Funded, in part, by DARPA under HR0011-08-090001. The views, opinions, and/or findings contained in this presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

Wavelength Stealing: An Opportunistic Approach to Channel Sharing in Multi-chip Photonic Interconnects

Arslan Zulfiqar
University of Wisconsin – Madison
MICRO-46

COLLABORATORS

- Oracle Labs
  - Pranay Koka
  - Herb Schwetman
  - Xuezhe Zheng
  - Ashok Krishnamoorthy
- University of Wisconsin – Madison
  - Mikko Lipasti

EXECUTIVE SUMMARY

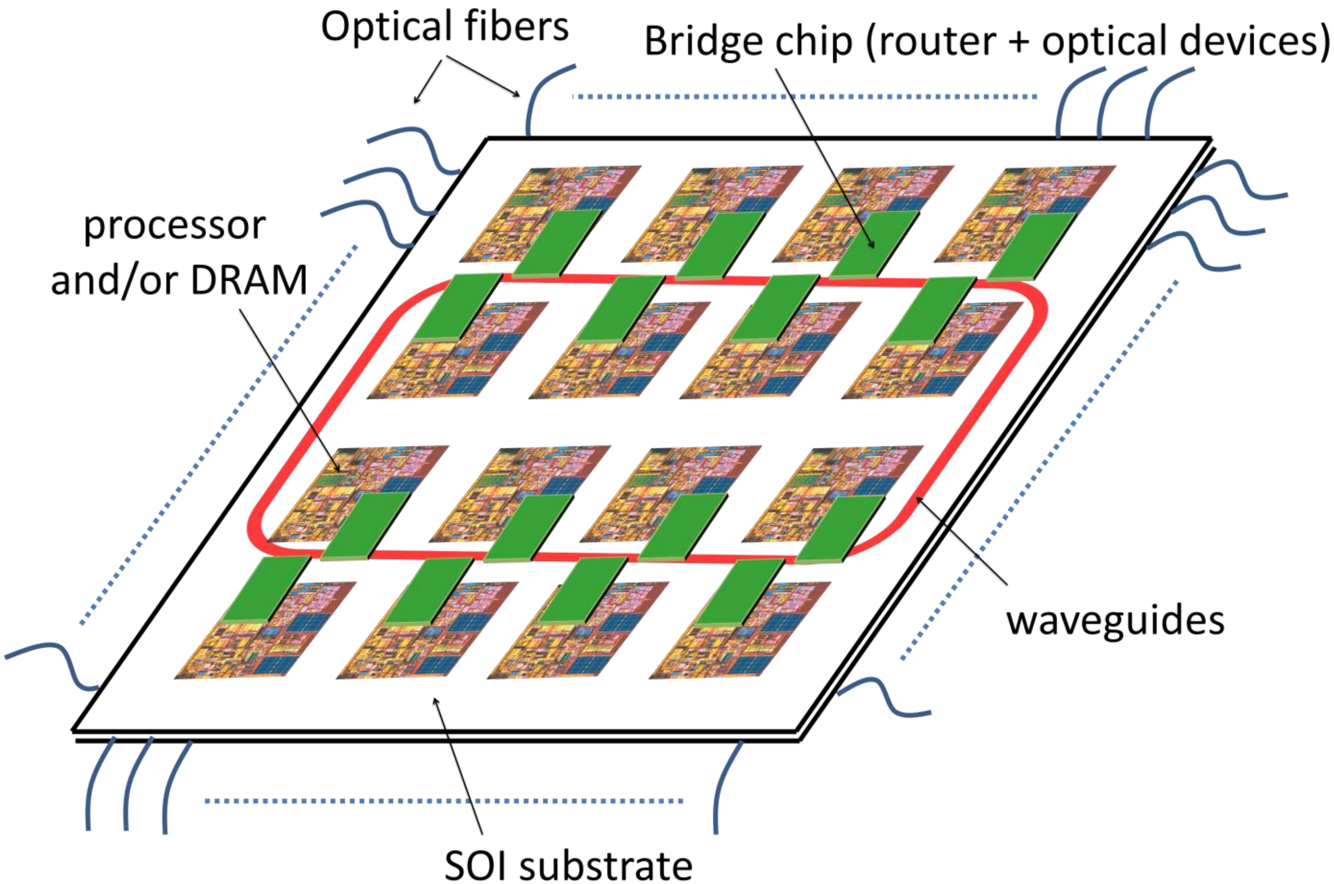
- Silicon-photonics offers integration of multiple chips
  – High sustainable performance
  – Improved process yields
  – Lower energy-per-bit consumption
- This work:
  – Focus on **channel sharing** nanophotonic designs
  – Model to determine limits and gains of channel sharing
  – Novel shared channel architecture: **Wavelength Stealing**
    - Arbitration-free
    - Strong fairness guarantees
    - Up to 28% better EDP than baseline on HPC workloads

MOTIVATION

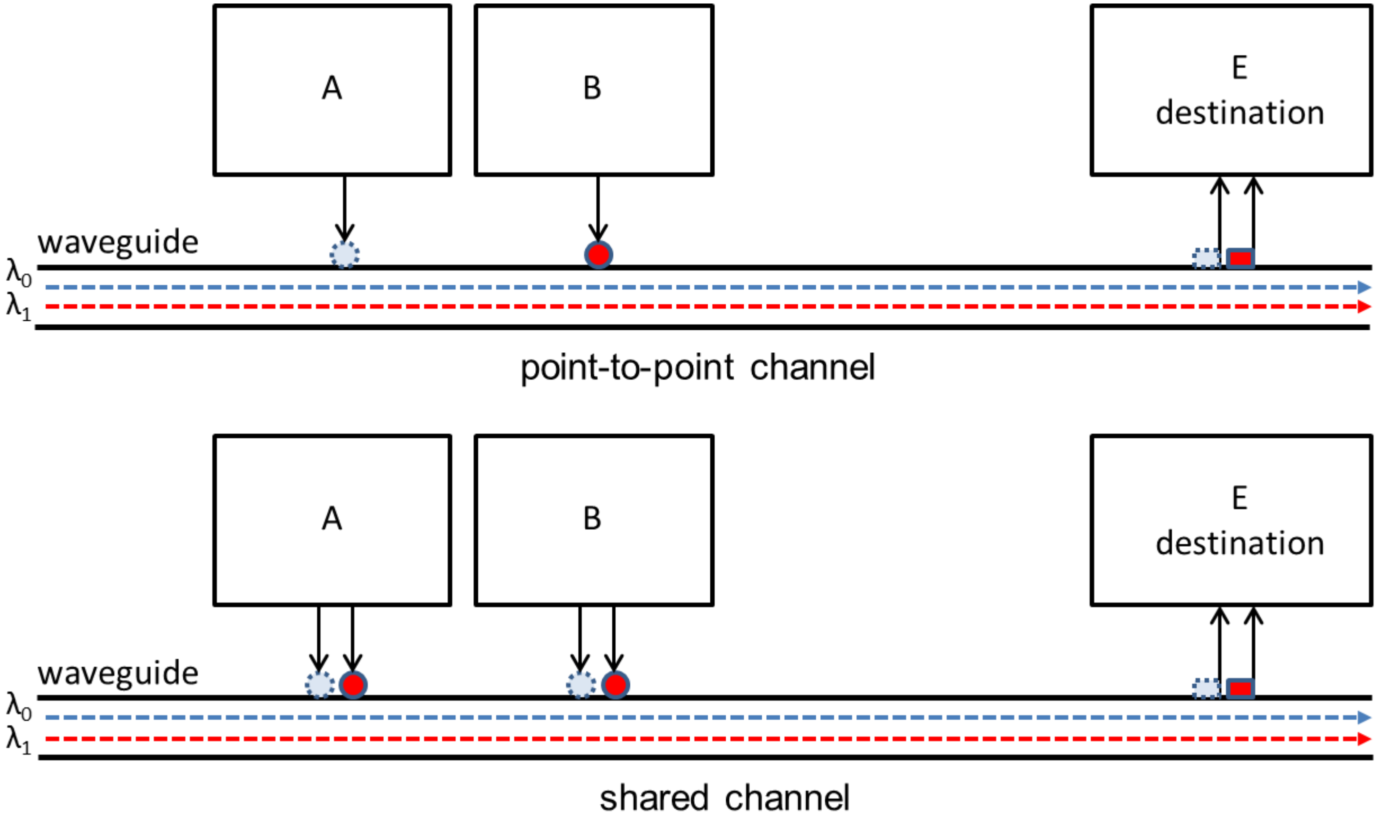
- Technology trend: More cores
- Scaling single chip systems
  - Increasing fabrication costs
  - Low process yields
  - Power delivery limitations
- Multichip systems – require enormous off-chip communication
  - Low bandwidth densities for off-chip I/O
  - High power consumption of serial links

ORACLE'S "MACROCHIP" VISION

- Aggregate multiple chips with photonic communication



WAVELENGTH SHARING LOSSES



Extra ring-resonators on shared wavelengths increase link-loss leading to higher laser power consumption

"Approved for Public Release. Distribution Unlimited"

