Plenary, Indium Phosphide and Related Materials Conference, May 15-18, Matsue, Japan

Frequency Limits of InP-based Integrated Circuits

Mark Rodwell , E. Lind, Z. Griffith, S. R. Bank, A. M. Crook
U. Singisetti, M. Wistey, G. Burek, A.C. Gossard
University of California, Santa Barbara

Sponsors

J. Zolper, S. Pappert, M. Rosker
DARPA (TFAST, SWIFT, FLARE)

D. Purdy, I. Mack
Office of Naval Research

Kwok Ng, Jim Hutchby
Semiconductor Research Corporation

Collaborators (HBT)

M. Urteaga, R. Pierson , P. Rowell, M-J Choe, B. Brar
Teledyne Scientific Company

X. M. Fang, D. Lubyshev, Y. Wu, J. M. Fastenau, W.K. Liu
International Quantum Epitaxy, Inc.

S. Mohney
Penn State University

Collaborators (III-V MOS)

A. Gossard, S. Stemmer, C. Van de Walle
University of California Santa Barbara

P. Asbeck, A. Kummel, Y. Taur,
University of California San Diego

J. Harris, P. McIntyre,
Stanford University

C. Palmstrøm,
University of Minnesota

M. Fischetti
University of Massachusetts Amherst

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

Specific Acknowledgements

(Prof.) Erik Lind

125 nm HBTs
process technology
theory / epi design

Dr. Zach Griffith

500 & 250 nm HBTs
150 GHz Logic
100 GHz op-amps

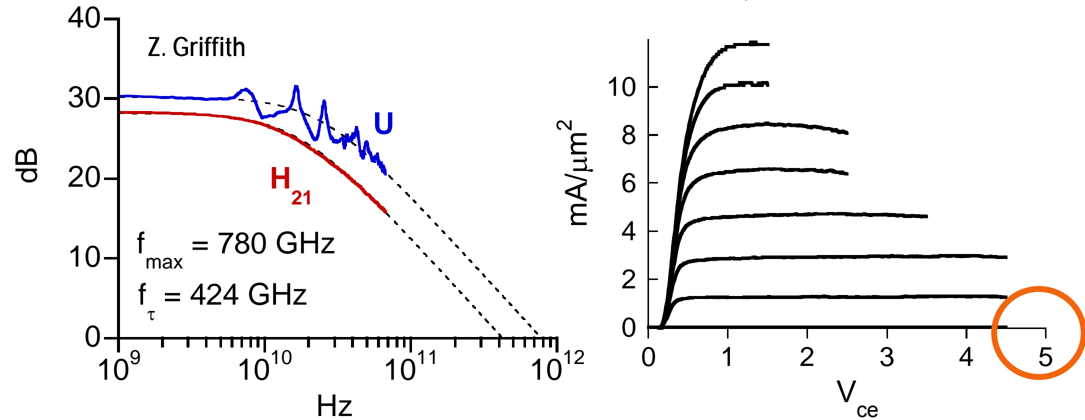
Dr. Mark Wistey

InGaAs MOSFET
process technology
theory / epi design

THz Transistors are coming soon; both InP & Silicon

InP Bipolars: 250 nm generation: → 780 GHz f_{max}, 424 GHz f_{τ}, 4-5 V BV_{CEO}

125 nm & 62 nm nodes
→ ~THz devices



IBM IEDM '06: 65 nm SOI CMOS → 450 GHz f_{max}, ~1 V operation

Intel Jan '07: 45 nm / high-K / metal gate

continued rapid progress
→ continued pressure on III-V technologies

Intel Demonstra
High-k + Metal Gate Transist
on 45 nm Microproc

Mark Bohr
Intel Senior Fellow
Logic Technology Development

Kaizad Mistry
45 nm Program Manager
Logic Technology Developme

1 Jan. 2007

Intel's Logic Technology Evolution

Process Name:	P1262	P1264	P1266	P1268	P1270
Lithography:	90 nm	65 nm	45 nm	32 nm	22 nm
1st Production:	2003	2005	2007	2009	2011

If you can't beat them, join them !
unclear if Si MOSFETs will work well at sub-22-nm gate length
InGaAs/InAs/InP channels under serious investigation for CMOS VLSI.

Datta, DelAlamo, Sadana, ...

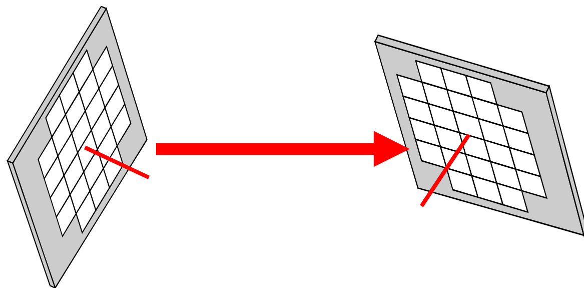
THz InP vs. near-THz CMOS: different opportunities

65 / 45 / 33 / 22 ... nm CMOS

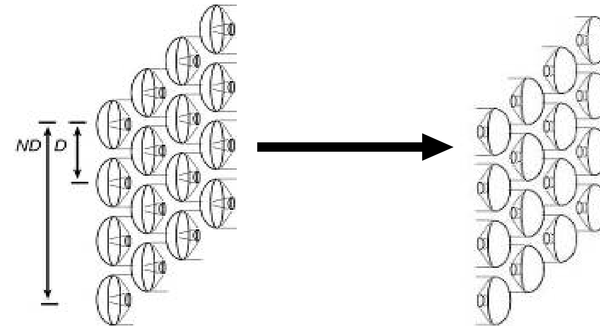
vast #s of very fast transistors

... having low breakdown, sloppy DC parameters

*what **NEW** mm-wave applications will this enable ?*



massive monolithic mm-wave arrays
→ 1 Gb/s over ~1 km

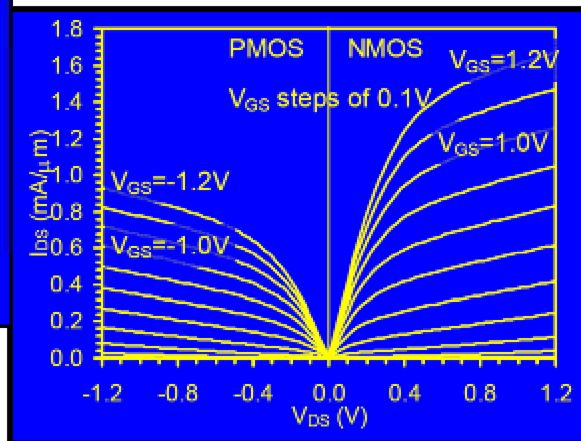


mm-wave MIMO

An advanced low power, high performance, strained channel 65nm technology

Portland Technology Development,
Intel Corporation, Hillsboro, OR, USA

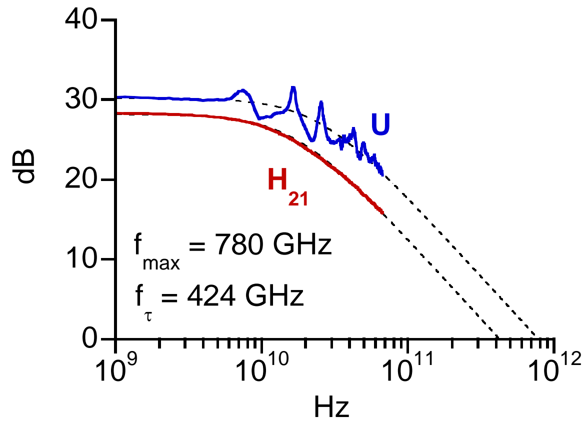
Intel Corporation



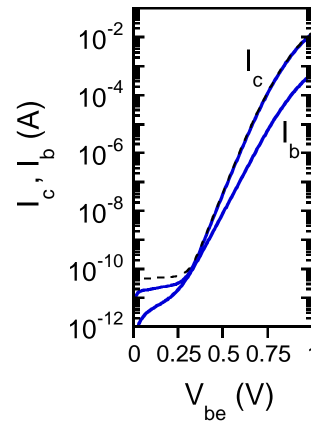
DC parameters
limit analog precision...

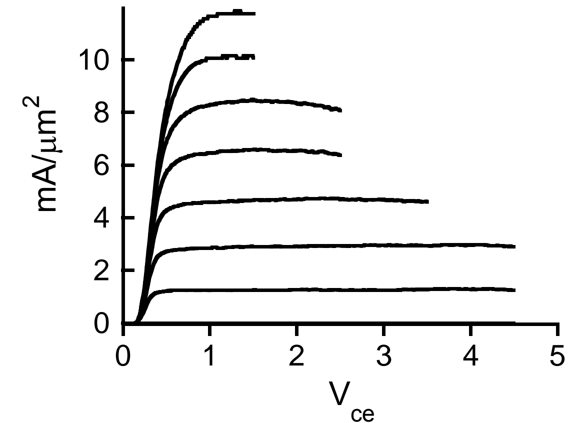
THz InP vs. near-THz CMOS: different opportunities

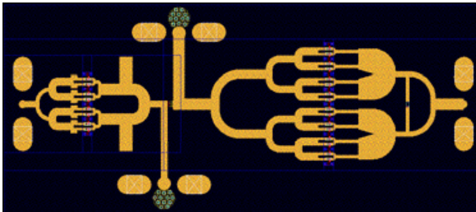
InP HBT: THz bandwidths, good breakdown, analog precision



&



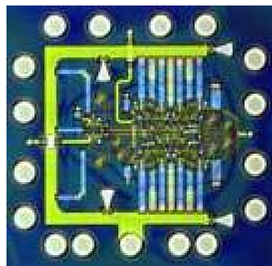




M. Jones

***340 GHz, 70 mW amplifiers** (design)*

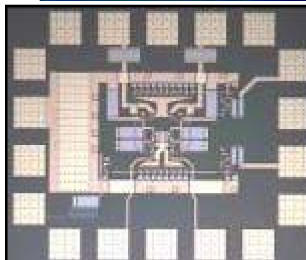
In future: 700 or 1000 GHz amplifiers ?



Z. Griffith

***200 GHz digital logic** (design)*

In future: 450 GHz clock rate ?



