THz Transistors, sub-mm-wave ICs, mm-wave Systems

Mark Rodwell
University of California, Santa Barbara

rodwell@ece.ucsb.edu   805-893-3244, 805-893-5705 fax

The End (of Moore's Law) is Near (?)

It's a great time to be working on electronics !

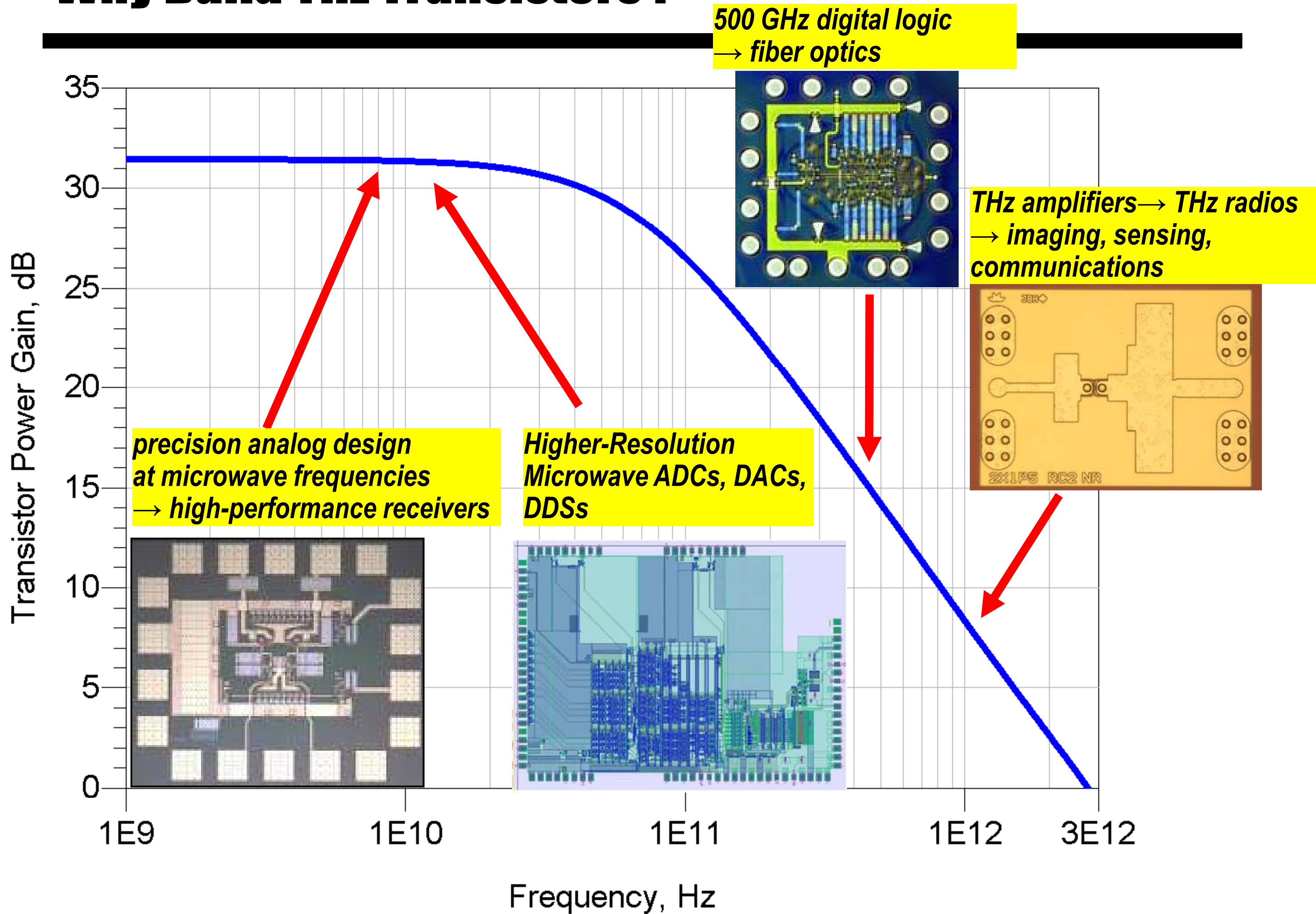
Things to work on:

- *InP transistors: extend to 3-4 THz→ GHz & low-THz ICs*
- *GaN HEMTs: powerful transmitters from 1-300 GHz*
- *Si MOSFETs: scale them past 16 nm*
- *III-V MOSFETs: help keep VLSI scaling (maybe)*
- *VLSI transistors: subvert Boltzmann→ solve power crisis*
- *mm-wave VLSI: massively complex ICs to re-invent radio*

Our focus today: THz transistors - how and why

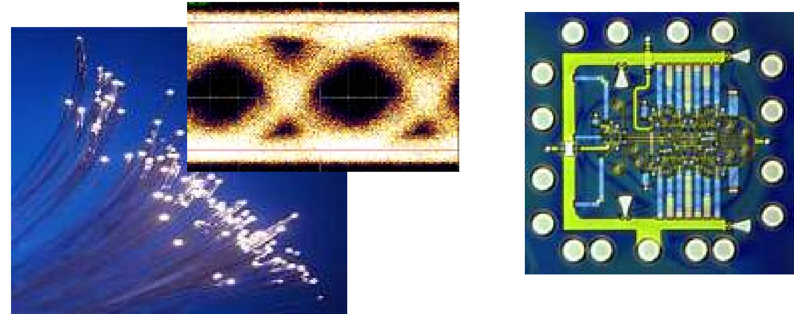
Why
THz Transistors ?

Why Build THz Transistors ?



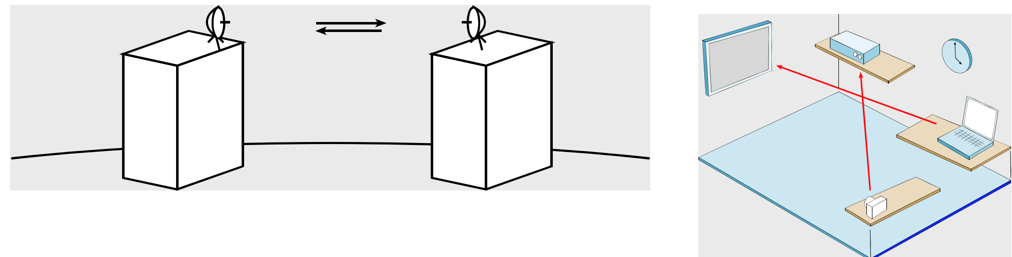
What Would You Do With a THz Transistor ?

640 Gb/s ETDM optical fiber links



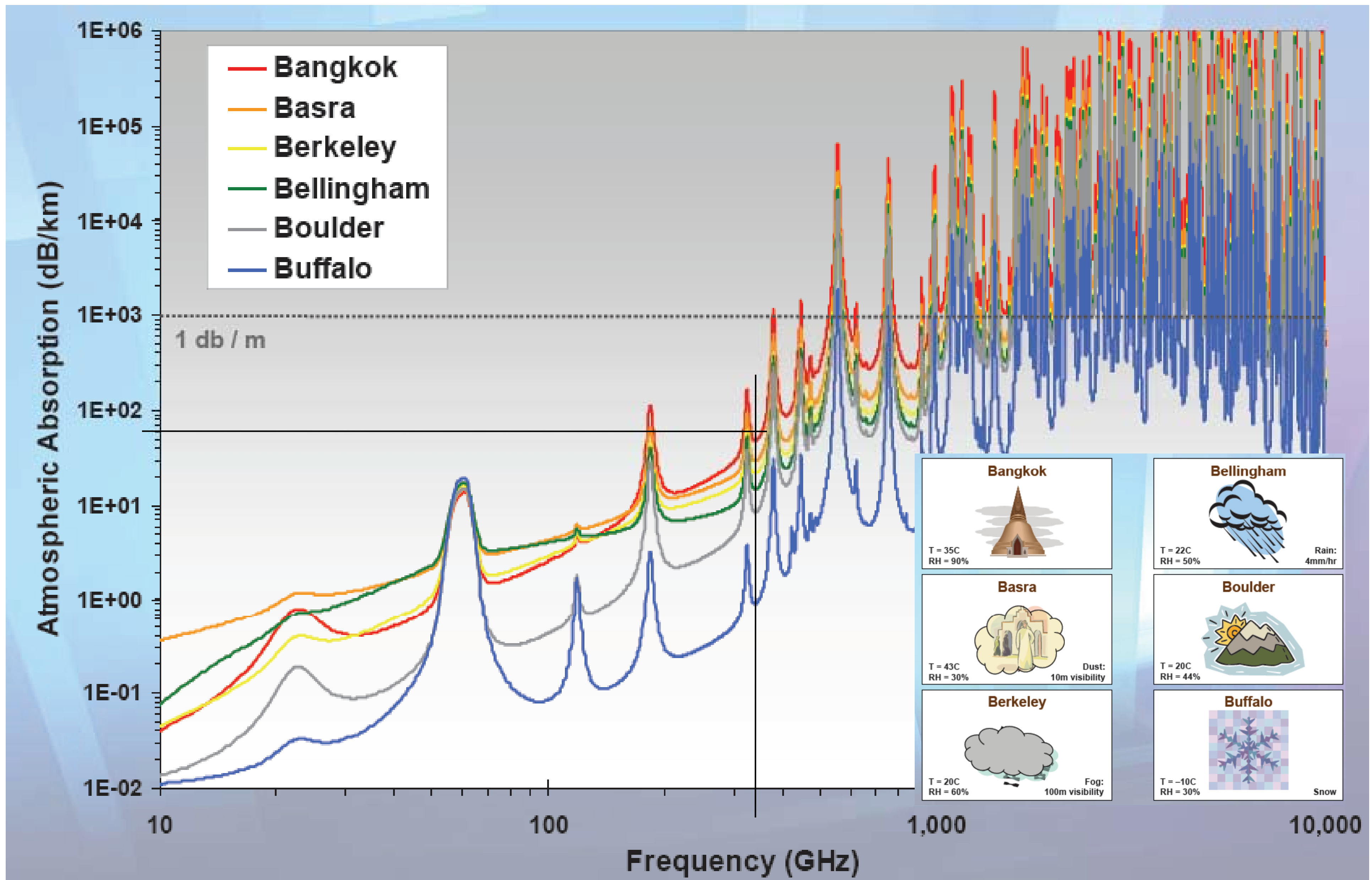
300-1000 GHz imaging systems

mm-wave communication links



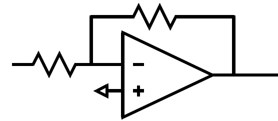
At High Frequencies The Atmosphere Is Opaque

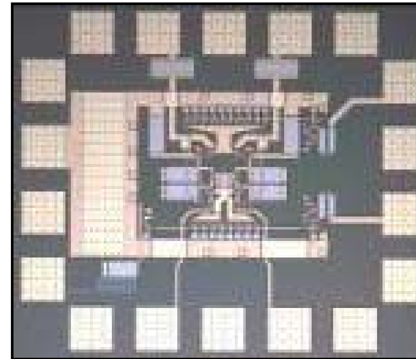
Mark Rosker
IEEE IMS 2007

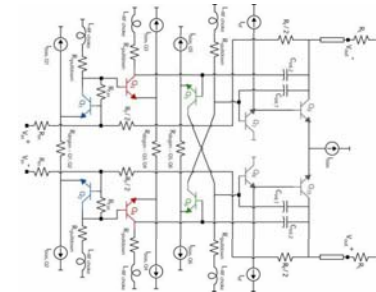


What Else Would You Do With a THz Transistor ?

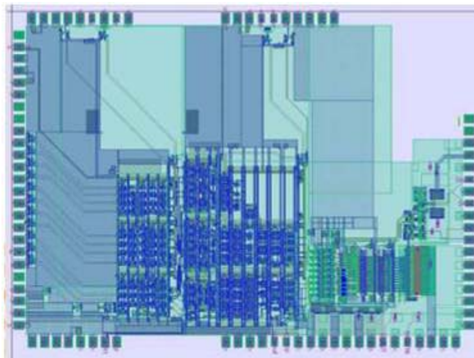
precision, high-performance analog microwave circuits

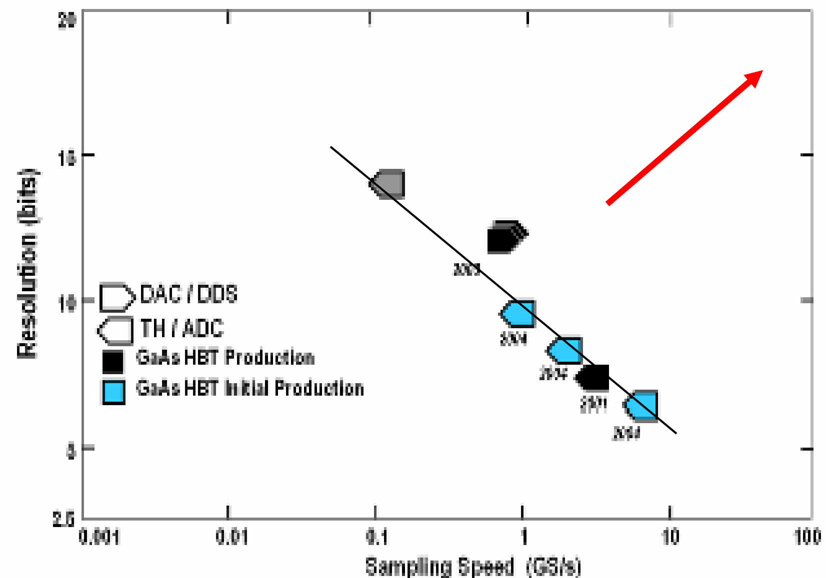






higher-resolution microwave ADCs, DACs, DDSs

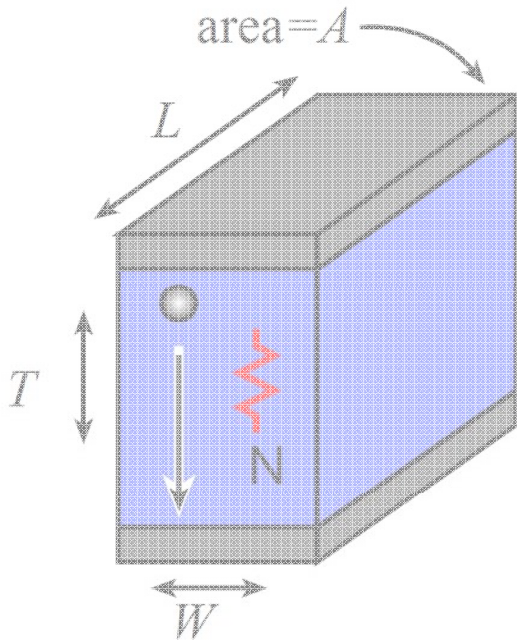




How to Make THz Transistors

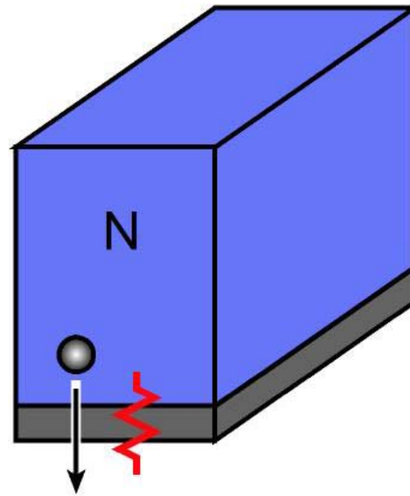
Simple Device Physics: Resistance

bulk resistance



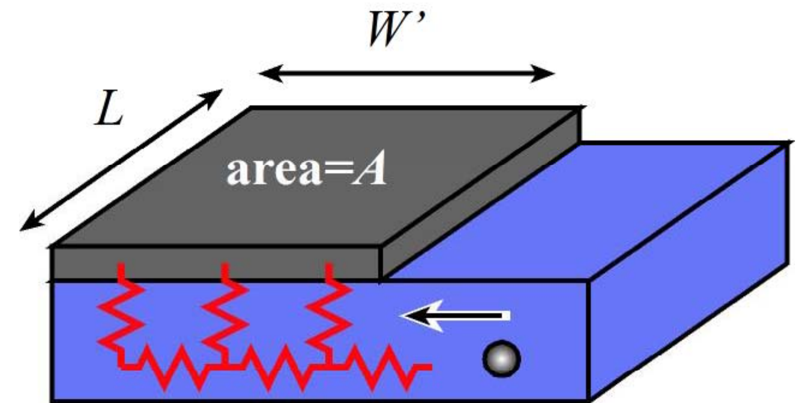
$$R = \frac{\rho_{bulk} \cdot T}{A}$$

contact resistance -perpendicular



$$R = \frac{\rho_{contact}}{A}$$

contact resistance - parallel

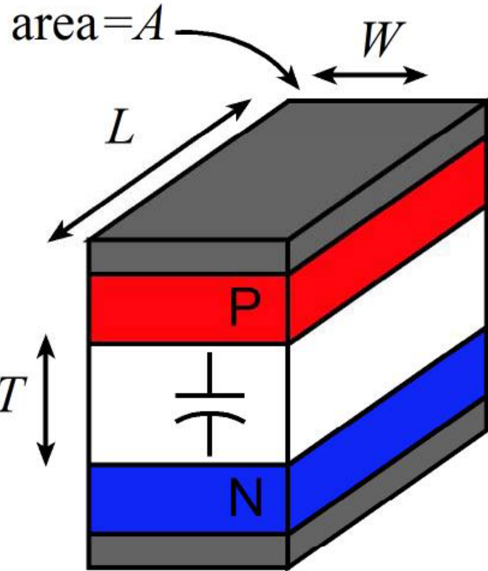


$$R = \frac{\rho_{contact}}{A} + \rho_{sheet} \cdot \frac{W'}{3L}$$

Good approximation for contact widths less than 2 transfer lengths.

