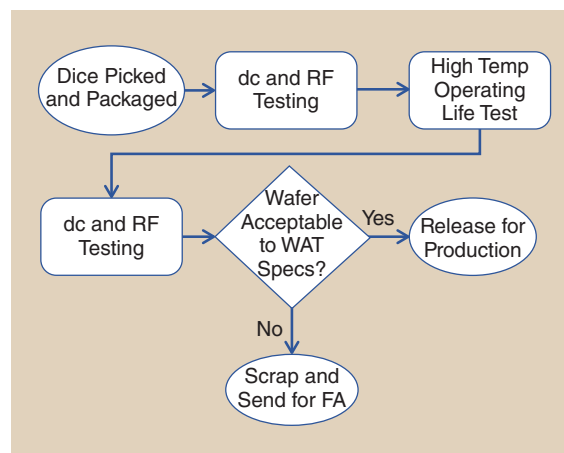


Figure 9. Flow chart of the WAT [15].

There is data statistically demonstrating that low voltage products have less need for ELF screening than products operating at higher voltages.

In addition to providing a reliability screen capability, maintaining a WAT process can also help with establishing baseline performance and lead to more tightly distributed final product performance. For

instance, maintaining traceability of every WAT die to its spatial location on a wafer provides an opportunity to uncover correlations to material and fabrication process metrics as well as within-wafer or even within-reticle variation.

As an example, the relationship between a wafer-level pinch-off measurement and the output power at the product-level was tracked and is shown (Figure 11) [15]. The relatively easy measurement of dc pinch-off voltage can then serve as a predictor of RF output performance or could be used to bin individual die into various performance categories.

GaN Early Life Failures

Modeling the failure rate of semiconductors in time has typically been broken up into three phases: infant mortality, operating life, and wear-out. This bathtub model has failures dropping to a nominal level quickly and then remaining stable until the design begins to reach the end of its useful lifetime.

Many end customers perform some form of life testing on high power amplifiers, and as such, it makes sense for GaN semiconductor suppliers to furnish some sort of early life failure (ELF) screening sample data to indicate how products may fare under these conditions. ELF screening can take many forms, including high temperature on-state soak or high temperature RF aging. Additionally, ELF conditions may be optimized over time, on a per-product basis, through extensive design of experiment (DOE) efforts.

There is data statistically demonstrating that low voltage products have less need for ELF screening than products operating at higher voltages [15]. The ELF rate has been found to be extremely low for 28 V products, but increasing with operating voltage (Figure 12). This indicates that high voltage stress can effectively be used to assist in screening weak devices and/or wafers. Including ELF screening data in the example WAT process described here has demonstrated the ability to drive improvements and identify root-cause for weak devices.

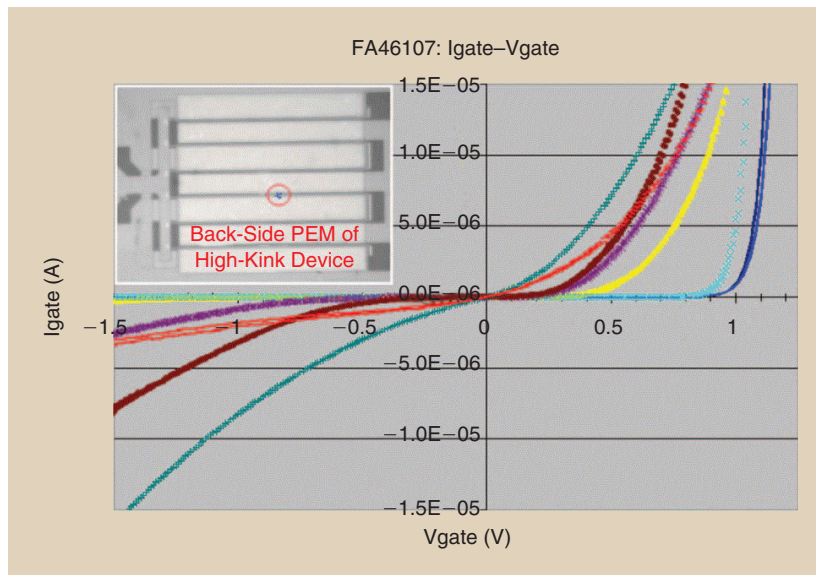


Figure 10. PEM indicating a die level defect and how such a device can be detected through subthreshold I-V curve variations and thus screened.