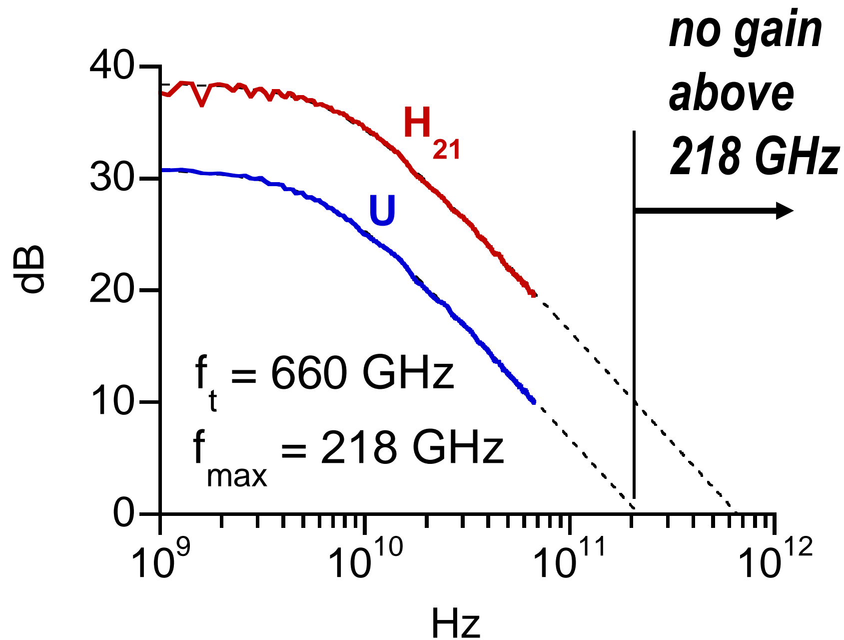
M. Urteaga (Teledyne)

Z. Griffith

***30-50 GHz gain-bandwidth op-amps→ low IM3 @ 2 GHz** In future: 200 GHz op-amps for low-IM3 10 GHz amplifiers?*

Transistor Benchmarks

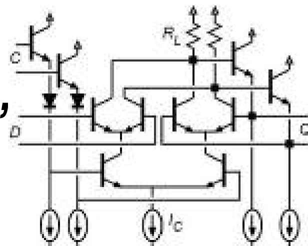
f_{max} matters



Tuned amplifiers: f_{max} sets bandwidth

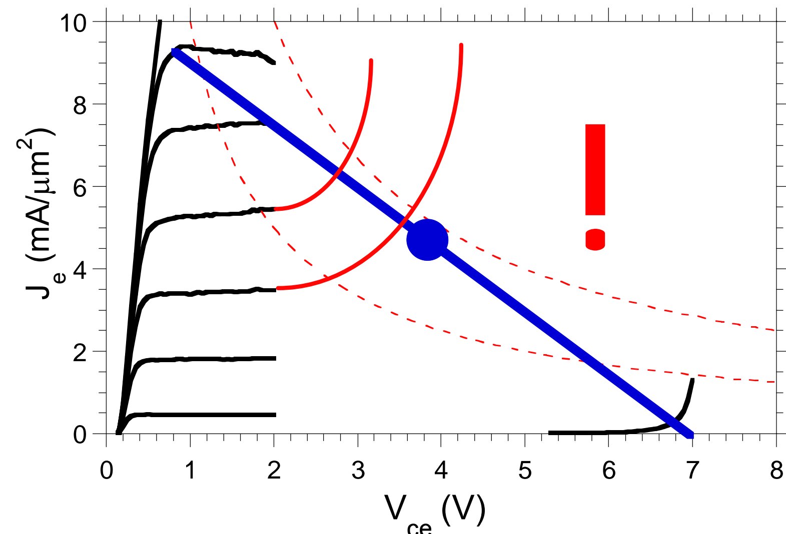
Mixed-signal:

$C_{cb}\Delta V/I_c$, $C_{je}\Delta V/I_c$, $R_{ex}I_c/\Delta V$, $R_{bb}I_c/\Delta V$, τ_f



Goal is >1 THz f_τ and f_{max}
<50 fs $C\Delta V/I$ charging delays

BVCEO is not the only voltage limit



Need Safe Operating Area
...at least $BV_{ceo}/2$ at $J_{max}/2$

thermal resistance,
high-current breakdown
high-temperature operation (~75 C) ?

→ emphasize InP-collector DHBTs

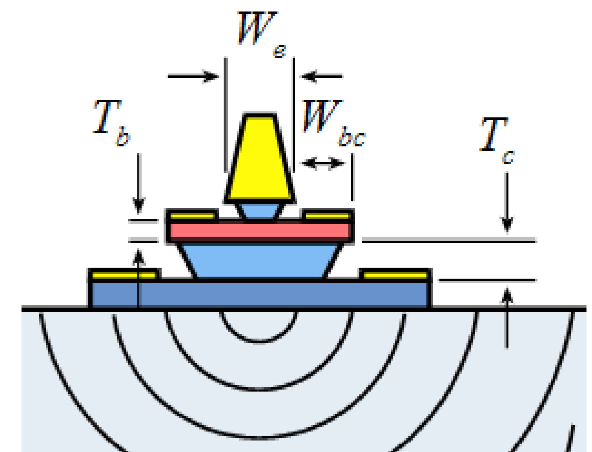
HBT Scaling Laws

HBT scaling laws

*Goal: double transistor bandwidth when used in **any** circuit*

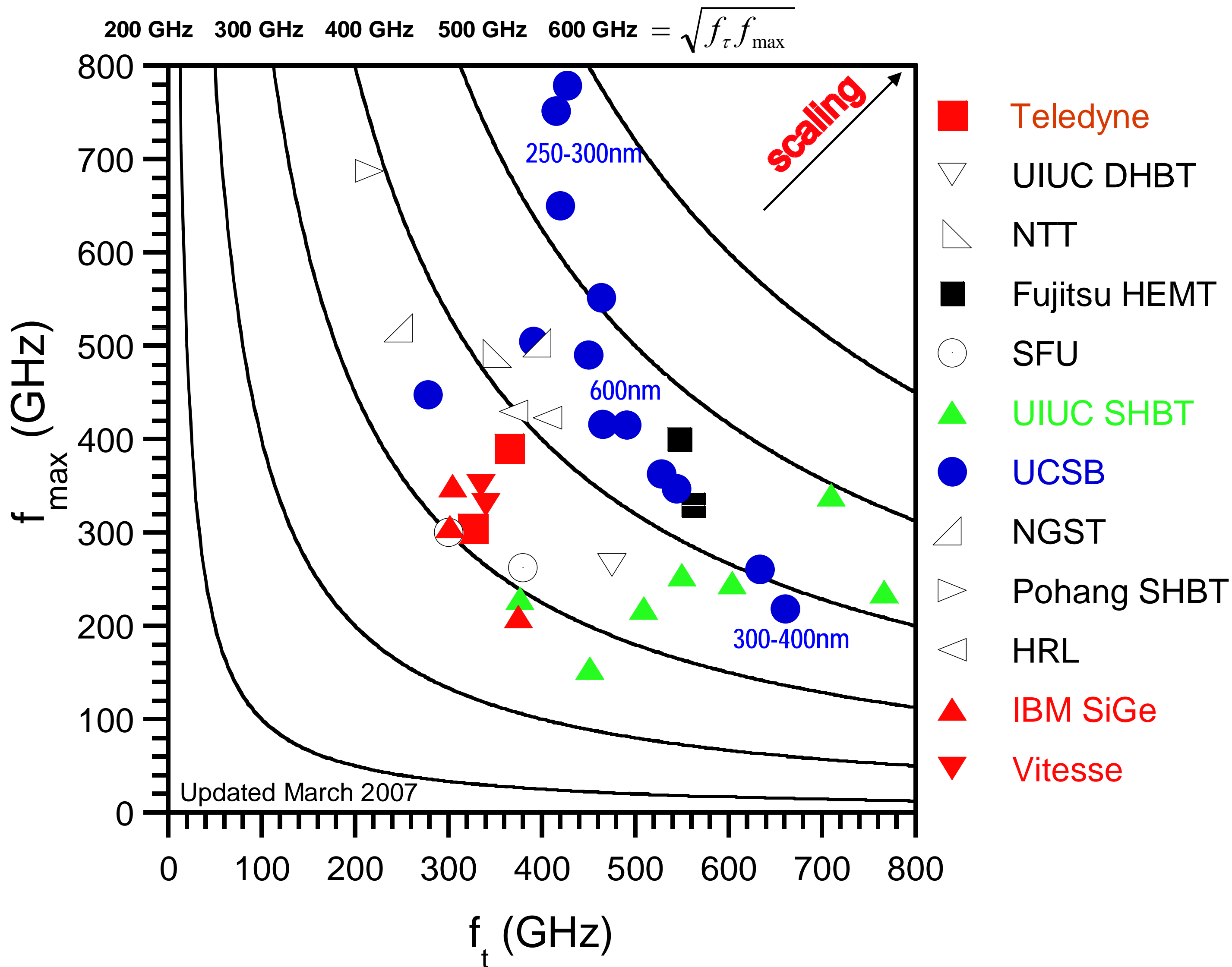
→ keep constant all resistances, voltages, currents

→ reduce 2:1 all capacitances and all transport delays



(emitter length L_E)

InP DHBTs: May 2007



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

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

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

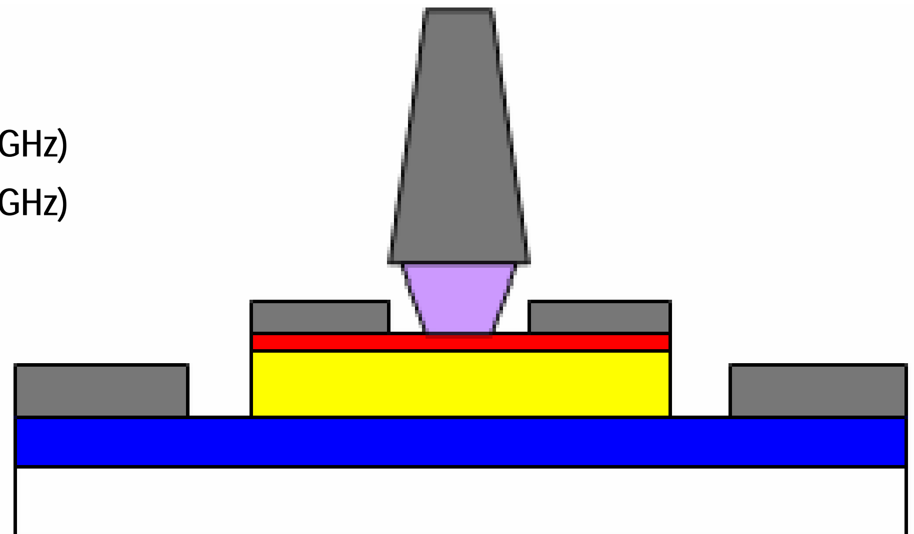
$(R_{bb} I_c / \Delta V)$,

$(\tau_b + \tau_c)$

HBT Scaling Roadmaps

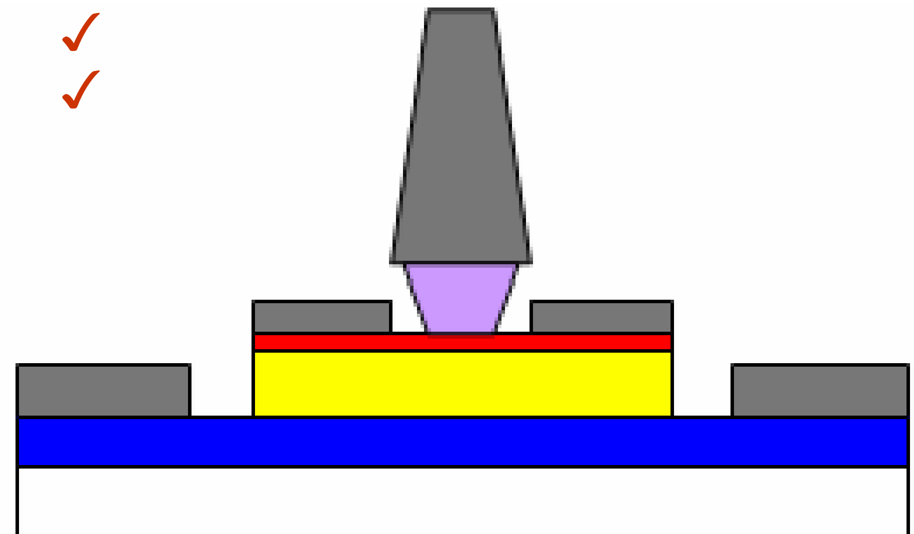
2005: InP DHBTs @ 500 nm Scaling Generation

emitter	500 nm width	✓
	16 $\Omega\cdot\mu m^2$ contact ρ	✓
base	300 width,	✓
	20 $\Omega\cdot\mu m^2$ contact ρ	✓
collector	150 nm thick,	✓
	5 mA/μm^2 current density	✓
	5 V, breakdown	✓
f_τ	400 GHz	✓
f_{max}	500 GHz	✓
power amplifiers	250 GHz	✓ (178 GHz)
digital clock rate (static dividers)	160 GHz	✓ (150 GHz)