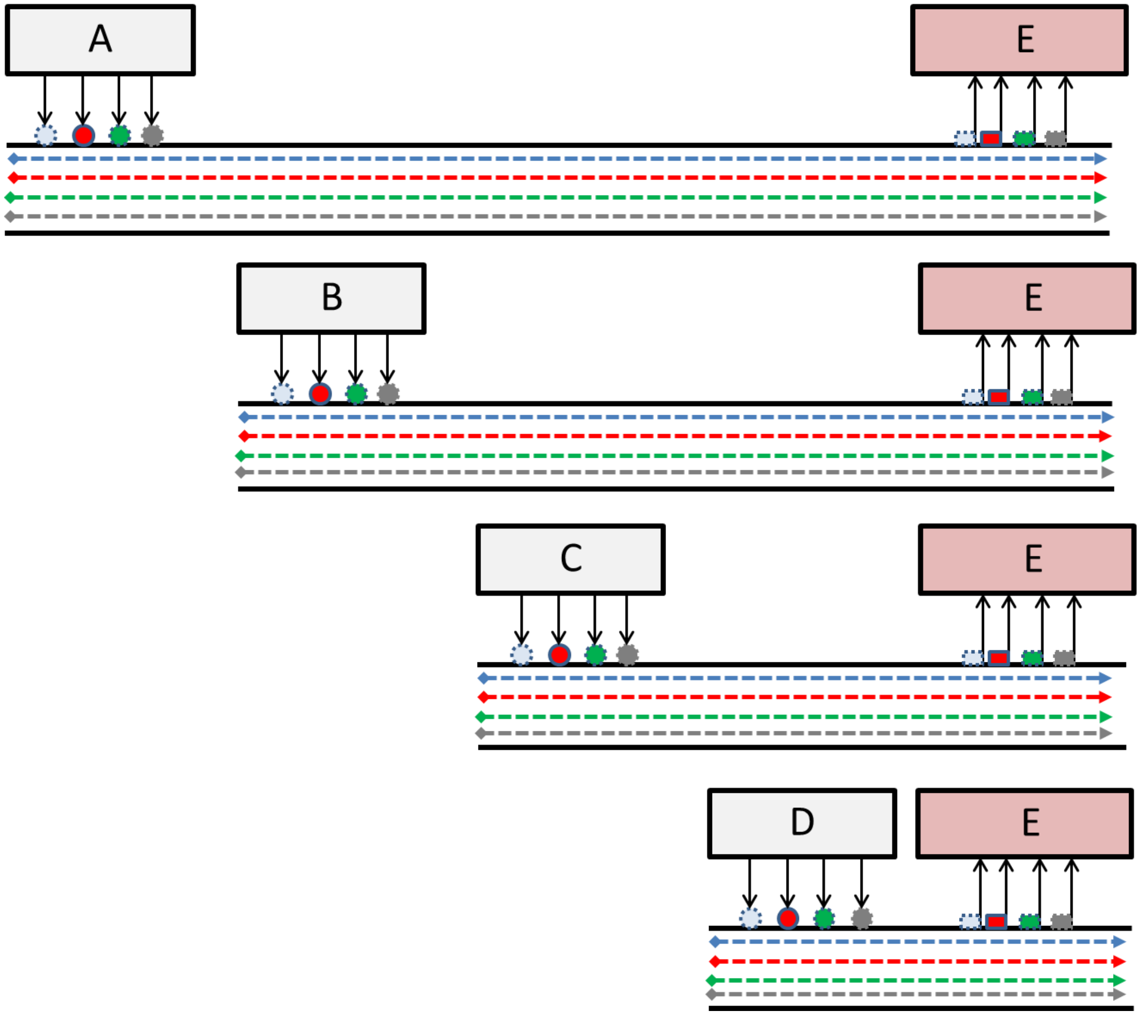
TECHNOLOGY IMPLICATIONS

- Photonic networks are static power dominated
  – Laser power
  – Ring tuning power
- Efficiencies of commercial WDM lasers: $\mathbf{1-5}\%$
  – Laser power consumption biggest contributor to static power
- Optimizing for laser power first-order design constraint

Fixed input laser-power budget for all designs

IMPLICATIONS OF LASER POWER BUDGET (I)

P2P (unshared)



All-to-All traffic

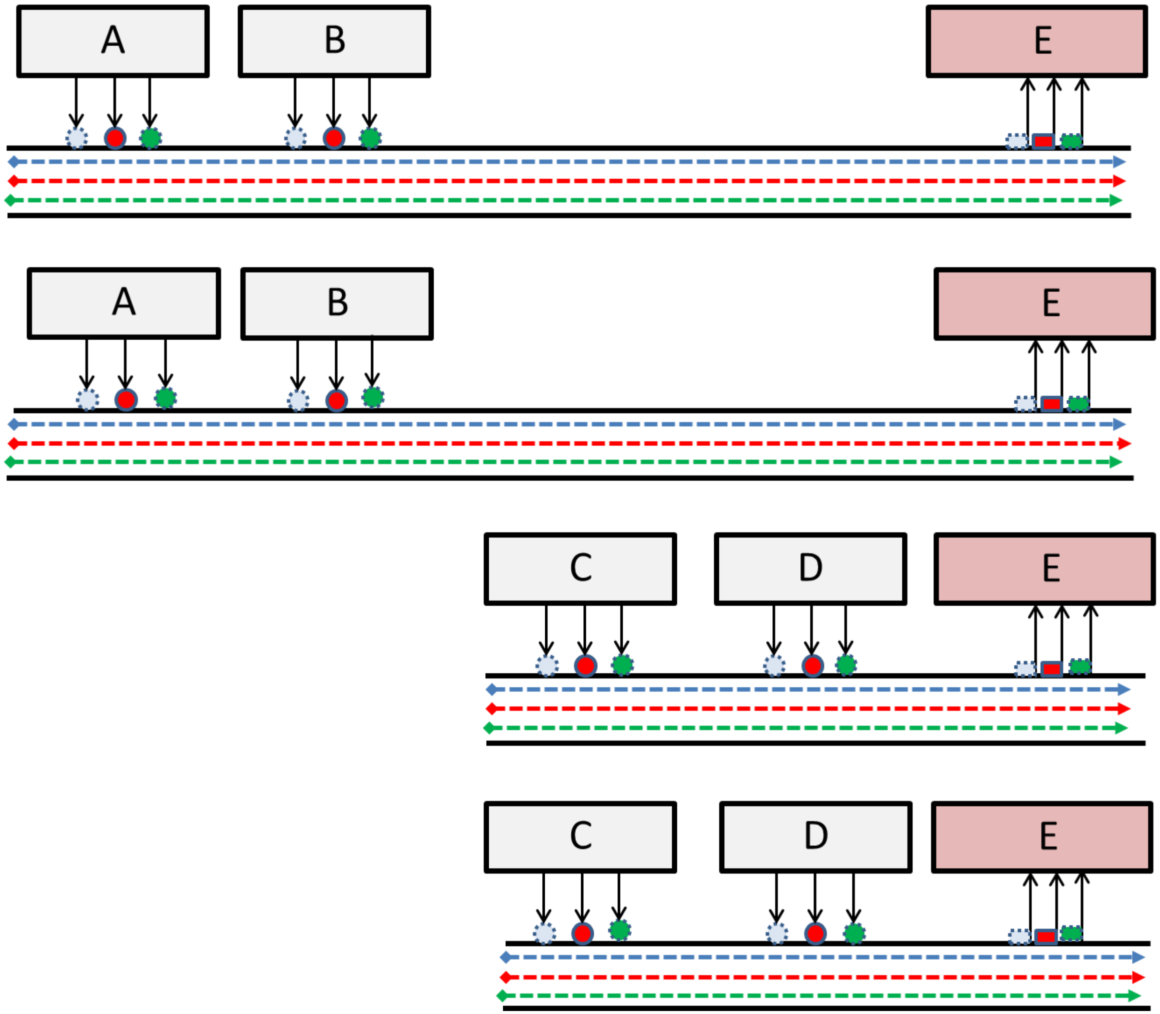
- 4 x 4b/cycle
  = 16b/cycle

Permutation traffic

- 4b/cycle

IMPLICATIONS OF LASER POWER BUDGET (II)

