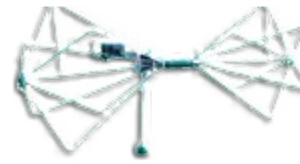
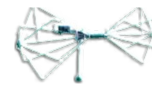


Consideration for Advancing Technology in Computer System Packaging

Dale Becker, Ph.D.
IBM Corporation, Poughkeepsie, NY

IEEE Distinguished Lecture Series
© 2014



Motivation

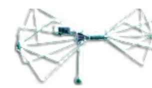
- Modern Computing is driven by
  - Cloud
  - Analytics
  - Mobile
  - Social
  - Security
- With many data inputs
- Demanding sophisticated analytics
- Sent back to distributed users
- Securely

➔More Data Bandwidth

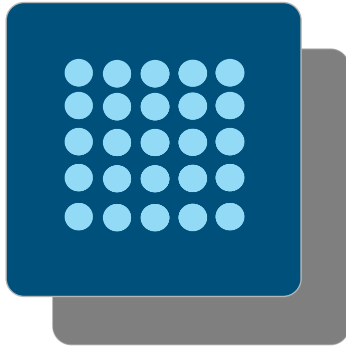
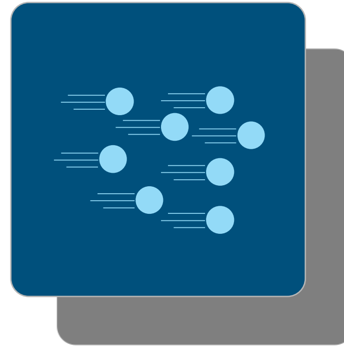
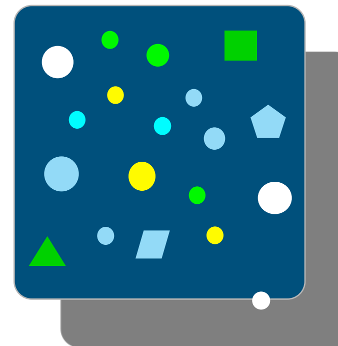
➔Less Data Latency

➔Higher integration of computing, networking and storage

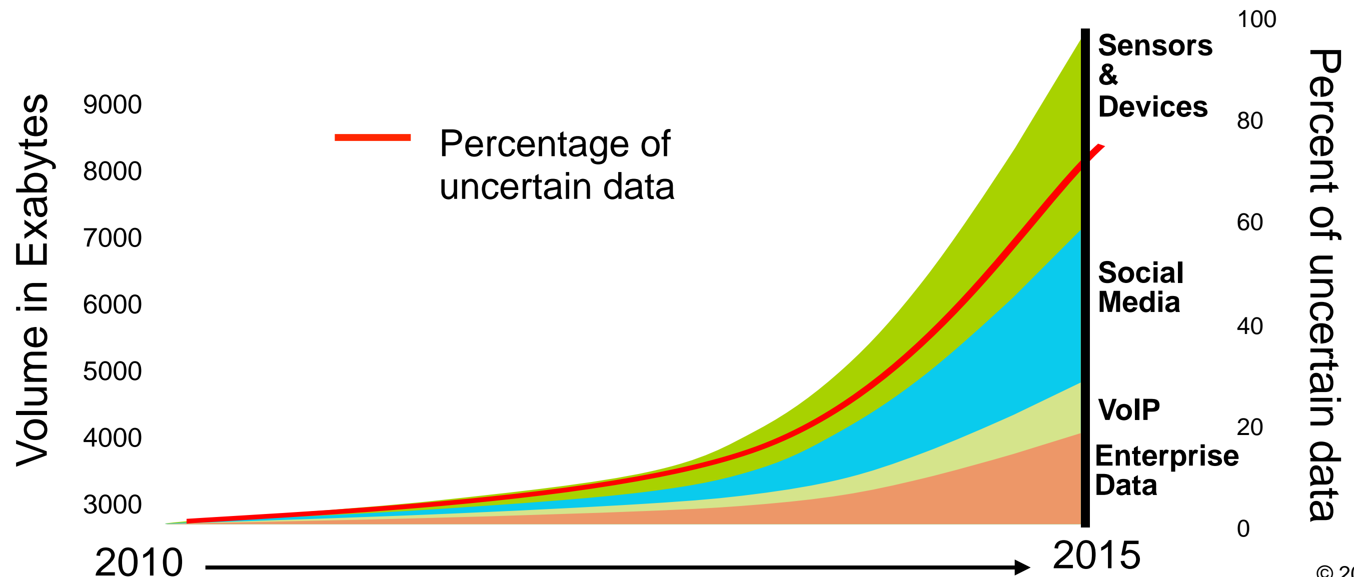
© 2015



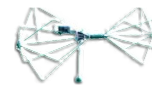
Big Data: Why we must move to a new era of computing.

Volume	Velocity	Variety	Veracity
			
Terabytes to exabytes of existing data to process	Streaming data, milliseconds to seconds to respond	Structured, unstructured, text & multimedia	Uncertainty from inconsistency, ambiguities, etc.

This is just the beginning.

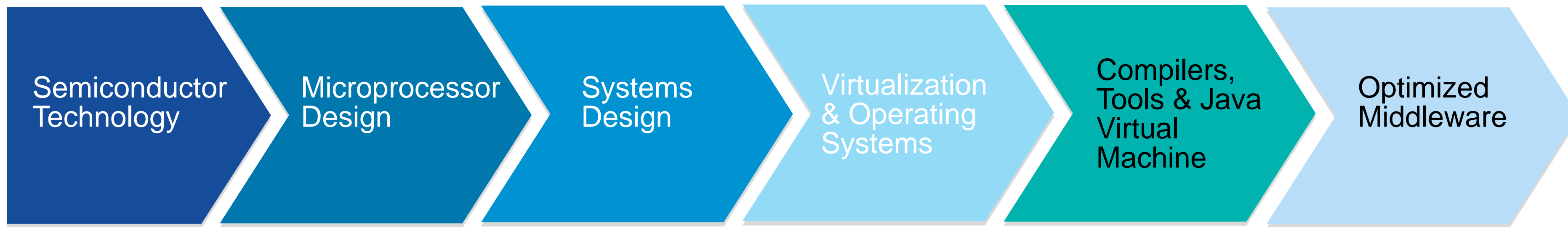


© 2015

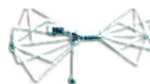


What is Technology?
zEnterprise EC12:

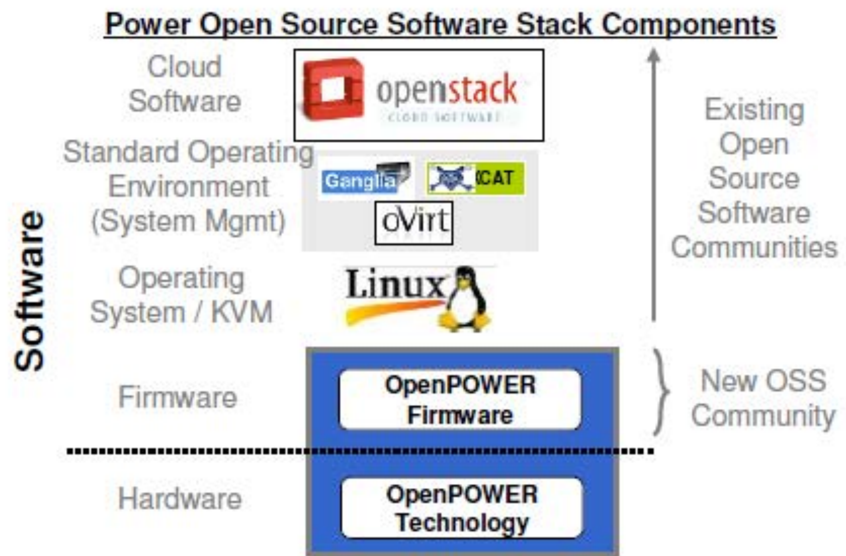
Linux
z/OS
z/TPF
z/VM
z/VSE
AIX
Linux
Java
Rational
WebSphere
Information Management
Tivoli
Lotus
Cognos