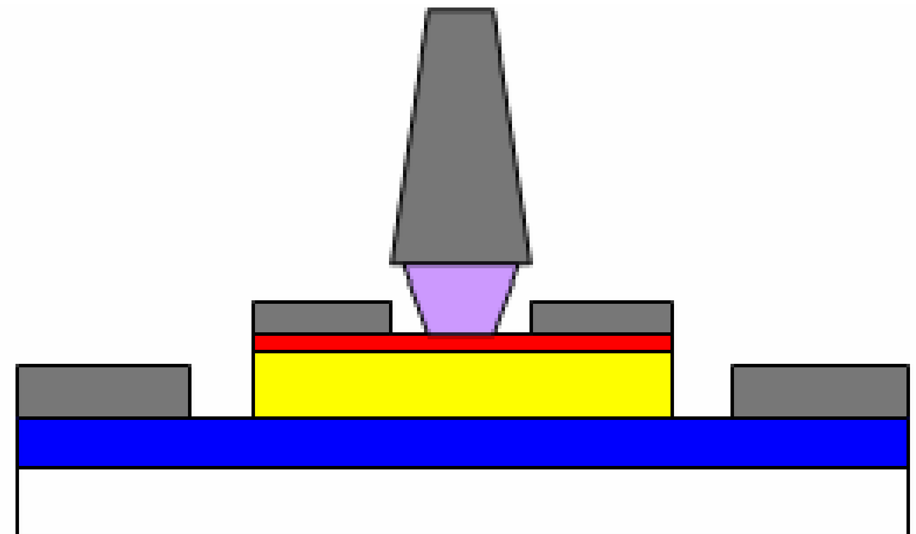
2006: 250 nm Scaling Generation, 1.414:1 faster

emitter	500	250 nm width		✓
	16	9 $\Omega\cdot\mu m^2$ access ρ		✓
base	300	150 width,		✓
	20	10 $\Omega\cdot\mu m^2$ contact ρ		✓
collector	150	100 nm thick,		✓
	5	10 $mA/\mu m^2$ current density		✓
	5	3.5 V, breakdown		✓
f_τ	400	500 GHz	(425 GHz)	✓
f_{max}	500	700 GHz	(780 GHz)	✓
power amplifiers	250	350 GHz		
digital clock rate (static dividers)	160	230 GHz		



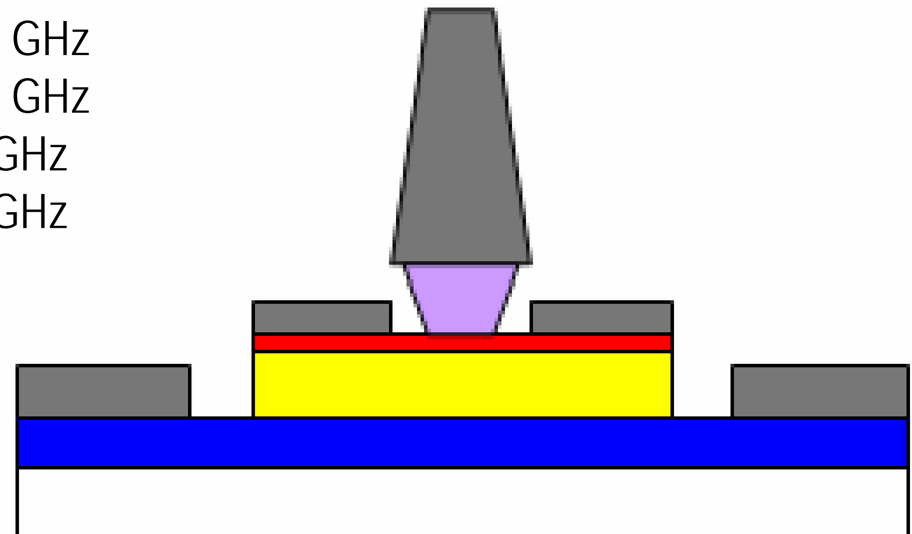
2007: 125 nm Scaling Generation → almost-THz HBT

emitter	500	250	125 nm width	✓
	16	9	4 $\Omega\cdot\mu m^2$ access ρ	✓
base	300	150	75 width,	✓
	20	10	5 $\Omega\cdot\mu m^2$ contact ρ	✓
collector	150	100	75 nm thick,	✓
	5	10	20 $mA/\mu m^2$ current density	✓
	5	3.5	3 V, breakdown	✓
f_τ	400	500	700 GHz	
f_{max}	500	700	1000 GHz	
power amplifiers	250	350	500 GHz	
digital clock rate (static dividers)	160	230	330 GHz	



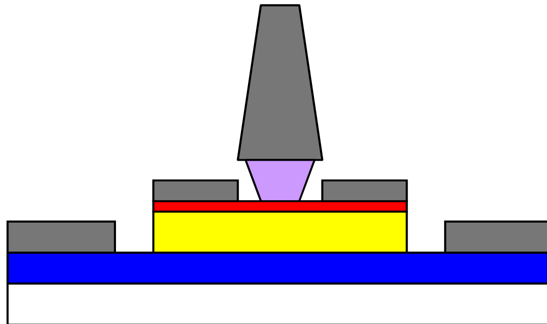
2008-9: 65 nm Scaling Generation→beyond 1-THz HBT

emitter	500	250	125	63 nm width
	16	9	4	2.5 $\Omega\cdot\mu m^2$ access ρ ✓
base	300	150	75	70 nm width,
	20	10	5	5 $\Omega\cdot\mu m^2$ contact ρ ✓
collector	150	100	75	53 nm thick,
	5	10	20	35 $mA/\mu m^2$ current density
	5	3.5	3	2.5 V, breakdown
f_τ	400	500	700	1000 GHz
f_{max}	500	700	1000	1500 GHz
power amplifiers	250	350	500	750 GHz
digital clock rate (static dividers)	160	230	330	450 GHz



HBT Scaling Challenges

Scaling challenges: What looks easy, what looks hard ?



key device parameter	required 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 resistance per unit emitter area	decrease 4:1
current density	increase 4:1
base contact resistivity (if contacts lie above collector junction)	decrease 4:1
base contact resistivity (if contacts do not lie above collector junction)	unchanged

Hard:
Thermal resistance (ICs)
~~Emitter contact + access resistance~~
Yield in deep submicron processes
Contact electromigration (?), dark-line defects (?)

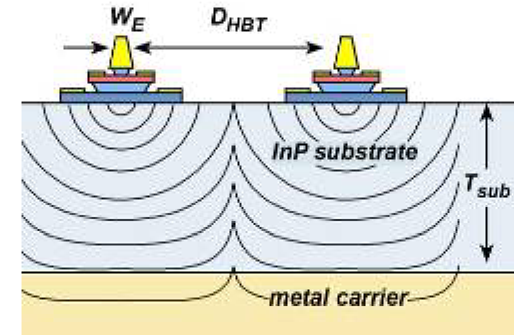
Probably not as hard :
Maintaining adequate breakdown for 3 V operation...

Temperature Rise: Transistor, Substrate, Package

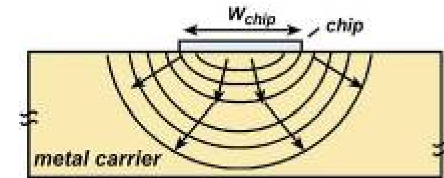
cylindrical heat flow near junction — spherical flow for $r > L_e$ — planar flow for $r > D_{HBT}/2$

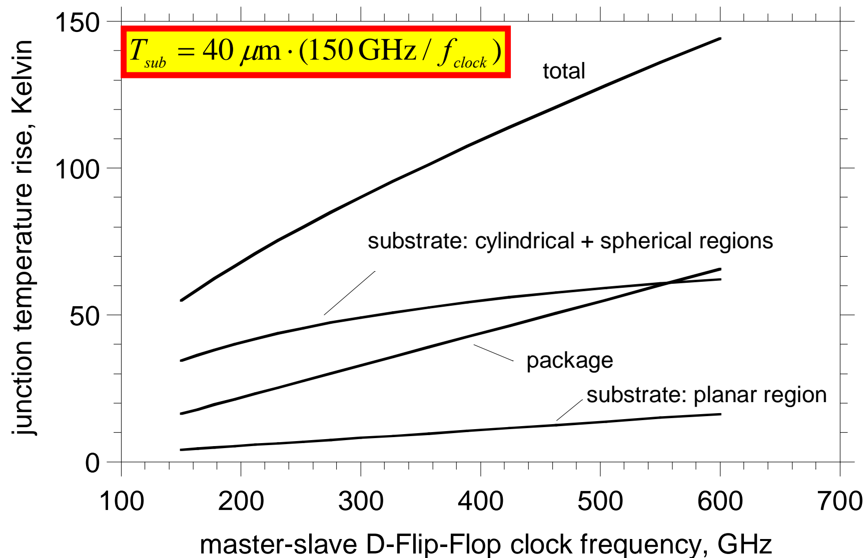
$$\Delta T_{\text{substrate}} \cong \frac{P}{\pi K_{InP} L_E} \ln\left(\frac{L_e}{W_e}\right) + \frac{P}{\pi K_{InP}}\left(\frac{1}{L_E} - \frac{1}{D}\right) + \frac{P}{K_{InP}} \cdot \left(\frac{T_{sub} - D/2}{D^2}\right)$$