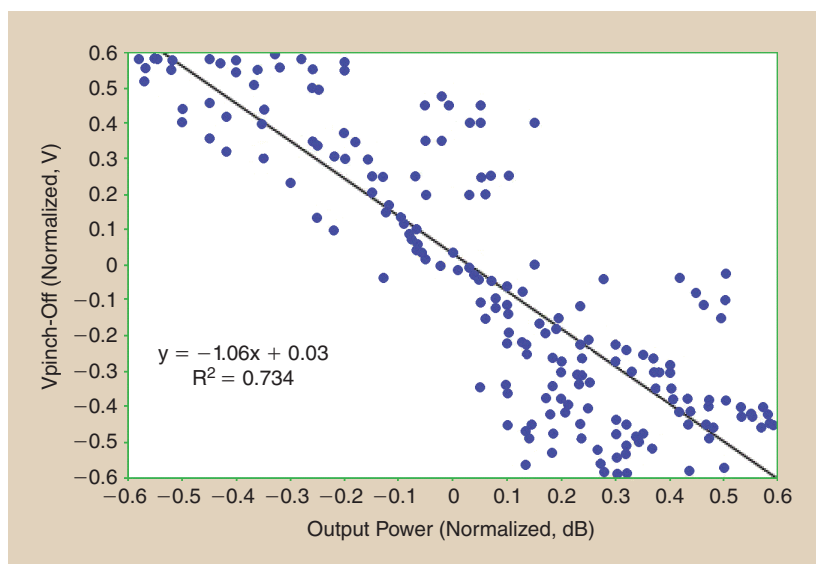


Figure 11. Normalized correlation of wafer-level pinch-off voltage to the RF output power of a packaged GaN amplifier.

With all new technologies, there are unanswered questions. For GaN products, these questions surround applicability of wear-out mechanisms and drift/aging characterization that can influence performance in uncertain ways. Although the WAT and ELF processes outlined here can represent a high level of overhead, following this sampling process through new technology introduction provides limitless data analysis potential to predict and respond to these concerns. Many GaN suppliers indicate significant amounts of reliability data showing product aging over time, but few advertise the extensive use of a WAT/ELF process on all GaN wafers produced to drive product improvement, robustness, and high volume manufacturability. It is strictly the desire to push new technology to its full performance realization that drives the need for continuous screening and improvement. Accepting the risk of new technology, at a less than optimal performance improvement, is rarely worth the investment.

GaN High Power Amplifiers Portfolio

Once the basic GaN power amplifier building block technology is qualified, the next step is to convert this into products to target specific market segments. At this time, the qualification and assembly efforts outlined here move from die/wafer and wafer foundry to sub-assemblies. Packaging becomes a major challenge as the increased power density (qualified 6-8 W/mm) must dissipate significantly more heat in a smaller area than entrenched semiconductor amplifier technologies. Careful attention to safe operating area and power/temperature de-rating curves must be considered during system integration.

Unmatched Power Amplifiers

For applications where a generic gain block, requiring 10–150 W over a narrow, less than 10%, operating bandwidth is required, and where discrete PCB matching is possible, a simple unmatched power amplifier is a likely candidate.

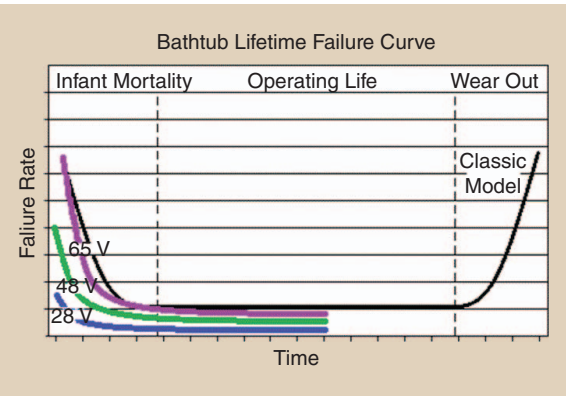


Figure 12. A typical bathtub failure curve which illustrates how the failure rate drops dramatically as infant failures are removed from the population [15].

