Systems and Technology Group

Signal Interconnect Trends and Challenges Inside the CEC

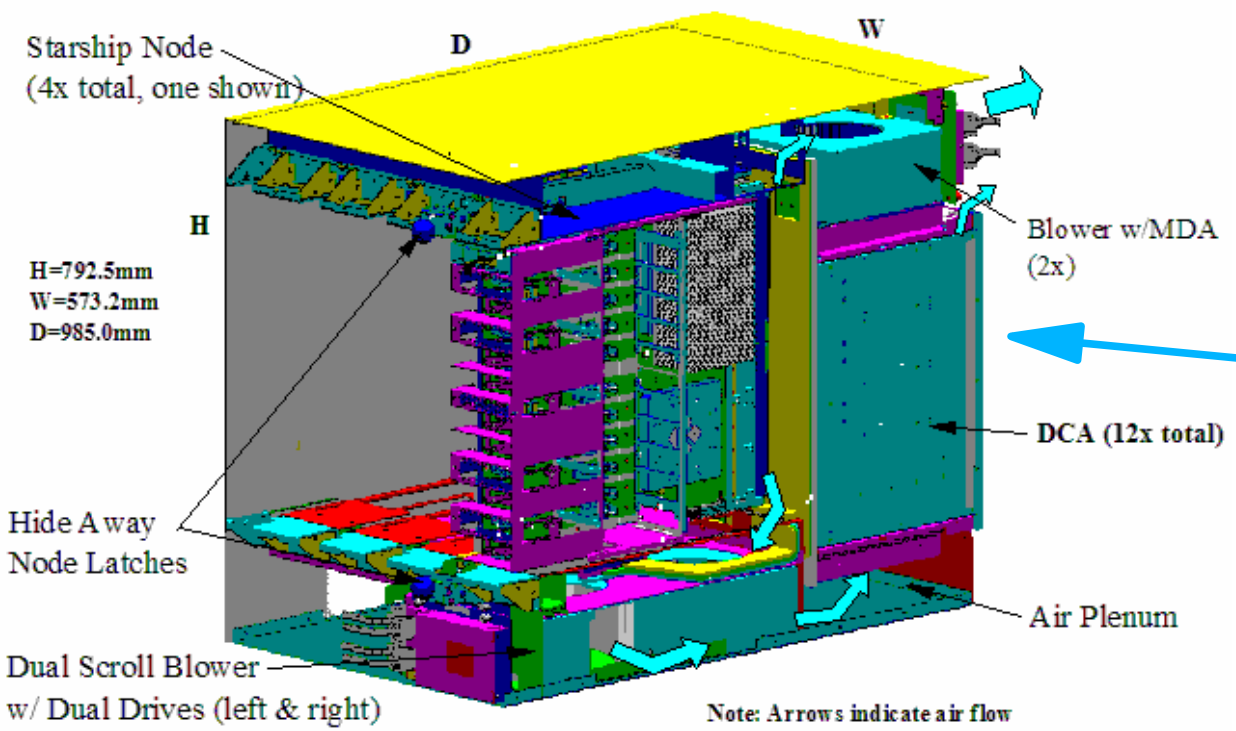
George A. Katopis
Distinguished Engineer STG
IBM Corp.

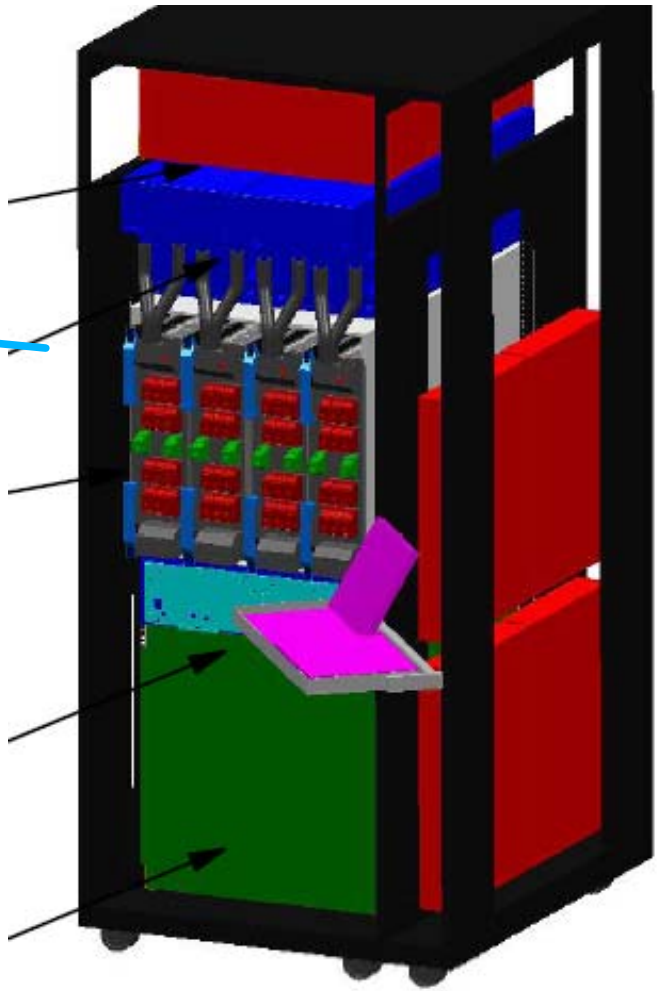
Contents

- ***CEC Description and Bandwidth requirements***
- ***The four Challenges***
- ***Comparisons Yardsticks and Tools***
- ***SE to DS I/O comparisons***
- ***The Socket world and the MCM technology***
- ***Conclusions***

