Brandenburg
University of Technology
Cottbus - Senftenberg

25th August 2015

DIAMOND RINGS

ACKNOWLEDGED EVENT PROPAGATION IN MANY-CORE PROCESSORS

Stefan Nürnberger, Randolf Rotta,
Gabor Drescher, Daniel Danner, Jörg Nolte

ACKNOWLEDGED EVENT PROPAGATION

Brandenburg
University of Technology
Cottbus - Senftenberg

What does it do?

- Make events observable in a networked system
- Make sure events are globally observable
- Enforce ordering of events

What is it good for?

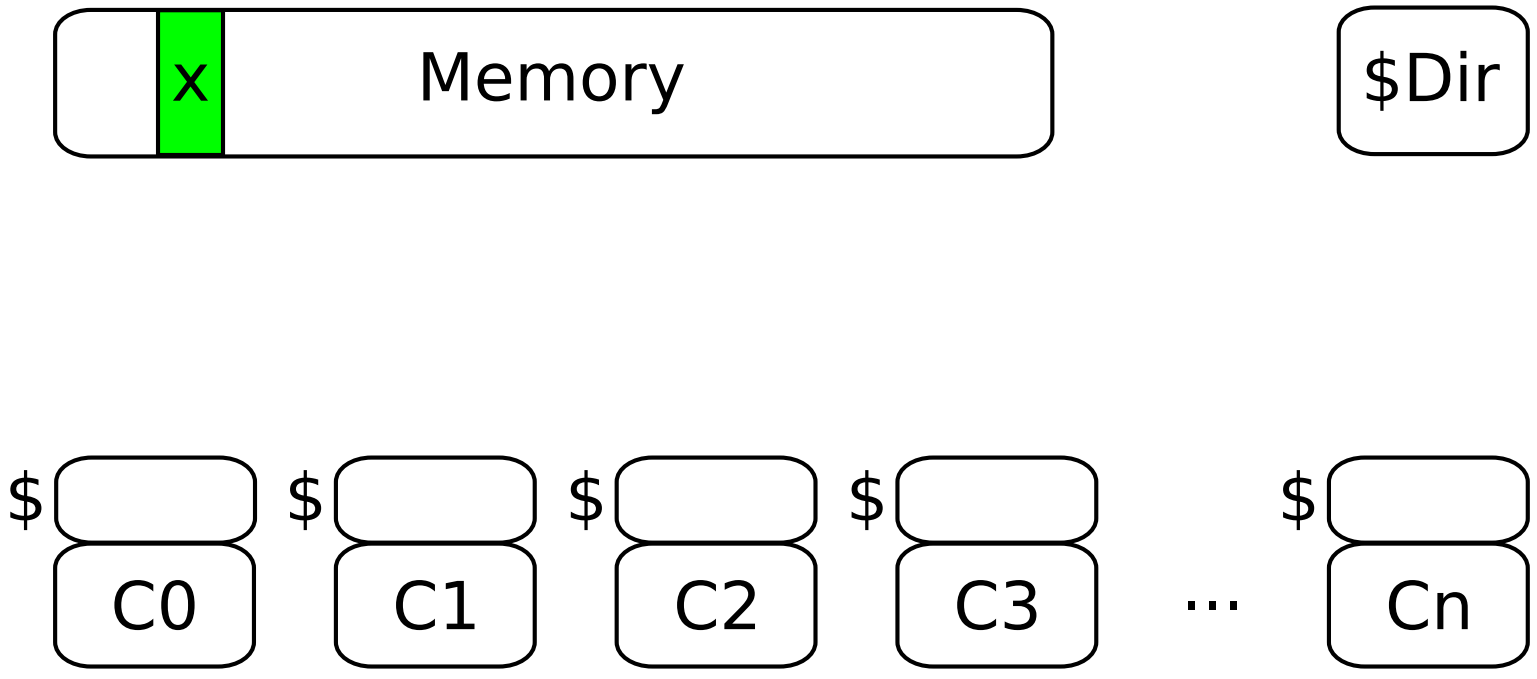
- Memory Consistency
- Coherence Protocols
- Atomic Operations

How to implement it?

“Just use broadcast with acknowledgement...”

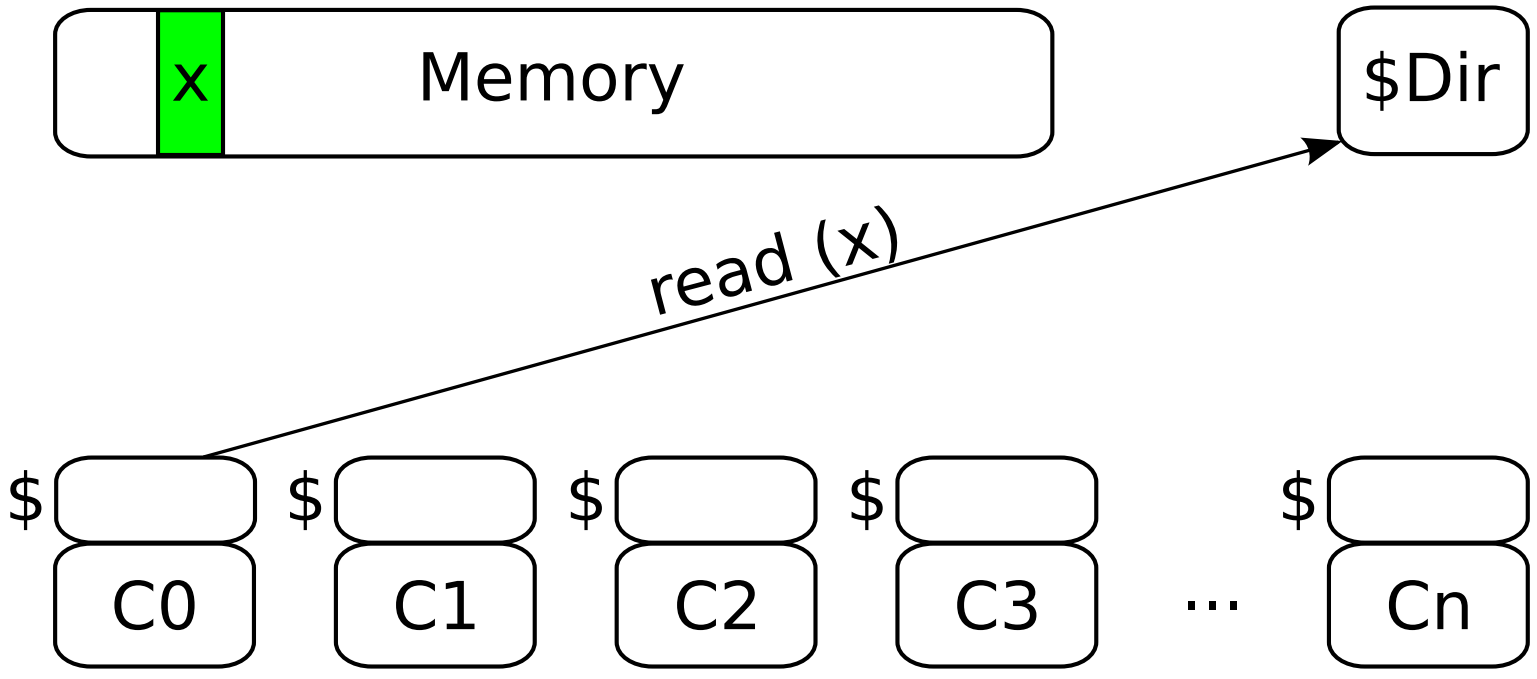
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



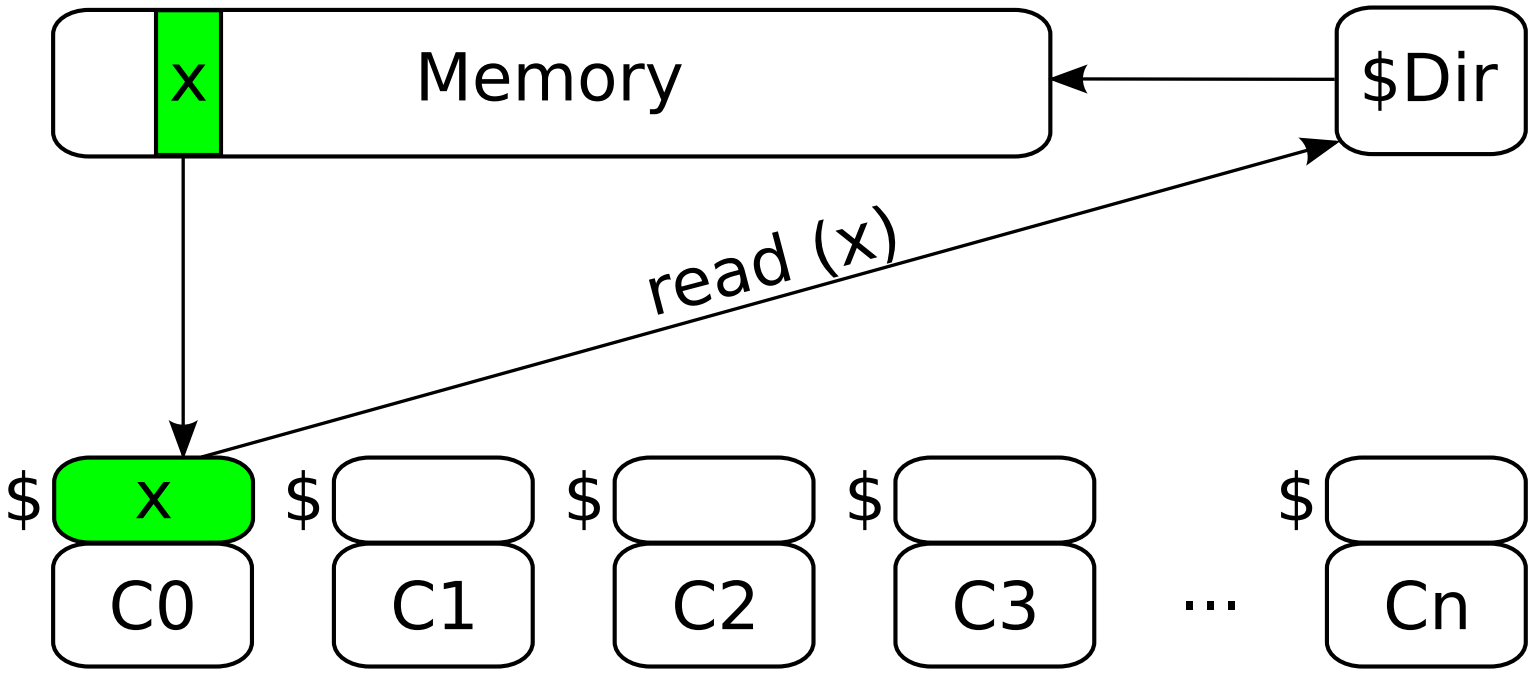
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



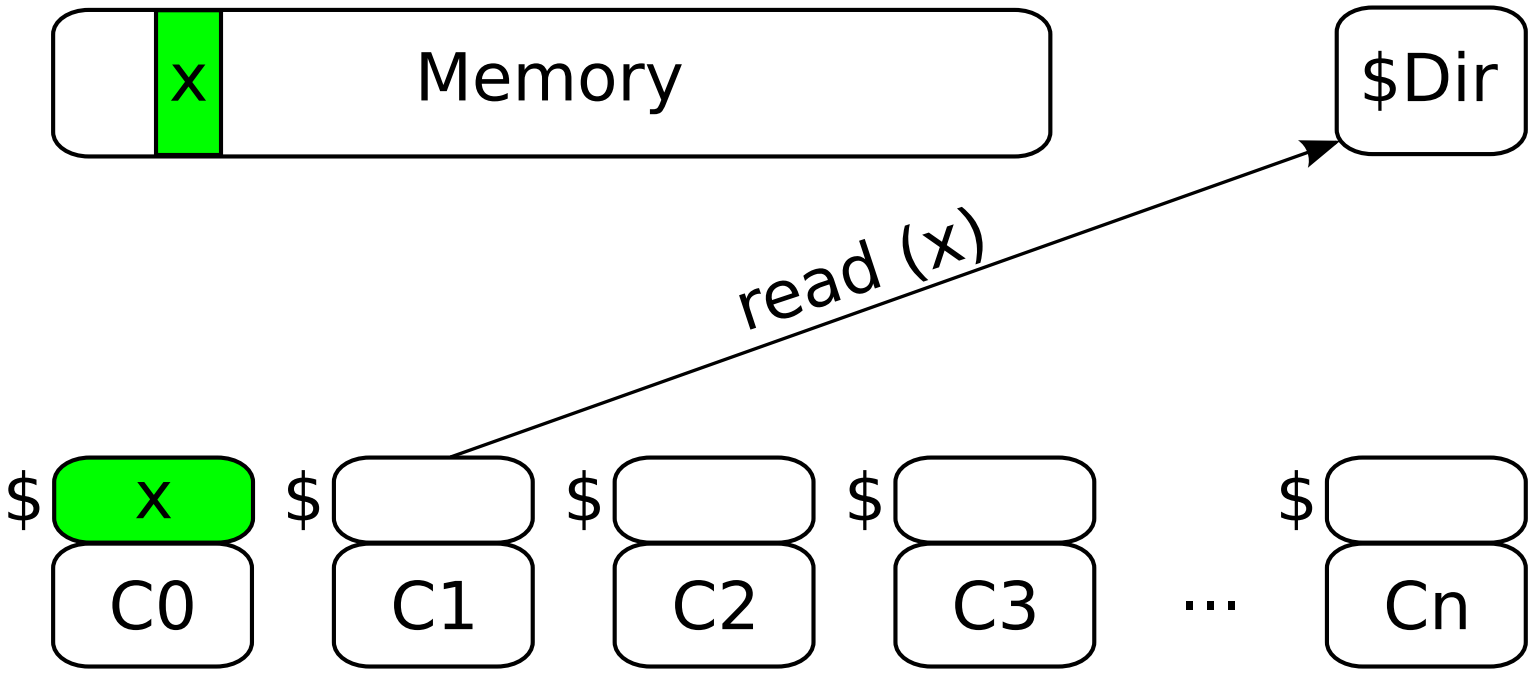
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



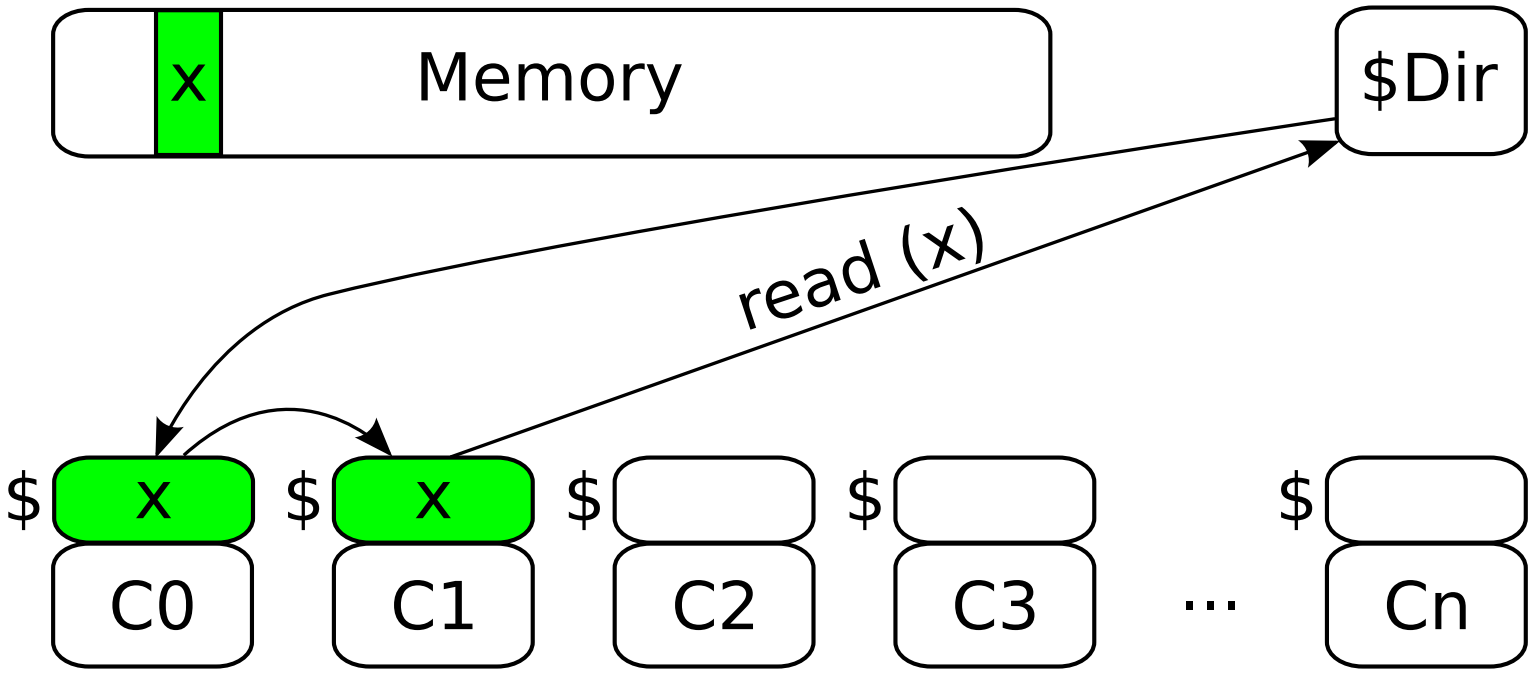
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



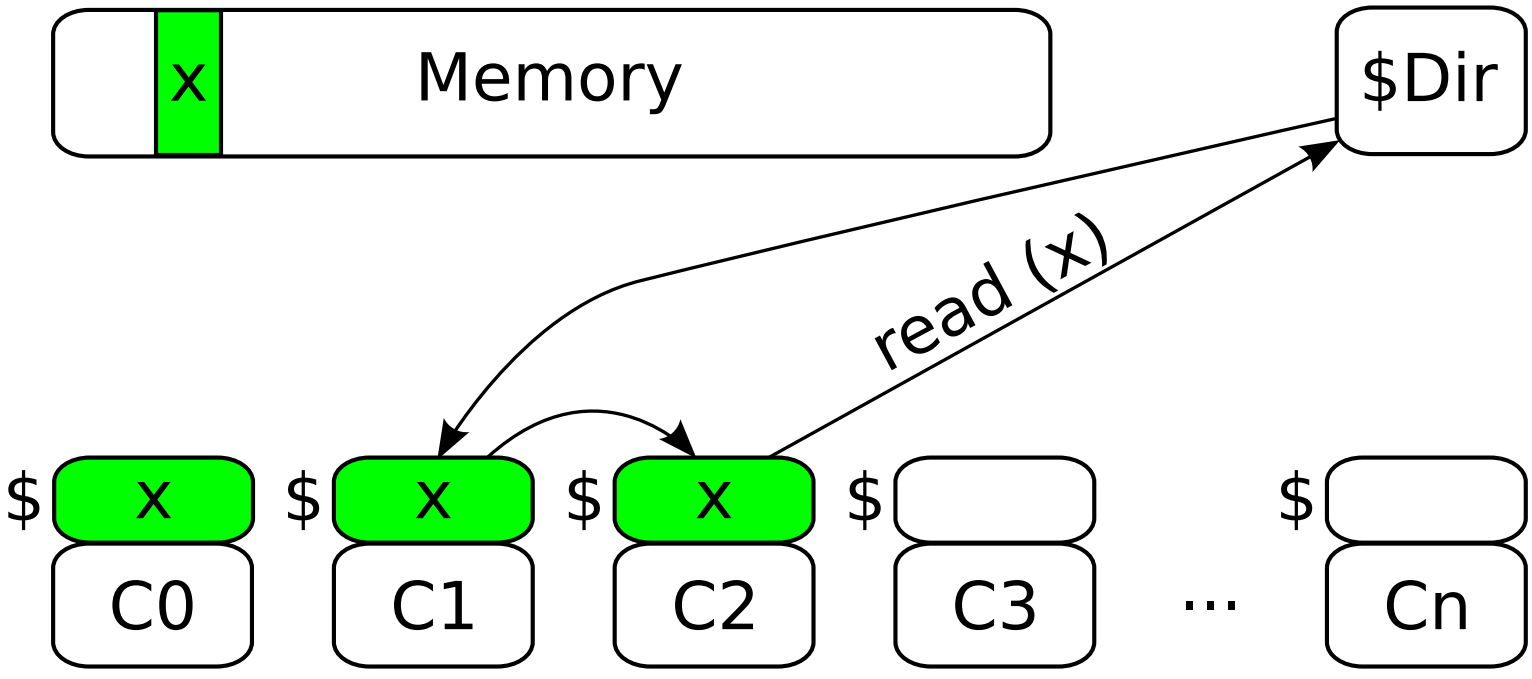
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



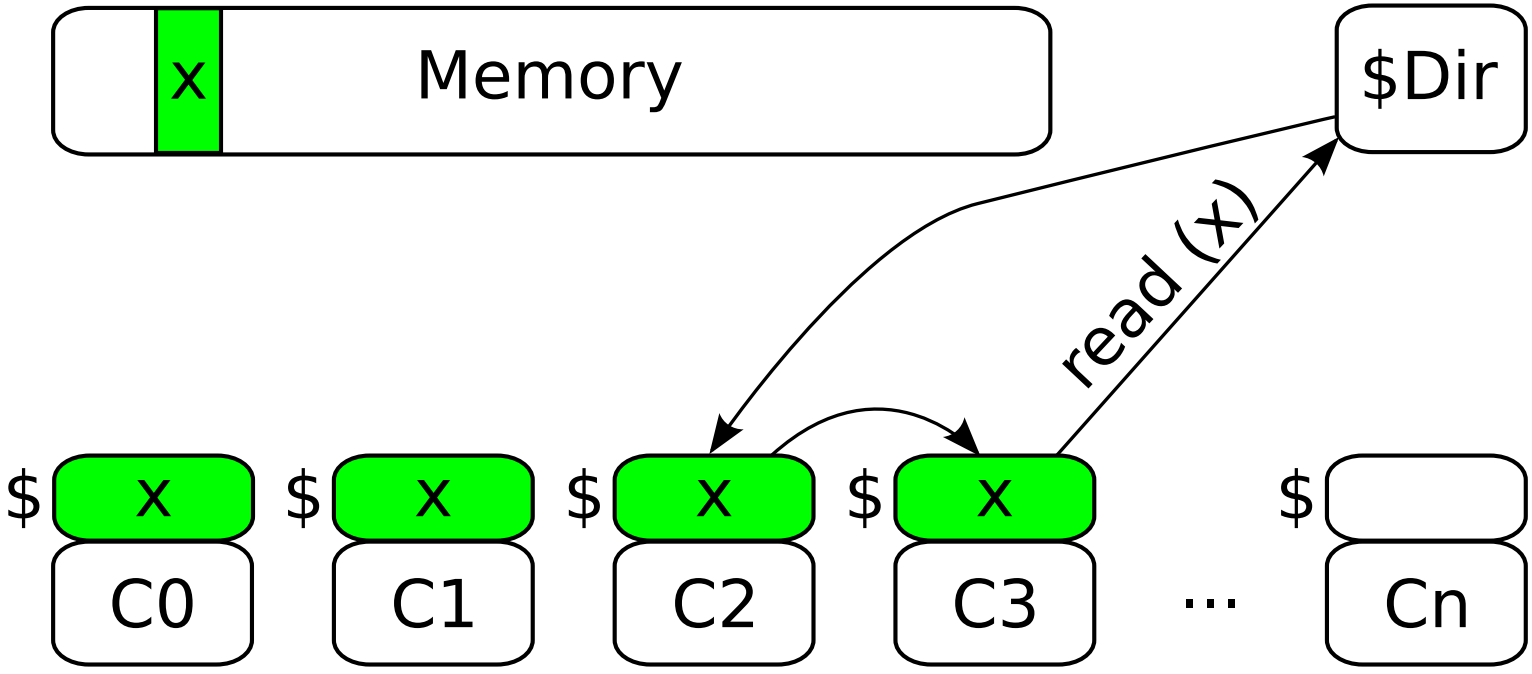
EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg

