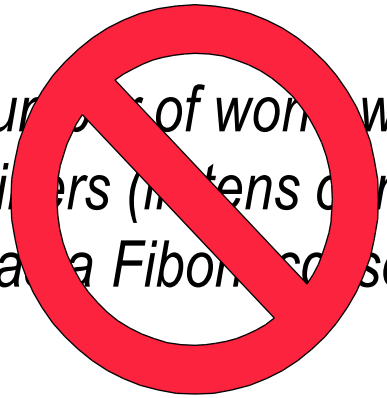
Tutorial Hot Chips 01

Silicon Architectures for Wireless Systems – Part 1

Jan M. Rabaey
BWRC
University of California @ Berkeley
http://www.eecs.berkeley.edu/~jan

Berkeley Wireless Research Center

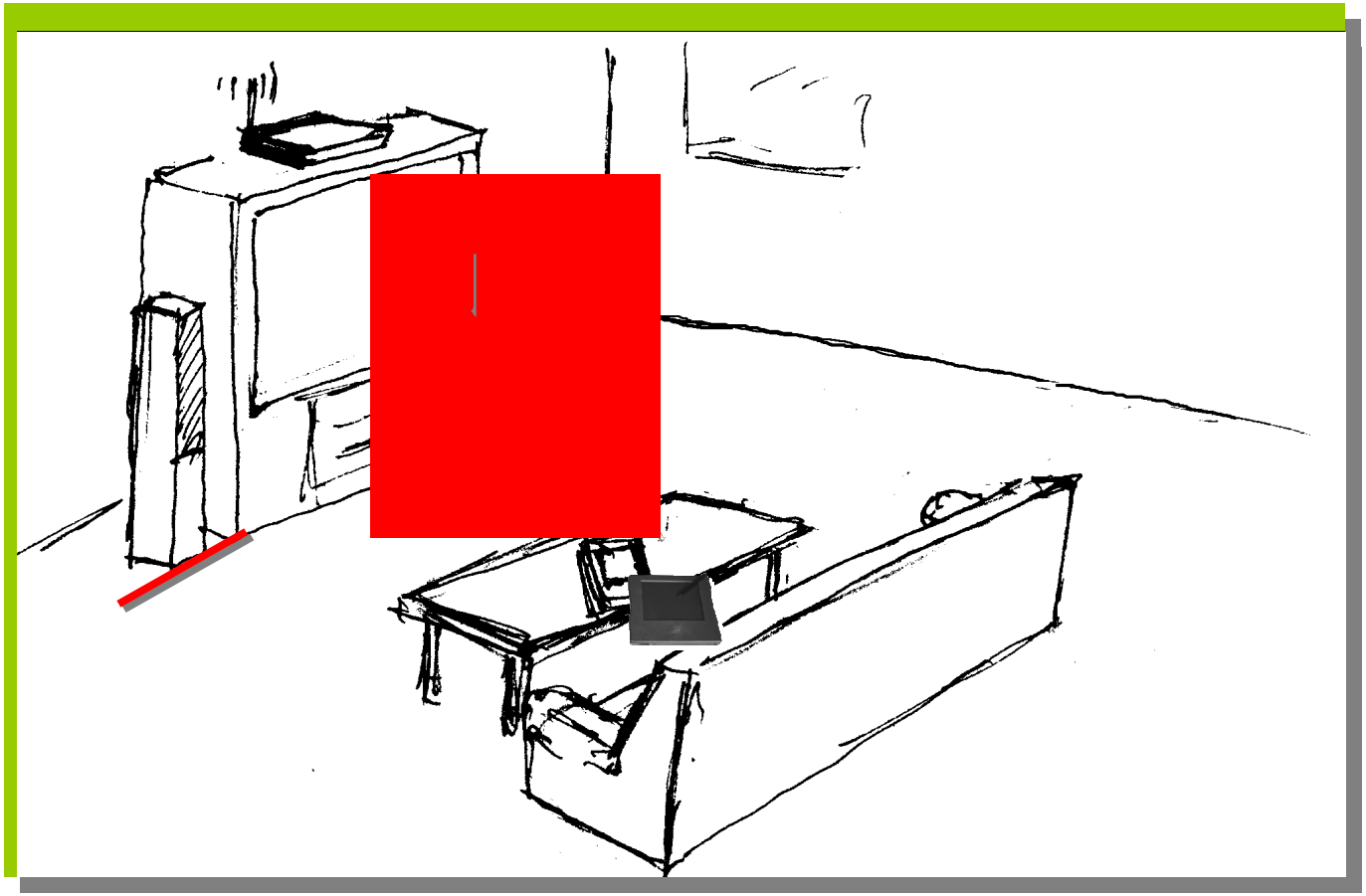
The Fibonacci Law on Wireless Growth



"The nu[illegible] of wor[illegible]wide wireless subscri[illegible]ers (i[illegible]tens o[illegible]millions) grows a[illegible]a Fibo[illegible]series"

Berkeley Wireless Research Center

From Handsets to Mobile Devices

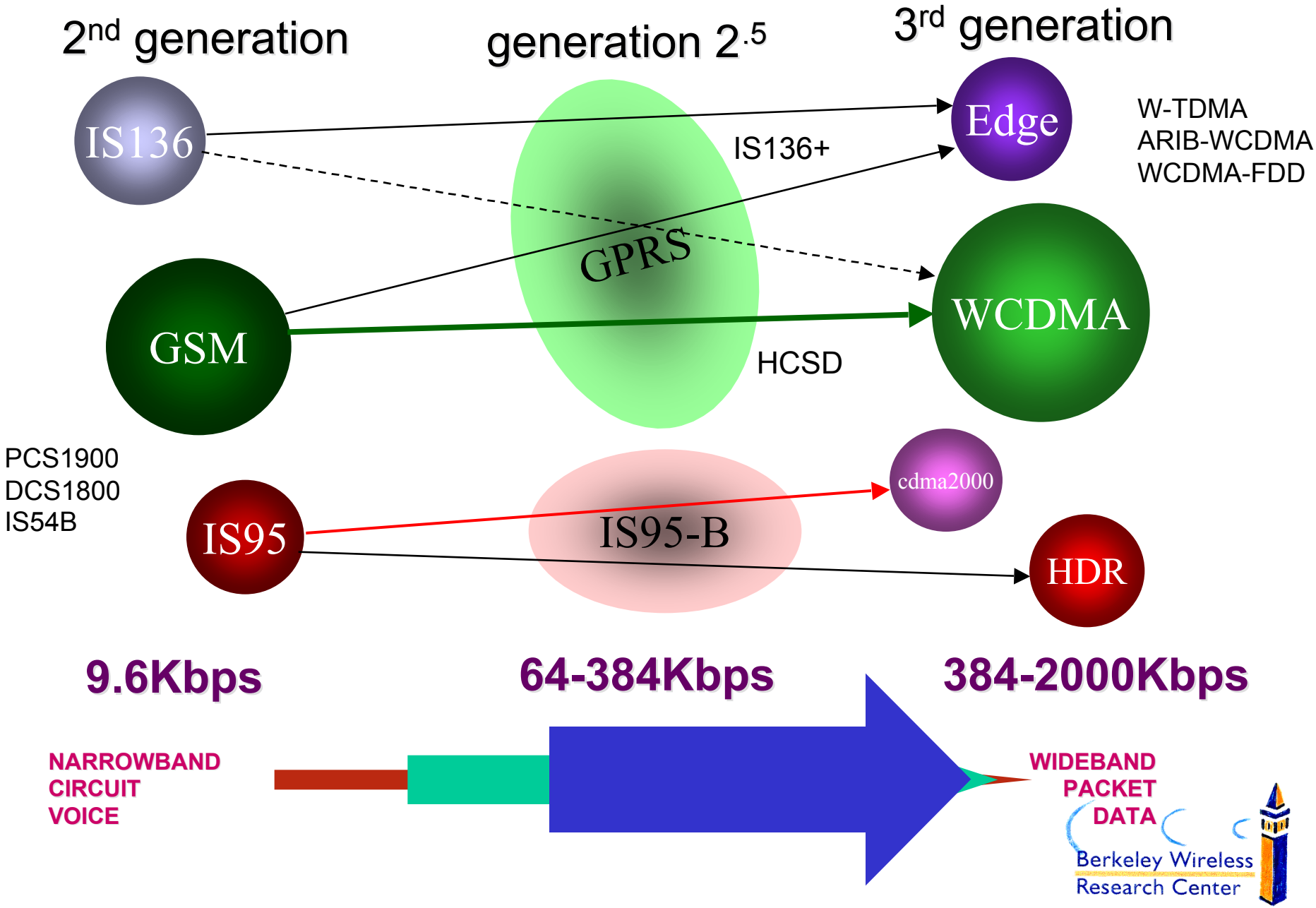


**Berkeley Infopad,
One of the first wireless
internet applicance
1990-1996**

**Internet access the most important driver
(text, graphics, multimedia)**

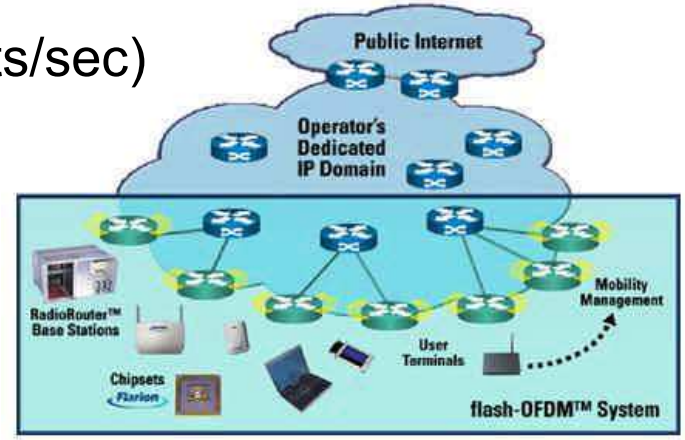
Berkeley Wireless
Research Center

A Smorgasboard of Choices



And the Alternatives

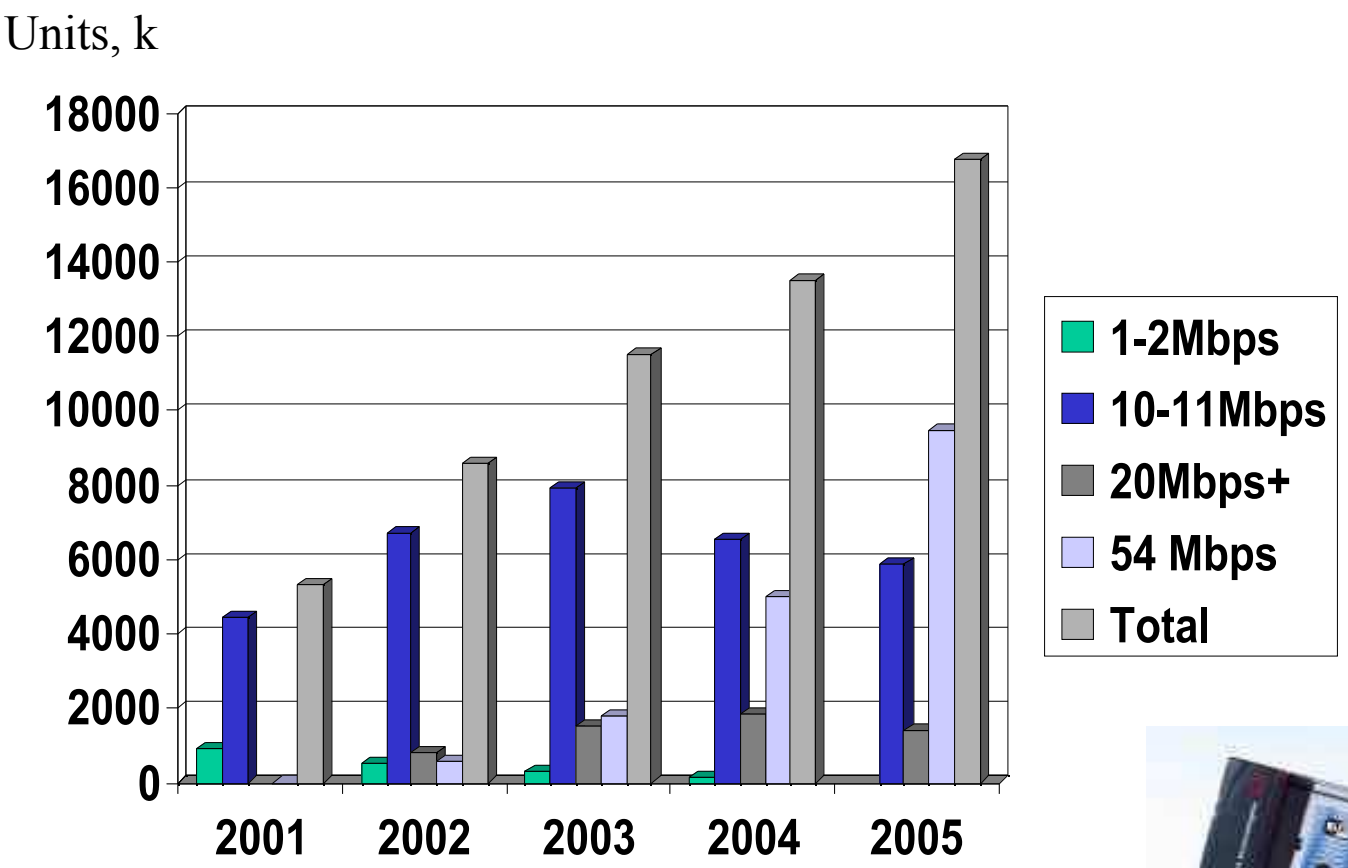
- **Metropolitan Wireless Networks**
  - Various proprietary solutions
    - Ricochet (up to 138 Kb/sec)
    - Flarion Flash-OFDM (> 384 kBits/sec)



- **Wireless LANs**
  - 802.11 (b,a); Hyperlan
  - from 1 to 56 Mbit/sec
  - Restricted to the 50 meter range (at present)

Research Center

(Projected) Growth in 802.11 WLAN

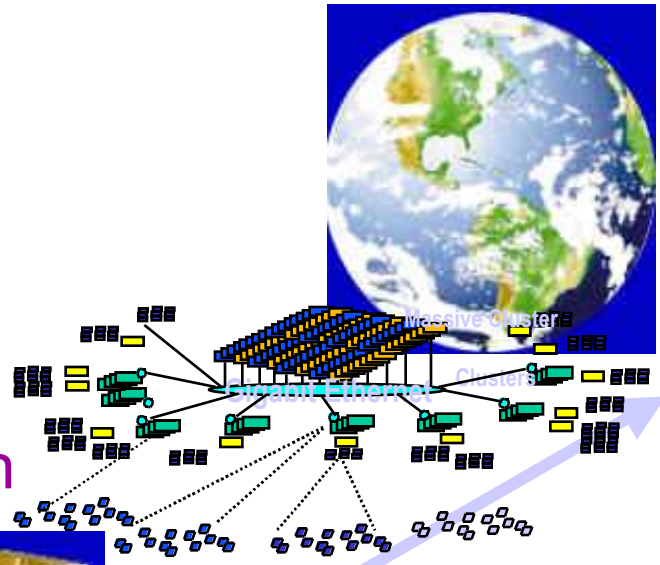


Source: Cahners In-Stat 2001

Berkeley Wireless Research Center

The New Internet

The 1990s:
Conquering the World
The Network revolution

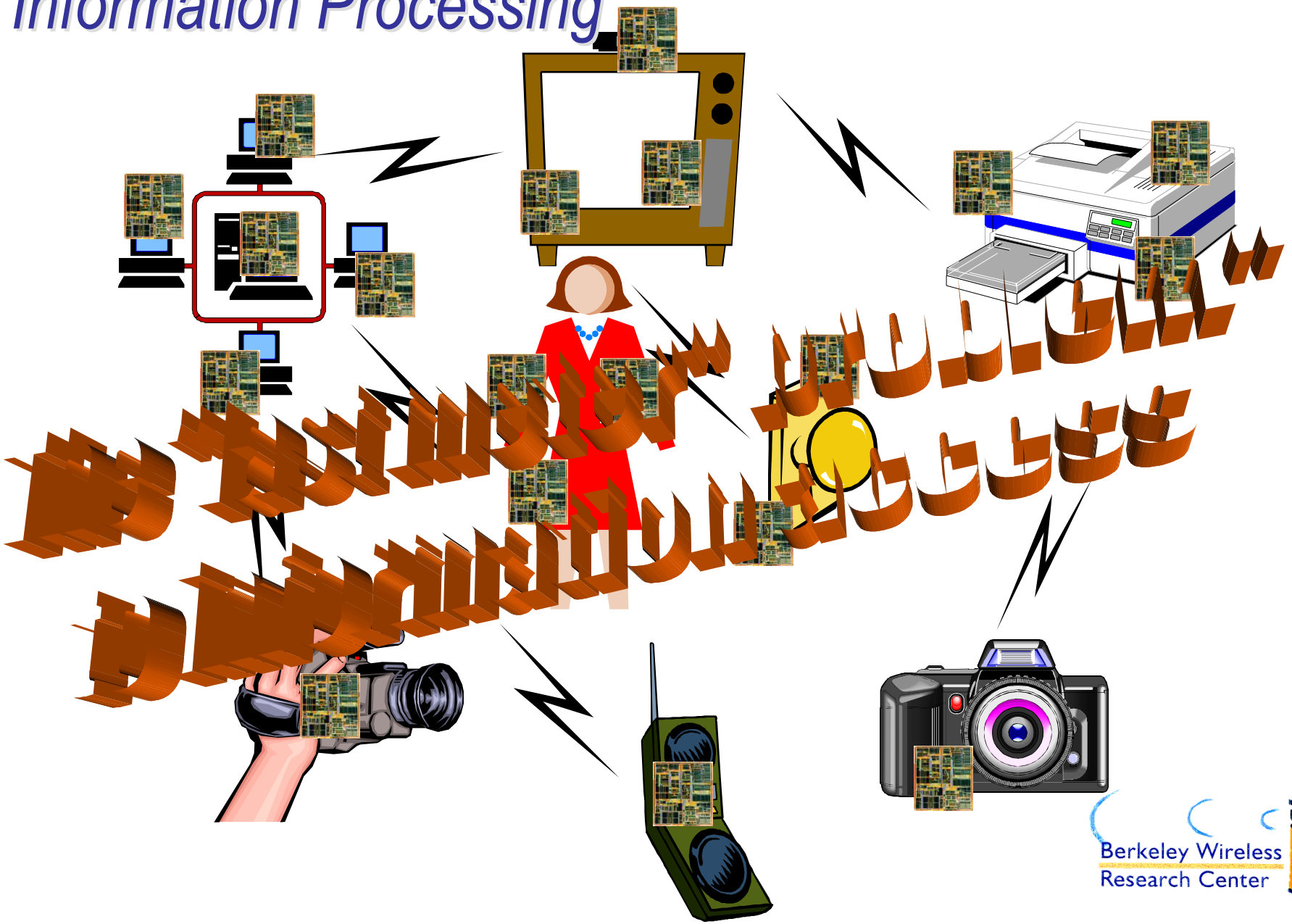


The 2000s:
Extending toward the Small
Enabled by integration
and wireless connectivity