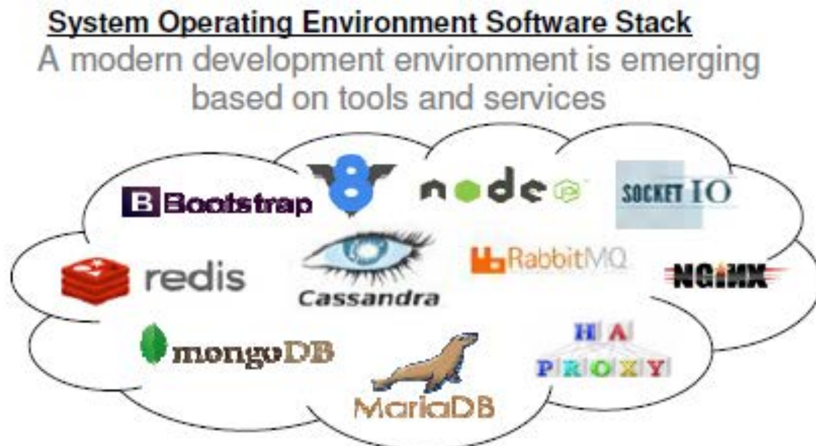


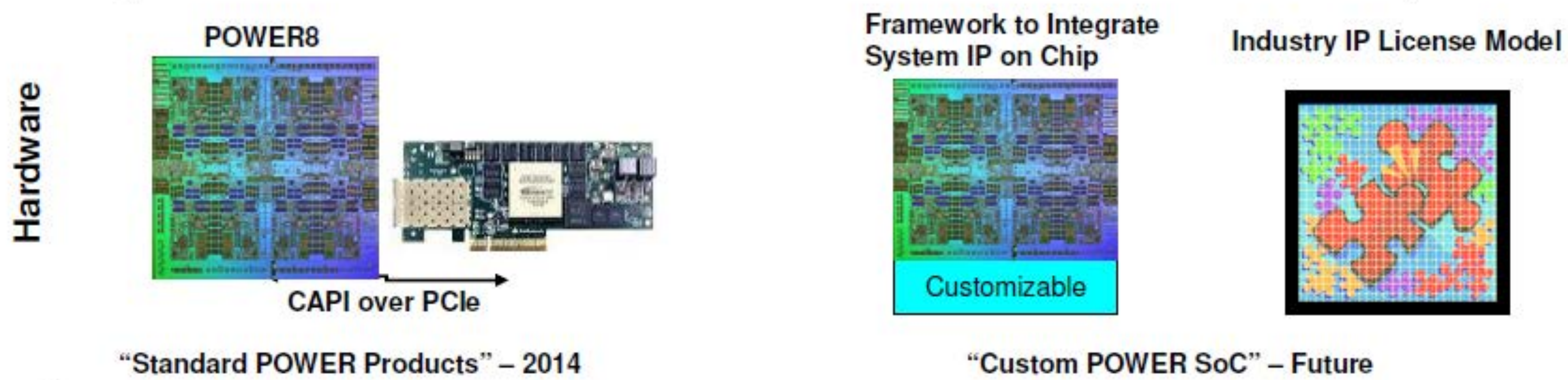
© 2015



Proposed Ecosystem Enablement





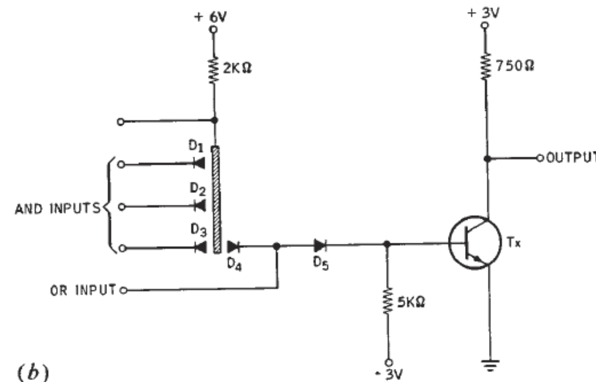


© OpenPOWER Foundation 2014

© 2015

50 Years of Mainframe – 1964 IBM S/360

Figure 1 AND/OR INVERT **logic module.** *(a) Completed AO1 module, without overcoating. (b) Logic circuit.*



Power dissipation in mW

	On	*Off*	
Resistors	28	19	(All *R*: 5%)
Transistor	7	0	
D_1, D_2, D_3	0	2	
D_4	1	1	
D_5	1	1	
Total	37	23	

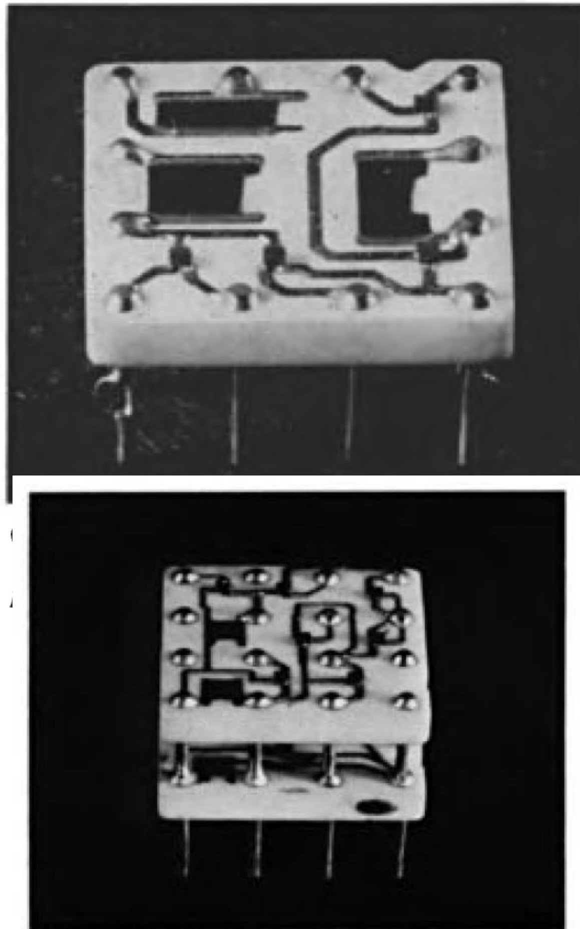
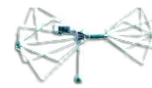


Figure 13 **Stacked module.**

Solid Logic Technology: Versatile, High-Performance Microelectronics Davis, E.M. ; Harding, W.E. ; Schwartz, R.S. ; Corning, J.J.
IBM Journal of Research and Development Volume: 8 , Issue: 2 Publication Year: 1964 , Page(s): 102 - 114

© 2015



POWER Die

Core Core Core Core Core Core
L2 L2 L2 L2 L2 L2
Local SMP Links
Accelerators
8M L3 Region
MemCtrl
L3 Cache and Chip Interconnect
MemCtrl
Remote SMP Links
PCI Gen 3 Links
L2 L2 L2 L2 L2 L2
Core Core Core Core Core Core

Figure 5.1.7: Annotated POWER8 die photo.

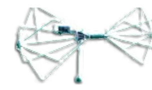
POWER8: A 12-core server-class processor in 22nm SOI with 7.6Tb/s off-chip bandwidth
Fluhr, E.J. ; et. al. Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International
Publication Year: 2014 , Page(s): 96 - 97