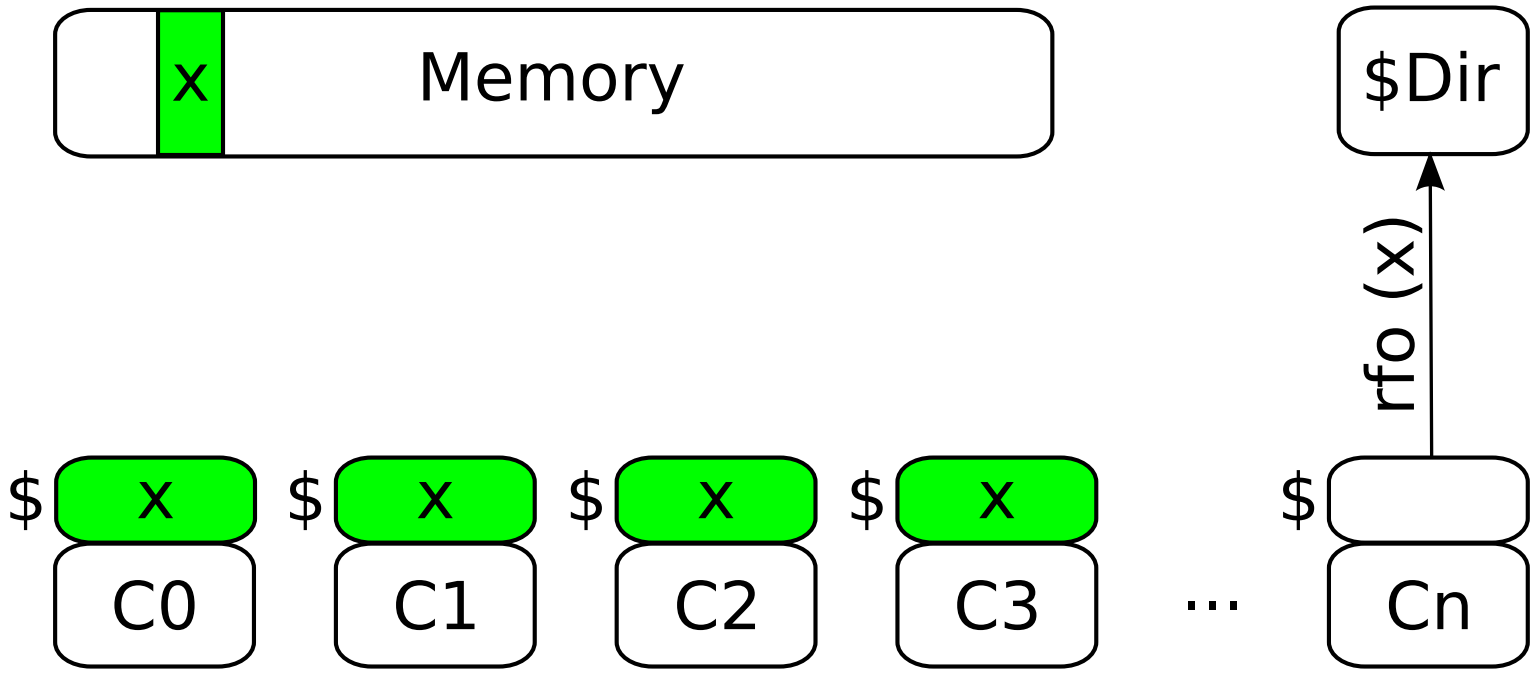


EXAMPLE: READ FOR OWNERSHIP

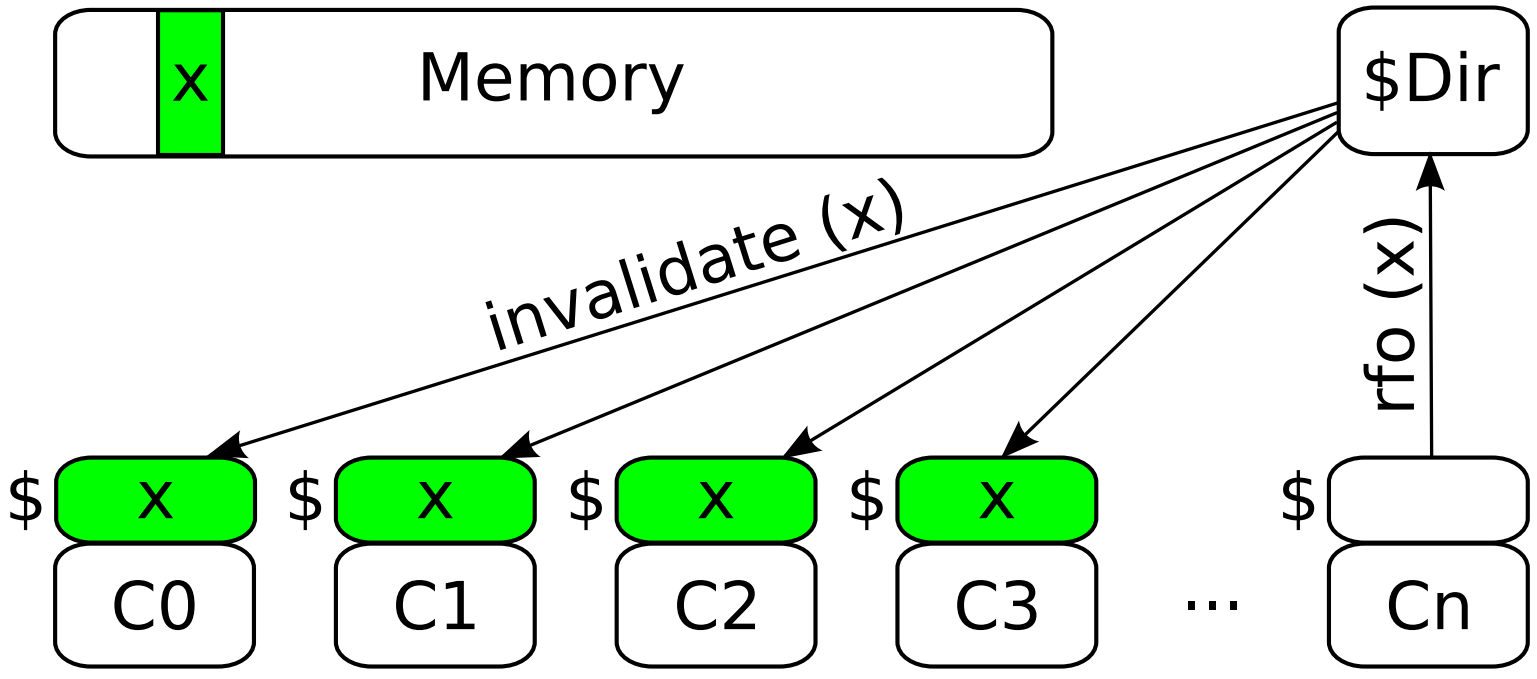
Brandenburg
University of Technology
Cottbus - Senftenberg



EXAMPLE: READ FOR OWNERSHIP

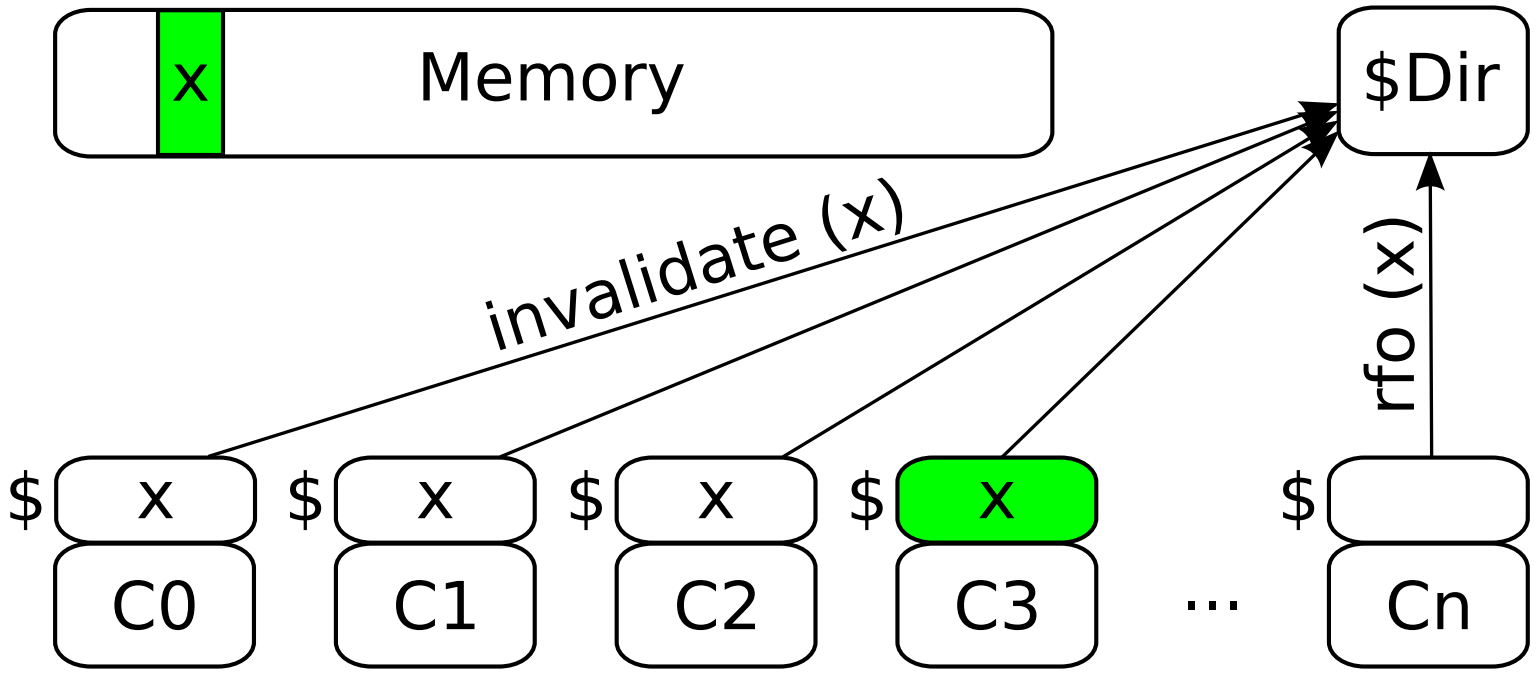


EXAMPLE: READ FOR OWNERSHIP

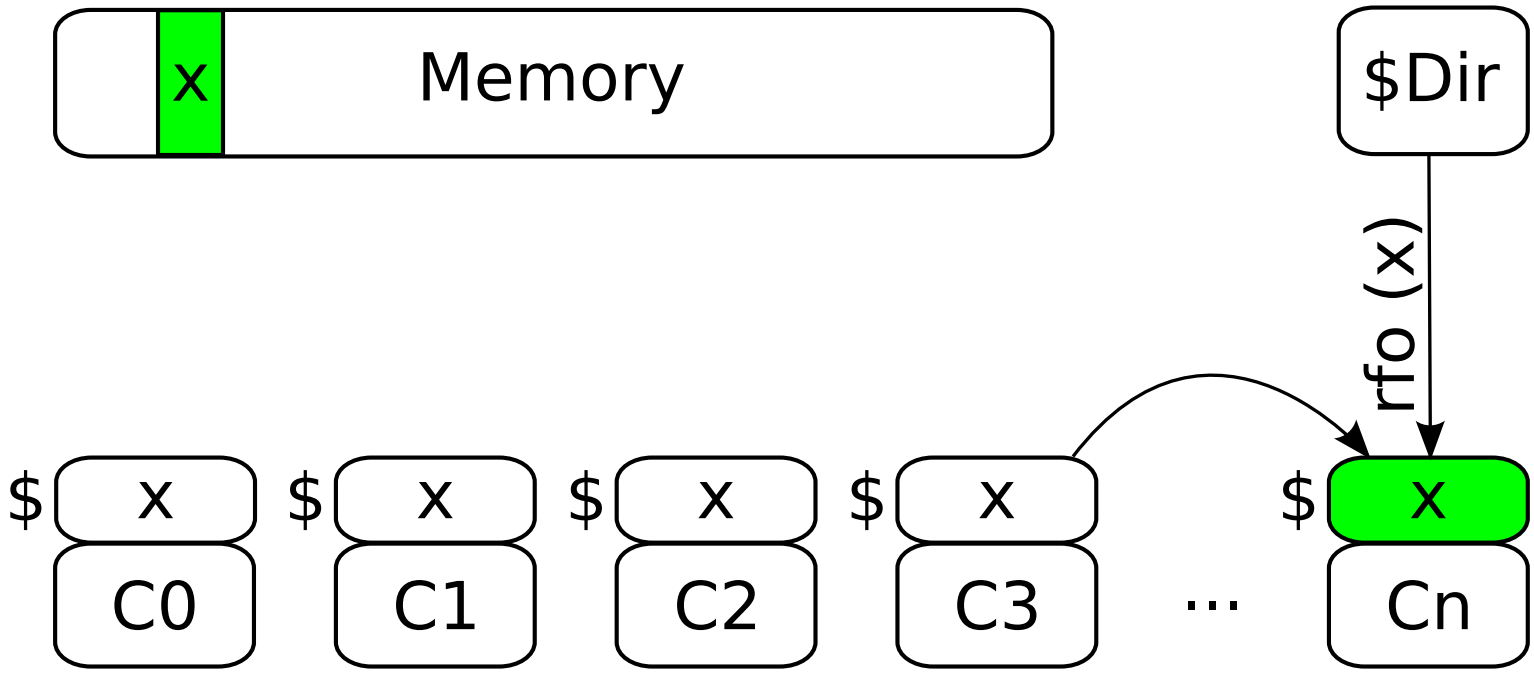


EXAMPLE: READ FOR OWNERSHIP

Brandenburg
University of Technology
Cottbus - Senftenberg



EXAMPLE: READ FOR OWNERSHIP



OUTLINE

Brandenburg
University of Technology
Cottbus - Senftenberg

1. Throughput & Latency of Broadcast
2. The Diamond Ring Topology
3. Evaluation

Latency

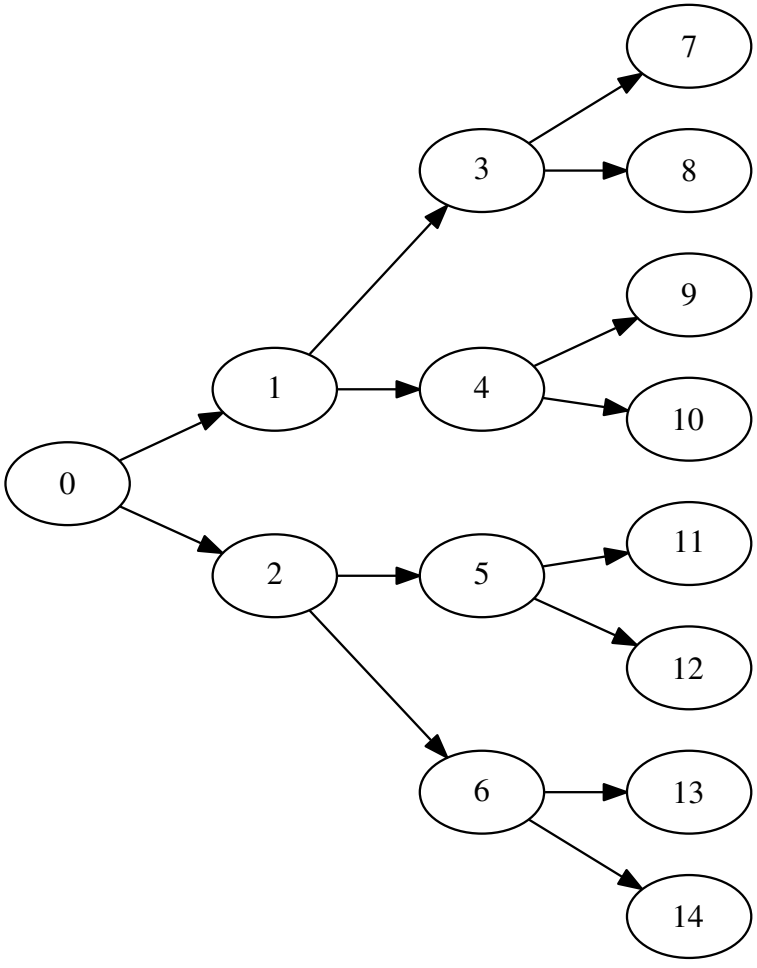
- time from sending out message to reception of acknowledgement
- determined by “longest path” (#hops + processing at each node)
- “lower is better”

Throughput

- number of messages processed within fixed time span
- determined by node with maximum overhead (i.e. bottleneck)
- requires pipelining of messages (latency hiding)
- “higher is better”

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

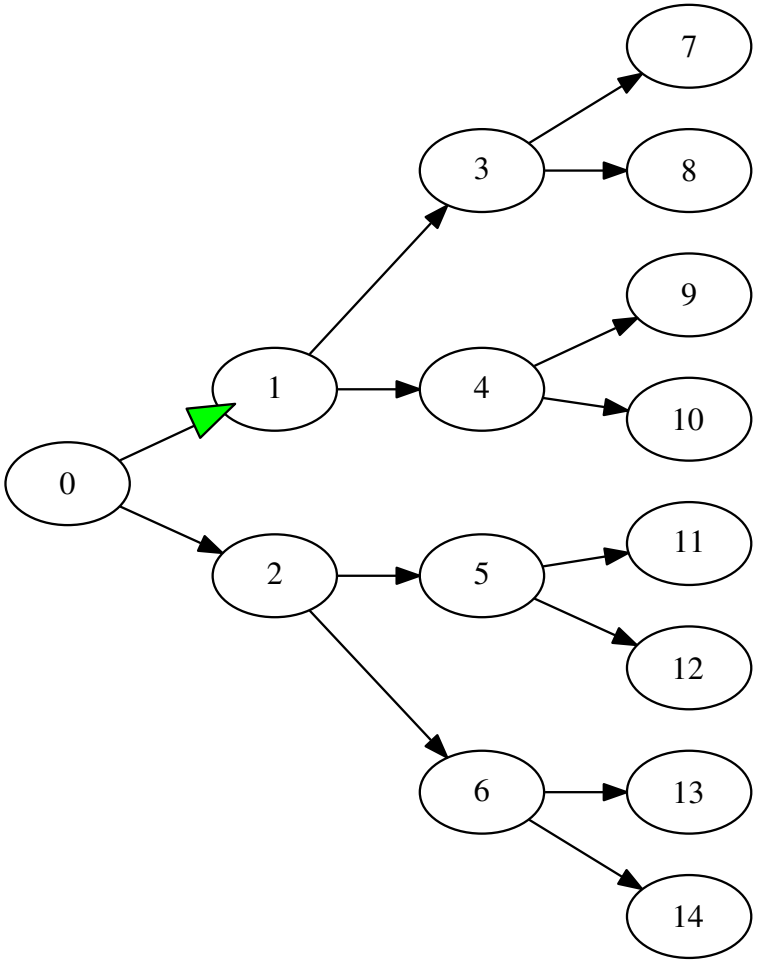
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

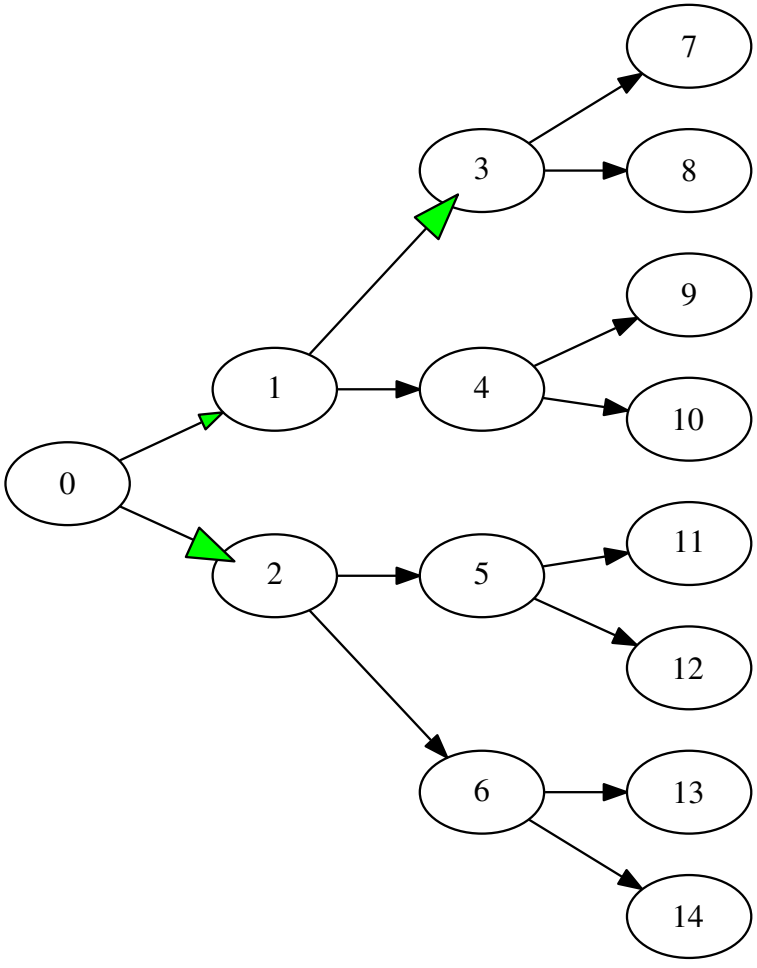
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

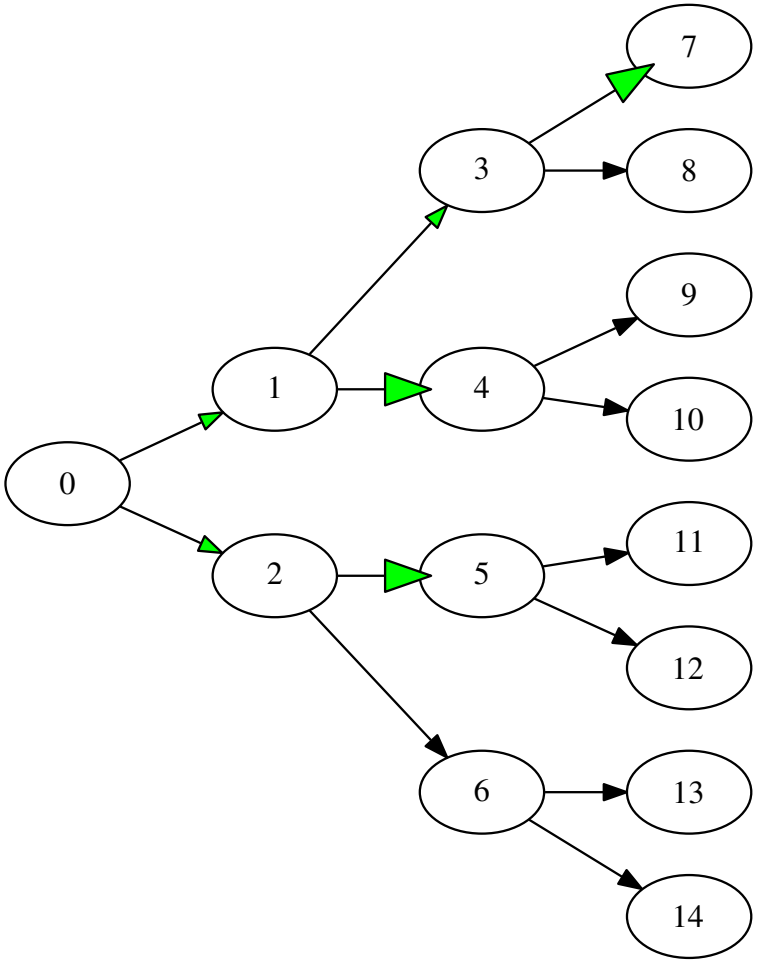
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

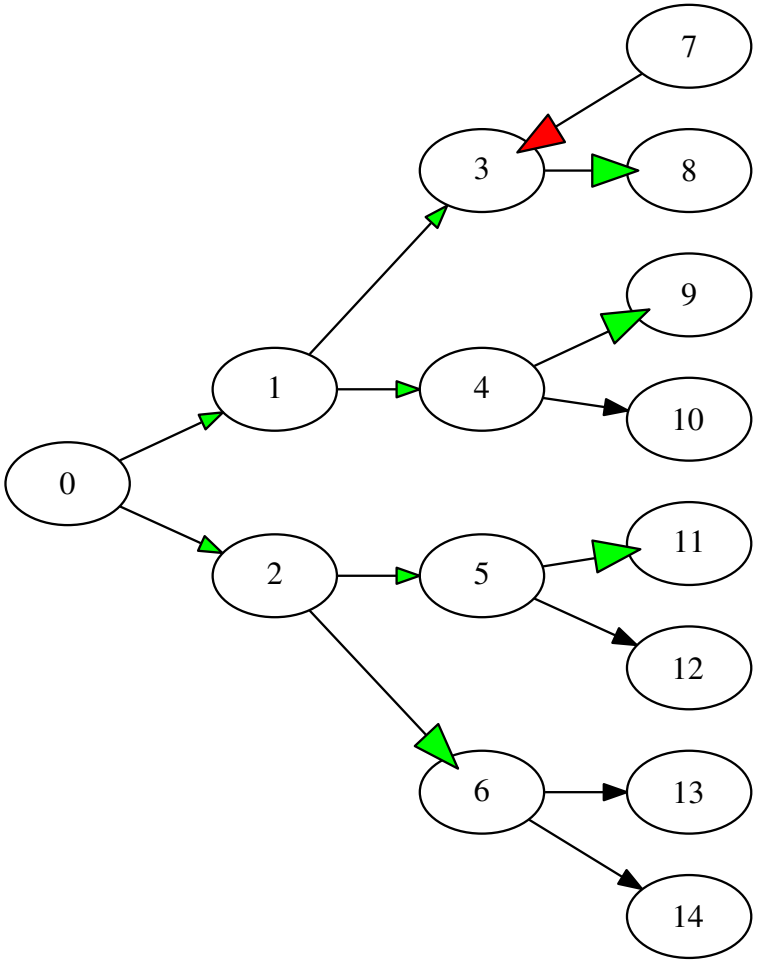
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

