Generalized DCell Structure for Load-Balanced Data Center Networks

Markus Kliegl, Swarthmore College
Jason Lee, Duke University
Jun Li, Fudan University
Xinchao Zhang, Shanghai Jiao Tong University
Chuanxiong Guo, Microsoft Research Asia
David Rincón, Universitat Politècnica de Catalunya

March 15, 2010

Data Center Networks

Data Center Networks

- Data centers provide popular online services such as email, IM, and search.
  - A form of Moore's law has been observed for data center size.
- There are several important traits that DCN's must have.

Data Center Networks

- Data centers provide popular online services such as email, IM, and search.
  - A form of Moore's law has been observed for data center size.
- There are several important traits that DCN's must have.
  - ★ scalability

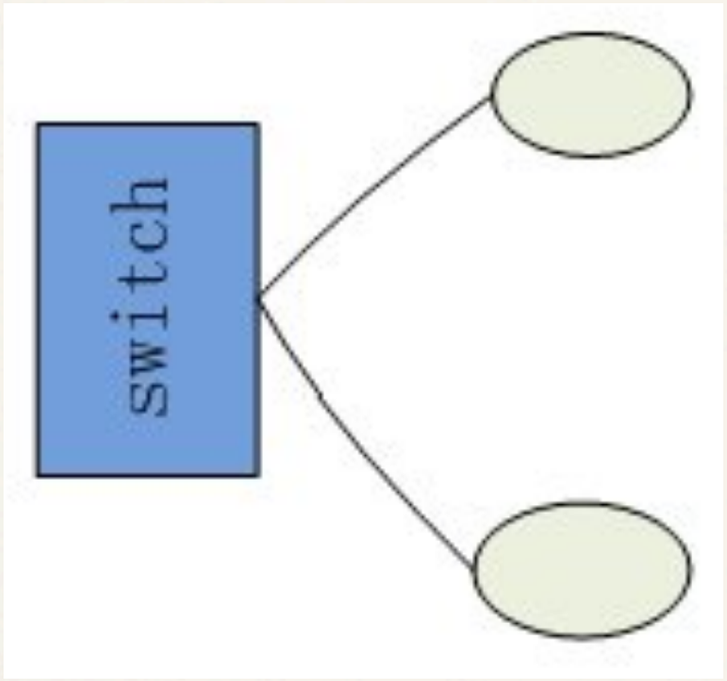
Data Center Networks

- Data centers provide popular online services such as email, IM, and search.
  - A form of Moore's law has been observed for data center size.
- There are several important traits that DCN's must have.
  - ★ scalability
  - ★ load-balance

Data Center Networks

- Data centers provide popular online services such as email, IM, and search.
  - A form of Moore's law has been observed for data center size.
- There are several important traits that DCN's must have.
  - ★ scalability
  - ★ load-balance
  - ★ cost-effectiveness

The DCell Network Structure [Guo 08]