© 2015

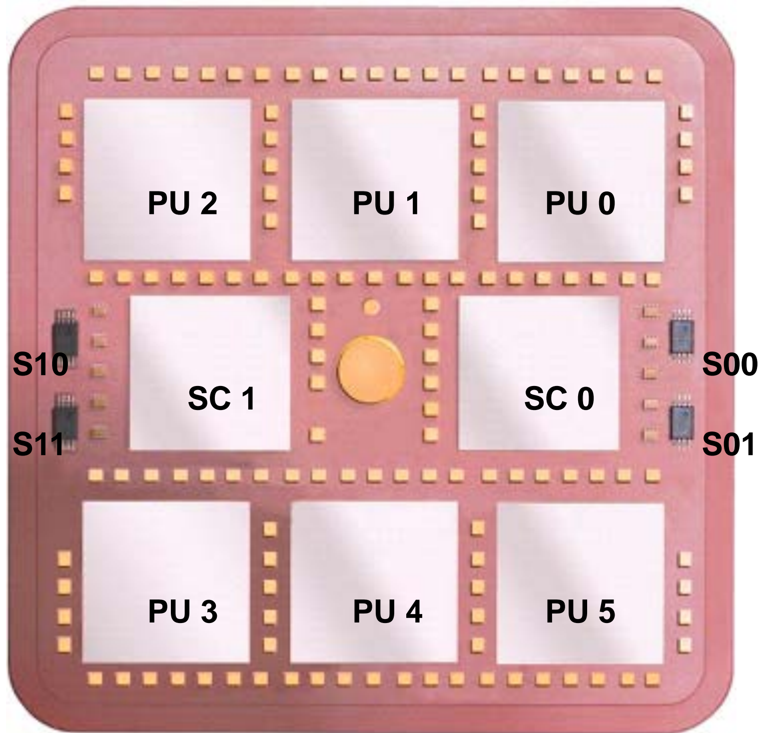
Ongoing Challenges for system packaging

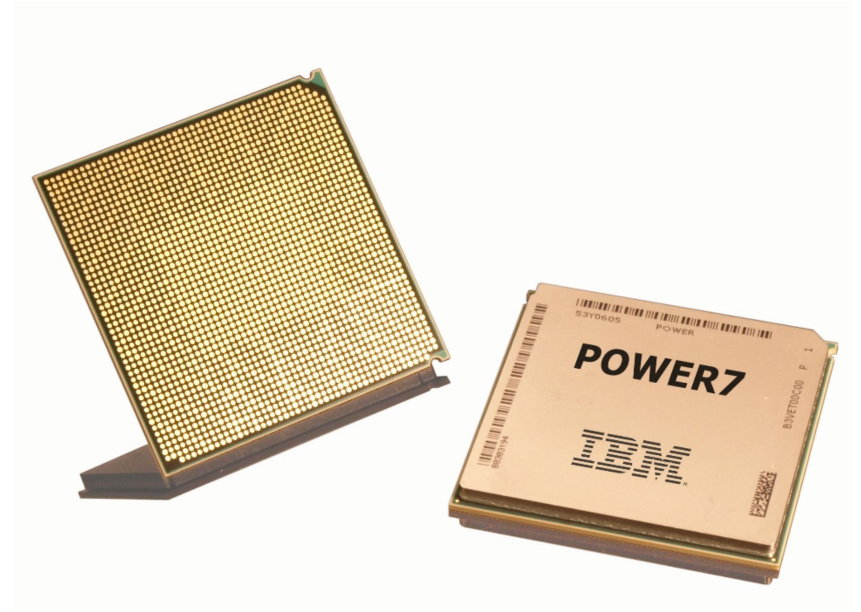
- Cost – development and product cost
- Physical form factor – Incremental changes
- Signal bandwidth density – Increasing quickly
- Voltage regulation (Power In) – Integrate closer to load
- Cooling (Power out) – Constant power density

© 2015



Processor Packaging

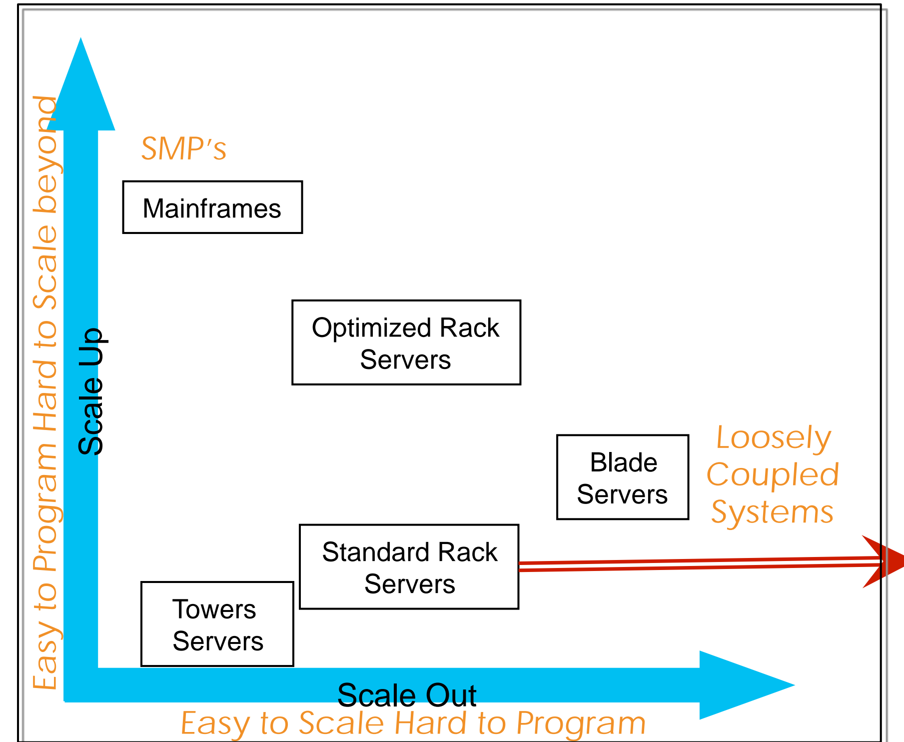




© 2015

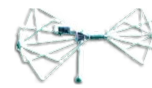
Traditional vs. "Cloudified" Hardware

- "Scale-Up"
  - Symmetrical Multiprocessing Systems
  - Large shared memory machines
  - Expensive to scale beyond a certain size
  - 4 / 8 / 16 / 32 sockets
  - 4U/10U/Rack Sized Systems
- "Scale-Out"
  - Loosely coupled systems
  - "Infinite" Scale
  - Mostly 1 & 2 sockets
  - 1U / 2U Form Factor (0.33/0.5/1 wide)



- Significant changes in programming & application paradigms
  → Hadoop/HDFS / NoSQL DB's….
- Open source software community driven
  → Linux / OpenStack …

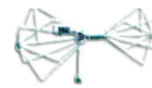
© 2015



Scale-out

PCIe G3 x8
2X SAS Ports
PCIe G3 x16
SAS Cntrl
PCIe G3 x16
Memory
Memory
POWER8
2xHMC Serial USB
PCIe G3 x8
Memory
POWER8
12 SFF Bays
6 SSD Slots
Slim DVD

http://www.ibm.com
© 2015



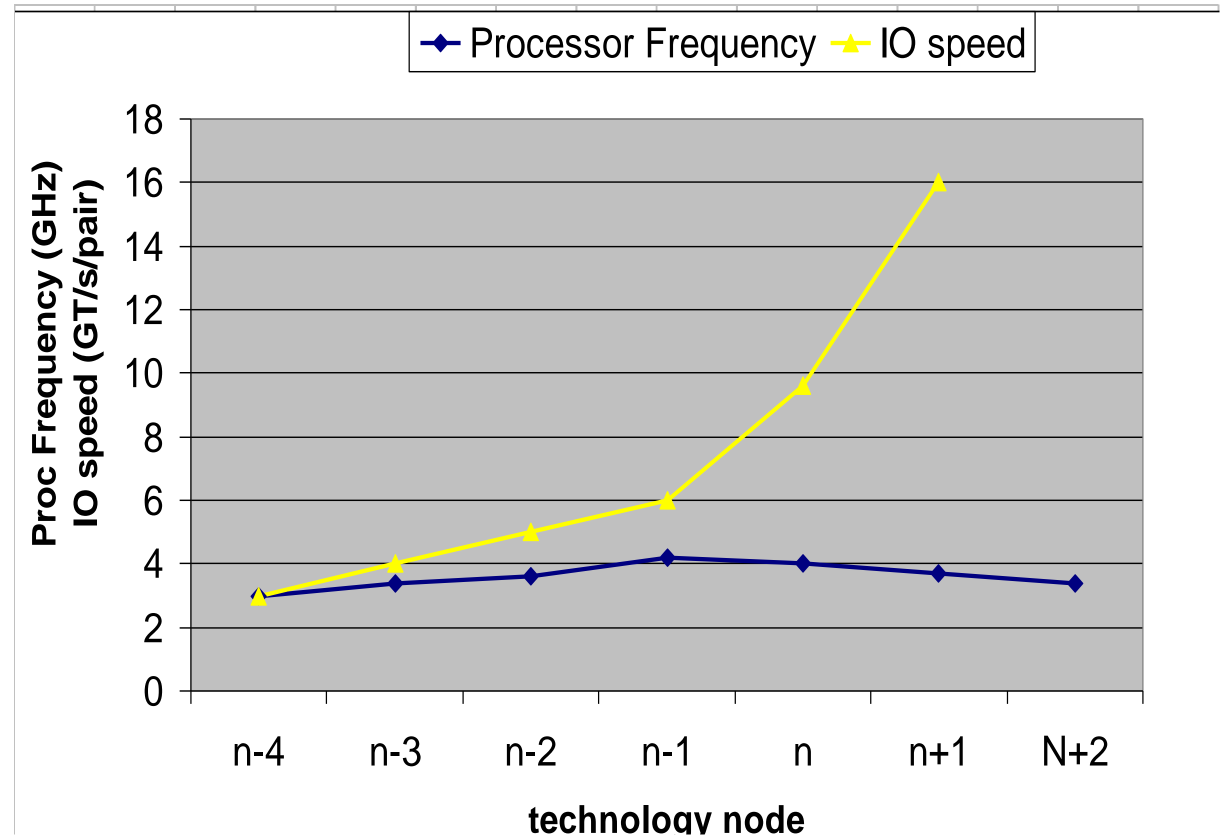
Scale-up

OSC
DCA
DCA
POL
Memory
POL
Memory
POL
POL
Memory
POL
POL
Memory
POL
POL
Memory
POL
Memory
POL
CP
CP
CP
CP
CP
CP
SC
Calypso
SC
5 x PCI-e
5 x PCI-e
3 x SMP
4 x GX
3 x SMP