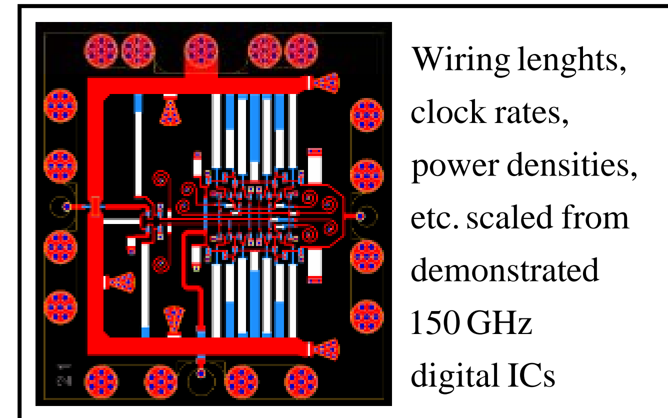
increases logarithmically — insignificant variation — increases quadratically if T_{sub} is constant



$$\Delta T_{package} \cong \left(\frac{1}{\pi} + \frac{1}{2}\right) \frac{P_{chip}}{K_{Cu} W_{chip}}$$







Wiring lenghts, clock rates, power densities, etc. scaled from demonstrated 150 GHz digital ICs

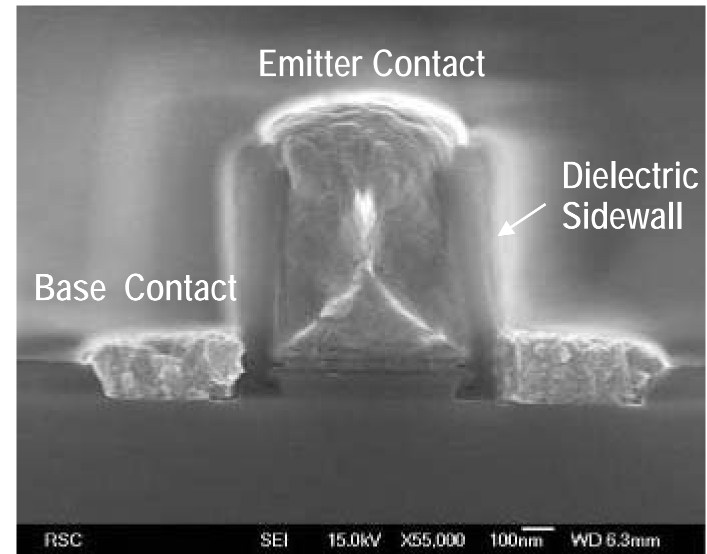
HBTs:

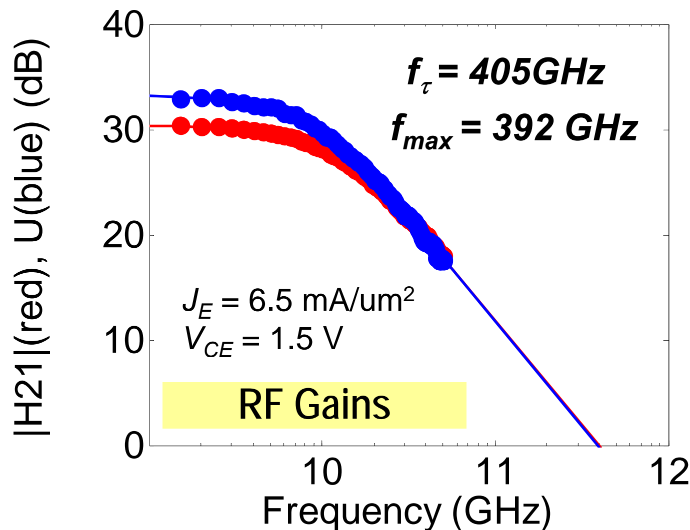
500 nm Generation

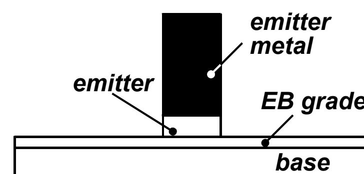
500 nm Generation in Manufacturing: Teledyne Self-aligned Dielectric Sidewall Process

No short-circuits from liftoff defects.

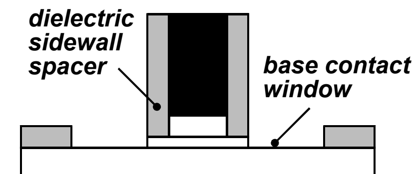
*Emitter can be **much** thinner*
→ small etch undercut.



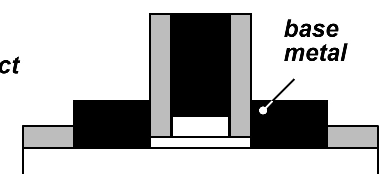




Electroplate emitter contact
Etch emitter semiconductor



Dielectric sidewall deposition
Base contact patterning



Selectively deposit base metal

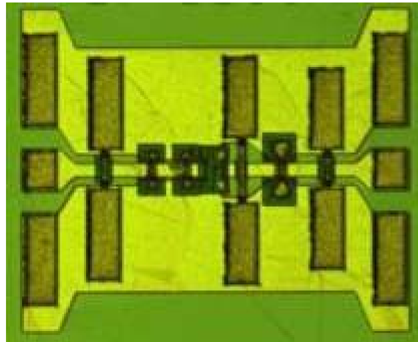
M. Urteaga ***et al***, 2004 IEEE Device Research Conference, June 21-23, 2004

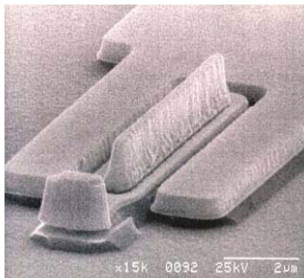
c.f. also Minh Le ***et al*** IEDM 2006 (Vitesse)

Example ICs in 500 nm HBT

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

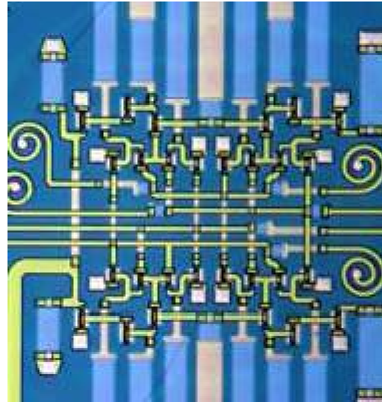
175 GHz, 7.5 mW medium-power amp.

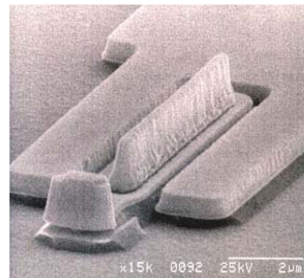




mesa HBT UCSB

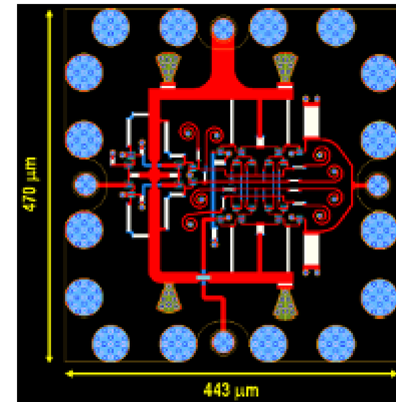
142 GHz, 800 mW master/slave latch

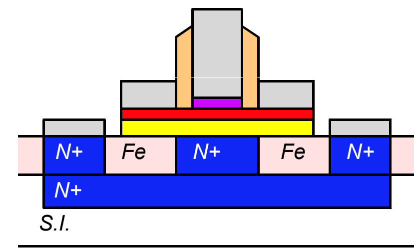




mesa HBT UCSB

128 GHz, 206 mW master/slave latch





sidewall /pedestal HBT Teledyne

Other Results:

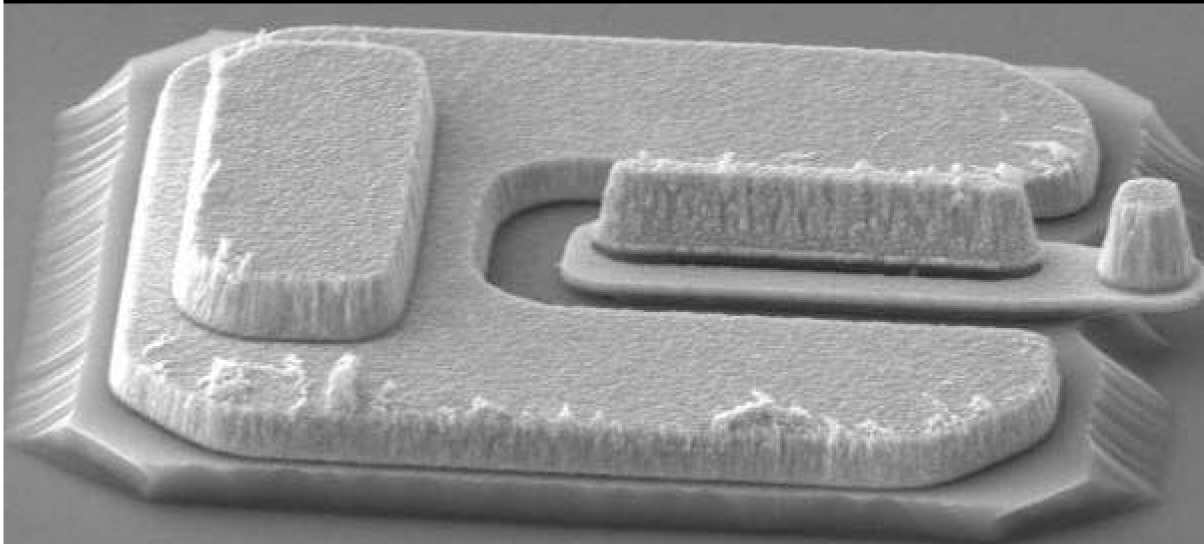
160 Gb/s multiplexer (T. Swahn et al, Chalmers / Vitesse)

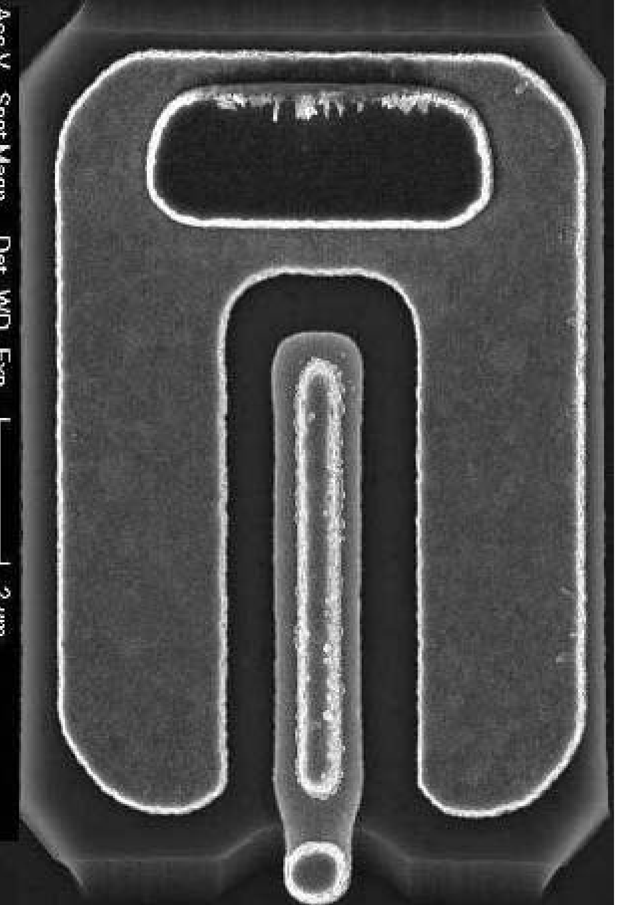
~5000-HBT direct-digital frequency synthesis ICs (Vitesse, Teledyne)