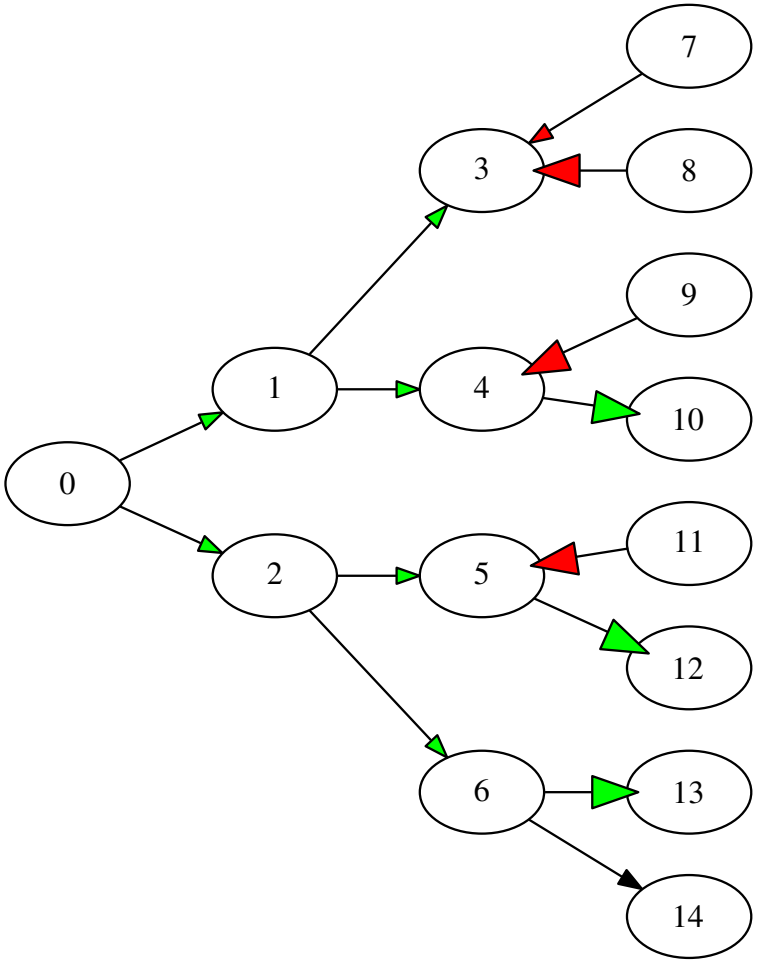
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

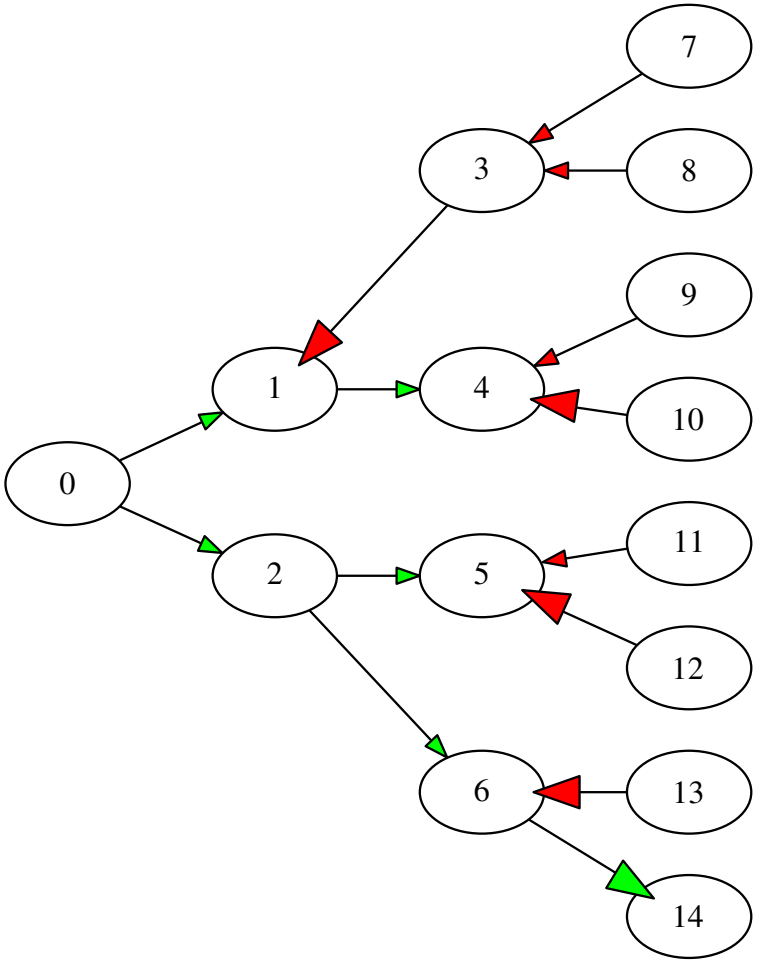
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

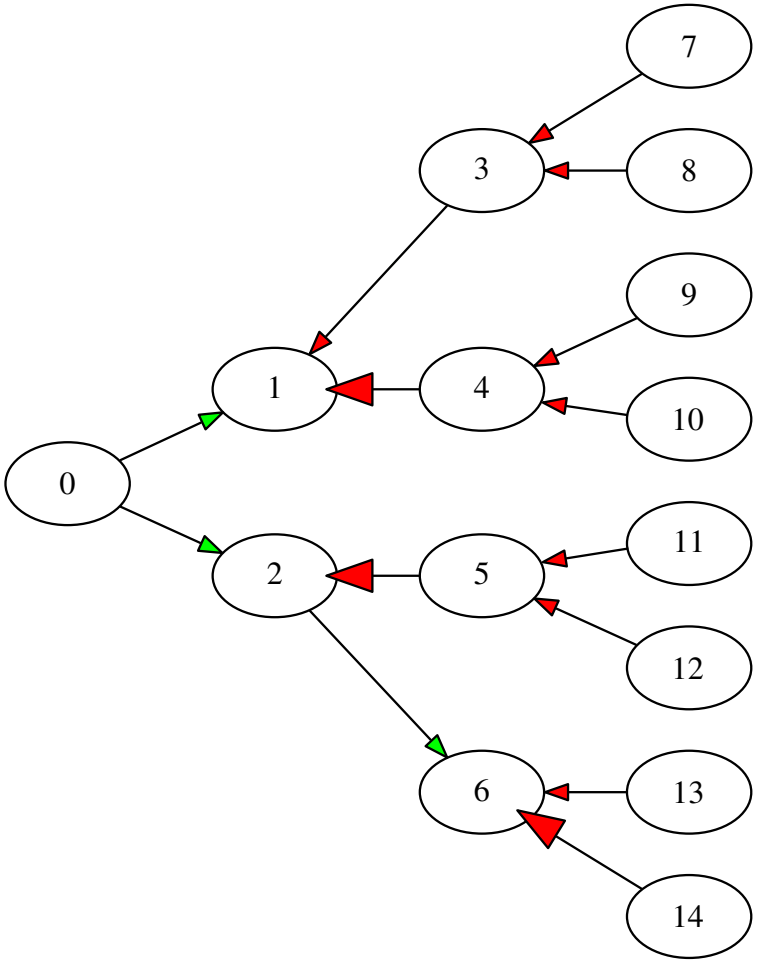
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

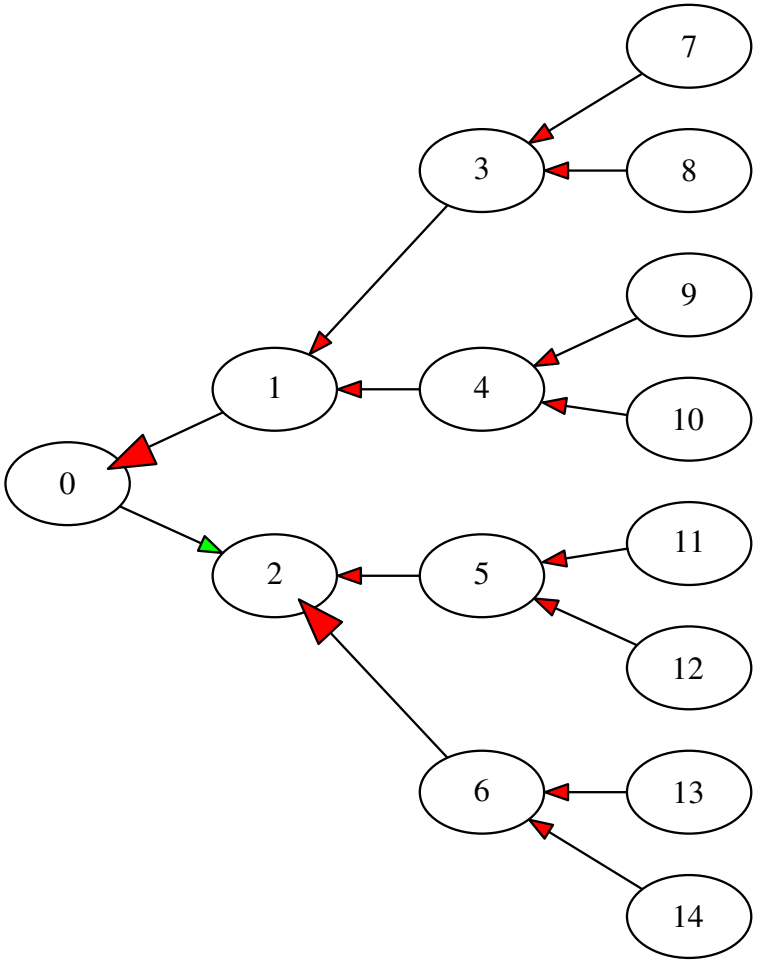
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

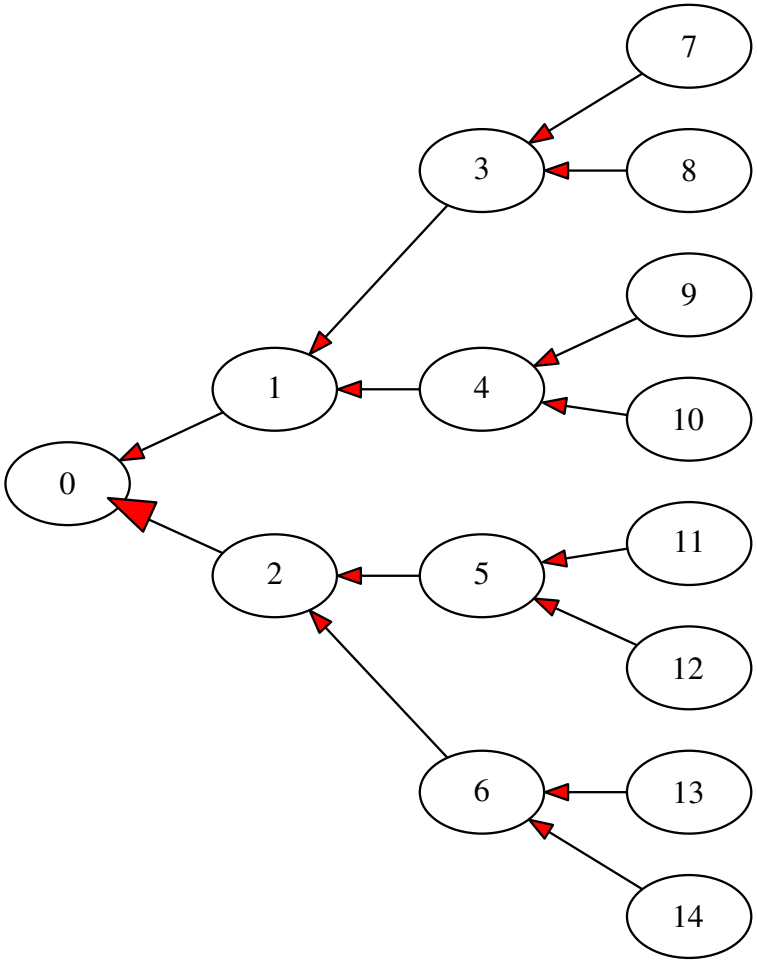
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING BALANCED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

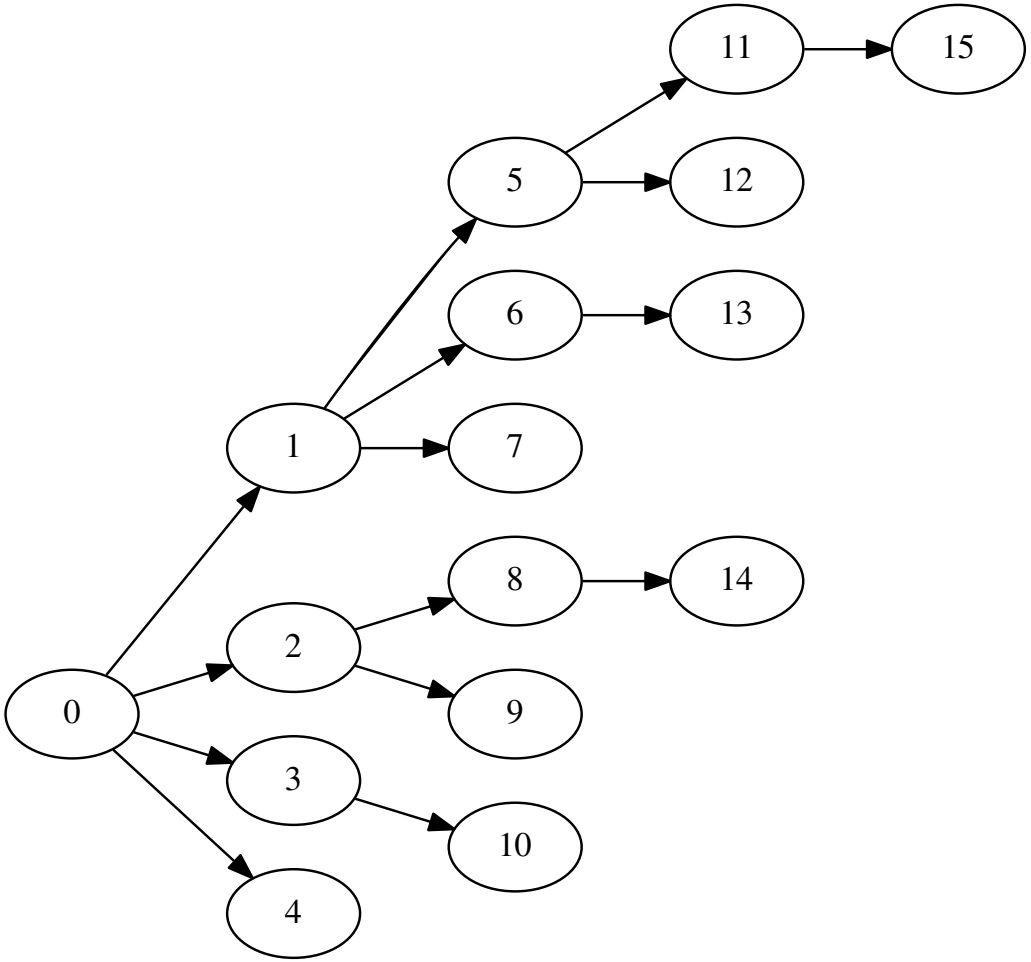
★★☆

Latency

★★☆

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

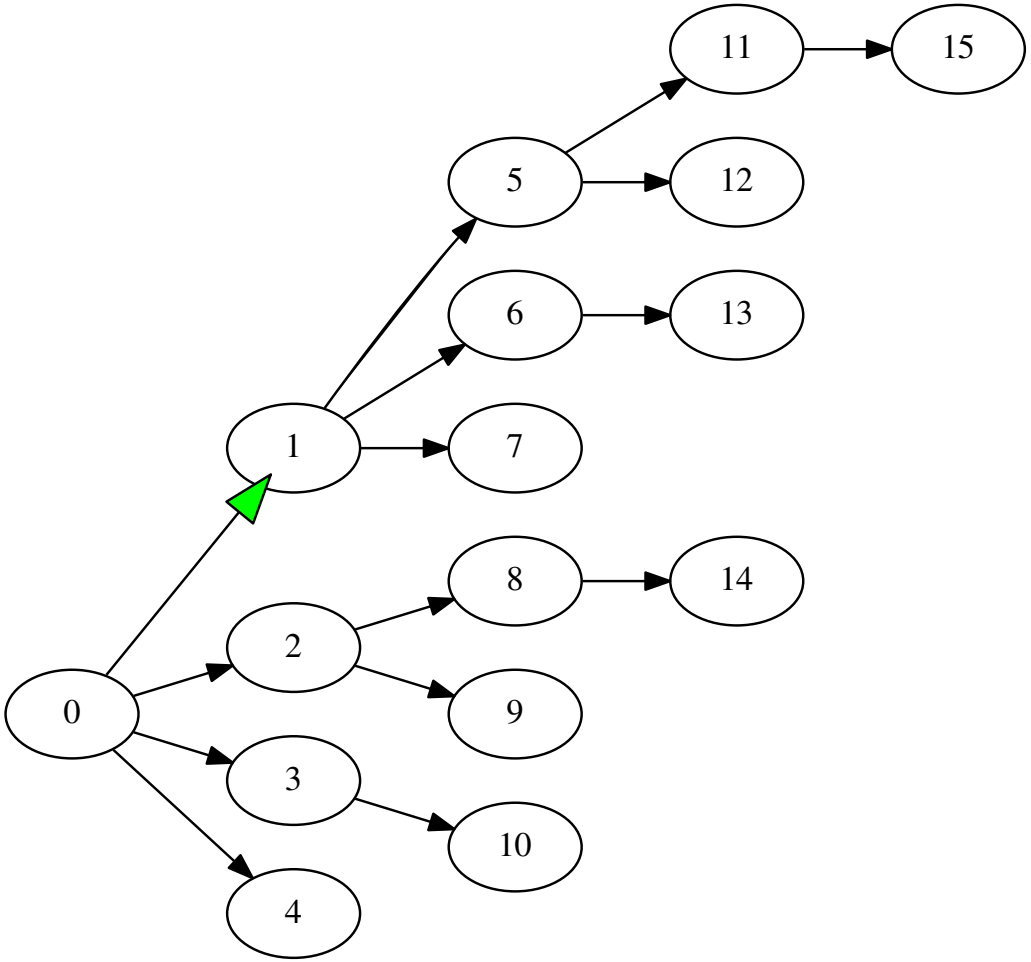
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

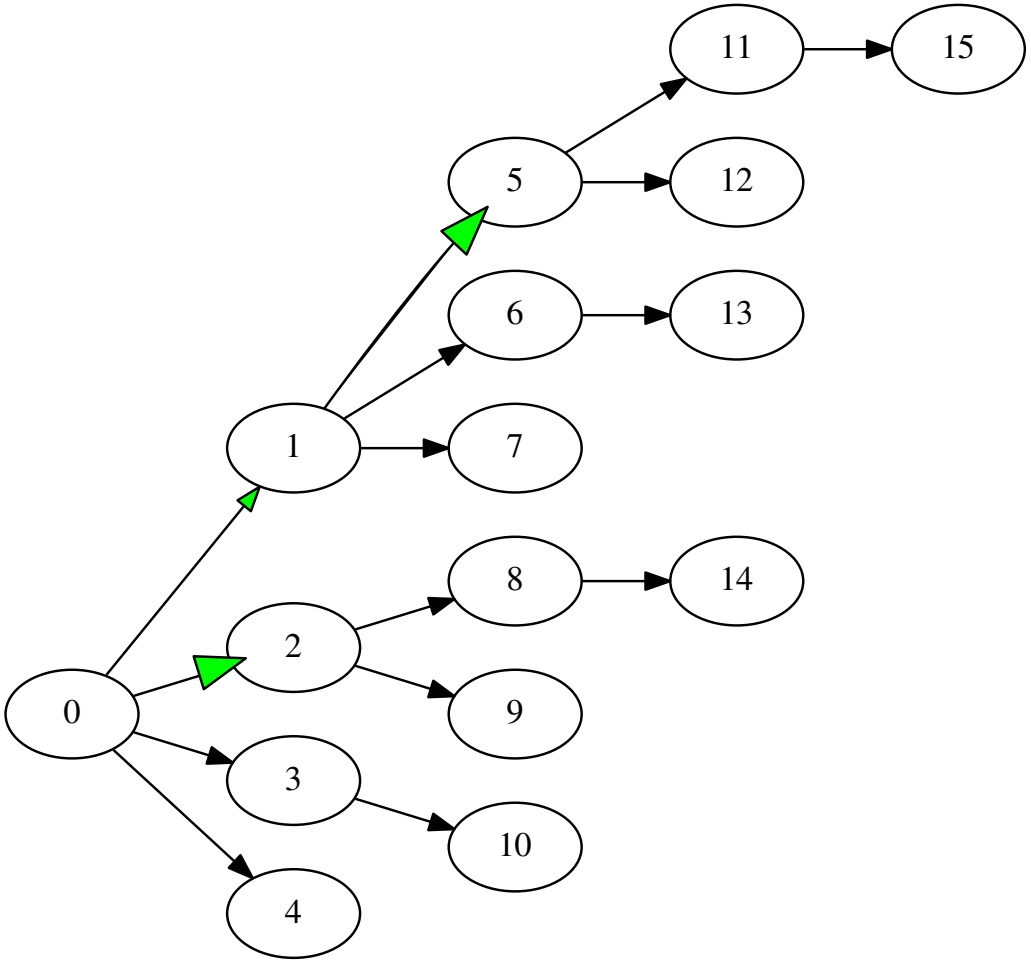
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

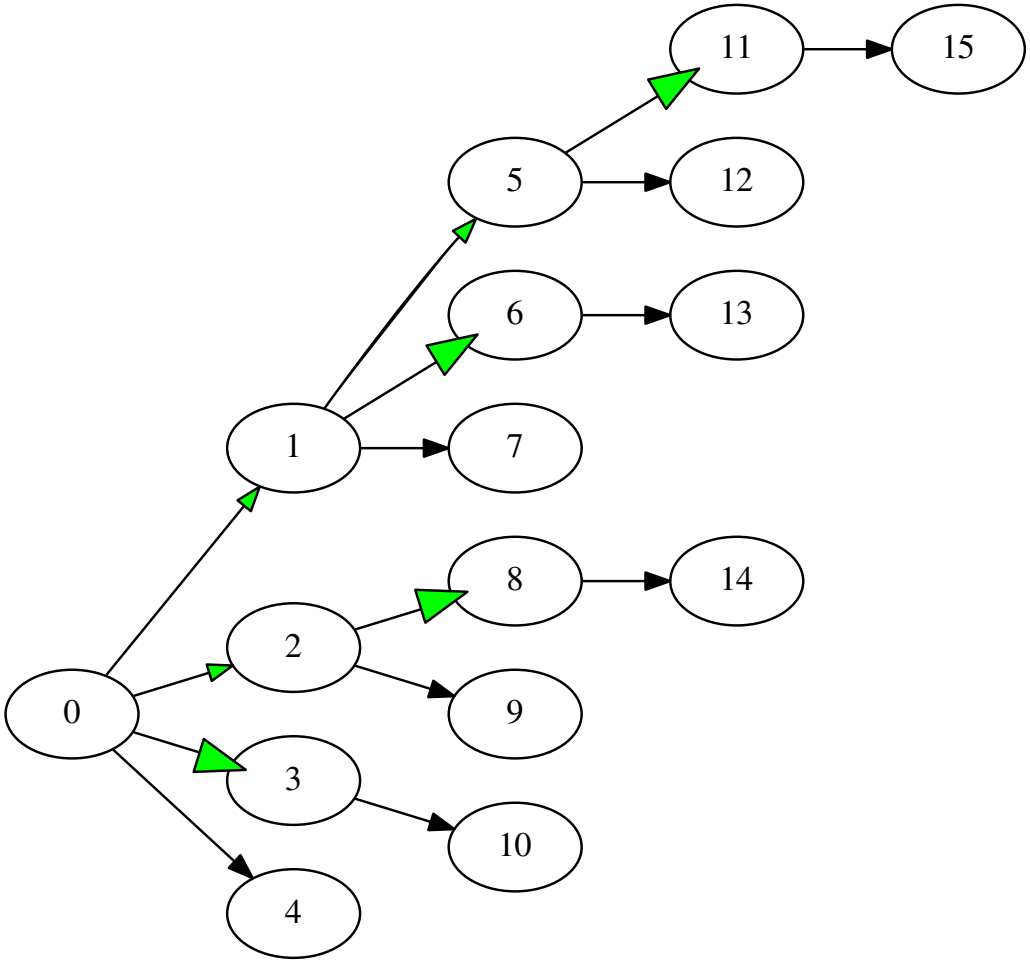
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