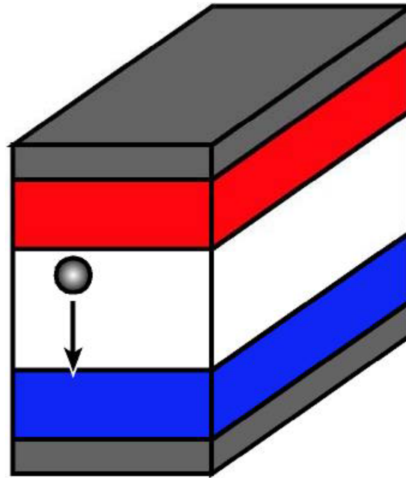
Simple Device Physics: Depletion Layers

capacitance



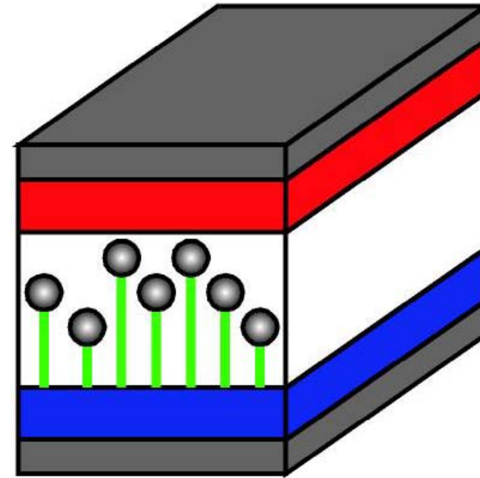
$$C = \varepsilon \cdot \frac{A}{T}$$

transit time



$$\tau = \frac{T}{2v}$$

space-charge limited current



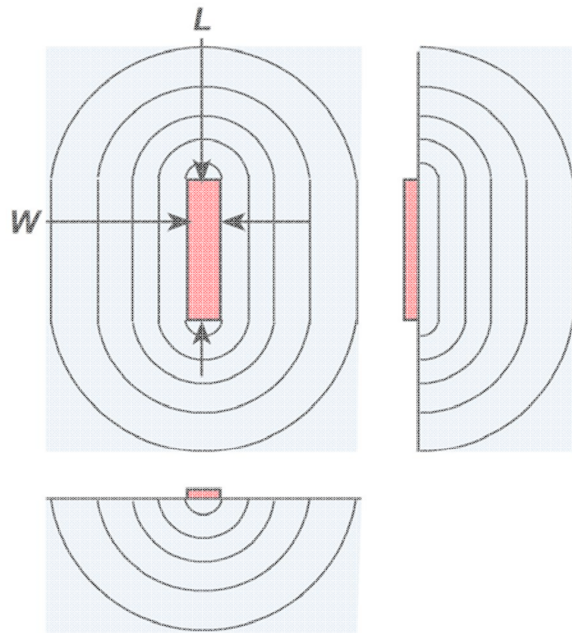
$$\frac{I_{\max}}{A} = \frac{2\varepsilon v}{T^2}\left(V_{applied} + V_{depletion} + 2\phi\right)$$

$$I = C\frac{\Delta V}{\Delta T} \text{ where } \frac{C}{I_{\max}} = \frac{\tau}{V_{applied} + V_{depletion} + 2\phi}$$

Simple Device Physics: Thermal Resistance

Exact

Carslaw & Jaeger 1959



$$R_{th} = \frac{1}{\pi K_{th} L}\sinh^{-1}\left(\frac{L}{W}\right) + \frac{1}{\pi K_{th} W}\sinh^{-1}\left(\frac{W}{L}\right)$$

Long, Narrow Stripe

HBT Emitter, FET Gate

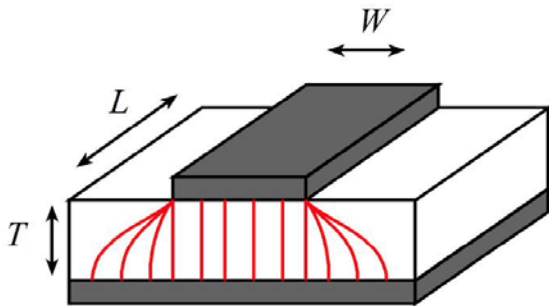
$$R_{th} \cong \underbrace{\frac{1}{\pi K_{th} L}\ln\left(\frac{L}{W}\right)}_{\text{cylindrical heat flow near junction}} + \underbrace{\frac{1}{\pi K_{th} L}}_{\text{spherical heat flow far from junction}}$$

Square (L by L)

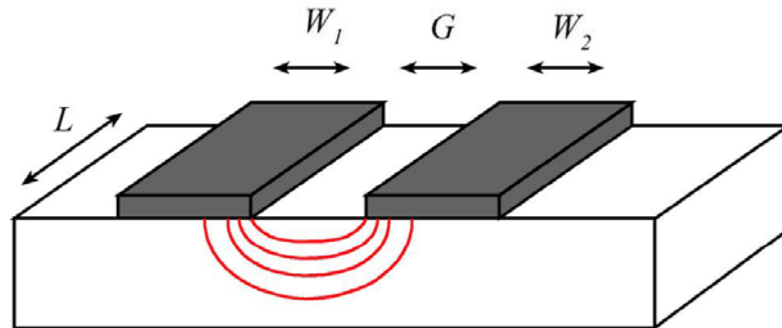
IC on heat sink

$$R_{th} \cong \underbrace{\frac{1}{4 K_{th} L}}_{\text{planar heat flow near surface}} + \underbrace{\frac{1}{\pi K_{th} L}}_{\text{spherical heat flow far from surface}}$$

Simple Device Physics: Fringing Capacitance



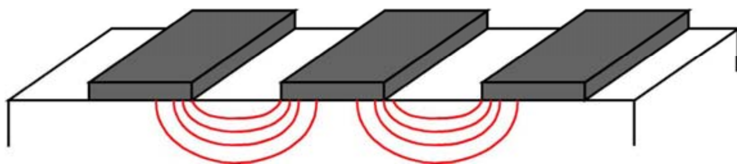
$$\frac{C}{L} \cong \underbrace{\varepsilon \cdot \frac{W}{T}}_{\text{parallel - plate}} + \underbrace{1.5 \cdot \varepsilon}_{\text{fringing}}$$



$$\frac{C}{L} \cong \varepsilon \cdot \left[\begin{array}{l} \text{slowly - varying function} \\ \text{of } W_1 / G \text{ and } W_2 / G \end{array} \right]$$

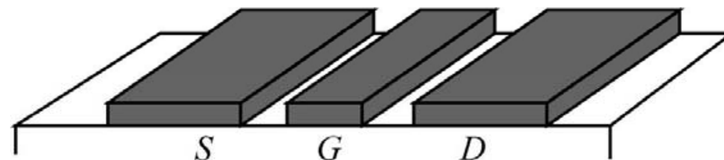
$$\approx (1 \text{ to } 3) \cdot \varepsilon$$

wiring capacitance



$$C / L > \varepsilon$$

FET parasitic capacitances



$$C_{parasitic} / L \sim \varepsilon$$

VLSI power-delay limits

FET scaling constraints

Frequency Limits and Scaling Laws of (most) Electron Devices

$$\tau \propto \text{thickness}$$

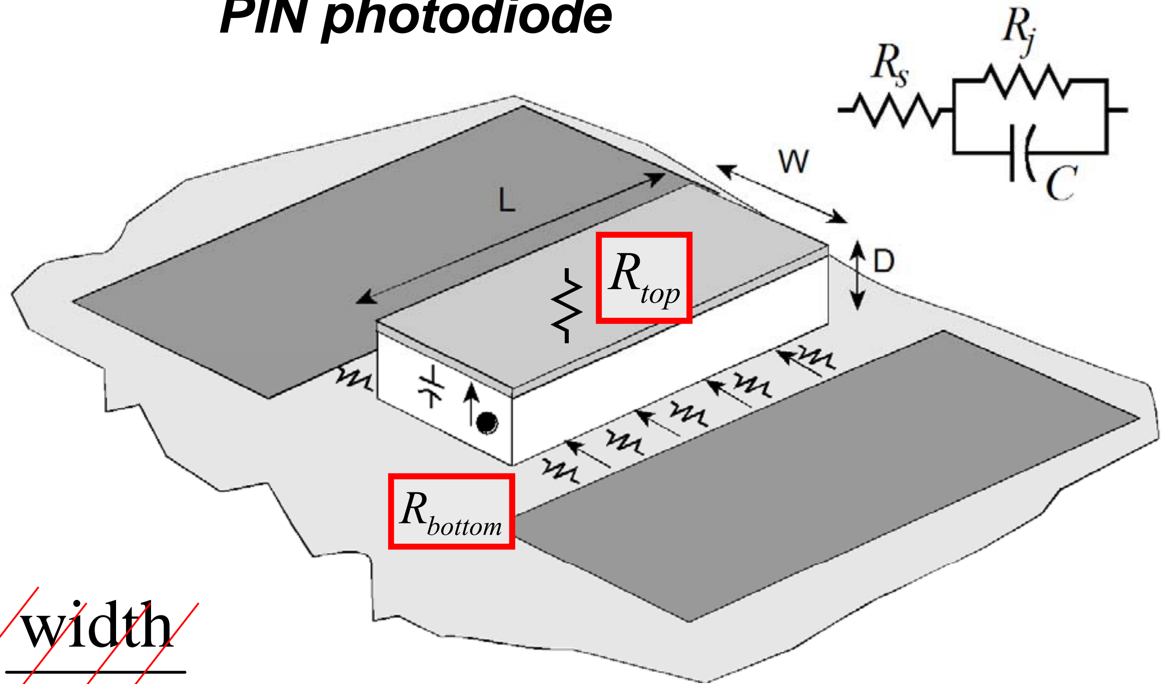
$$C \propto \text{area} / \text{thickness}$$

$$R_{top} \propto \rho_{contact} / \text{area}$$

$$R_{bottom} \propto \frac{\rho_{contact}}{\text{area}} + \frac{\rho_{sheet}}{4} \cdot \frac{\text{width}}{\text{length}}$$

$$I_{\text{max, space-charge-limit}} \propto \text{area} / (\text{thickness})^2$$

$$\Delta T \propto \frac{\text{power}}{\text{length}} \times \log\left(\frac{\text{length}}{\text{width}}\right)$$



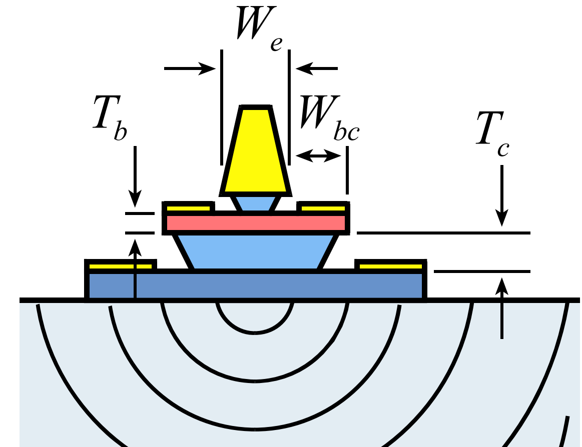
To double bandwidth,

reduce thicknesses 2:1 Improve contacts 4:1

reduce width 4:1, keep constant length

increase current density 4:1

Bipolar Transistor Scaling Laws



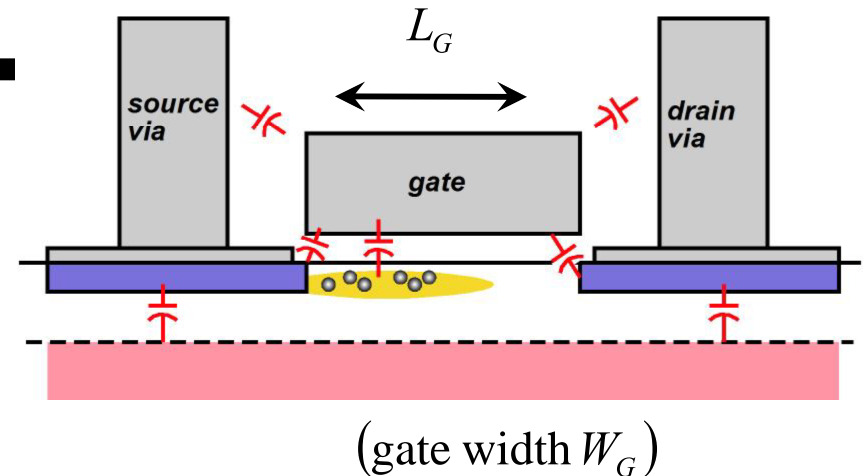
Changes required to double transistor bandwidth:

(emitter length L_E)

parameter	change
collector depletion layer thickness	decrease 2:1
base thickness	decrease 1.414:1
emitter junction width	decrease 4:1
collector junction width	decrease 4:1
emitter contact resistance	decrease 4:1
current density	increase 4:1
base contact resistivity	decrease 4:1

Linewidths scale as the inverse square of bandwidth because thermal constraints dominate.

FET Scaling Laws



$(\text{gate width } W_G)$

Changes required to double transistor bandwidth:

parameter	change
gate length	decrease 2:1
gate dielectric capacitance density	increase 2:1
gate dielectric equivalent thickness	decrease 2:1
channel electron density	increase 2:1
source & drain contact resistance	decrease 4:1
current density (mA/μm)	increase 2:1

Linewidths scale as the inverse of bandwidth because fringing capacitance does not scale.

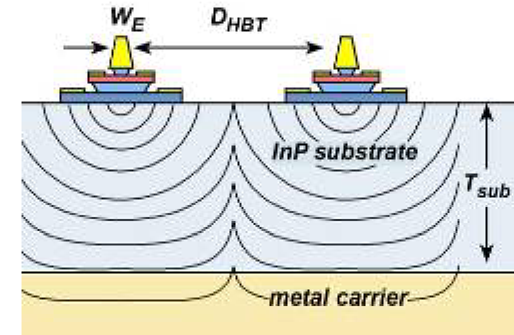
Thermal Resistance Scaling : Transistor, Substrate, Package

$$\Delta T_{\text{substrate}} \cong \underbrace{\frac{P}{\pi K_{InP} L_E} \ln\left(\frac{L_e}{W_e}\right)}_{\text{cylindrical heat flow near junction}} + \underbrace{\frac{P}{\pi K_{InP}}\left(\frac{1}{L_E} - \frac{1}{D}\right)}_{\text{spherical flow for } r > L_e} + \underbrace{\frac{P}{K_{InP}} \cdot \left(\frac{T_{sub} - D/2}{D^2}\right)}_{\text{planar flow for } r > D_{HBT}/2}$$

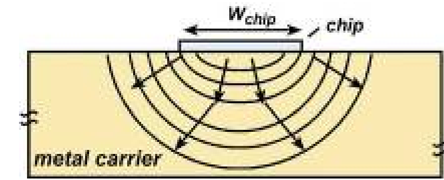
cylindrical heat flow near junction: increases logarithmically

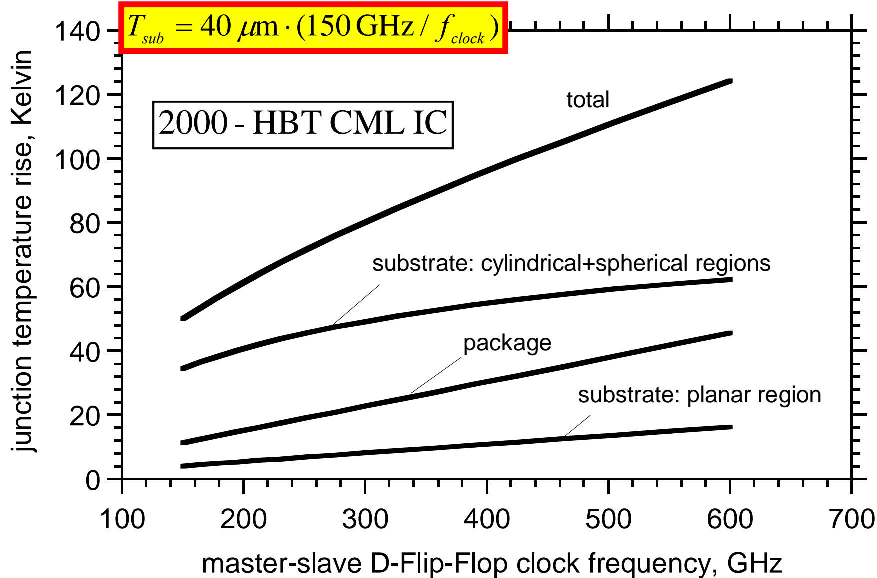
spherical flow for $r > L_e$: insignificant variation

planar flow for $r > D_{HBT}/2$: increases quadratically if T_{sub} is constant



$$\Delta T_{package} \cong \left(\frac{1}{4} + \frac{1}{\pi}\right) \frac{P_{chip}}{K_{Cu} W_{chip}}$$





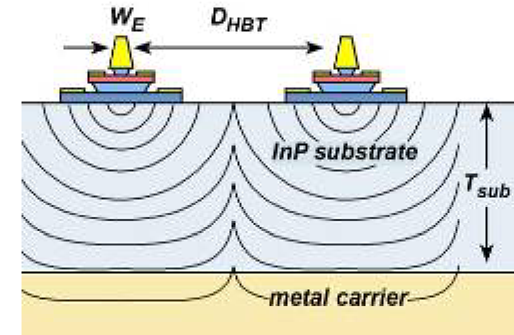
Wiring lenghts scale as 1/bandwidth. Power density, scales as $(\text{bandwidth})^2$.