Once the basic GaN power amplifier building block technology is qualified, the next step is to convert this into products to target specific market segments.

With increased power densities from GaN, 10–150 W solutions in a single, small form factor package, without any intentional impedance transformation network, other than required interconnect, are available. The basic unmatched power amplifier makes no frequency optimization choices for the end user, instead it simply provides the shortest interconnect possible, and allows use in any desired frequency band. Figure 13 shows the internal interconnect design of a simple 60 W unmatched amplifier.

The low output capacitance, driven by reduced CDs of GaN technology, provides the end user with a fairly simplistic, low Q, impedance matching task, while the input impedance is much more challenging due to the higher input impedance Q. Still, the end user is allowed, with this unmatched topology, to accept the performance tradeoffs and operate the device anywhere within its usable bandwidth range.

Typical narrow-band amplifiers using straight forward lumped or distributed matching networks can easily obtain 5% operating bandwidth with optimized performance. 20–50% bandwidth is also achievable utilizing the same matching topologies, but system specifications, such as input VSWR usually require balanced or quadrature combined

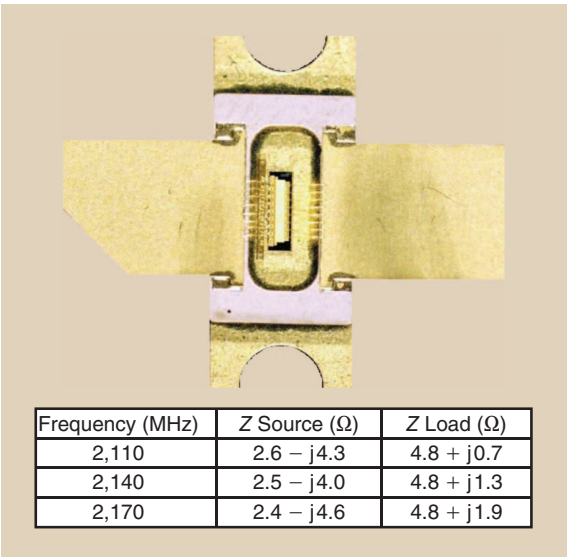


Figure 13. A typical unmatched 60-W, 48-V GaN HEMT and its associated input/output impedances. Devices in this power family typically demonstrate >14-dB power gain and drain efficiency >65% at P-3 dB.

The significant benefit of fully matched amplifiers is the small form factor achievable at the full circuit level.

topologies that can ease the impedance matching challenge or provide isolation to hide the typical high input reflected power when matching high Q input impedances over a large bandwidth.

Partially Matched Power Amplifiers

For systems requiring above 75 W and operational bandwidths of 15% or higher, the shortfall of the unmatched amplifier must be addressed. In these cases, precise printed circuit board (PCB) matching becomes much more difficult and combinational (or balanced) topologies are not always desirable. To simplify the impedance transformation, internal pre-match circuitry using MOS capacitors, or other equivalent components, may be incorporated.

Sub-2Ω impedance system matching difficulties are eased with partial matching where the first and, possibly, the second matching stages are implemented inside the package, attempting to balance both the requirements for a higher magnitude and lower-Q impedance target. Unlike equivalent Si-LDMOS or GaAs solutions, which face both input and output impedance challenges, the main difficulty lies with the input impedance. Packaged solutions that provide some sort of pre-matching on the input give a more simplistic solution to the end user (Figure 14). Typical GaN solutions in this product range do not offer nor need any output pre-matching. This allows system designers to exploit memory effect reduction circuitry or drain voltage modulation without adversely affecting the optimal load impedance or performance of the amplifier.