2-way sharing



All-to-All traffic

- 4 x 3b/cycle
  = 12b/cycle

Permutation traffic

- 6b/cycle

vs

All-to-All traffic

- 4 x 4b/cycle
  = 16b/cycle

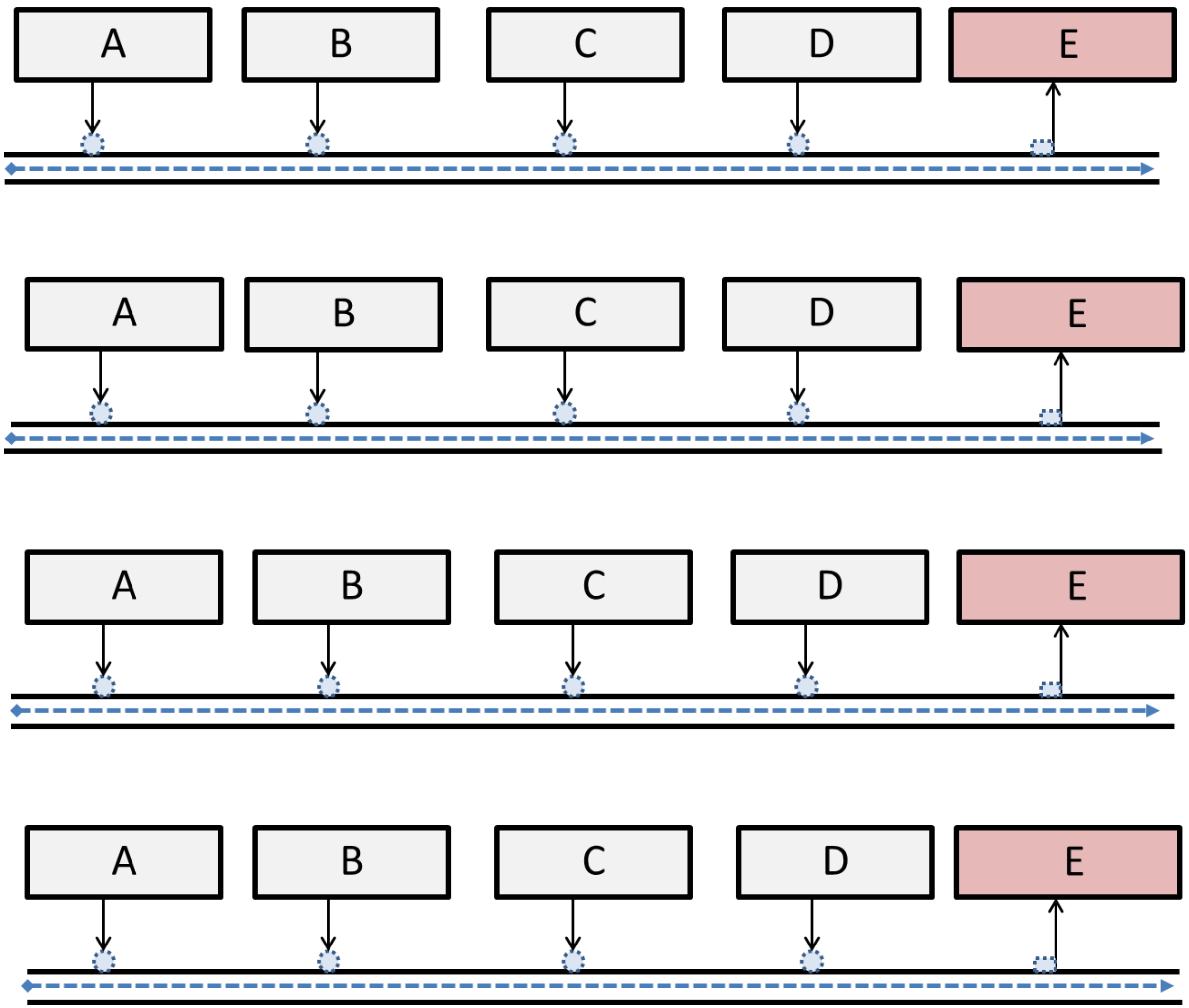
Permutation traffic

- 4b/cycle

ORACLE

IMPLICATIONS OF LASER POWER BUDGET (III)

4-way sharing



All-to-All traffic

- 4 x 1b/cycle
  = 4b/cycle

Permutation traffic

- 4b/cycle

vs

All-to-All traffic

- 4 x 3b/cycle
  = 12b/cycle

Permutation traffic

- 6b/cycle

IMPLICATIONS OF LASER POWER BUDGET (IV)

- Increasing sharing degree: 's'
  - **Reduces effective capacity**
    - Lower performance on all-to-all (uniform random) traffic than P2P
  - **Increases peak node-node BW followed by drop-off**
    - Potentially better performance on permutation traffic than P2P

- **Can we estimate ideal sharing degree 's_{ideal}' and ideal node-node BW gain over P2P? Yes**

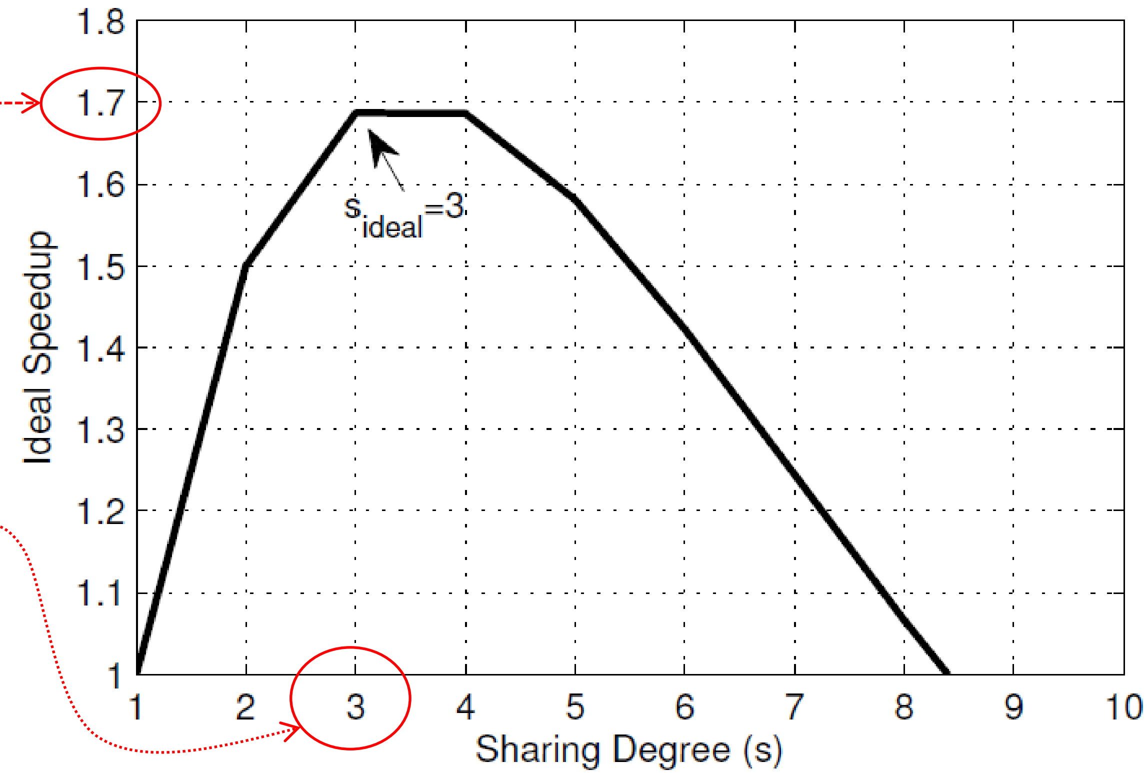
IDEAL SHARING GAINS

Optimal sharing gain

~ 1.70x

Optimal sharing degree

s_{ideal} = 3



- Ignore wavelength/ time overheads of sharing
- Conservative device assumptions

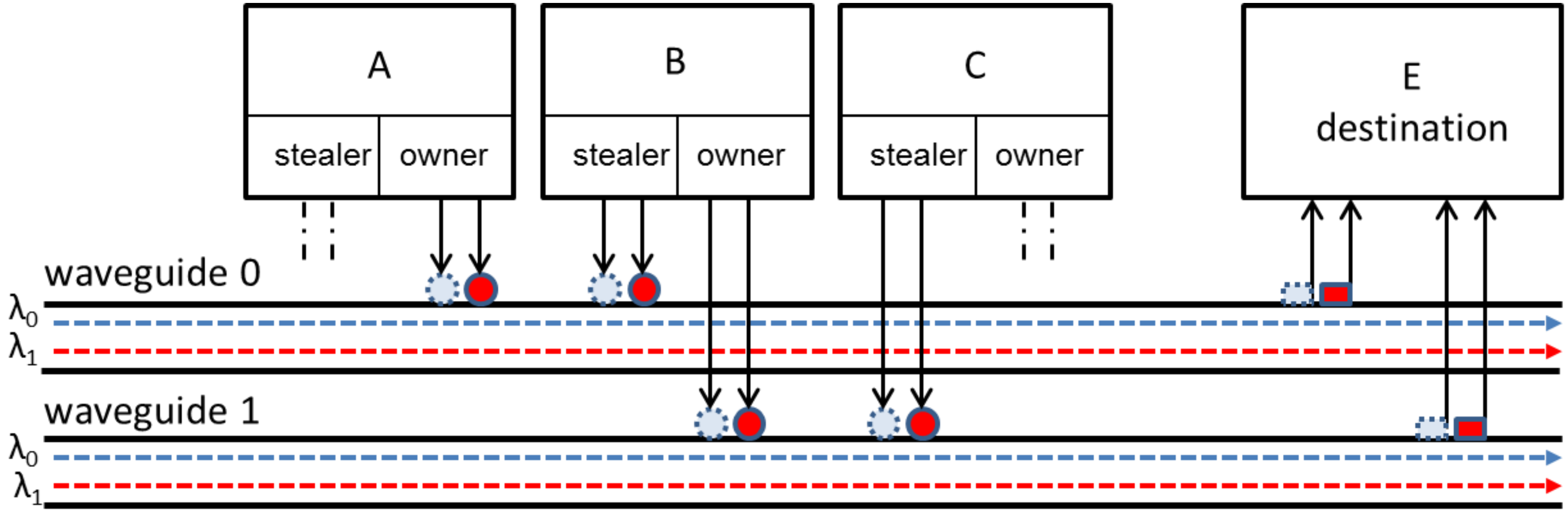
IDEAL SHARING GAINS SUMMARY

P2P	Channel Sharing
(+) High capacity	(+) High N-N BW only when $2 \leq s \leq 3$
➢High-radix traffic	➢Speedup $\leq 1.70\times$
	➢Low-radix traffic
(-) Low N-N BW	(-) Low capacity
➢Low-radix traffic	➢High-radix traffic

WAVELENGTH STEALING

- Same topology as the P2P network: N^2 channels
- Every channel has **one owner** and **one or more stealers**

