Transistor Performance Scaling: The Role of Virtual Source Velocity and Its Mobility Dependence

Ali Khakifirooz and Dimitri A. Antoniadis

Microsystems Technology Laboratories,
Massachusetts Institute of Technology

Supported by MARCO MSD.

Outline

- New Performance Metric
- Historical Trend of the Performance
- The Role of Velocity Evolution
- Velocity-Mobility Relationship
- How Close to Ballistic?
- Future Directions
- Conclusions

Performance Metric

$$\tau = \frac{\Delta Q_G}{I_{\text{eff}}}$$

ΔQ_G Charge difference between the ON and OFF states.

$$I_{\text{eff}} = ((I_D(V_G = V_{DD}, V_D = 0.5V_{DD}) + I_D(V_G = 0.5V_{DD}, V_D = V_{DD}))/2$$

$$I_D = WC_{\text{inv}}(V_{GS} - V_T)v$$

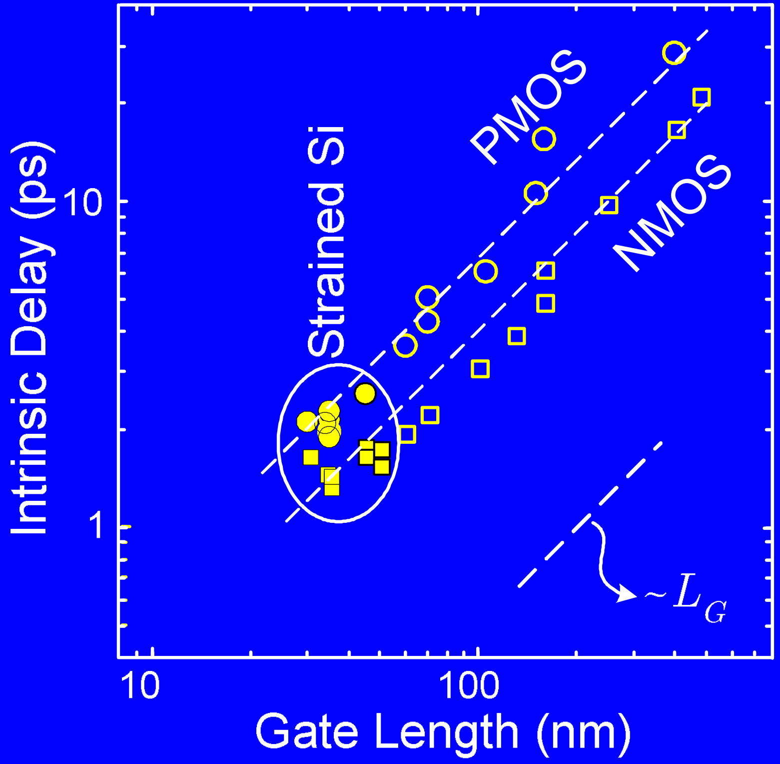
$$V_T(V_{DS}) = V_{T0} - \delta V_{DS}$$

$$\tau = \frac{(1-\delta)V_{DD} - V_T + (C_f^* / C_{\text{inv}} L_G)V_{DD}}{(3-\delta)V_{DD}/4 - V_T} \frac{L_G}{\boxed{v}}$$

D. A. Antoniadis, *et al.*, *IBM J. Res. Dev.*, vol. 50., p. 363, 2006.

Historical Trend of Performance

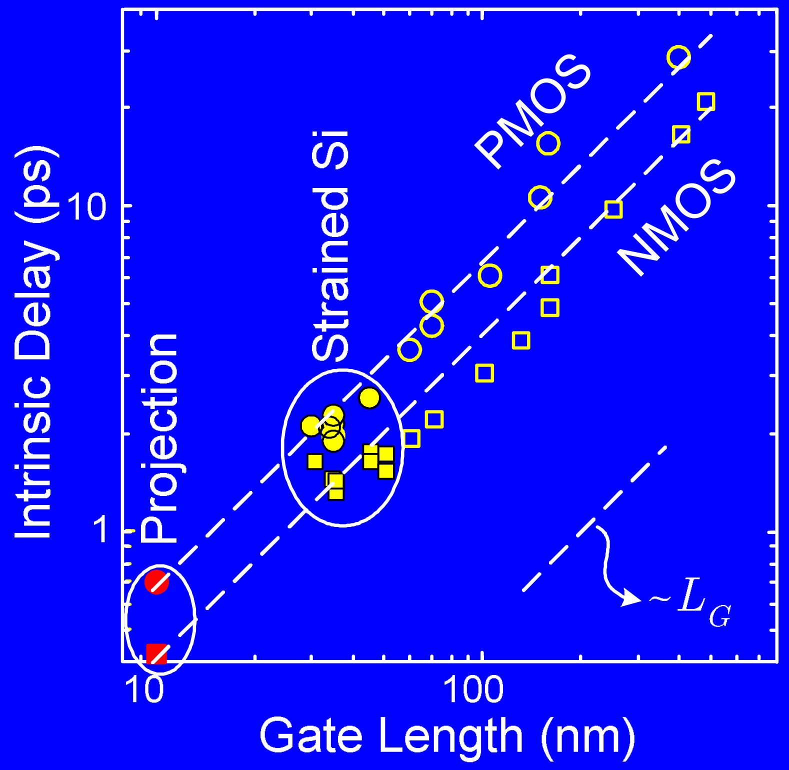
- **Over the past 20 years, MOSFET intrinsic delay has remarkably followed the gate length scaling.**



C_f^* = 0.5 fF/μm independent of technology.

Historical Trend of Performance

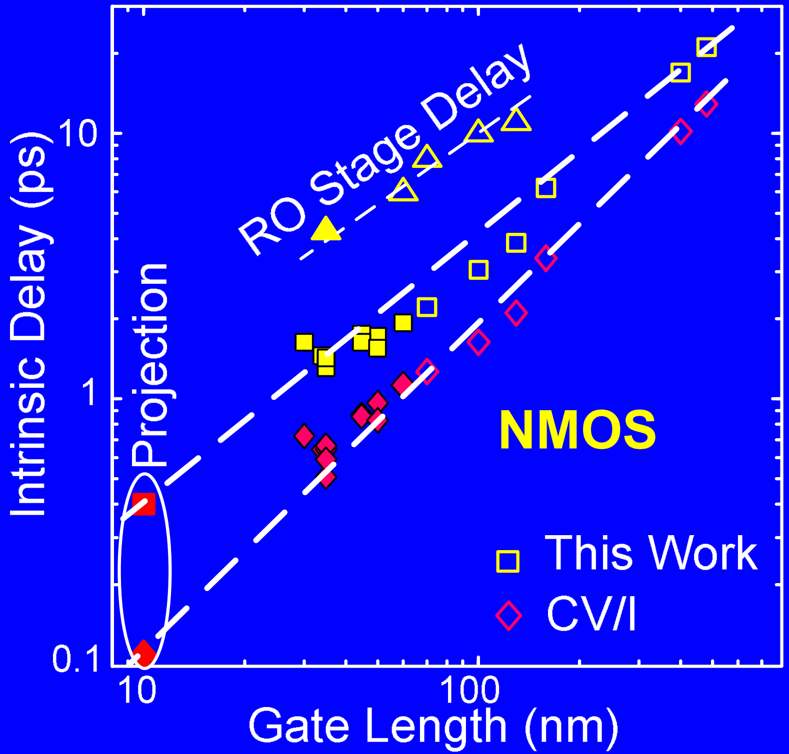
- Over the past 20 years, MOSFET intrinsic delay has remarkably followed the gate length scaling.