G. A. Katopis

CEC description

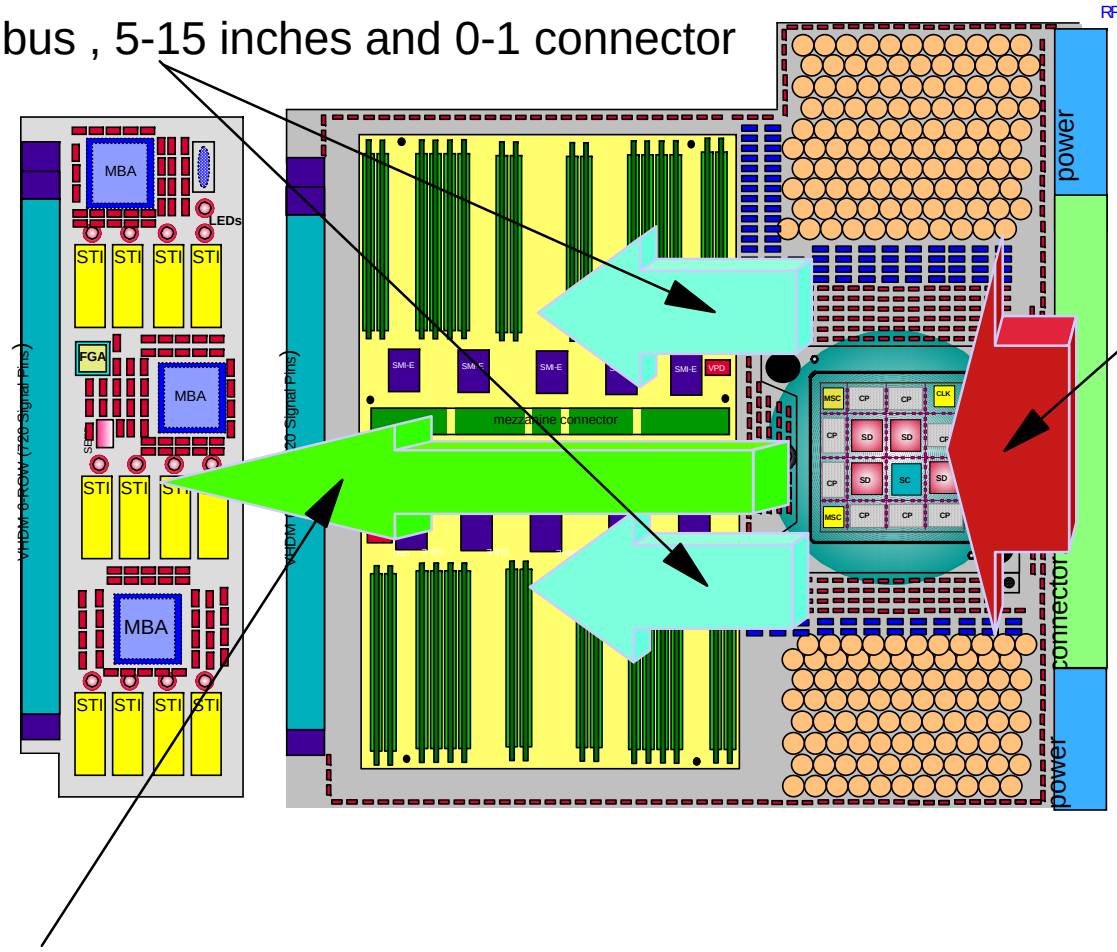




The buses and their topology on a node

200 Bytes off MCM , 256 bytes on MCM for z-servers

The memory bus , 5-15 inches and 0-1 connector

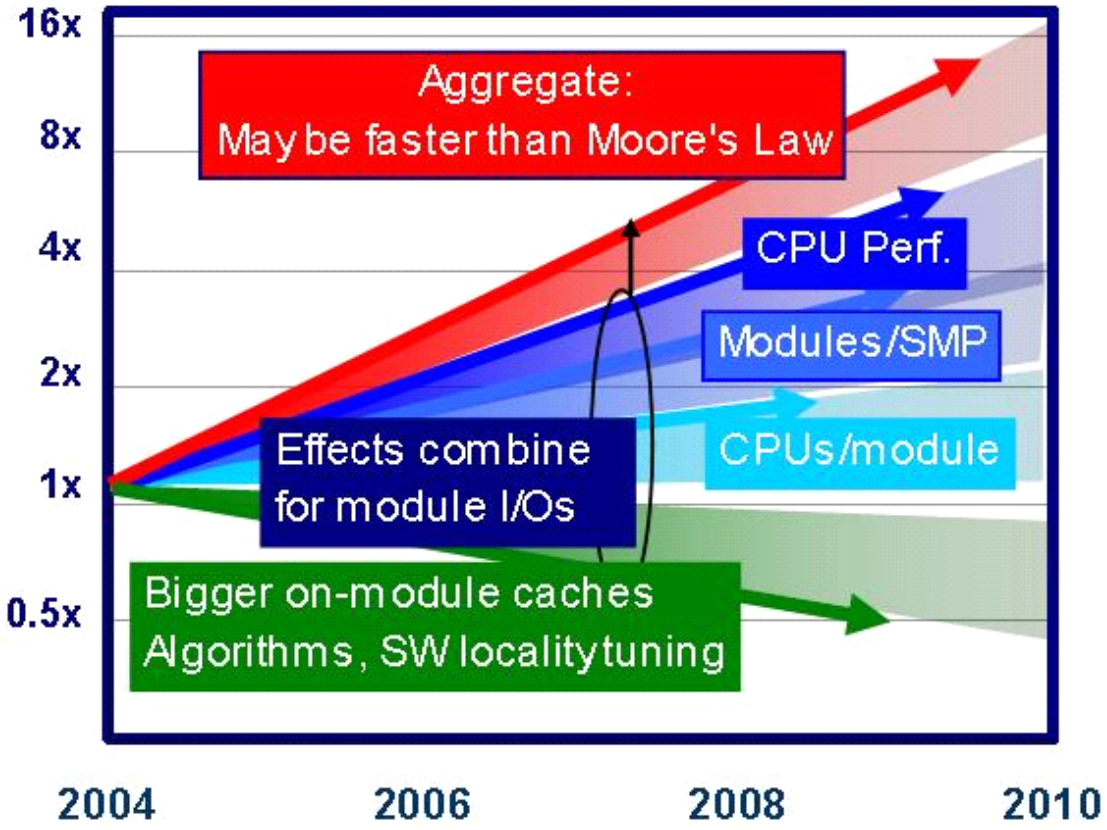


The ring bus
connecting blades in CEC
30 inches with 2 connectors

The I/O hib bus, 15 inches one connector