© 2015

Data rates are increasing

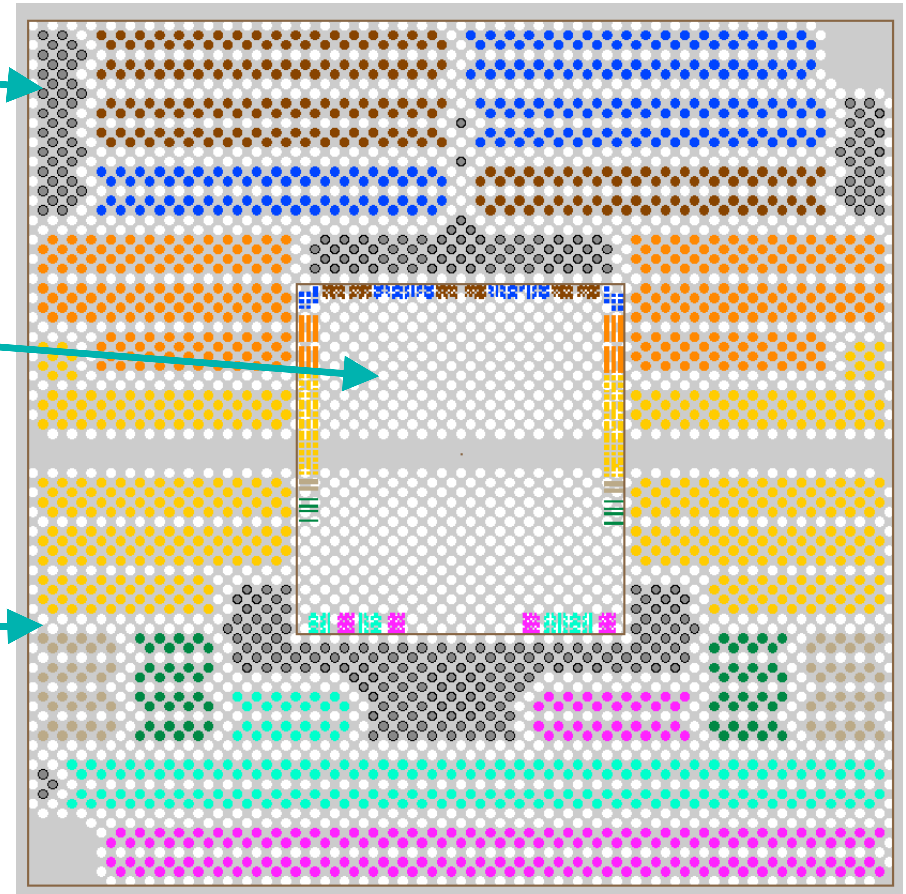
- Today
  - Proprietary – 12.4 GT/s
  - PCIe Gen3 – 8 GT/s
  - DDR3 – 2133 MT/s
  - SAS – 6 GT/s
  - Optics – 10 GT/s
- Soon
  - PCIe Gen4 – 16 GT/s
  - Proprietary 28 GT/s
  - DDR – 3200 MT/s
  - SAS – 12 GT/s
  - Optics – 25 GT/s
- R&D
  - ~50 GT/s
  - PAM4 vs NRZ
  - Tighter optics integration



© 2015

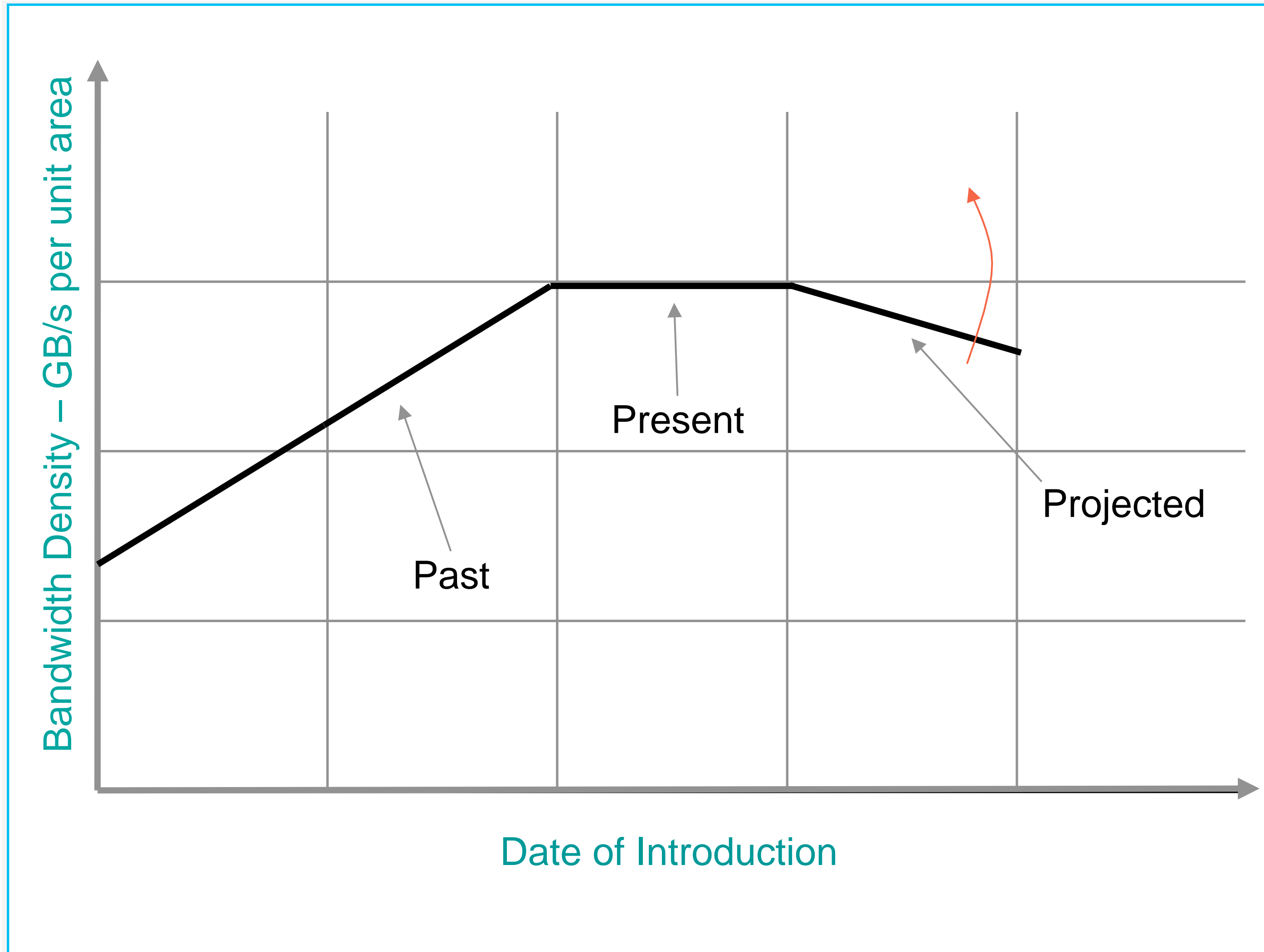
Socket Pin Assignment

- Signal Pins
  - More pins = more bandwidth
- Power Pins
  - Lower voltage levels require more pins
- Reference Pins
  - Higher frequencies require better isolation between signal pins