Pre-1990:
Client-Server Systems

Berkeley Wireless
Research Center

The Post-PC Era: The Distributed Approach to Information Processing

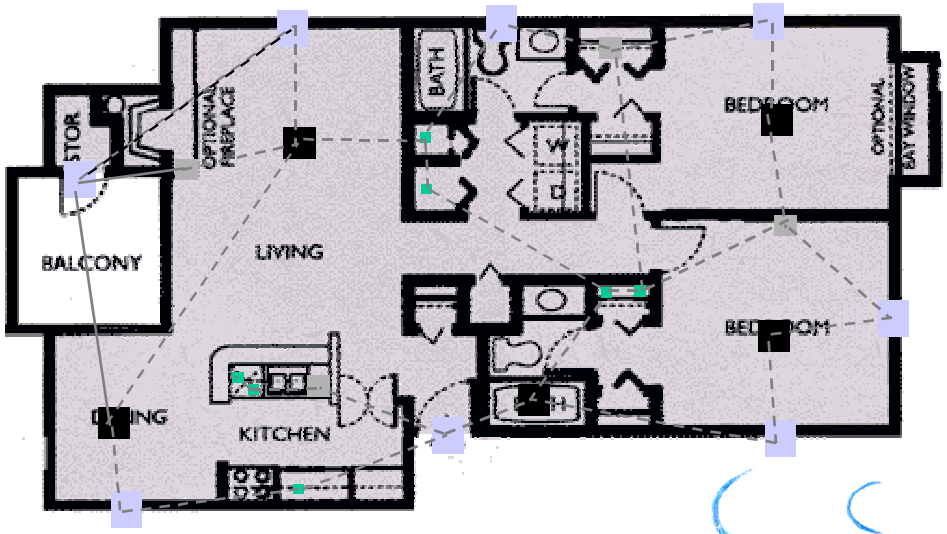


Berkeley Wireless Research Center

The emergence of ad-hoc wireless networks – the wire replacement

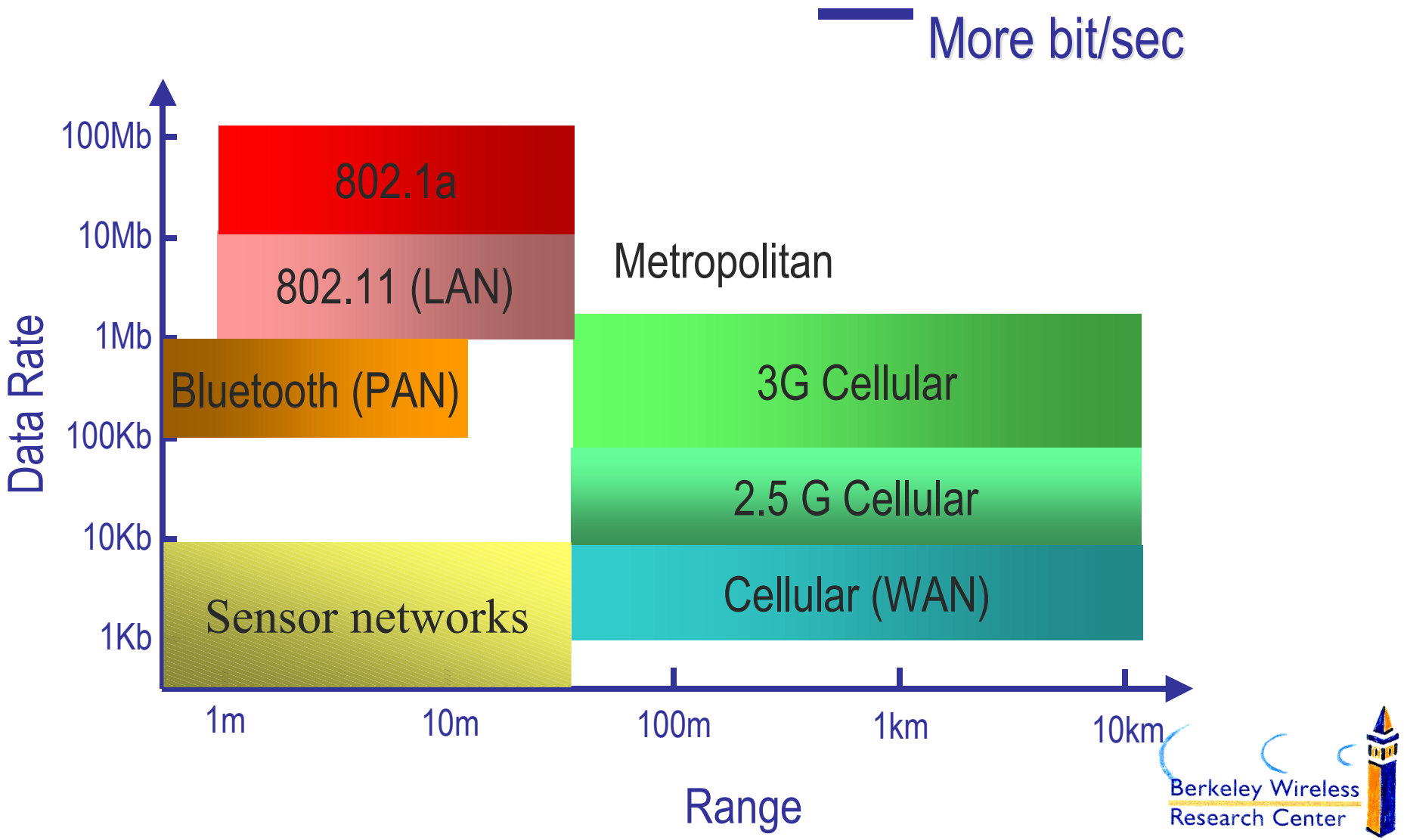
- Bluetooth, HomeRF
  - ✓ up to 800 kBit/sec

- Sensor networks
  - ✓ Low data-rates



Berkeley Wireless Research Center

The Evolving Wireless Scene



Compelling Issues in Wireless (1)

Ubiquitous services put wireless spectrum at a premium

- **Effective use of aether hampered by standardization and fragmentation**
- **Current spectral efficiency far below theoretical limits**
- **Emerging Solutions**
  - Adoption of better spectrum utilization techniques (interference cancellation, multi-path fading mitigation and exploitation)
  - multi-functional, adaptive systems
- **But … huge appetite for computations**

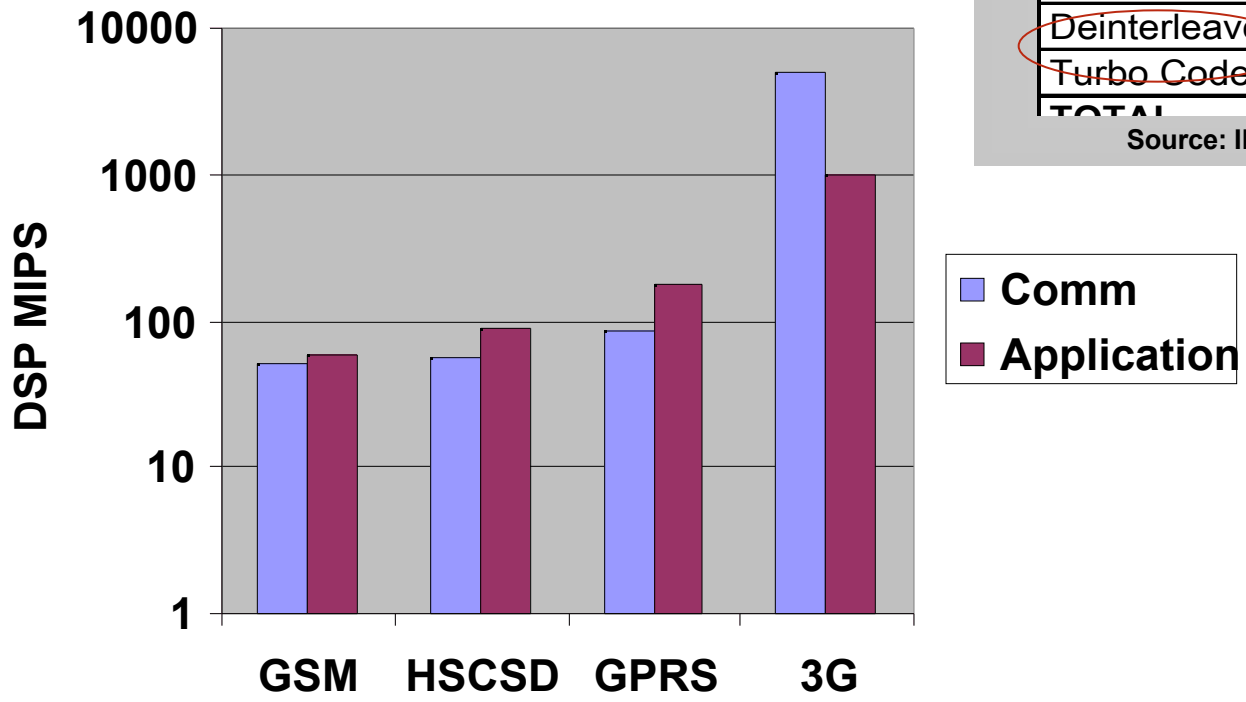
Berkeley Wireless Research Center

Evolution of MOPS Requirements in Cellular

Single 384 kbps UTRA W-CDMA Channel

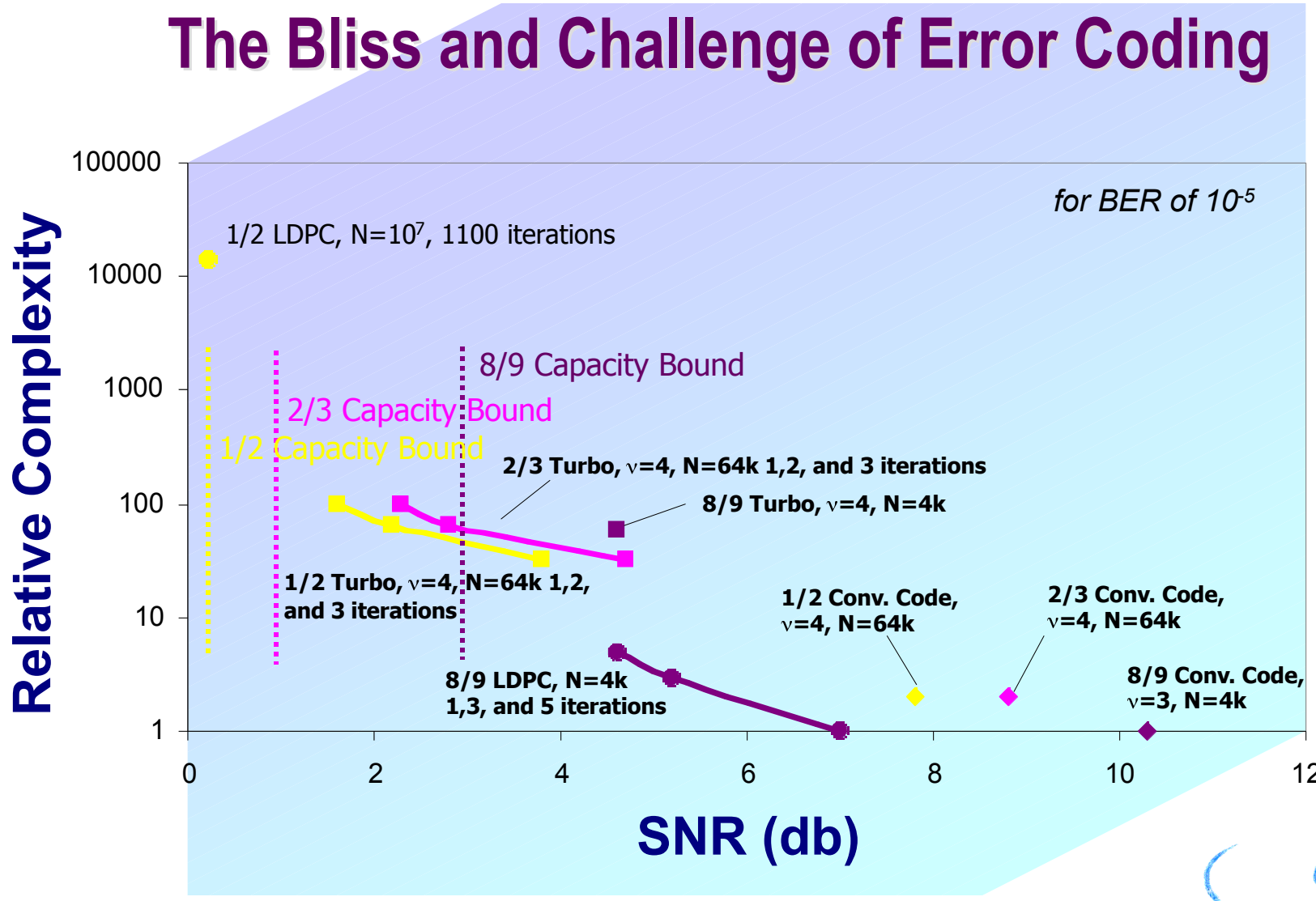
Function	MOPS
Digital RRC Channel	3600
Searcher	2100
RAKE	2050
Maximal Ratio Combiner	24
Channel Estimator	12
AGC, AFC	10
Deinterleaver	15
Turbo Coder	90
TOTAL	7901

Source: IEEE Comm Theory Workshop, May 1999



Berkeley Wireless Research Center

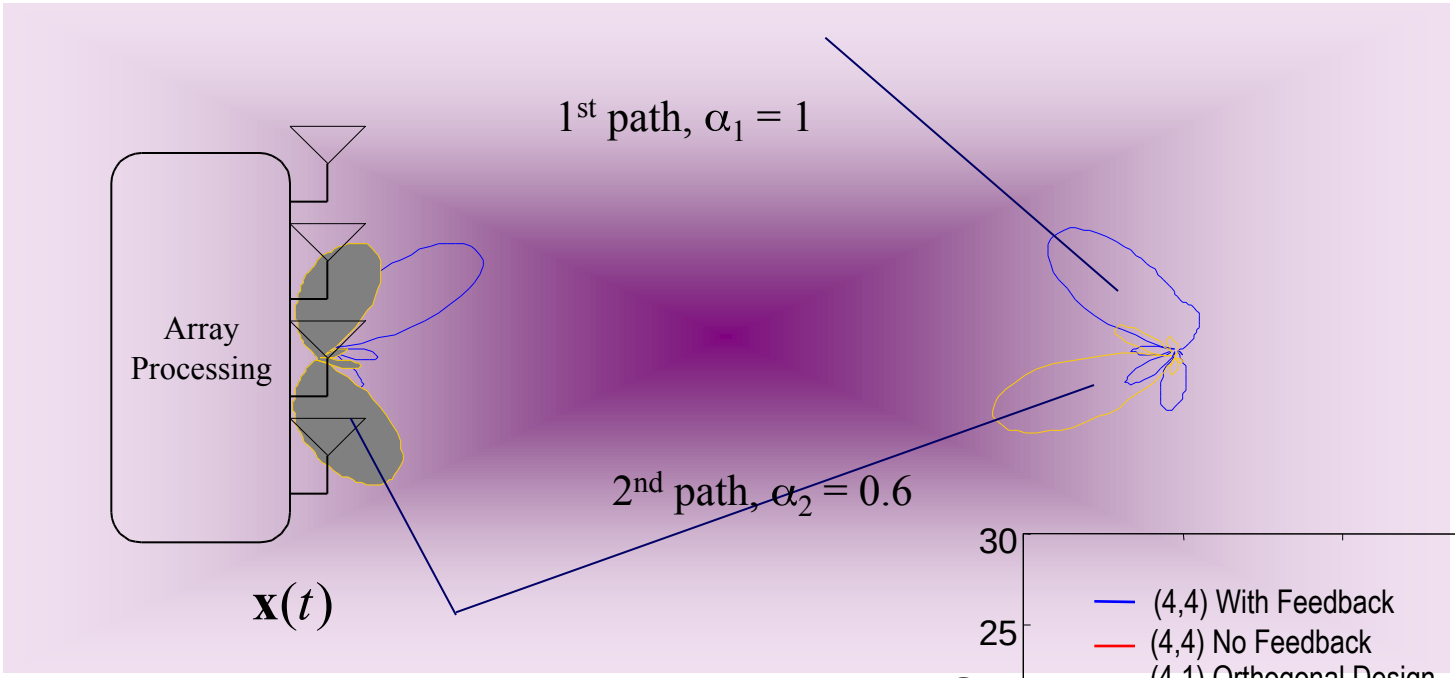
The Cost of Approaching Shannon's Bound



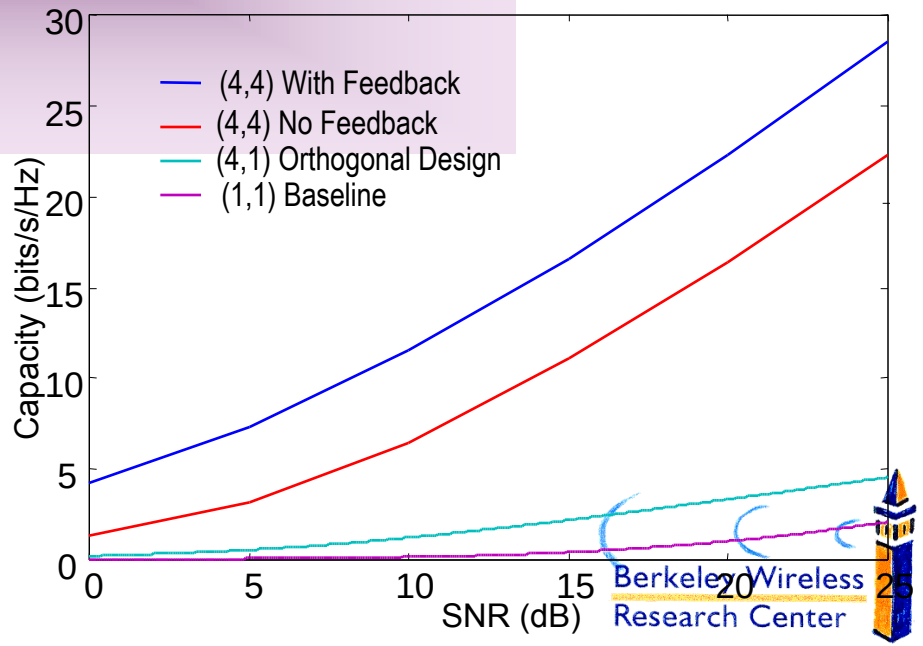
Courtesy Engling Yeo, UCB

Berkeley Wireless Research Center

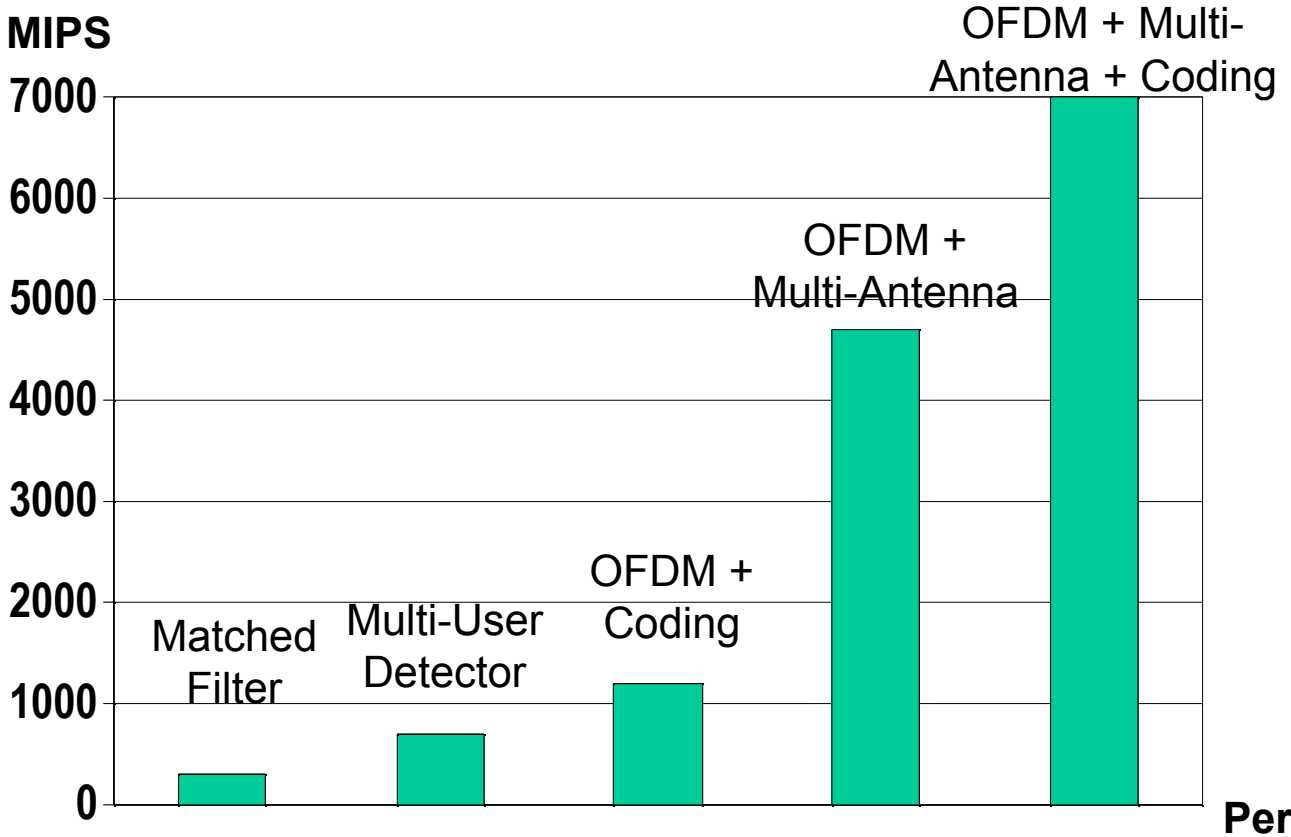
Dealing with Non-ideal Channels (e.g., fading)