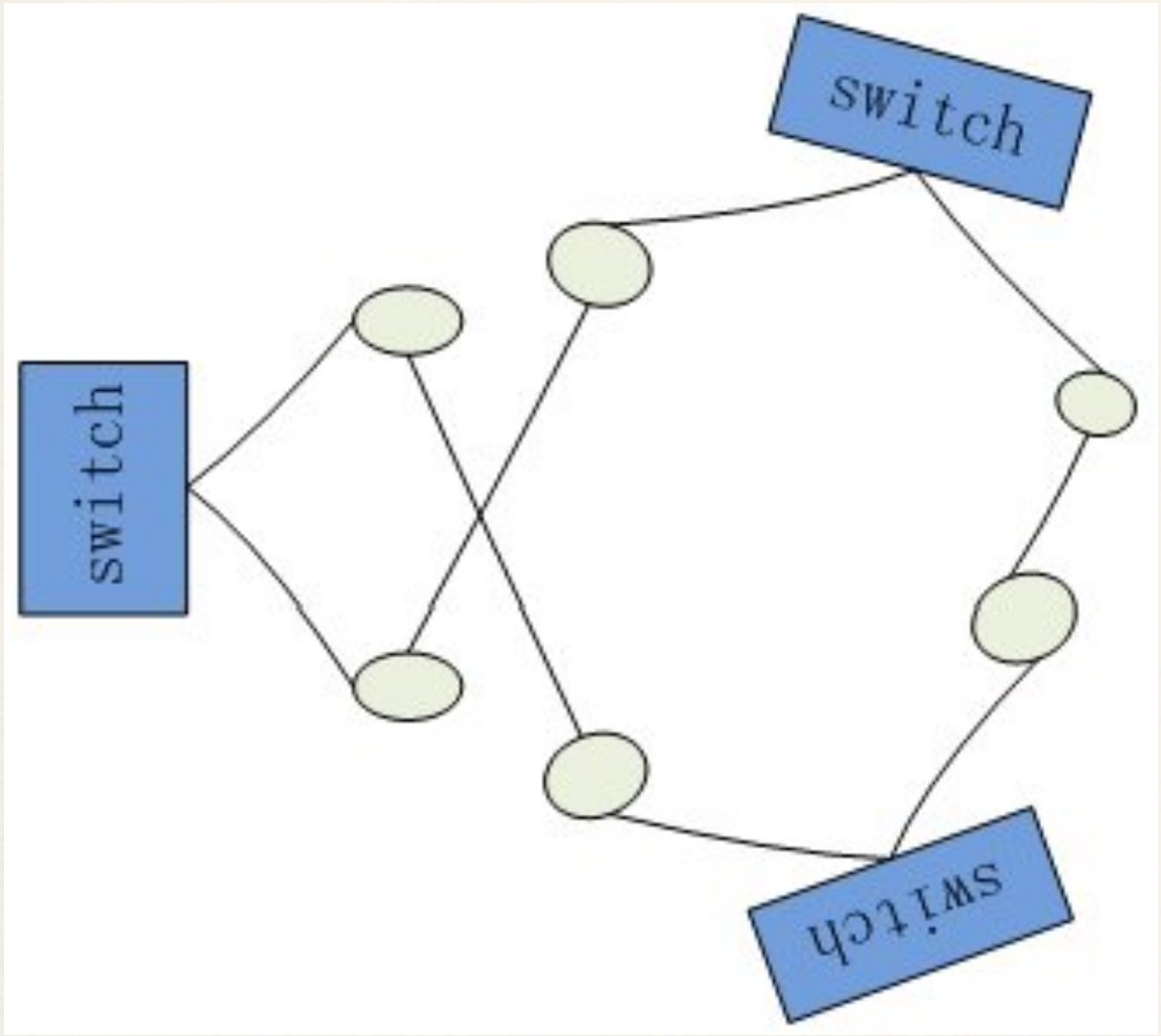


n=2, k=0/1/2

- High scalability
  - Doubly exponential growth
    - Recursive structure
- Commodity switches and servers
- Fault-tolerant routing
- **Unevenly** distributed traffic

[Guo 08] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu. DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers. In ACM SIGCOMM'08, 2008.

The DCell Network Structure [Guo 08]