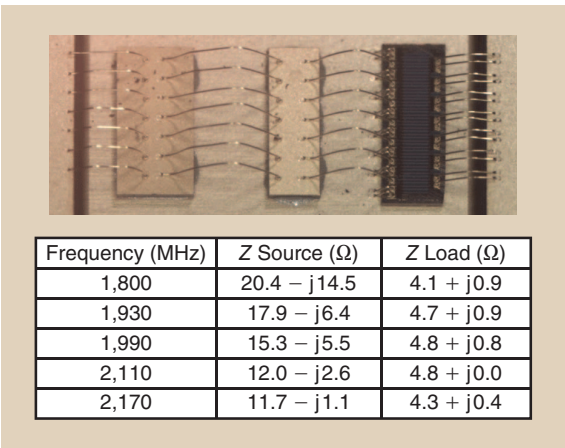


Figure 14. A typical input partially matched 90-W, 48-V GaN HEMT and its associated input/output impedances. Devices in this power family typically demonstrate >15-dB power gain and drain efficiency >65% at P-3 dB.

Typical 15% bandwidth amplifiers are easily implemented when the impedance requirements of the power amplifier (PA) are simplified using this approach. Figure 15 shows a 120 W, 15% bandwidth amplifier from Cree. The distributed matching network fits within a 4 × 1.5 in. area.

Fully Matched Power Amplifiers

As power capability increases above 200 W, the challenge of impedance matching forces even higher integration at the package level. Additionally, some users of these solutions desire to maximize the power output within as small a space as possible and be provided the most simplistic matching challenge. To meet this need, the use of high terminal impedance matched power transistors is common. Amplifier designers prefer 35Ω (or higher) terminal impedances to keep the radar amplifiers as small as possible.

The typical fully matched amplifier topology (schematic in Figure 16 and realization from Sumitomo Electronic Device Innovations shown in Figure 17) uses two devices combined using a Wilkinson power divider/combiner on the input and output respectively. This topology achieves wider bandwidth than would be obtained using a single higher power device [17]. Along with the power division/combination function, the Wilkinson design also performs the impedance transformation required to provide the optimum source and load impedance to the GaN devices and 35Ω (or higher) terminal impedances.

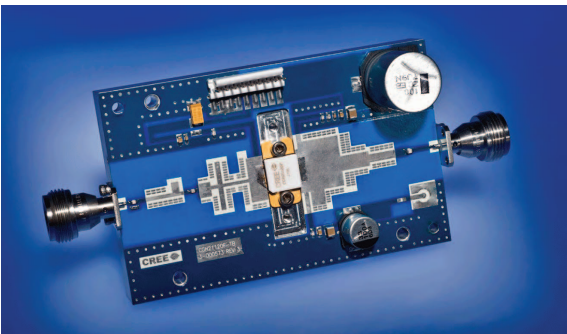


Figure 15. Compact distributed solution for a 120-W peak power, 15% operational bandwidth PA from Cree [16]. (Photo used with permission of Cree, Inc.)

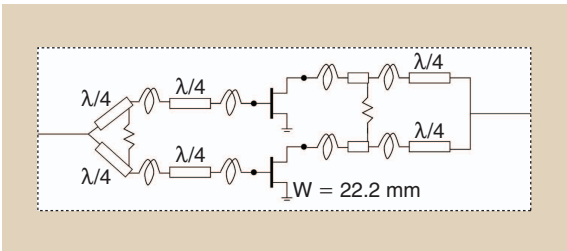


Figure 16. Schematic of a high power fully matched amplifier [17].

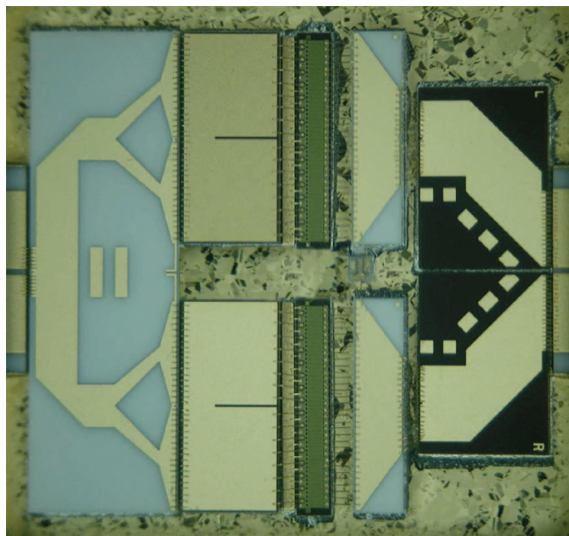


Figure 17. Realization of a high power fully matched PA architecture from SEDI [18].