2-way stealing



IMPLEMENTATION REQUIREMENTS

- **Owner node**
  - Guaranteed non-blocking access
- **Stealer node**
  - Arbitration-free access on an owner's channel: possible packet corruption
  - Notification to halt stealing when channel busy
- **Destination node**
  - Valid phit: identify sender (owner or stealer?)
  - Corrupted phit: perform correction

IMPLEMENTATION MECHANISM

ERASURE CODING

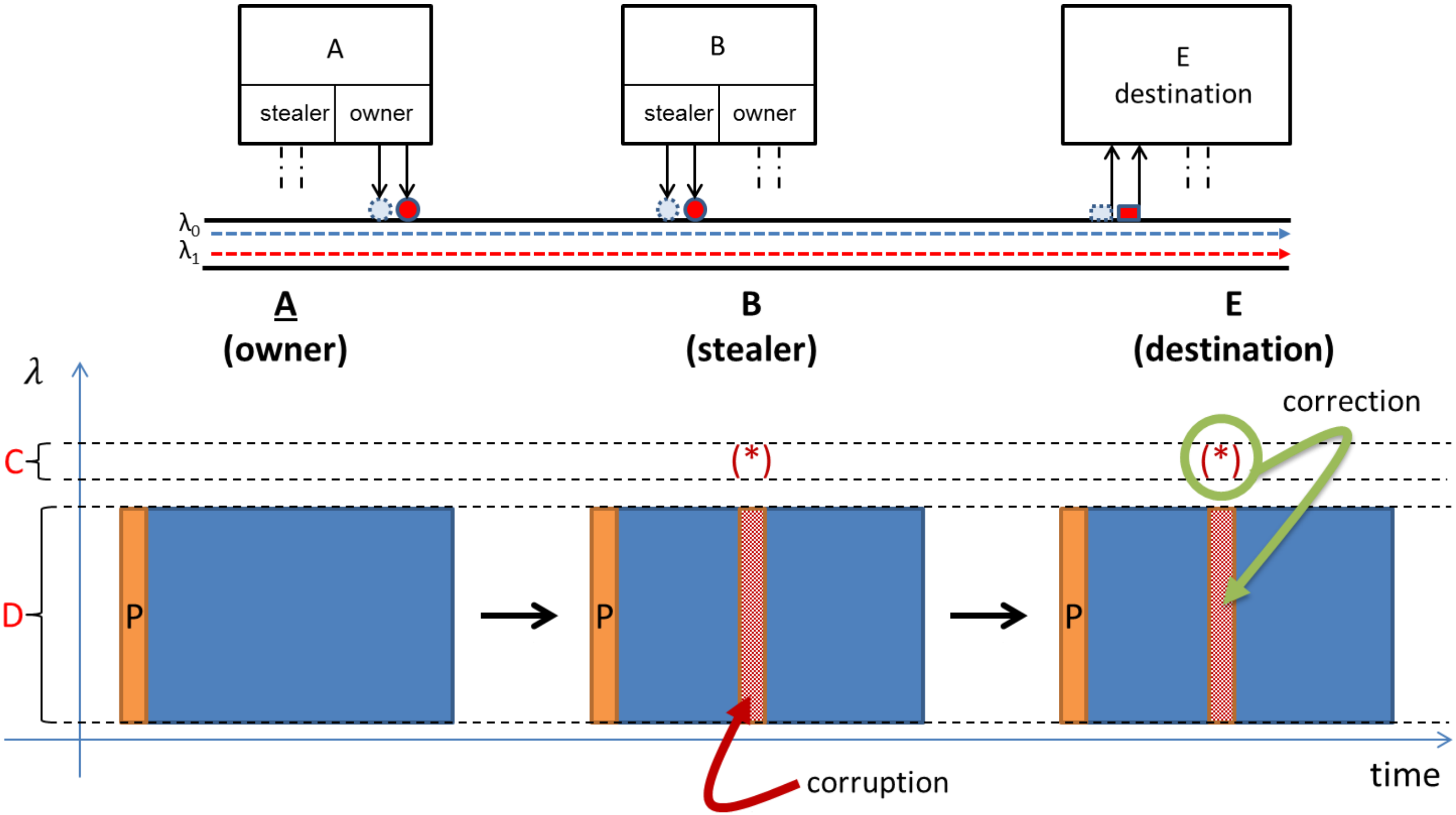
+

CONTROL WAVELENGTHS PER CHANNEL

ORACLE

ERASURE CODING

- Erasure coding is used at the destination to correct corruptions due to a collision

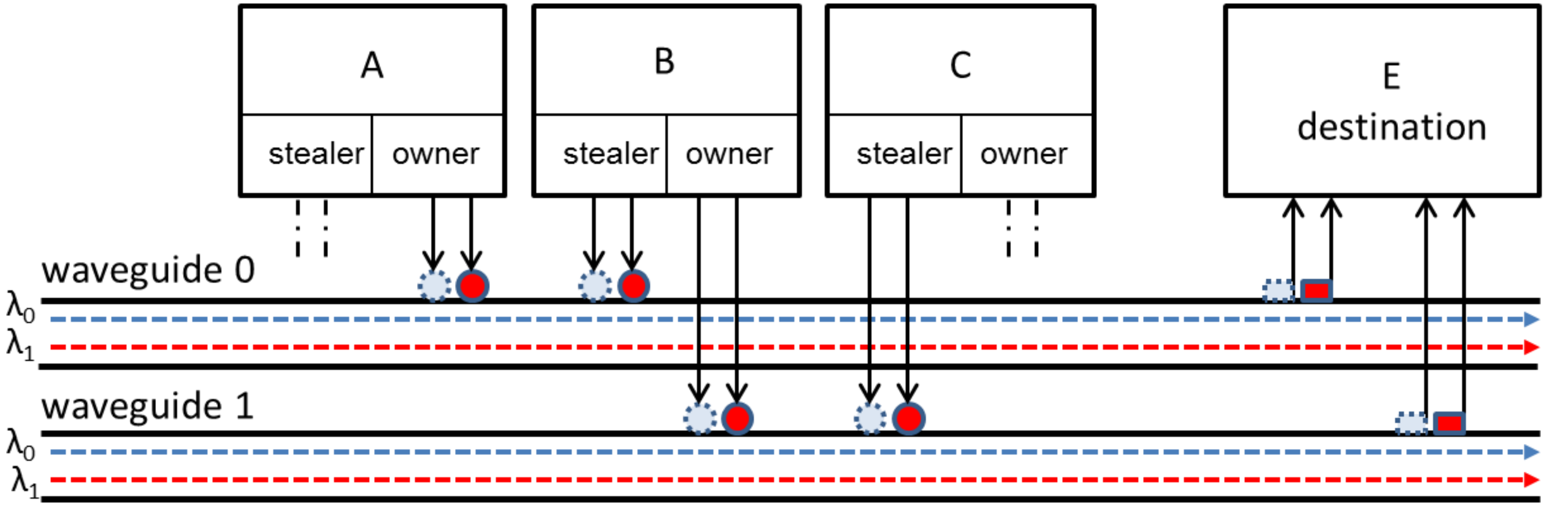


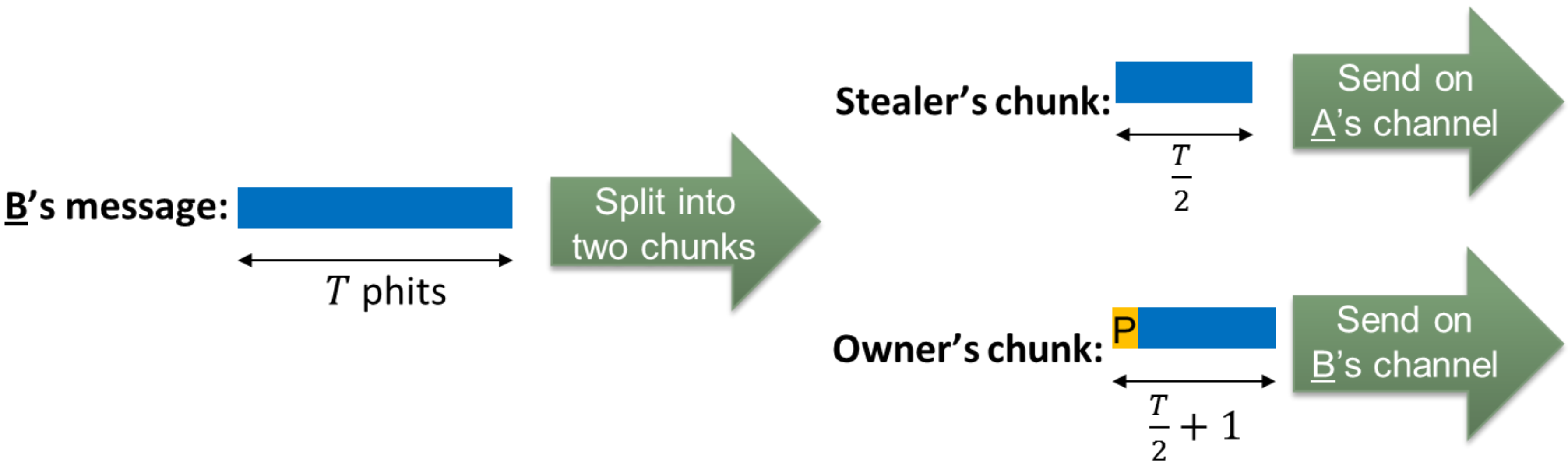
CONTROL WAVELENGTHS

- **Functionality**
  – Mark location of corruption for erasure correction
  – Inform stealer to halt stealing when owner becomes active
  – Inform destination of the ID (owner, stealer, corrupted) of the received communication (phit)
- **Two designs – different trade-offs**
  – **Abort**
  – **Sense**