HBTs:

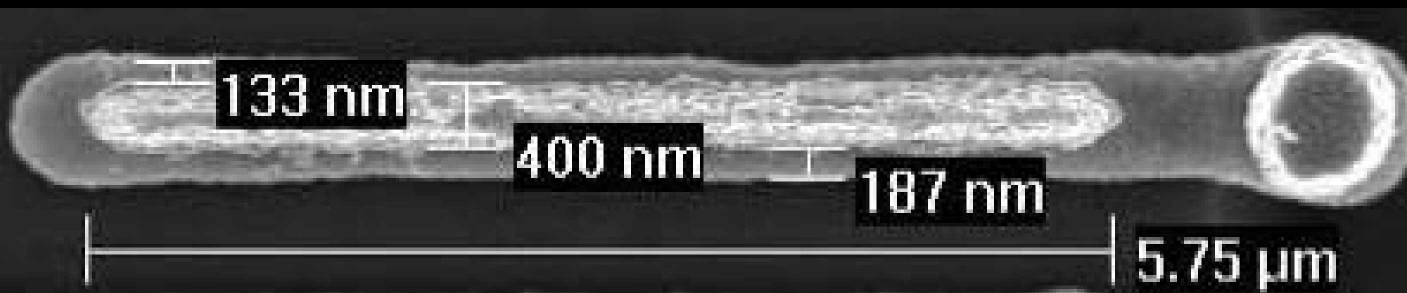
250 nm Generation

250 nm scaling generation InP DHBTs





Emitter contact resistance 5 $\Omega\cdot\mu m^2$
Base contact resistance is < 5 $\Omega\cdot\mu m^2$

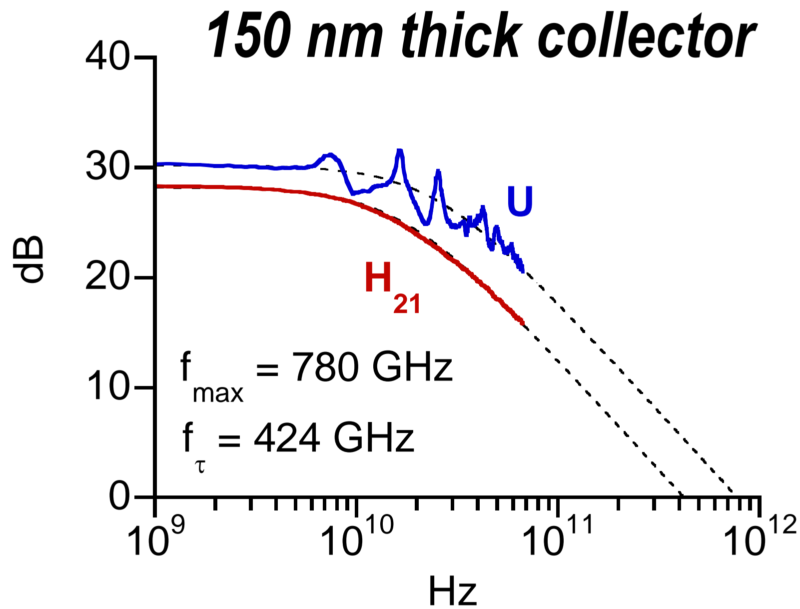


Z. Griffith
E. Lind

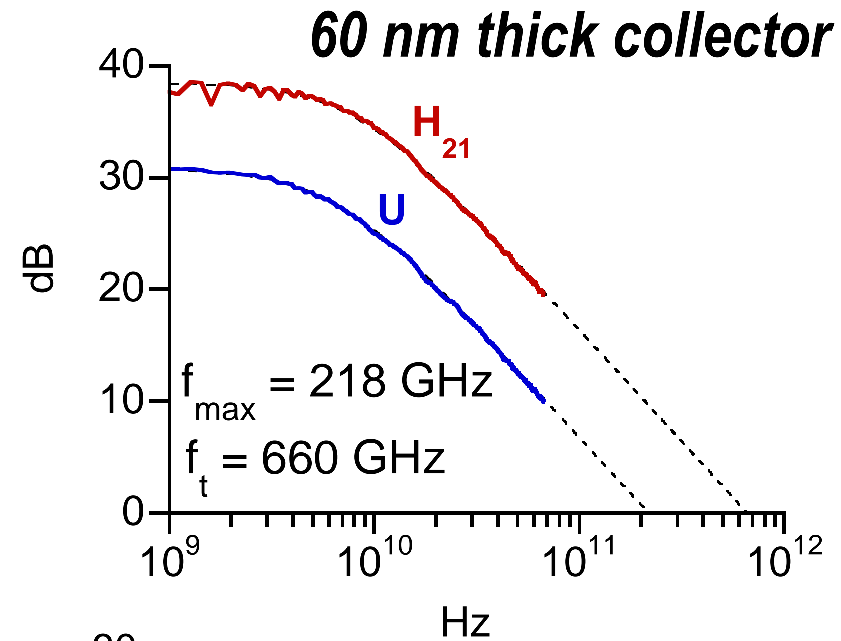
DHBTs: 250 nm Scaling Generation

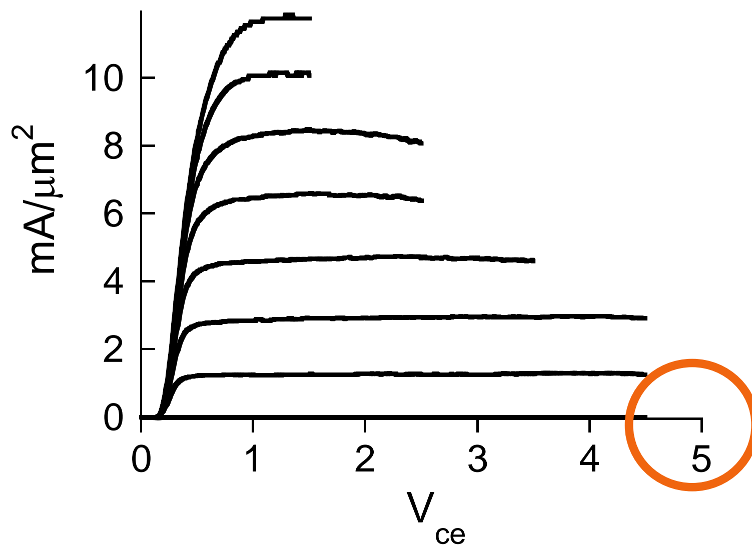
Z. Griffith
E. Lind

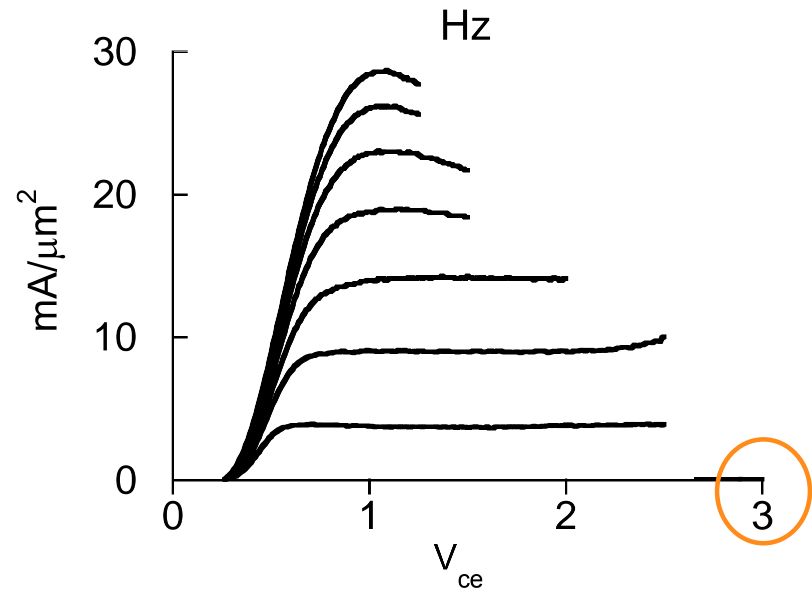
150 nm thick collector



60 nm thick collector





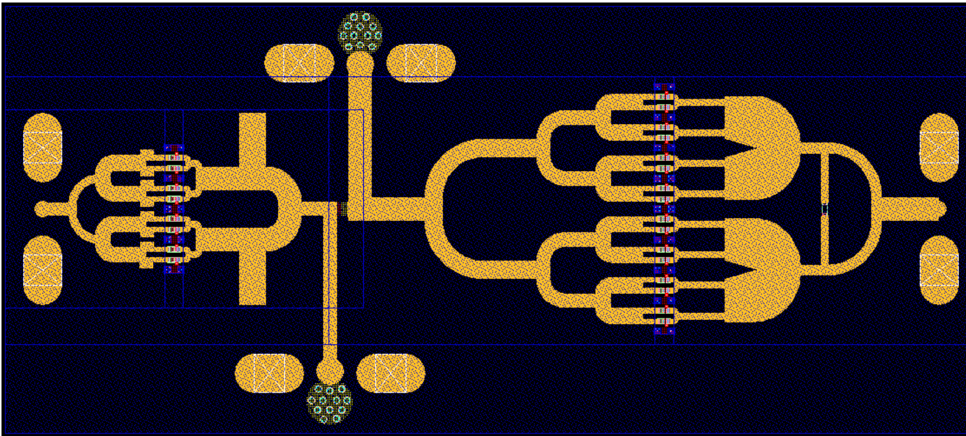


Emitter access: 5.1 Ω·µm²
Base contact: 6.3 Ω·µm²

Example IC Designs in 250 nm HBT

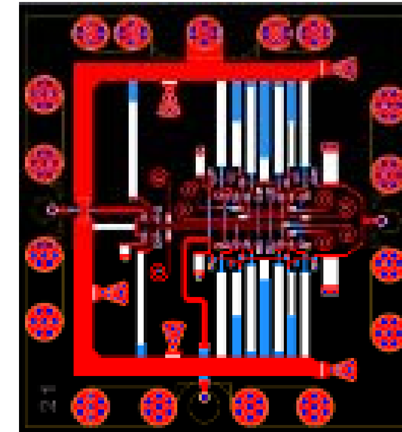
M. Jones
Z. Griffith

340 GHz, 70 mW,
medium-power amplifiers



...fabrication planned summer/fall 2007

200 GHz
master-slave latches



...fabrication on hold...