The significant benefit of fully matched amplifiers is the small form factor achievable at the full circuit level. Figure 18, shows an example from Cree demonstrating typical simplistic matching circuitry needed for fully matched amplifiers that enable high powers of 250–400 W while requiring minimal circuit board matching components.

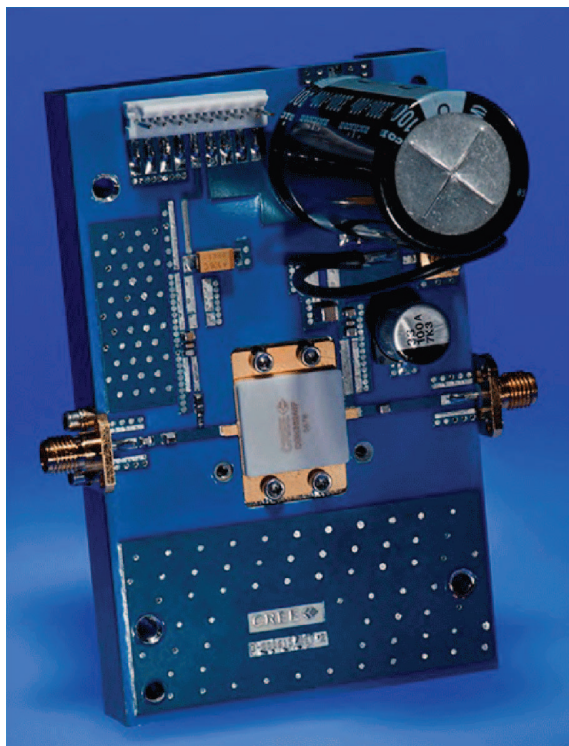


Figure 18. Compact S-band PA module from Cree [19]. (Photo used with permission of Cree, Inc.)

A significant benefit of a WAT process is the vast amount of useful information contained in the measurements generated.

To enable higher power devices, many GaN power amplifiers for L and S-band radar applications are developed utilizing 28–50 V supply rails, but the true benefit occurs when higher supply voltages are employed. As mentioned previously, qualification at 65 V can start to provide these benefits.

The increase in power afforded by higher voltage operation can be seen as in Figure 19. Through simple scaling, a 325 W previously designed power transistor was used as the test vehicle. RF power performance versus drain voltage was tested and a linear correlation was verified (Figure 19) with a statistical confidence R^2 fit of 0.88. For this large device, a $0.09 \Delta \text{ dBm/Vd}$ offset allows for a 50 V device to be operated at 65 V, provide 1.35 dB more output power and achieving 400 W over the same bandwidth without any further load line optimization [15].

In addition to easing impedance matching, in high-volume production (such as phased array amplifiers), high terminal impedances of the matched power transistor help to absorb power variation from all sources of variation in the design. High volume testing of a typical matched power amplifier product, including the variation of the amplifying die, passive components, assembly, and testing shows power variation of the amplifier is limited to $\pm 0.15 \text{ dB}$ capturing 68% of the distribution. This is based on testing of more than 1400 devices, from over 25 wafers, (Figure 20) using pulsed RF at 3 dB or more gain compression (at or near power saturation). [15] From this data the advantages of the fully matched amplifier solution can be seen.

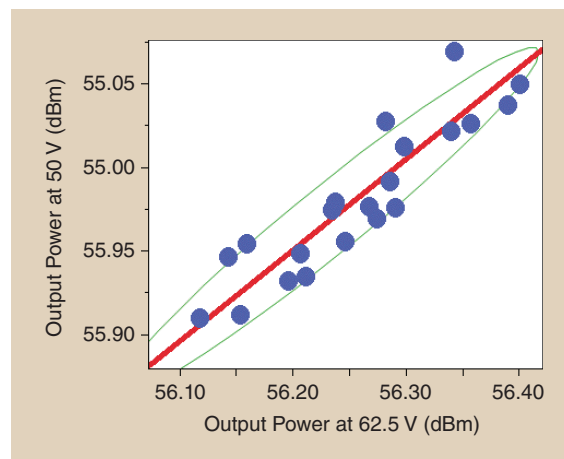


Figure 19. Correlation between output power at 50 V and 62.5 V [15].