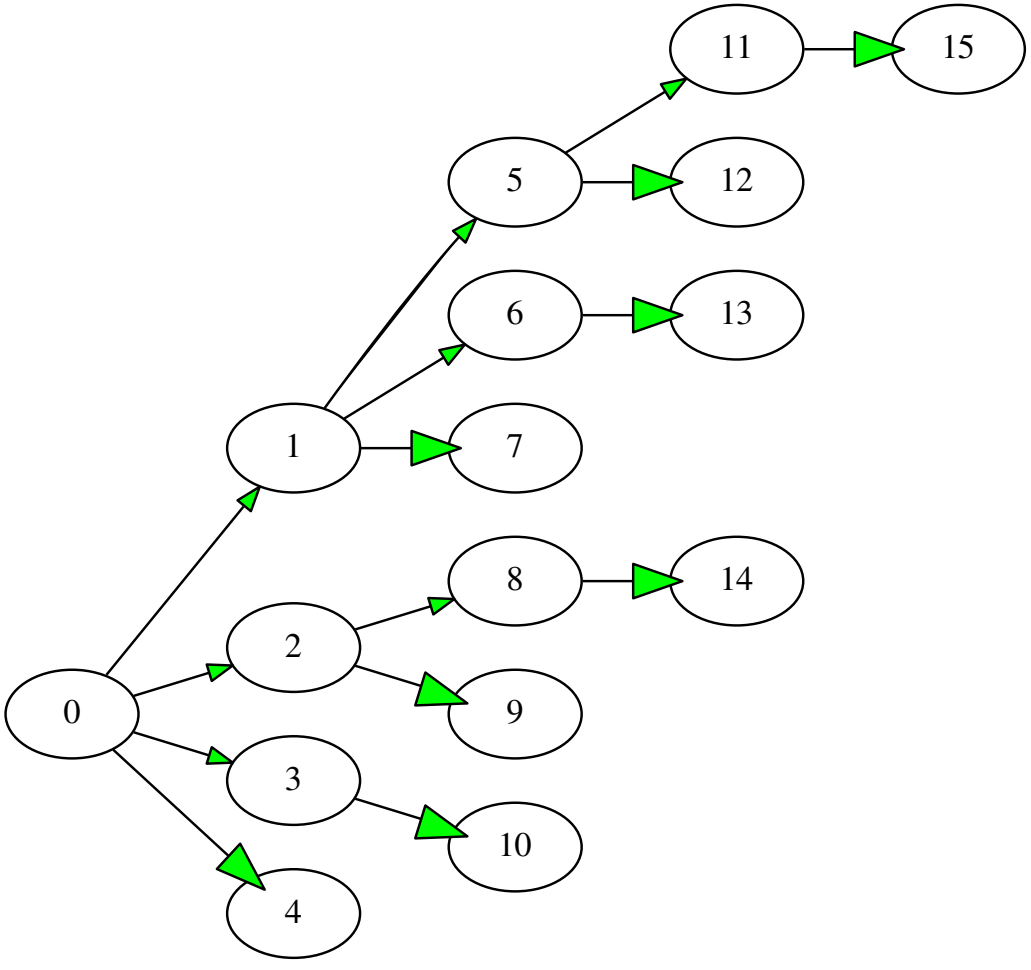
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

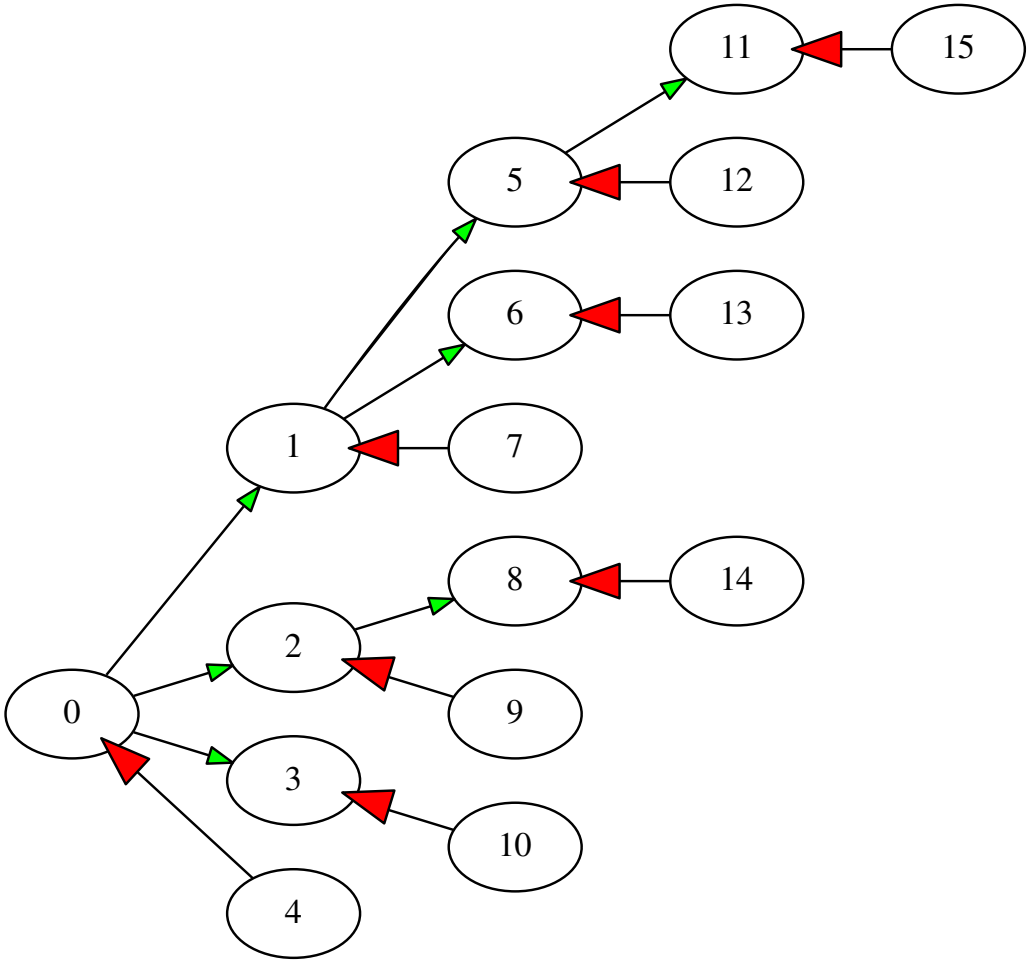
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

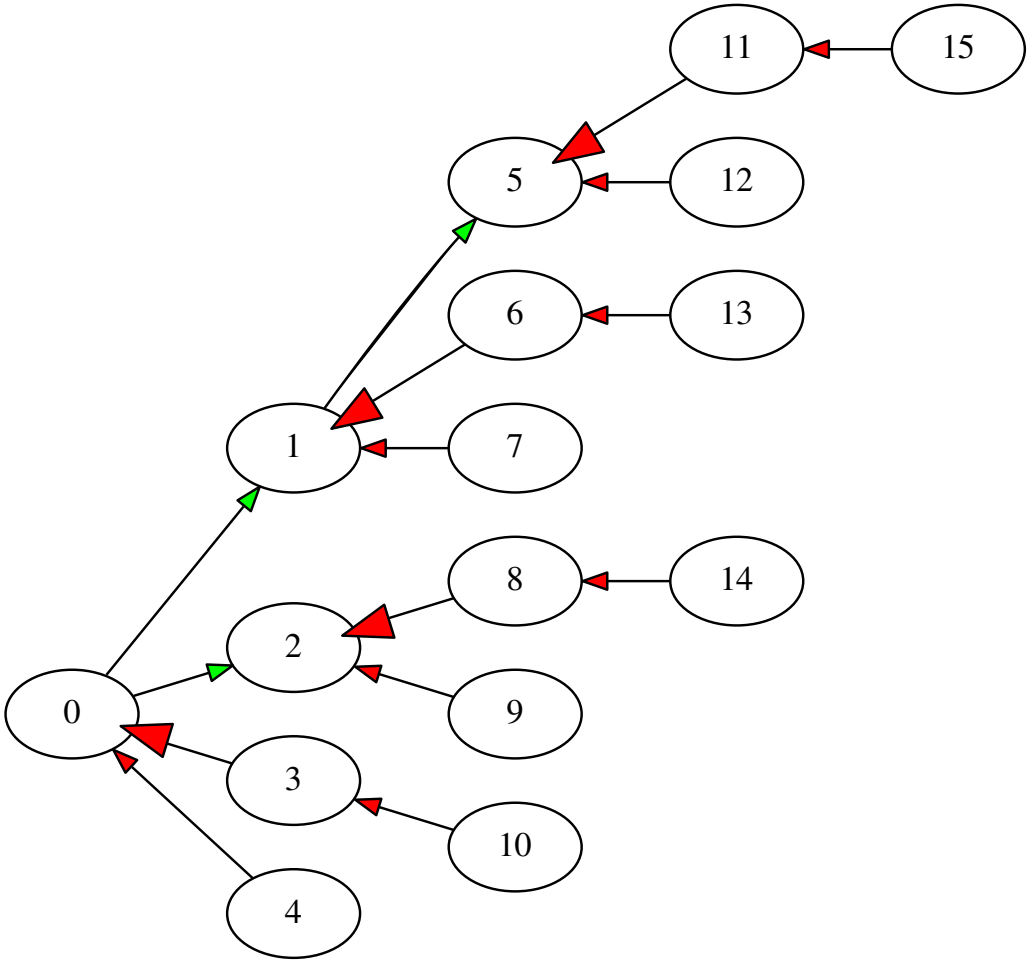
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

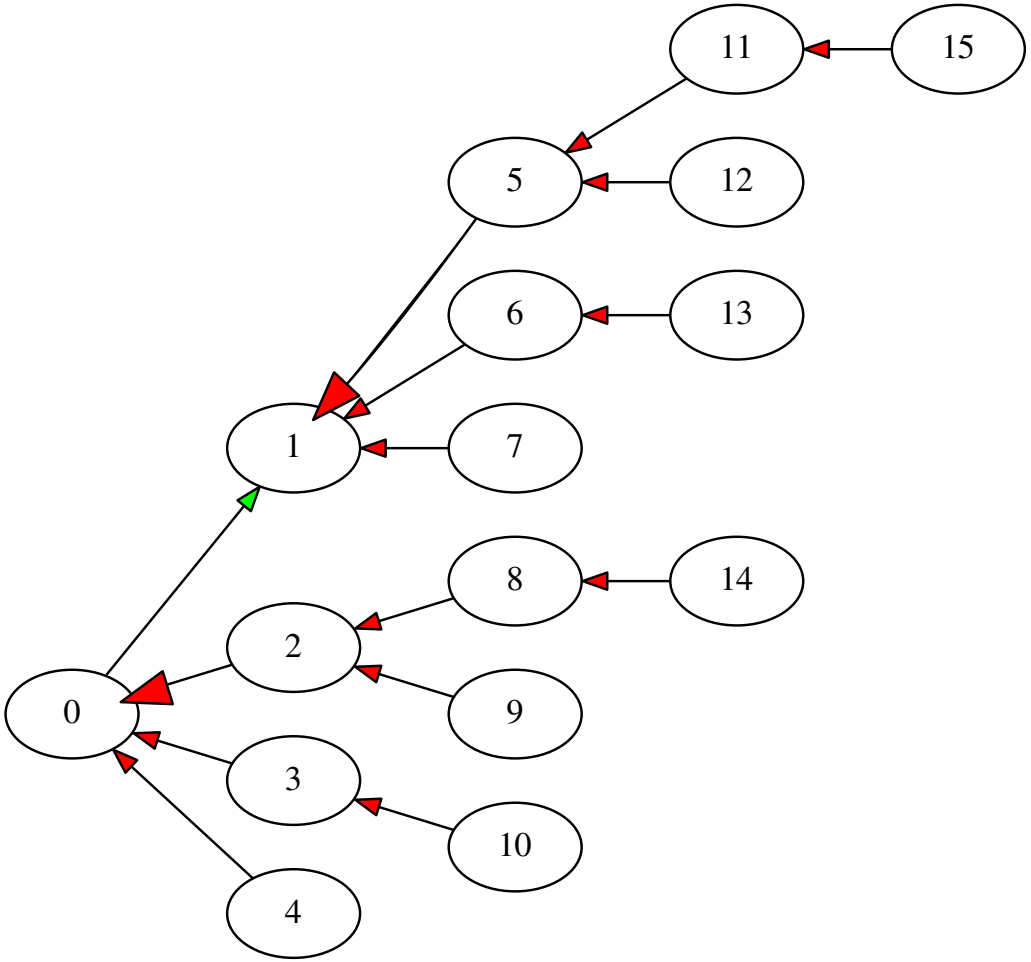
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

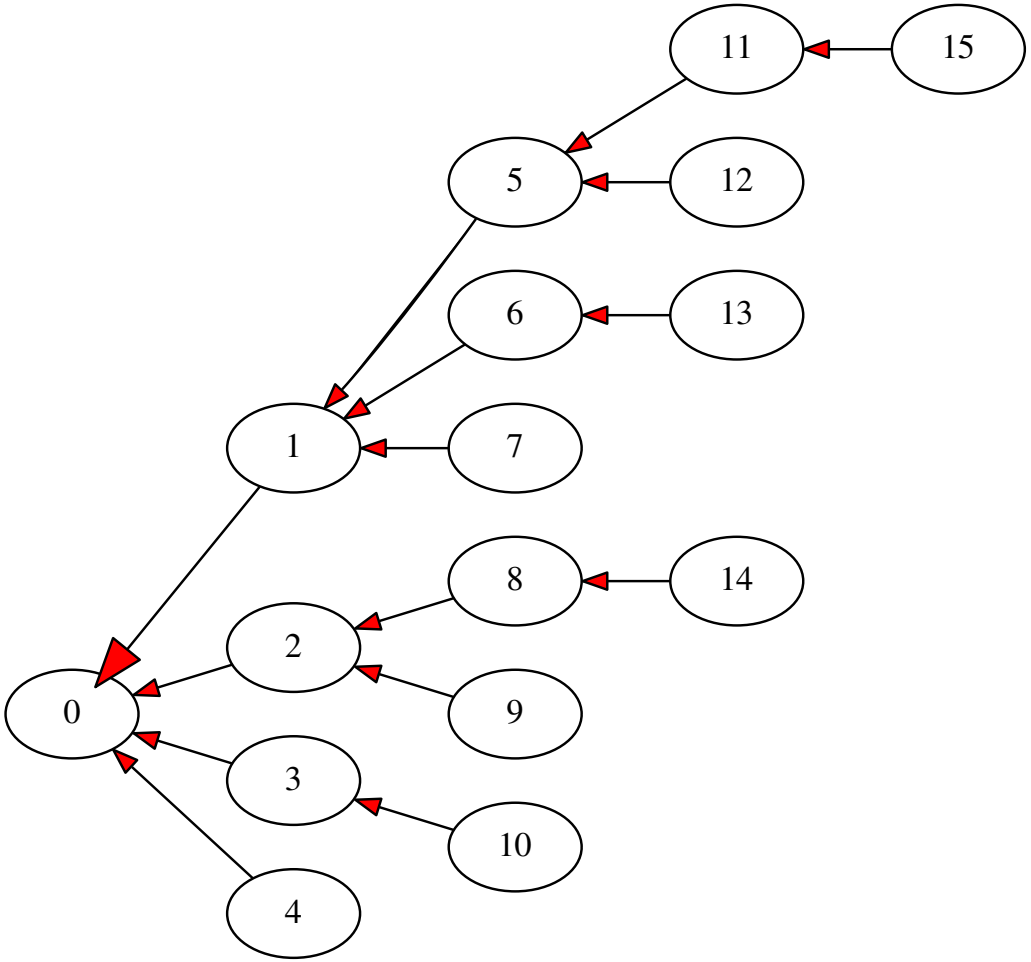
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING SKEWED TREES

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

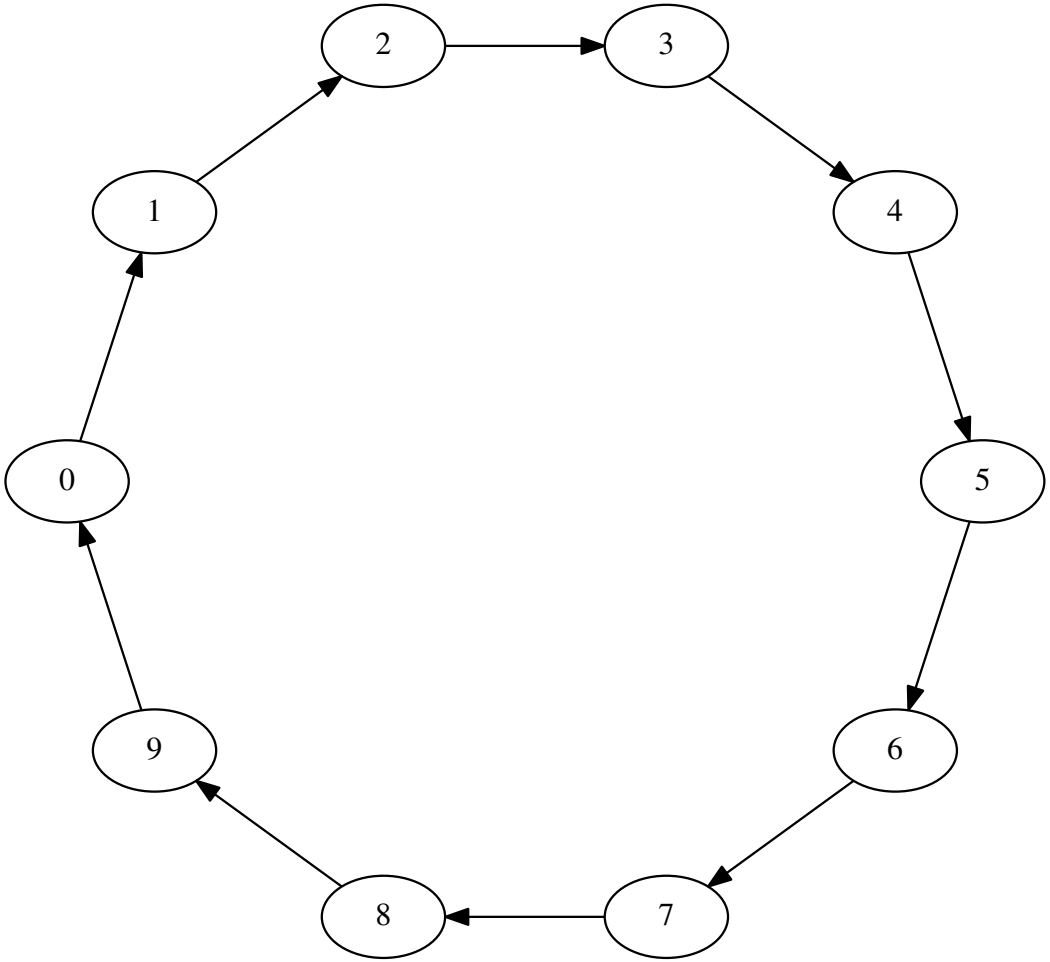
★☆☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

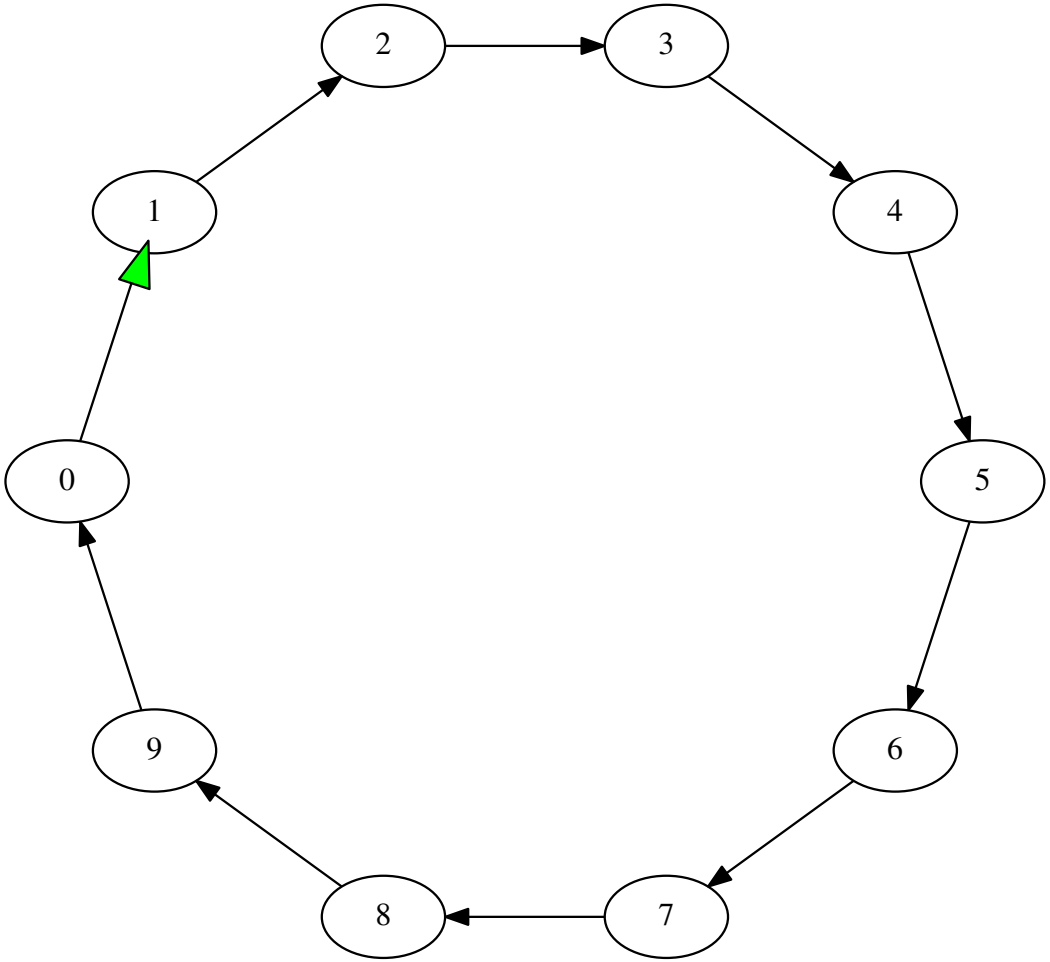
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

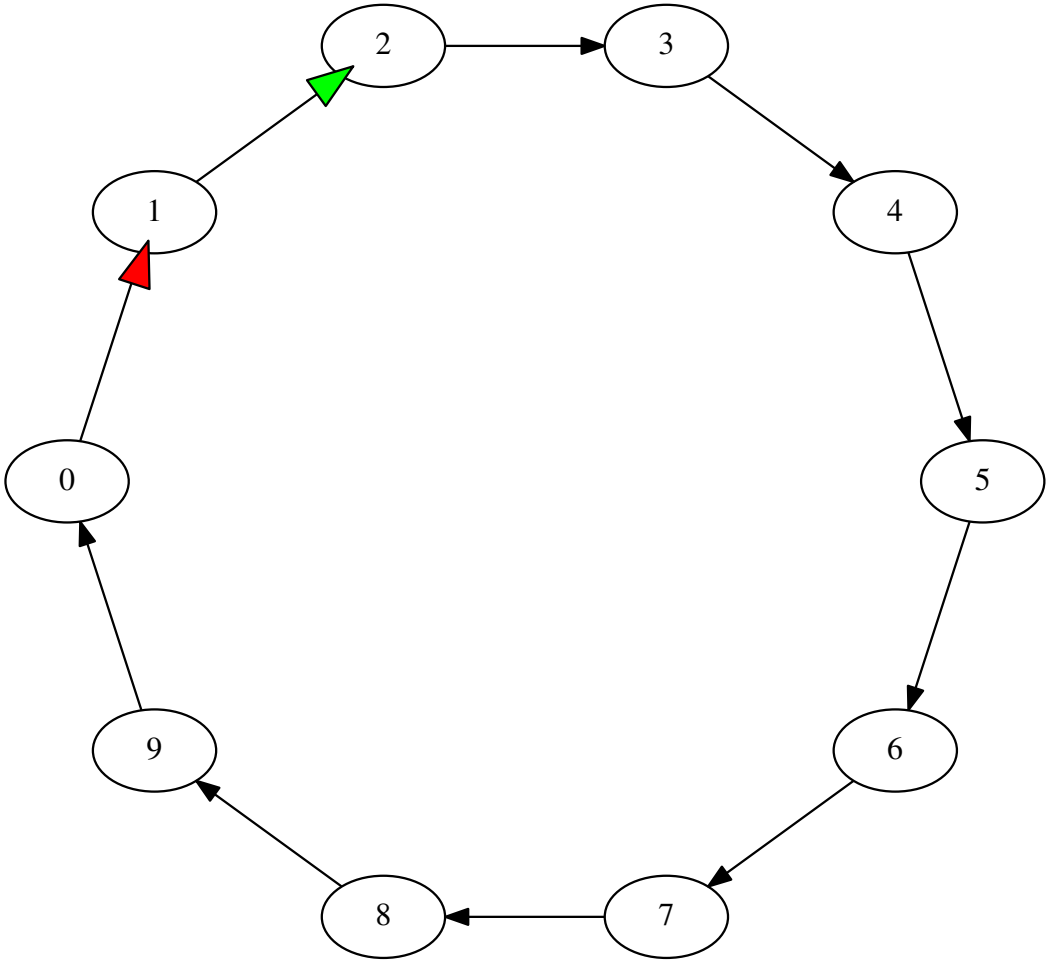
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