C_f^* = 0.5 fF/μm independent of technology.

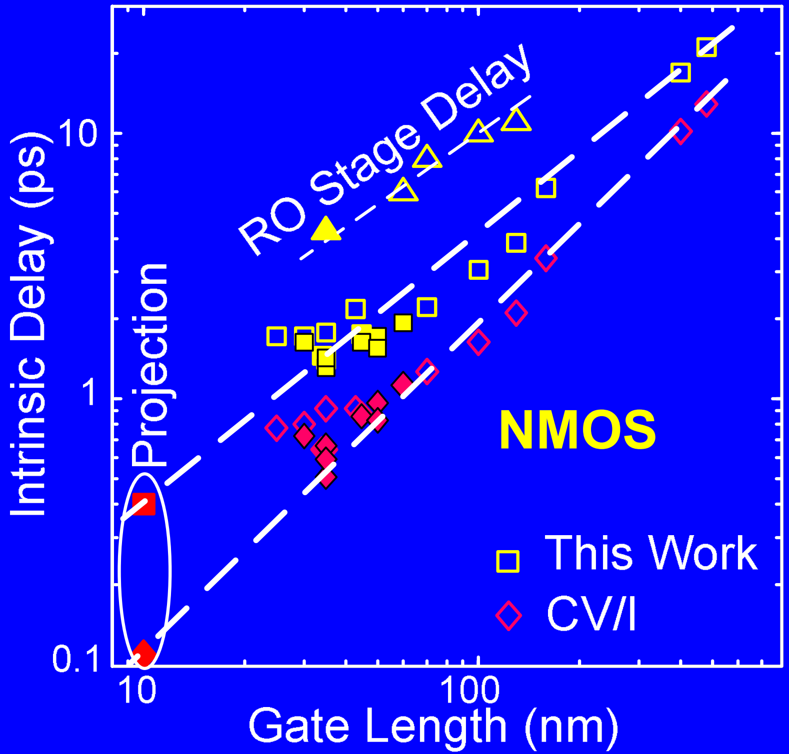
New Performance Metric vs. CV/I

- $C_{inv}V_{DD}/I_D$ does not depend on V_T and hence DIBL and does not include C_f^*.
- Our metric better follows the measured ring oscillator delay.



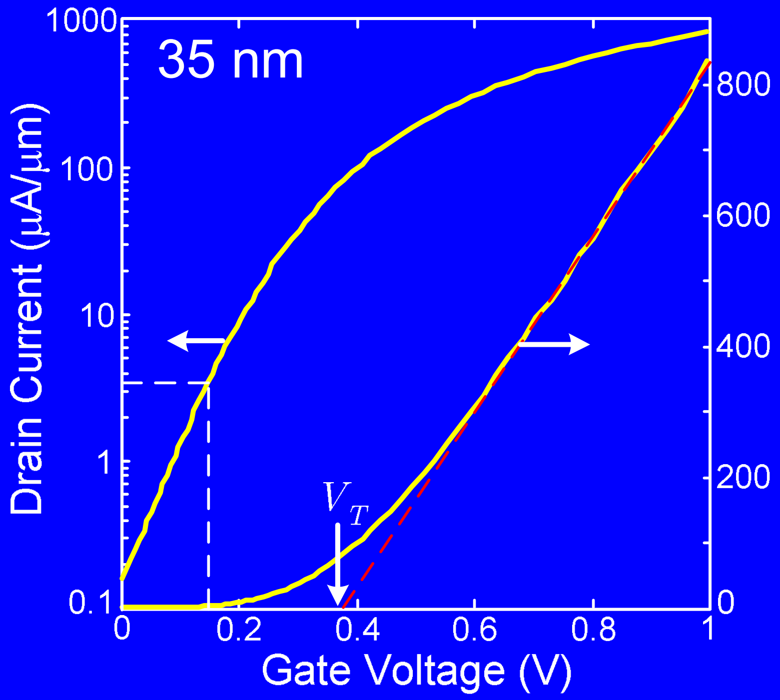
New Performance Metric vs. CV/I

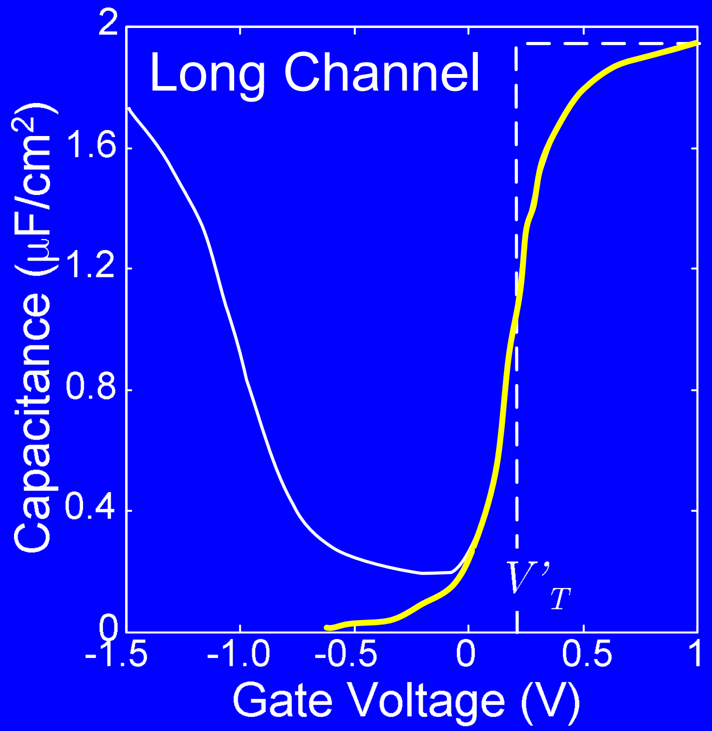
- $C_{inv}V_{DD}/I_D$ does not depend on V_T and hence DIBL and does not include C_f^*.
- Our metric better follows the measured ring oscillator delay.



Effective Velocity – How to Extract?

$$I_D = W C_{\text{inv}} (V_{GS} - V_T) v$$





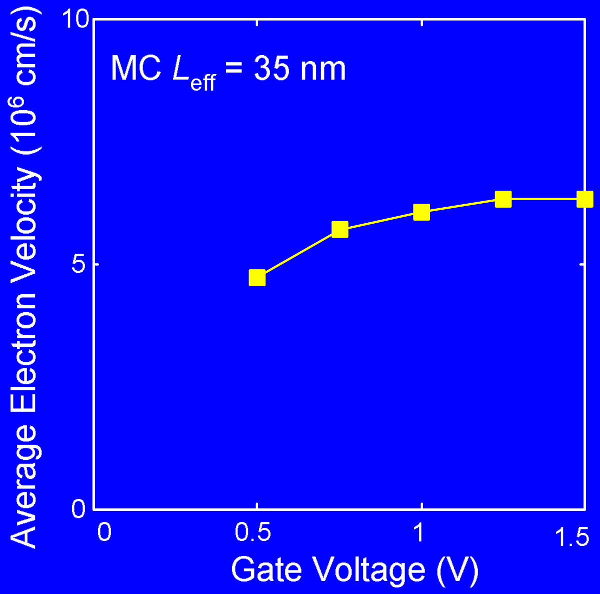
Data From: K. Adachi, *VLSI Symp.*, p. 142, 2005.