G. A. Katopis

The CEC bandwidth component and requirements

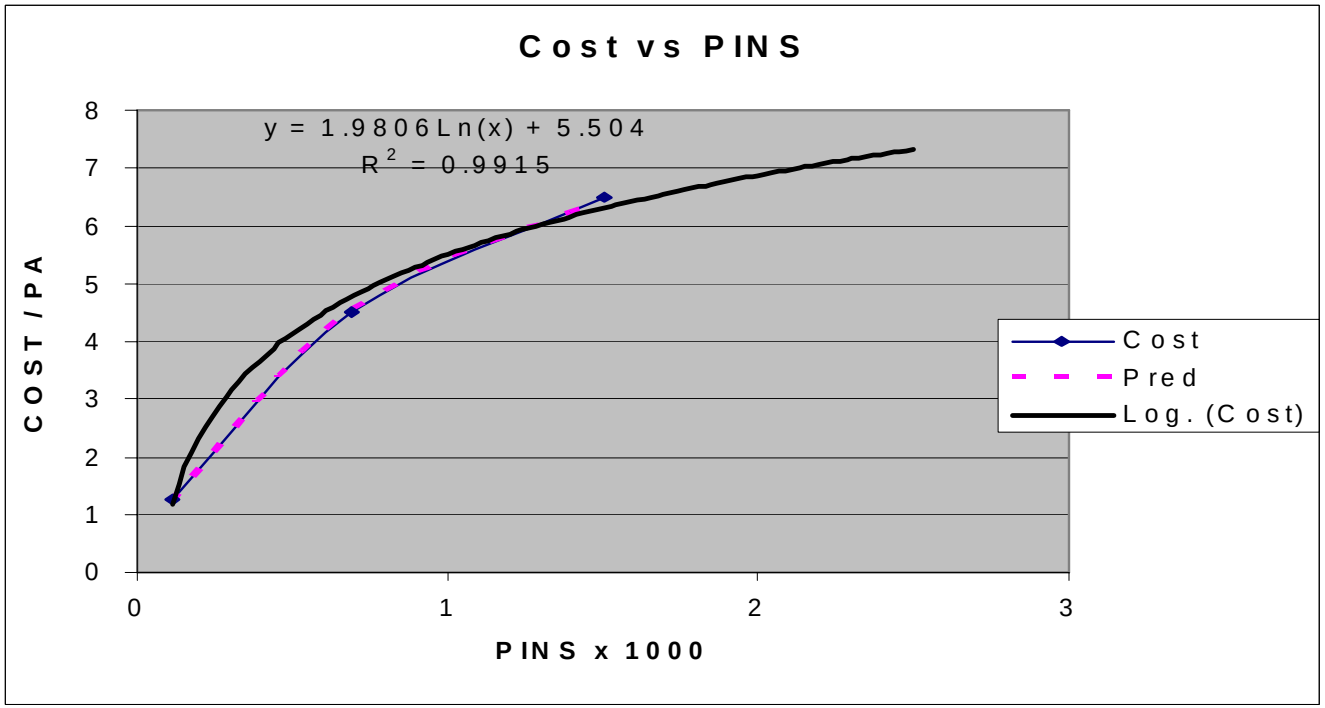


TWO ways to do it

More pins
and/or
higher frequency of interconnect

Figure 1: MCM signal I/O requirement trends

The relative cost challenge of wider buses



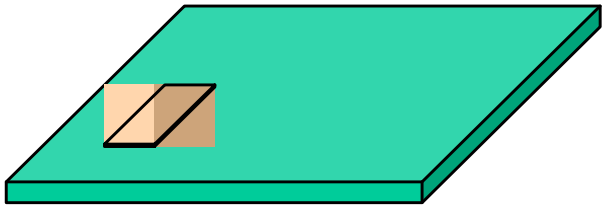
1.4 cost increase per unit bandwidth increase