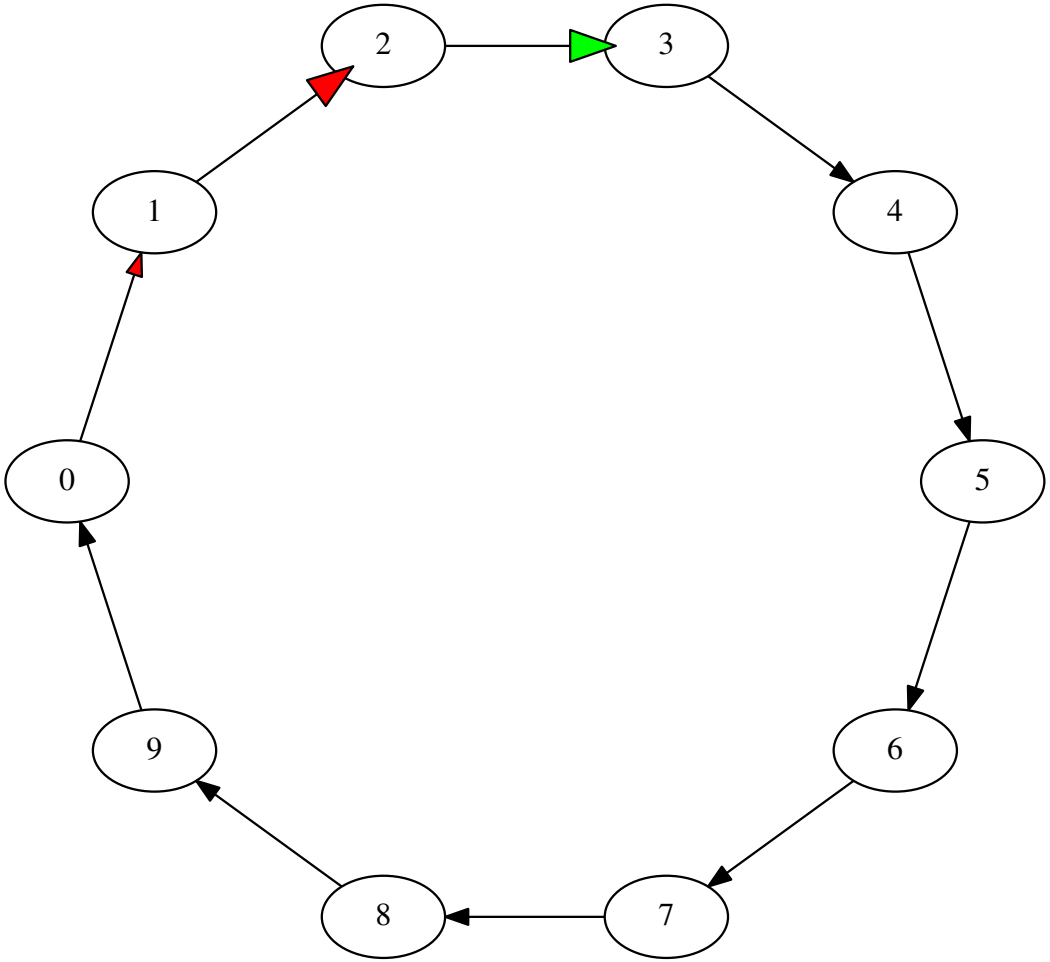
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

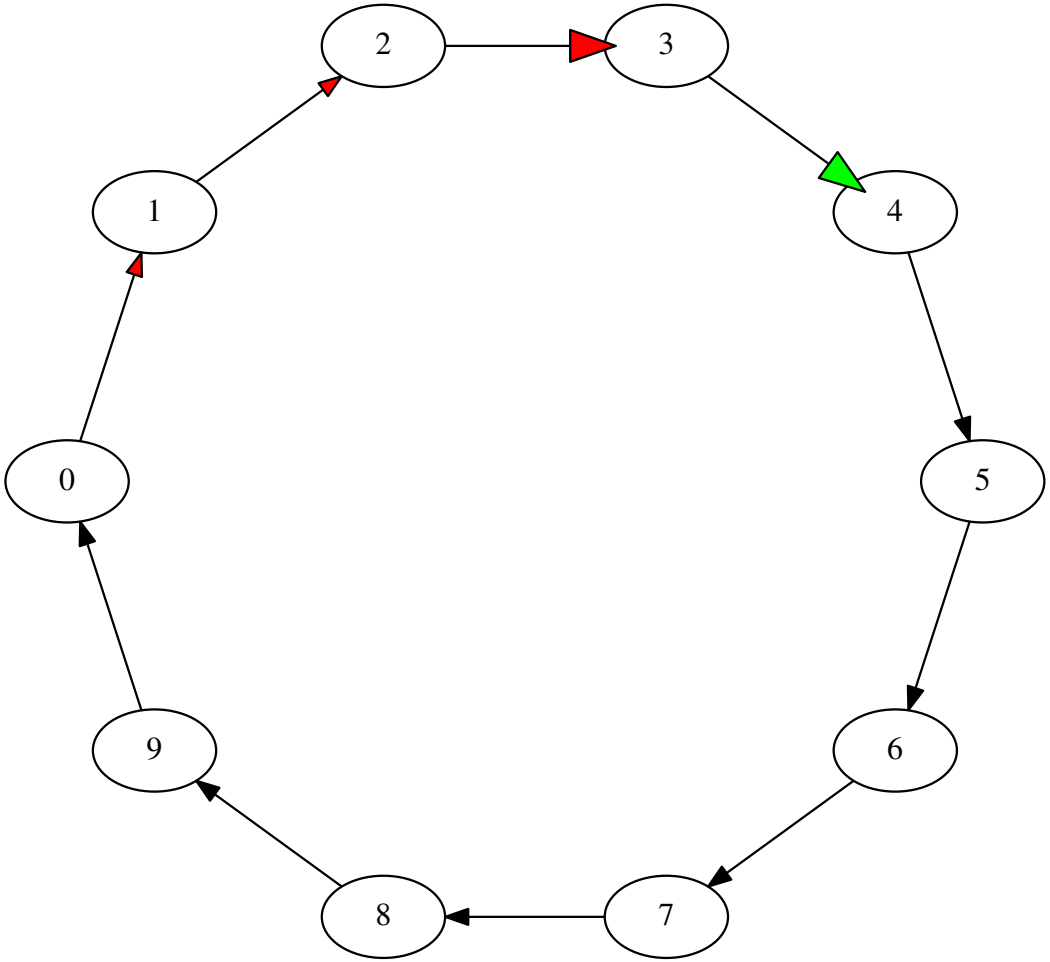
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

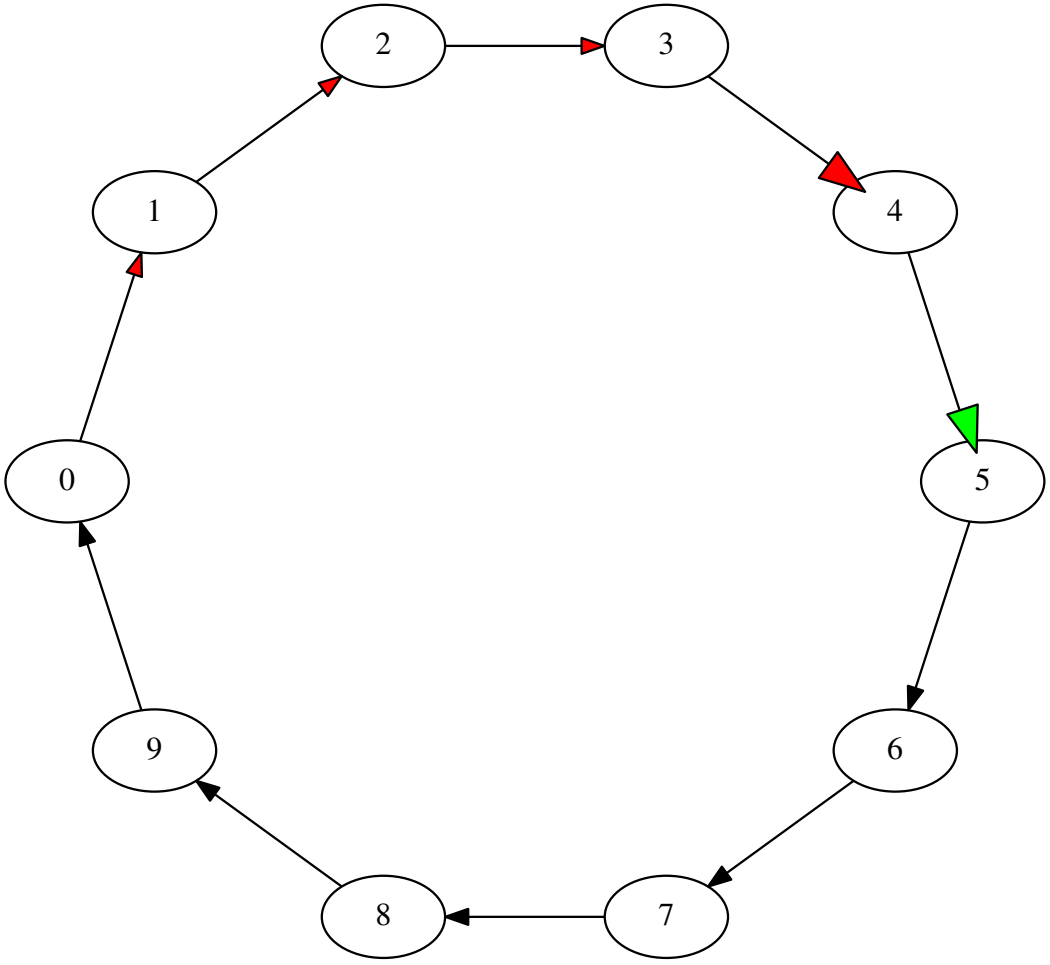
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

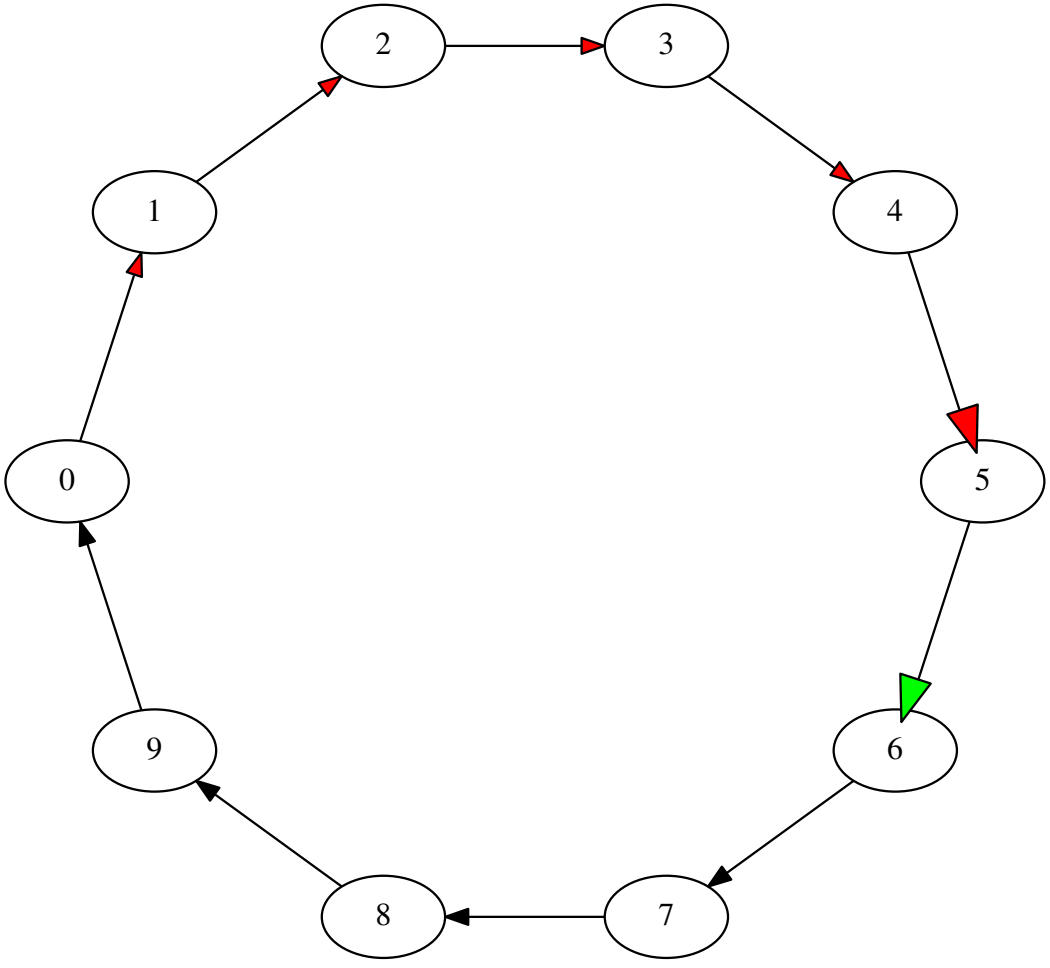
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

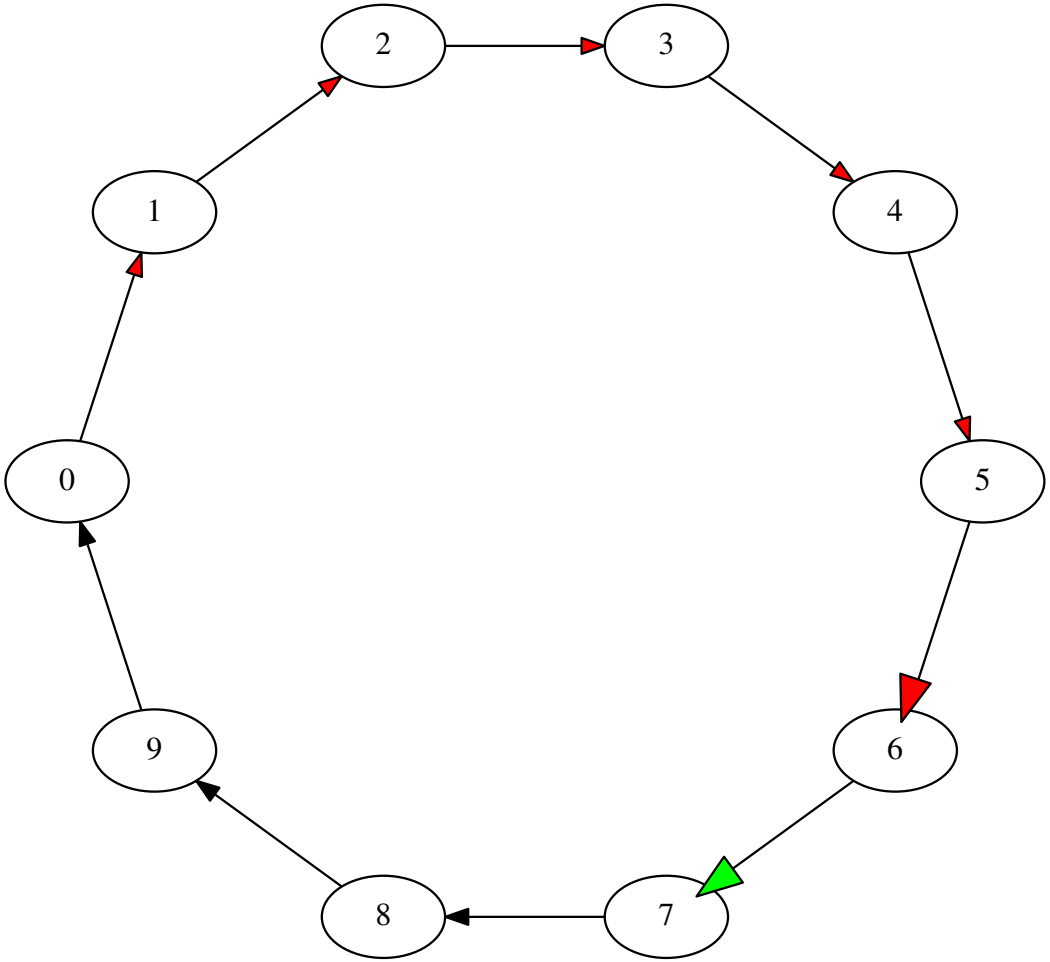
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

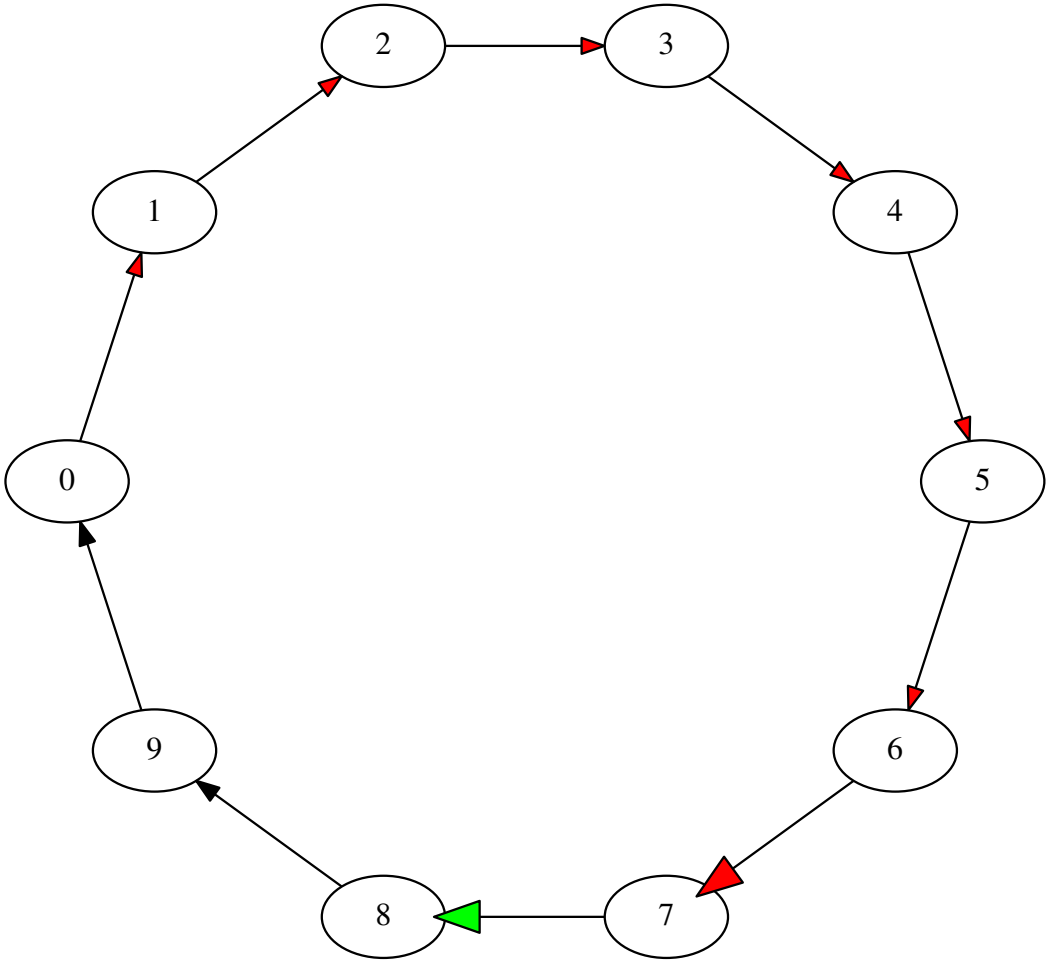
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

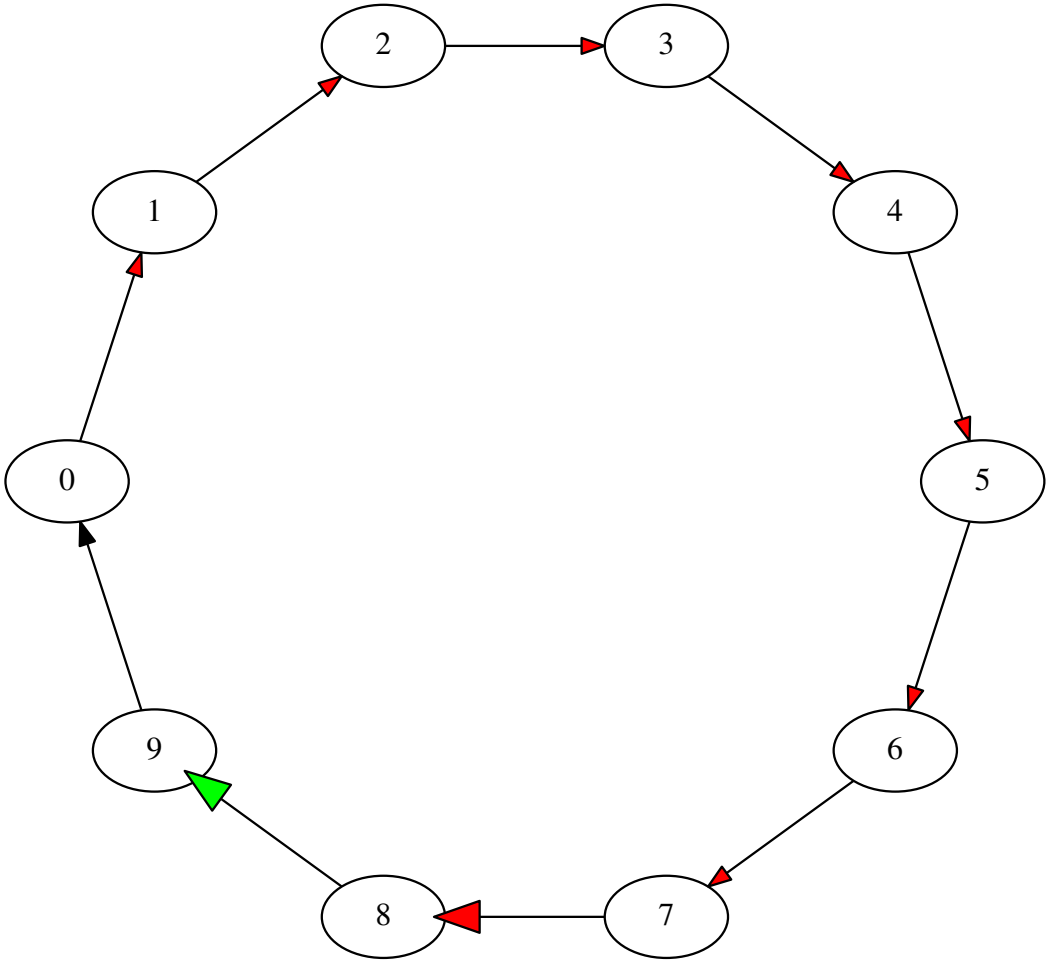
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

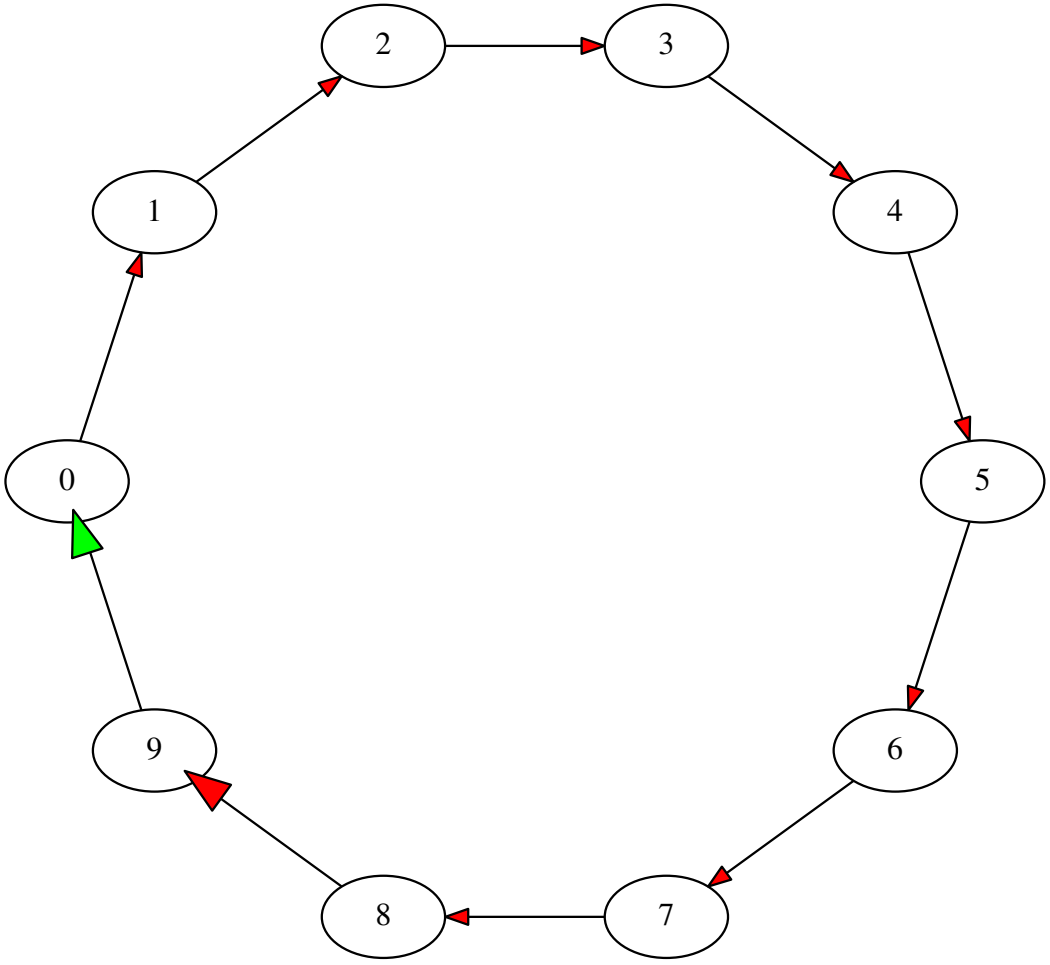
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