125 nm InP HBT development

Emitter Access Resistance

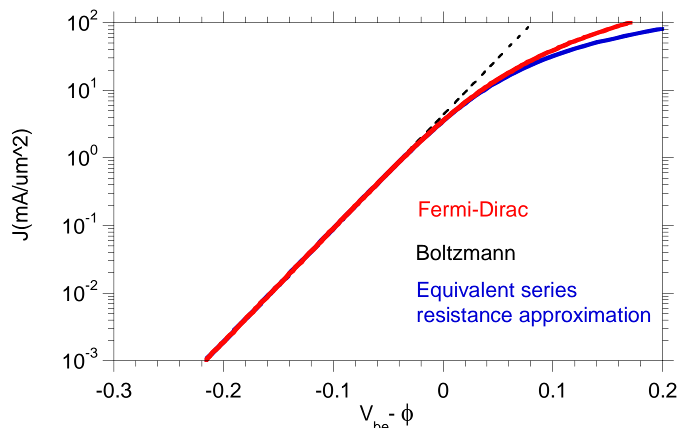
Erik Lind
Adam Crook
Seth Bank
Uttam Singisetti

125 nm generation requires 5 Ω - µm² emitter resistivities

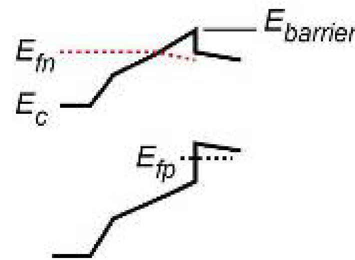
65 nm generation requires 1-2 Ω - µm²

Recent Results:		
ErAs/Mb	*MBE in-situ*	*1.5 Ω - µm²*
Mb	*MBE in-situ*	*0.6 Ω - µm²*
TiPdAu	*ex-situ*	*0.5 Ω - µm²*
TiW	*ex-situ*	*0.7 Ω - µm²*

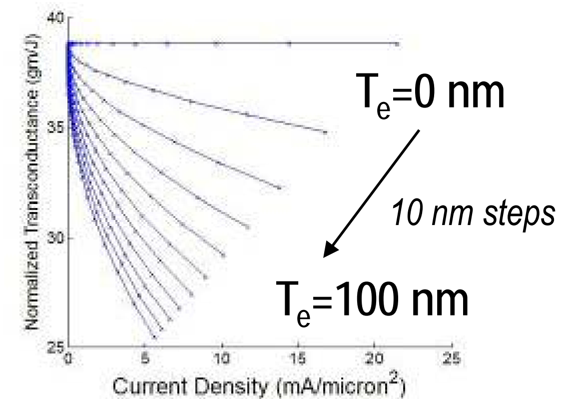
Degeneracy contributes 1 Ω - µm²



20 nm emitter-base depletion layer contributes 1 Ω - µm² resistance



$$\frac{\partial E_{fn}(x)}{\partial x} = \frac{-J}{qn(x)}$$

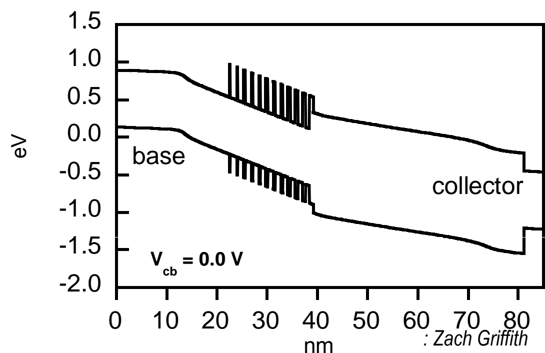


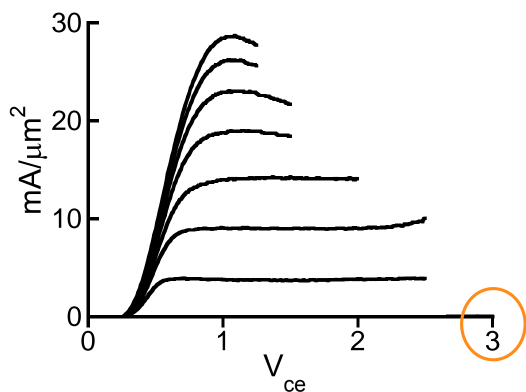
Epitaxial Layer Development for 125 nm Generation

Erik Lind

InGaAs base: low sheet resistivity, low transit time, but collector must be graded

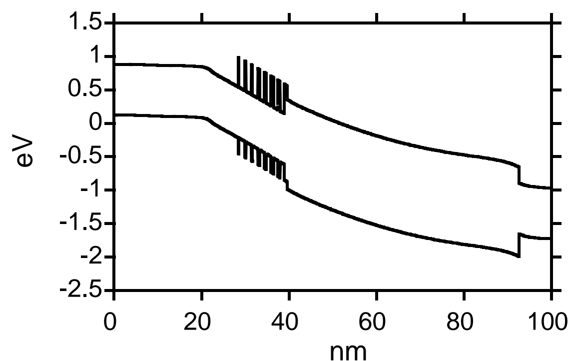
low-current breakdown dominated by tunneling in setback layer

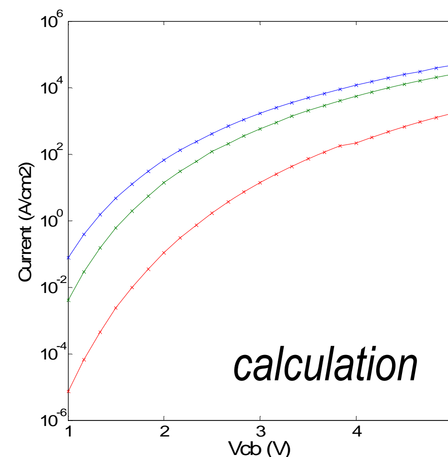




B-C grade redesign: thin the setback, thin the grade

1) less superlattice periods...





2) thinner (sub-monolayer) superlattice periods → random alloy grade

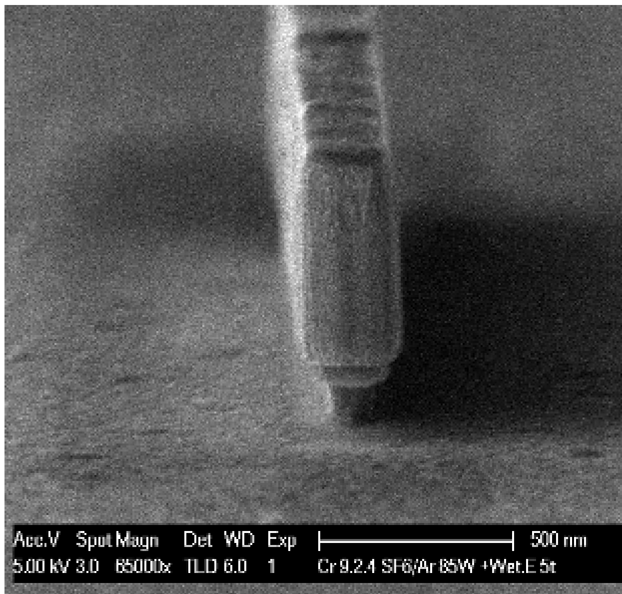
3) thin GaAs/InGaAs strained-layer grade

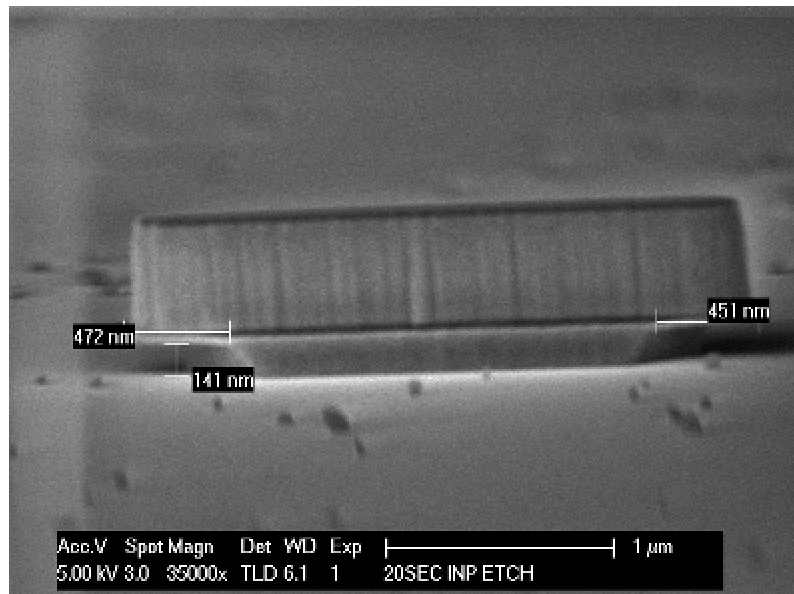
DC data shows expected increase in breakdown. Transport (RF) data is pending.

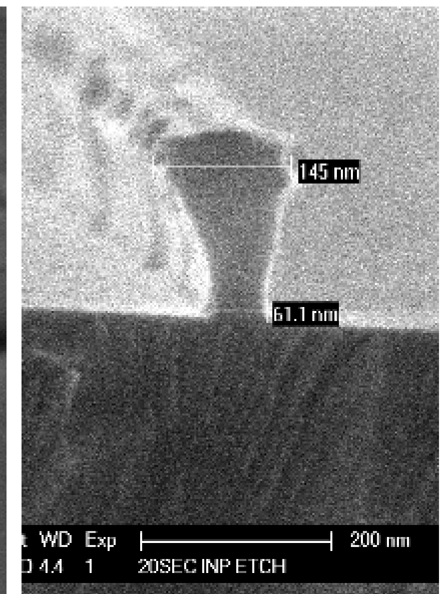
Erik Lind

125 nm Emitter Process

Blanket sputter deposition TiW emitter contact metal
Optical lithography → ICP reactive-ion etching
ICP RIE etch of InGaAs/InP semiconductor,
Selective wet etch to base