The area challenge and extendability thoughts

Type of TX/ Rxcircuit	*Topology of TX/RX circuit*	*CMOS node*	*Drive distance in inches*	*mm^2/lane*	*max Gb/s/lane*	*Gb/s/mm^2*
EI1	**SE**	**135**	**20**	**0.025**	**1.0**	**40**
EI2	**SE**	**95**	**30**	**0.012**	**1.25 2.5(MCM)**	**104 208(MCM)**
EI3	**SE**	**65**	**30**	**0.039**	**3.2**	**144**
Serdes	**DS**	**135**	**6**	**0.4**	**2.8**	**7**
Typical Serdes	**DS**	**95**	**6**	**0.25**	**5.4**	**22**

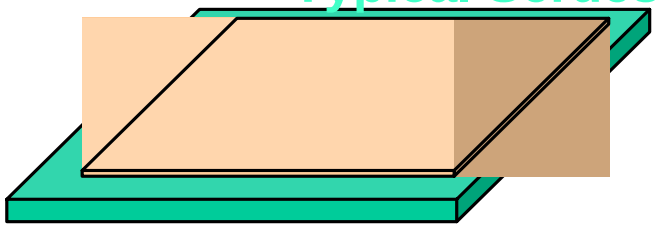
400 mm^2 chip
1800 SIO

EI3



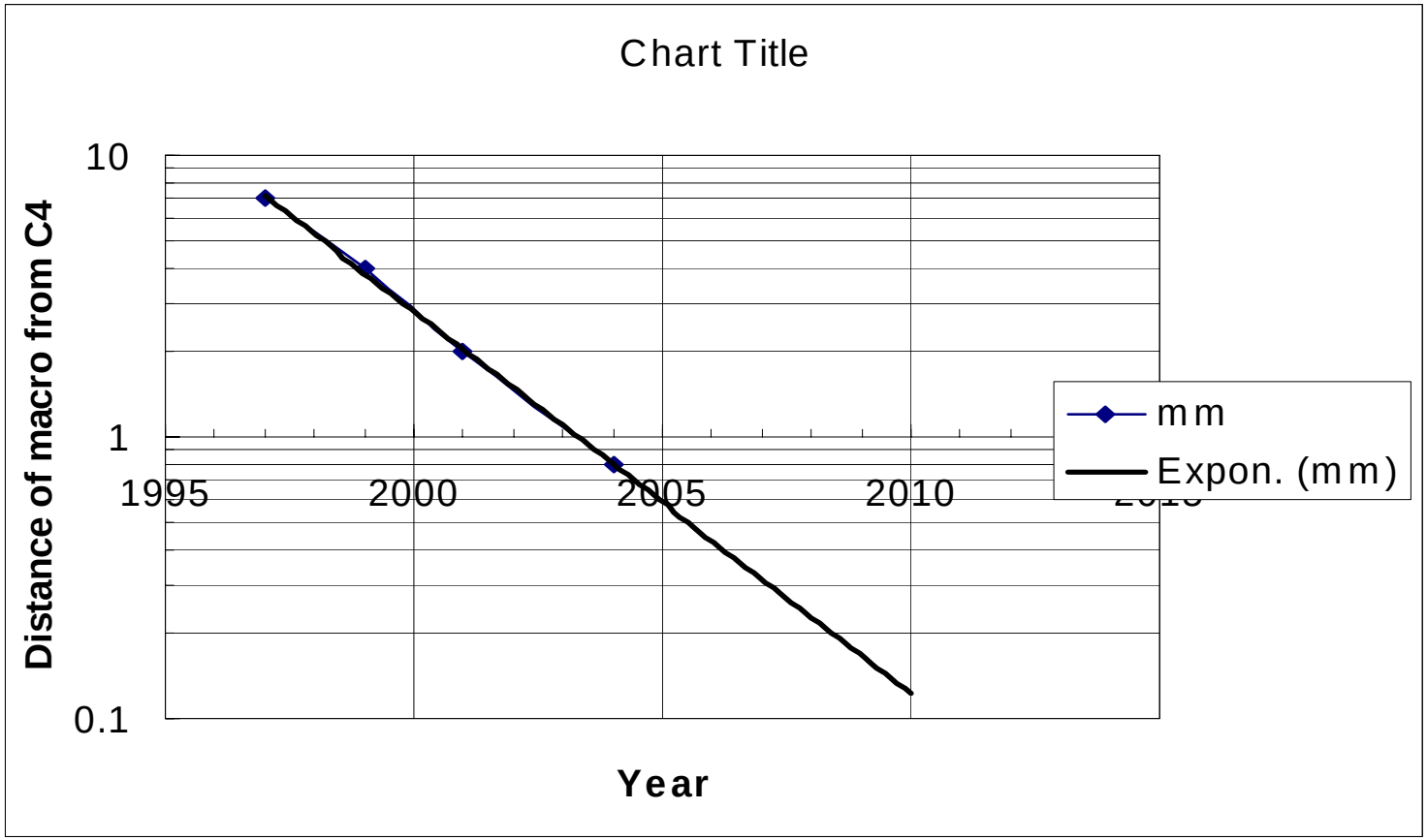