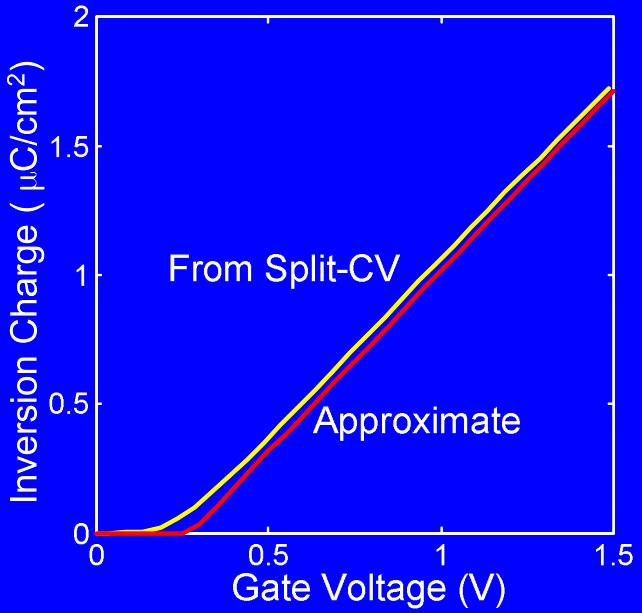
Velocity Extraction – Assumptions

$$I_D = WC_{\text{inv}}(V_{GS} - V_T)v$$

Assumptions:

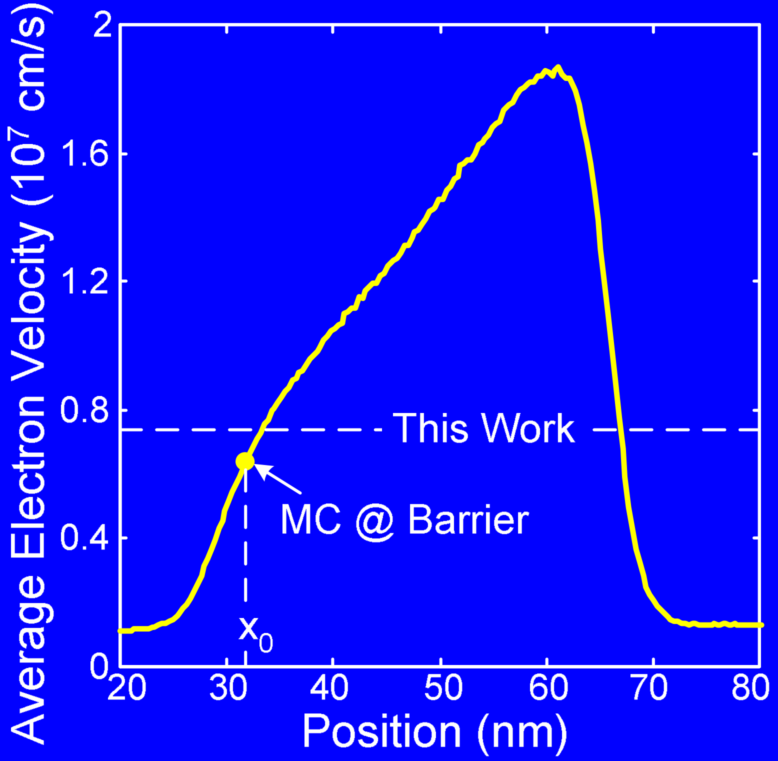
- Velocity is independent of V_G.
- Inversion charge is linear for $V_{GS} > V_T$.





Velocity – Comparison with MC

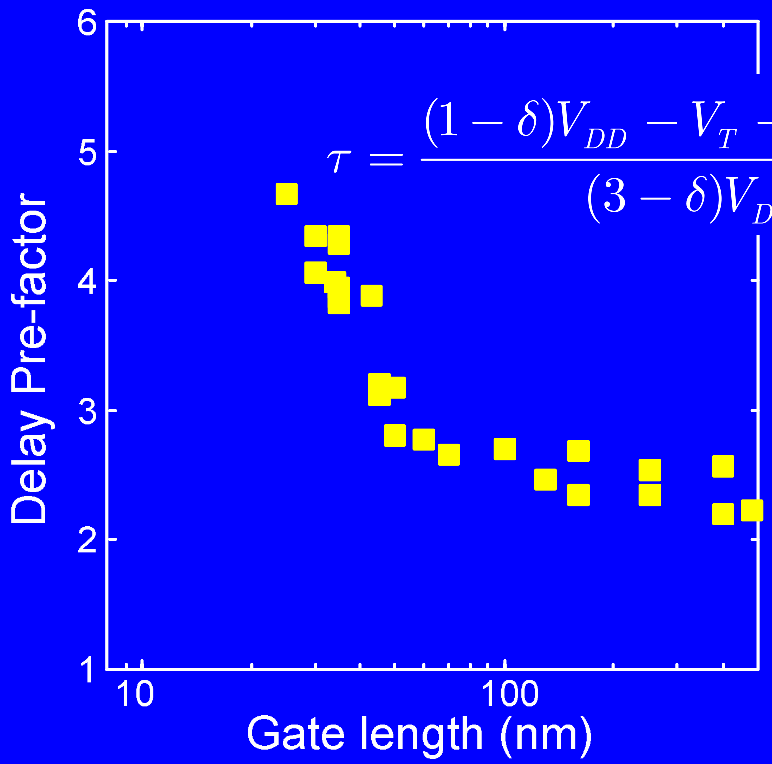
- **Roughly 10% overestimation of velocity compared to calibrated MC simulations on inverse-modeled devices.**



For details of MC see: O. M. Nayfeh, *et al.*, SISPAD, 2004.

Impact of Pre-factor

- Velocity needs to increase to cancel out the increase in the pre-factor.
- The main contributor is parasitic capacitance.

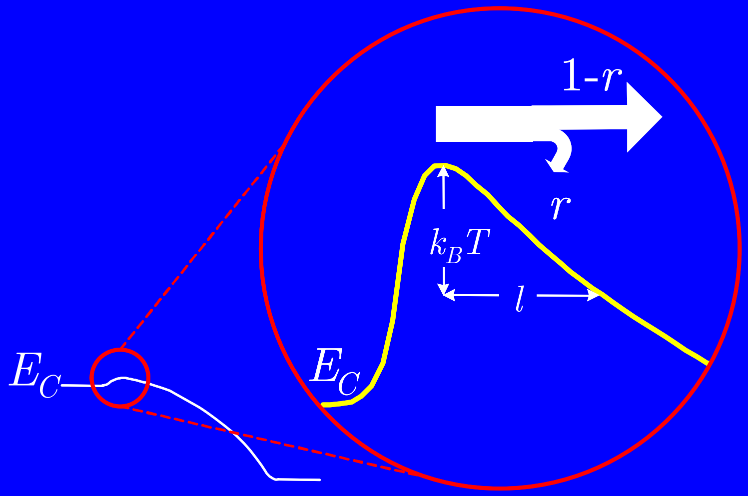


$$\tau = \frac{(1-\delta)V_{DD} - V_T + (C_f^* \,/\, C_{\text{inv}} L_G)V_{DD}}{(3-\delta)V_{DD} \,/\, 4 - V_T} \frac{L_G}{v}$$

Effective Velocity – Physical?