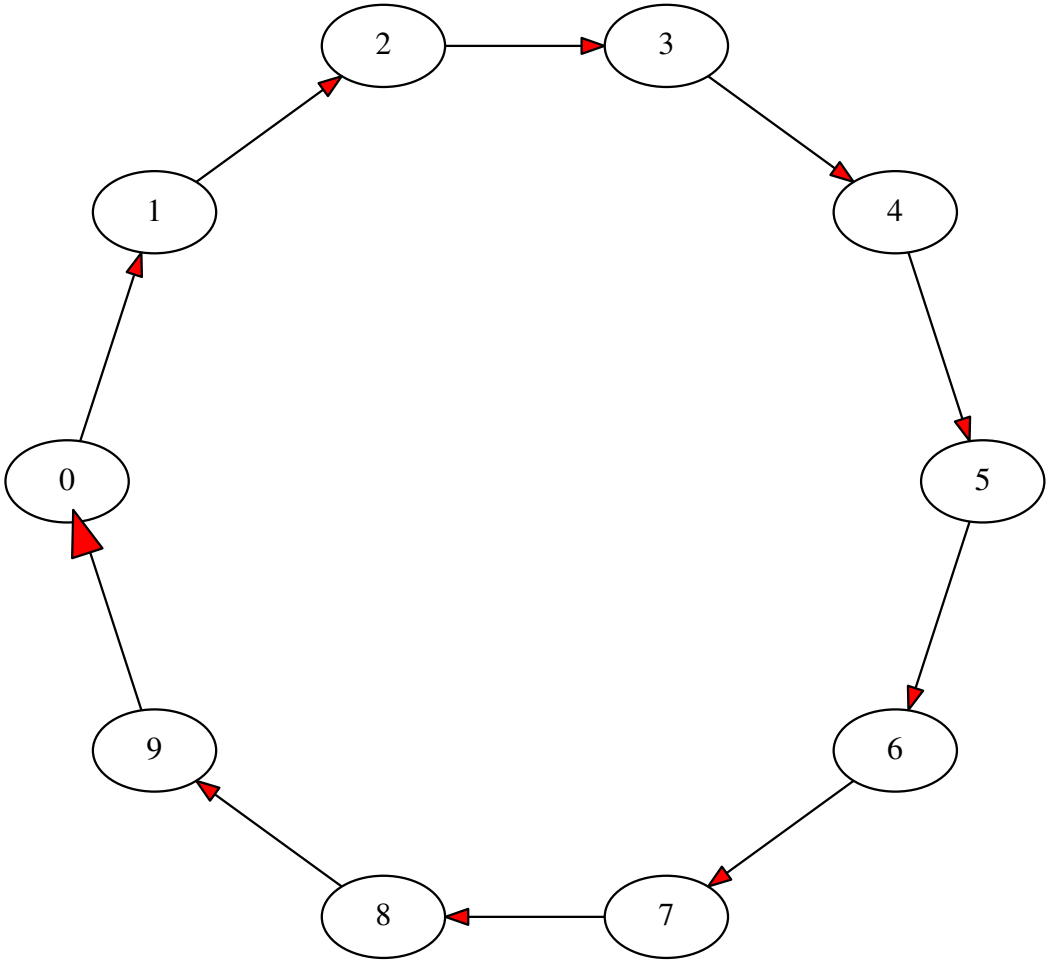
★★★

Latency

★★★

ACKNOWLEDGED BROADCAST USING RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

★★★

Latency

★★★

FORWARD – PROCESS – ACK

Message Forwarding as Acknowledgement

- possible in ring structures
- halve number of sent messages (network contention)
- may increase latency (processing time at node)

Ring Structure

1. Receive Message
2. Process Message
3. Forward Message (Ack)

Tree Structure

1. Receive Message
2. Forward Message (except leaves)
3. Process Message
4. Receive Ack (except leaves)
5. Forward Ack

Not an issue if only message reception needs acknowledgement.

OUTLINE

Brandenburg
University of Technology
Cottbus - Senftenberg

1. Throughput & Latency of Broadcast

2. **The Diamond Ring Topology**

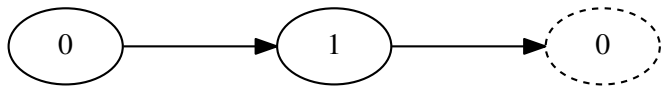
3. Evaluation

THE DIAMOND RING TOPOLOGY

Brandenburg University of Technology Cottbus - Senftenberg

Combine Ring and Balanced Tree

- Logarithmic path length for low latency
- Forwarding is acknowledgement
- Parallel message propagation
- Computable topology



Diamond Ring: Directed Graph D_k^l

k Arity of tree nodes

l Levels of tree scattering

- Based on a balanced tree B_k^l
- Mirrored at the leaves
- Closed to ring at the root

$$|D_k^l| = \frac{(k+1)k^l-(k+1)}{k-1}$$

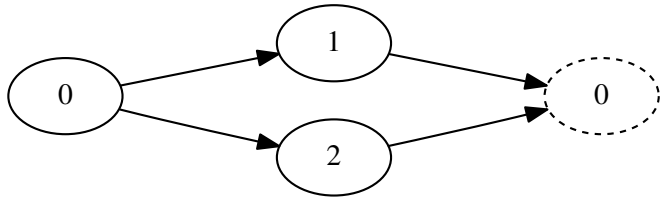
$$|D_k^{l+1}| = |D_k^l| + k^l + k^{l+1}$$

THE DIAMOND RING TOPOLOGY

Brandenburg University of Technology Cottbus - Senftenberg

Combine Ring and Balanced Tree

- Logarithmic path length for low latency
- Forwarding is acknowledgement
- Parallel message propagation
- Computable topology



Diamond Ring: Directed Graph D_k^l

k Arity of tree nodes

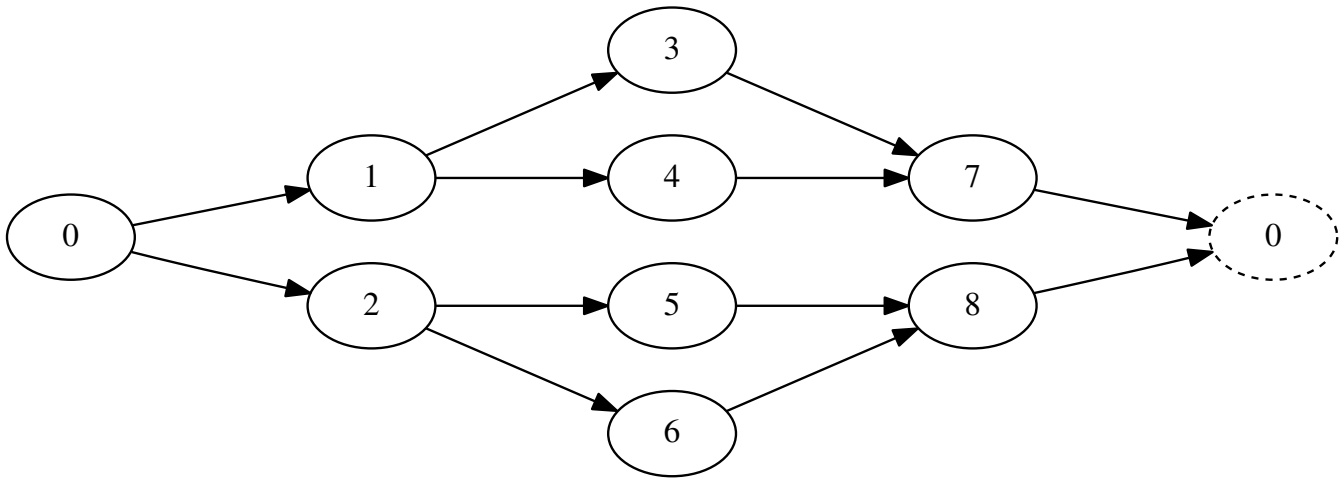
l Levels of tree scattering

- Based on a balanced tree B_k^l
- Mirrored at the leaves
- Closed to ring at the root

$$|D_k^l| = \frac{(k+1)k^l-(k+1)}{k-1}$$

$$|D_k^{l+1}| = |D_k^l| + k^l + k^{l+1}$$

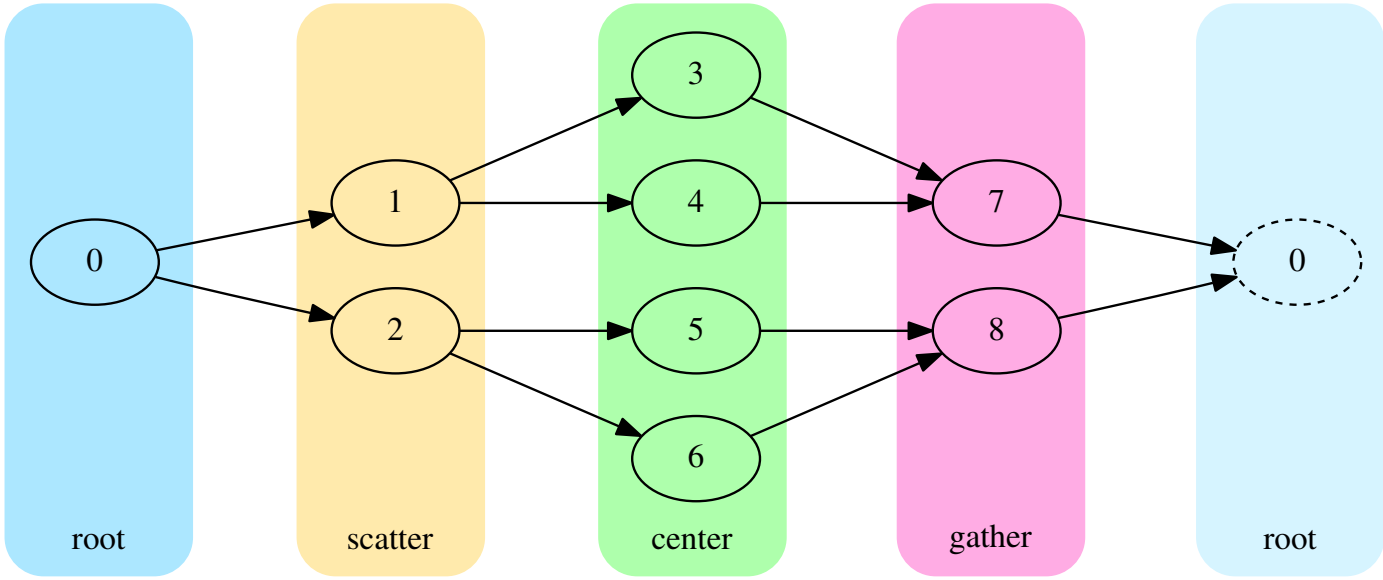
D_2^2 - diamond ring with 9 nodes



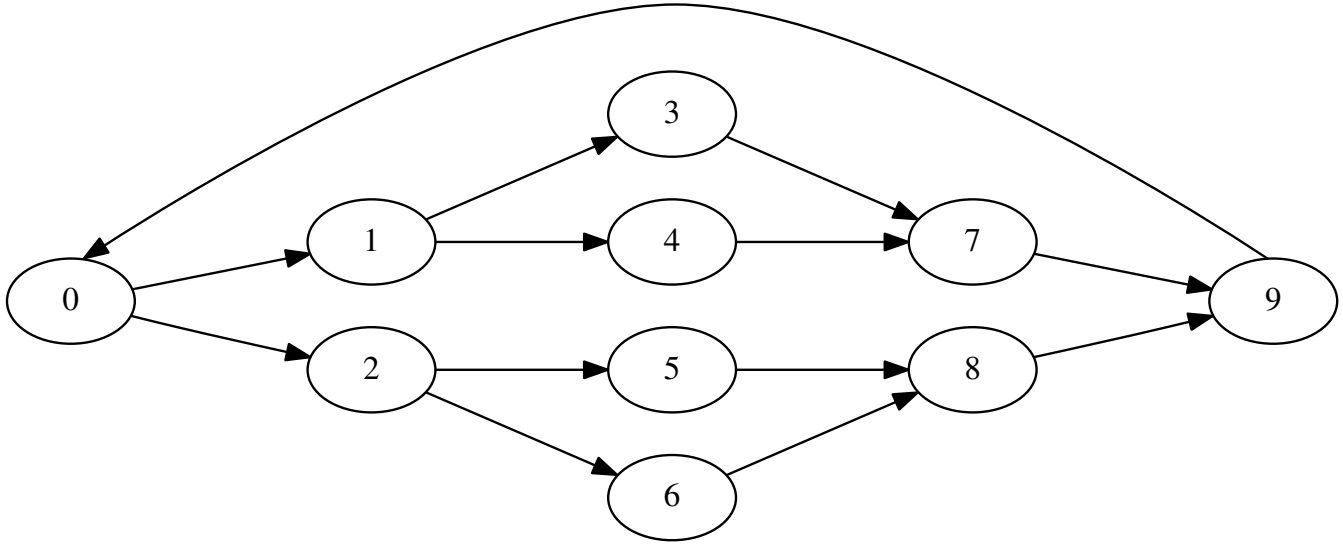
THE PERFECT DIAMOND RING

Brandenburg University of Technology Cottbus - Senftenberg

D_2^2 - diamond ring with 9 nodes



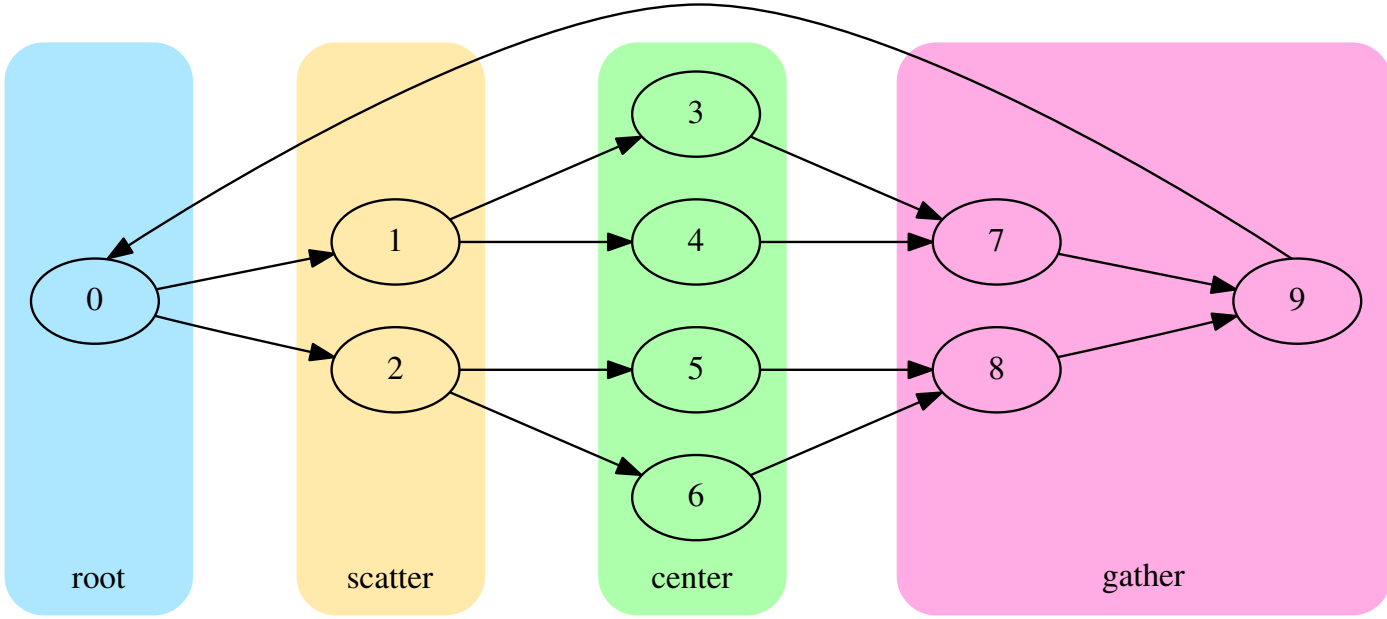
D_2^2 - diamond ring with 9 nodes +1 (no bottleneck version)



THE PERFECT DIAMOND RING

Brandenburg University of Technology Cottbus - Senftenberg

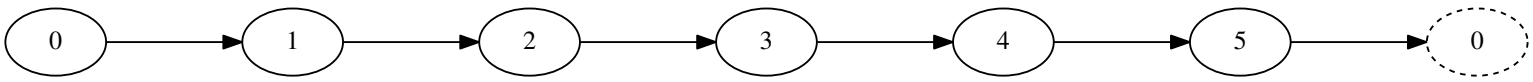
D_2^2 - diamond ring with 9 nodes +1 (no bottleneck version)



SOME MORE EXAMPLES

Brandenburg
University of Technology
Cottbus - Senftenberg

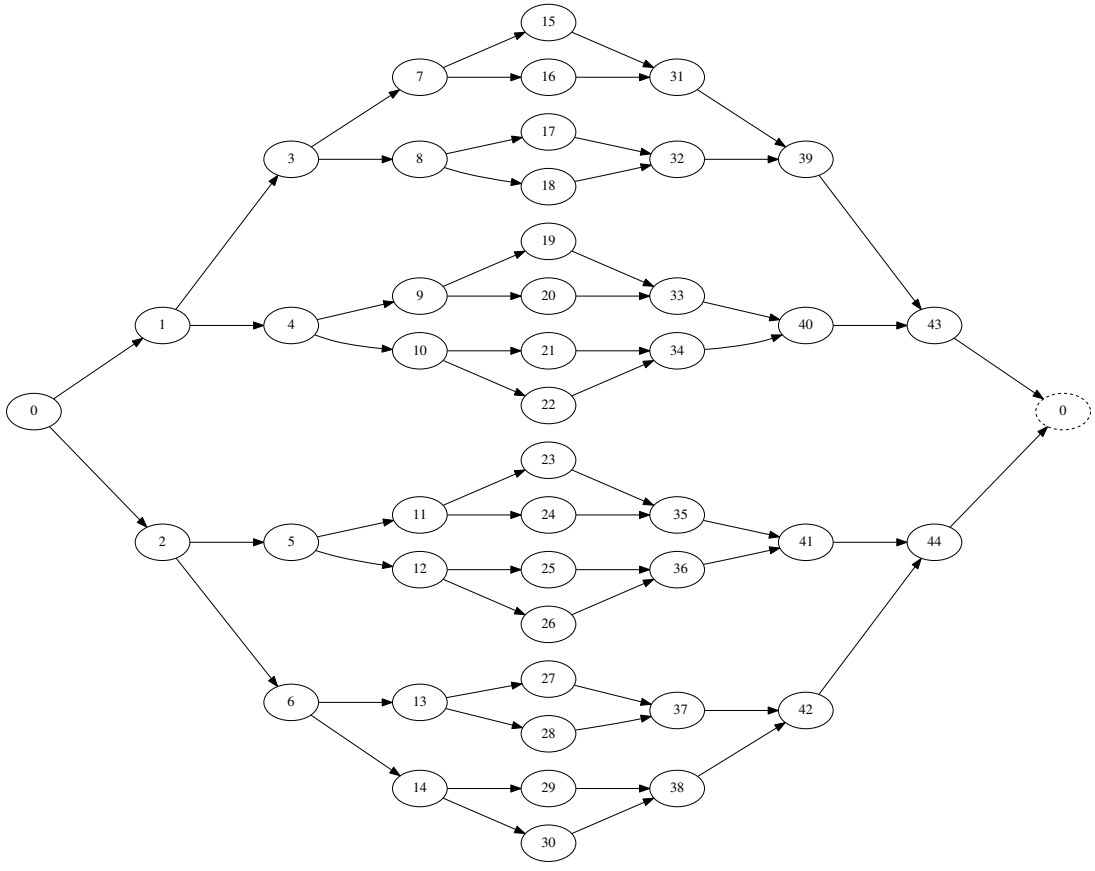
D_1^3 - diamond ring with 6 nodes



SOME MORE EXAMPLES

Brandenburg University of Technology Cottbus - Senftenberg

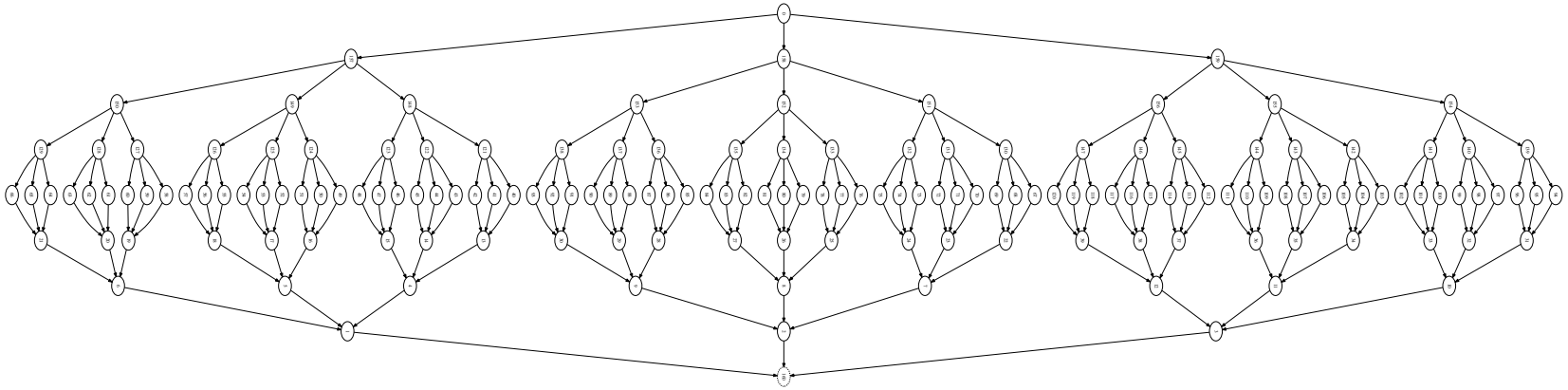
D_2^4 - diamond ring with 45 nodes



SOME MORE EXAMPLES

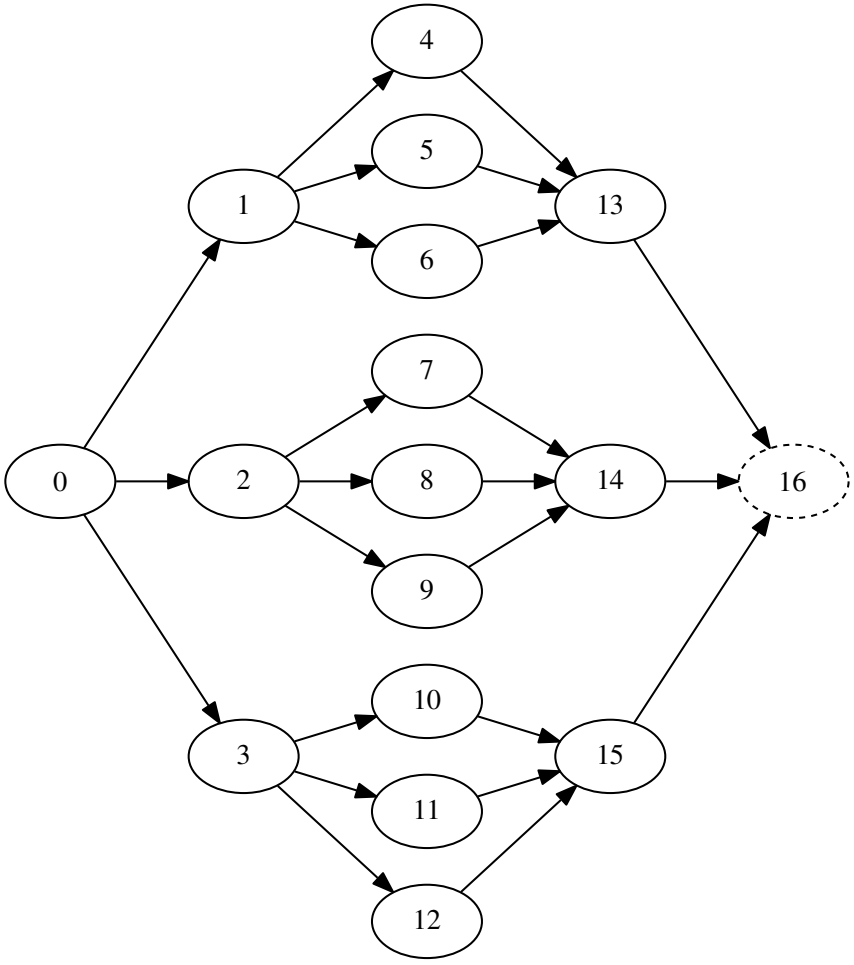
Brandenburg
University of Technology
Cottbus - Senftenberg