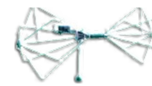


© 2015

Reaching a bandwidth breakpoint at the socket level

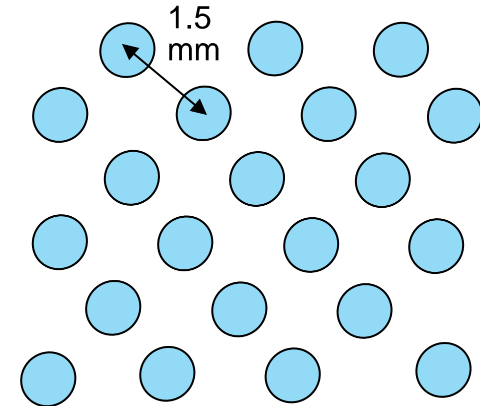


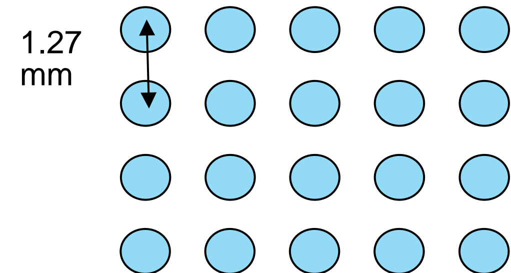
© 2015

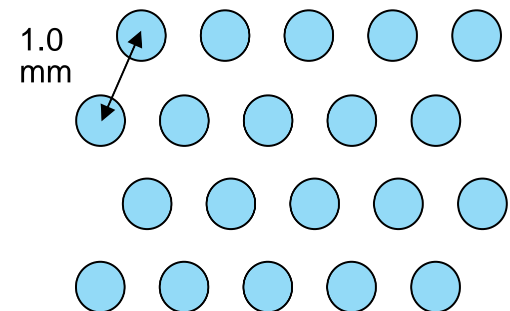


Pin Density Increases Incrementally

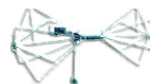
- Over 20 years
  - 1.5 mm min pitch interstitial
  - 50 mil (1.27 mm) square
  - 1 mm square
  - 1 mm min pitch hexagonal
- Sockets are pin limited
- Crosstalk needs to be managed



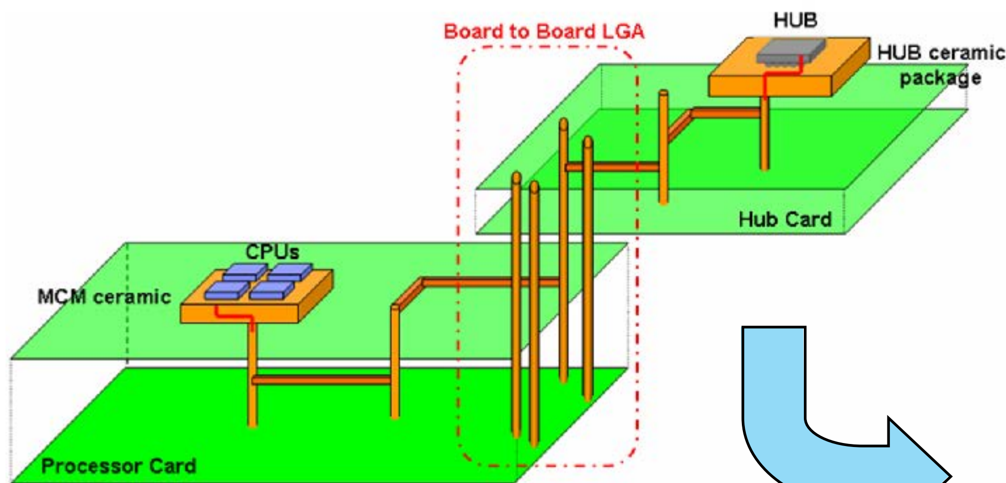




© 2015



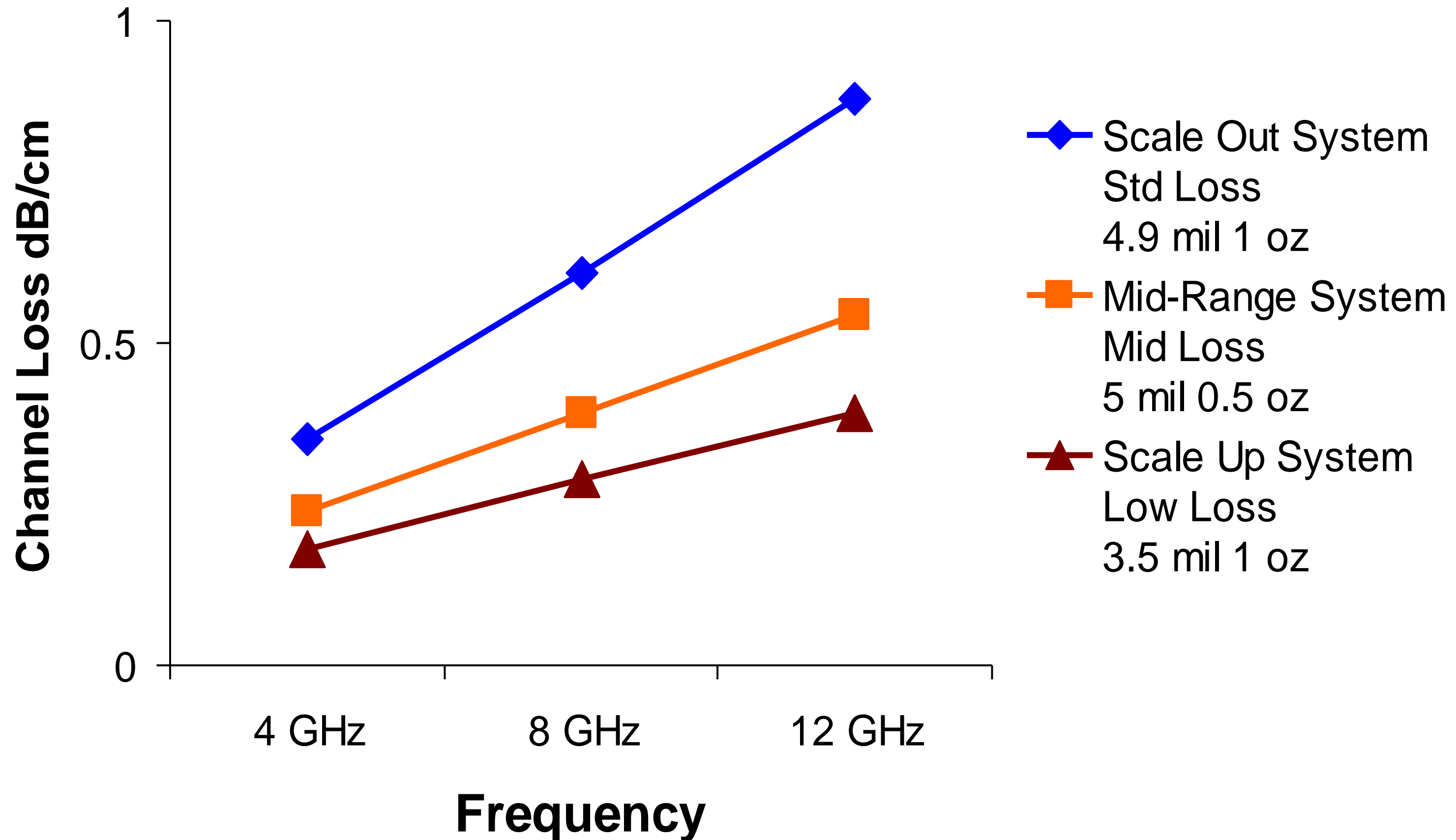
Form factor limitations



- *The system under analysis is composed by two PCBs, two MCMs and three connectors*
- *To represent it adequately 52 models are needed:*
- *1. W-elements to model the TL portions*
- *2. S-parameters (Touchstone) for the 3D parts (Vias and connectors).*
- *3. Mpilog Precompensation Driver macromodel*
- *4. Frequency step for touchstone: 50 MHz*
- Total channel length ~ 70cm