- **Multi-antenna approach exploits multi-path fading by sending data along good channels**
- **Results in large theoretical improvements in bandwidth efficiency for fading channels**
- **But…computationally hungry**



The Cost of Dealing with Non-ideal Channels



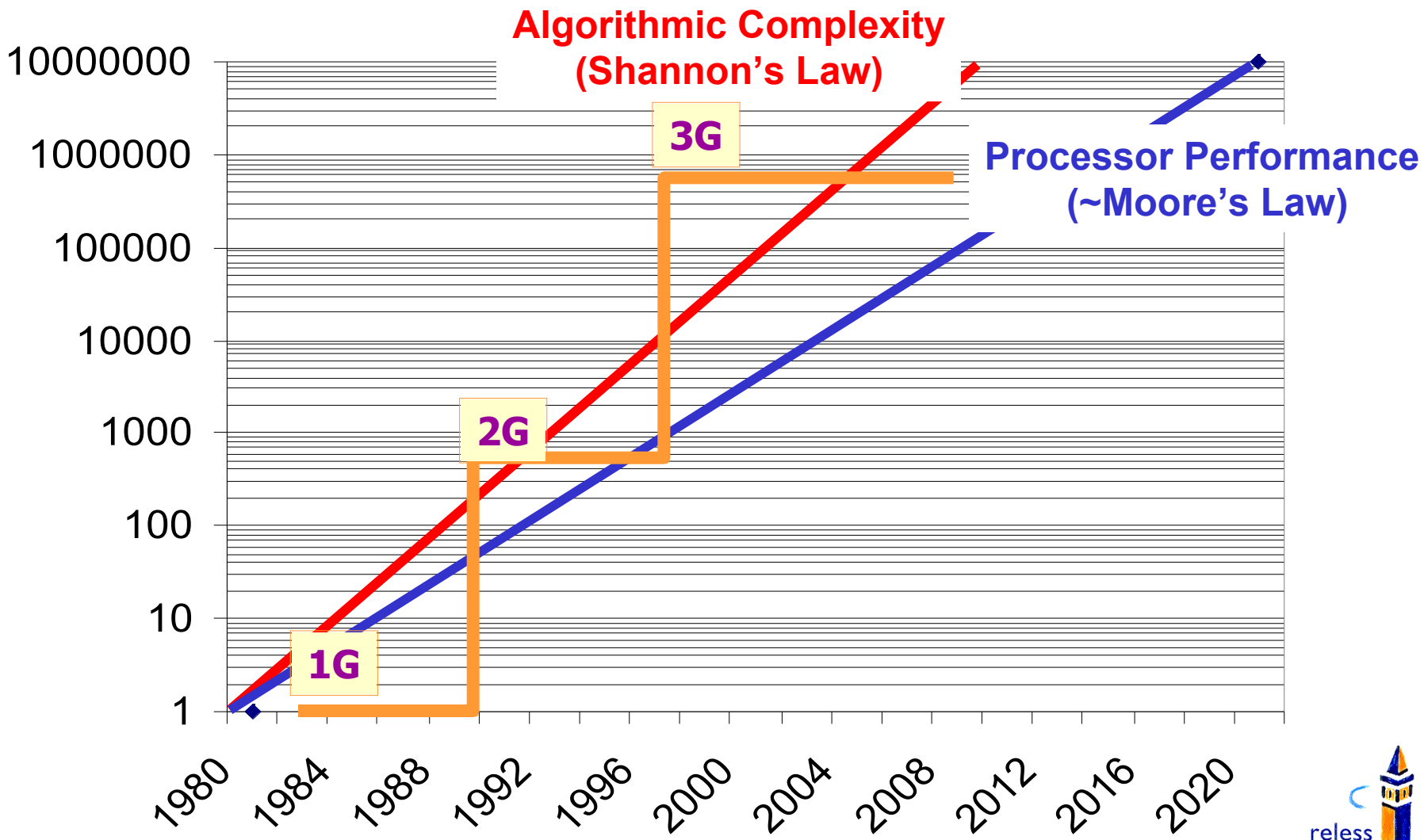
Data rate per user	0.8 Mb/s	1.6 Mb/s	1.9 Mb/s	3.8 Mb/s	5.6 Mb/s
Spectral efficiency	0.9 b/s/Hz	1.8 b/s/Hz	2.1 b/s/Hz	4.2 b/s/Hz	6.3 b/s/Hz

* Assume 25 MHz bandwidth and 28 users

Source: Ning Zhang, UCB

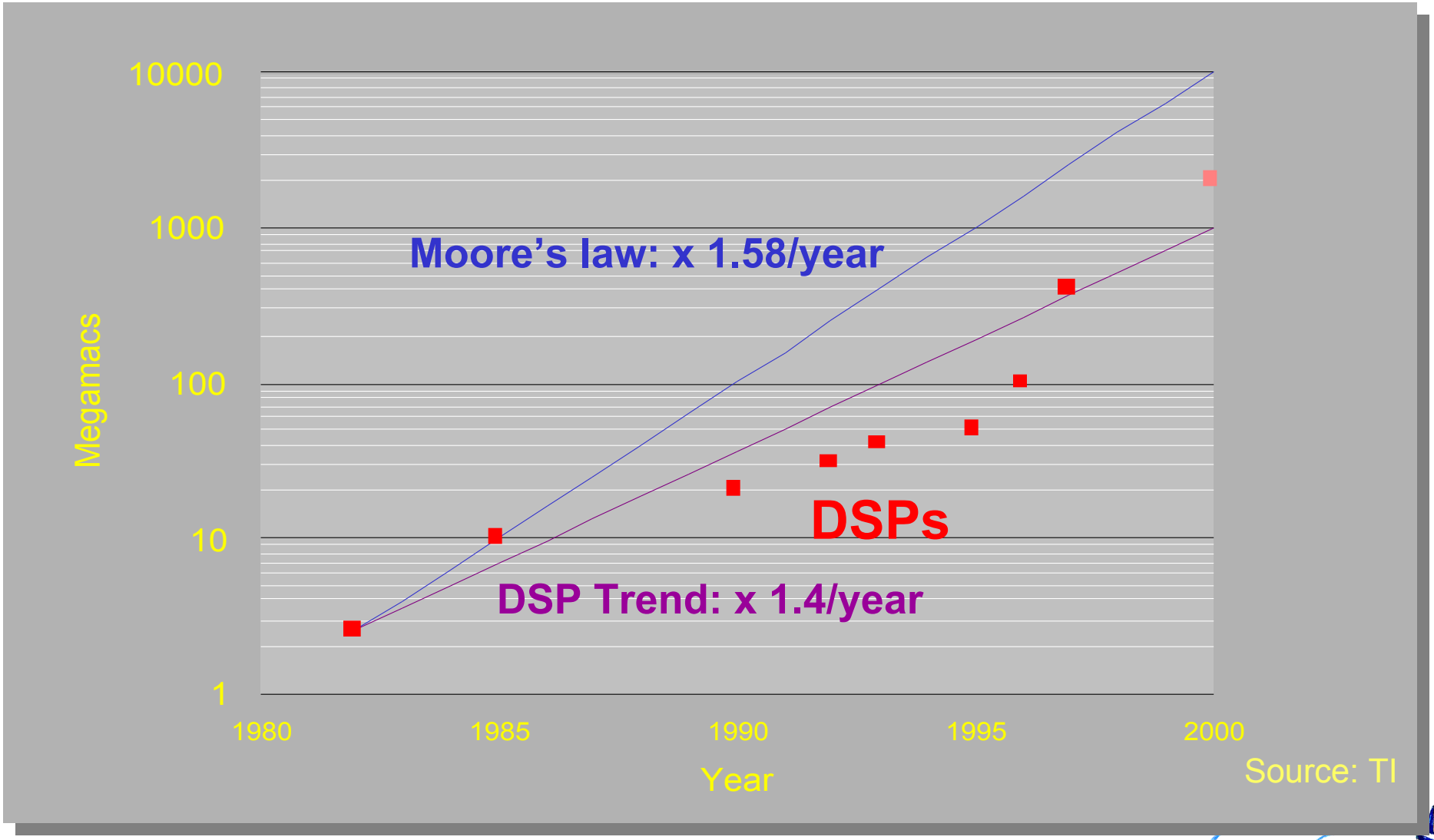
Berkeley Wireless Research Center

Shannon beats Moore's law



Courtesy: Ravi Subramanian (Morphics)

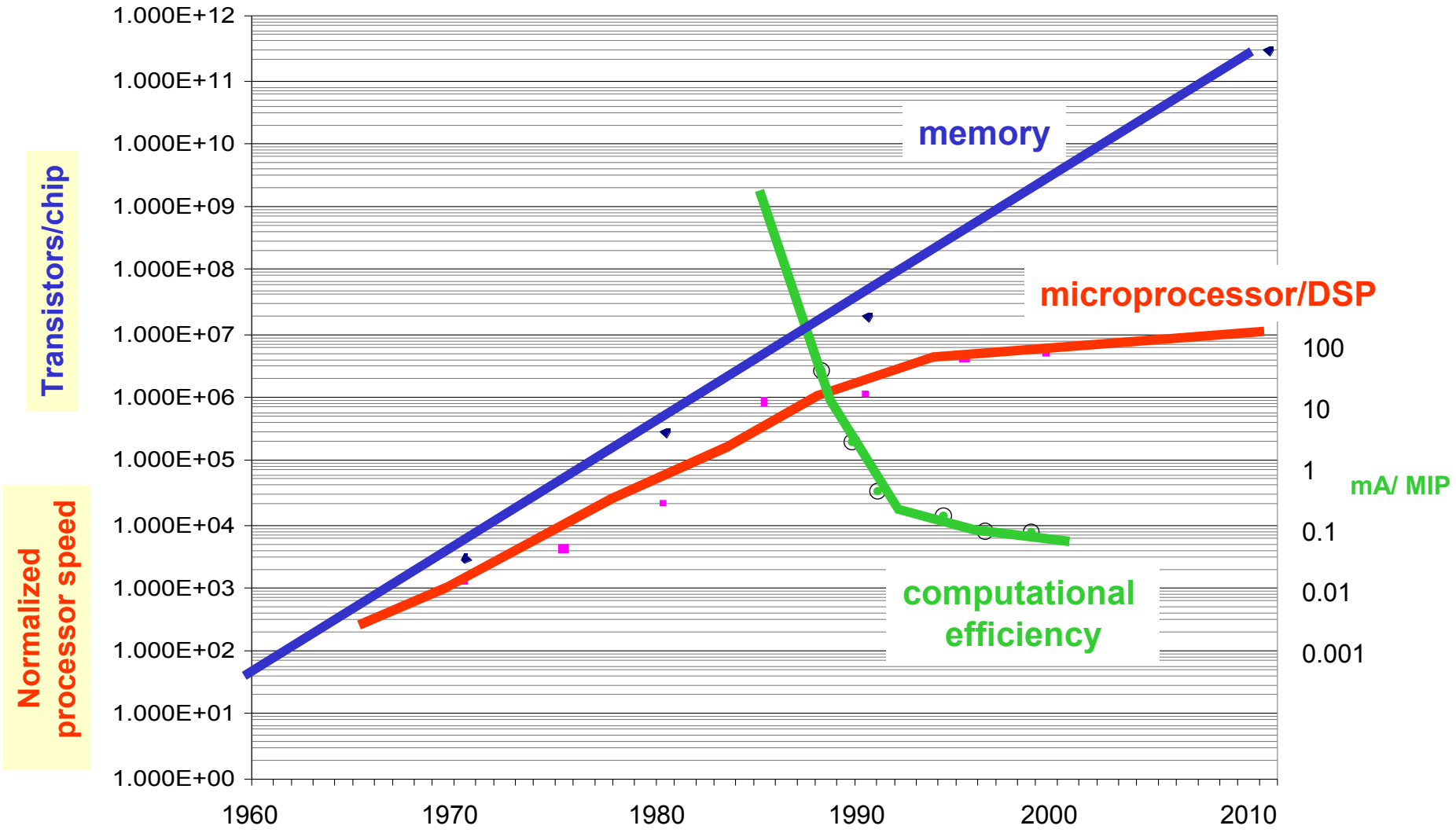
Single-Chip DSPs are Lagging ...



While algorithms are beating Moore's law!

Berkeley Wireless Research Center

Digital Processor Performance



Courtesy of Ravi Subramanian (Morphics)

Berkeley Wireless Research Center

The Law of Diminishing Returns

- **More transistors are being thrown at improving general-purpose CPU and DSP performance**

- **Fundamental bounds are being pushed**
  - limits on instruction-level parallelism
  - limits on memory system performance

- **Returns per transistor are diminishing**
  - new architectures realizing only 2-3 instructions/clock
  - increasingly large caches to hide DRAM latency

Berkeley Wireless Research Center

Some observations

- **Von-Neuman style instruction set architectures were perceived when switching devices and interconnections were extraordinarily expensive, and multiplexing-in-time provided the most economical solution**
  - Intel 4004: 2000 transistors, 1 MHz clock frequency, 1 metal layer
- **This led to the "clock-speed" affixation, which in fact is only a secondary measure of performance**
- **Power is rapidly becoming a limiting factor**
  - Newest processors are including thermal sensors and automatic slow-down (throttling) using pipeline bubbles and nop's to combat overheating and meltdown

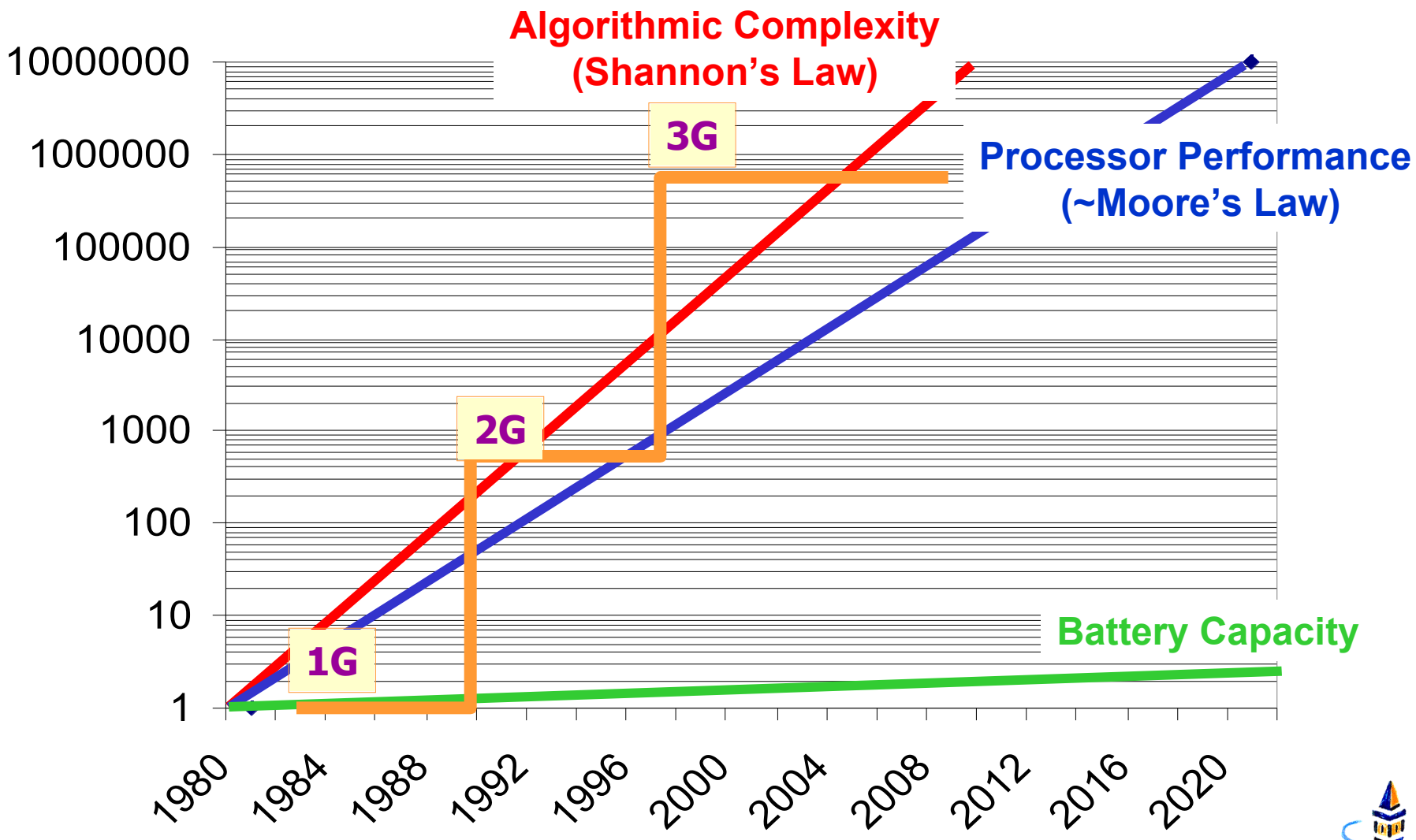
Compelling Issues in Wireless (2)

"The Last Meter Problem"
Ubiquitous wireless networking requires steep reduction in cost and energy dissipation

- **To be acceptable, radio cost has to be below 1$**
- **Frequent battery replacement on 100's of devices unacceptable**
- **Technology not likely to be of major help**