9% of area
20% of power budget of 200W

Typical Serdes

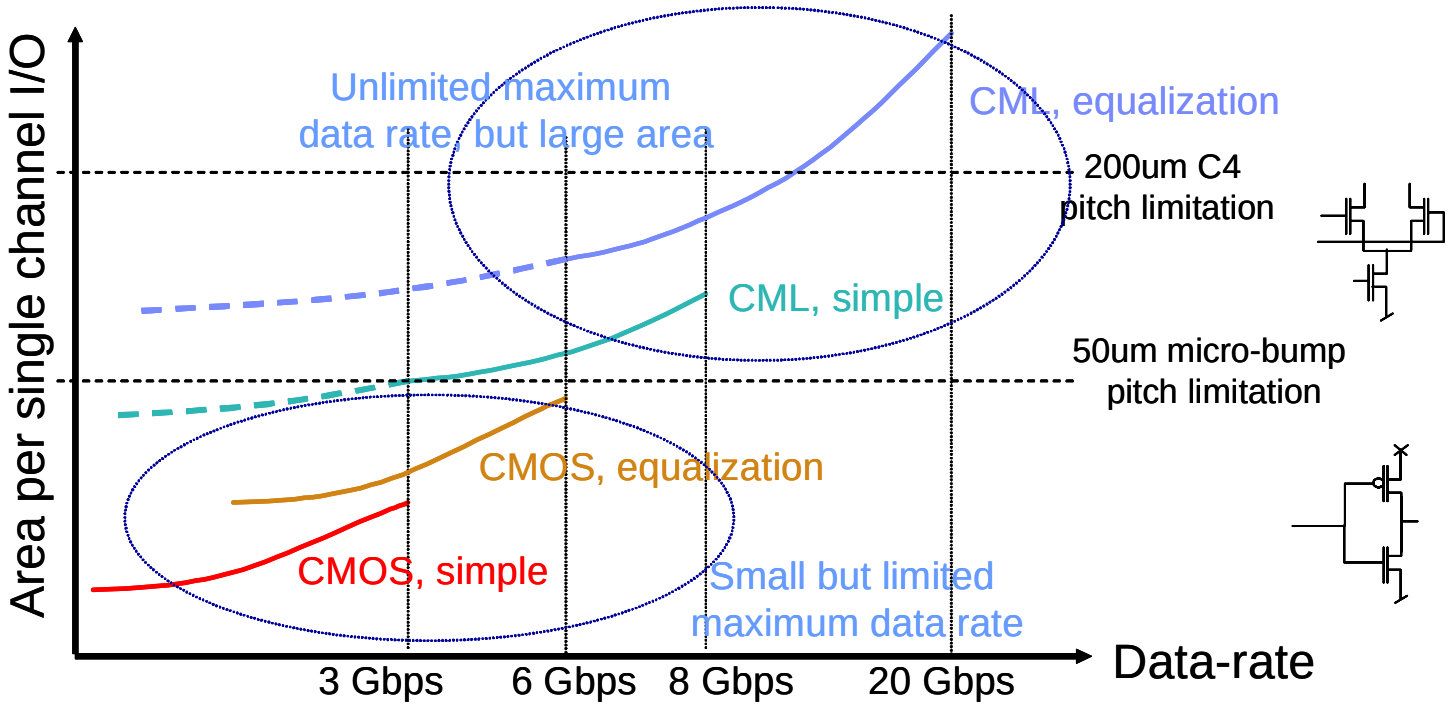


83% of area
43%

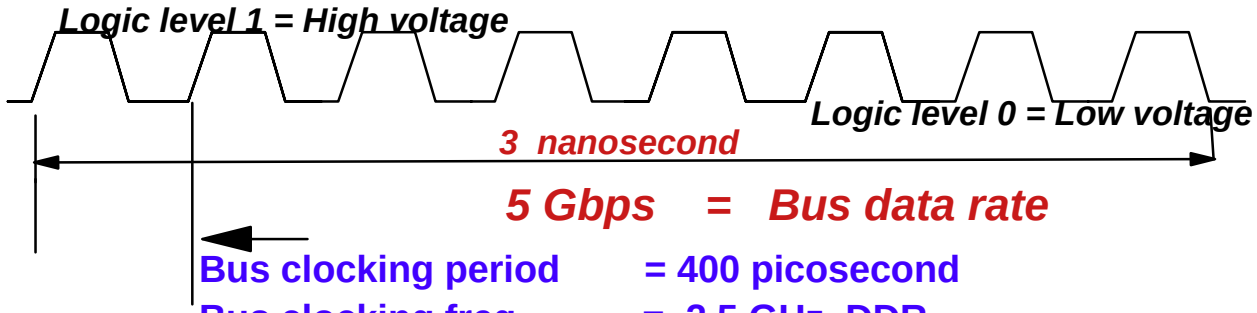
The challenge of a level field is the peripheral I/Os



Looking at a technical crystal ball



The EM modeling challenge



Eq. PU Freq. of Oper.. GHz	Time	Data rate for 2:1 bus gear Single Ended	Data rate for 1:1 bus gear Differential	Bus freq. in GHz	Interc. Model freq. GHz
2	Now	1 Gbps	2 Gbps	0.5-1	5-10
5	Next Gen.	2.5 Gbps	5 Gbps	1.25-2.5	12.5-25
10	2010	5 Gbps	10 Gbps	2.5-5	25-50