SiC is the preferred substrate due to its superior thermal conductivity, low-temperature-related memory effects, and high-power-density compatibility.

GaN with Enhanced Linearity

As an extension to the above product families, high power GaN technologies that are truly optimized for saturated power applications achieving >7 W/mm do not perform well when backed-off in power in an attempt to achieve linear performance. In particular, the nonlinearity of these high power optimized technologies stems from the steep nonlinearity of gate to drain capacitance (C_{gd}) at low drain voltages (Figure 21). It is worth noting that various correction techniques, both analog and digital, have been found to adequately linearize these high power optimized devices.

To optimize for (noncorrected) linearity, however, it is necessary to minimize the changes in parasitic capacitance over the region of operation for the device. During operation, voltage is continually varying at the device terminals (on the load line), and thus by minimizing capacitance change versus voltage, linearity is improved. As mentioned previously, it is possible to optimize the device structure for linearity through optimization of the capacitance versus voltage transfer curves as shown in Figure 21. The gradual change in C_{gd} versus drain voltage significantly changes the linearity signature of the amplifying device.

As a result of this optimization, the two-tone linearity can be improved substantially. The plot in Figure 22

shows how intermodulation distortion using a 1 MHz tone separation versus output power can be impacted. The case shown compares power optimized GaN-1 and linearity optimized GaN-2 processes from RFMD. Although the data plotted is for third-order intermodulation distortion, this same linearity improvement can be seen in adjacent channel power performance for noise-like modulated signals such as commercial Interim Standard 95, the first CDMA-based digital cellular standard, or the Third Generation Partnership Project (3 GPP).

The linearity optimization outlined in Figure 22 is not necessarily considered a topology for power amplifiers, but instead represents an alternate optimization for GaN device technologies. This technology can be deployed into each of the topologies previously presented (unmatched, partially, and fully matched) and offers a bandwidth advantageous solution compared to existing semiconductor devices, such as Si-LDMOS/GaAs, without adversely affecting linearity performance.

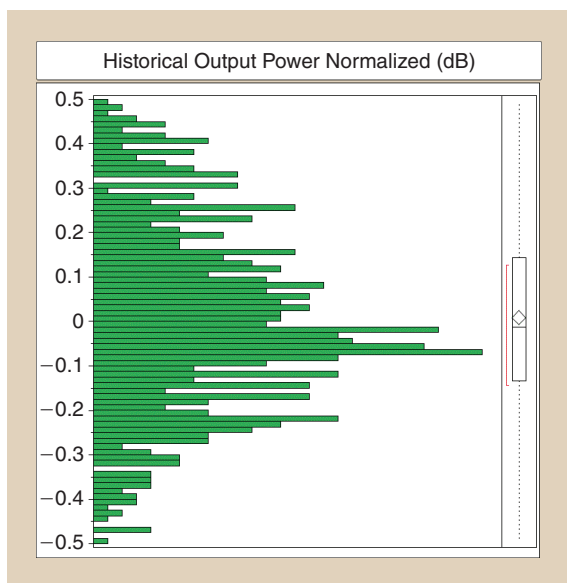


Figure 20. Output power distribution of more than 1,400 S-band power transistors.

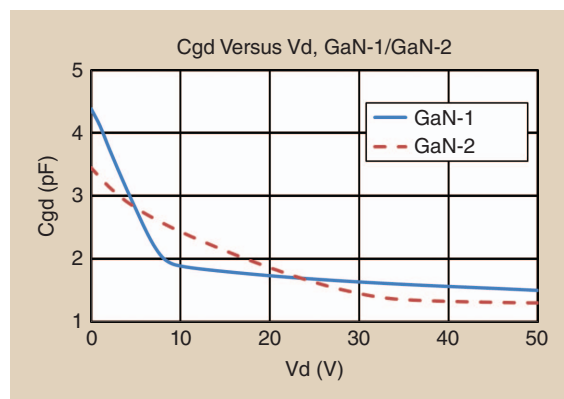


Figure 21. Gate to drain parasitic capacitance (C_{gd}) versus drain voltage (power optimized GaN-1 versus linearity optimized GaN-2) for an equivalent powered device.