Thermal Resistance Scaling : Transistor, Substrate, Package

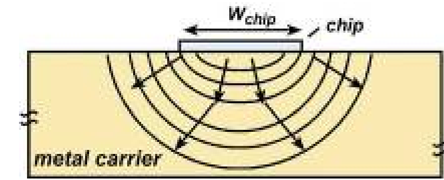
cylindrical heat flow near junction — spherical flow for $r > L_e$ — planar flow for $r > D_{HBT}/2$

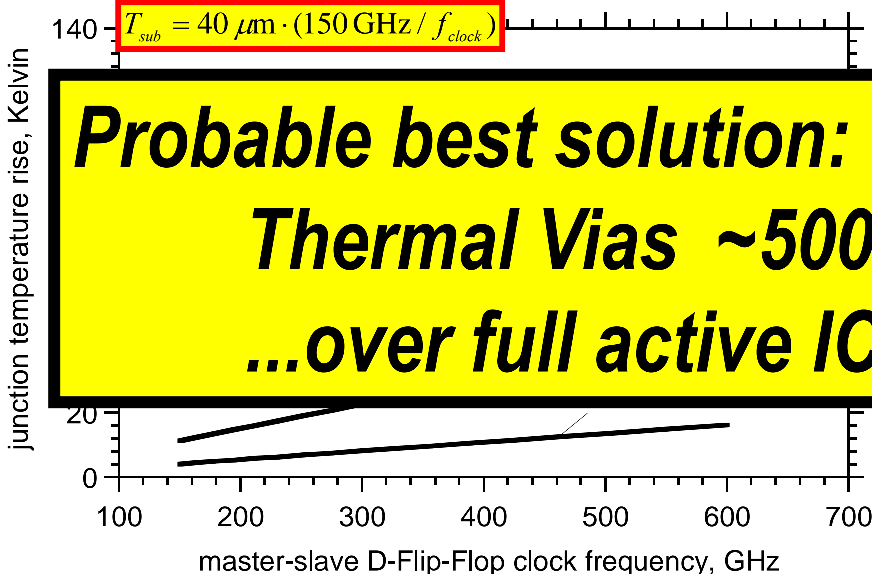
$$\Delta T_{\text{substrate}} \cong \frac{P}{\pi K_{InP} L_E} \ln\left(\frac{L_e}{W_e}\right) + \frac{P}{\pi K_{InP}}\left(\frac{1}{L_E} - \frac{1}{D}\right) + \frac{P}{K_{InP}} \cdot \left(\frac{T_{sub} - D/2}{D^2}\right)$$

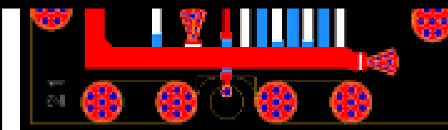
increases logarithmically — insignificant variation — increases quadratically if T_{sub} is constant



$$\Delta T_{package} \cong \left(\frac{1}{4} + \frac{1}{\pi}\right) \frac{P_{chip}}{K_{Cu} W_{chip}}$$



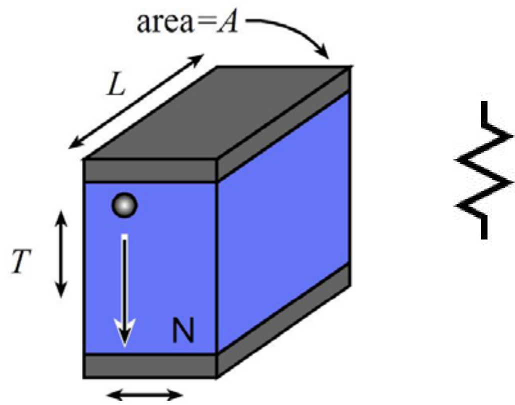




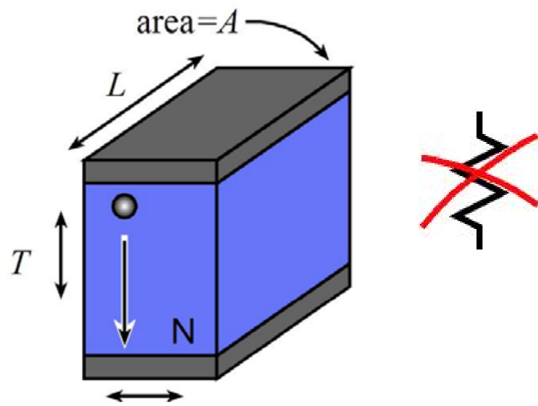
(bandwidth)^2.

Probable best solution:
Thermal Vias ~500 nm below InP subcollector
...over full active IC area.

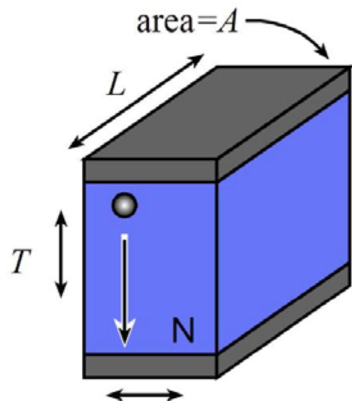
Electron Plasma Resonance: Not a Dominant Limit

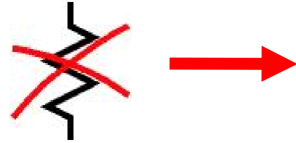


Electron Plasma Resonance: Not a Dominant Limit



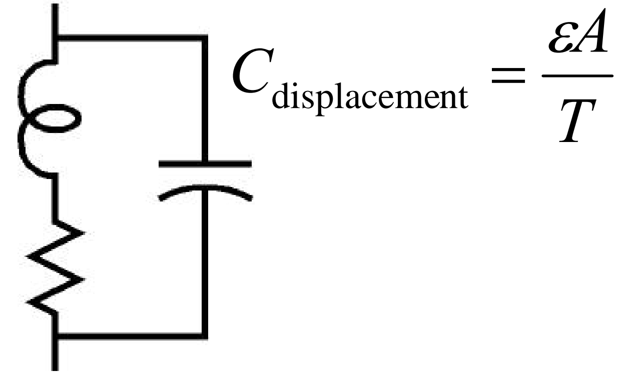
Electron Plasma Resonance: Not a Dominant Limit





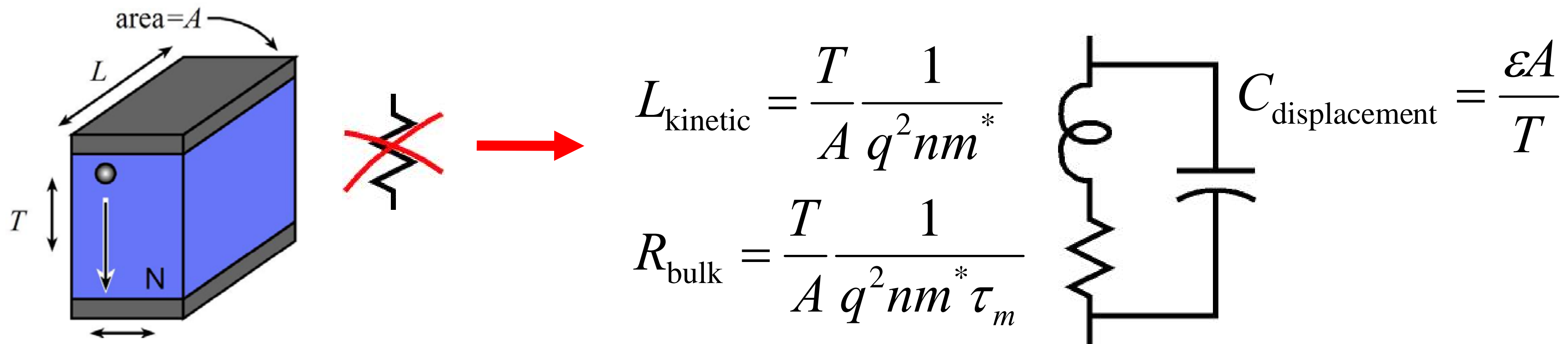
$$L_{\text{kinetic}} = \frac{T}{A}\frac{1}{q^2 n m^*}$$

$$R_{\text{bulk}} = \frac{T}{A}\frac{1}{q^2 n m^* \tau_m}$$



$$C_{\text{displacement}} = \frac{\varepsilon A}{T}$$

Electron Plasma Resonance: Not a Dominant Limit



	dielectric relaxation frequency	scattering frequency	plasma frequency
	$f_{dielecic} = \dfrac{1/2\pi}{C_{\text{displacement}} R_{\text{bulk}}} = \dfrac{1}{2\pi}\dfrac{\sigma}{\varepsilon}$	$f_{scattering} = \dfrac{1}{2\pi}\dfrac{R_{\text{bulk}}}{L_{\text{kinetic}}} = \dfrac{1}{2\pi\tau_m}$	$f_{plasma} = \dfrac{1/2\pi}{\sqrt{L_{\text{kinetic}} C_{\text{displacement}}}}$
n - InGaAs $3.5 \cdot 10^{19} / \text{cm}^3$	800 THz	7 THz	74 THz
p - InGaAs $7 \cdot 10^{19} / \text{cm}^3$	80 THz	12 THz	31 THz

THz & nm Transistors: it's all about the interfaces

Metal-semiconductor interfaces (Ohmic contacts):
__very low resistivity__

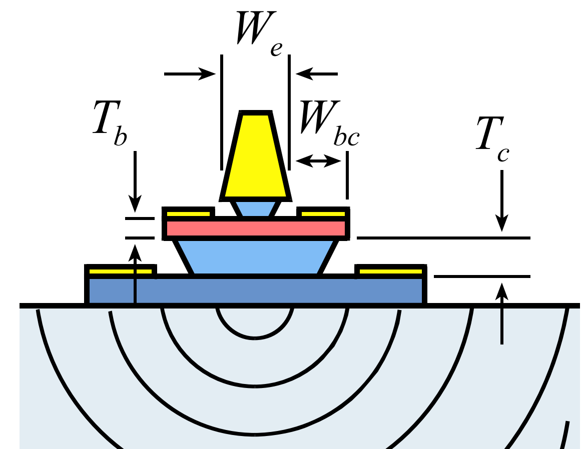
Dielectric-semiconductor interfaces (Gate dielectrics):
__very high capacitance density__

Transistor & IC thermal resistivity.

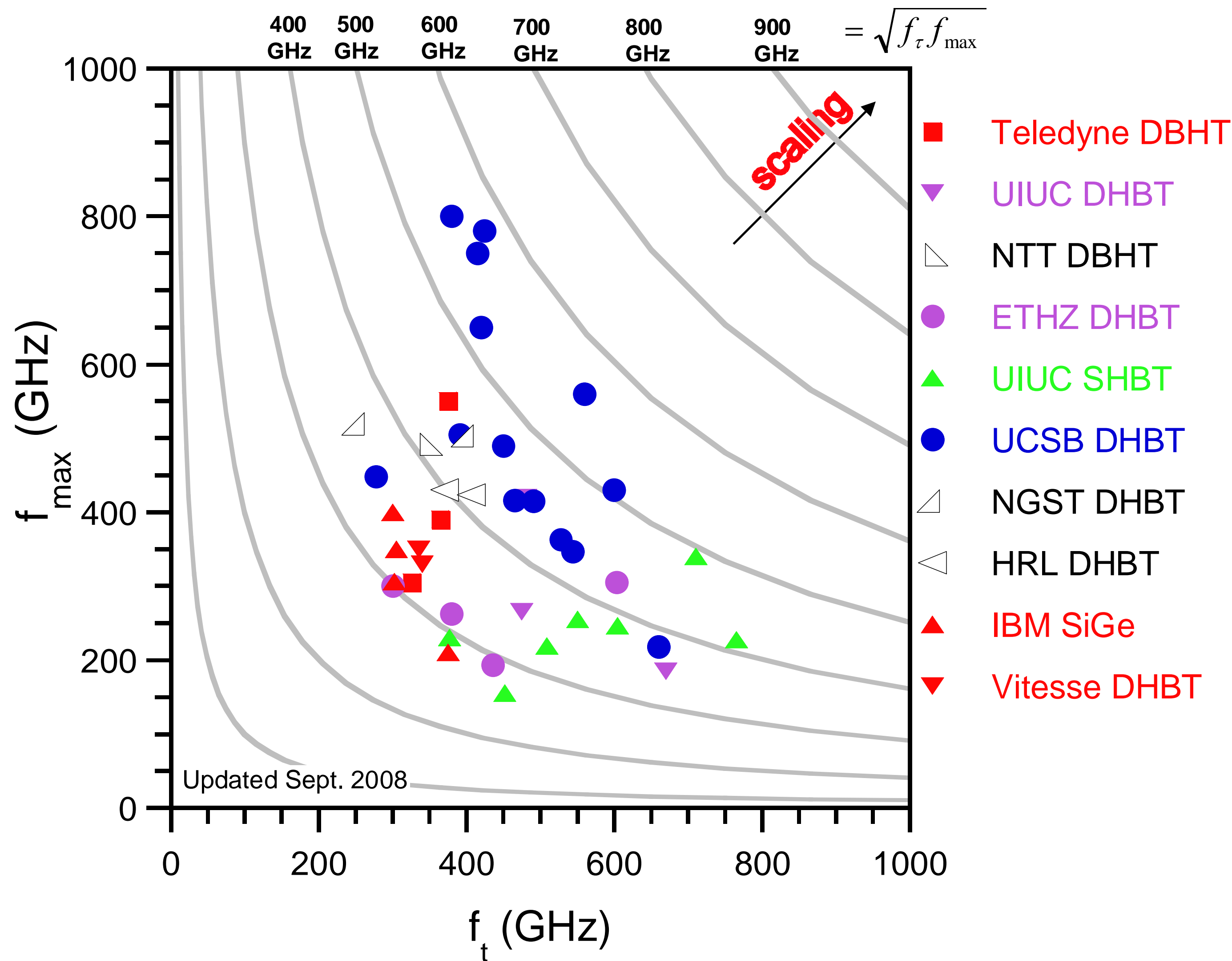
THz
Bipolar Transistors

InP Bipolar Transistor Scaling Roadmap

	industry	university →industry	university 2007-9	appears feasible	maybe
emitter	512	256	128	64	32 nm width
	16	8	4	2	1 $\Omega\cdot\mu m^2$ access ρ
base	300	175	120	60	30 nm contact width,
	20	10	5	2.5	1.25 $\Omega\cdot\mu m^2$ contact ρ
collector	150	106	75	53	37.5 nm thick,
	4.5	9	18	36	72 $mA/\mu m^2$ current density
	4.9	4	3.3	2.75	2-2.5 V, breakdown
f_τ	370	520	730	1000	1400 GHz
f_{max}	490	850	1300	2000	2800 GHz
power amplifiers	245	430	660	1000	1400 GHz
digital 2:1 divider	150	240	330	480	660 GHz



InP DHBTs: September 2008



popular metrics :

f_τ or f_{max} alone

$(f_\tau + f_{max})/2$

$\sqrt{f_\tau f_{max}}$

$(1/f_\tau + 1/f_{max})^{-1}$

much better metrics :

power amplifiers:

PAE, associated gain,

mW/μm

low noise amplifiers:

F_{min}, associated gain,

digital:

f_{clock}, hence

$(C_{cb}\Delta V / I_c)$,

$(R_{ex} I_c / \Delta V)$,

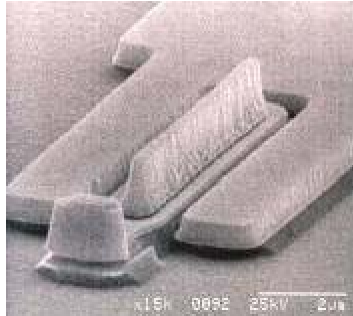
$(R_{bb} I_c / \Delta V)$,

$(\tau_b + \tau_c)$

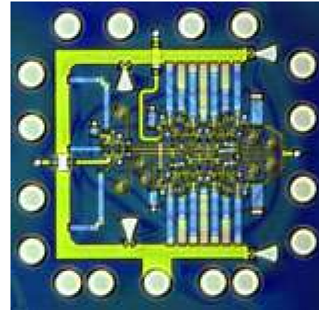
512 nm InP DHBT

Laboratory Technology

500 nm mesa HBT

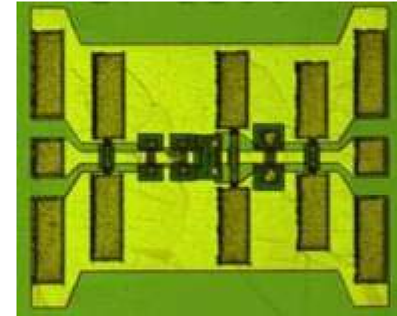


150 GHz M/S latches



UCSB / Teledyne / GCS

175 GHz amplifiers

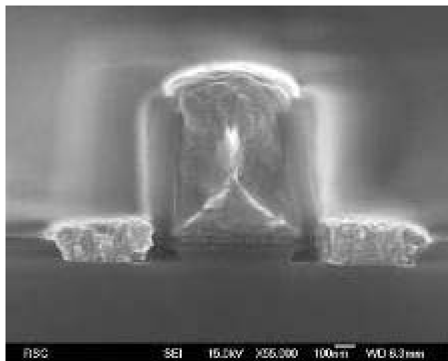


UCSB

Production

(Teledyne)

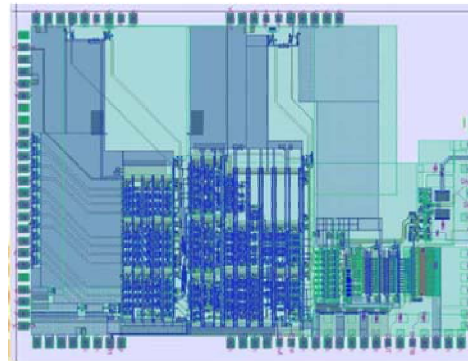
500 nm sidewall HBT



Teledyne

$f_\tau = 405$ GHz
$f_{max} = 392$ GHz
$V_{br,\,ceo} = 4$ V

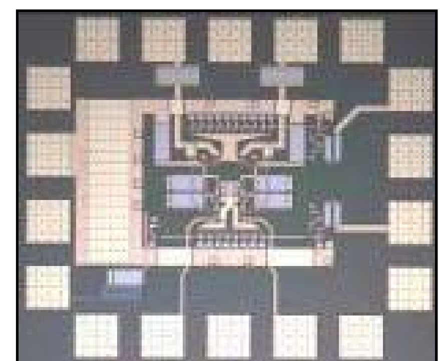
DDS IC: 4500 HBTs



Teledyne / BAE

20 GHz clock

20-40 GHz op-amps

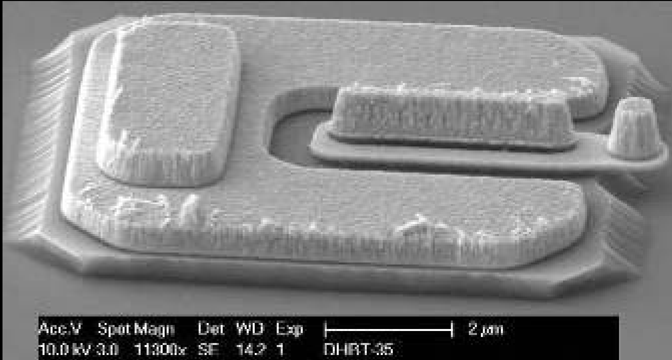


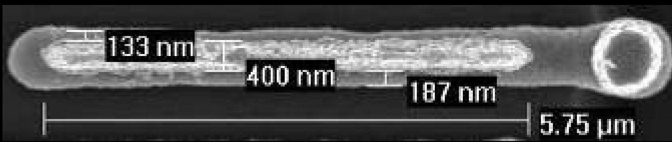
Teledyne / UCSB

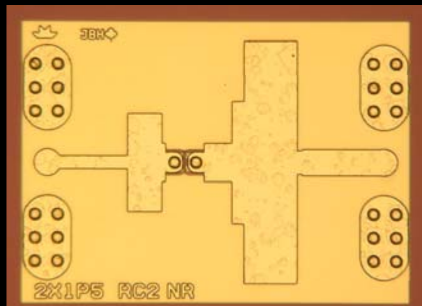
53-56 dBm OIP3 @ 2 GHz with 1 W dissipation