- **The effective velocity is related to the virtual source velocity, v_{x0}:**

$$v = \frac{v_{x0}}{1 + \underbrace{W R_s C_{\text{inv}} (1 + 2\delta) v_{x0}}_{\sim 0.2}}$$



$$v_{x0} = B v_\theta = \frac{1 - r}{1 + r} v_\theta$$

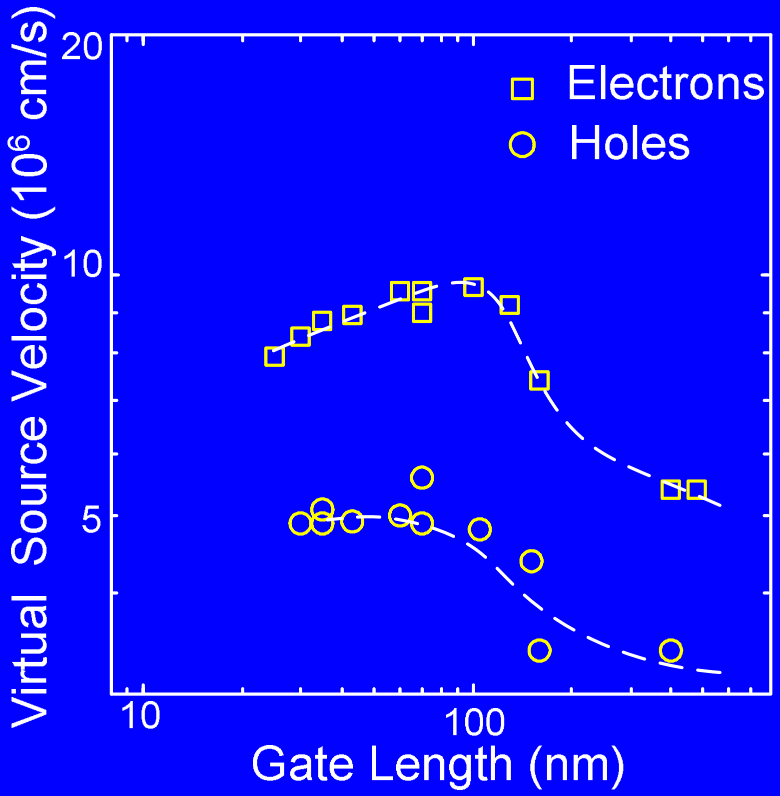
$$r = l / (\lambda + l)$$

l **critical length for backscattering**

λ **mean free path**

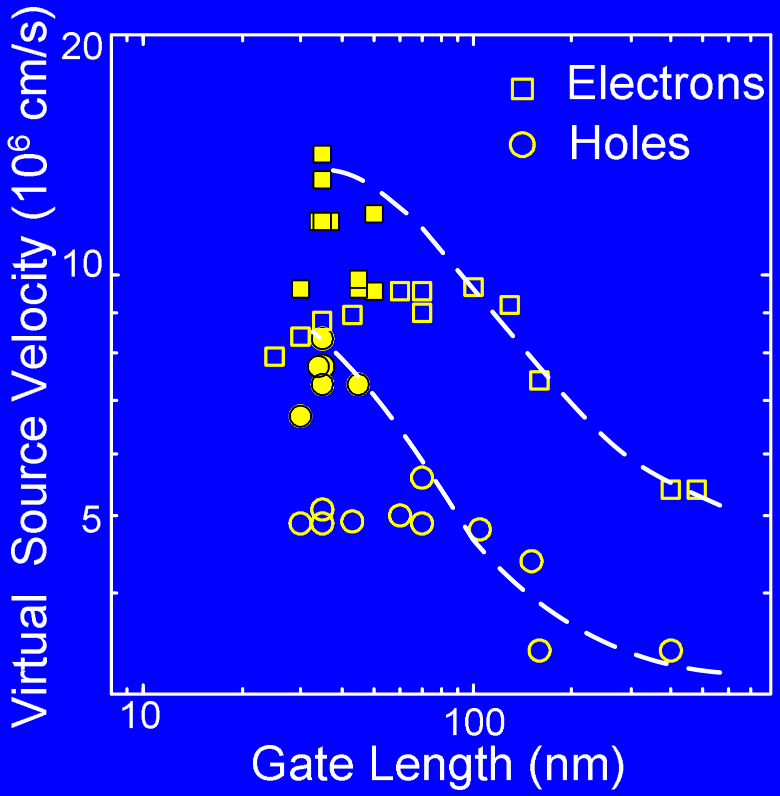
Velocity Evolution

- **Virtual source velocity significantly increased to compensate the increase in the parasitics.**



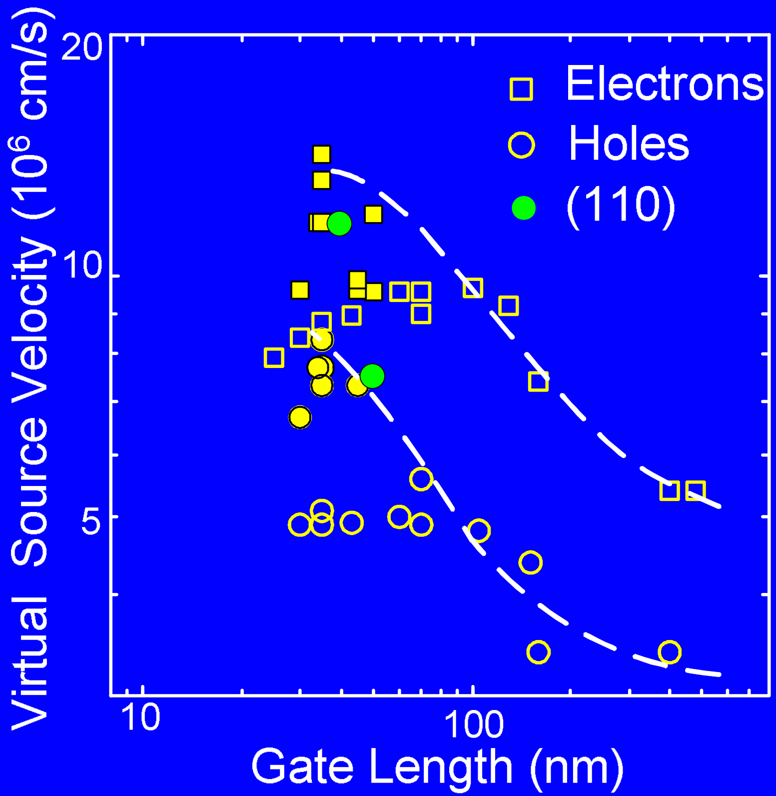
Velocity Evolution

- **Virtual source velocity significantly increased to compensate the increase in the parasitics.**



Velocity Evolution

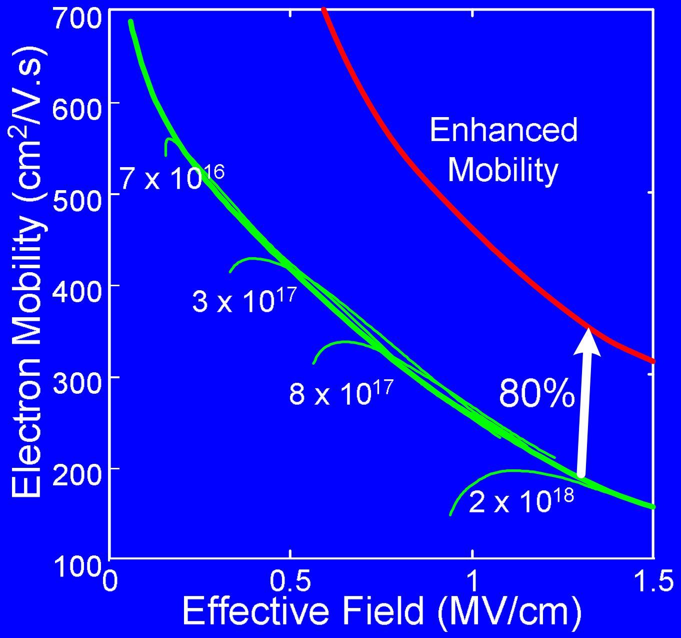
- **Virtual source velocity significantly increased to compensate the increase in the parasitics.**