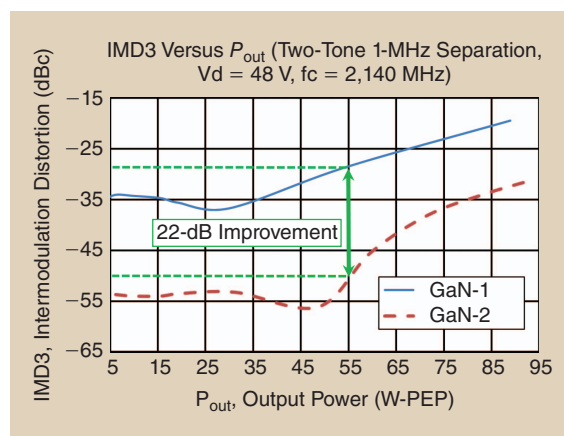


Figure 22. Two-tone intermodulation distortion versus output power for GaN-1 and GaN-2. At 55 W-PEP, a 22 dB improvement in IMD3 is observed.

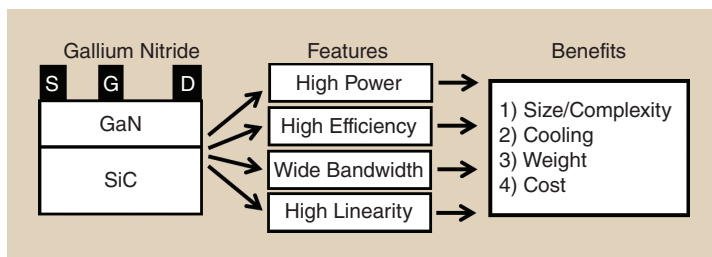


Figure 23. Summary of features and benefits of GaN RF power technology [15].

System Considerations

As originally envisioned by the semiconductor pioneers exploring the potential of GaN, many market spaces can benefit from the realized product capabilities. A summary of the features and associated system benefits offered by GaN is provided (Figure 23). For applications in the defense market, GaN offers significant advantages over existing Silicon, GaAs, and vacuum electronic device-based amplifiers, including higher voltage and broadband performance with high drain efficiency. Furthermore, radio communications and data backhaul applications can also benefit from these same characteristics. Recently, alternate technology optimizations also allow for low spectral regrowth applications and, in this case, GaN can now provide linear power, wide bandwidth, and high efficiency to enable new architectures for next generation data systems. Altogether, the increased linear power capability from a single GaN PA translates to size and weight reduction with greater combining efficiencies due to fewer number of PA modules.

System designers continue to balance performance benefits with the maturity of a technology. The data set for application of GaN in different market spaces continues to develop. Publications demonstrating usefulness for different applications continue to reduce technology level concerns by systematically addressing fundamental material issues [20] and highlighting advancements in reliability [15].

Conclusion

This article presented a brief history of GaN development, from early promising conceptual phases to robust product development. Details about first generation GaN processes as well as enhancement possibilities for next generation, market specific, processes have been presented. One example of 65 V process qualification, which is furthering the promise of GaN market differentiation, is shown and manufacturability with an eye on continuous improvement has demonstrated how the GaN process will continue to mature as new product portfolios are released and adopted.