© 2015

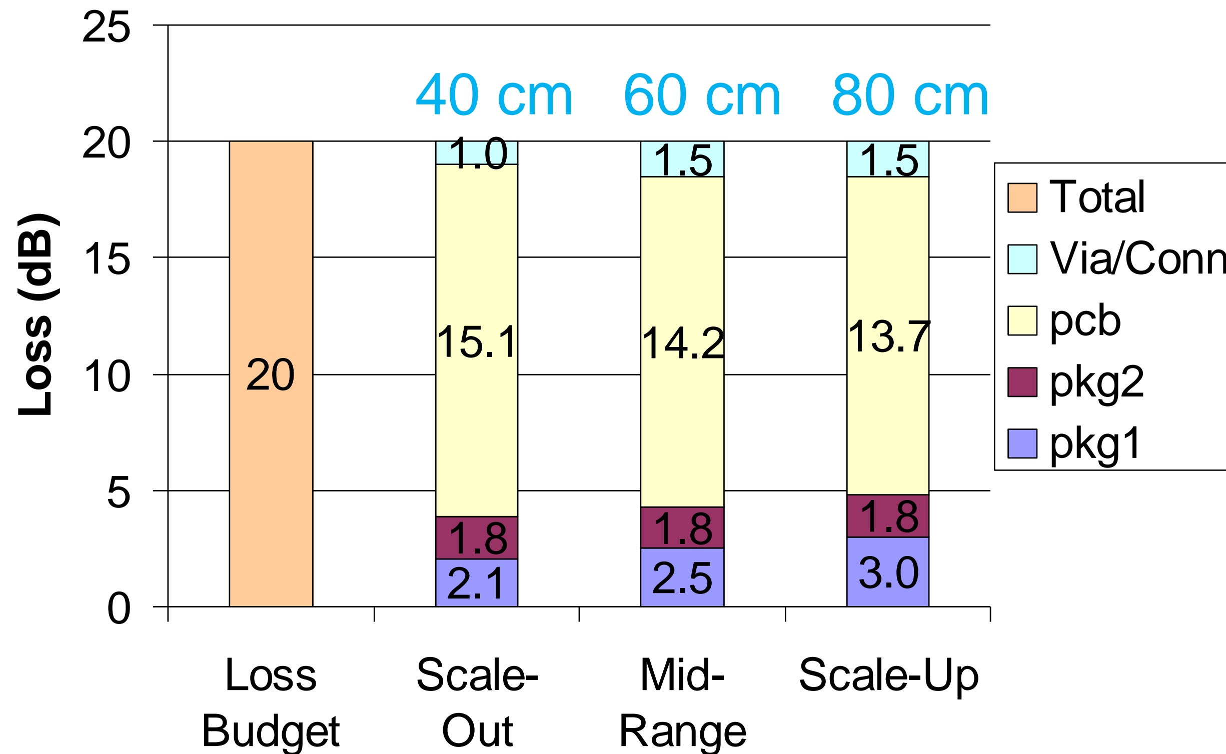
PCB Technology Choices
Same speed, different technology



All systems are 8 GT/s meet 20dB total channel loss

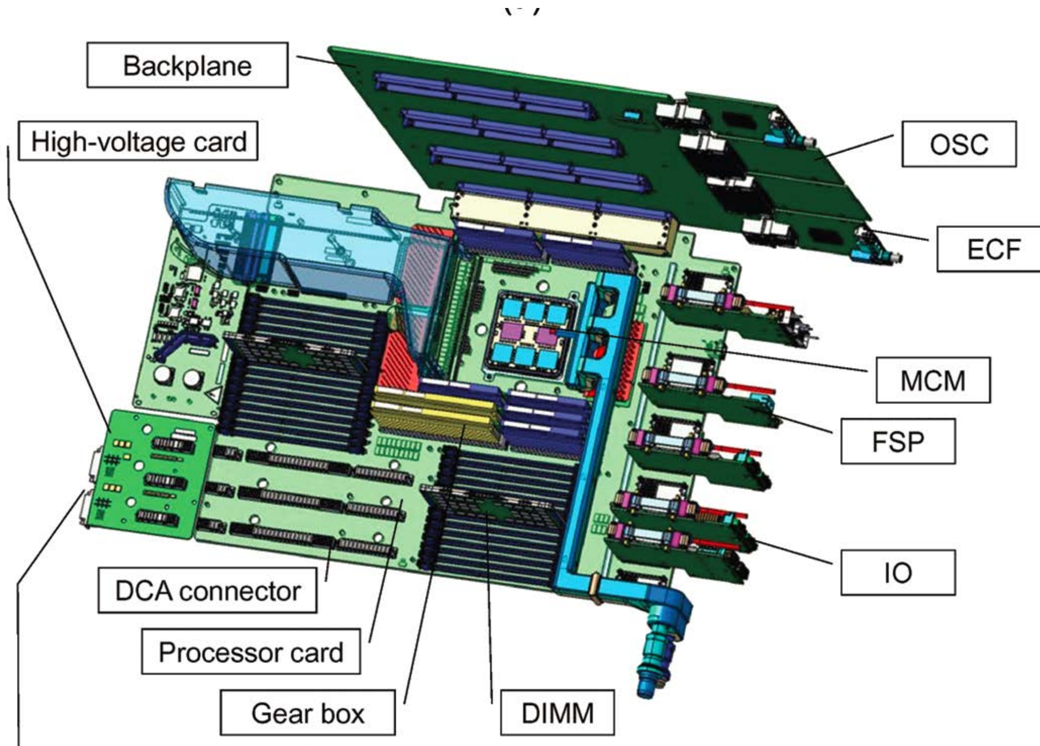
© 2015

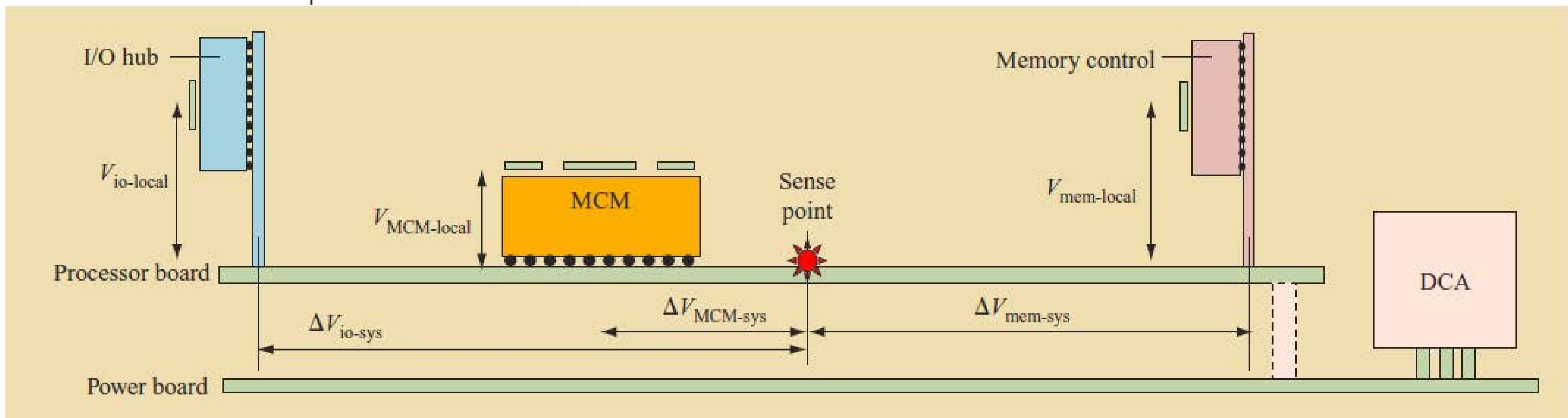
Impact of PCB loss – Size of system, length of PCB trace



© 2015

z196/EC12 Compute Cage

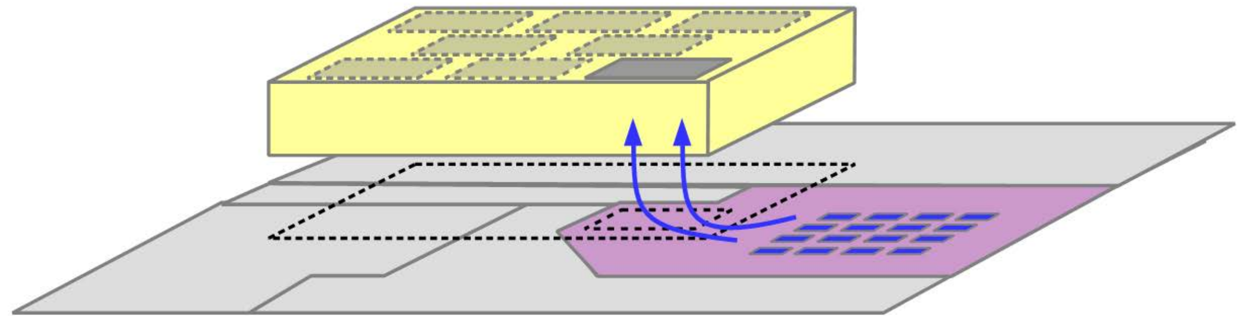




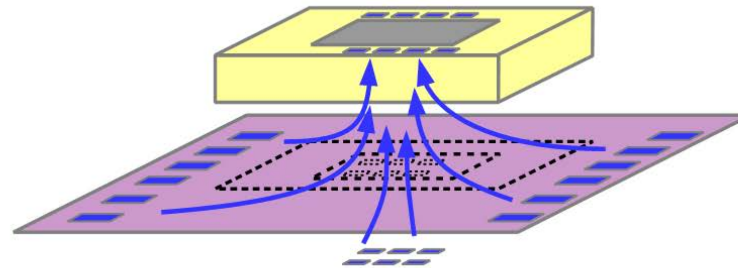
Decoupling improvement

Improved power distribution often counters increased density

zEC12



z13



© 2015

Deep trench technology – DT capacitors

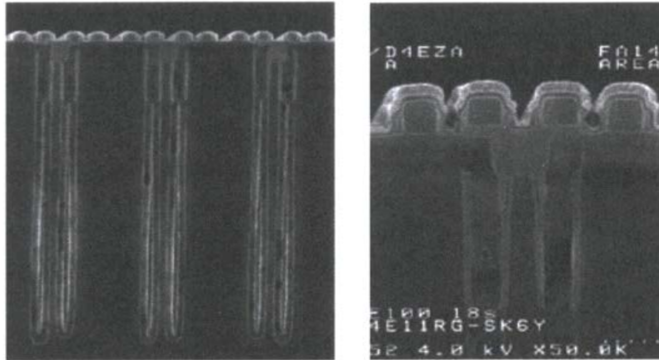


Figure 2 – Cross section of DT Capacitor (with Word Line and Passing Word Line on top)