D_3^4 - diamond ring with 160 nodes



ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

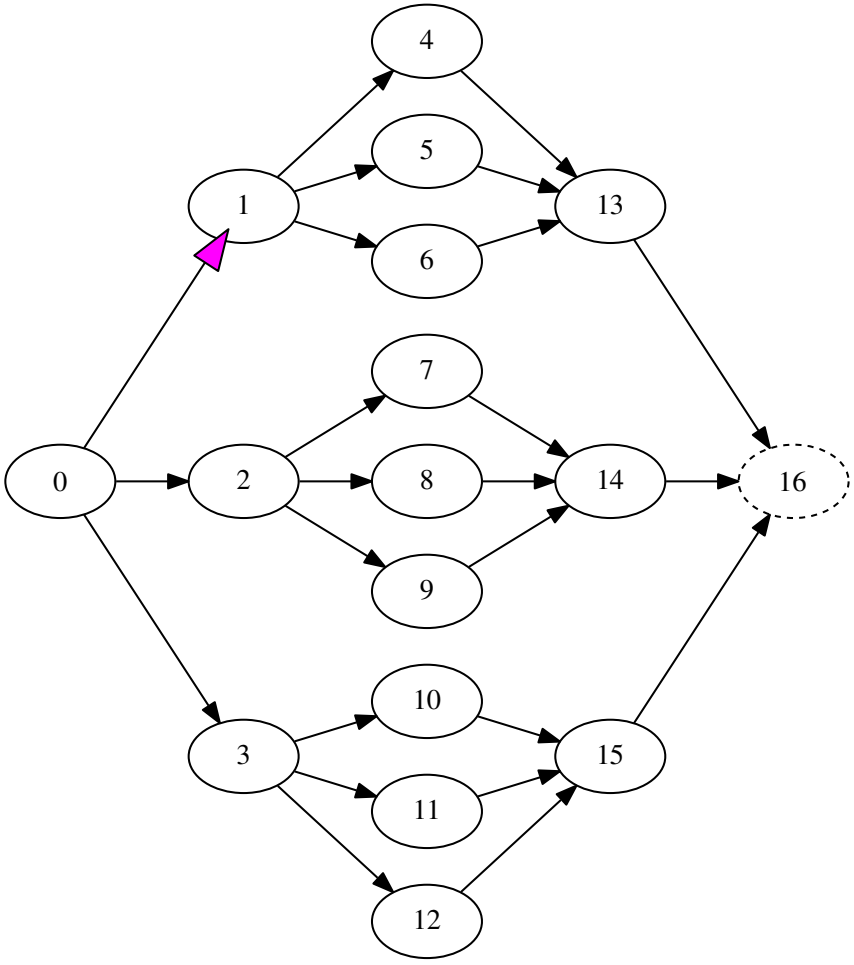
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

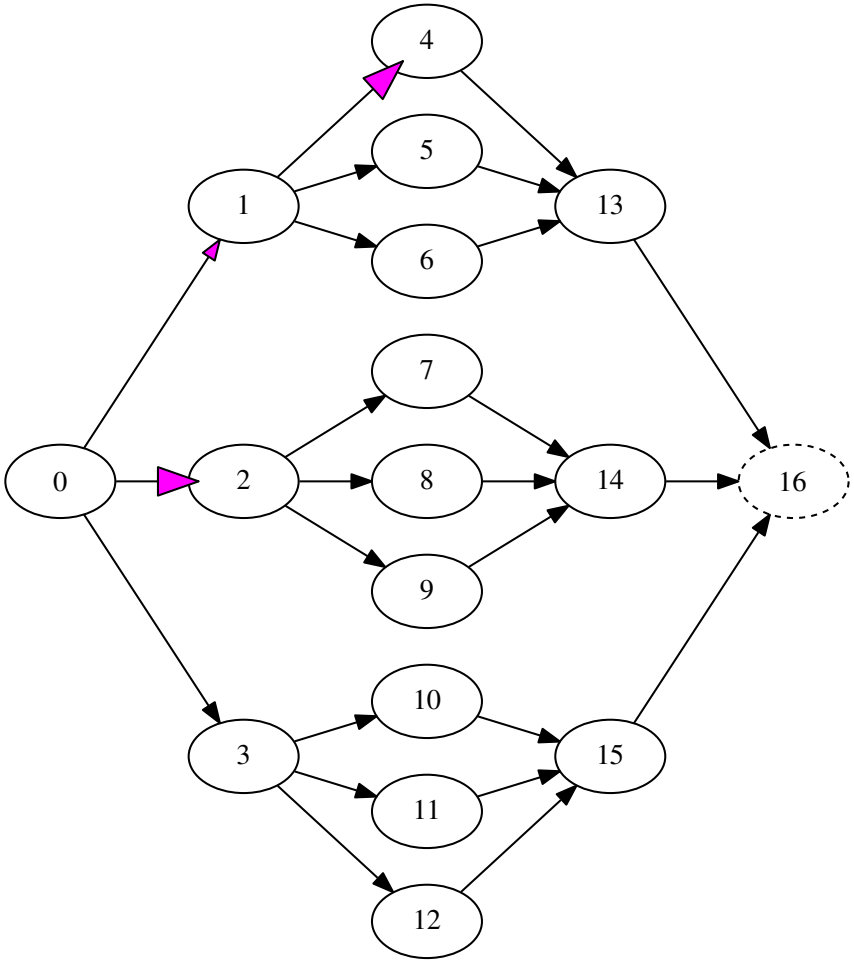
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

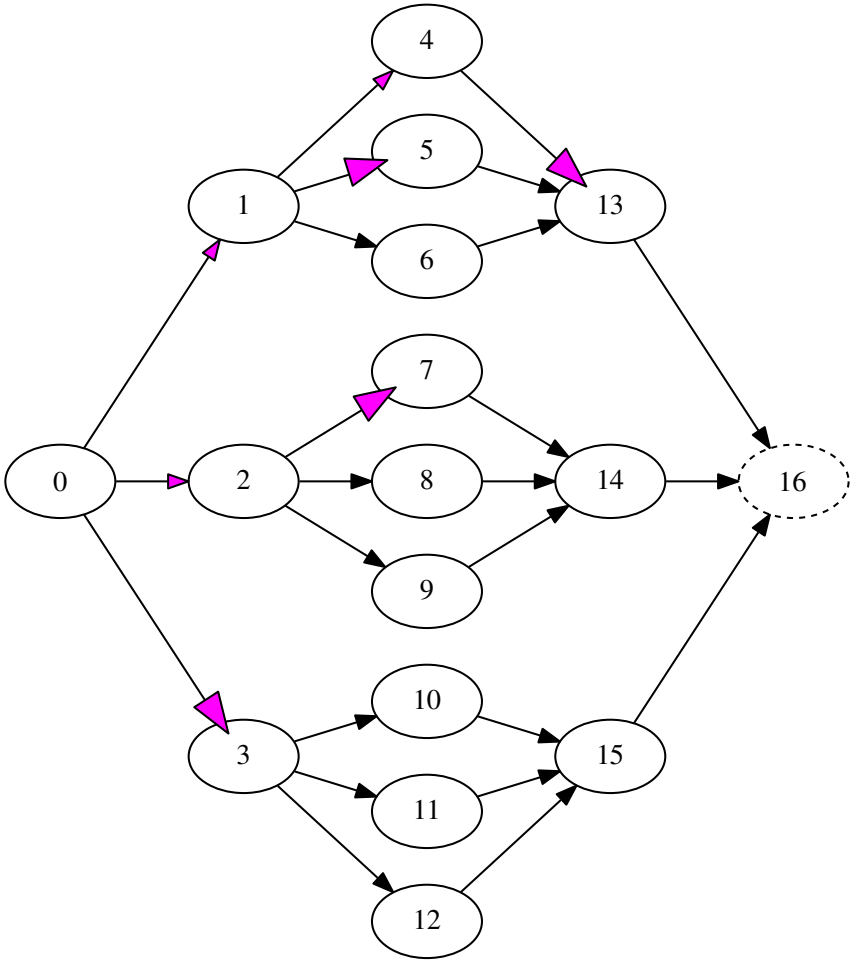
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

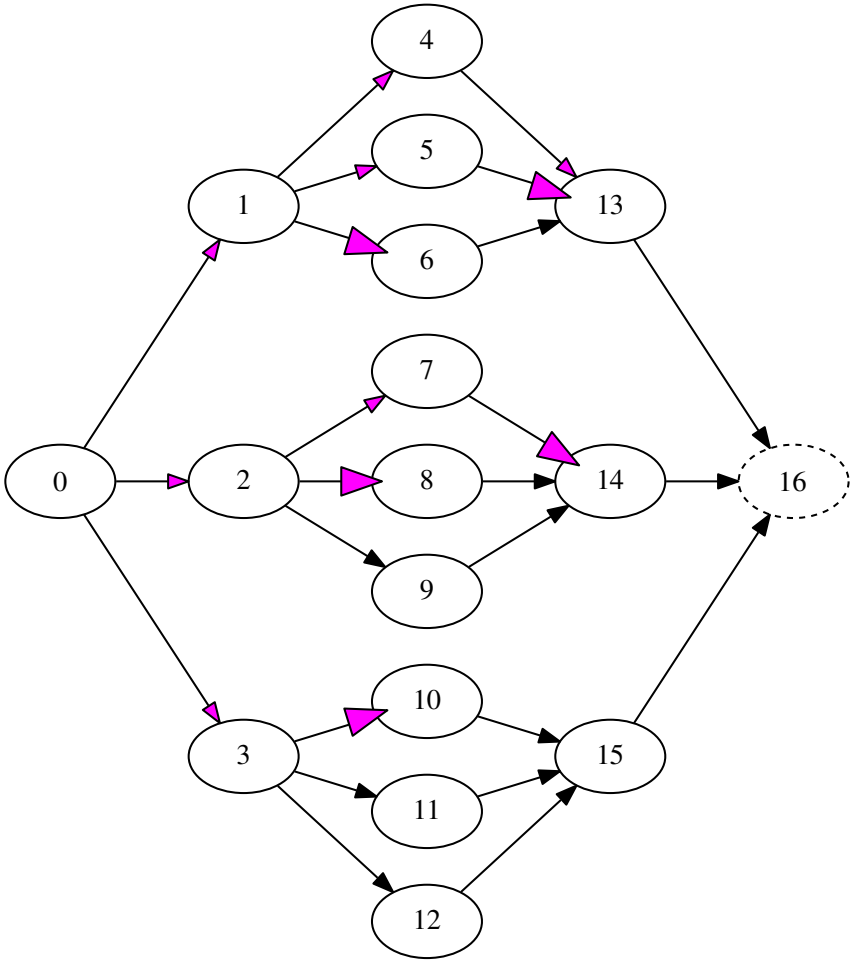
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg University of Technology Cottbus - Senftenberg



Throughput

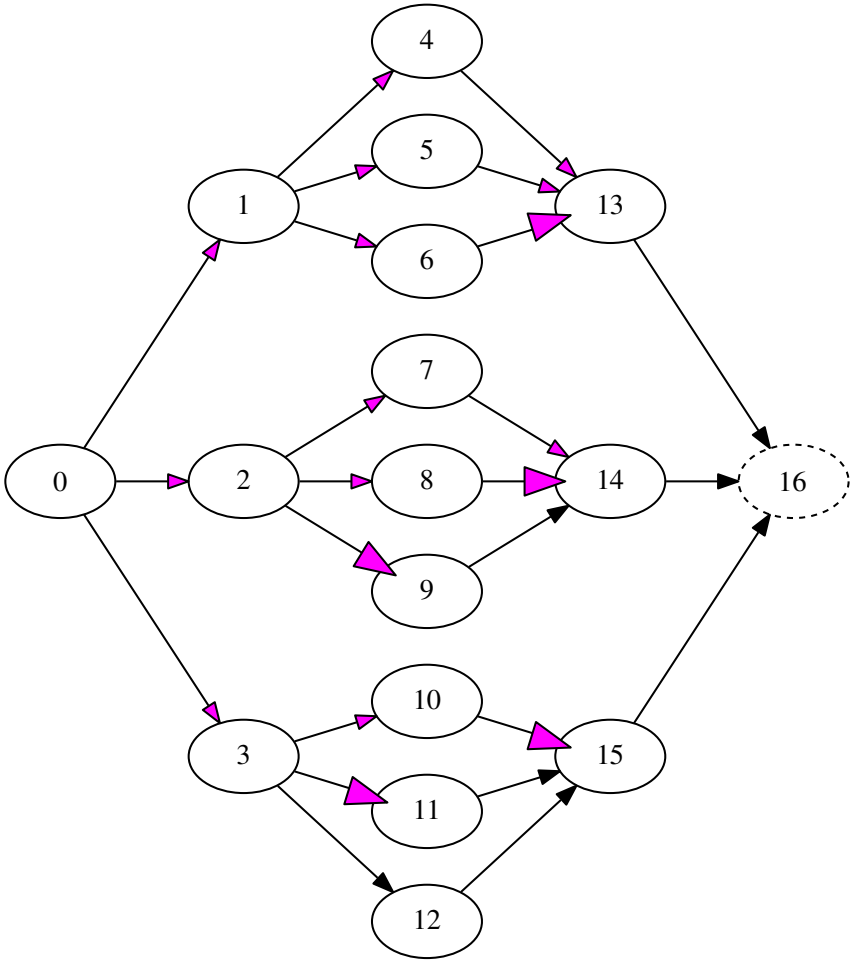
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

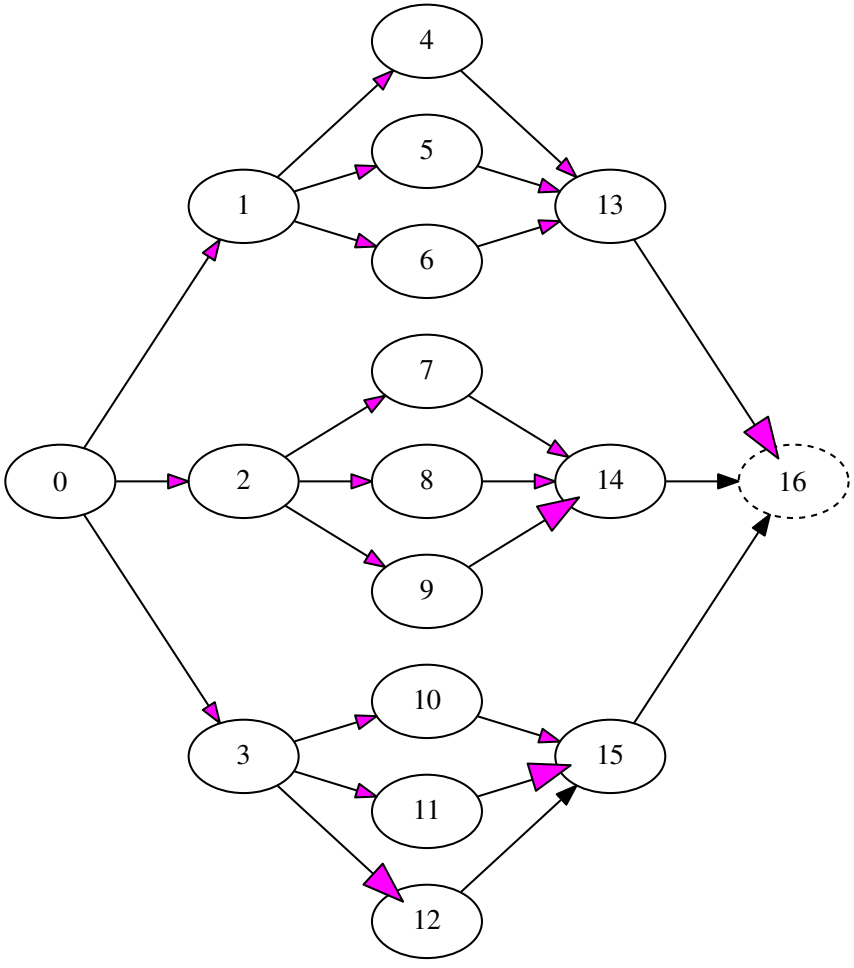
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

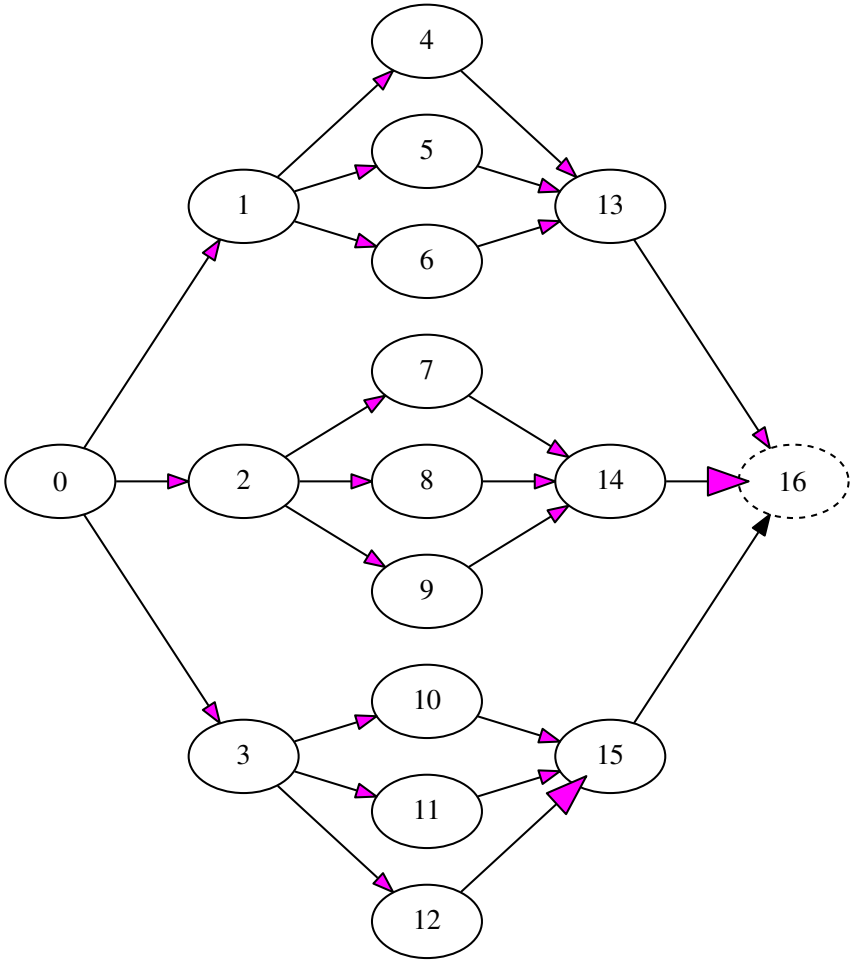
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

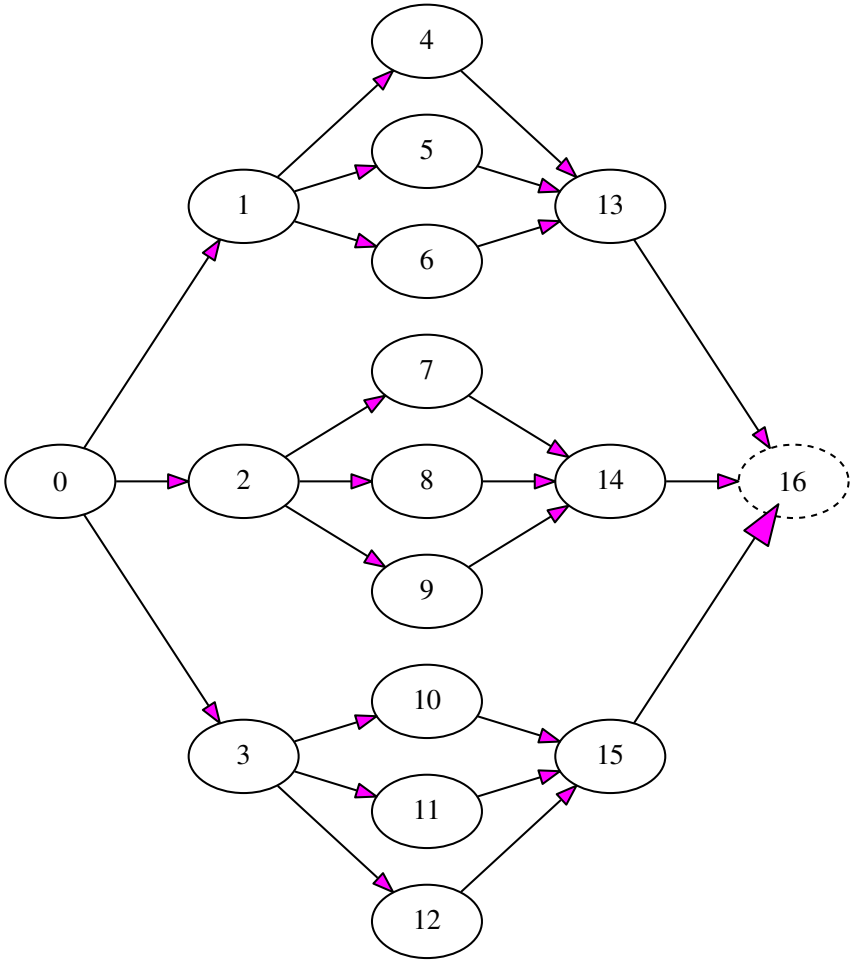
★★☆

Latency

★★★

ACKNOWLEDGED BROADCAST USING DIAMOND RINGS

Brandenburg
University of Technology
Cottbus - Senftenberg



Throughput

★★☆

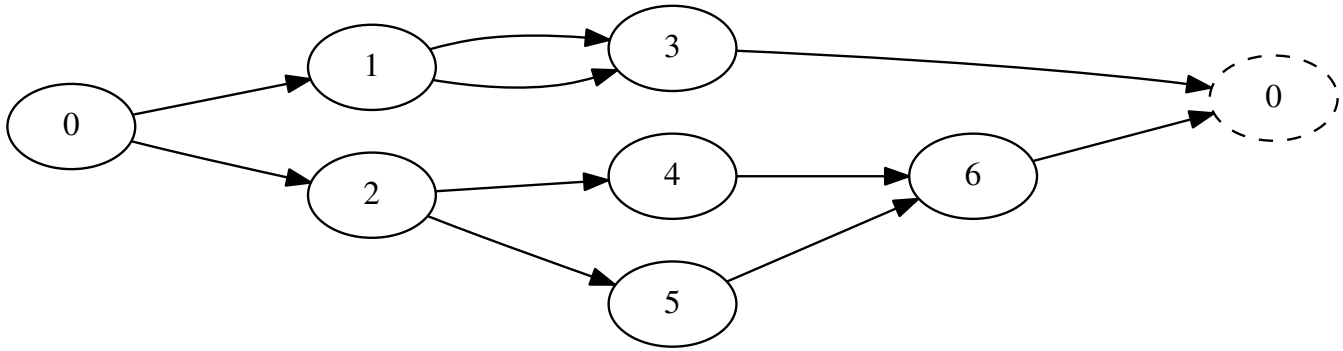
Latency

★★★

DEALING WITH ODD NODE COUNTS

Brandenburg University of Technology Cottbus - Senftenberg

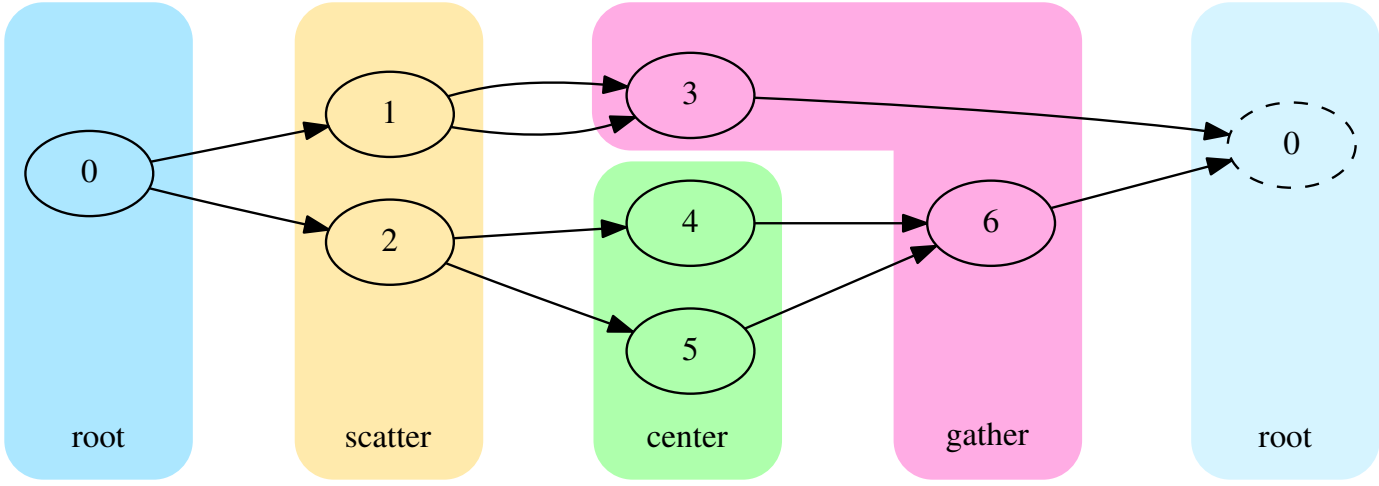
D_2^2 - diamond ring with 7 nodes (-2 nodes)