Z. Griffith
M. Urteaga
P. Rowell
D. Pierson
B. Brar
V. Paidi

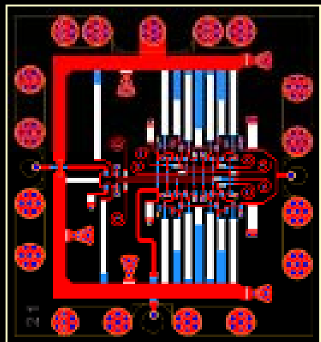
256 nm Generation InP DHBT







324 GHz Amplifier

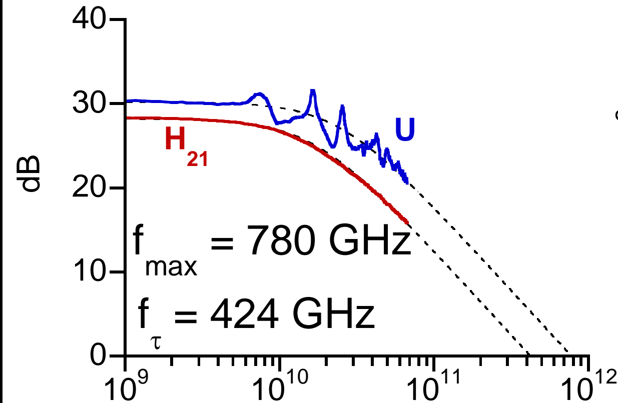


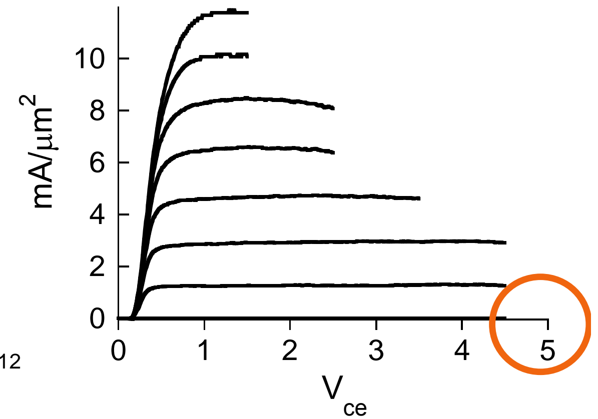
200 GHz master-slave latch design

Z. Griffith, E. Lind

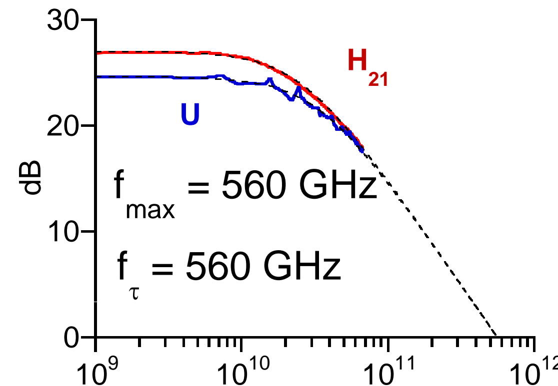
J. Hacker, M. Jones

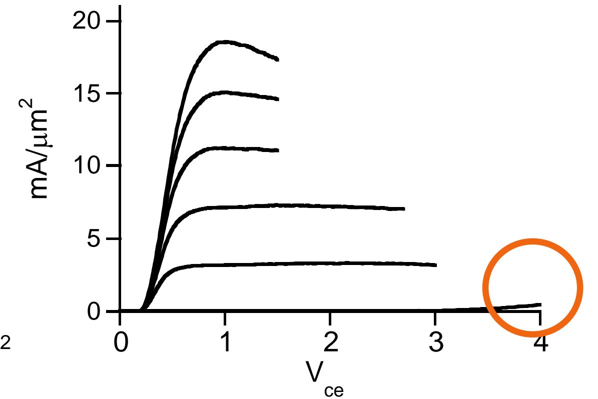
150 nm thick collector



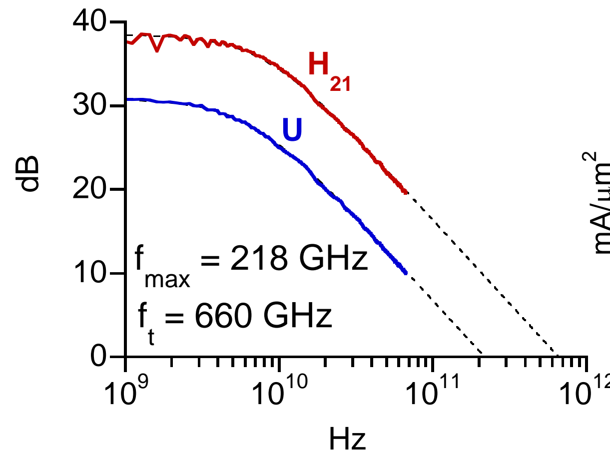


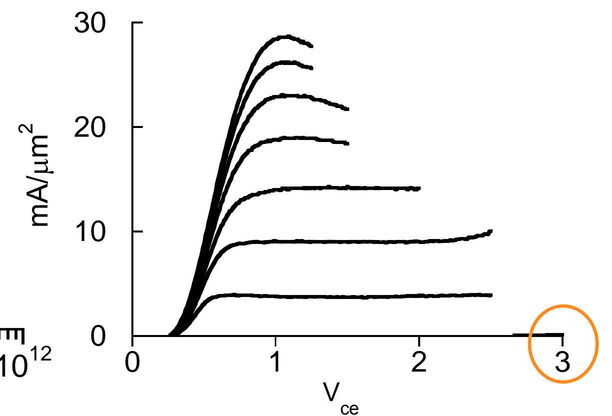
70 nm thick collector





60 nm thick collector





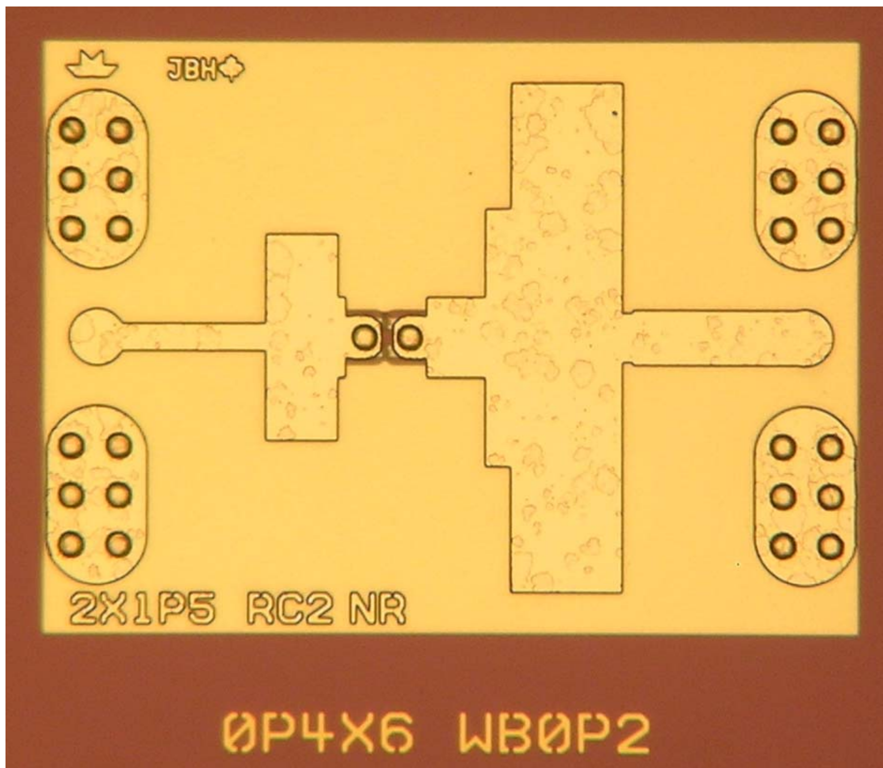
324 GHz Medium Power Amplifiers in 256 nm HBT

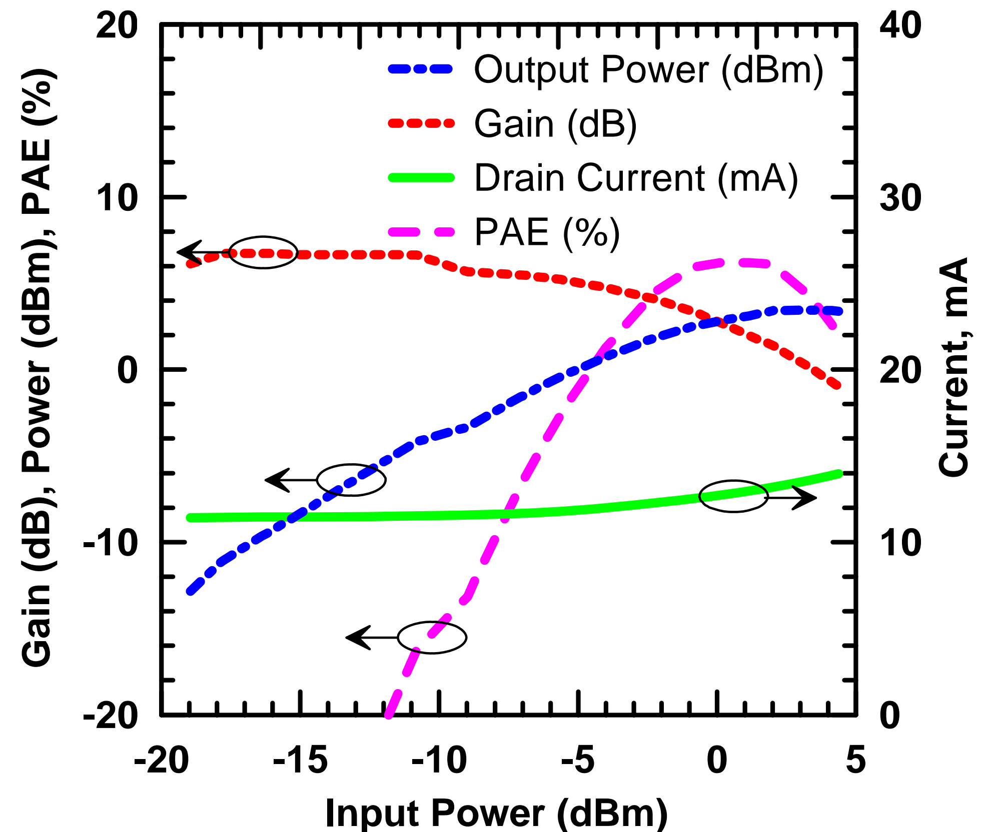
ICs designed by Jon Hacker / Teledyne

Teledyne 256 nm process flow-

Hacker et al, 2008 IEEE MTT-S

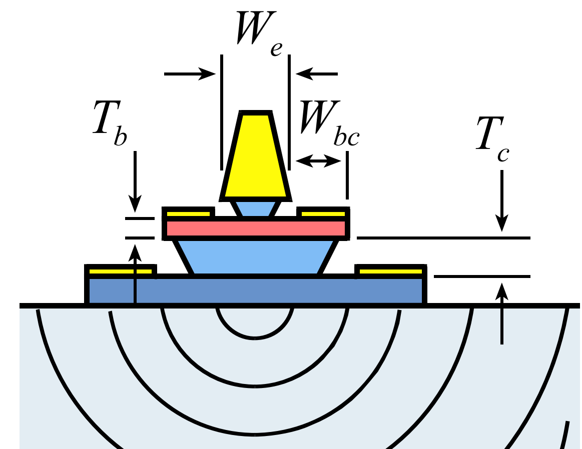
~2 mW saturated output power





InP Bipolar Transistor Scaling Roadmap

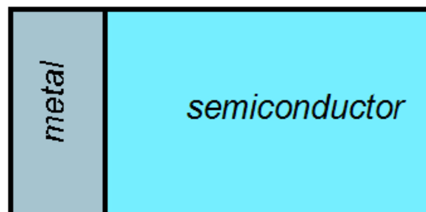
	industry	university →industry	university 2007-9	appears feasible	maybe
emitter	512	256	128	64	32 nm width
	16	8	4	2	1 $\Omega\cdot\mu m^2$ access ρ
base	300	175	120	60	30 nm contact width,
	20	10	5	2.5	1.25 $\Omega\cdot\mu m^2$ contact ρ
collector	150	106	75	53	37.5 nm thick,
	4.5	9	18	36	72 $mA/\mu m^2$ current density
	4.9	4	3.3	2.75	2-2.5 V, breakdown
f_τ	370	520	730	1000	1400 GHz
f_{max}	490	850	1300	2000	2800 GHz
power amplifiers	245	430	660	1000	1400 GHz
digital 2:1 divider	150	240	330	480	660 GHz

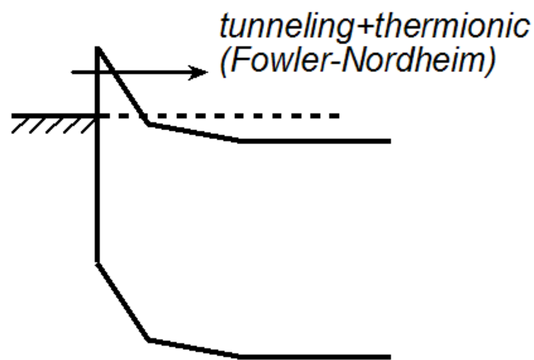


Conventional ex-situ contacts are a mess

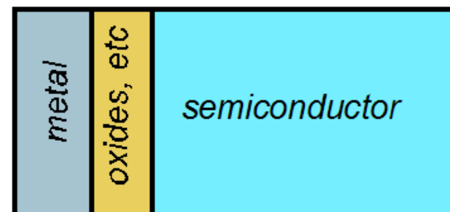
THz transistor bandwidths: very low-resistivity contacts are required

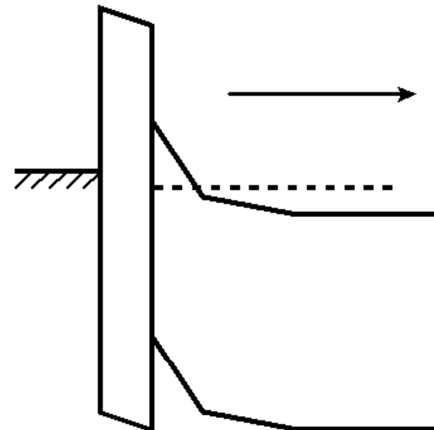
textbook contact



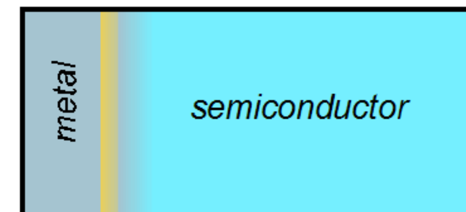


with surface oxide





with metal penetration



Interface barrier → resistance

Further intermixing during high-current operation → degradation

Ohmic Contacts Good Enough for 3 THz Transistors

64 nm (2.0 THz) HBT needs ~ 2 Ω - μm^2 contact resistivities

32 nm (2.8 THz) HBT needs ~ 1 Ω - μm^2

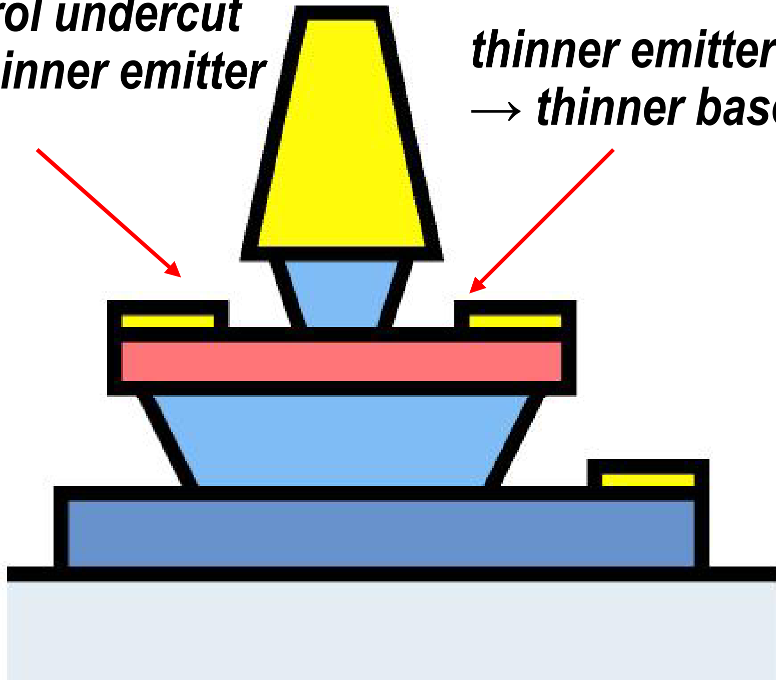
Contacts to N-InGaAs:*		
Mo	*MBE in-situ*	*2.2 (+/- 0.5) Ω - μm^2*
TiW	*ex-situ*	*~2.2 Ω - μm^2*

Contacts to P-InGaAs:		
Mo	*MBE in-situ*	*below 2.5 Ω - μm^2*
Pd/...	*ex-situ*	*~1 (+/- ?) Ω - μm^2*