ORACLE

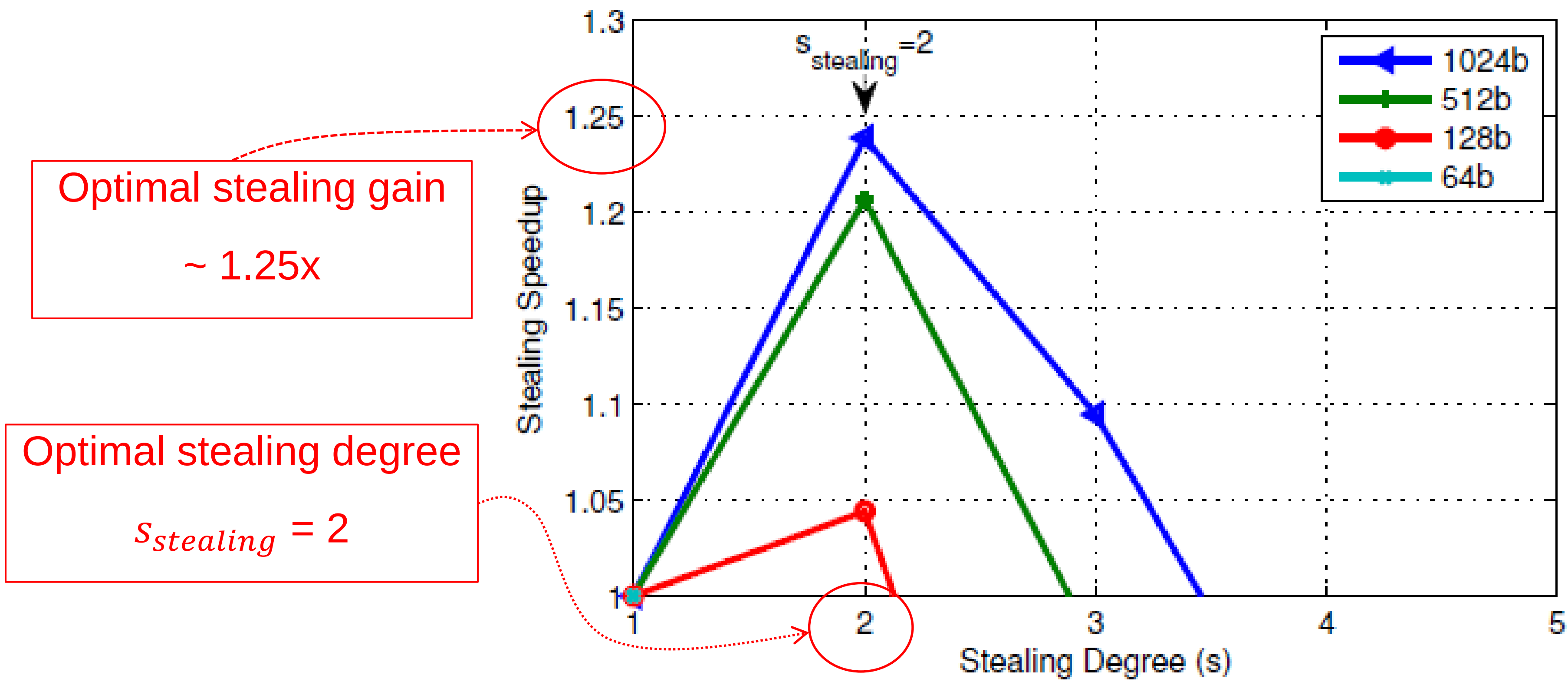
PROTOCOL OPERATION





"Approved for Public Release. Distribution Unlimited"

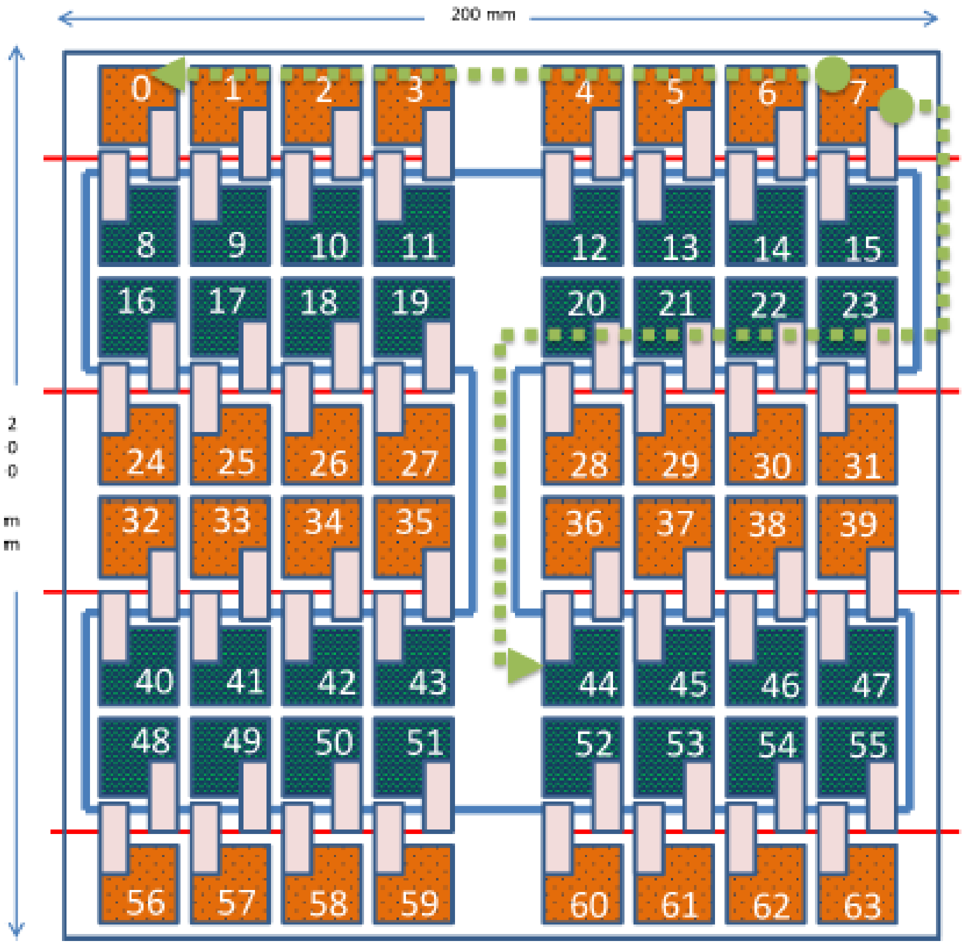
WAVELENGTH STEALING GAINS



- Loss in performance due to
  - Control (wavelength) overheads $\propto \frac{1}{data\ wavelengths}$
  - Coding (time) overheads $\propto \frac{1}{message\ size}$

EVALUATION – SETUP

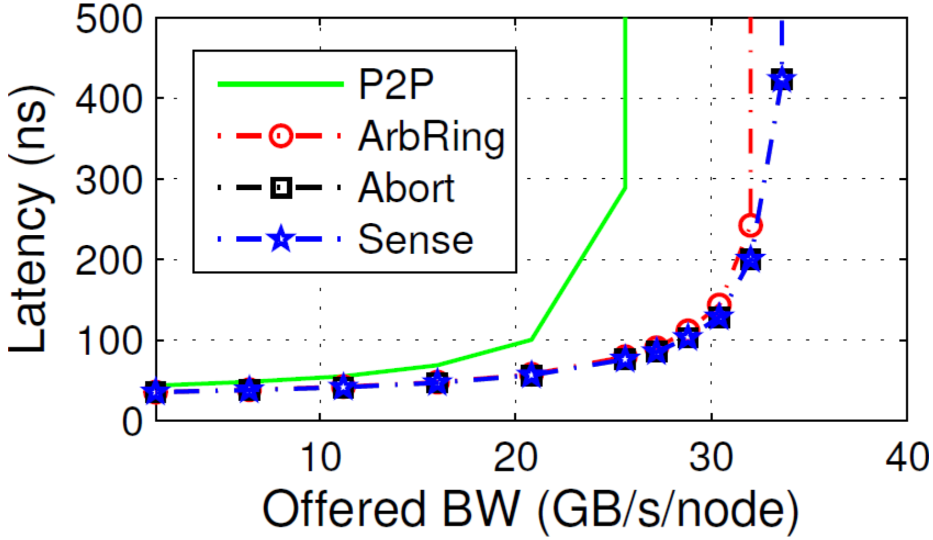
8 x 8 Macrochip System



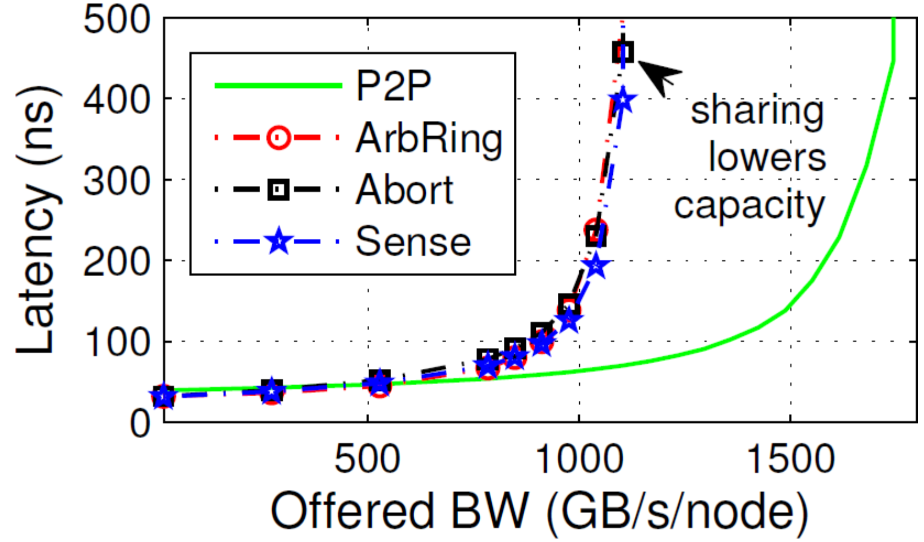
- Synthetic workloads
  - Uniform random
  - Permutation
  - Asymmetric
- Application workloads: NAS
  - BT: Block tri-diagonal solver
  - CG: Conjugate gradient kernel
  - DT WH: “White Hole” graph analysis
  - DT BH: “Black Hole” graph analysis
  - DT SH: “Shuffle” graph analysis