The tools and verification measurements must be acurate for up to 50GHz in 2 years

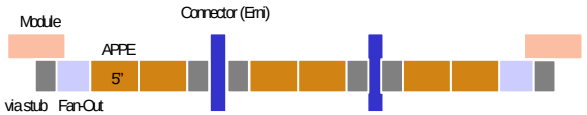
Comparison Metrics or why no comparisonn is possible

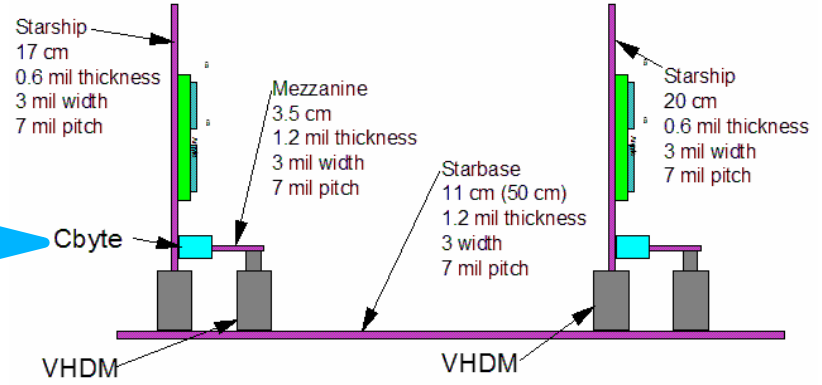
- **Link Performance**
  - ***Bit rate per lane (in Gbps)***
  - ***Aggregate bandwidth for 16 bits (Gbps) or for a given area (fixed number of pins) at the bottom surface of module***
- **Additional link attributes considered**
  - ***Power per lane (in mW)***
  - ***Power efficiency (in mW/Gbps)***
  - ***Chip Area per lane (mm^2)***
  - ***Chip Area per 16 bits (mm^2)***
  - ***Pins required per 16 wide bus***
- **Attributes not considered because of lack of time or design specificity**
  - ***Asynchronous link architectures***
  - ***Clock subsystem issues and requirements***
  - ***Impact of switching noise and ground return noise***
  - ***Design specific***
    - Number, value, tolerance of required voltages
    - Number of required pins for voltage and ground, and any placement restrictions
    - Number and type of sideband signals (strobe, clock, Vref forwarding, etc.) required
    - Number of layers of last metal required
    - Precision of on-chip resistors required
    - Number or type of thick(er) oxide layers required
    - Requirements for additional on-chip decoupling and ESD protection

The physical comparison structures

CEC Test Channels for Single/Diff Compare

1. CEC-CEC Link : ~30"





2. CEC-MEM Link : ~15"

4. CEC-CEC on Board Link : ~5"

1st level packages description

Max Freq 10GHz	Max Freq 10GHz
Package designed as differential: 2 diff pairs	Package designed as single ended 2 SE lines
Line length: 50 mm **Line width: 38um/50 um** **Line pitch within pair/ between pairs: 280 / 450 um** **C4 Via: 64 um and 90 um dia** **Top Via height: 0.22mm / 0.710 mm** **Sig-Sig TOP Via pitch between pairs: 450 um /Controlled Impedance Via (CIV)** **BSM pitch: 1mm**	Line length: 50 mm Line width 66 um Line pitch: 403.2 um C4 Via 90um diameter Via height: 1.0 mm Sig. to Sig. top side Via pitch: 200 um BSM pitch: 1mm Layer thickness 45 layers at 150um/layer
Er = 9.8, tand = 0.0005	Er=9.5, Rho=10.0e-6 uohm-cm, tand=0

Generic I/O core model

Parameter	Value (2-level)
Voltage Driver level at Tx	600
Receiver bandwidth	Data rate / 2
Receive latch error threshold	15
Receiver non-deterministic amplitude noise	2mV average RMS
Receiver excess deterministic noise	0mV p-p
Receiver AGC level	280
Maximum AGC gain	3
System non-deterministic time jitter (RJ)	0.62 UI RMS
System deterministic time jitter (SJ)	10 UI p-p
CDR algorithm	2 sample/UI DLL
Transmit and Receive IC parasitics	None
FFE taps	4 or 1, 1 precursor tap
DFE taps	5 or 0

G. A. Katopis

Comparison results for maximum data rates vs distance

Signal	Xtalk	EQU	5	Ratio	15	Ratio	30	Ratio
Single	FEXT	NO	5.5Gb/s		2.9Gb/s		2.4Gb/s	
Diff	FEXT	NO	7.0Gb/s	1.3	3.8Gb/s	1.4	2.5Gb/s	1.1
Single	FEXT	4	5.7Gb/s		4.3Gb/s		3.6Gb/s	
Diff	FEXT	4	17.4Gb/s	3.1	10.6Gb/s	2.5	6.5Gb/s	1.9
Single	FEXT	45	6.4Gb/s		4.9Gb/s		4.3Gb/s	
Diff	FEXT	45	21.0Gb/s	3.3	14.6Gb/s	3	9.1Gb/s	2.2
Single	OFF	45	25Gb/s		15.5Gb/s		10.2Gb/s	
Diff	OFF	45	21.5Gb/s	0.9	15.2Gb/s	1.0	9.6Gb/s	1.0