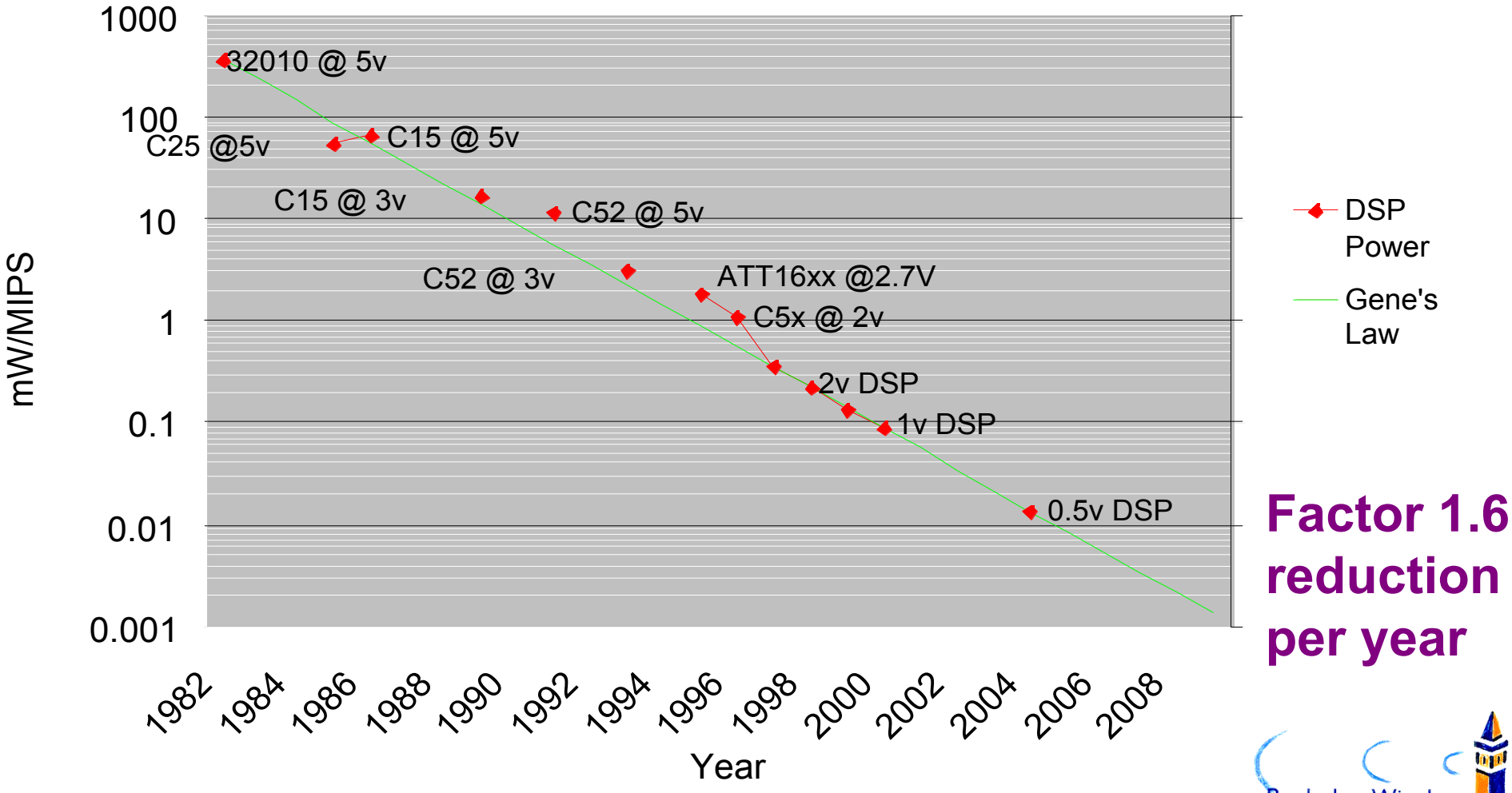
Berkeley Wireless Research Center

Energy to Play a Major Role



Courtesy: Ravi Subramanian (Morphics)

Energy Trends in DSPs



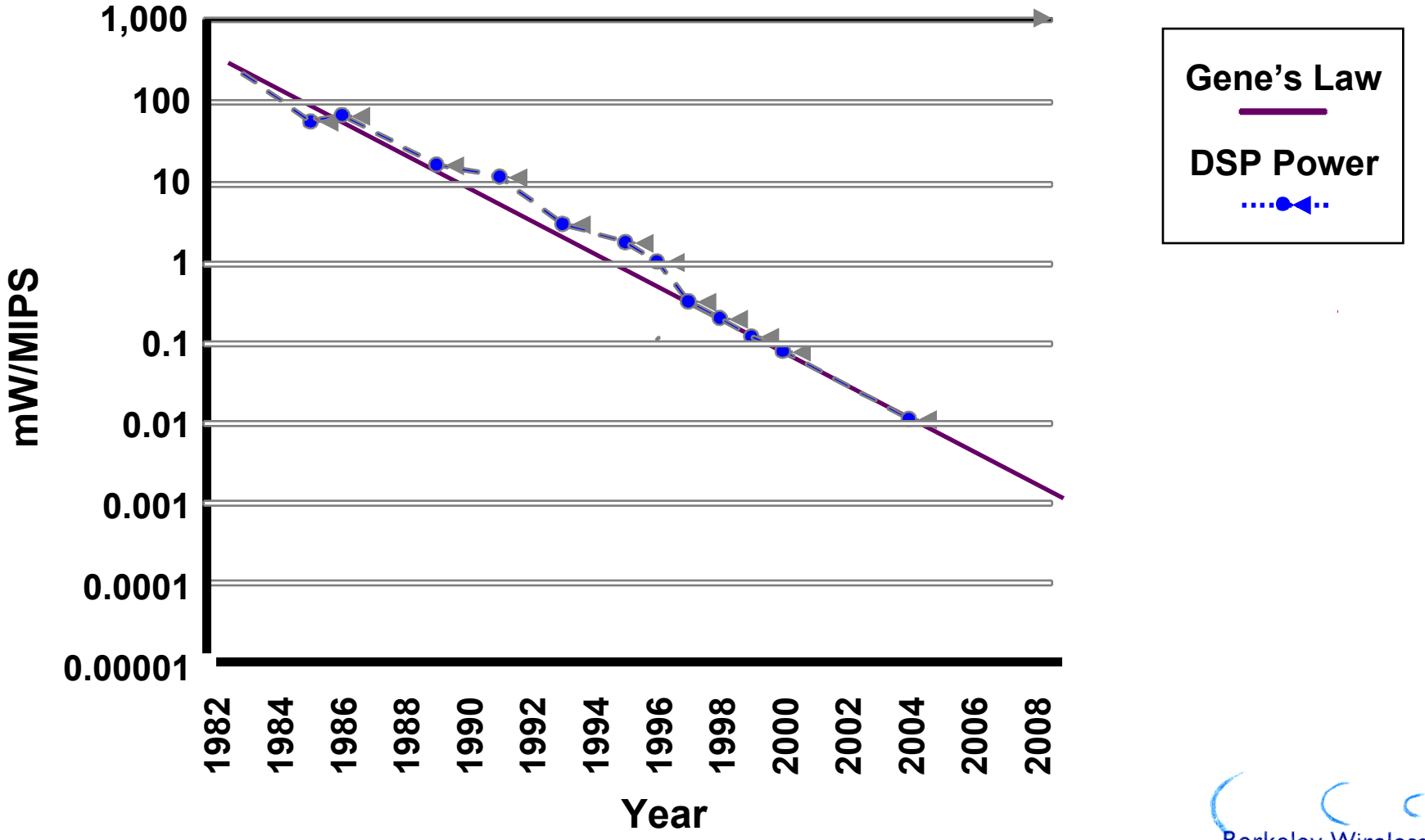
Factor 1.6 reduction per year

Source: TI

Berkeley Wireless Research Center

Energy Trends in DSPs

Gene (Frantz)'s Law



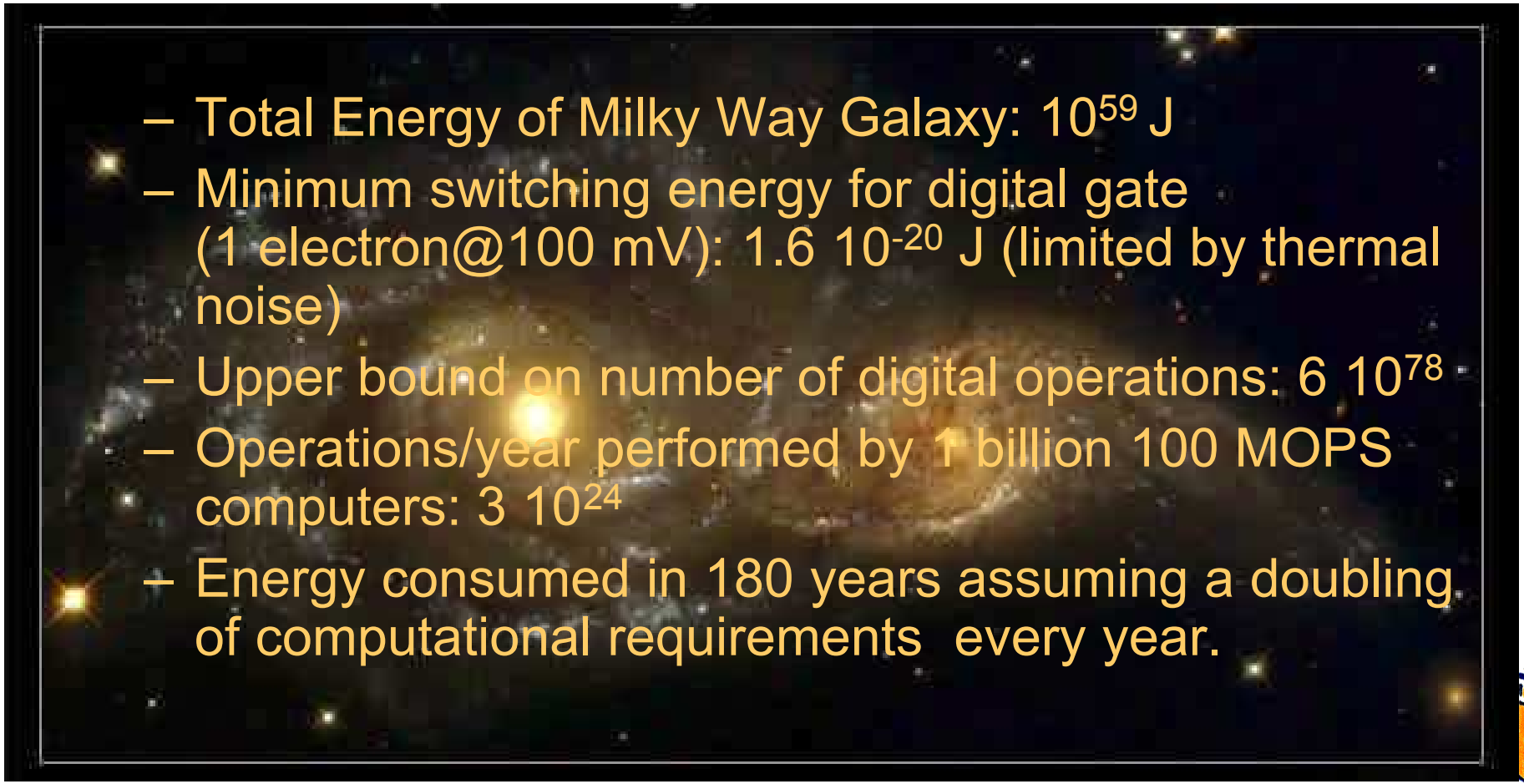
Source: Gene Frantz (TI)

Berkeley Wireless Research Center

Energy to Play a Major Role

A holistic perspective

Energy = upper bound on the amount of available computation



- Total Energy of Milky Way Galaxy: 10^{59} J
- Minimum switching energy for digital gate (1 electron@100 mV): 1.6 10^{-20} J (limited by thermal noise)
- Upper bound on number of digital operations: 6 10^{78}
- Operations/year performed by 1 billion 100 MOPS computers: 3 10^{24}
- Energy consumed in 180 years assuming a doubling of computational requirements every year.

Putting energy in perspective

- **Energy cost of digital computation**
  - 1999 (0.25µm): 1pJ/op (custom) … 1nJ/op (µproc)
  - 2004 (0.1µm): 0.1pJ/op (custom) … 100pJ/op (µproc)
    - Factor 1.6 per year; Factor 10 over 5 years
    - Assuming reconfigurable implementation: 1 pJ/op
- **Energy cost of communication**
  - 1999 Bluetooth
    - 1 nJ/bit transmission energy (thermal limit 30 pJ/bit)
      - 2.4 GHz band, 10 m distance
    - Overall energy: 170 nJ/bit reception / 150 nJ/bit transmission (!)
    - Standby power: 300 µW
  - 2004 Radio (10 m)
    - Only minor reduction in transmission energy
    - Reduce transceiver energy with at least a factor 10-50

The Changing Metrics

- **Power and/or Energy have become dominant drivers**
  - Limiting factor for performance and reliability in wall-plugged applications
  - Enabler for wide-spread use of distributed computing and data access
- **Energy reduction requires joint optimization process between application and implementation**

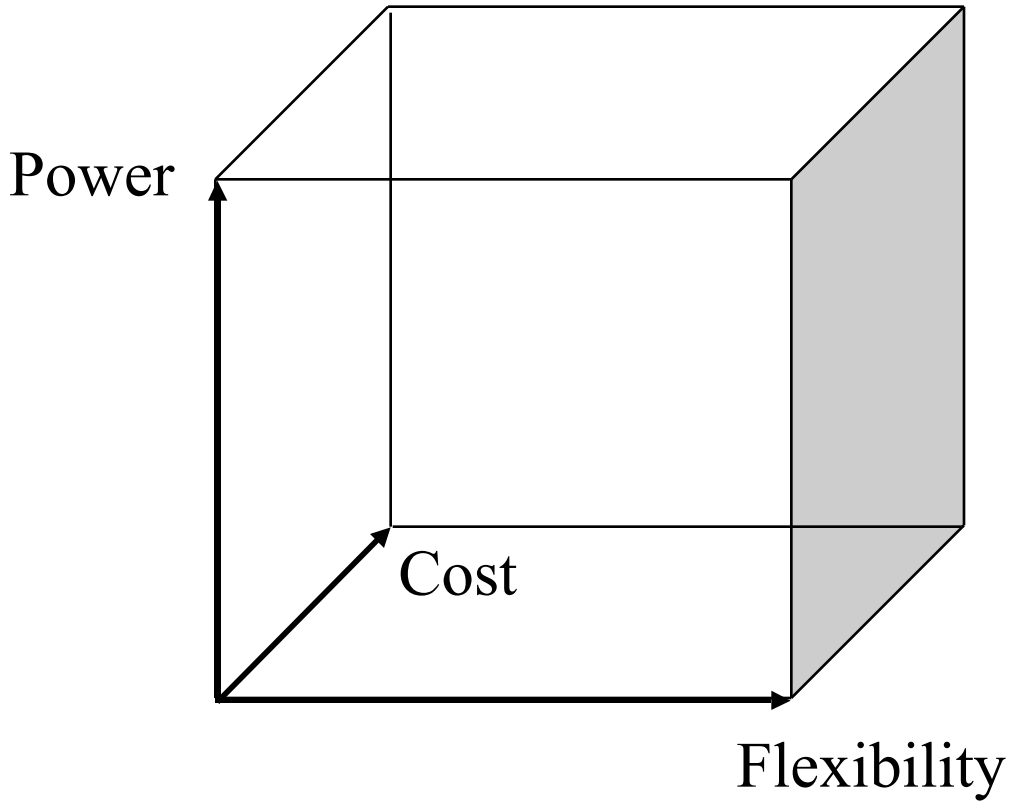
The Changing Metrics

- **Design complexity, and "context complexity" is sufficiently high that design verification is a major limitation on time-to-market**
- **Cost of fabrication facilities and mask making has increased significantly**
  - NRE cost of new design has increased significantly
- **Physical effects (parasitics, reliability issues, power management) are increasingly significant in the design process**
  - These must now be considered explicitly at the circuit level

Towards Fewer, but more Flexible and Reusable Silicon Platforms

Berkeley Wireless Research Center

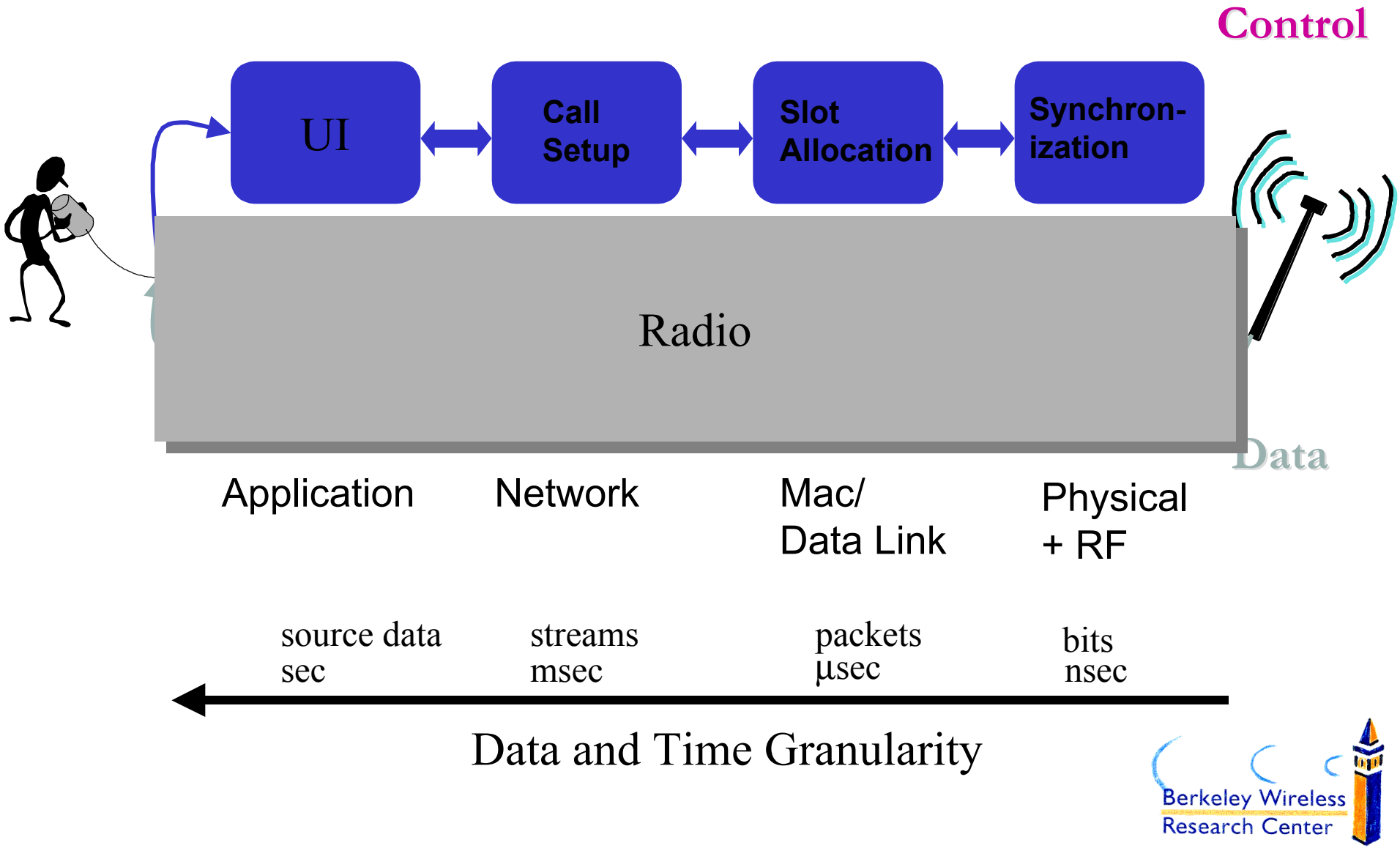
The Changing Metrics



Performance as a Functionality Constraint
("Just-in-Time Computing")