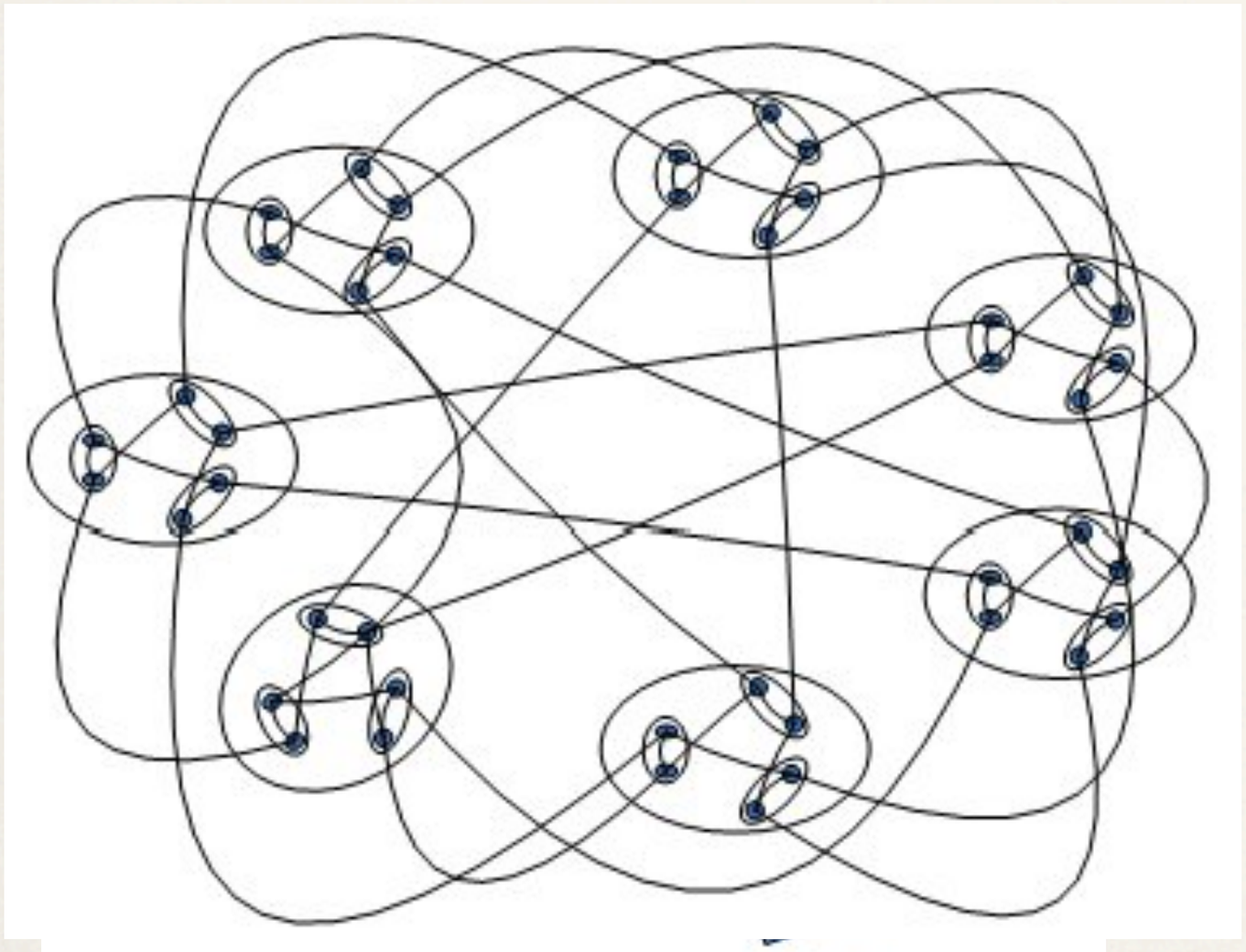


n=2, k=0/1/2

- High scalability
  - Doubly exponential growth
    - Recursive structure
- Commodity switches and servers
- Fault-tolerant routing
- **Unevenly** distributed traffic

[Guo 08] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu. DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers. In ACM SIGCOMM'08, 2008.