DEALING WITH ODD NODE COUNTS

Brandenburg University of Technology Cottbus - Senftenberg

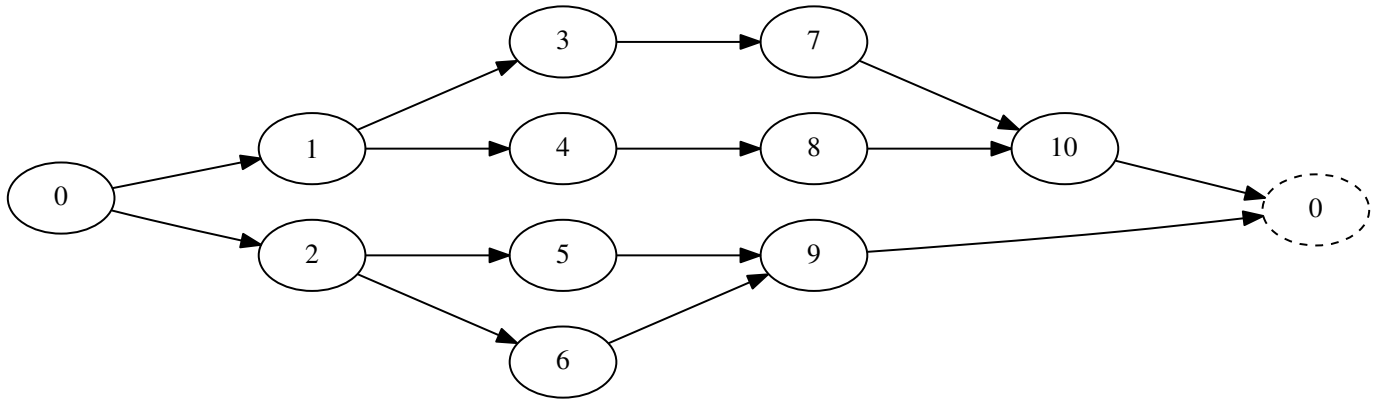
D_2^2 - diamond ring with 7 nodes (-2 nodes)



DEALING WITH ODD NODE COUNTS (2)

Brandenburg University of Technology Cottbus - Senftenberg

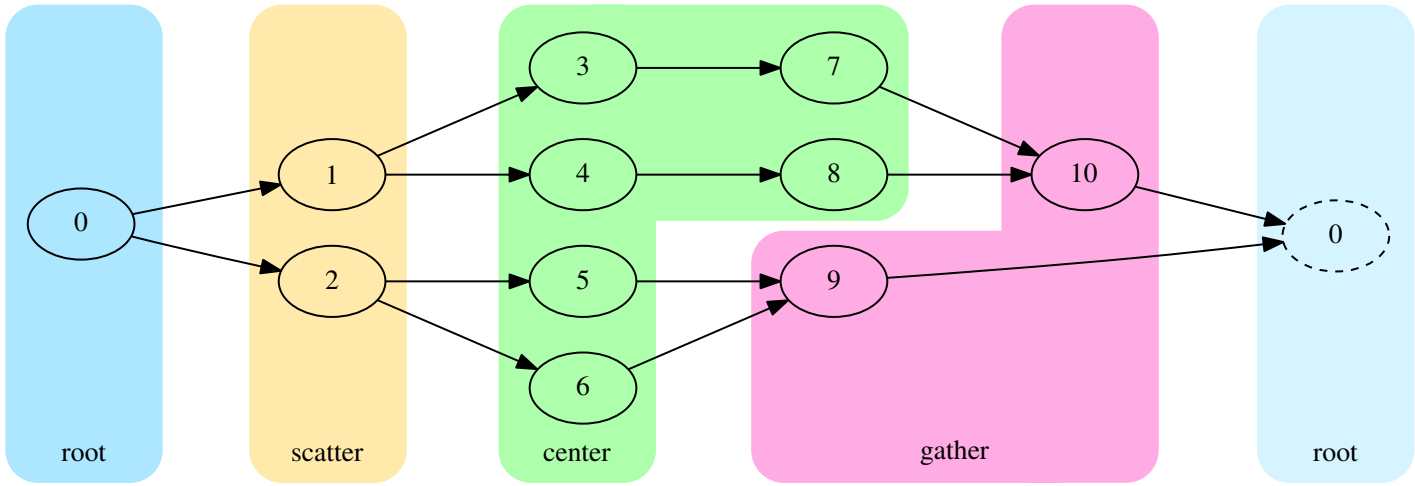
D_2^2 - diamond ring with 11 nodes (+2 nodes)



DEALING WITH ODD NODE COUNTS (2)

Brandenburg University of Technology Cottbus - Senftenberg

D_2^2 - diamond ring with 11 nodes (+2 nodes)



Latency and Throughput

Latency is reduced due to shorter longest path

Throughput is increased since nodes have less communication partners

Contention on the network is reduced due to less messages sent

	Balanced Tree	Diamond Ring	Ring
Longest Path	$2log_k(n)$	$2log_k(n)-2$	n
Max. Overhead	$2(k+1)$	$2k$	2
Messages sent	$2(n-1)$	$\frac{2k}{k+1}n$	n

Latency and Throughput

Latency is reduced due to shorter longest path

Throughput is increased since nodes have less communication partners

Contention on the network is reduced due to less messages sent

	Balanced Tree	Diamond Ring	Ring
Longest Path	$2log_k(n)$	$2log_k(n)-1$	n
Max. Overhead	$2(k+1)$	$k+1$	2
Messages sent	$2(n-1)$	$\frac{2k}{k+1}n+1$	n

OUTLINE

Brandenburg
University of Technology
Cottbus - Senftenberg

1. Throughput & Latency of Broadcast

2. The Diamond Ring Topology

3. **Evaluation**

EVALUATION OF DIAMOND RINGS

Brandenburg University of Technology Cottbus - Senftenberg

Hypothesis

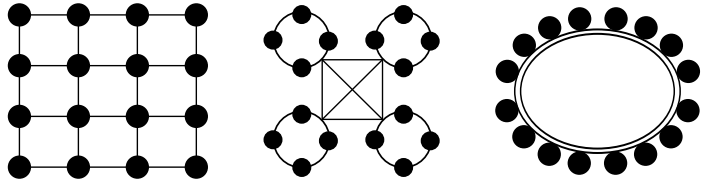
Acknowledged broadcasts using diamond rings should have...

1. lower latency,
2. higher throughput

... than balanced trees.

Benchmark Setup

- Custom active message framework
- Messages in shared memory
- Topologies: Balanced Tree (BT), Diamond Ring (DR), Sequenced Diamond Ring (SDR)
- Three different evaluation platforms



EVALUATION PLATFORMS

Brandenburg University of Technology Cottbus - Senftenberg

EZ-Chip Tilera TILE-Gx72

- 72 Cores @1Ghz (in-order)
- Low-Latency Mesh Network (UDN)

Intel Xeon E5 4640v2

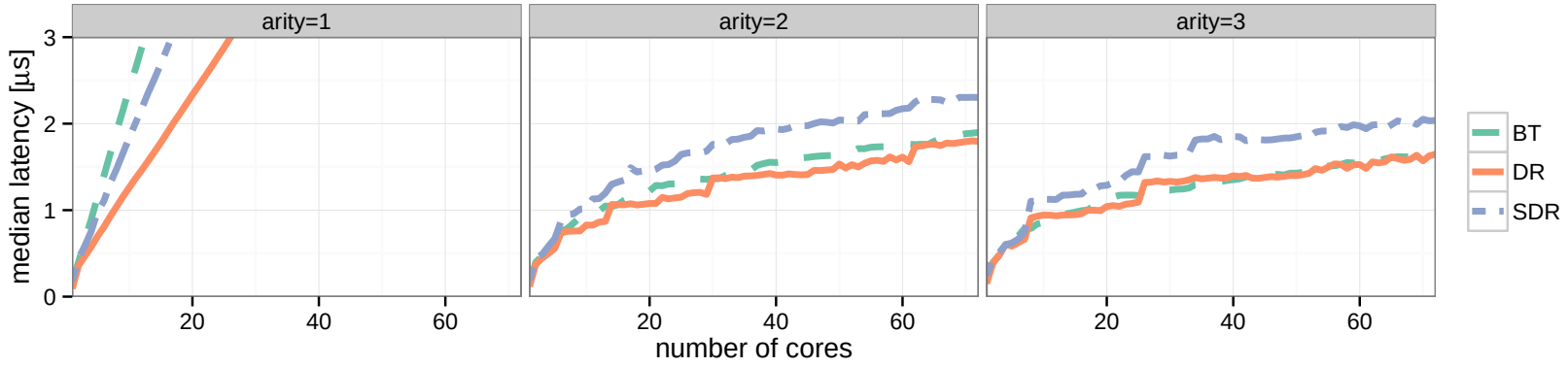
- 4 Sockets, 40 Cores, 80 HW-Threads @2.2Ghz (out-of-order)
- Slotted Rings, QPI between Sockets

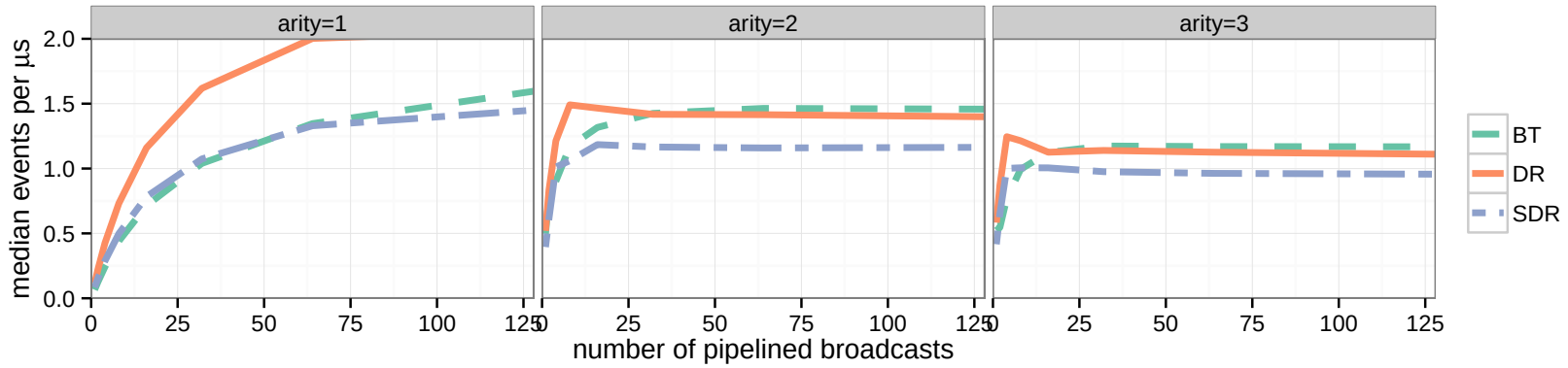
Intel Xeon Phi 5110P

- 60 Cores, 240 HW-Threads @1GHz (in-order)
- Slotted Ring Network

EZ-CHIP TILERA TILE-GX72

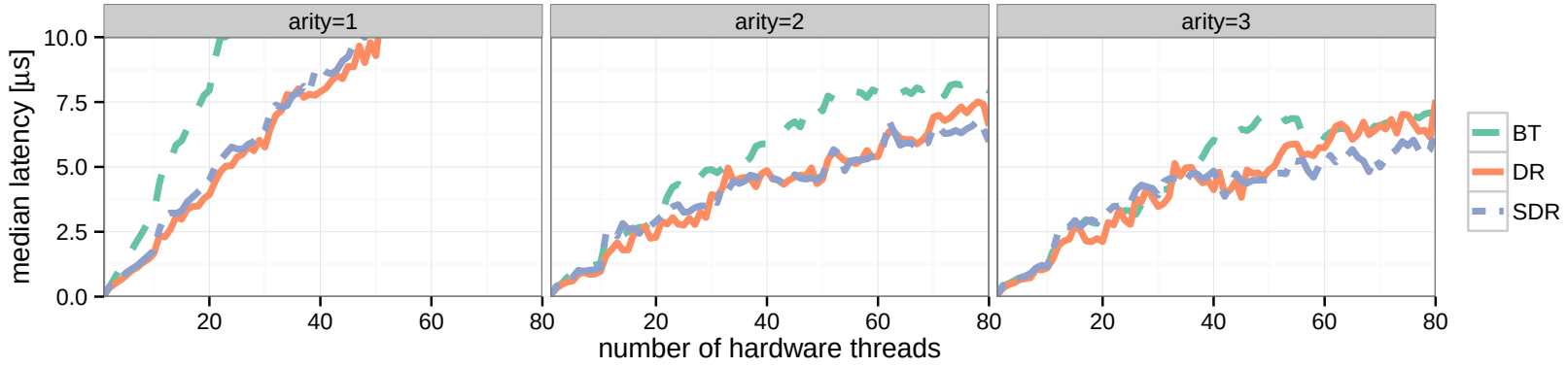
Brandenburg University of Technology Cottbus - Senftenberg

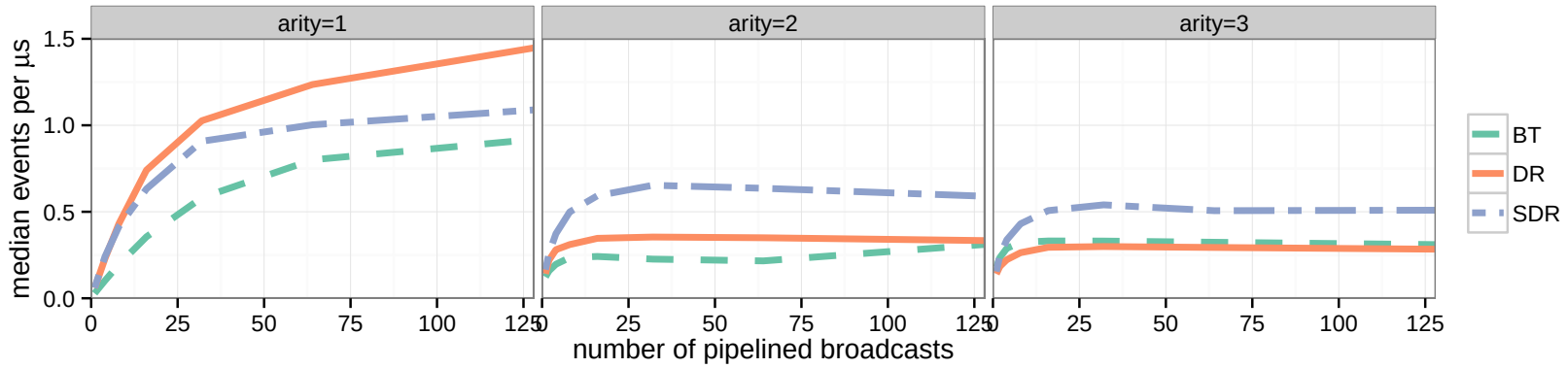




INTEL XEON 4640V2

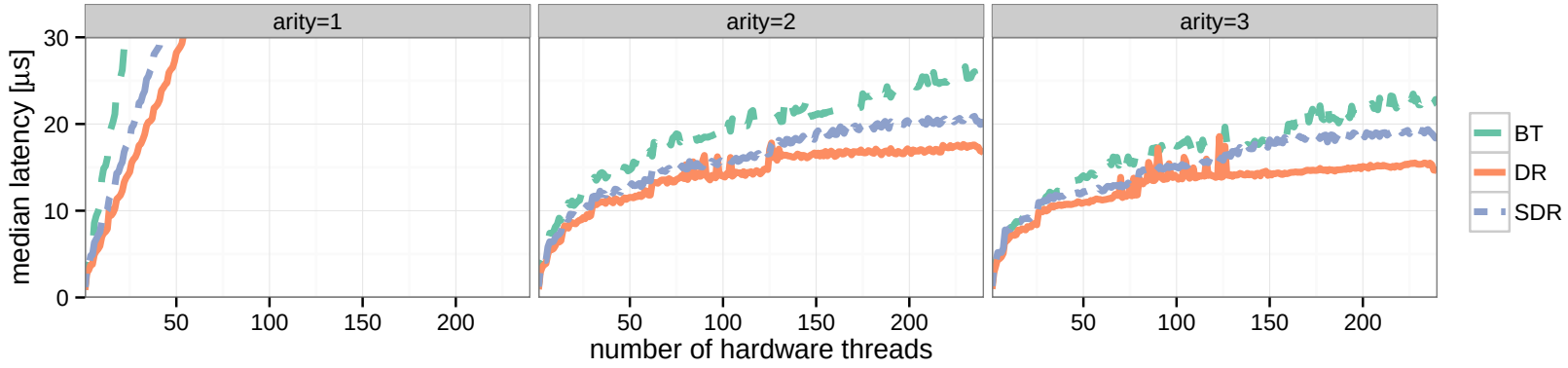
Brandenburg University of Technology Cottbus - Senftenberg

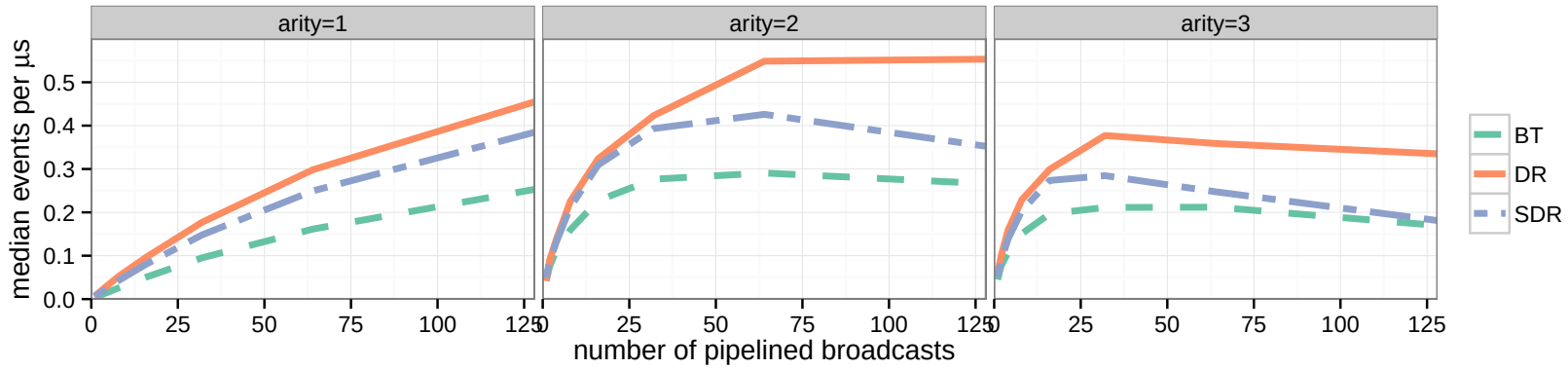




INTEL XEON PHI 5110P

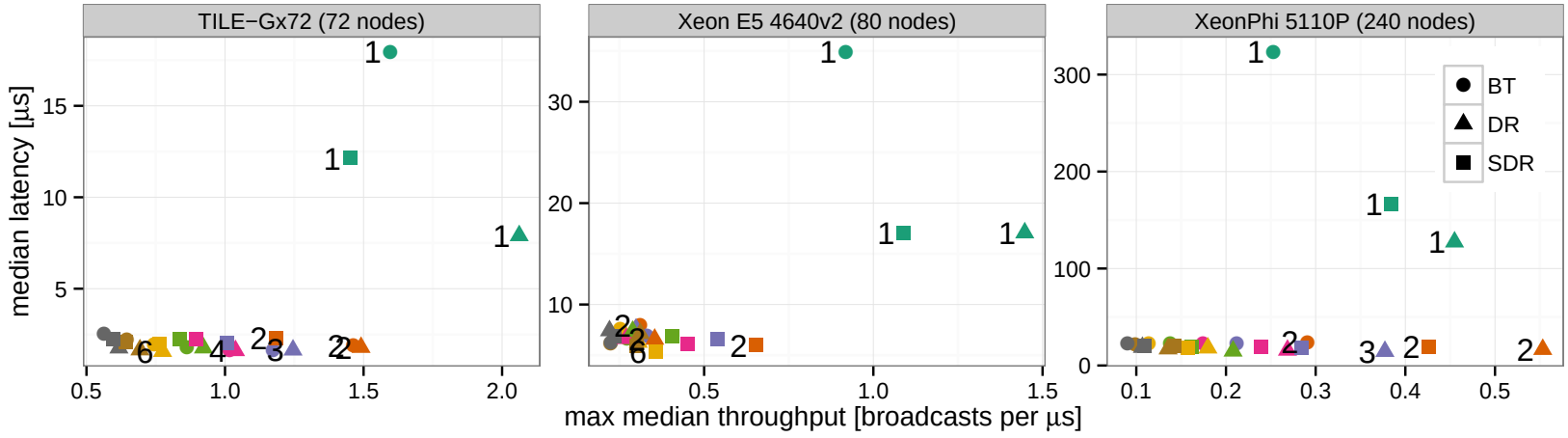
Brandenburg University of Technology Cottbus - Senftenberg





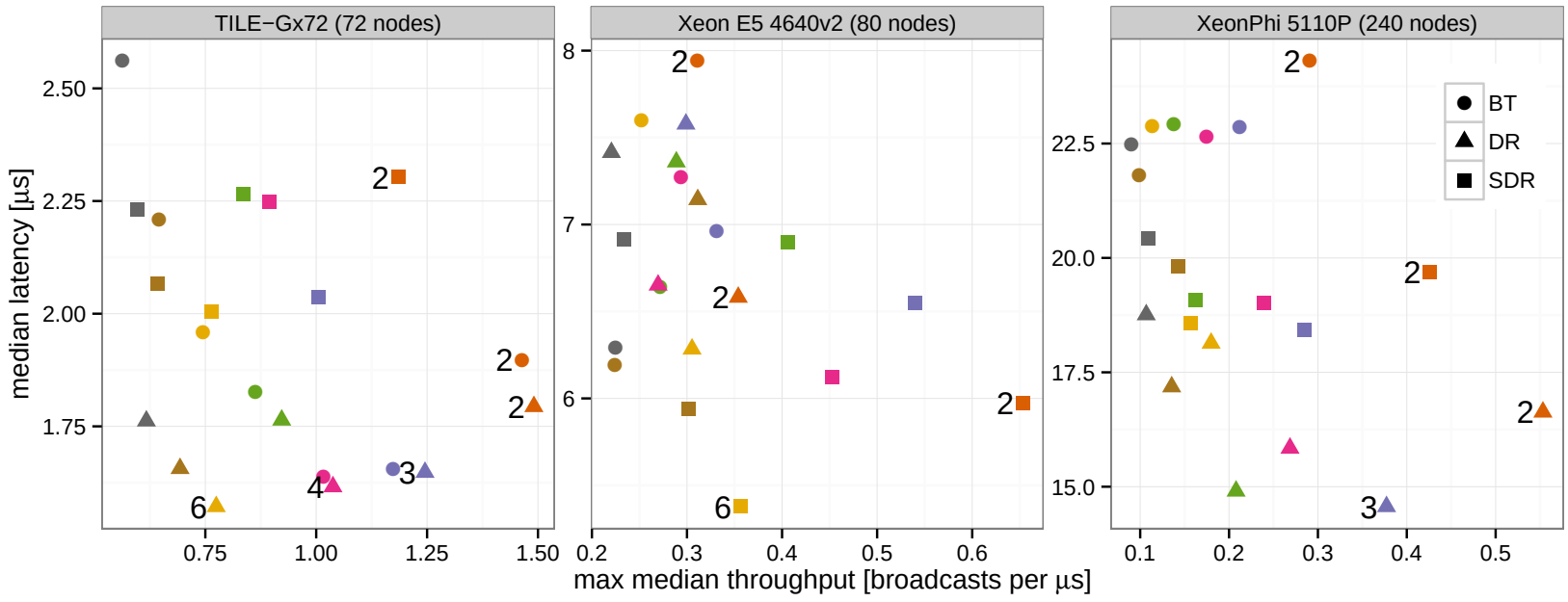
RESULTS OVERVIEW

Brandenburg University of Technology Cottbus - Senftenberg



RESULTS OVERVIEW

Brandenburg University of Technology Cottbus - Senftenberg



SUMMARY

Brandenburg University of Technology Cottbus - Senftenberg

Acknowledged Event Propagation

- is very important in consistency management.

Throughput and Latency

- require a trade-off.

Diamond Rings

- offer a better trade-off than balanced trees.
- are acknowledged broadcast's best friend.

Thank you for your attention!

- Questions?

This work was supported by the German Research Foundation (DFG) under grant no. NO 625/7-1 and SCHR 603/10-1