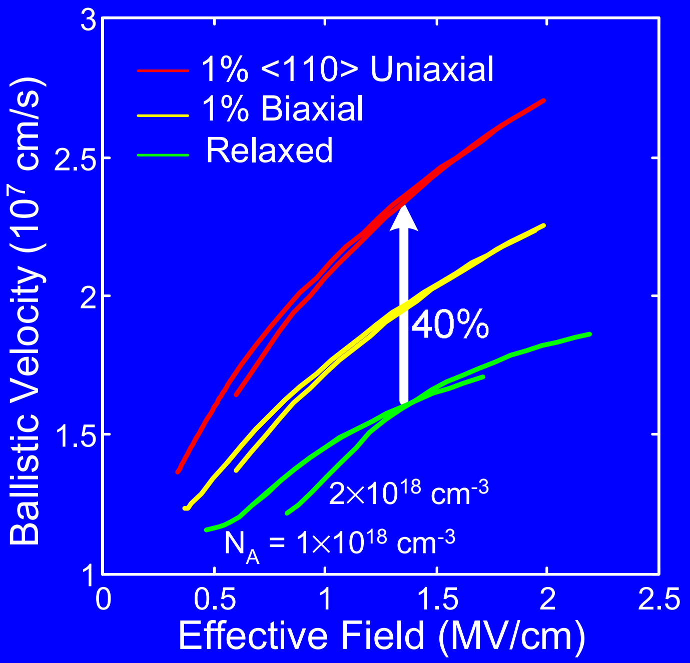


How to Increase v_{x0}?

$$v_{x0} = Bv_\theta = \frac{\lambda}{\lambda + 2l} v_\theta$$

1. **Higher B (higher λ ~ mobility)**
2. **Higher ballistic velocity.**

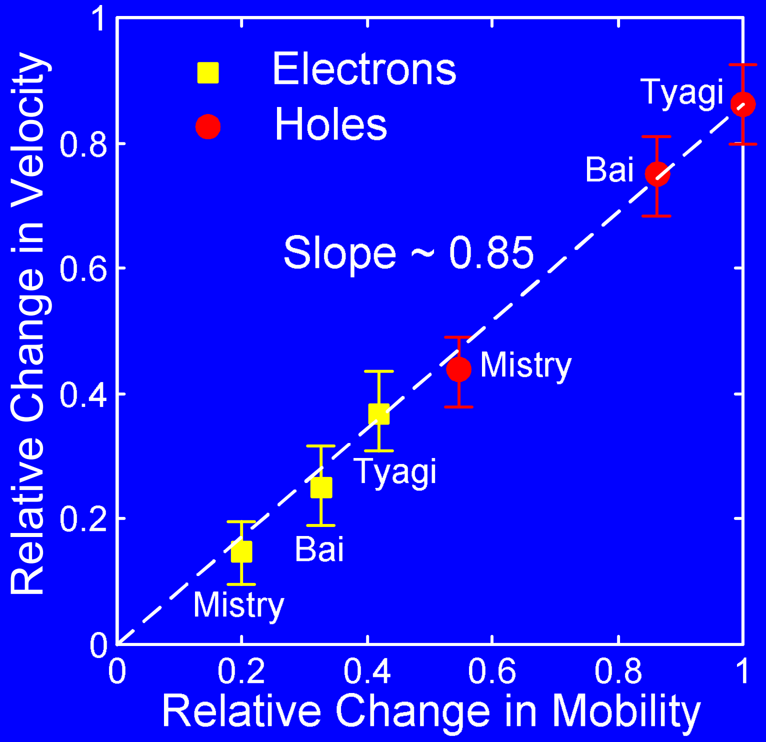




Mobility – Velocity Relationship

- **In short channel (L_G < 45 nm) devices with process-induced strain, velocity depends on mobility more strongly than previously thought.**

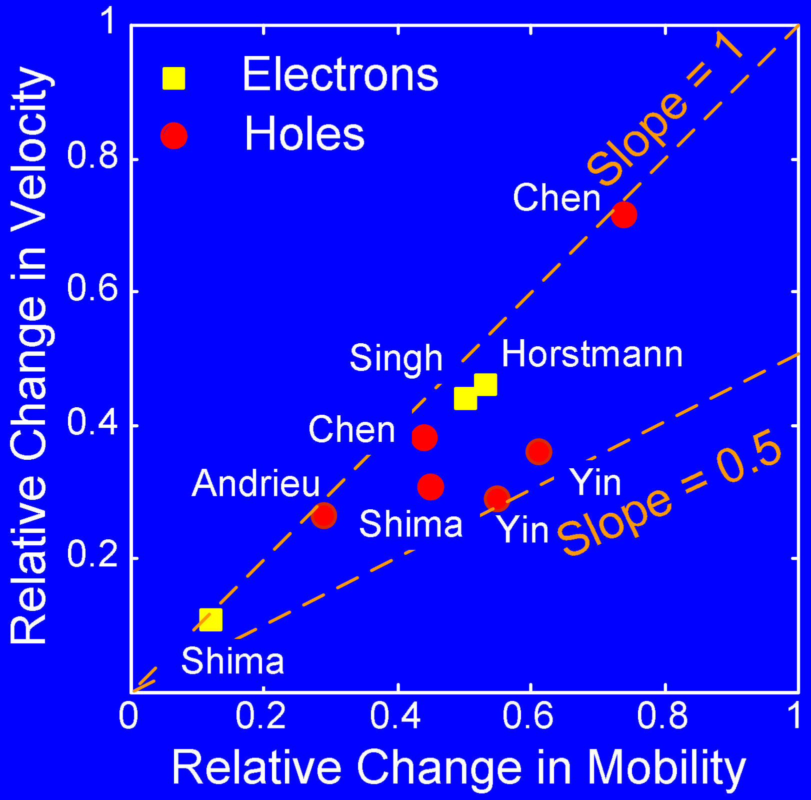
$$\frac{\partial v_{x0}}{v_{x0}} \Big/ \frac{\partial \mu}{\mu} \simeq 0.85$$



Mobility data from: S. Tyagi, *et al.*, *IEDM*, p. 1070, 2005.