No stubs at end of nets
- No delta -I considered
- At 30 inches length
- differential drive has

no or little advantage over SE

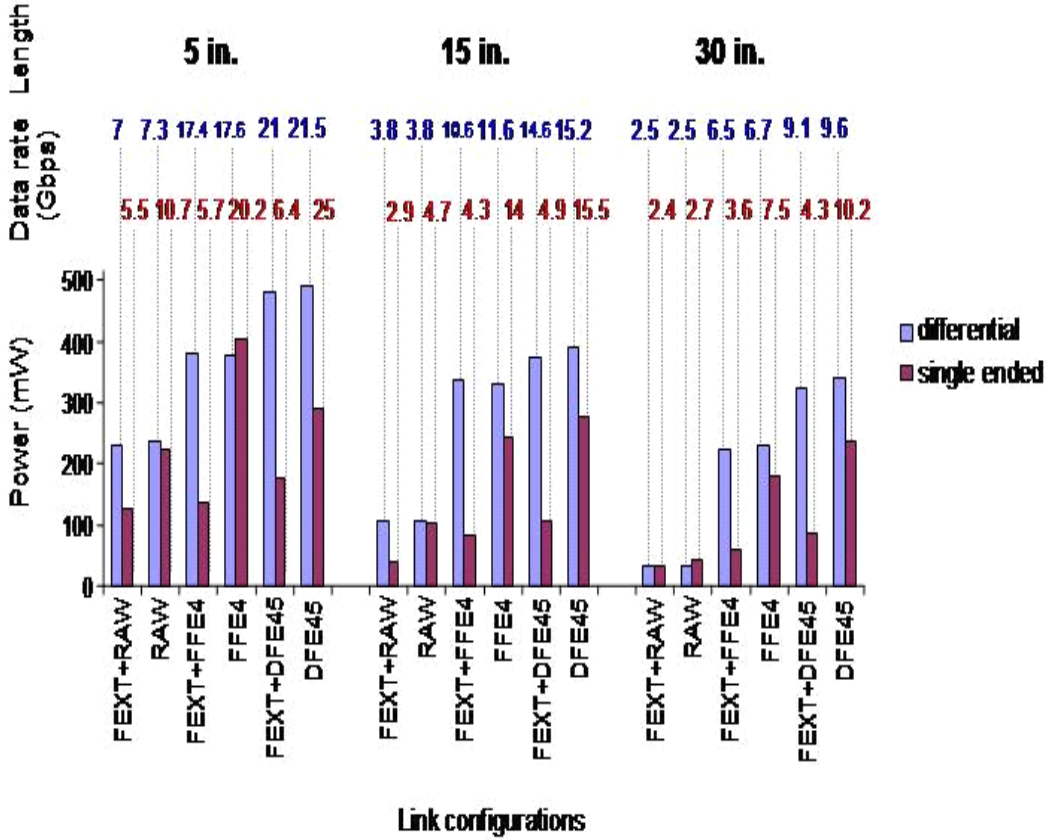
When line loss dominates noise impact
DS no better than SE in data rate per pin