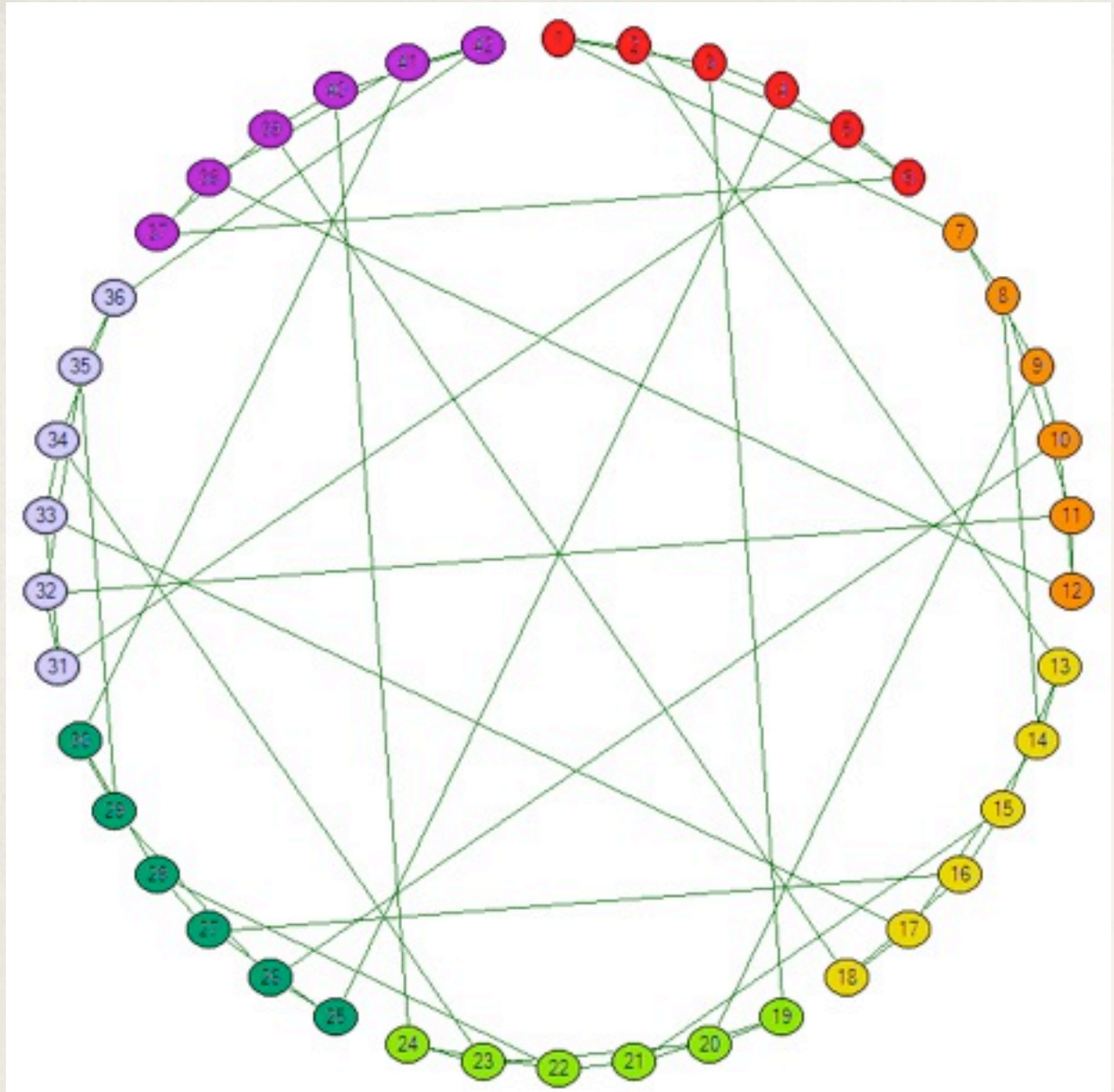
The DCell Network Structure [Guo 08]