EVALUATION – SYNTHETIC WORKLOADS (I)

Bit-Complement (No Contention)



Unif. Random (Unif. Contention)

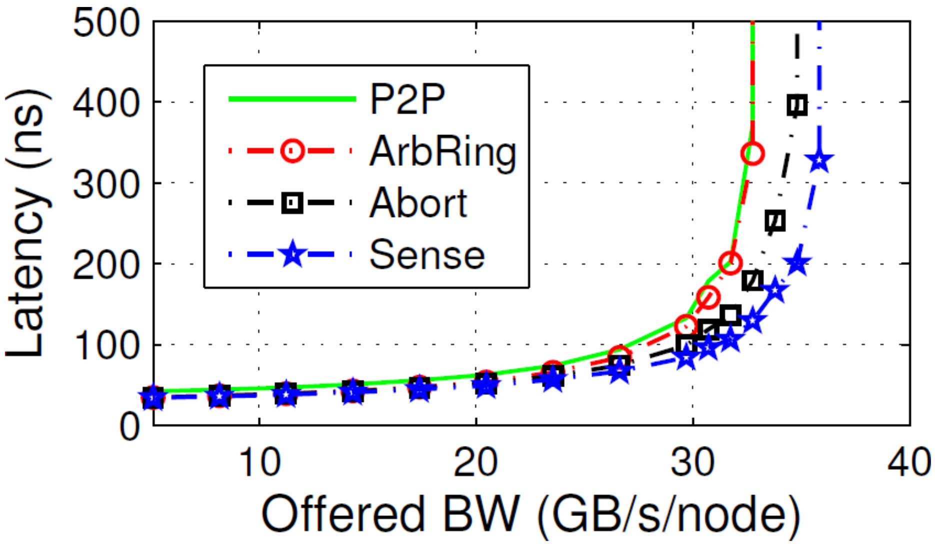


- Sharing designs provide higher (lower) throughput than P2P in the absence (presence) of contention
- Sharing designs exhibit lower capacity

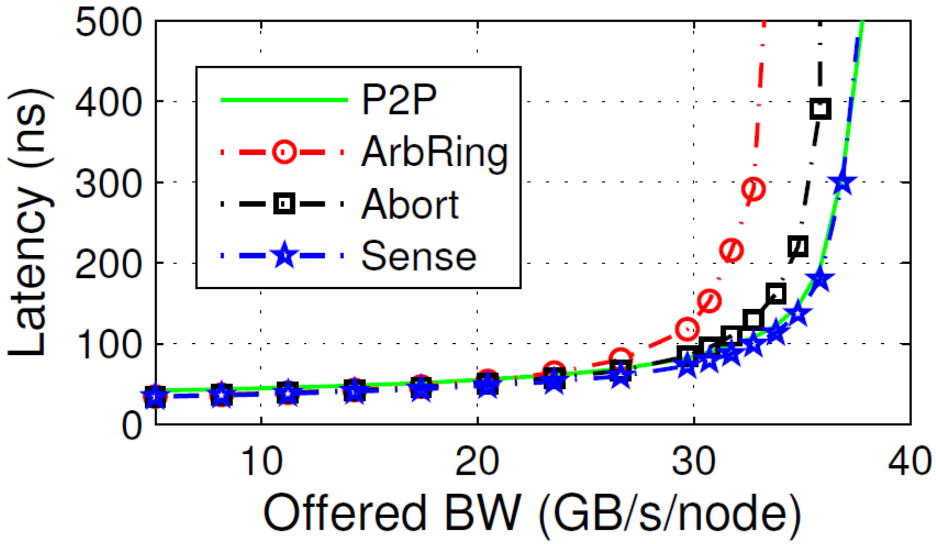
EVALUATION – SYNTHETIC WORKLOADS (II)

Asymmetric K (Variable Contention)

K = 20

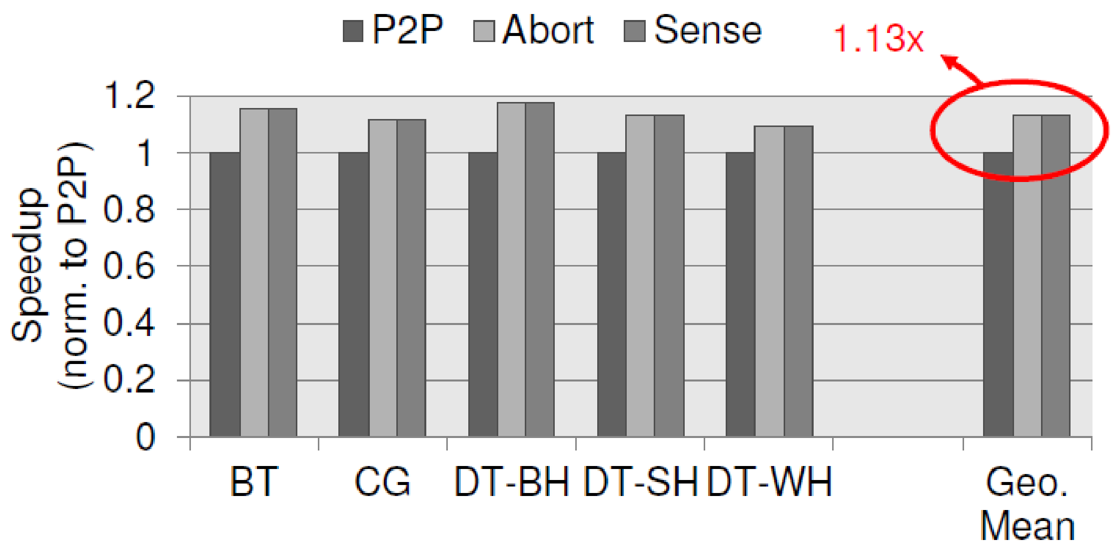


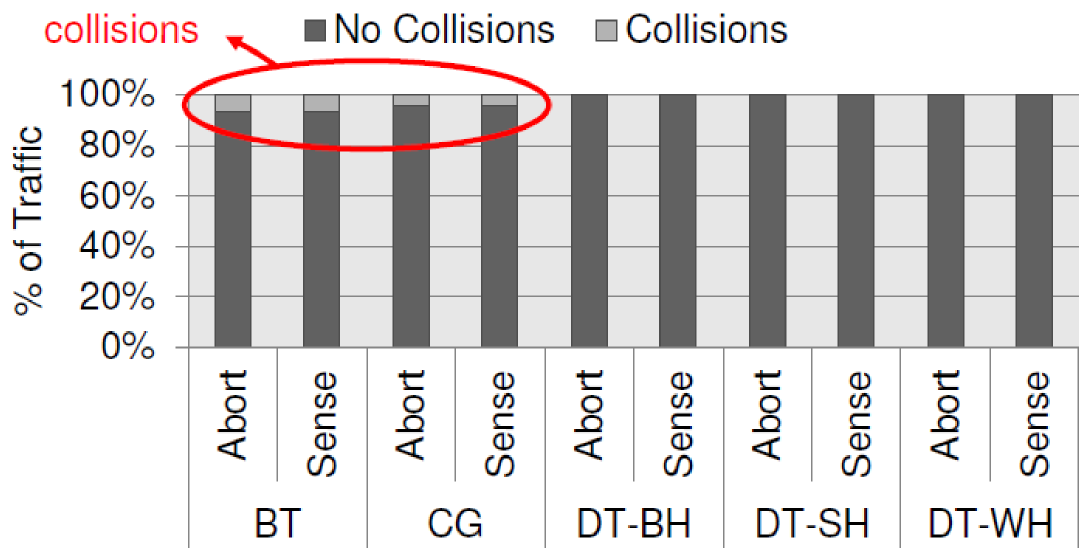
K = 30



- As contention is increased
  - Wavelength stealing provides better throughput than Token-ring design
  - Sense design outperforms abort design
  - Throughput performance of P2P improves