125 nm emitter

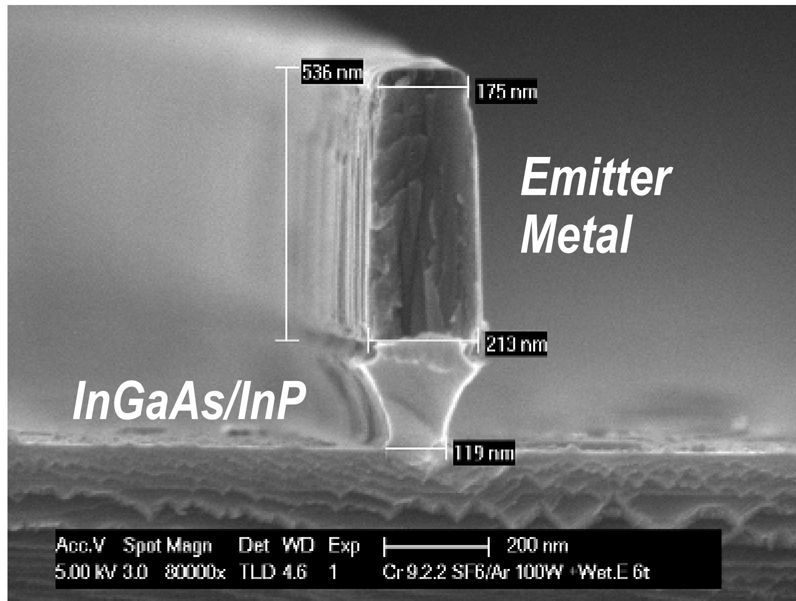
500 nm undercut at emitter ends

61 nm junction: 40 nm lateral undercut

UCSB 125 nm DHBT Development

Erik Lind
Adam Crook

125 nm emitter process

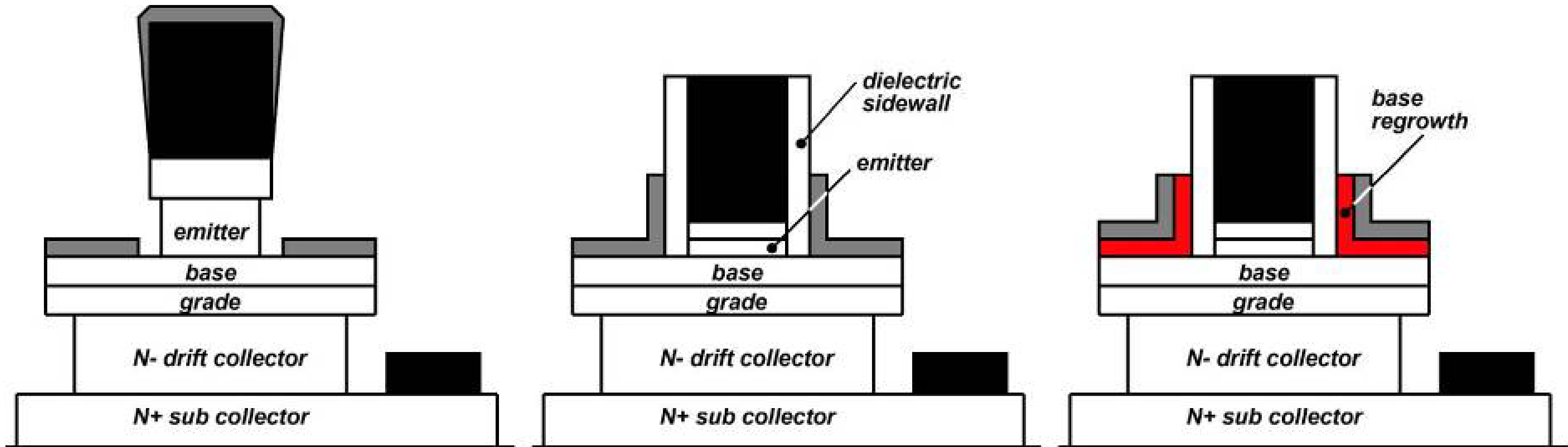


emitter contact resistivity
~ 0.7 Ω - μm^2

base contact resistivity
~ 3-5 Ω -μm^2

Target performance ~ 700-900 GHz simultaneous f_t & f_{max} ,
3-4 V breakdown

How might we build the 62.5 nm HBT ?



Mesa process: control of etch undercut with dry+wet process

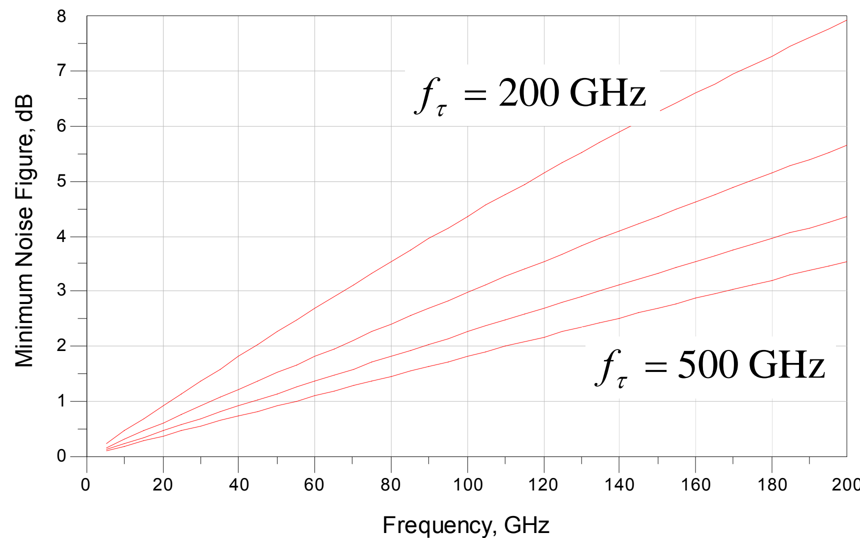
Alternatives:
- dielectric sidewall process
- sidewall process with extrinsic base regrowth: allows thinner base

InP-based FETs;

MOSFETs & HEMTs

InP-based HEMTs & MOSFETs : Why ?

InGaAs/InP HEMTs: mm-wave low-noise amplifiers



$$F_{\min} \approx 1 + \sqrt{g_{mi}(R_s + R_g + R_i)\Gamma} \cdot \left(\frac{f}{f_\tau} \right)$$

A ~2.5:1 f_τ / f_{signal} ratio provides 3 dB noise figure.

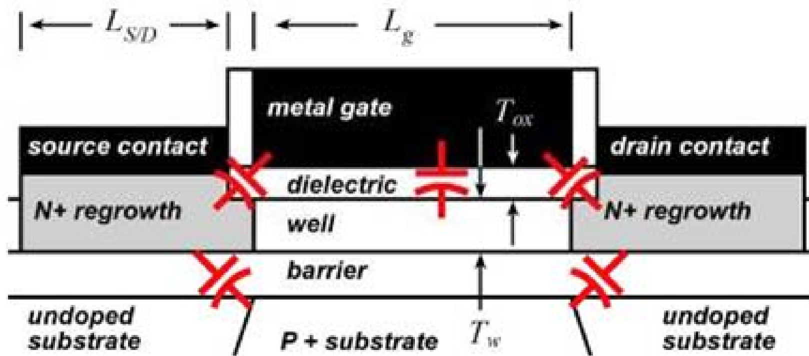
Low-noise 100-300 GHz preamplification is a key application for 1-THz-f_τ HEMTs

InGaAs/InP MOSFETs: post-22-nm VLSI (?)

Higher mobility and peak electron velocity than in Silicon

→ higher (I_d / W_g) and lower ($C\Delta V$ / I) at sub-22-nm scaling (?)

Back-of-Envelope FET Scaling



*Goal double transistor bandwidth when used in **any** circuit*
→ reduce 2:1 all capacitances and all transport delays
→ keep constant all resistances, voltages, currents

FETs no longer scale well

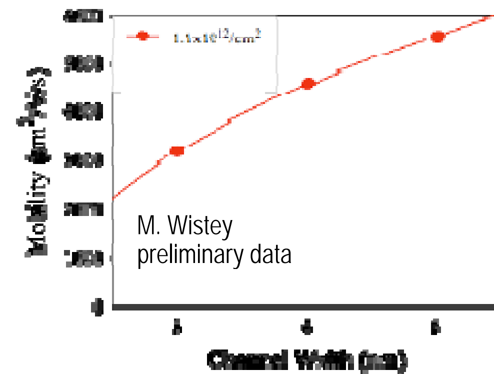
tunneling through oxide → *high-K dielectrics (if feasible)*

Some Encouraging Initial Data . . .

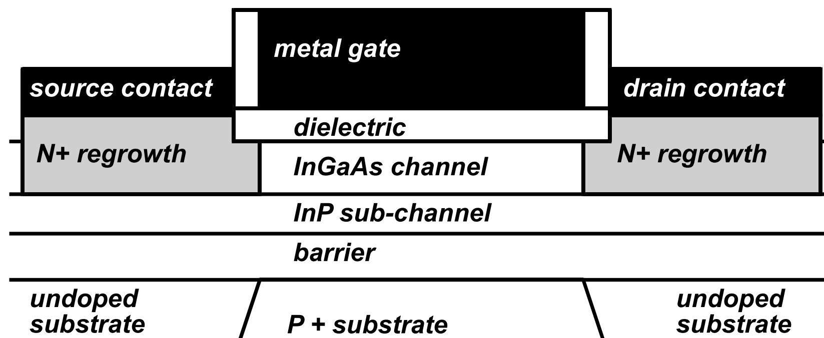
-- non-parabolic bands (variable m*) significantly increase feasible sheet charge

Asbeck / Fischetti / Taur simulate drive currents much larger than for constant-m model*

-- mobilities seem to be acceptable even in thin wells



. . . and our current device designs . . .



well: 2.5 nm InGaAs, 2.5 nm InP

N+ InGaAs/InAs extrinsic source & drain by regrowth

device design and fabrication:
Asbeck group: UCSD
Taur group: UCSD
Fischetti group: U. Mass
Rodwell group: UCSB
Palmstrøm group: U. Minn

Frequency Limits of InP-based Integrated Circuits

InP Bipolar Transistors

Scaling limits: contact resistivities, device and IC thermal resistances.

62 nm (1 THz f_τ, 1.5 THz f_{max}) scaling generation is feasible.

700 GHz amplifiers, 450 GHz digital logic

Is the 32 nm (1 THz amplifiers) generation feasible ?

InP Field-Effect Transistors

Low electron effective mass → difficulties with further scaling

Guarded optimism regarding 22 nm generation for VLSI

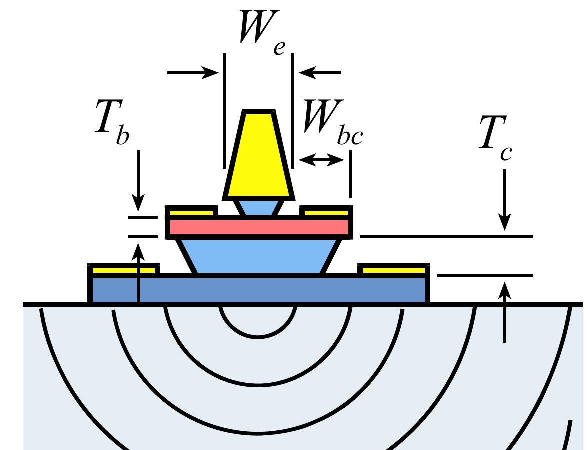
Serious difficulties beyond.