**measured emitter resistance remains higher than that of contacts.*

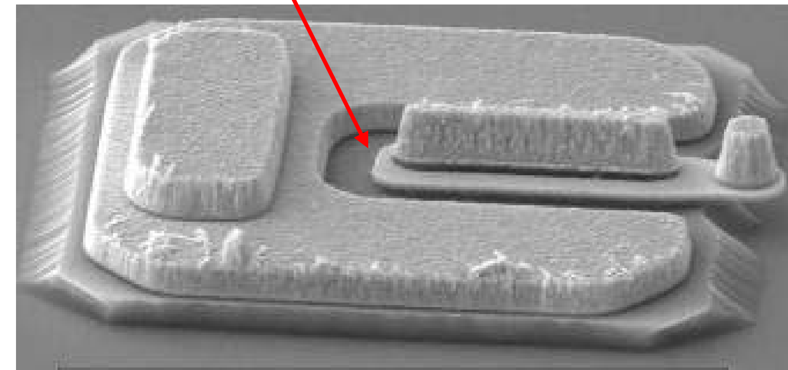
Process Must Change Greatly for 128 / 64 / 32 nm Nodes

control undercut
→ thinner emitter



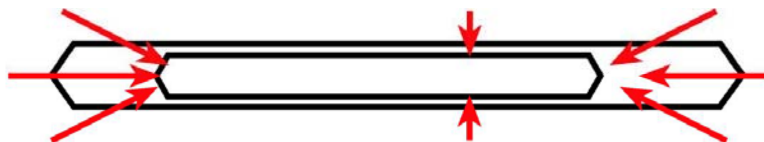
thinner emitter
→ thinner base metal

thinner base metal
→ excess base metal resistance

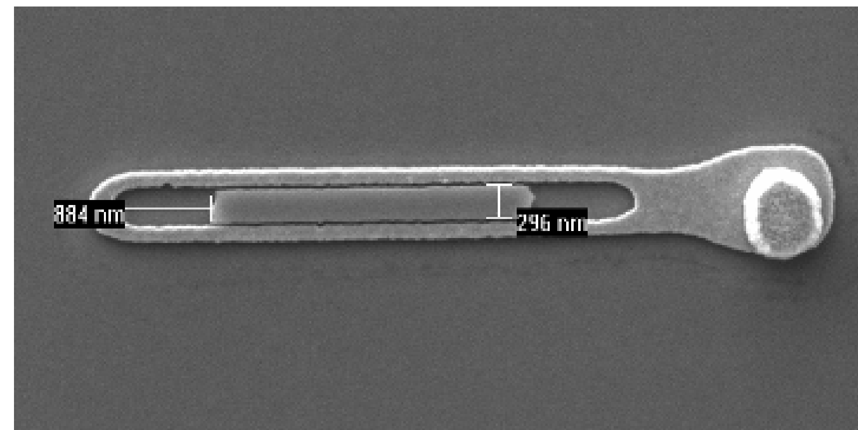


Undercutting of emitter ends

{101}A planes: fast



{111}A planes: slow



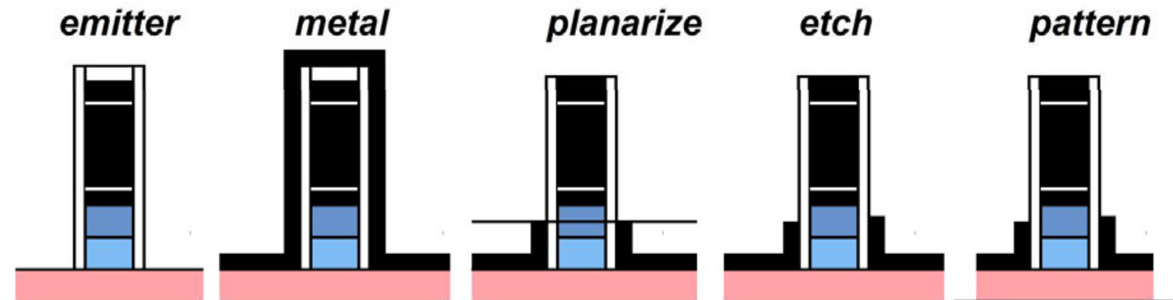
128 / 64 nm process: where we are going

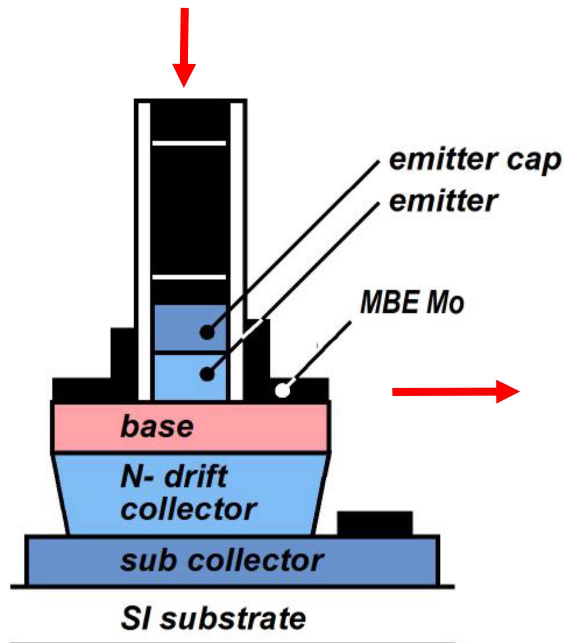
Developing scalable self-aligned base process

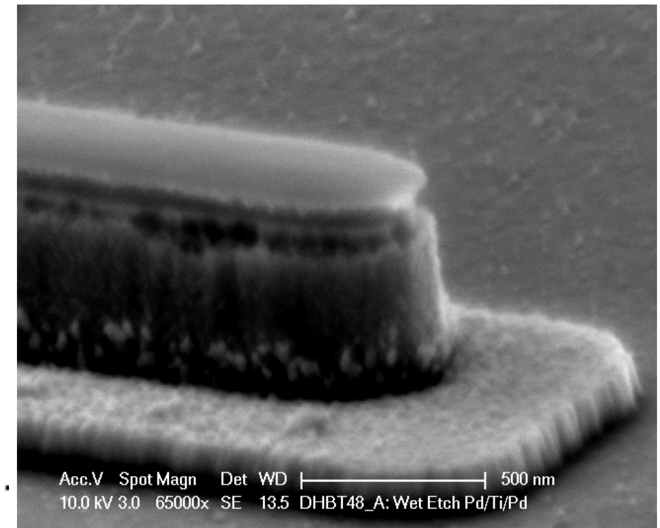
0.3 Ω-μm^2 resistivity emitter contacts
- *in-situ, in MBE*

2 Ω-μm^2 resistivity base contacts
- *in-situ, in MBE*

target ~2000 GHz device



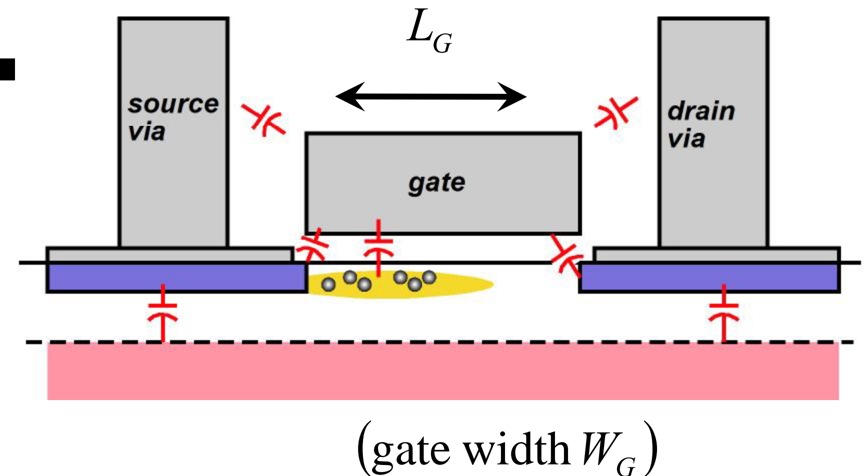




THz
Field-Effect Transistors

(THz HEMTs)

FET Scaling Laws



$(\text{gate width } W_G)$

Changes required to double transistor bandwidth:

parameter	change
gate length	decrease 2:1
gate dielectric capacitance density	increase 2:1
gate dielectric equivalent thickness	decrease 2:1
channel electron density	increase 2:1
source & drain contact resistance	decrease 4:1
current density (mA/μm)	increase 2:1

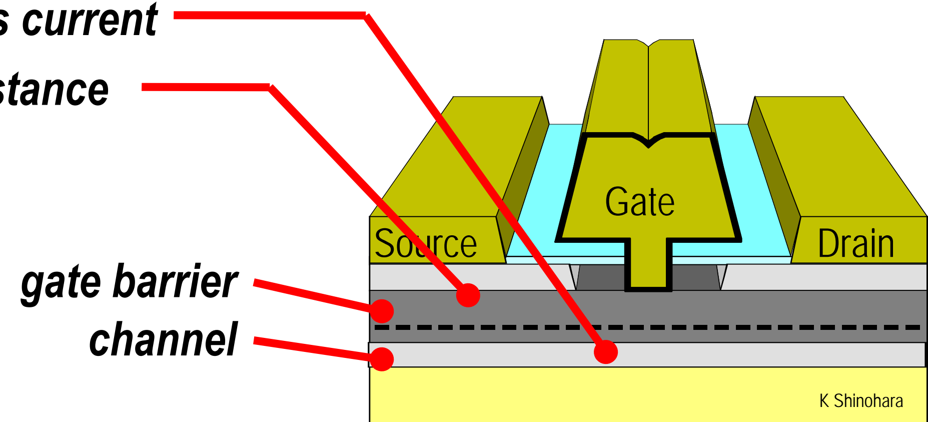
InGaAs HEMTs are best for mm-wave low-noise receivers...
...but there are difficulties in improving them further.

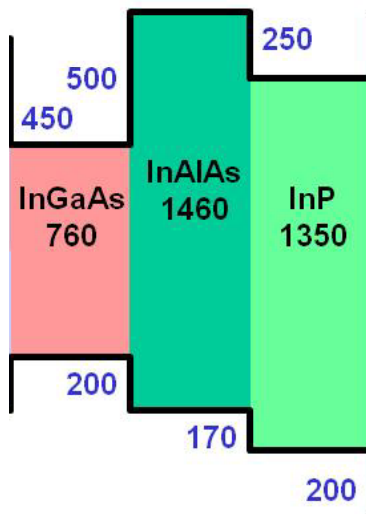
Why HEMTs are Hard to Improve

1st challenge with HEMTs: reducing access resistance

low electron density under gate recess→ limits current

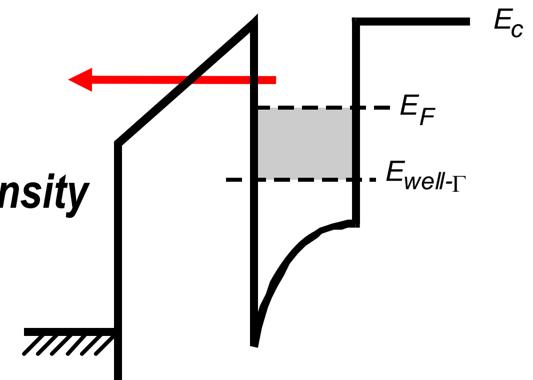
gate barrier lies under S/D contacts → resistance





2nd challenge with HEMTs: low gate barrier

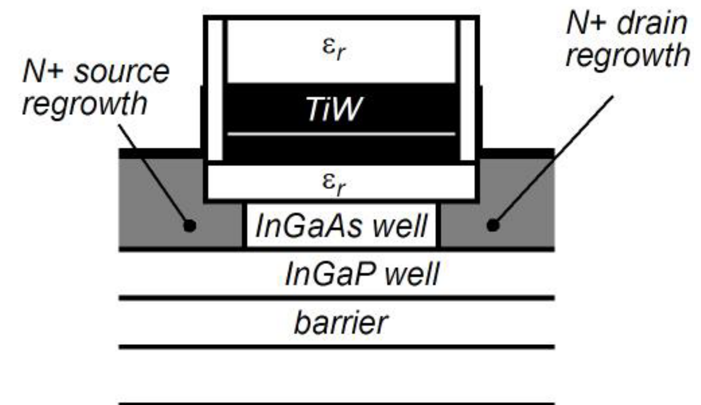
high tunneling currents with thin barrier
high emission currents with high electron density



III-V MOSFETs do not face these scaling challenges

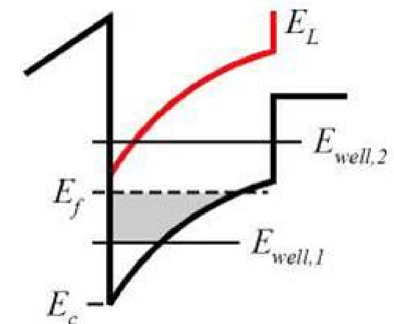
III-V MOSFETs for VLSI→ Also Helps HEMT Development

What is it ?
MOSFET with an InGaAs channel



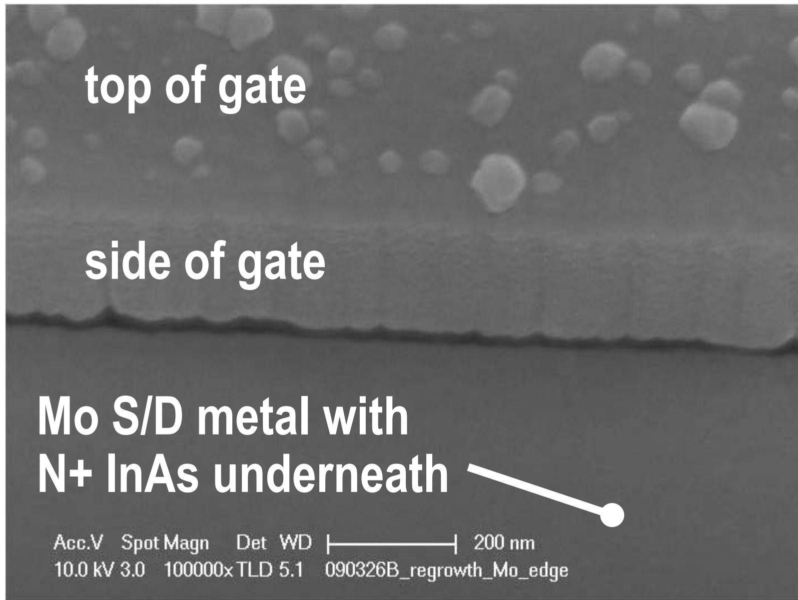
Why do it ?
low electron effective mass→ higher electron velocity
more current, less charge at a given insulator thickness & gate length
very low access resistance

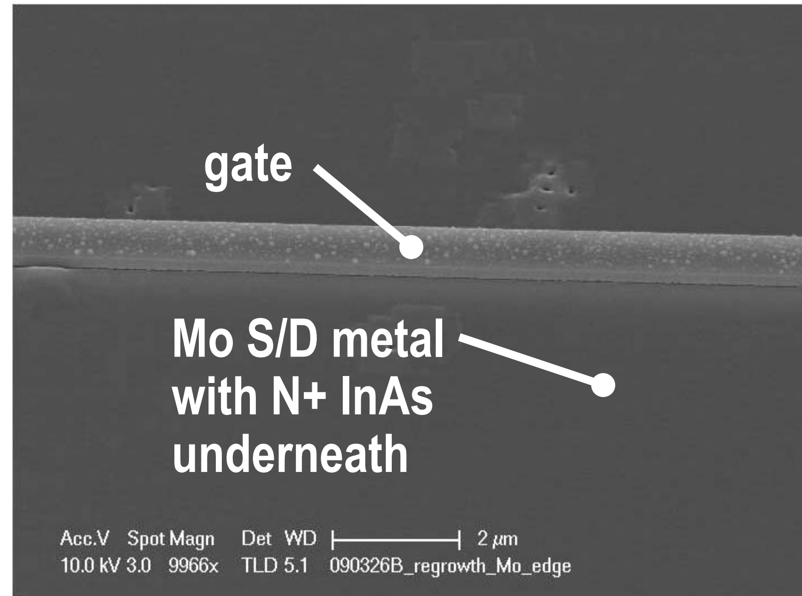
What are the problems ?
low electron effective mass→ constraints on scaling !
must grow high-K on InGaAs, must grow InGaAs on Si

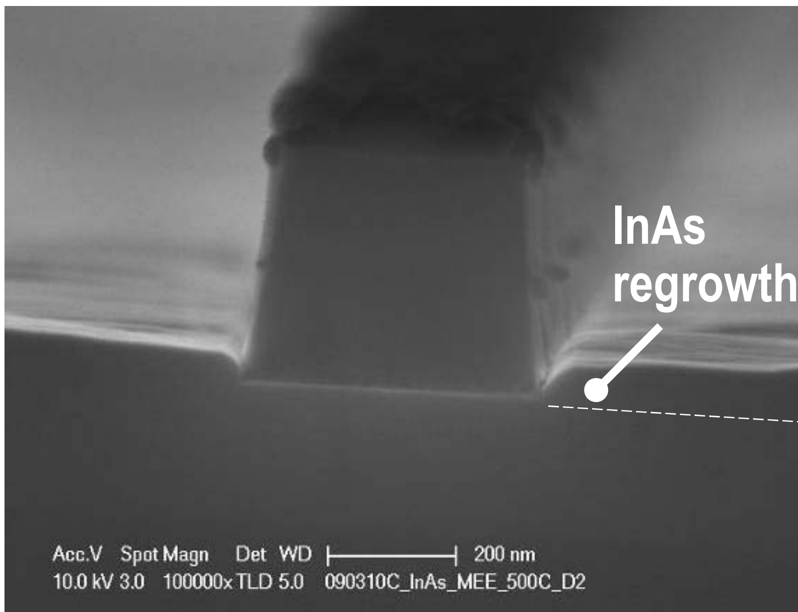


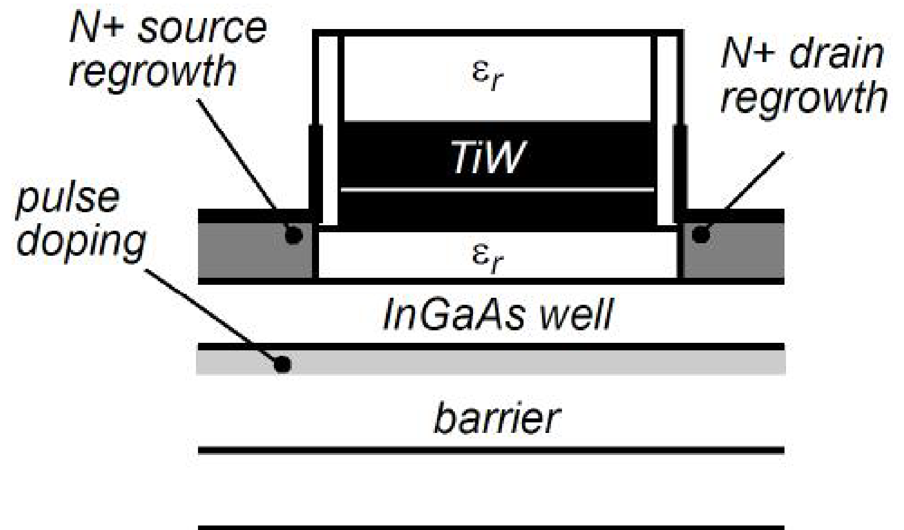
III-V MOSFETs in Development

Wistey
Singisetti
Burek
Lee
Rodwell
Gossard

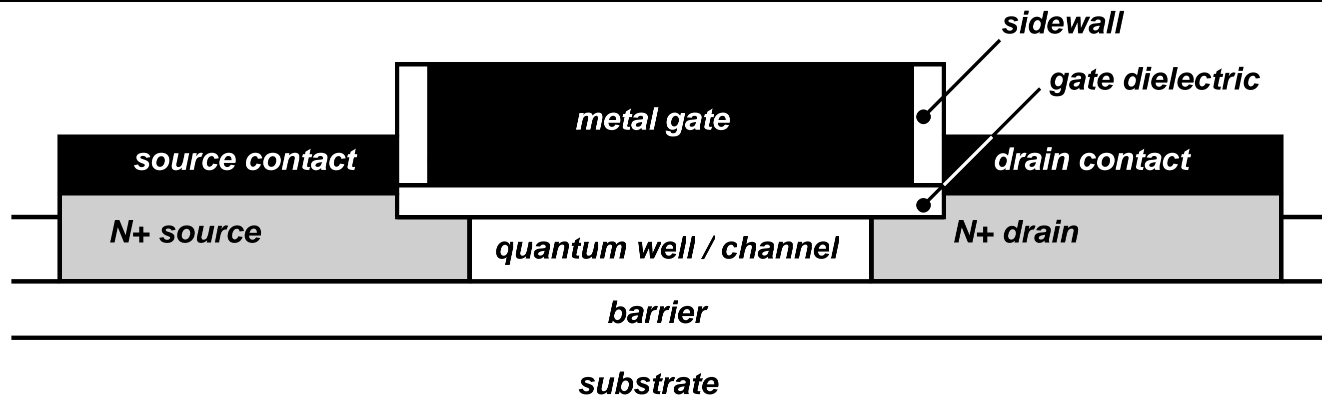








III-V MOSFETs as a Scaling Path for THz HEMTs



Why ?