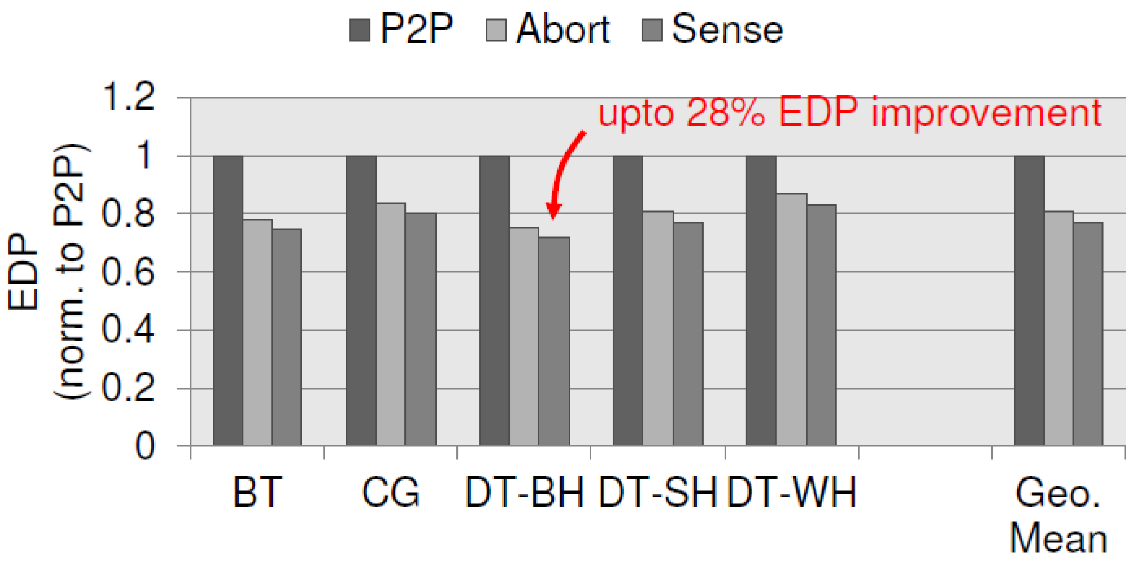
EVALUATION – APPLICATION WORKLOADS (I)





- All workloads exhibit speedups
  - Max: **1.17x**
  - Average: **1.13x**
- Differences in speedups due to
  - Traffic patterns
  - Message sizes
  - Message frequencies

EVALUATION – APPLICATION WORKLOADS (II)



- Wavelength stealing architectures achieve up to 28% lower EDP than the P2P network
- Average EDP improvement: 20% for Abort, 23% for Sense
  – Sense uses fewer ring-resonators

CONCLUSION

- Channel sharing improves peak node-node BW compared to P2P but at the cost of reduced capacity

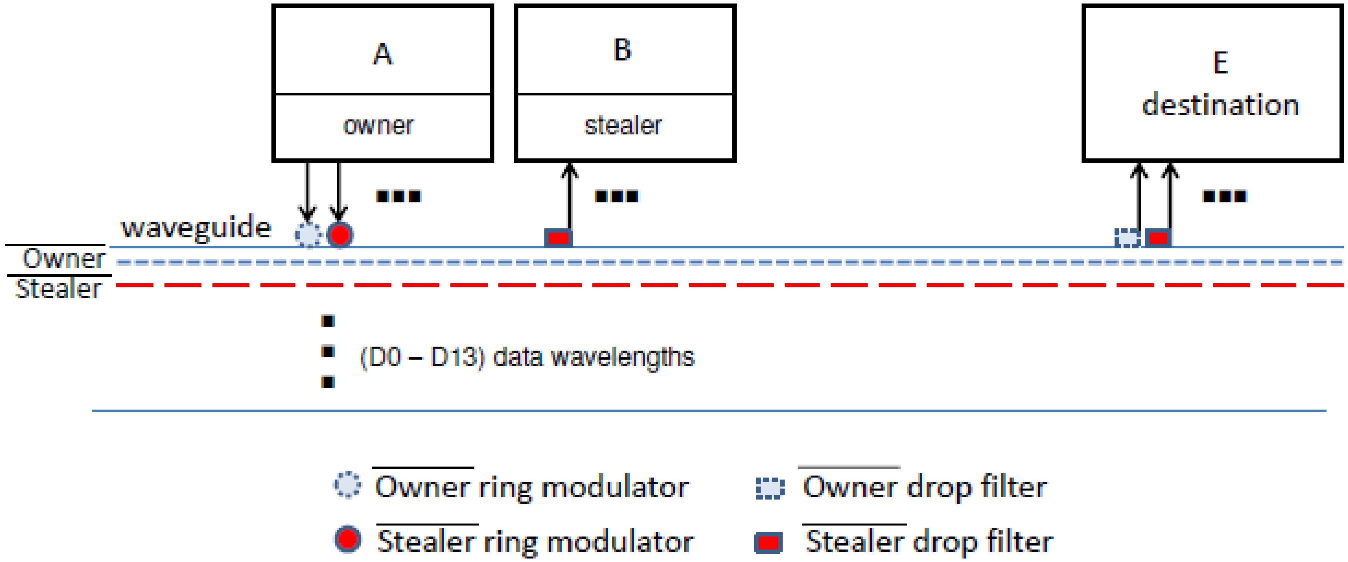
- Developed an analytical model to quantify limits and gains of channel sharing
  – sharing degree ≤ 3
  – sharing gain ≤ 1.70x

- Wavelength Stealing architecture
  – Arbitration-free accesses
  – Strong fairness guarantees
  – **Guaranteed gains on VMs**

ORACLE

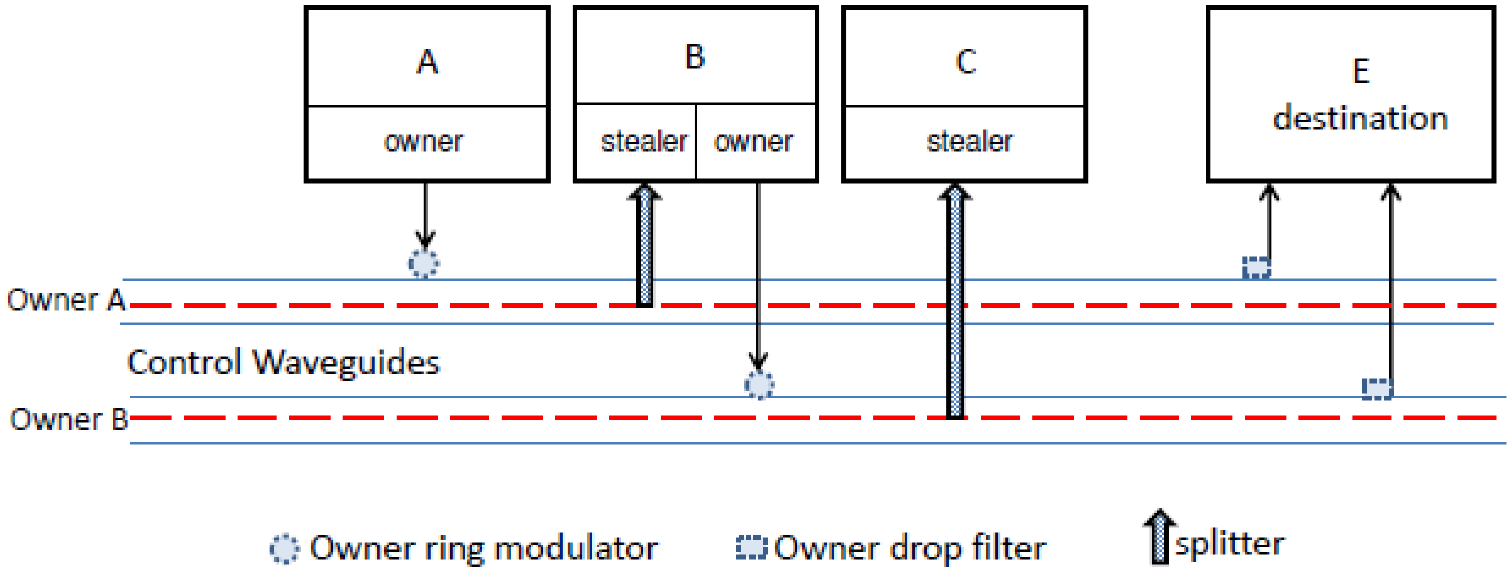
Backup Slides

ABORT DESIGN



Active Sender	A		B	E		Received
	$\overline{Own.}$	$\overline{St.}$	$\overline{St.}$	$\overline{Own.}$	$\overline{St.}$	
A	0	1	–	0	1	A
B	1	0	0	1	0	B
A, B	0	1	1	0	0	$Collision$
(Invalid)	1	1	–	1	1	(Invalid)

SENSE DESIGN

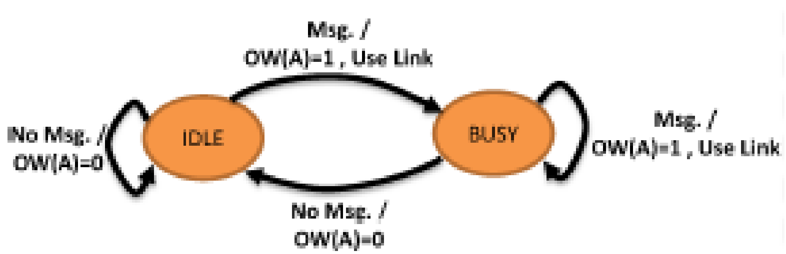


- Employs broadband splitters (non-destructive reads)
- Implemented using state machines at
  - Owner
  - Stealer
  - Destination