(end)

non-animated versions of the three key scaling slides

HBT scaling laws

*Goal: double transistor bandwidth when used in **any** circuit*
→ keep constant all resistances, voltages, currents
→ reduce 2:1 all capacitances and all transport delays



(emitter length L_E)

$\tau_b = T_b^2/2D_n + T_b/v$ → *thin base ~1.414:1*

$\tau_c = T_c/2v$ → *thin collector 2:1*

$C_{cb} \propto A_c/T_c$ → *reduce junction areas 4:1*

$R_{ex} = \rho_c/A_e$ → *reduce emitter contact resistivity 4:1*

$I_{c,Kirk} \propto A_e/T_c^2$ *(current remains constant, as desired)*

$$\Delta T \cong \frac{P}{\pi K_{InP} L_E} \ln\left(\frac{L_e}{W_e}\right) + \frac{P}{\pi K_{InP} L_E}$$

need to reduce junction areas 4:1
reduce widths 2:1 & reduce length 2:1 → doubles ΔT ✗
reducing widths 4:1, keep constant length→ small ΔT increase ✓

$$R_{bb} \cong \frac{\rho_s W_e}{12 L_e} + \frac{\rho_s W_{bc}}{6 L_e} + \frac{\rho_c}{A_{contacts}}$$

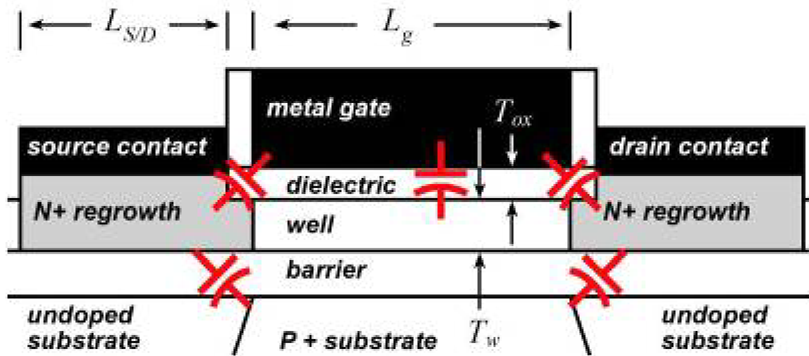
→ *reduce base contact resistivity 4:1*

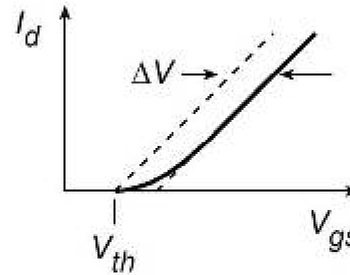
reduce widths 2:1 & reduce length 2:1 → constant R_{bb} ✓
reducing widths 4:1, keep constant length → reduced R_{bb} ✓✓

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

Back-of-Envelope FET Scaling

*Goal double transistor bandwidth when used in **any** circuit*
→ reduce 2:1 all capacitances and all transport delays
→ keep constant all resistances, voltages, currents





$$I_d \sim c_{eq} v_{exit} W_g (V_{gs} - V_{th} - \Delta V)$$

$$v_{exit} \sim (kT/m^*)^{1/2}$$ *(non-degenerate)*

$$\Delta V \sim v_{exit} L_g / \mu$$

$1/c_{eq} \sim T_{ox}/\varepsilon_{ox} + T_w / 2\varepsilon_{well}$ — thin layers 2:1 → c_{eq} doubled

$g_m \sim c_{eq} v_{exit} W_g$ — reduce W_g 2:1 → g_m, I_d held constant ✓

$C_{gs} \sim c_{eq} L_g W_g + \alpha_1 W_g$ — reduce L_g 2:1 → C_{gs} reduced 2:1 ✓

(dielectric: $c_{eq} L_g W_g$; fringing: $\alpha_1 W_g$)

$(C_{gd}, C_{s-b}, C_{d-b})$ all proportional to W_g → $(C_{gd}, C_{s-b}, C_{d-b})$ all reduced 2:1 ✓

$$R_s = \frac{\rho_c}{W_g L_{S/D}} + \frac{\rho_s L_{S/D}}{W_g}$$

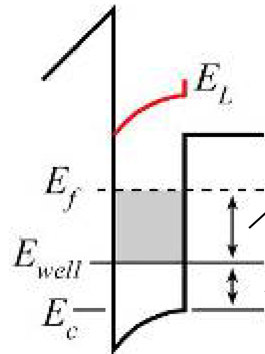
reduce $L_{s/d}$ 2:1, reduce ρ_c 4:1 → (R_s, R_d) held constant ✓

2:1 vertical scaling → 2:1 increased (g_m / W_g) → 2:1 reduced W_g
→ <u>2:1 reduced fringing capacitances</u>

FETs no longer scale well

tunneling through oxide *→ high-K dielectrics (if feasible)*

Thin layers & low effective mass limit channel sheet charge density



$$\left(E_f - E_{well}\right) \propto qn_s / m^*$$

$$\left(E_{well} - E_c\right) \propto 1/T_{well}^2\ m^*$$ infinite well approximation

high sheet-charge in thin well
→ populate higher-mass band minima

Low density of states limits drive current Solomon & Laux, 2001 IEDM

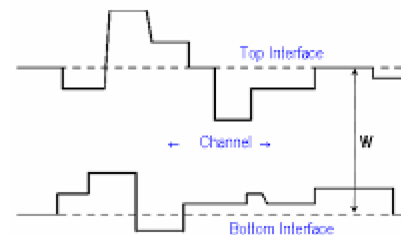
$$g_m \sim c_{eq} v \cdot W_g \text{ where } 1/c_{eq} \sim (2)\pi\hbar^2 / q^2 m_e^* + T_{ox} / \varepsilon_{ox} + T_w / 2\varepsilon_{well}$$

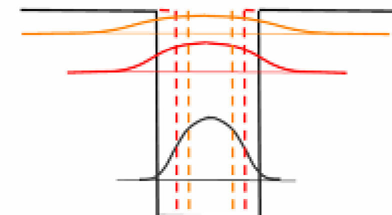
density-of-states term dominates , limits (g_m / W_g)and (I_d / W_g)
→ fringing & substrate capacitances no longer scale, can dominate over C_{gs}

Thin quantum wells have low mobility

Li SST 2005; Gold *et al*, SSC 1987; Sakaki *et al*, APL 1987

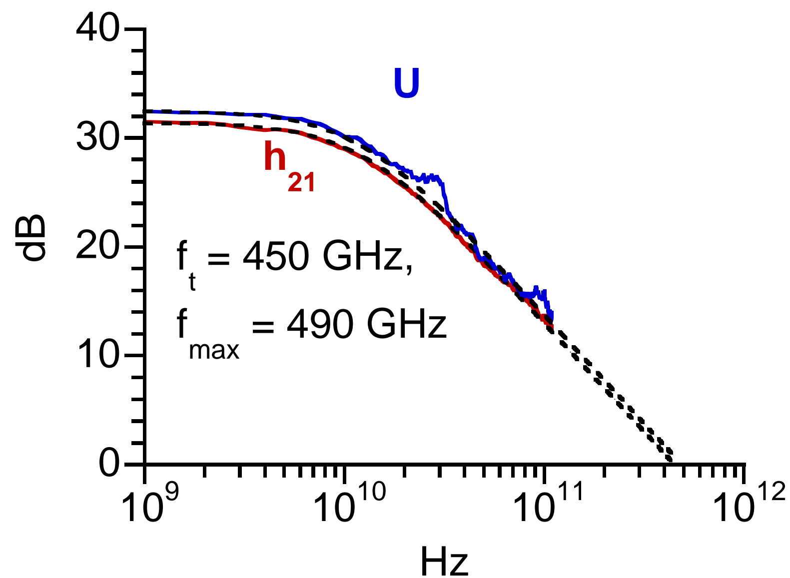
$$\mu \propto \left(\partial E / \partial W\right)^{-2} \propto T_{well}^6$$

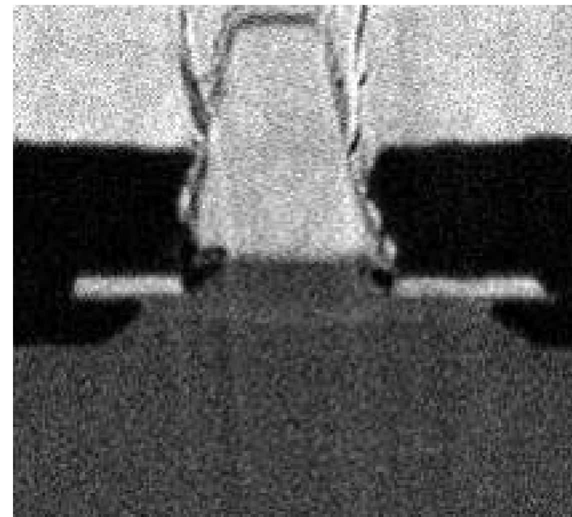


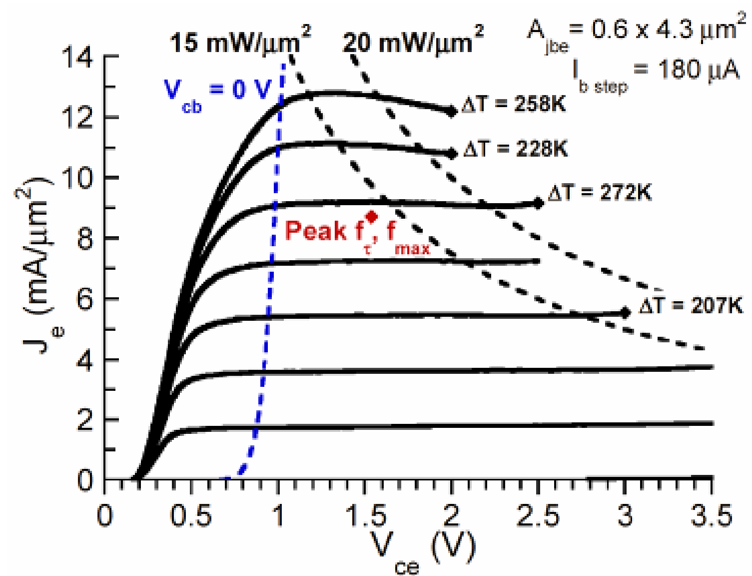


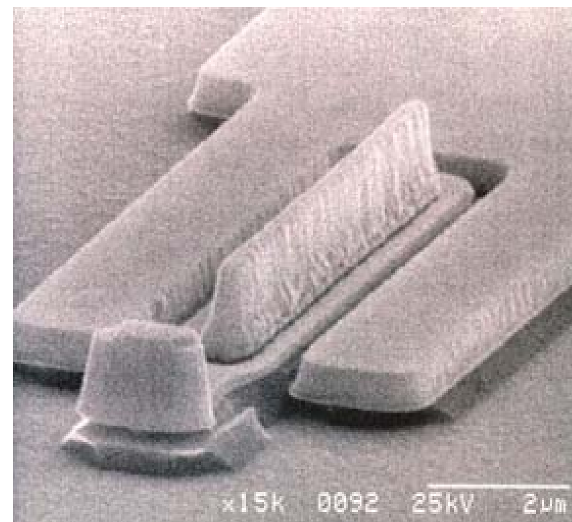
InP DHBT: 500 nm Scaling Generation

Z. Griffith









600 nm wide emitter, 120 nm thick collector, 30 nm thick base