Much lower access resistance in S/D regions

Higher gate barrier→ higher feasible electron density in channel

Higher gate barrier→ gate dielectric can be made thinner

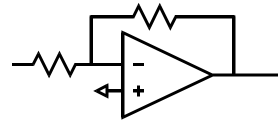
Estimated Performance (?)

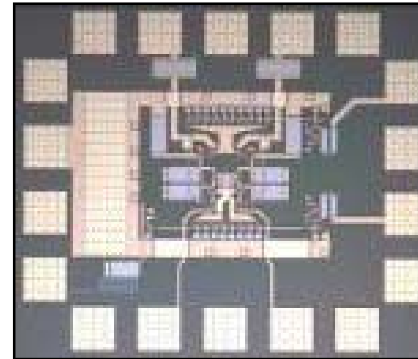
2 THz cutoff frequencies at 32 nm gate length

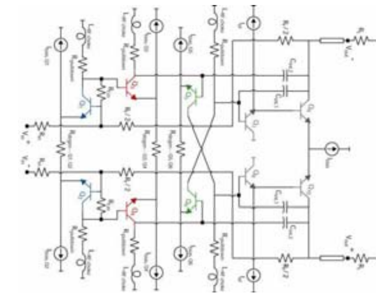
Applications of THz III-V Transistors

What Else Would You Do With a THz Transistor ?

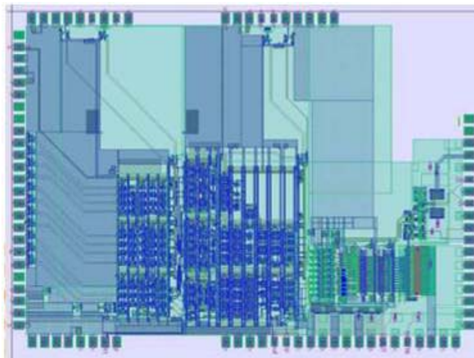
precision, high-performance analog microwave circuits

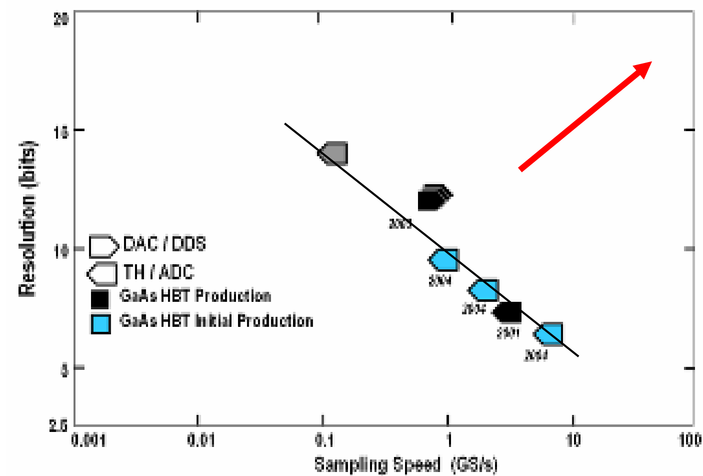






higher-resolution microwave ADCs, DACs, DDSs

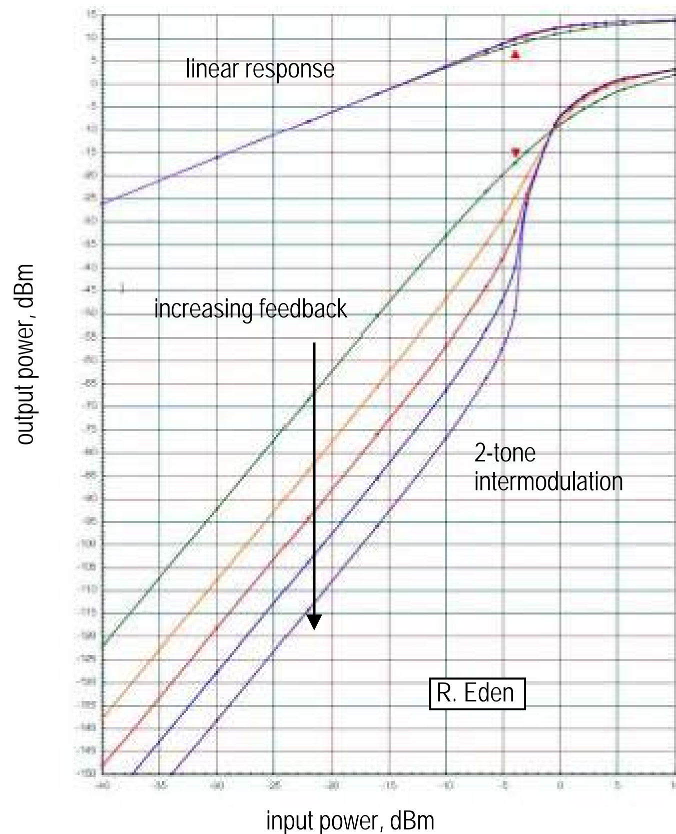


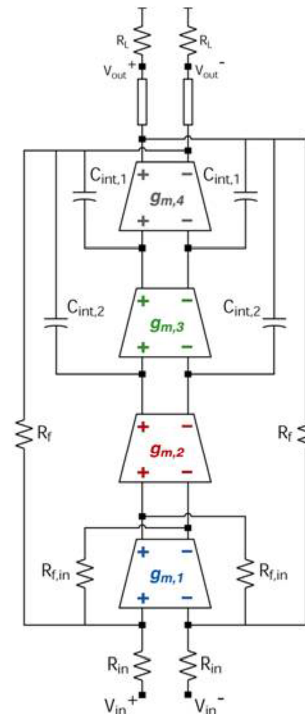


mm-wave Op-Amps for Linear Microwave Amplification

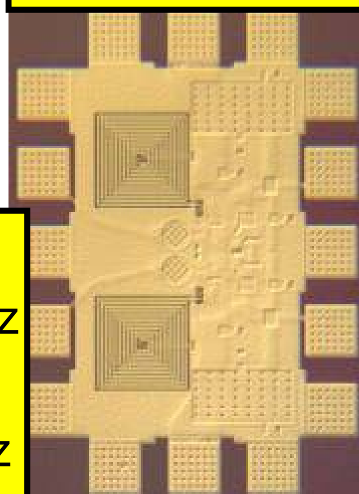
DARPA / UCSB / Teledyne FLARE: Griffith & Urteaga

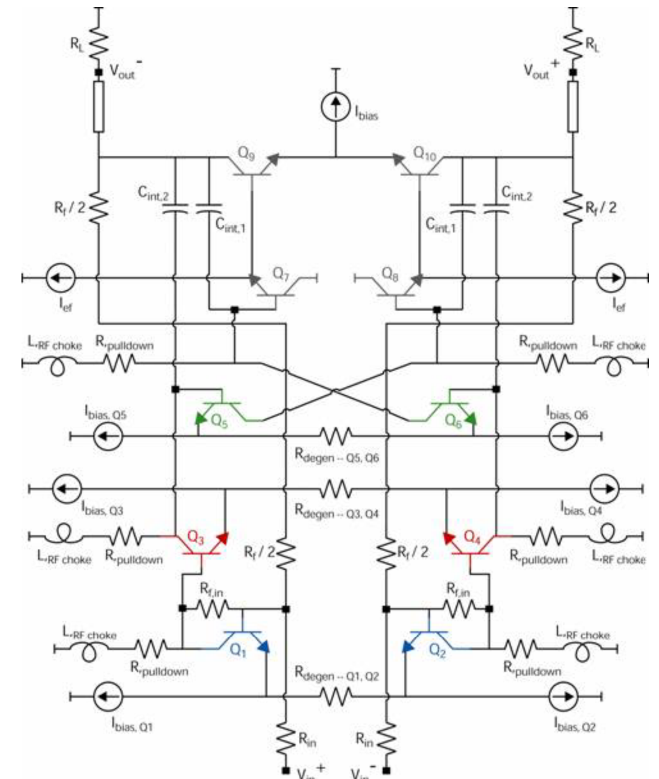
Reduce distortion with strong negative feedback

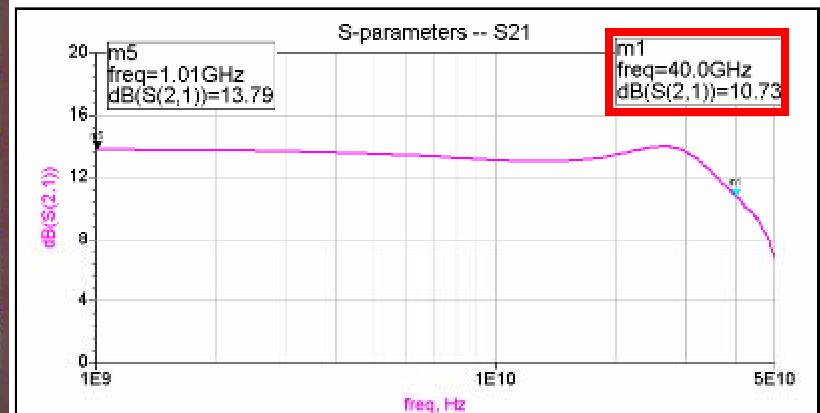




300 GHz / 4 V InP HBT





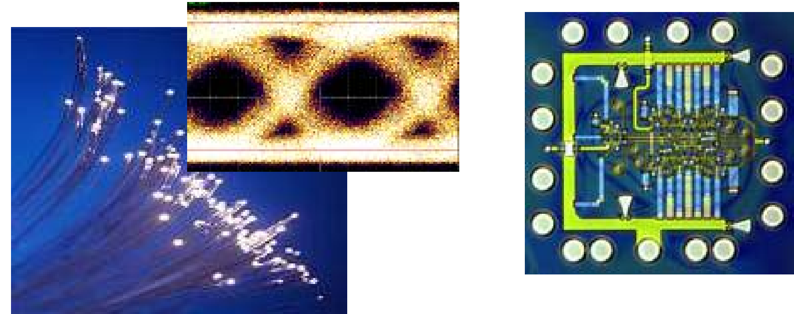


measured 20-40 GHz bandwidth
measured **54 dBm OIP3** @ 2 GHz

new designs in fabrication
simulated **56 dBm OIP3** @ 2 GHz

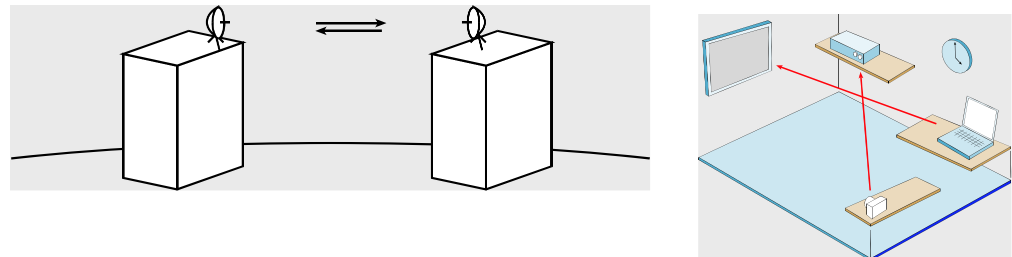
What Would You Do With a THz Transistor ?

640 Gb/s ETDM
optical fiber links



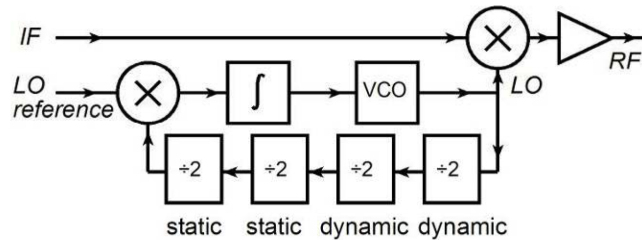
300-1000 GHz
imaging systems

mm-wave
communication links

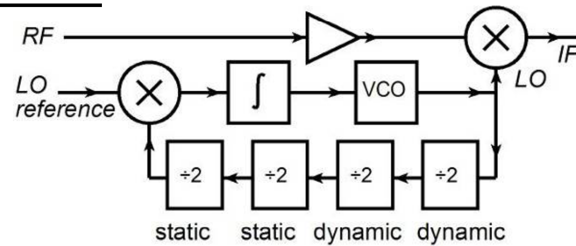


670 GHz Transceiver Simulations in 128 nm InP HBT

transmitter exciter



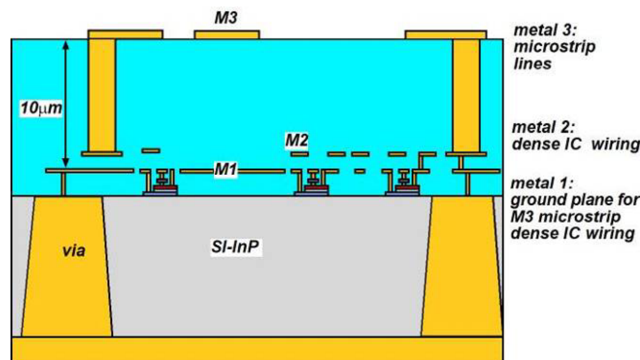
receiver



3-layer thin-film THz interconnects

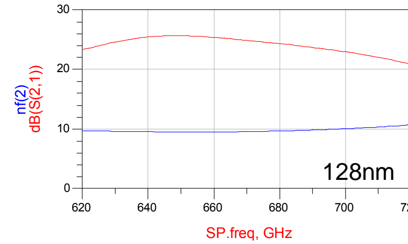
thick-substrate--> high-Q TMIC

thin -> high-density digital

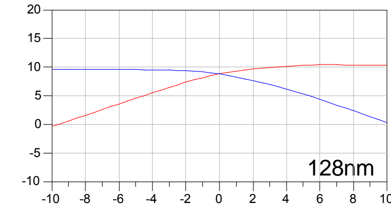


Simulations @ 670 GHz (128 nm HBT)

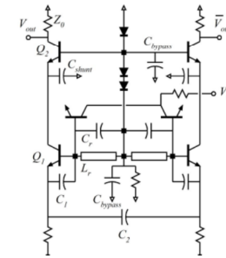
LNA: 9.5 dB Fmin at 670 GHz

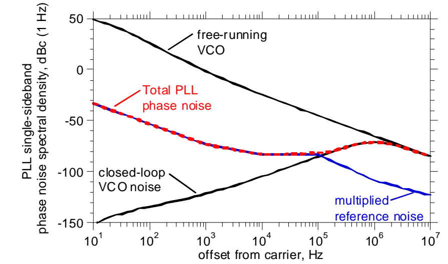


PA: 9.1 dBm Pout at 670 GHz

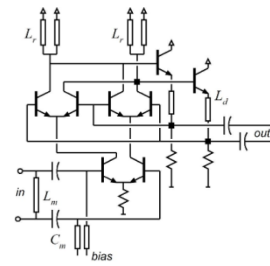


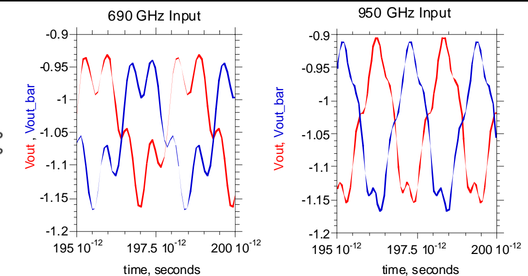
VCO:
-50 dBc (1 Hz)
@ 100 Hz offset
at 620 GHz (phase 1)



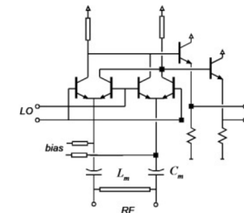


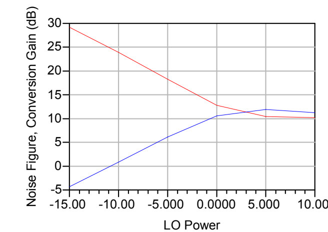
Dynamic divider:
novel design,
simulates to 950 GHz





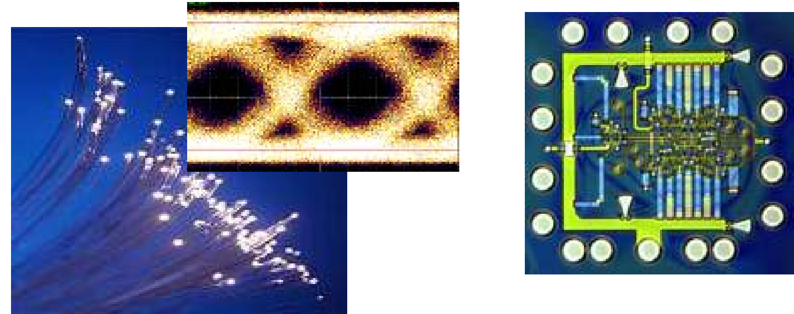
Mixer:
10.4 dB noise figure
11.9 dB gain





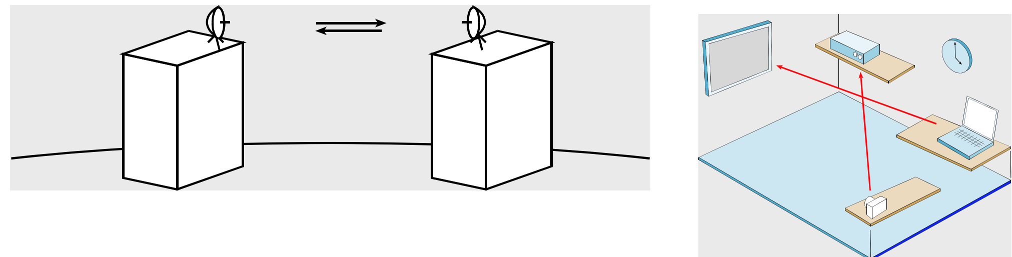
What Would You Do With a THz Transistor ?