Mobility – Velocity Relationship

- **Similar results are obtained for other strain-engineered devices.**

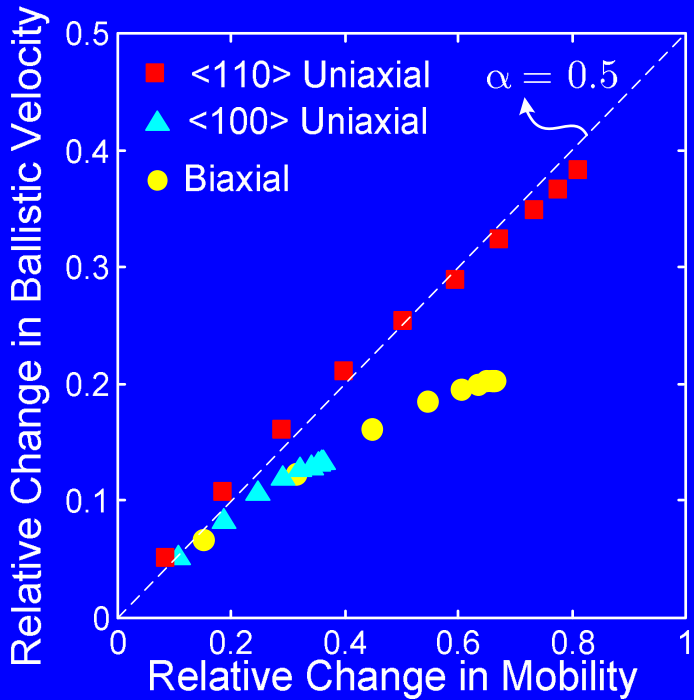


L_G < 50 nm

* Mobility enhancement deduced from the slope of R_{tot}-L_G plots.

Velocity – Mobility: Theory

- In uniaxially strained Si, mobility enhancement is mostly due to reduction in the effective mass.



$$\mu \propto \frac{1}{m_C}\frac{1}{m_D}$$

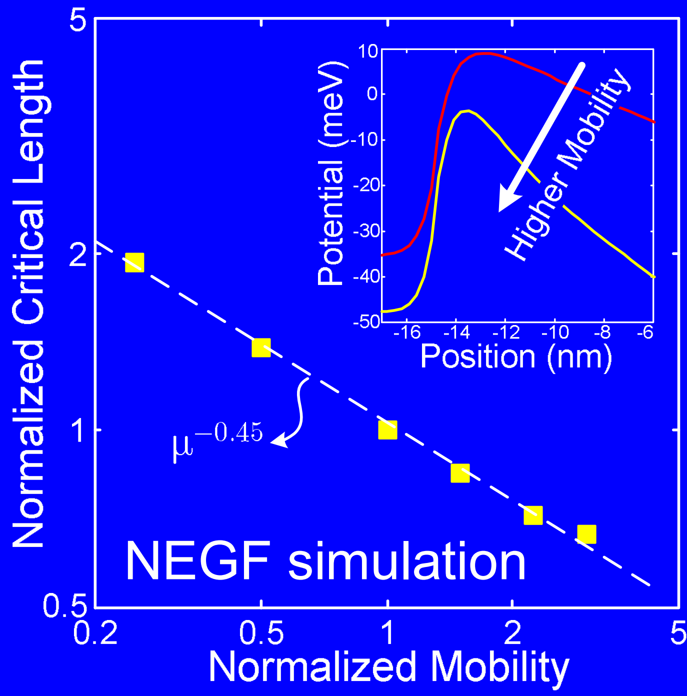
$$v_\theta \propto \frac{1}{\sqrt{m_C}}\frac{1}{\sqrt{m_D}}$$

$$v_\theta \propto \mu^{0.5}$$

K. Uchida, *et al.*, *IEDM*, p. 135, 2005.

Velocity – Mobility: Theory

- **With steeper velocity increase in the channel, charge distribution drops more steeply in the channel. Potential profile changes accordingly to accommodate for this steeper charge drop.**



$$l \propto \mu^{-0.45}$$

Velocity – Mobility: Theory

$$v_{x0} = B v_\theta = \frac{\lambda}{\lambda + 2l} v_\theta$$

$$\frac{\partial v_{x0}}{v_{x0}} = [\alpha + (1 - B)(1 - \alpha + \beta)] \frac{\partial \mu}{\mu}$$

Compare to Lundstrom's scattering theory*:

$$\frac{\partial v_{x0}}{v_{x0}} = (1 - B) \frac{\partial \mu}{\mu}$$

* M. Lundstrom, *IEEE EDL*, p. 293, 2001.

How Close to Ballistic?

- **Mobility – velocity relation gives an alternative way to estimate ballistic efficiency, B :**

$$\frac{\partial v_{x0}}{v_{x0}} = [\alpha + (1 - B)(1 - \alpha + \beta)]\frac{\partial \mu}{\mu}$$

For Intel L_G = 35 nm strain-engineered devices*

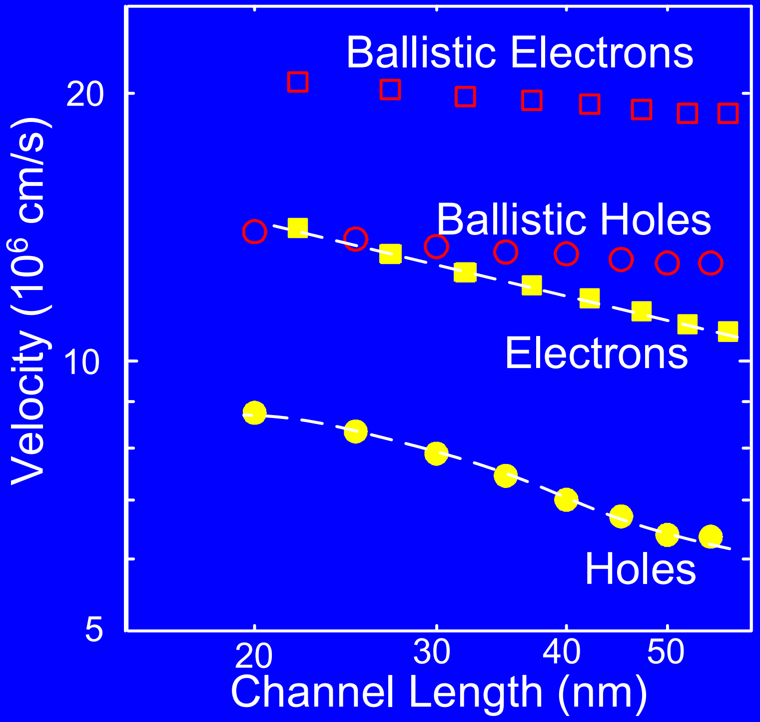
$$\alpha \approx 0.5 \qquad \beta \approx -0.45 \qquad \partial v_{x0} / v_{x0} \approx 0.85\,\partial \mu / \mu$$

$$\boxed{B \approx 0.65}$$

* S. Tyagi, *et al.*, *IEDM*, p. 1070, 2005.

How Close to Ballistic?

- **With calculated ballistic velocity at similar E_{eff} and strain level, ballistic efficiency is directly estimated as 60 – 65%.**



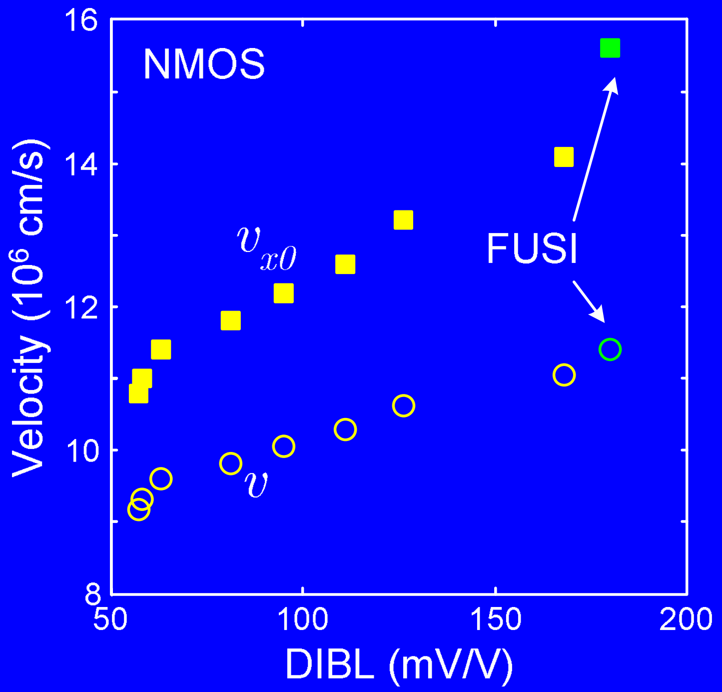
Ballistic velocity calculated at E_{eff} and strain estimated for each device.