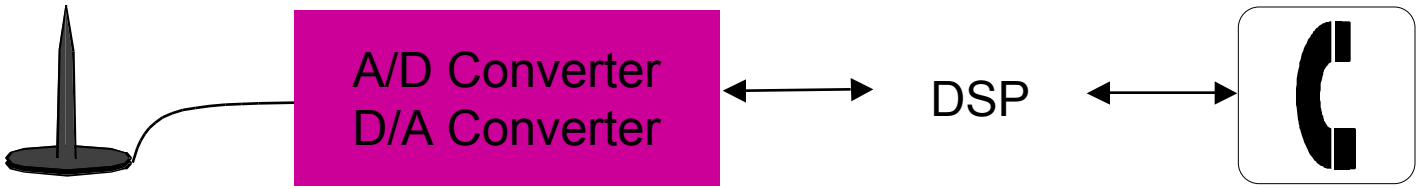
Berkeley Wireless Research Center

Challenges in Single-Chip Radio Design



Berkeley Wireless Research Center

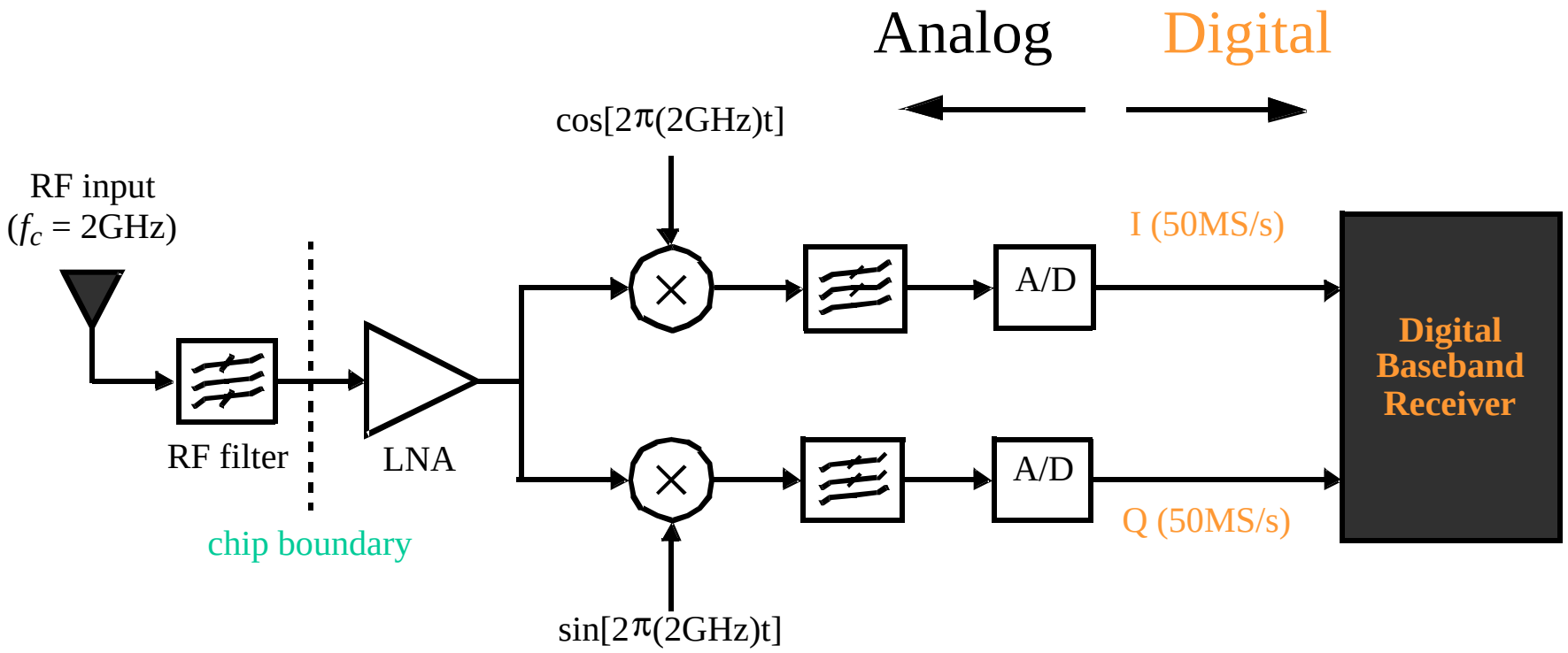
The Software Radio



- **Idea: Digitize (wideband) signal at antenna and use signal processing to extract desired signal**
- **Leverages of advances in technology, circuit design, and signal processing**
- **Software solution enables flexibility and adaptivity, but at huge price in power and cost**
- **16 bit A/D converter at 2.2 GHz dissipates 1 to 10 W**

Berkeley Wireless Research Center

The Mostly Digital Radio



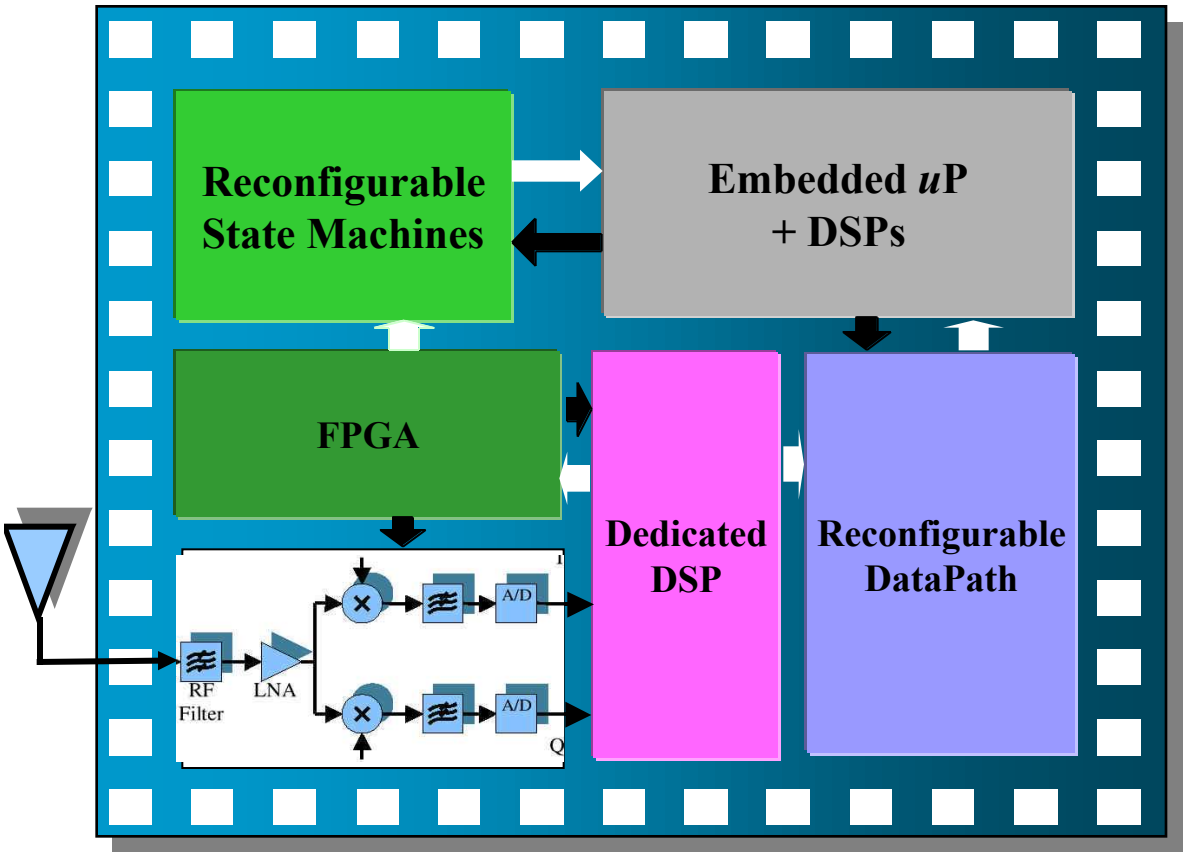
Berkeley Wireless Research Center

The Opportunities …

- **Scaling of technology, of course**
  - Performance doubles every 18 months
  - Energy reduced by 10 every 5 years
- **The system-on-a-chip**
  - Integration of heterogeneous functions on a single die leads to new solutions
- **Novel architectures**
  - Processor architectures exploiting concurrency (e.g. VLIW)
  - Reconfigurable hardware
  - Network-on-a-chip
- **Novel circuit solutions**
  - Voltage as a design variable
- **Novel design methodologies**
  - Communication-based Design
  - Platform-based Design

The Ideal "Radio-on-a-Chip" Platform

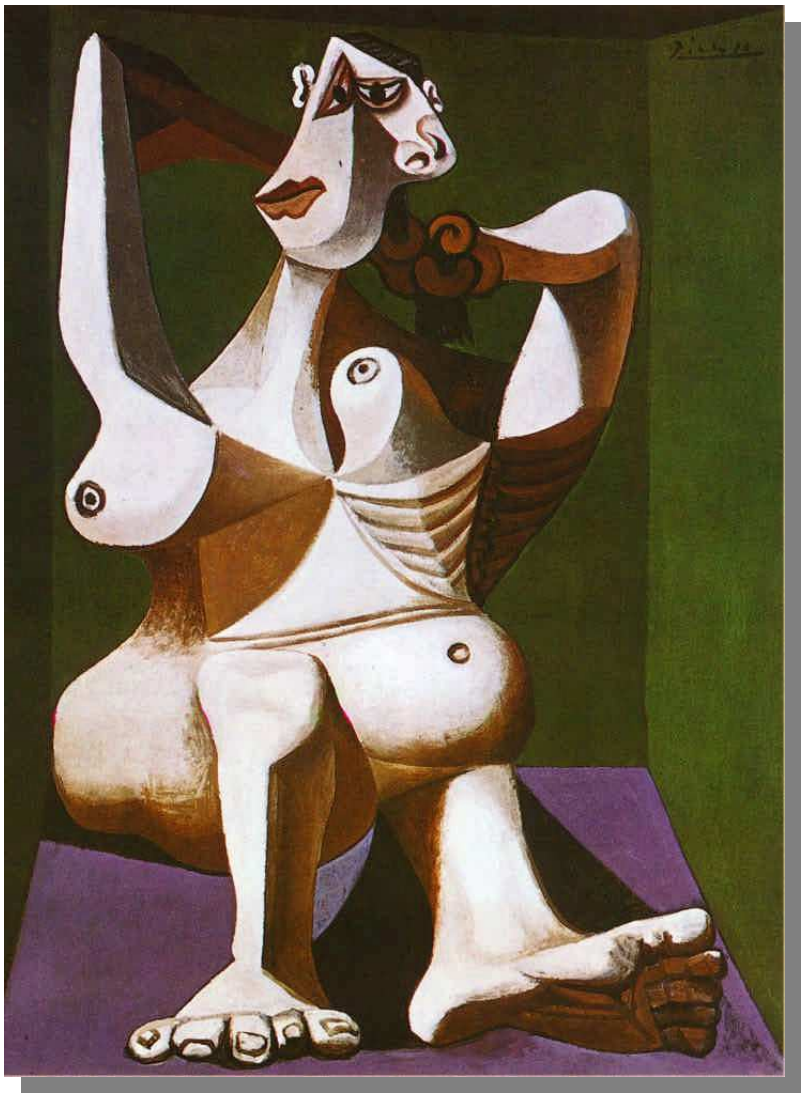
Combines performance, flexibility and energy-efficiency



- Heterogeneous
- Supports massive concurrency
- Matches the computational model
- Operates at minimum supply voltage and clock frequency
- Provides flexibility only where needed and desirable and at the right granularity