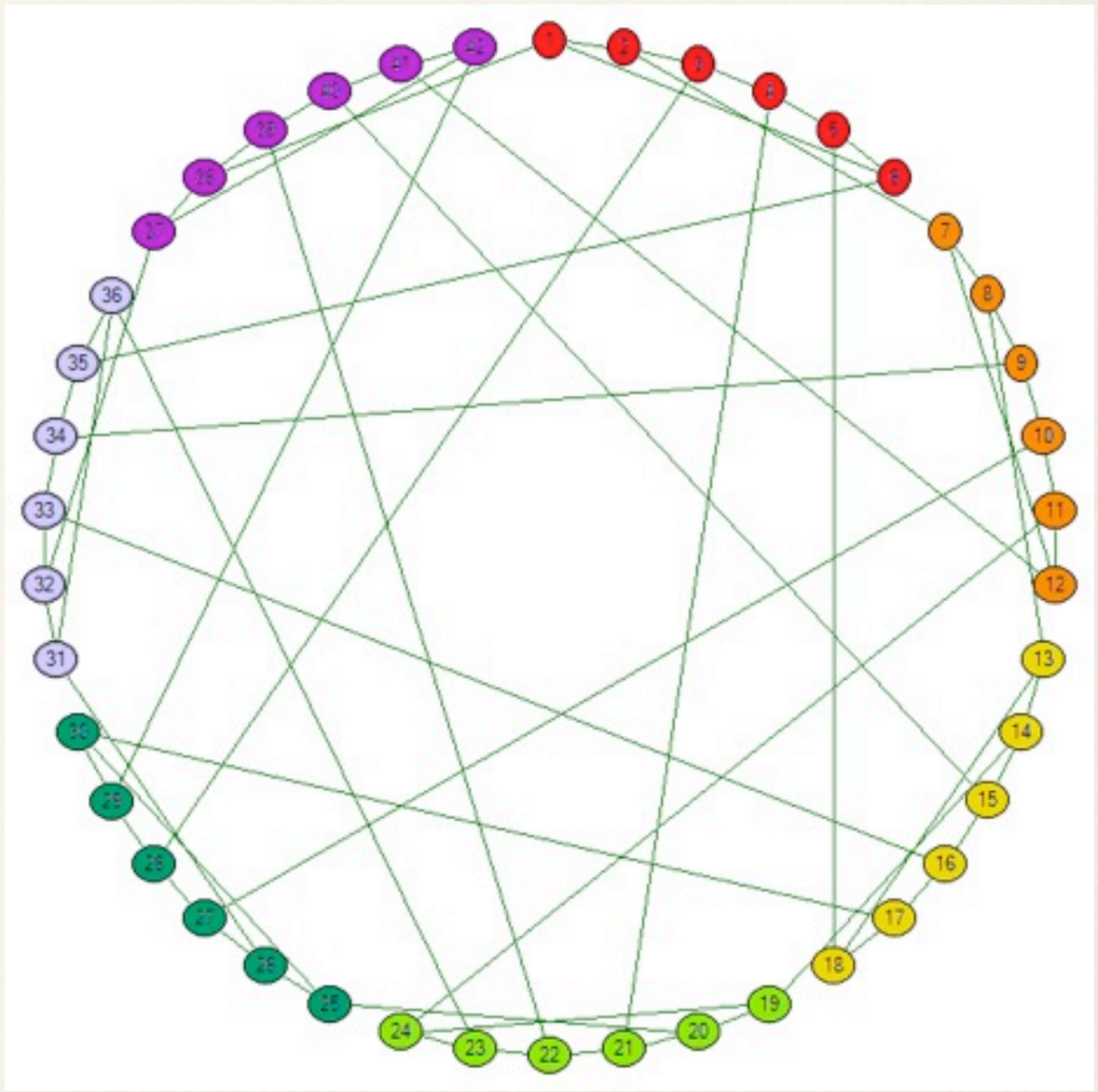
n=2, k=0/1/2

- High scalability
  - Doubly exponential growth
  - Recursive structure
- Commodity switches and servers
- Fault-tolerant routing
- **Unevenly** distributed traffic

[Guo 08] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu. DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers. In ACM SIGCOMM'08, 2008.