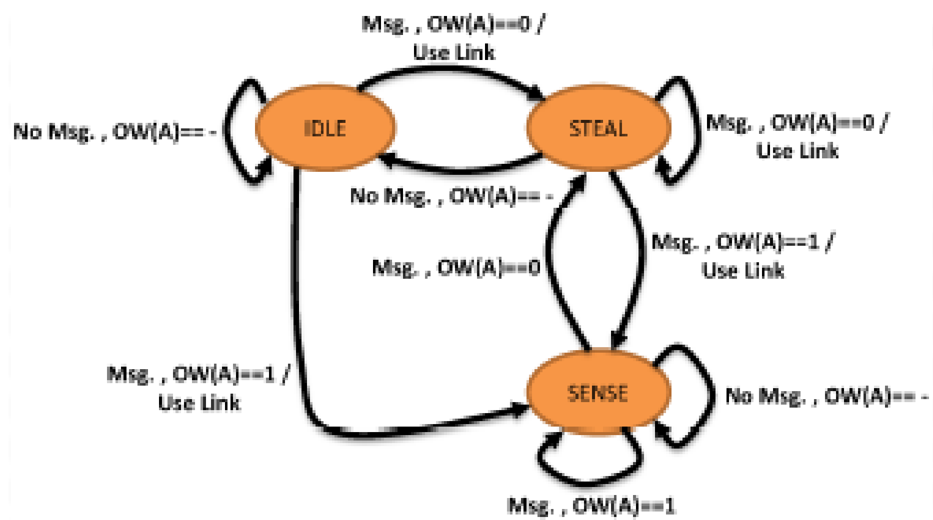
SENSE DESIGN FUNCTIONALITY (I)

Terminology

Condition(s)/ Action(s)	If the Condition(s) are satisfied, make a state transition and perform the Action(s)	OW(A)=1	The Owner A puts a 1 on its control wavelength	OW(A)==1	If the value of the control wavelength on Owner A's link is 1	OW(A)== -	The value of the control wavelength on Owner A's link is irrelevant for the state transition

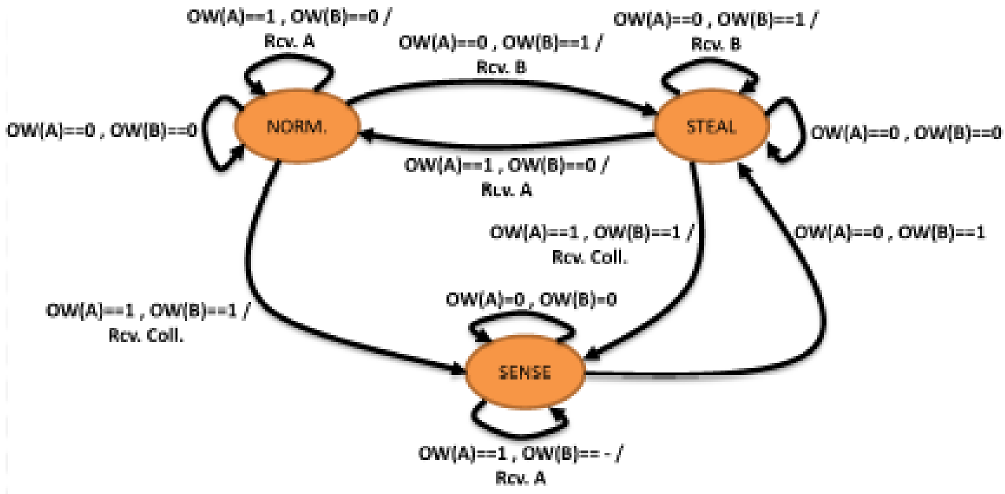


Owner (A) State Machine



Stealer (B) State Machine

SENSE DESIGN FUNCTIONALITY (II)



Destination (E) State Machine

ABORT VS. SENSE TRADE-OFFS

- Abort
  - (+) Fewer waveguides
  - (–) Conservative performance
  - (–) More ring-resonators
- Sense
  - (+) Aggressive performance
  - (+) Fewer ring-resonators
  - (–) More waveguides

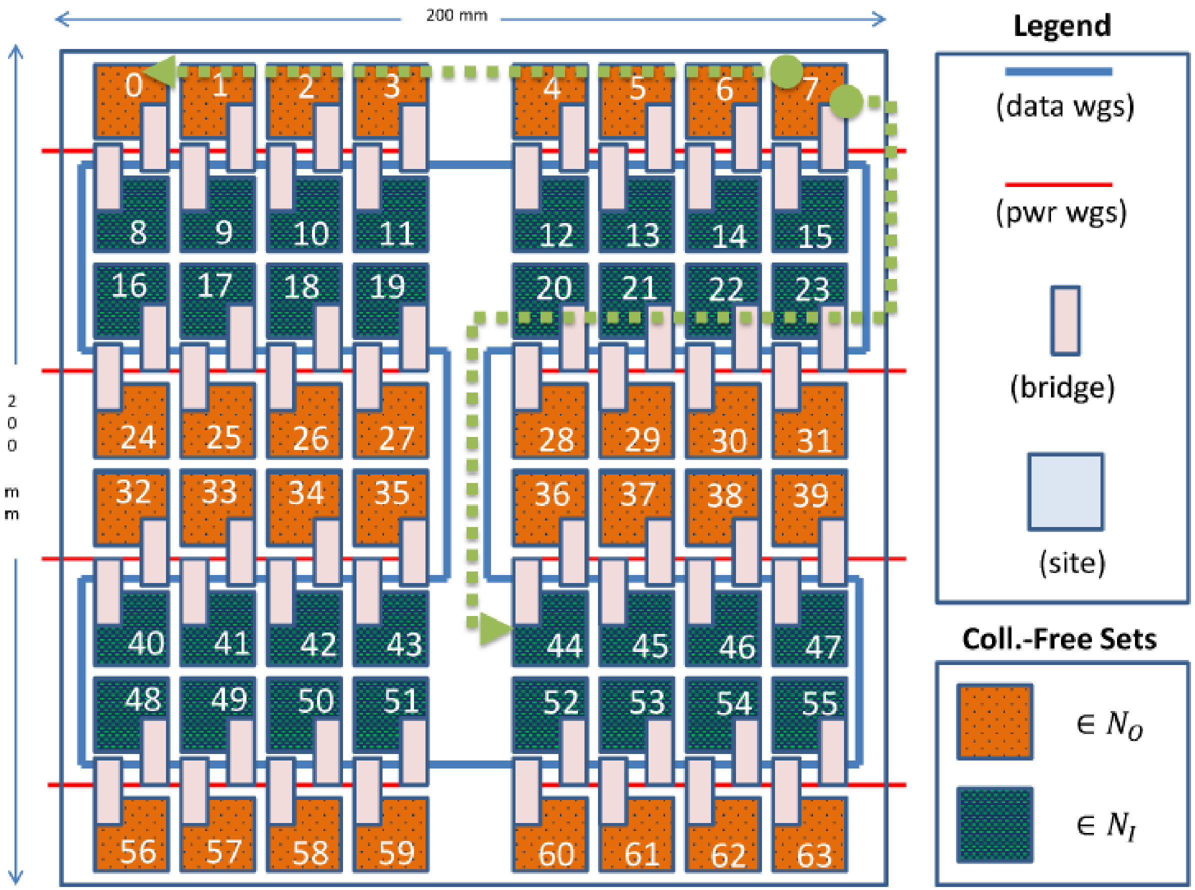
OPTICAL DEVICE PARAMETERS

Parameter	Assumption
Mod. (Insertion) Ring Loss	$4dB$
Inactive Mod. Ring Loss	$0.5dB$
Active Drop-Filter Ring Loss	$1dB$
Passive Ring Loss	$0.05dB$
Waveguide Loss	$0.05dB/cm$
Bridge Chip Waveguide Loss	$1dB$
Coupler Loss	$2dB$
Receiver Sensitivity Margin	$4dB$
Receiver Sensitivity Level	$-21dBm$
Ring Tuning Power	$0.3mW/ring$
Mod. Driver	$35fJ/bit$
Detector Driver	$65fJ/bit$
Max. Fiber WDM-Factor	32
Max. Waveguide WDM-Factor	16
Max. Port Fibers	2500
Power per Fiber	$32mW$

ORACLE

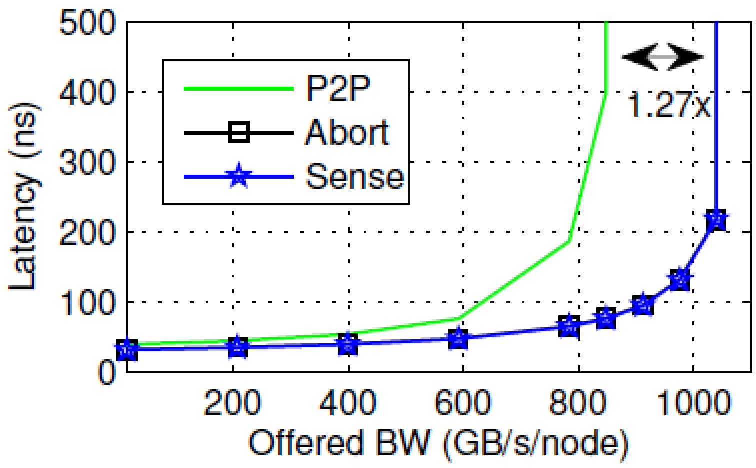
VIRTUALIZATION GAINS (I)

- Virtualization: many VMs share the system
  - Better utilization of system resources



VIRTUALIZATION GAINS (II)

Domain Uniform Random



Four 16-Node VMs