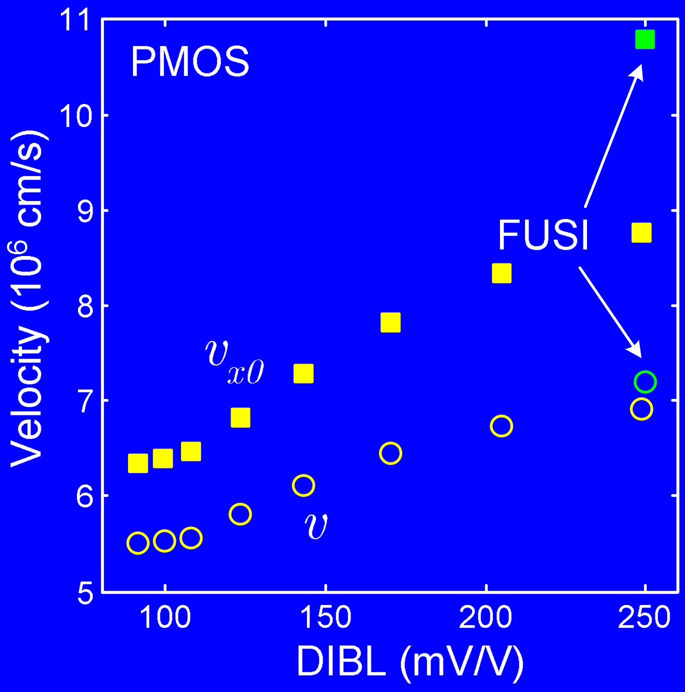
Change in the effective mass was calculated based on the experimental mobility data.

Devices from S. Tyagi, *et al.*, *IEDM*, p. 1070, 2005.

Impact of Metal Gate

- **With less doping in the channel, v_{x0} is higher. However, effective velocity, v, is comparable to poly gate devices because of higher C_{inv}.**

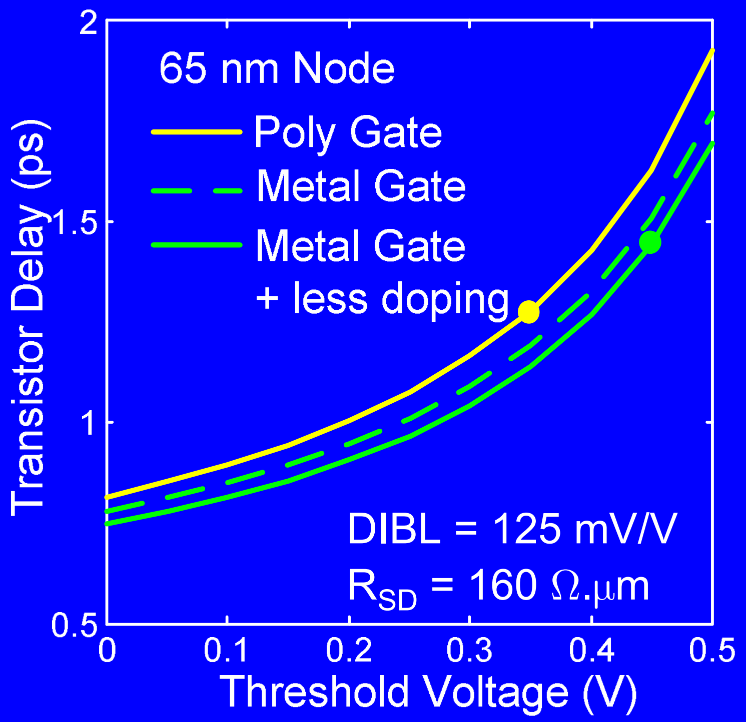




Data from: P. Ranade (IEDM'05), S. Tyagi (IEDM'05)

Impact of Metal Gate

- **Slight improvement in velocity is canceled out by high DIBL and inappropriate threshold voltage.**

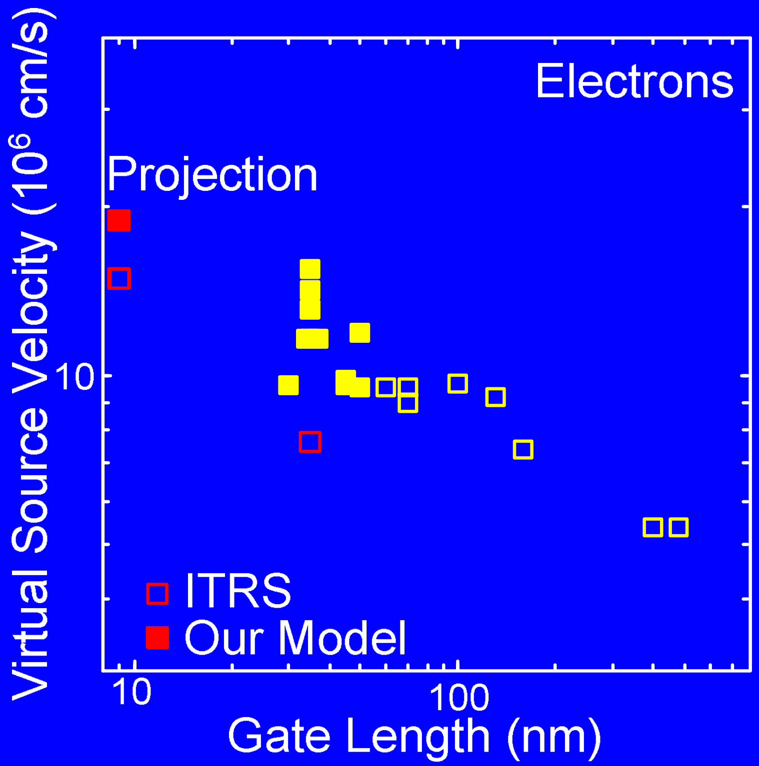


Data points from: P. Ranade (IEDM’05), S. Tyagi (IEDM’05)

Future Directions (ITRS L_G~18-9 nm*)

$v_{x0} = 1.9\times10^7$ cm/s with optimistic ITRS R_{SD}.

Requires $v_\theta = 2.9\times10^7$ cm/s with B = 0.65!