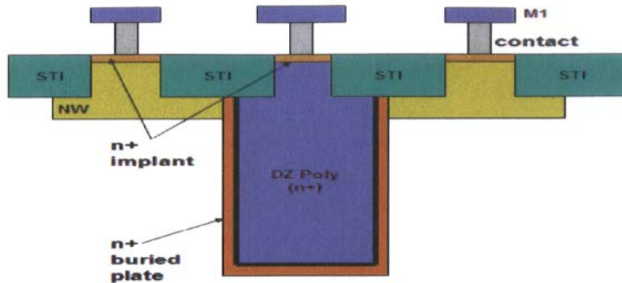
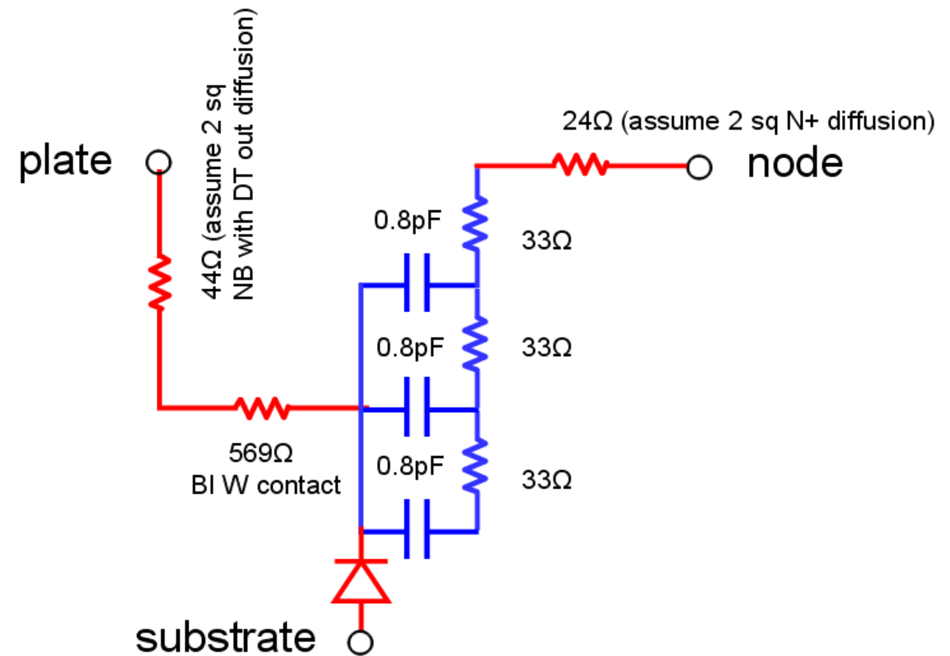


Figure3: Schematic cross section of DZ capacitor

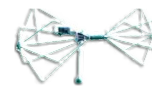


Circuit model for a book having 30 base cells

C. Pei, "A novel, low-cost deep trench decoupling capacitor for high-performance, low-power bulk CMOS applications," 9th ICSECT, pp. 1146-1149, 2008

© 2015

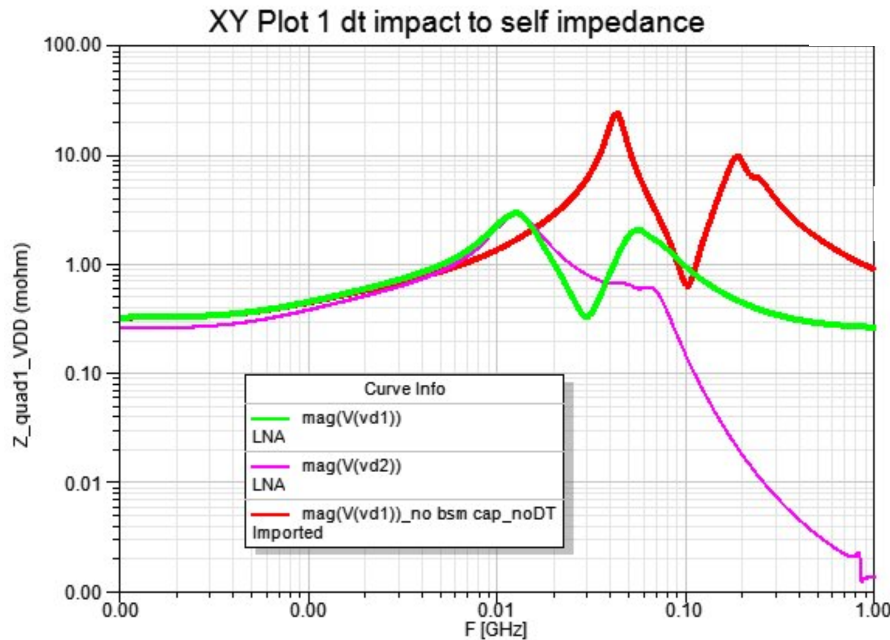
IEEE Advancing Technology for Humanity

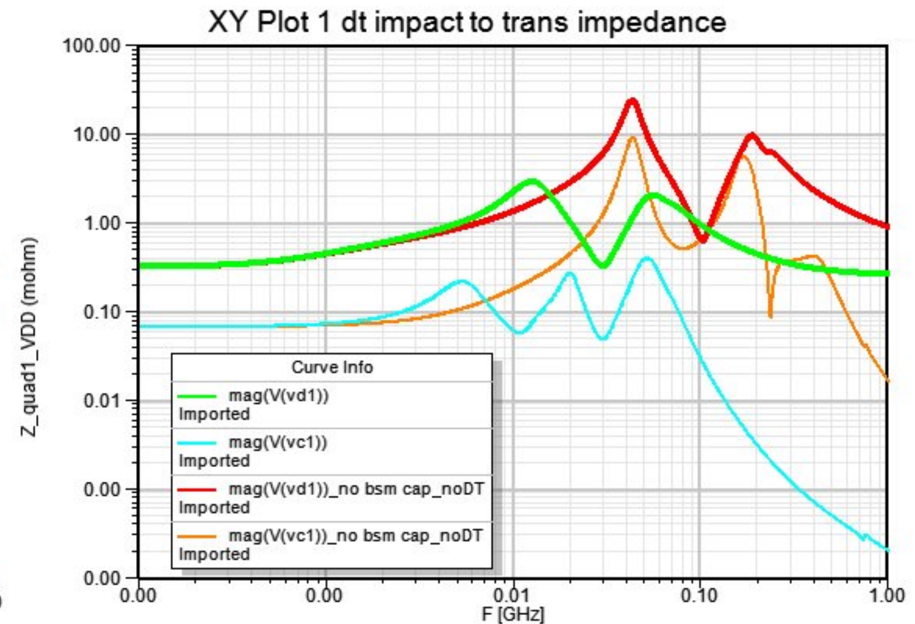


IBM

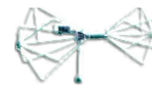
Power issues : On-chip Caps – Deep Trench

- Deep Trench capacitors
  - Enables significant increase (~20uF) in on-chip decoupling.
  - Mid/high-freq noise significantly reduced
  - Facilitates on-chip voltage regulation