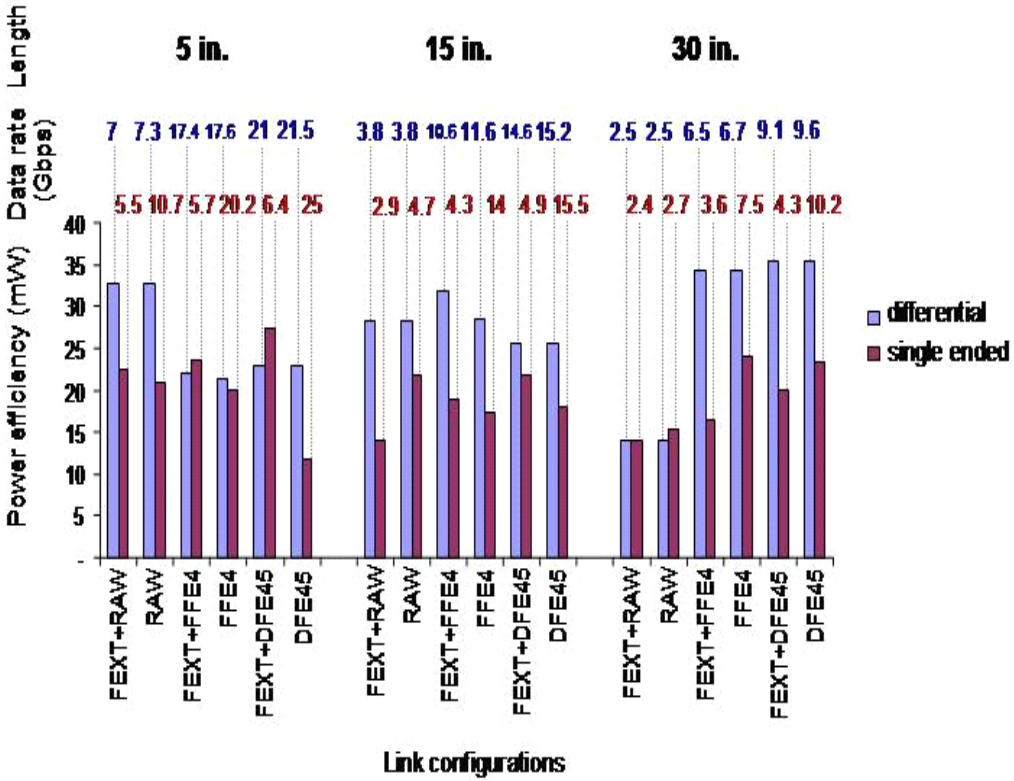
Confirmation of principle

Signal	Xtalk	EQU	5"	Ratio	15"	Ratio	30"	Ratio
single	FEXT	NONE	3.8Gb/s		2.7Gb/s		2.5Gb/s	
diff	FEXT	NONE	6.3Gb/s	1.7x	3.6Gb/s	1.4x	2.4Gb/s	1.0x
single	FEXT	FFE4	3.8Gb/s		3.5Gb/s		2.8Gb/s	
diff	FEXT	FFE4	8.6Gb/s	2.3x	6.4Gb/s	2.0x	4.6Gb/s	1.7x
single	FEXT	DFE45	5.5Gb/s		4.0Gb/s		3.3Gb/s	
diff	FEXT	DFE45	9.8Gb/s	1.8x	7.4Gb/s	1.9x	6.4Gb/s	2.0x
single	OFF	DFE45	11.2Gb/s		6.8Gb/s		5.6Gb/s	
diff	OFF	DFE45	11.1Gb/s	1.0x	8.8Gb/s	1.3x	7.1Gb/s	1.3x

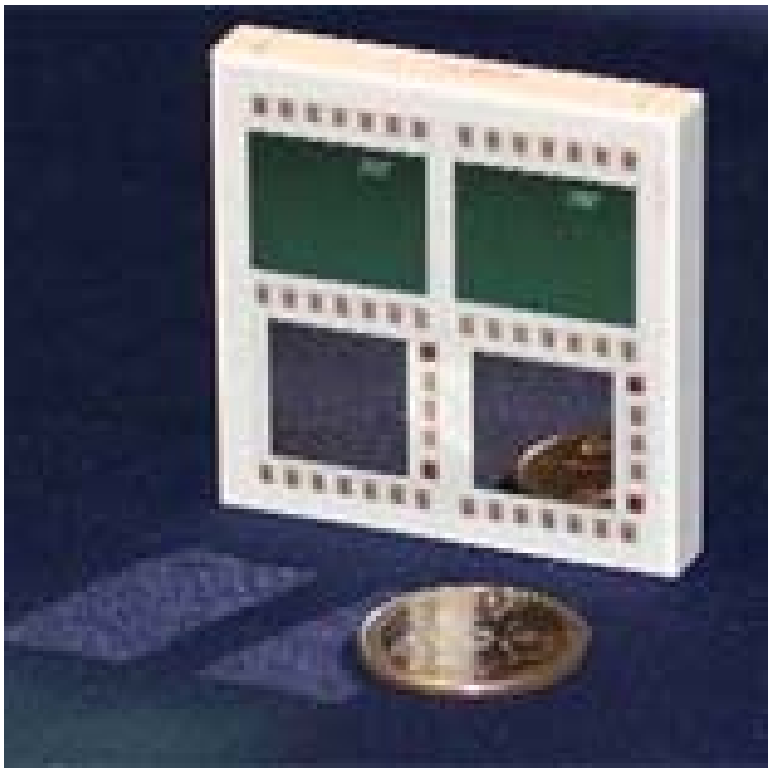
Same comparisons using stubs at the end of nets and increasing the line loss

Power comparisons without stubs





Sockets and rentry of MCM technology



IBM 's Power 5+, 4 core socket
An industry's first

Do not call it MCM though, it is a QCM !!
QCM : quad core module

MCM technology and its promise

- ***Even for entry level blade systems multiple sockets are envisioned***
  - if a socket contains more than 1 chip (PU and L3 memory for example) there is an environment on which the SI engineer can control the noise
  - if the material is such that the noise is small or highly attenuated (organic or Si) then
    - ✓ *SE I/O designs can provide many connections on MCM at high frequencies and little area and power cost*
    - ✓ *For plastic MCM the "ring buses" have minimum escape discontinuity but have long lenghts on Board (e.g. a more or less uncontrolled noise environment) hence differential I/O can be the topology of choice for these I/O designs*
- ***MCM technology reduces board space by 40-60% compared to SCM implementation BUT***
  - has the cost of known good die
  - KGD can be minimized by appropriate use of chip set on an MCM
    - ✓ *only one expensive chip and memory chips possibly on CSP (CISCO)*

Conclusions

- ***The Good News***
  - **Packaging Engineers have a lot of work to do from modeling to TX/RX design**
    - ***Develop better EM modeling tools for higher frequencies***
    - ***Develop better design tools and techniques in frequency domain***
    - ***Devevelop new packaging structures cheaper than SLC***
  - **Multisocket systems will be benefited from MCM/ SOP technology**
    - ***Which can be made cost effective for the right chip complement***
- ***The wish list***
  - **A DS Tx/Rx design that has the perfromance advantage of the serdes but without the area penalty**
    - ***Academia has a big role to play***
  - **A cheap , but chip friendly package material**
    - ***Academia has a big role to play***
- ***The Bad news***
  - **Short time to meet the system's architects needs**
    - ***Requires collaborative development wherever possible during pre competition phase***
    - ***Strengething the consortia focus on interconnect challenges***
    - ***3S workshop (SOC, SOP,SIP) generate the required dialog***