Berkeley Wireless Research Center

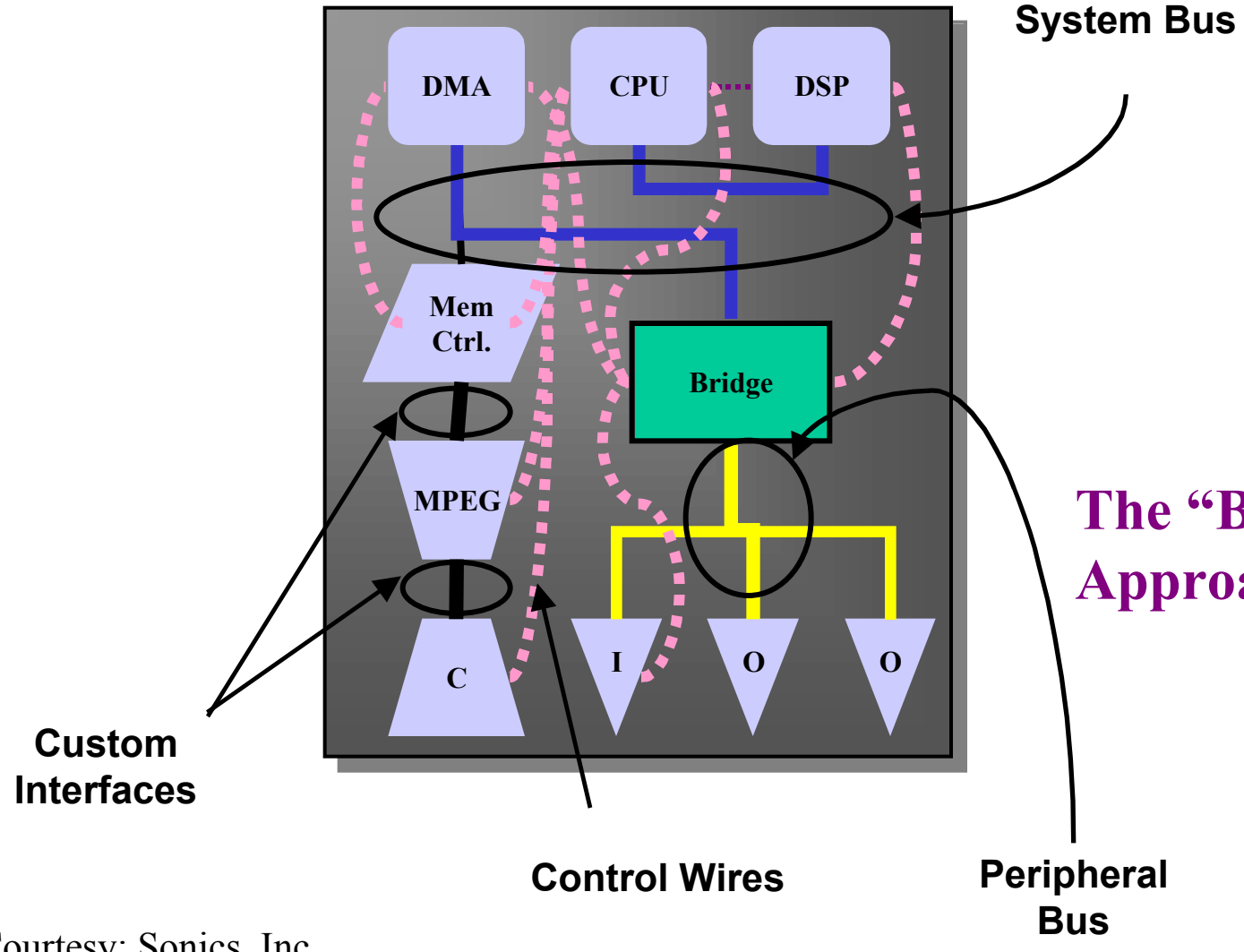
The System-on-a-Chip Nightmare



"Femme se coiffant"
Pablo Ruiz Picasso
1940

Berkeley Wireless
Research Center

The System-on-a-Chip Nightmare

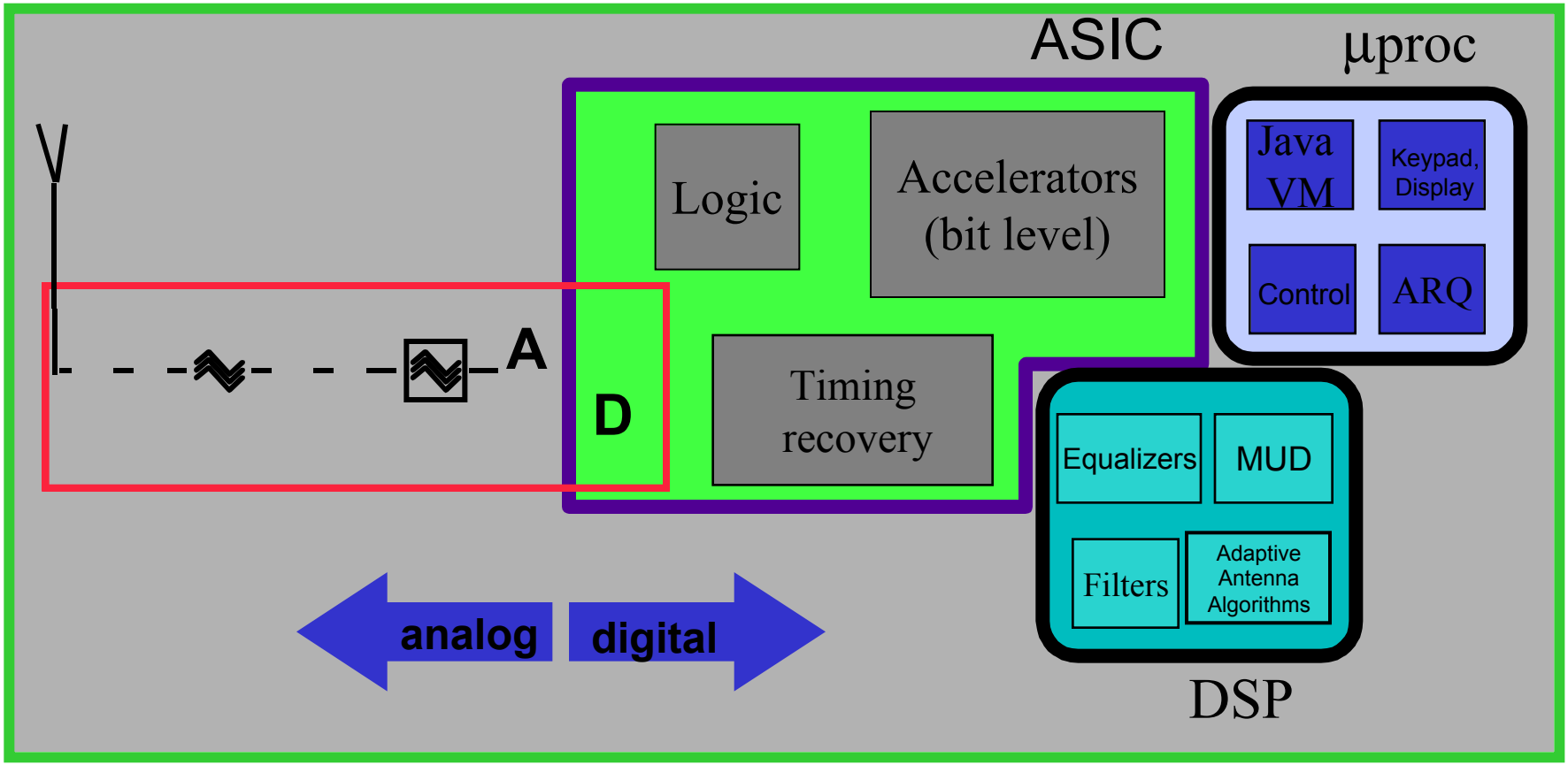


The "Board-on-a-Chip" Approach

Courtesy: Sonics, Inc

Berkeley Wireless Research Center

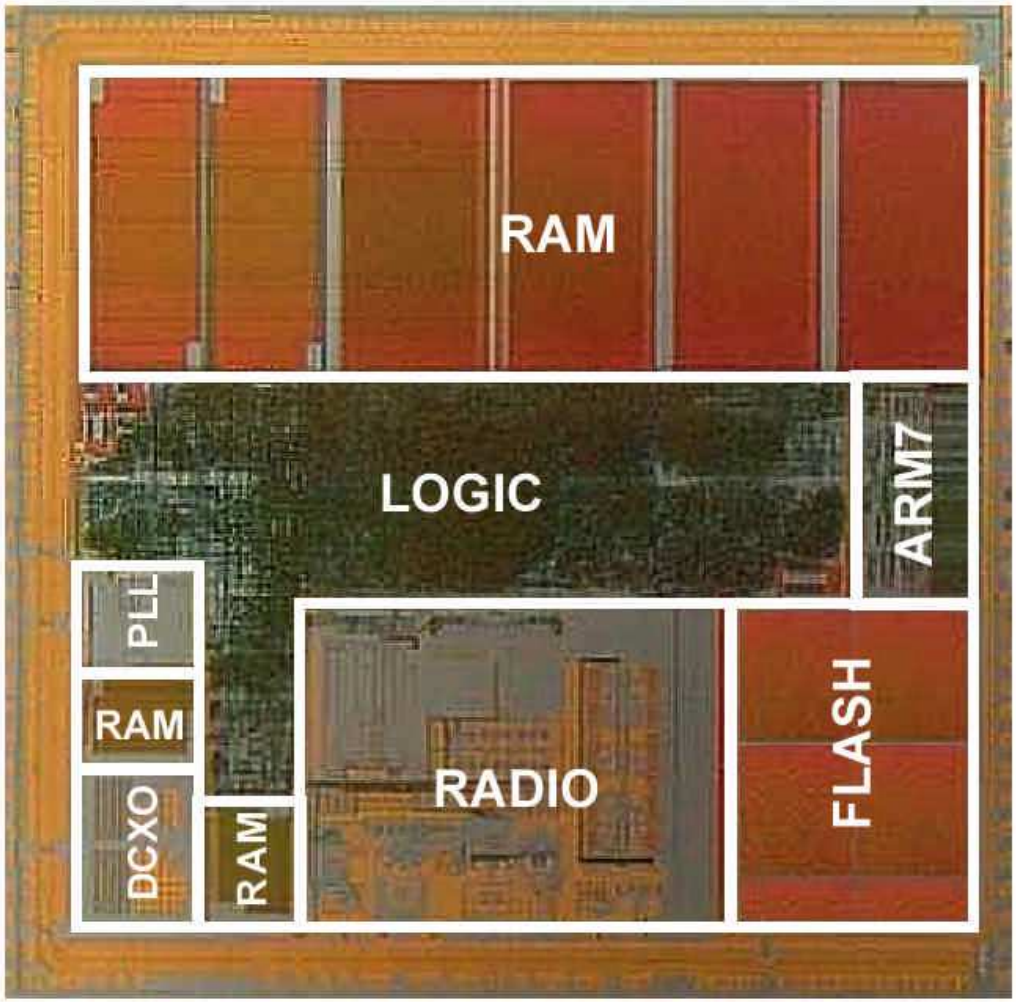
Current Integrated Wireless Transceivers



A "Board-on-a-Chip" Approach

Berkeley Wireless
Research Center

Example: Single-Chip Bluetooth



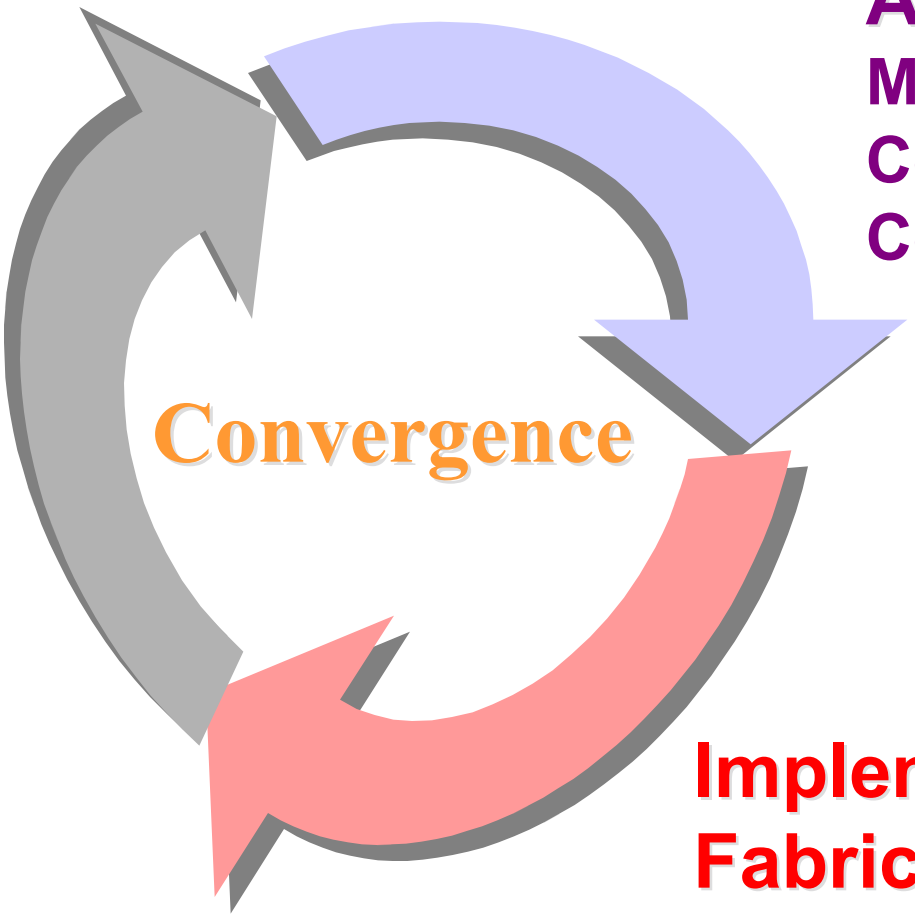
40 mm²
CMOS 0.25µm
5 metals
MiM-capacitors
200Ω/□ Resistors
Flash

Source: Alcatel, ISSCC 2001

Berkeley Wireless Research Center

System-on-a-Chip
A Renaissance in Design

Applications
Multimedia
Consumer
Communications



Design Methodology
Hard+Soft

Implementation Fabrics
Silicon reuse
Silicon networks

Aart De Geus
DAC'99

Berkeley Wireless Research Center

A Central Theme: Raising the Reuse Factor

Reuse comes in generations

Generation	Reuse element	Status
1st	Standard cells	Well established
2nd	IP blocks	Being introduced
3rd	Architecture	Emerging
4th	IC	Early research

Source: Theo Claasen (Philips) – DAC 00

Berkeley Wireless Research Center

Platform-Based Design

"Only the consumer gets freedom of choice; designers need freedom *from* choice"
(Orfali, et al, 1996, p.522)

- **A platform is a restriction on the space of possible implementation choices, providing a well-defined abstraction of the underlying technology for the application developer**
- **New platforms will be defined at the architecture-micro-architecture boundary**
- **They will be component-based, and will provide a range of choices from structured-custom to fully programmable implementations**
- **Key to such approaches is the representation of communication in the platform model**

Source:R.Newton

Berkeley Wireless Research Center

Hardware Platforms

Hardware Platform: not only a fully specified SoC but also *a family of architectures that share some common feature*:

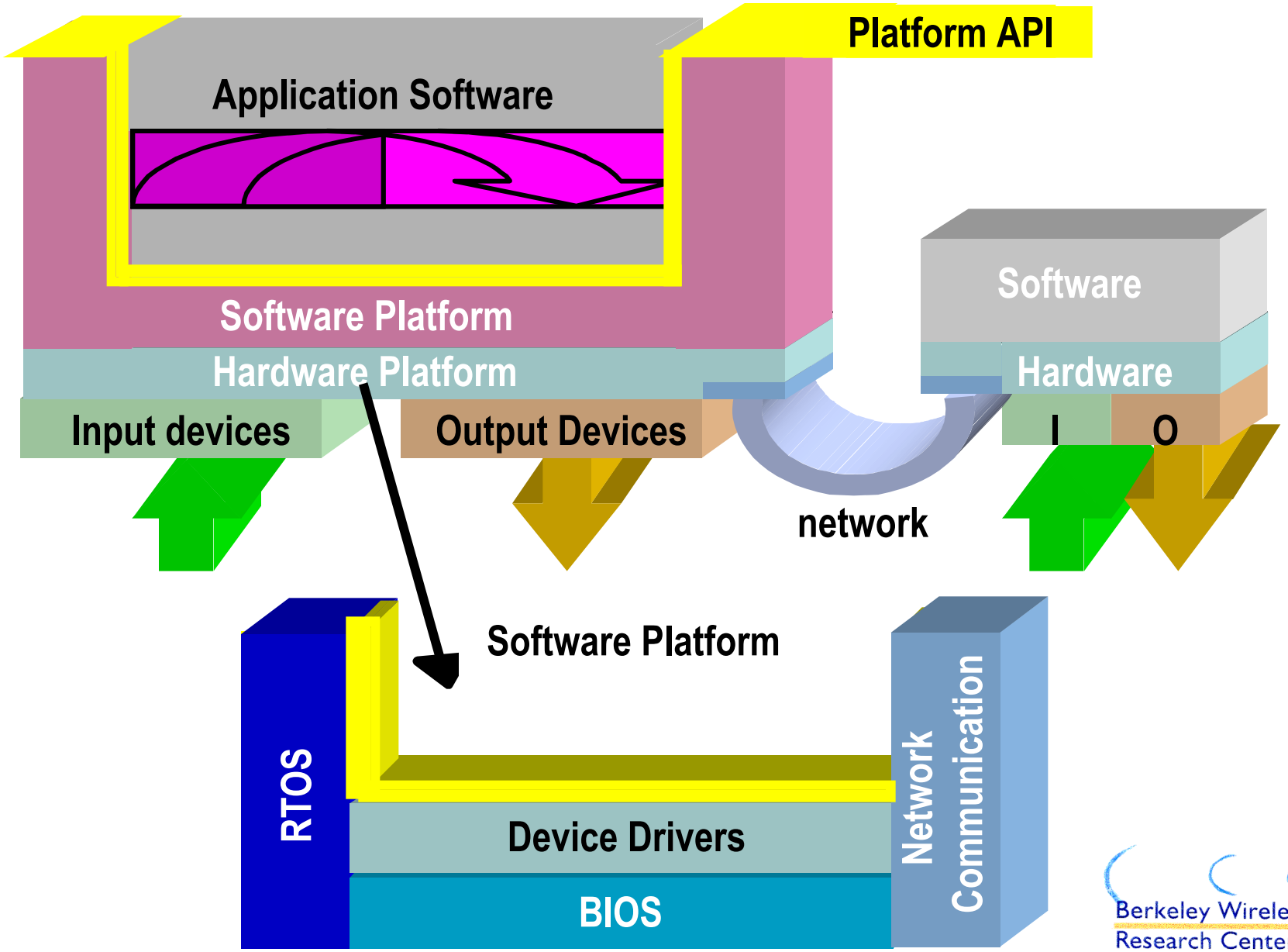
A Hardware Platform is a family of architectures that satisfy a set of architectural constraints imposed to allow the re-use of hardware and software components.

The stronger the constraints the more component re-use but stronger constraints imply fewer architectures to choose from!

Hardware Platforms Not Enough!

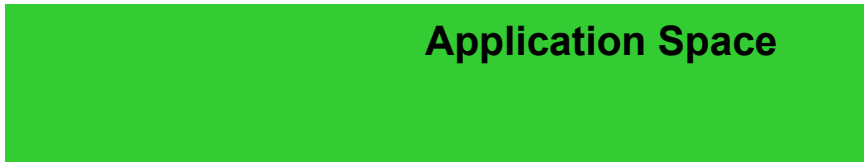
- **Hardware platform has to be abstracted**
- **Interface to the application software is API**
- **Software layer performs abstraction:**
  - Programmable cores and memory subsystem with RTOS
  - I/O subsystem via Device Drivers

Software Platforms



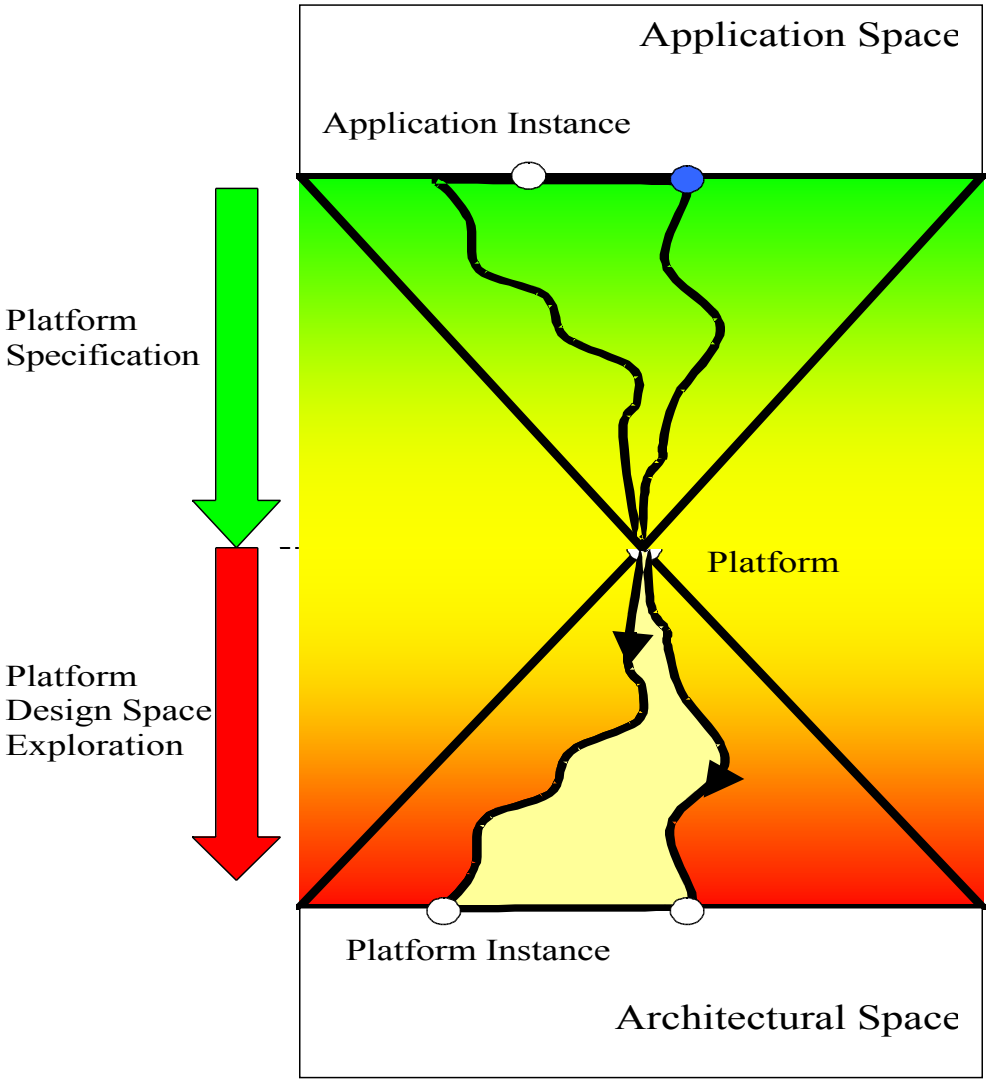
Berkeley Wireless Research Center

The Platform Tension



Berkeley Wireless Research Center

The Platform Approach



Source:
Alberto Sangiovanni-Vincentelli

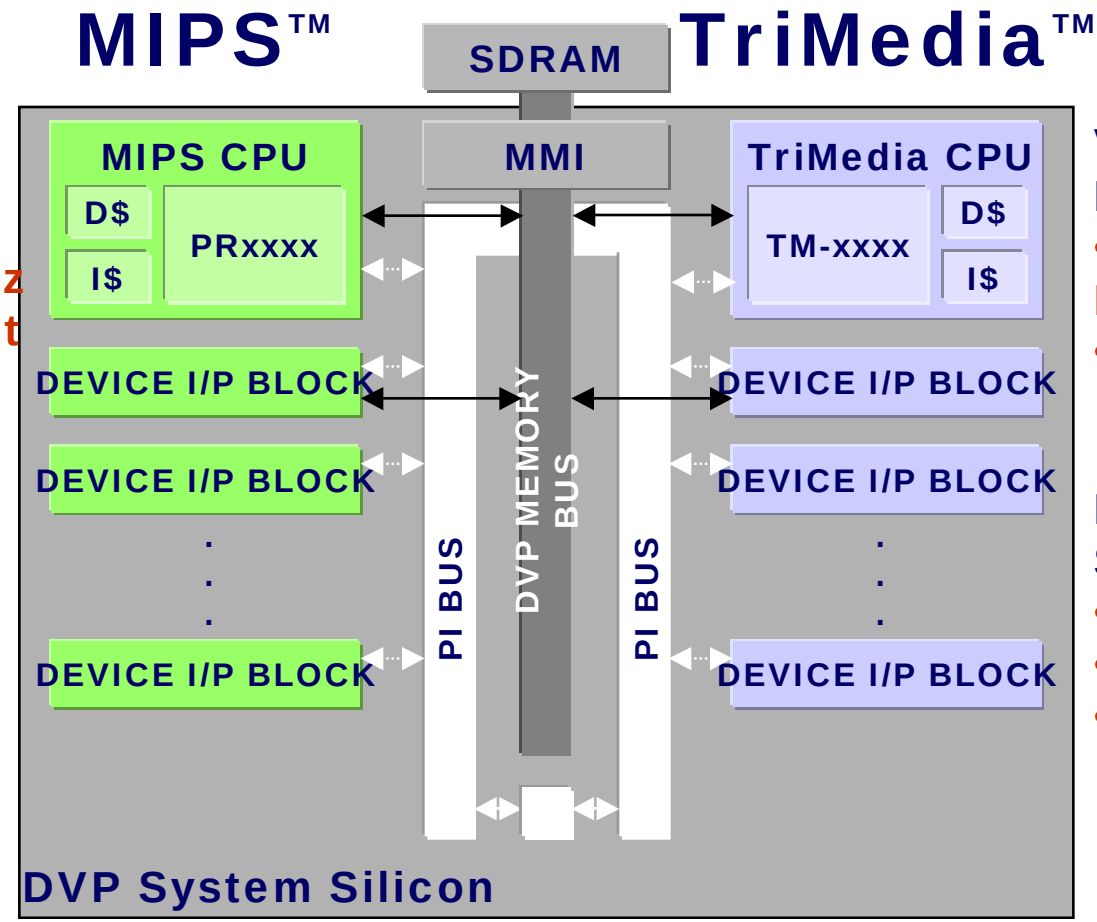
Berkeley Wireless
Research Center

Example: Philips Nexperia™ DVP

General Purpose RISC Processor
- 50 to 300+ MHz
- 32-bit or 64-bit

Library of Device Blocks
- Image coprocessors
- DSPs
- UART
- 1394
- USB

- *...and more*



VLIW Media Processor:
- 100 to 300+ MHz
- 32-bit or 64-bit

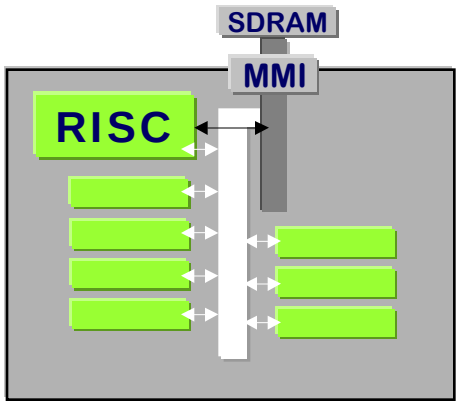
Nexperia System Busses
- PI bus
- Memory bus
- 32-128 bit