References

- [1] J. Zolper, "Advanced device technologies for defense systems," in *Proc. Device Research Conf.*, 2012, pp. 9–12.
- [2] U. Mishra, "Status of AlGaIn/GaN HEMT technology—A UCSB perspective," in *Proc. IEEE Gallium Arsenide Other Semiconductor Application Symp.*, 2005, pp. 21–27.
- [3] J. Zolper, "Wide bandgap semiconductor microwave technologies: From promise to practice," in *Proc. IEEE Int. Electron Devices Meeting*, 1999, pp. 389–392.
- [4] Y.-F. Wu, A. Saxler, M. Moore, T. Wisleder, U. K. Mishra, and P. Parikh, "Field-plated GaN HEMTs and amplifiers," in *Proc. IEEE Compound Semiconductor Integrated Circuit Symp.*, 2005, pp. 1–4.
- [5] A. Maekawa, T. Yamamoto, E. Mitani, and S. Sano, "A 500 W push-pull AlGaIn/GaN HEMT amplifier for L-band high power application," in *IEEE Int. Microwave Symp. Dig.*, 2006, pp. 722–725.
- [6] K. Krishnamurthy, T. Driver, R. Vetury, and J. Martin, "100 W GaN HEMT power amplifier module with > 60% efficiency over 100–1000 MHz bandwidth," in *IEEE Int. Microwave Symp. Dig.*, 2010, pp. 940–943.
- [7] *JESD91A*, JEDEC standard 91A, 2003.
- [8] J. L. Jimenez and U. Chowdhury, "X-band GaN FET reliability," in *Proc. IEEE Int. Reliability Physics Symp.*, 2008, pp. 429–435.
- [9] T. Ohki, T. Kikkawa, Y. Inoue, M. Kanamura, N. Okamoto, K. Makiyama, K. Imanishi, H. Shigematsu, K. Joshin, and N. Hara, "Reliability of GaN HEMTs: Current status and future technology," in *Proc. IEEE Int. Reliability Physics Symp.*, 2009, pp. 61–70.
- [10] F. Yamaki, K. Inoue, N. Ui, A. Kawano, and S. Sano, "A 65 % drain efficiency GaN HEMT with 200 W peak power for 20 V to 65 V envelope tracking base station amplifier," in *IEEE Int. Microwave Symp. Dig.*, 2011, pp. 1–4.
- [11] Y. Inoue, S. Masuda, M. Kanamura, T. Ohki, K. Makiyama, N. Okamoto, K. Imanishi, T. Kikkawa, N. Hara, H. Shigematsu, and K. Joshin, "Degradation-mode analysis for highly reliable GaN-HEMT," in *IEEE Int. Microwave Symp. Dig.*, 2007, pp. 639–642.
- [12] (2012, June 14). Cree unveils new gallium nitride RF MMIC process technologies to enable lower cost, higher performance telecom and radar systems. [Online]. Available: www.cree.com/news-and-events/cree-news/press-releases
- [13] RF Micro Devices. (2013). [Online]. Available: www.rfmd.com/foundry/
- [14] R. McIvor, *Managing for Profit in the Semiconductor Industry*. Englewood Cliffs, NJ: Prentice-Hall, 1989, pp. 2–19.
- [15] J. B. Shealy, R. Vetury, B. Trabert, and D. W. Runton, "A manufacturable, high power RF gallium nitride (GaN) technology portfolio with 65 V operation and enhanced linearity," in *Proc. IEEE Int. Conf. Microwaves, Communications, Antennas Electronics Systems*, 2011, pp. 1–8.
- [16] (March 2012). CGH21120F, 120 W, 1800–2300 MHz, GaN HEMT for WCDMA, LTE, and WiMAX. [Online]. Available: www.cree.com
- [17] K. Krishnamurthy, J. Martin, B. Landberg, R. Vetury, and M. J. Poulton, "Wideband 400 W pulsed power GaN HEMT amplifiers," in *IEEE Int. Microwave Symp. Dig.*, 2008, pp. 303–306.
- [18] A. Kawano, N. Adachi, Y. Tateno, S. Mizuno, N. Ui, J. Nikaido, and S. Sano, "High-efficiency and wide-band single-ended 200 W GaN HEMT power amplifier for 2.1 GHz W-CDMA base station application," in *Proc. Asia-Pacific Conf. Microwave*, 2005.
- [19] (April 2012). CGH35240F, 240 W, 3100–3500 MHz, 50-ohm input/output matched, GaN HEMT for S-Band radar systems. [Online]. Available: www.cree.com
- [20] F. Fornetti, M. A. Beach, and K. A. Morris, "Time and frequency domain analysis of commercial GaN HEMTs," in *Proc. Asia-Pacific Conf. Microwave*, 2009, pp. 1601–1604.