Projections with:

L_G ITRS

T_{inv} = 1 nm

R_{SD} = 80 Ω.μm

I_{off} = 300 nA/μm

V_{DD} ITRS

DIBL = 100 mV/V

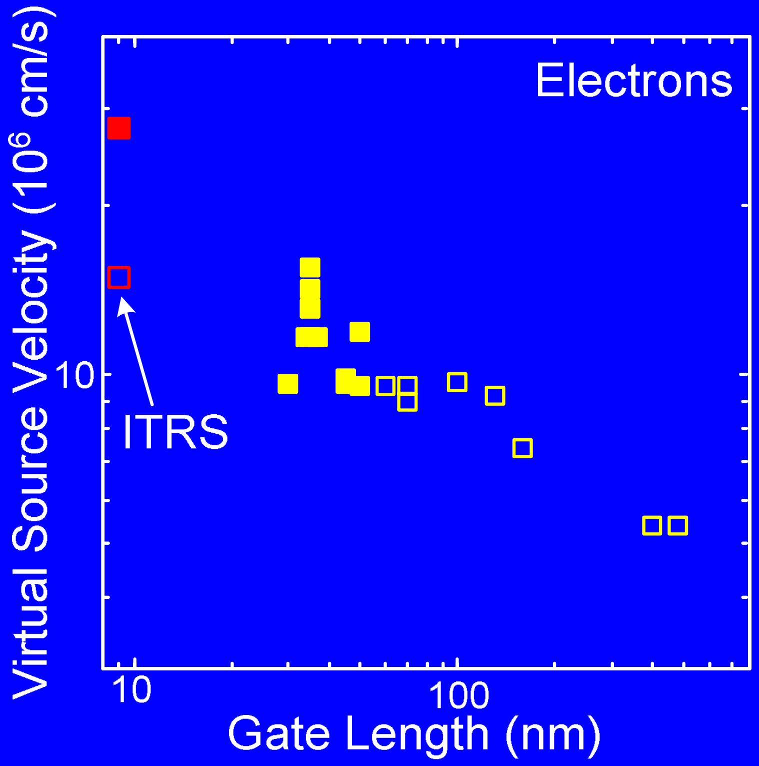
S = 100 mV/dec

C_f* = 0.5 fF/μm

* Note: ITRS – L_G projections will likely change.

Future Directions (NFET)

$v_{x0} = 2.7\times10^7$ cm/s if R_{SD} is not scaled below its current value (~160 Ω.μm).



Projections with:

L_G ITRS

T_{inv} = 1 nm

R_{SD} = 160 Ω.μm

I_{off} = 300 nA/μm

V_{DD} ITRS

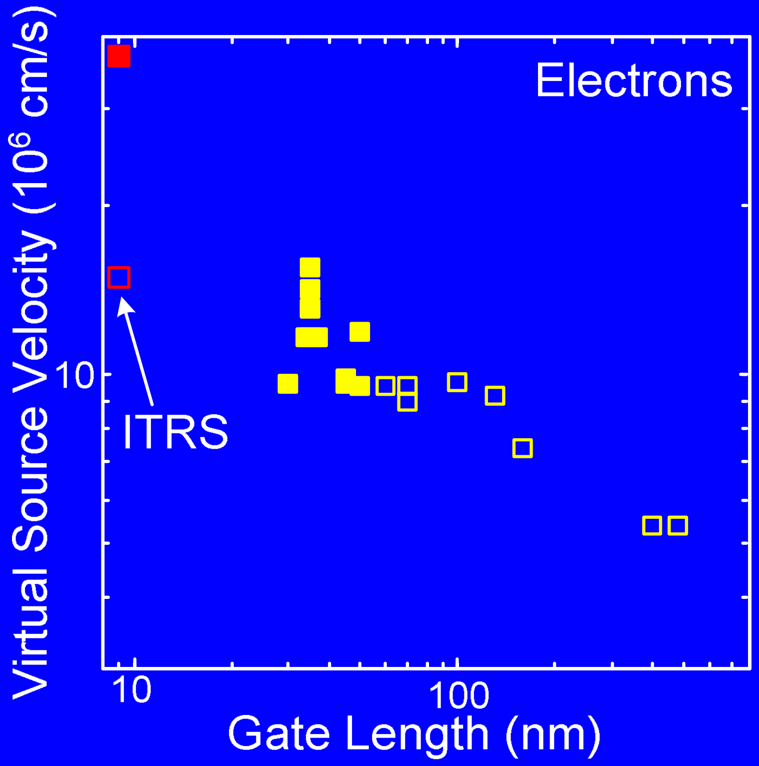
DIBL = 100 mV/V

S = 100 mV/dec

C_f^* = 0.5 fF/μm

Future Directions (NFET)

v_{x0} = 3.7×10^7 cm/s with realistic short channel effects. This is out of reach of strained Si.



Projections with:

L_G ITRS

T_{inv} = 1 nm

R_{SD} = 160 Ω.μm

I_{off} = 300 nA/μm

V_{DD} ITRS

DIBL = 130 mV/V

S = 130 mV/dec

C_f^* = 0.5 fF/μm

Conclusions

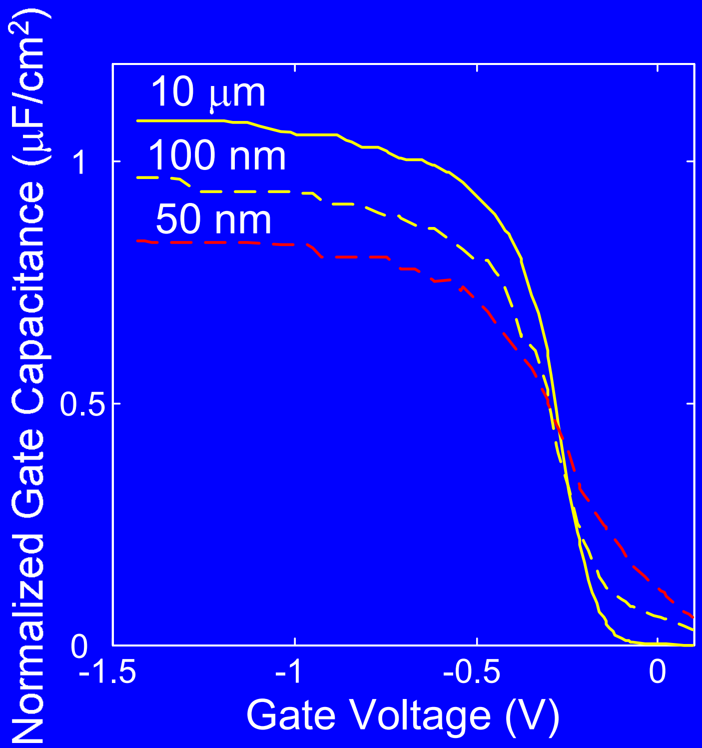
- A new transistor performance metric was defined as the intrinsic MOSFET delay.
- The traditional CV/I metric is overly optimistic as it misses effects of fringing cap and V_T.
- Process-induced strain introduced in time for intrinsic delay continuous decrease with L_G.
- ITRS'05 velocity prediction is optimistic as it relies on incorrect V_T, R_{SD}, and especially C_f^*.
- High mobility and velocity enhancement, to follow historical performance scaling trend, will require new channel materials.

Acknowledgements

- Osama Nayfeh (MIT)
- Scott Yu (TI)
- Steve Laux (IBM)
- Hassan Nayfeh (IBM)
- Pushkar Ranade (Intel)
- Tahir Ghani (Intel)
- François Andrieu (CEA-LETI)
- Haizhou Yin (IBM)
- Xiangdong Chen (IBM)

Charge Estimation – How Accurate?

- **C-V characteristics stretches out due to short channel effects.**



Corrections needed:
- **DIBL**
- **V_T roll-off**
- **Series resistance**
- **Subthreshold slope**
- **Poly depletion**

Data From: K. Romanjek, *et al.*, *IEEE EDL*, p. 583, 2004.