Flexible architecture for digital video applications

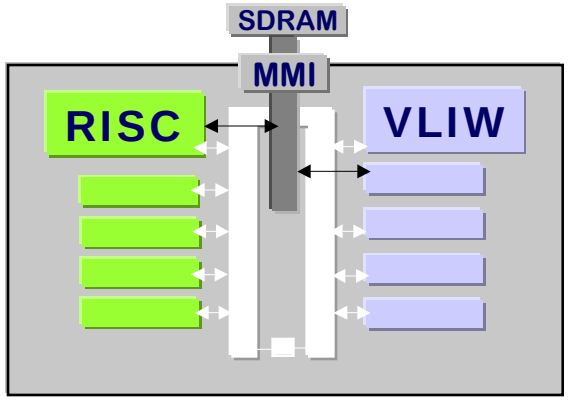
Source: Theo Claasen (Philips) – DAC 00

Berkeley Wireless Research Center

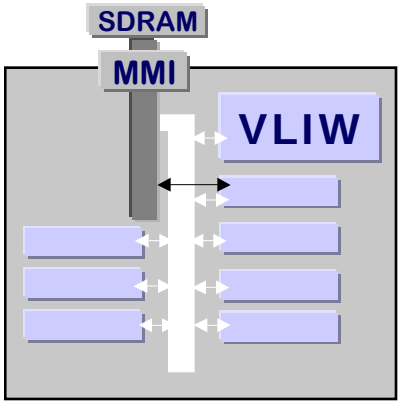
Nexperia™ Scalability



MIPS CPU + Device blocks + Software



MIPS CPU + Trimedia CPU replacing some Device blocks



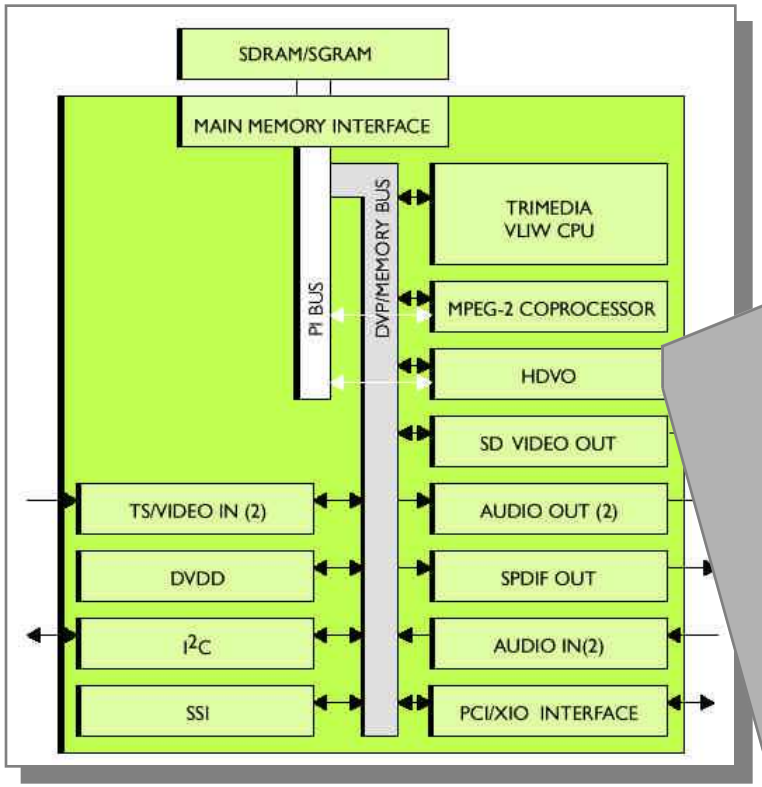
TriMedia CPU + Device blocks when control functions are minimal

- **Single architecture, multiple product configurations**
  - Processor core options - TM32, TM64, MIPS32, MIPS64 ...
  - Device block options
- **Highly programmable to weakly programmable**

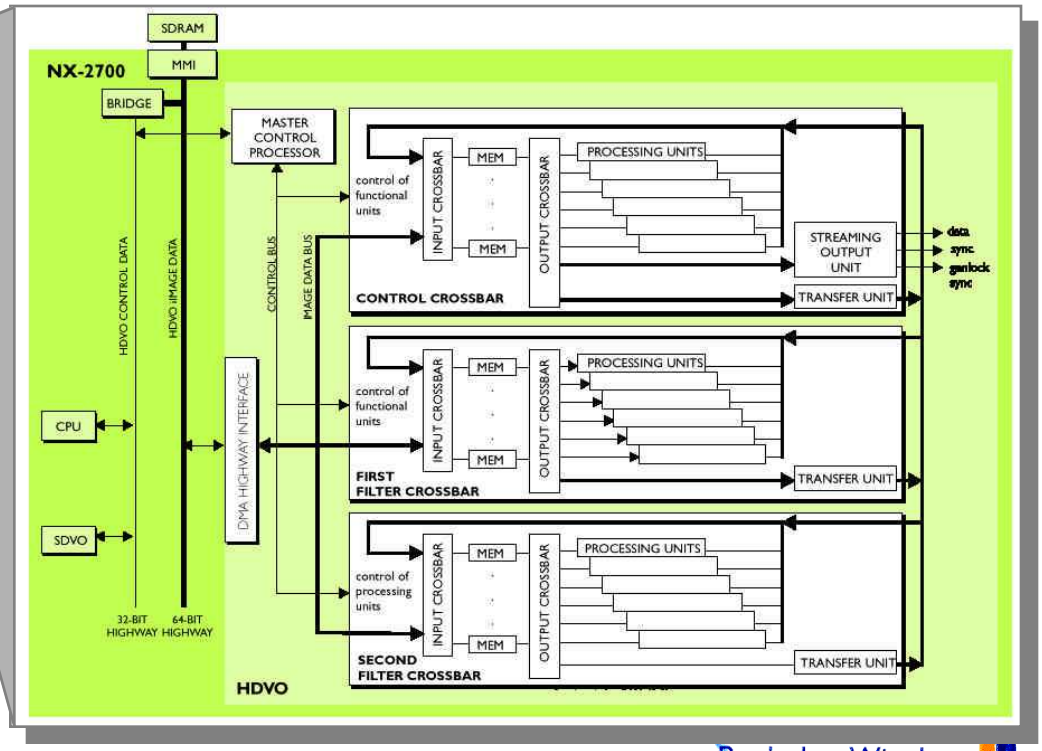
Source: Theo Claasen (Philips) – DAC 00

Berkeley Wireless Research Center

An Example of an Instantiation



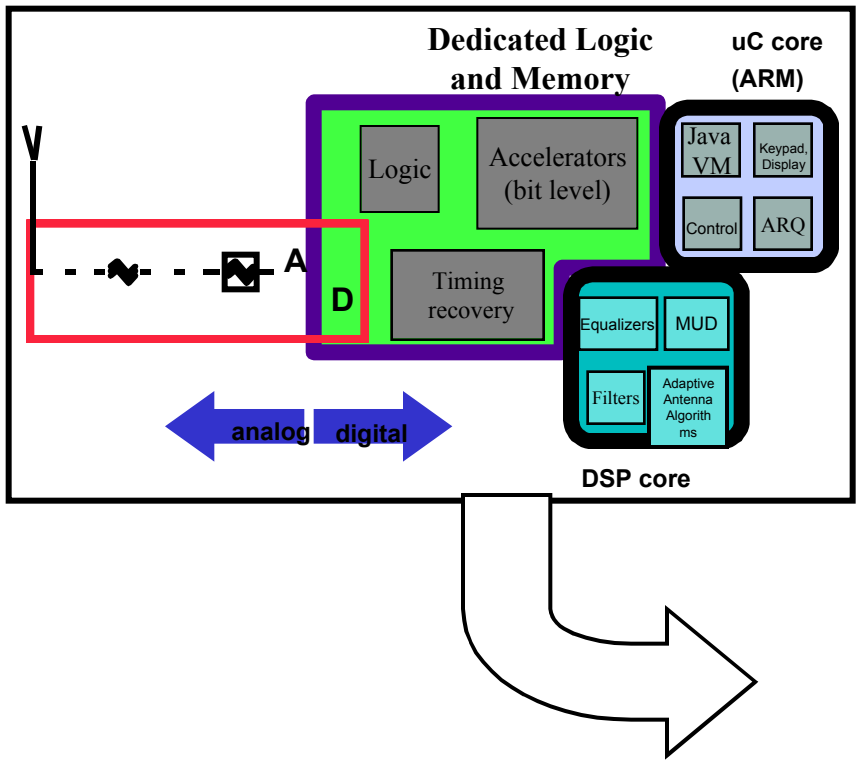
Combines Trimedia VLIW with Configurable media co-processors