640 Gb/s ETDM
optical fiber links

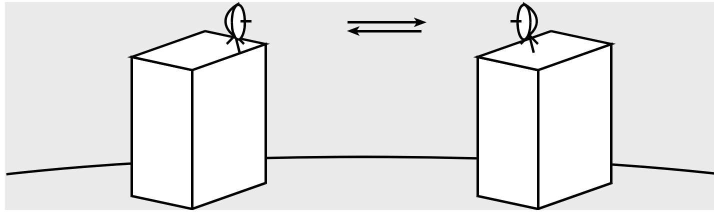


300-1000 GHz
imaging systems

mm-wave
communication links



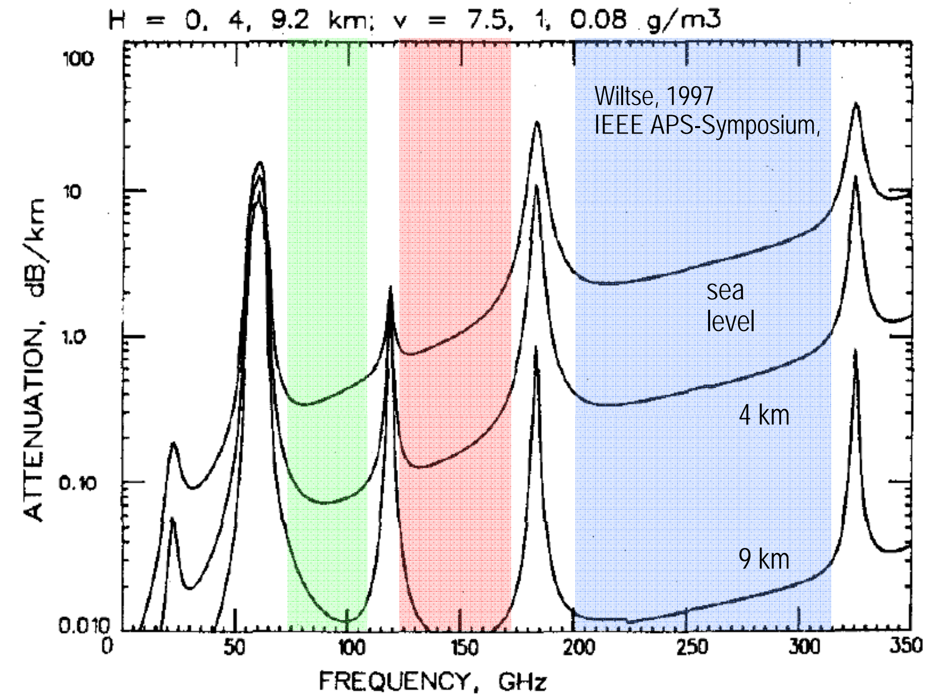
150 & 250 GHz Bands for 100 Gb/s Radio ?



$$P_{received} / P_{trans} = (D_t D_r / 16\pi^2)(\lambda / R)^2$$

$$P_{received\,(4QPSK)} = Q^2 \cdot kTFB\,; \quad Q \cong 6$$

$$D = 4\pi A_{eff} \big/ \lambda^2$$



125-150 GHz, 200-300 GHz:
enough bandwidth for 100 Gb/s QPSK

<u>150 GHz carrier, 100 Gbs/s QPSK radio:</u>
30 cm antennas, 10 dBm power, fair weather→ 1 km range

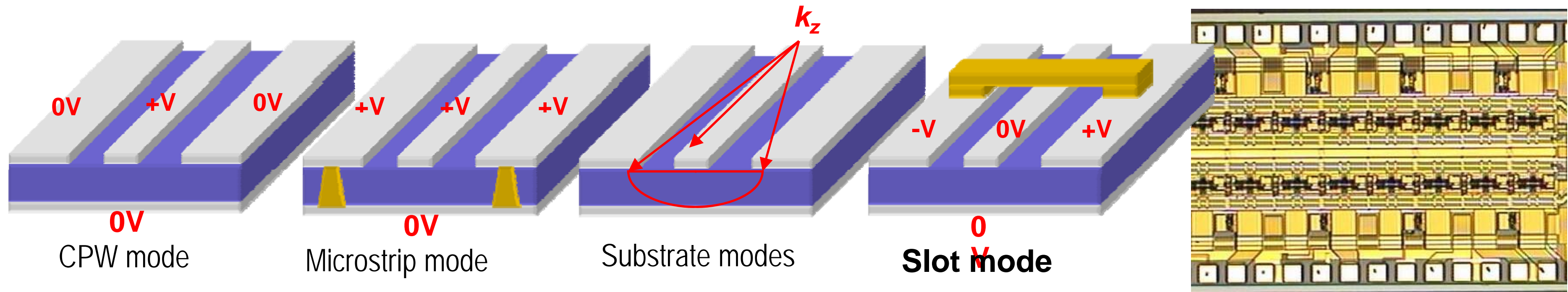
150 GHz band: Expect ~10-20 dB/km attenuation for rain
300 GHz band: expect ~20-30 db/km from 90% humidity

Interconnects

within

high-speed ICs

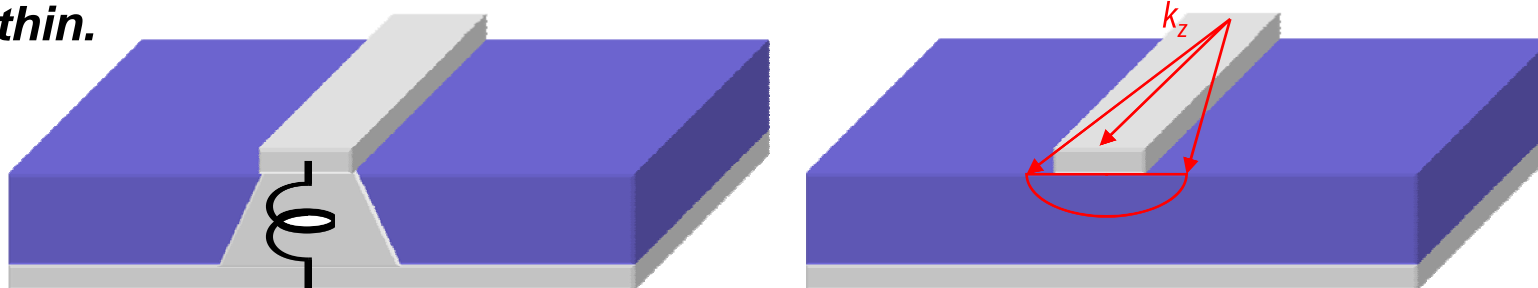
CPW has parasitic modes, coupling from poor ground plane integrity



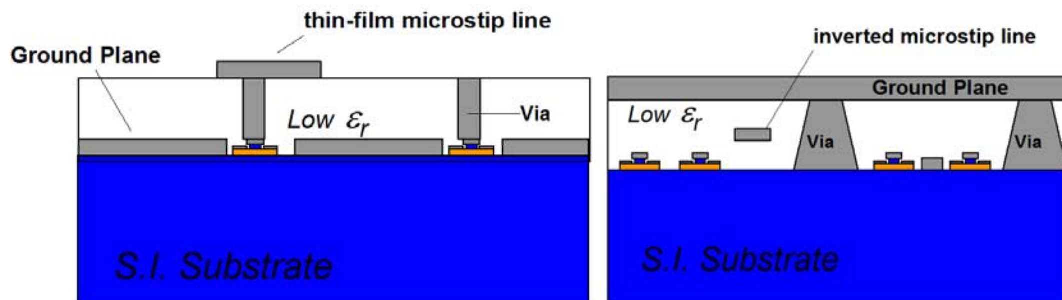
ground straps suppress slot mode, but multiple ground breaks in complex ICs produce ground return inductance
ground vias suppress microstrip mode, wafer thinning suppresses substrate modes

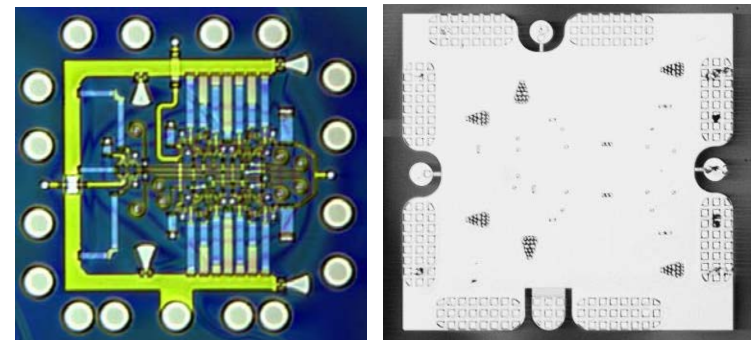
---

Microstrip has high via inductance, has mode coupling unless substrate is thin.



We prefer (credit to NTT) thin-film microstrip wiring, inverted is best for complex ICs



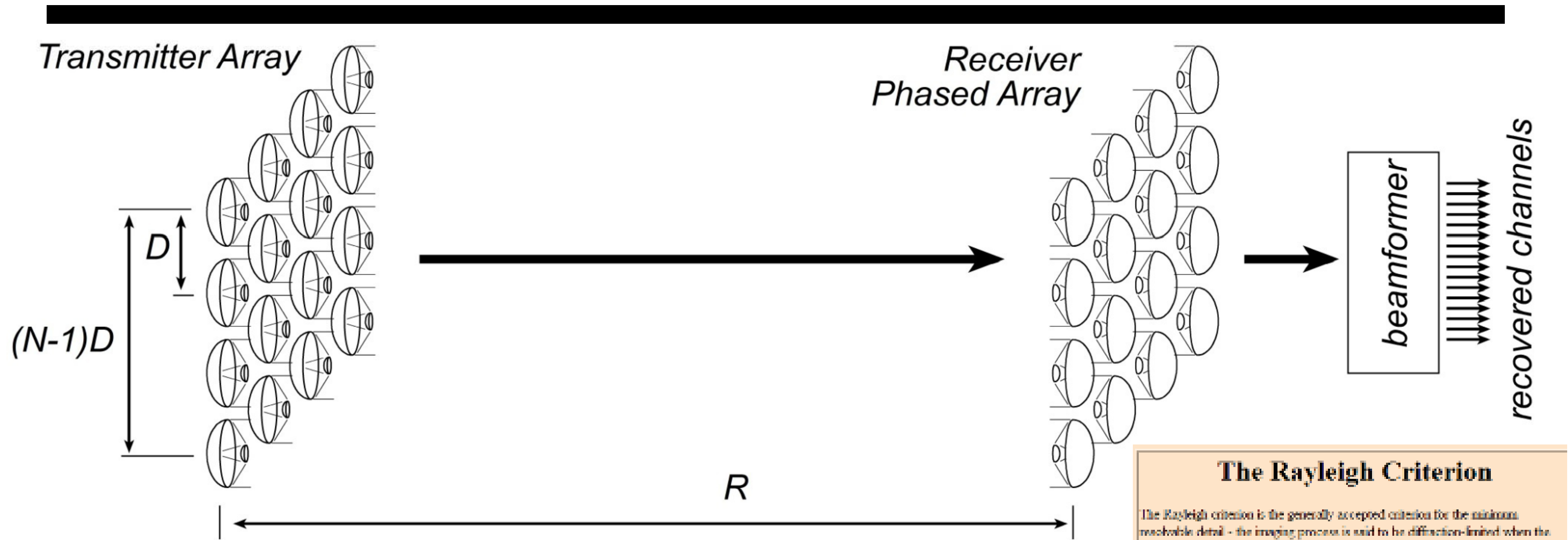


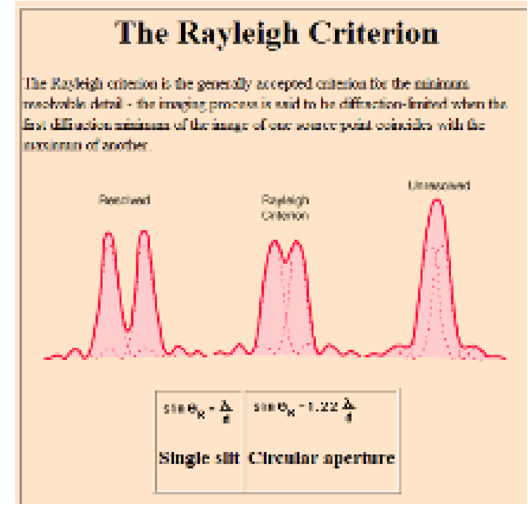
M. Urteaga, Z. Griffith, S. Krishnan

mm-wave MIMO:

→ Wireless Links at 100's of Gb/s

mm-wave (60-80 GHz) MIMO → wireless at 40+ Gb/s rates ?





Rayleigh Criterion:

Spatial angular separation of adjacent transmitters: $\delta\theta_t = D/R$

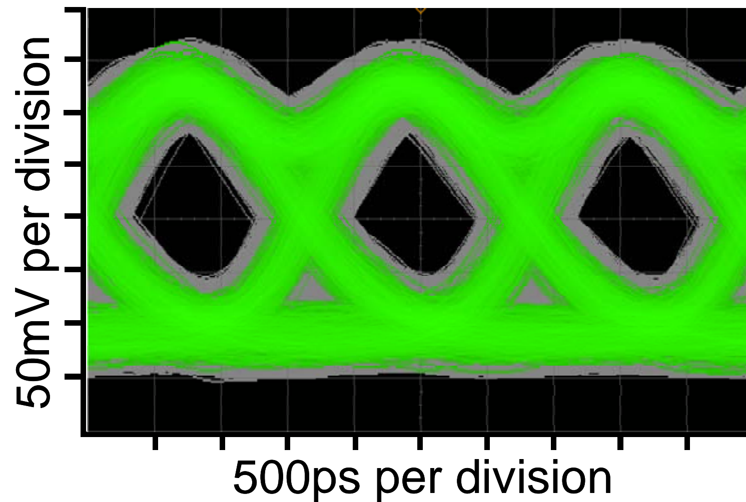
Receive array angular resolution: $\delta\theta_r = \lambda/(N-1)D$

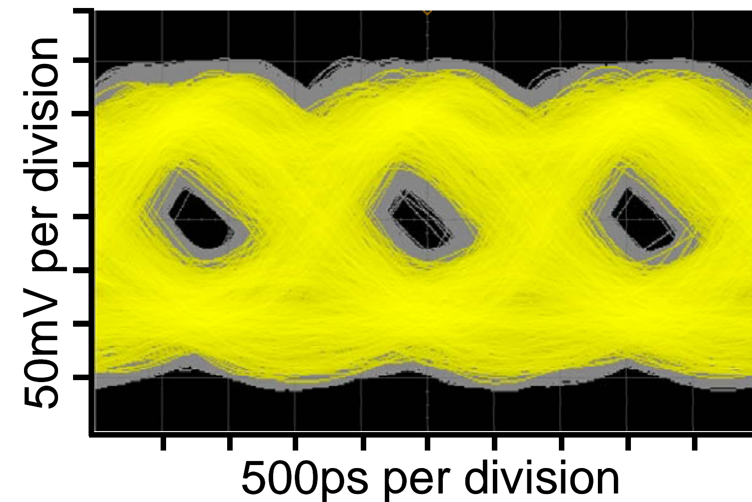
To resolve adjacent channels, $\delta\theta_r \leq \delta\theta_r \Rightarrow (N-1)D = \sqrt{\lambda R(N-1)}$

70 GHz, 1 km, 16 elements, 2 polarizations, 3.6 x 3.6 meter array, 2.5 GBaud QPSK
→ 160 Gb/s digital radio ?

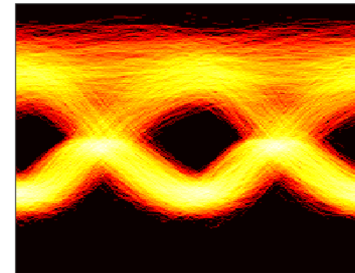
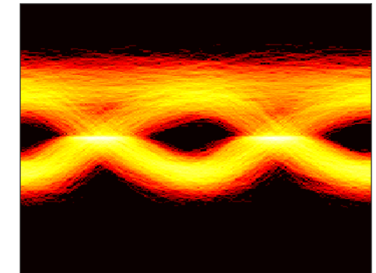
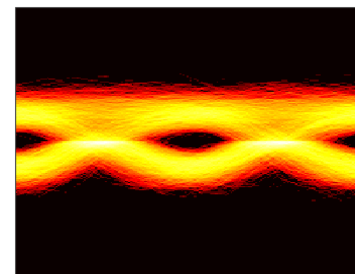
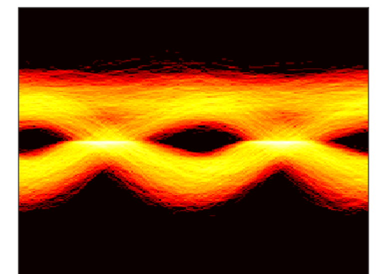
mm-wave MIMO: 2-channel prototype, 60 GHz, 40 meters

	Channel Number	1	2
BER	Single Active Transmitter	$<10^{-6}$	$<10^{-6}$
	Two Active Transmitters	$<10^{-6}$	1.8×10^{-6}
Channel Suppression Ratio (dB)	10Mbps per Channel	27.8	28.8
	600Mbps per Channel	21.1	9.7





mm-wave MIMO: 4-channel prototype, 60 GHz, Indoor

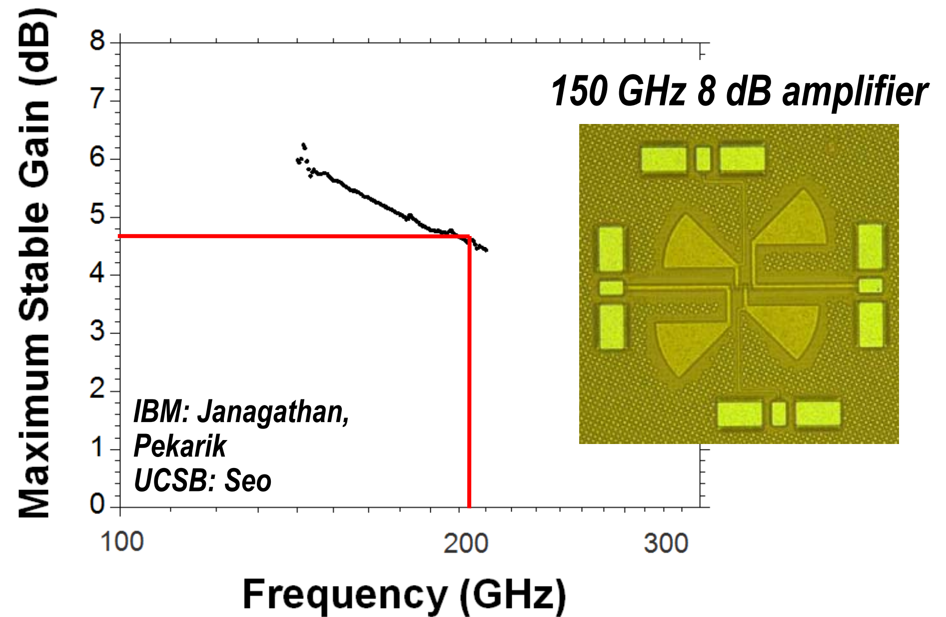
4 x 622 Mb/s
60 GHz carrier

VSLI for
mm-wave &
sub-mm-wave systems

Billions of 700-GHz Transistors→ Imaging & Arrays

65 nm CMOS: ~5 dB gain @ 200 GHz

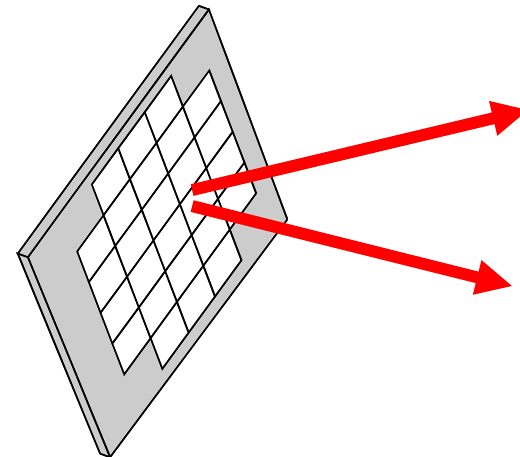
22 nm will be much faster yet.