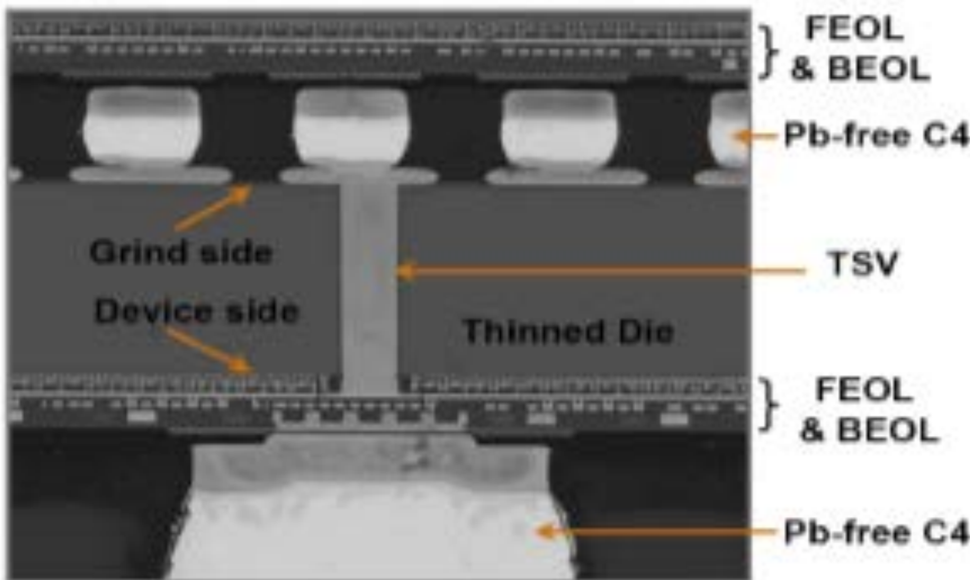


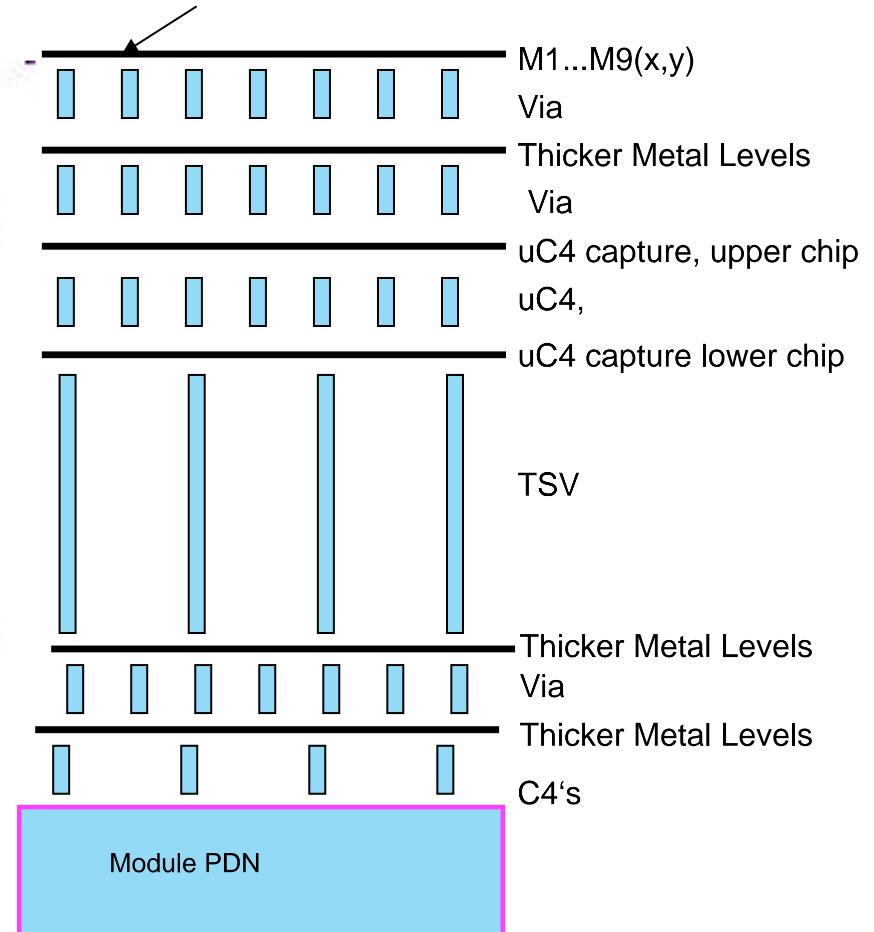


© 2015



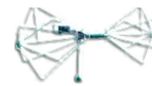
3D chip stacking modeling





Knickerbocker, et. al.
2.5D and 3D Technology Challenges and Test Vehicle Demonstrations
ECTC 2012

© 2015



The Case for Optics

Cable Length Limitation

High Attenuation loss

Higher power consumption per bit

High EMI , Cross Talk and Latency

Very low attenuation loss

Low power consumption per bit

No EMI, Cross talk, low latency

Long distance cable support

Economics of Bandwidth and Distance

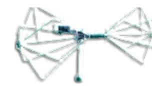
Copper

Optics

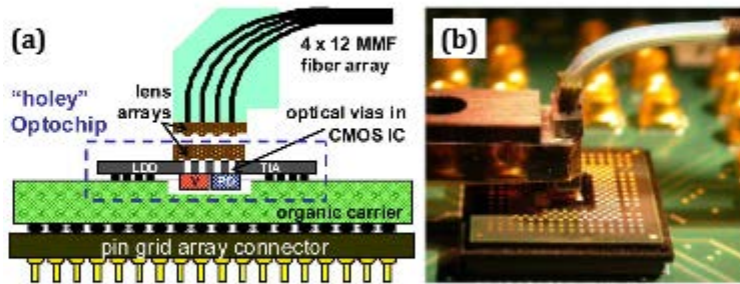
Up to 80Km for Ethernet, 100Gbps Bandwidth at low power

- Increasing benefits with optical, but products generally cost more than copper
- Optics less expensive when integrated with silicon - Silicon Photonics
    - Photonics integrated into silicon base
    - Reduces cost and provides higher bandwidth

© 2015

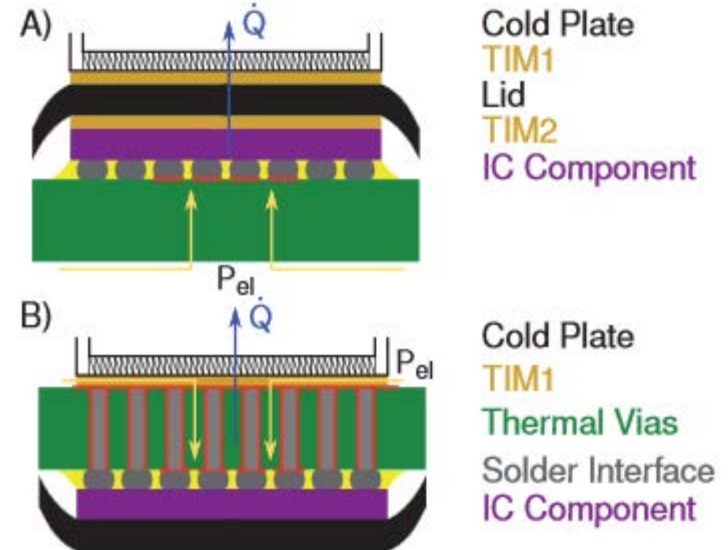


Future Integration



Increasing Bandwidth Density in Future Optical Interconnects

B. G. Lee, C. Baks, F. E. Doany, C. Jahnes, R. John, D. M. Kuchta, P. Pepeljugoski, A. V. Rylyakov, C. L. Schow, S. Assefa, W. M. J. Green, Y. A. Vlasov, J. A. Kash
IBM T. J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598
bglee@us.ibm.com



CIPS 2014, February, 25 – 27, 2014, Nuremberg/Germany

Laminate with Thermal - Power Insert for Efficient Front-Side Heat Removal and Power Delivery

Dominic Gschwend*, Timo Tick*, Stefano Oggioni†, Stephan Paredes*, Keiji Matsumoto‡, Manish K. Tiwari¶, Dimos Poulikakos¶ and Thomas Brunschwiler*§

© 2015

Ongoing Challenges for system packaging

- Cost
  - –Technology Reuse
- Physical form factor
  - – Form factors change slowly
  - – Scale-in functional integration
- Signal bandwidth density
  - – Frequency per lane
- Voltage regulation (Power In)
  - – Regulation moves closer to load
  - – More effective decoupling capacitors
- Cooling (Power out)

© 2015

Conclusions

- IT boundaries are becoming less clearly defined
  - Processor, Storage, Networking Integrated
- Systems
  - Cloud, Analytics, Mobile, Social, Security
- System Hardware
  - Drives system capacity with cores and computing capacity
  - Drives interconnect bandwidth at processor, node, system and network level
- All electrical performance elements must be balanced
  - Bandwidth
  - Power Distribution
  - Thermal
- Technology enables the innovation that systems provide, we must choose wisely.

© 2015