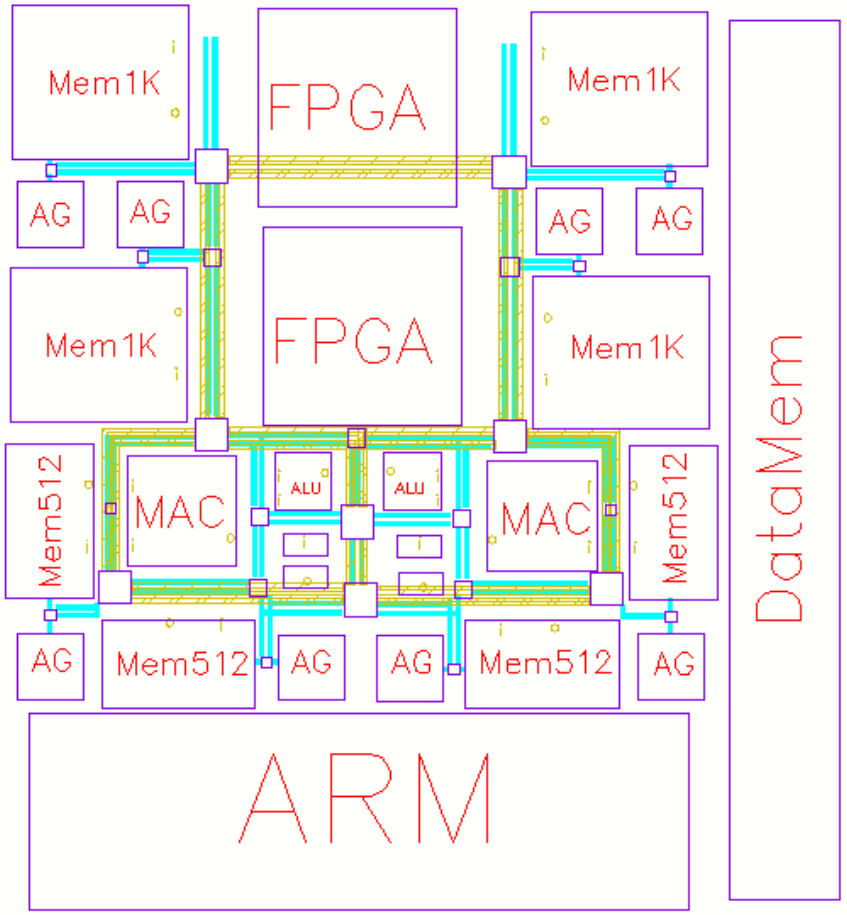


Philips Nexperia NX-2700
A programmable HDTV media processor

Berkeley Wireless Research Center

Pleiades:
Digital Wireless Platform

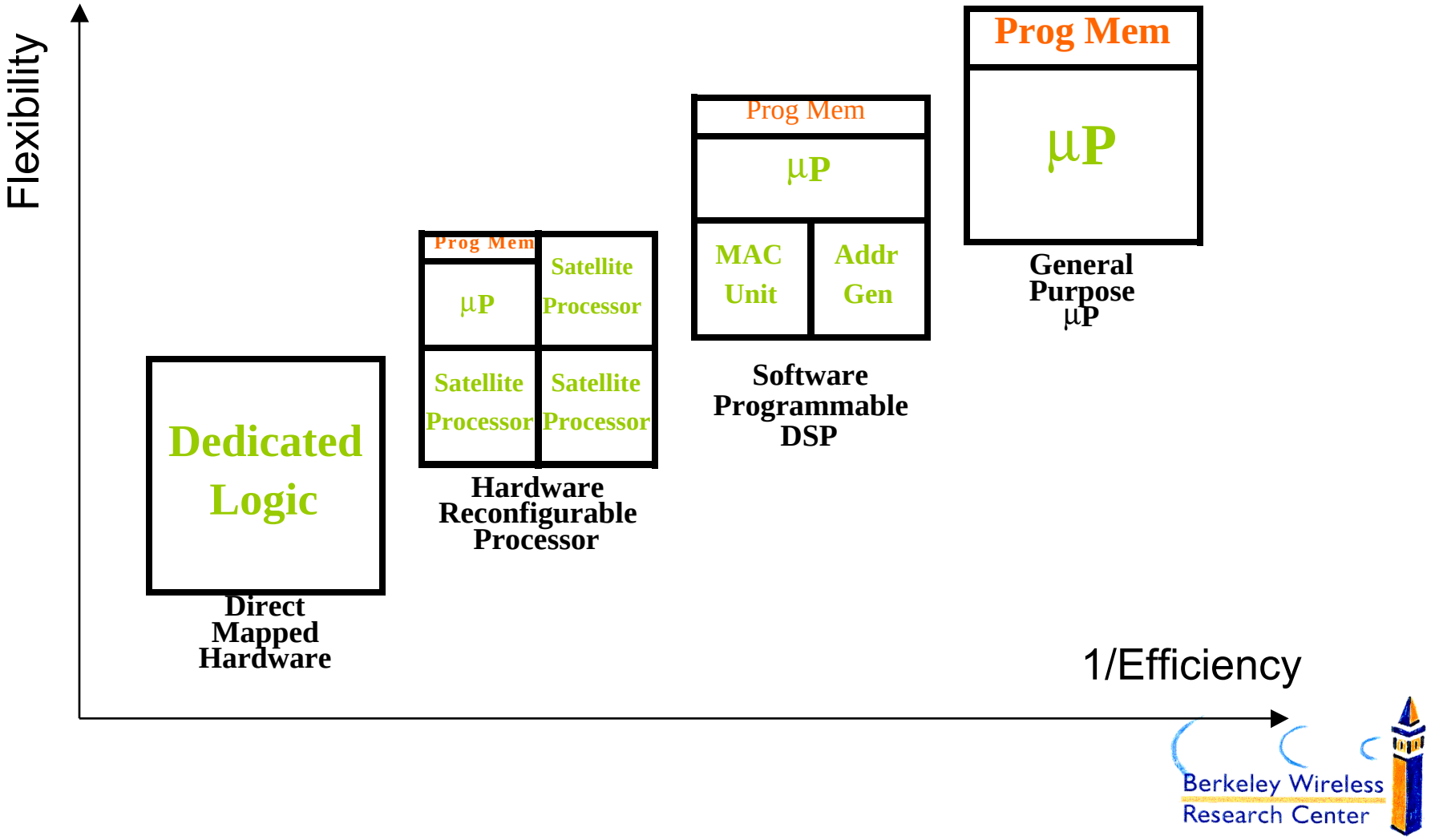




Source: Berkeley Wireless Research Center

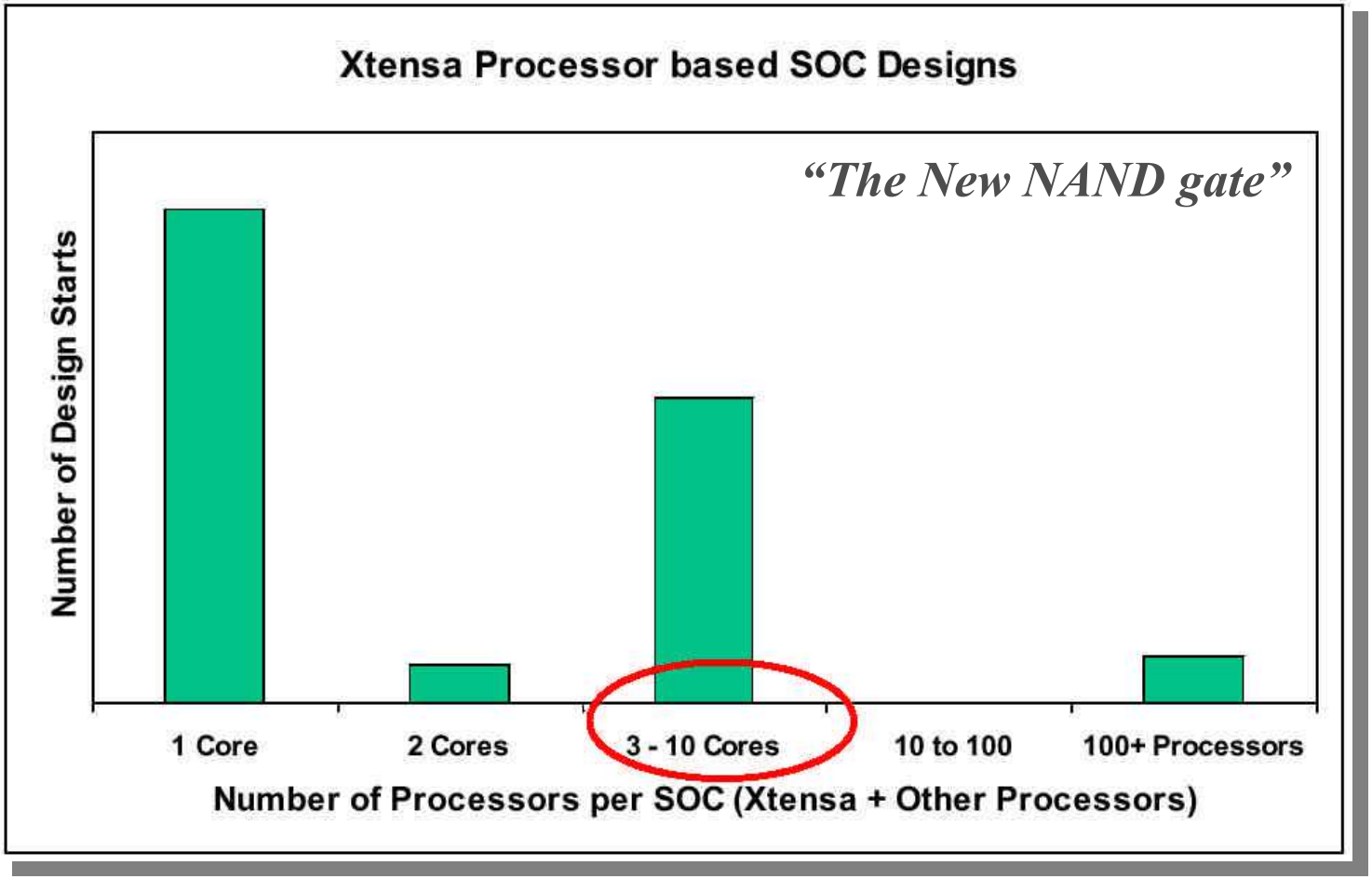
Berkeley Wireless Research Center

The Architectural Choices



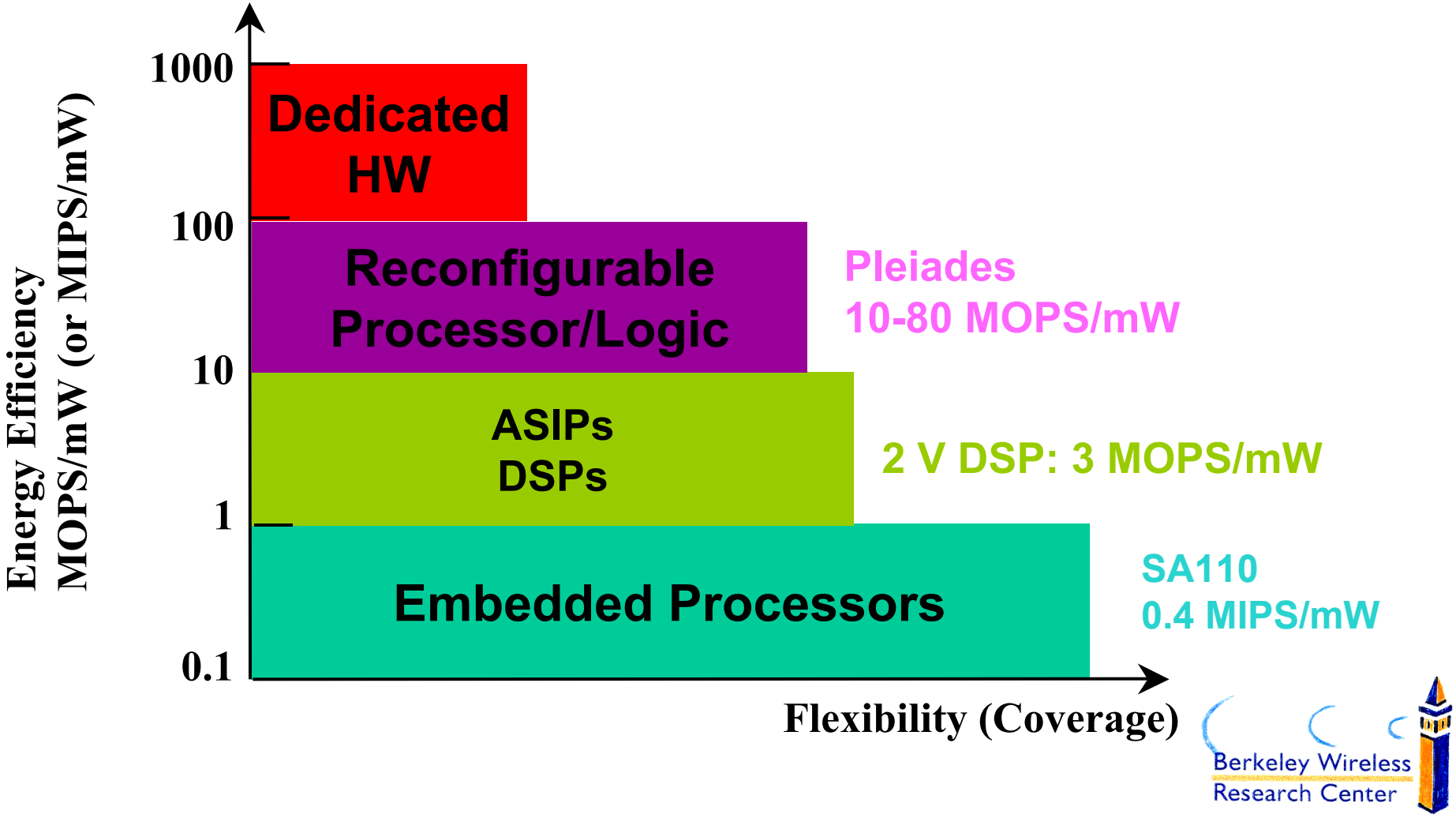
Berkeley Wireless Research Center

An Attractive Option: Multi-Processor System-on-a-Chip



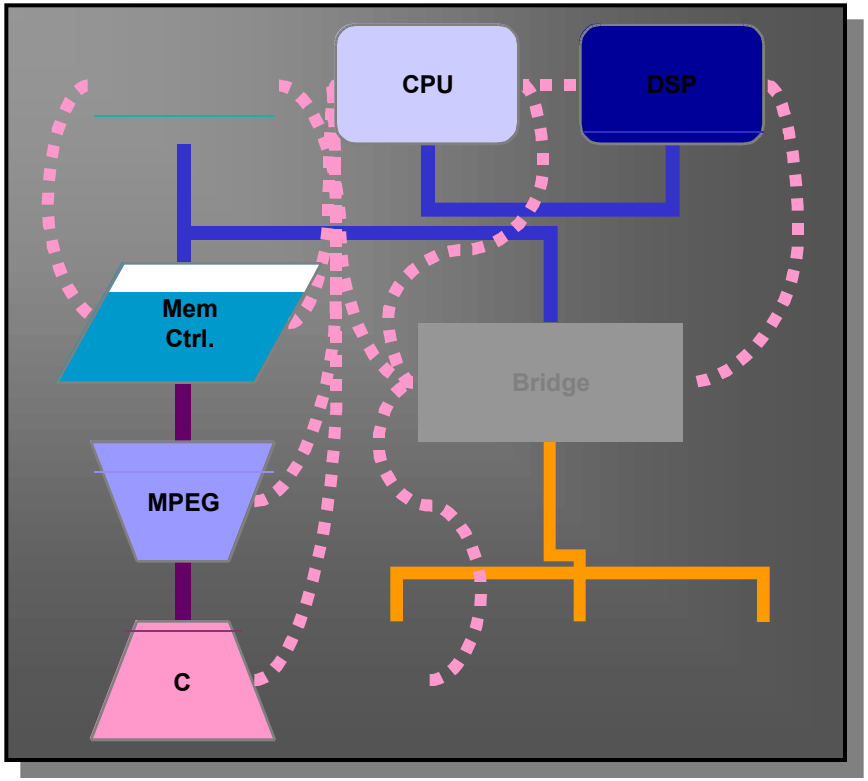
Courtesy: Chris Rowen, Tensilica

The Energy-Flexibility Gap



Berkeley Wireless Research Center

Communication-Based Design



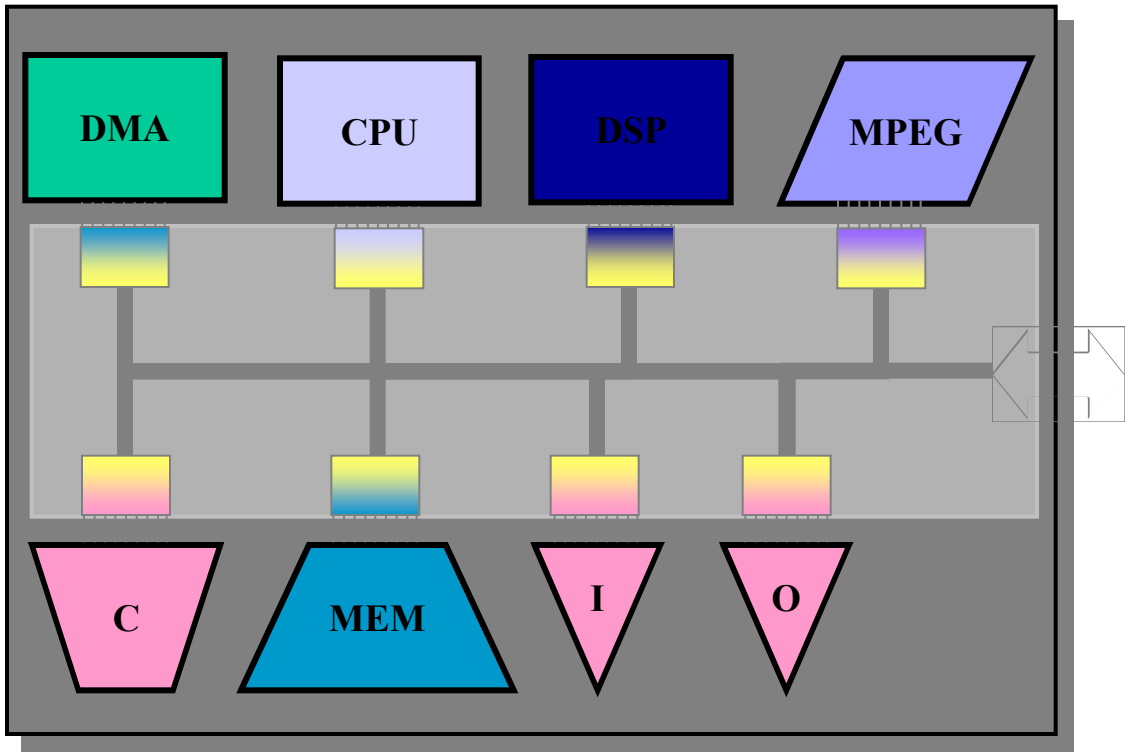
Berkeley Wireless
Research Center

Orthogonalizing Communication from Behavior

- **Historically lots of work on Behavior**
  - hierarchy well established
  - several descriptions available (with variable levels of precision)
  - synthesis available
- **Communication less well investigated**
  - hard to separate from behavior, usually intertwined
  - telecomm protocols are the best existing example
- **Need to understand Formalism, Abstraction, and Decomposition for communication**

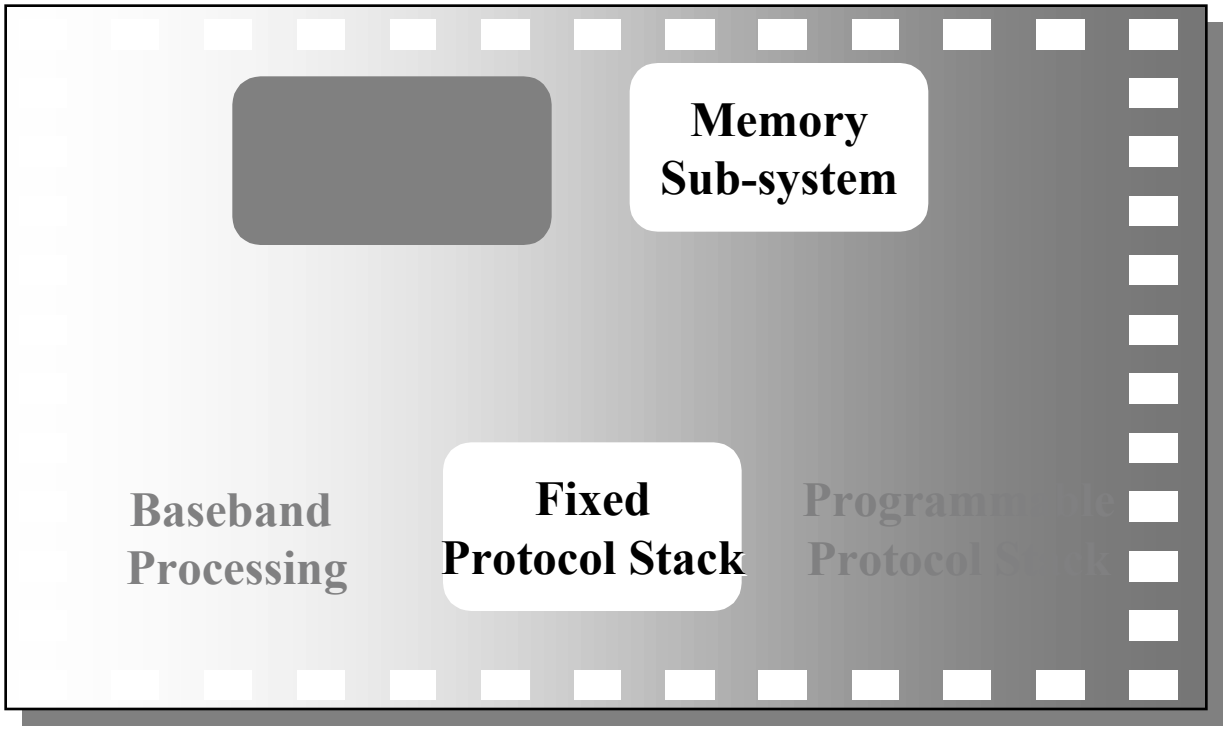
  DEFINED BY MODELS OF COMPUTATION!

Communication-based Design



Berkeley Wireless Research Center

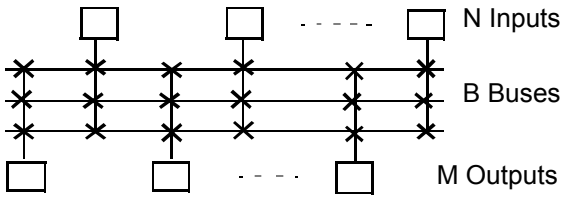
An Integrated Radio Processor (TCI)



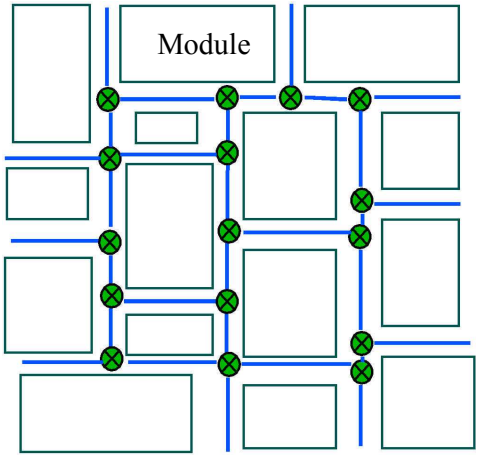
Berkeley Wireless
Research Center

The "Network-on-a-Chip"

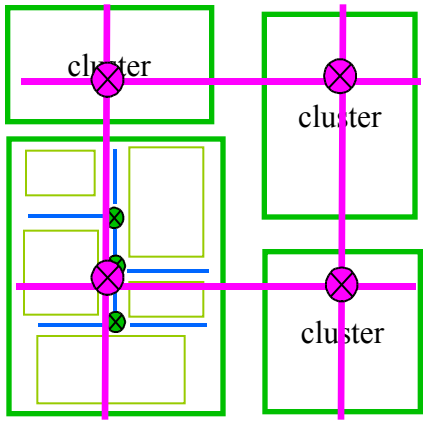
Multi-Bus



Mesh

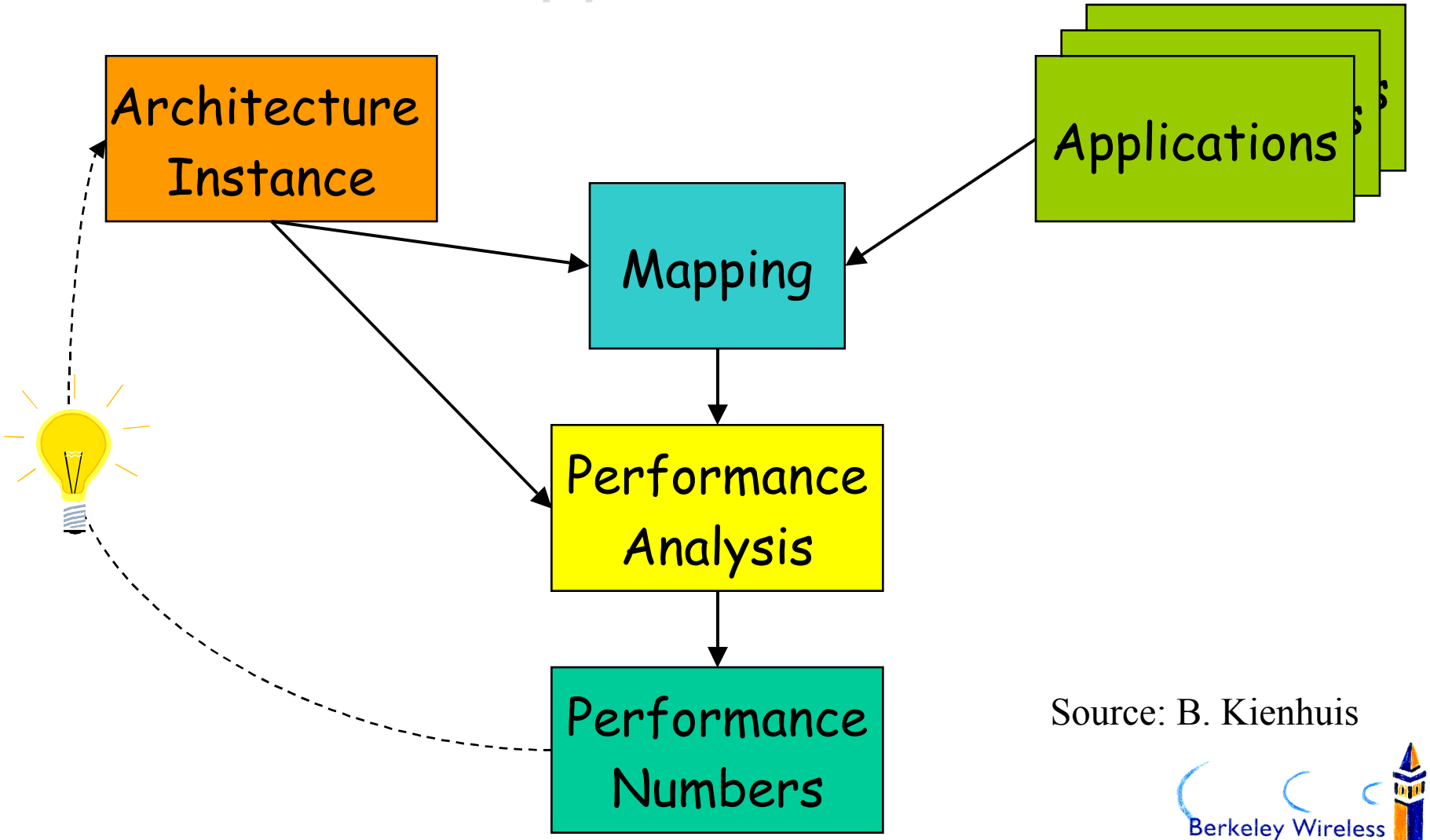


Hierarchical Mesh



		dot_ product	vector sum w/ scalar mult.	IIR
Multi-bus		50	50	138
Mesh	*Best*	8.7	5.2	24.6
	Worst	17.7	14.7	43.4
H. Mesh	***Best***	**4.7**	**3.8**	**18.8**
	Worst	**11.1**	**10.2**	**31.3**

Platform Exploration
The Y-Chart Approach



Source: B. Kienhuis

Berkeley Wireless Research Center

Summary and Perspective

- **Technology scaling is redefining the term "complexity"**
- **System-on-a-Chip fosters renaissance in processor architecture, opening the door for new models and combinations thereof:**
  Platform and Communication Based Design
- **SOC for wireless driven by new set of metrics: how to simultaneously optimize flexibility, cost, energy, and performance?**
- **Application-Architecture Exploration is Focal Part of Implementation Methodology**

Berkeley Wireless Research Center