What can you do with a few billion 700-GHz transistors ?

Build Transmitter / Receiver Arrays

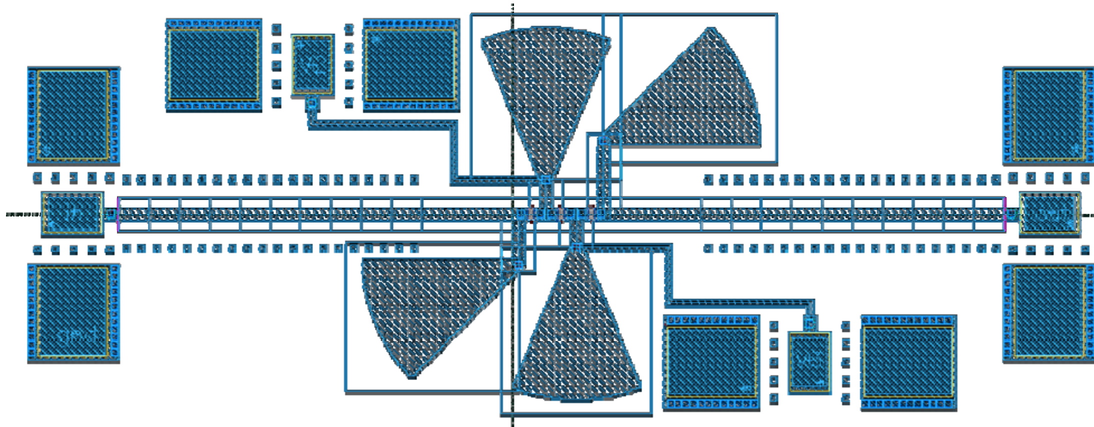
100's or 1000's of transmitters or receivers
...on < 1 cm^2 IC area
...operating at 100-500 GHz.



3-stage 250GHz Amplifier Design

IBM: Pekarik, Jaganathan
UCSB: M. Seo

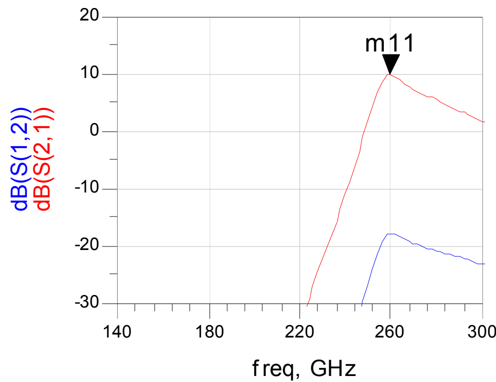
Designs in Fabrication
32 nm CMOS

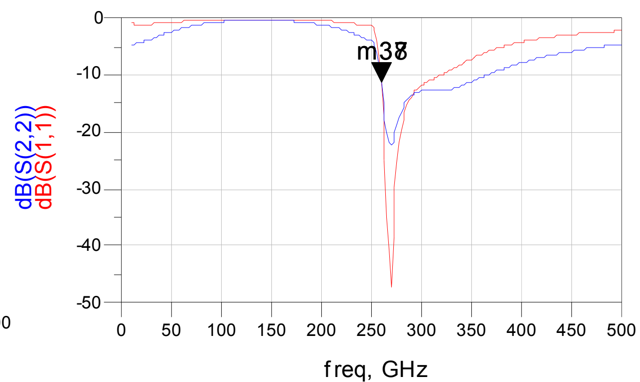


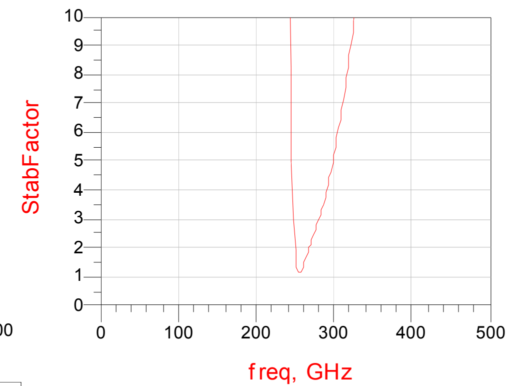
Fcenter= 260GHz

Gain= 9.9dB

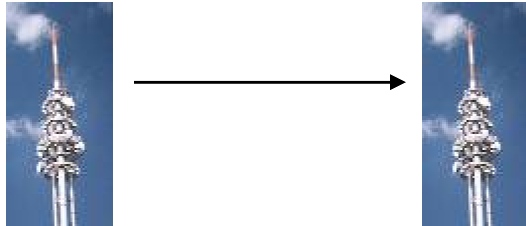
P_{DC}= 15mW







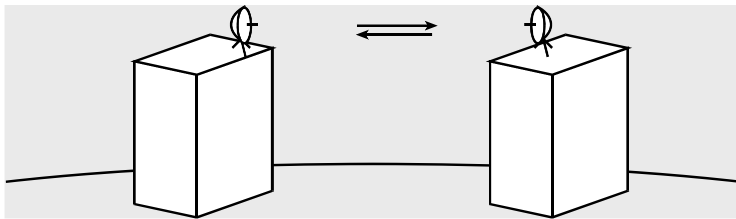
millimeter-wave spectrum: new solutions needed



mm-wave Bands → Lots of bandwidth

$$\left(\frac{P_{received}}{P_{transmitted}}\right) = \left(\frac{1}{16\pi^2}\right)\left(\frac{\lambda^2}{R^2}\right) e^{-\alpha R}$$

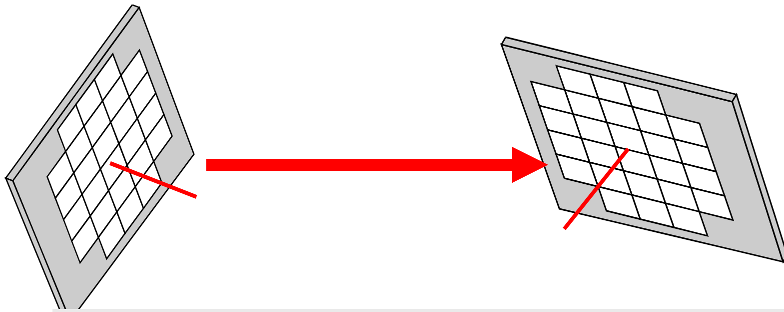
short wavelength→ weak signal → short range



highly directional antenna → strong signal→ long range

$$\left(\frac{P_{received}}{P_{transmitted}}\right) = \left(\frac{D_t D_r}{16\pi^2}\right)\left(\frac{\lambda^2}{R^2}\right) e^{-\alpha R}$$

narrow beam→ must be aimed → no good for mobile



monolithic beam steering arrays→ strong signal, steerable

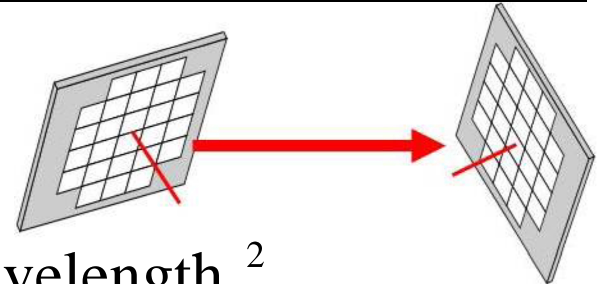
$$\frac{P_{received}}{P_{transmit}} = \frac{N_{receive} N_{transmit}}{16} \frac{\lambda^2}{R^2} e^{-\alpha R}$$

32 x 32 array → 60-90 dB increased SNR→ vastly increased range

→ multi-Gigabit mobile communications

Billions of 700-GHz Transistors→ Imaging & Arrays

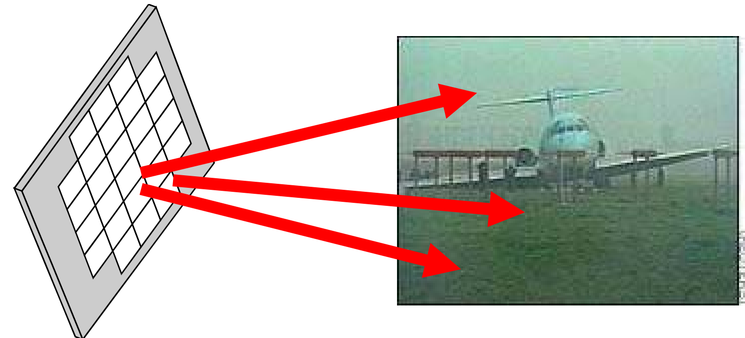
Arrays for point-point radio links:



$$\text{bit rate} \cdot \text{distance}^2 \propto (\#\ \text{array elements})^2 \cdot \text{wavelength}^2$$

Arrays for (sub)-mm-wave imaging :

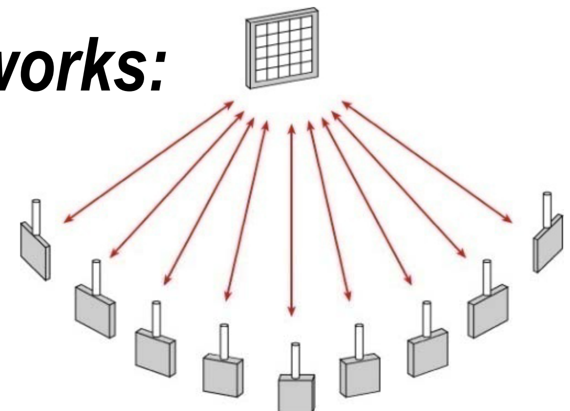
$$\#\ \text{resovable pixels} = \#\ \text{array elements}$$



Arrays for Spatial-Division-Multiplexing Networks:

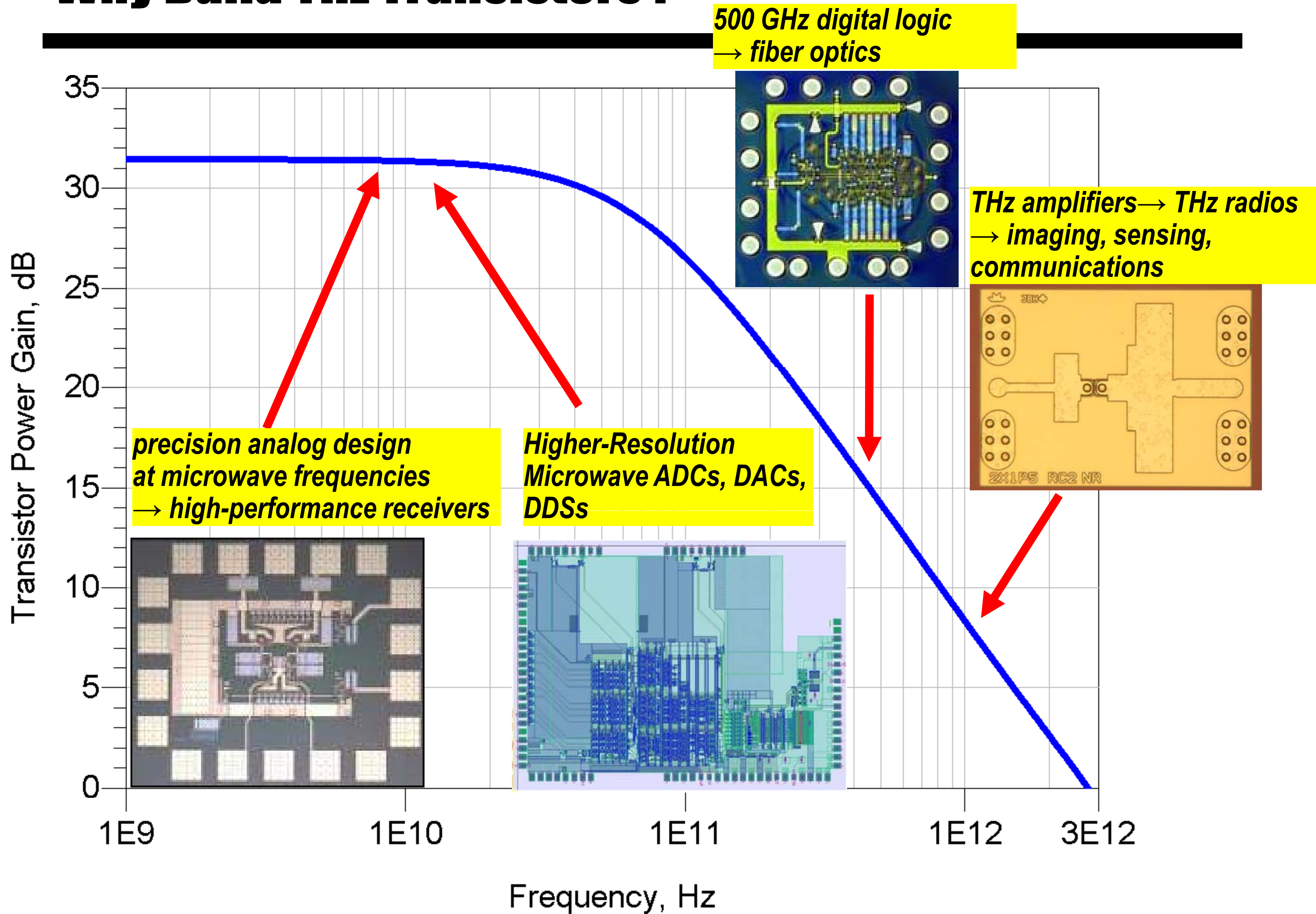
$$\#\ \text{independen t beams} = \#\ \text{array elements}$$

$$\leq \frac{4 \cdot \text{array area}}{\text{wavelength}^2}$$



THz Transistors

Why Build THz Transistors ?



THz Integrated Circuits

Device scaling (Moore's Law) is not yet over.

Scaling→ multi-THz transistors.

Challenges in scaling:
contacts, dielectrics, heat

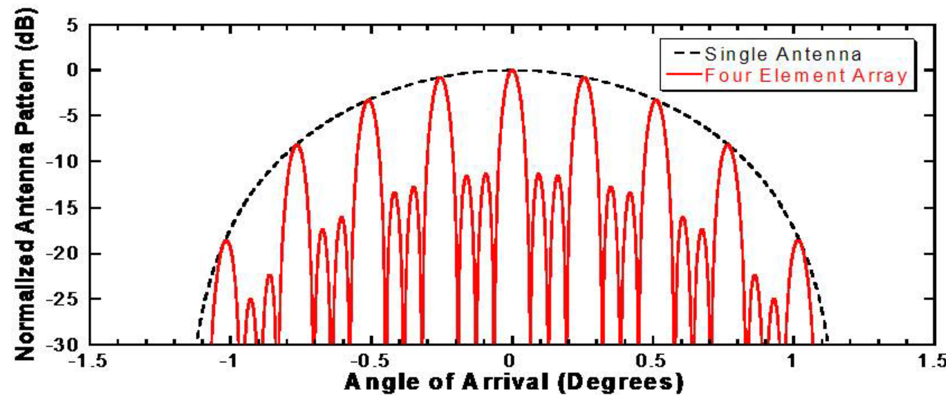
Multi-THz transistors:
for systems at very high frequencies
for better performance at moderate frequencies

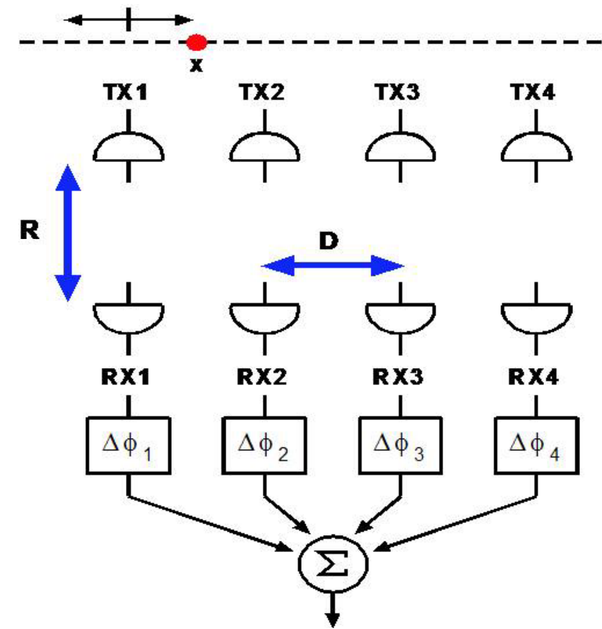
Vast #s of THz transistors
complex systems
new applications.... imaging, radio, and more

end

MIMO Link: a Subsampled Multi-Beam Phased Array

array places nulls at interfering transmitters





Range (m)	Carrier Frequency (GHz)	N	Array Length (m)	Data Rate[1] (Gb/s)
1000	83.5	2	1.34	32
		3	2.19	72
		4	2.84	128
100	83.5	2	0.42	32
		3	0.69	72
		4	0.90	128
10	60.5	2	0.16	45
		3	0.26	101
		4	0.33	179

[1] Assuming *N* x *N* square arrays

$$ND^2 = \lambda R$$