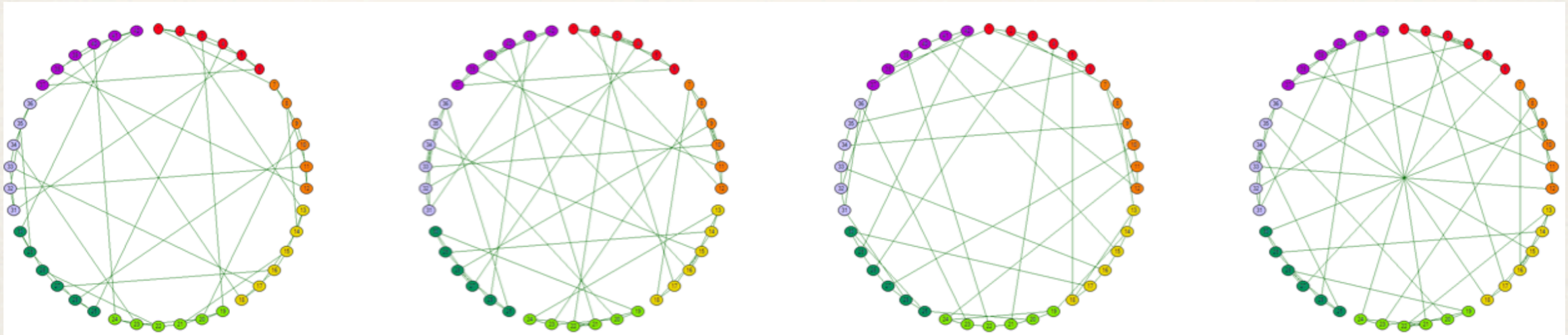


$$\alpha. \quad [a,b] \leftrightarrow \begin{cases} [b+1,a] & \text{if } a \leq b, \\ [b,a-1] & \text{if } a > b. \end{cases}$$



$$\gamma. \quad [a,b] \leftrightarrow \begin{cases} [a+b \pmod{t_{k-1}+1}, b-1] & \text{if } b \text{ is odd}, \\ [a-(b+1) \pmod{t_{k-1}+1}, b+1] & \text{if } b \text{ is even}. \end{cases}$$

Our Contribution: Generalized DCell Structure



(a) α-DCell (b) β-DCell (c) γ-DCell (d) δ-DCell

- Why generalized DCell?
  - Smaller diameter
  - Faster routing, better load balancing
  - Very different symmetries

Graph Properties

Graph Properties

- Scalability: $t_k = t_{k-1}(t_{k-1}+1)$

  We find $t_k = \lfloor c^{2^k} \rfloor$, where

$$c = \left(n + \frac{1}{2}\right) \prod_{i=0}^{\infty} \left(1 + \frac{1}{4\left(t_i + \frac{1}{2}\right)^2}\right)^{1/2^{i+1}}$$

Graph Properties

- Scalability: $t_k = t_{k-1}(t_{k-1}+1)$

  We find $t_k = \lfloor c^{2^k} \rfloor$, where

$$c = \left(n + \frac{1}{2}\right) \prod_{i=0}^{\infty} \left(1 + \frac{1}{4\left(t_i + \frac{1}{2}\right)^2}\right)^{1/2^{i+1}}$$

- Diameter

  For $k \leq n^2 + 1$, the diameter D is bounded below by $D \geq 2^{k-1}$.

  In [Guo 08], it is shown that $D \leq 2^{k+1} - 1$.

So, for realistic values of *k* and *n*, these bounds determine the diameter to within a factor of 4.

Graph Properties

- Scalability: $t_k = t_{k-1}(t_{k-1}+1)$

  We find $t_k = \lfloor c^{2^k} \rfloor$, where

$$c = \left(n + \frac{1}{2}\right) \prod_{i=0}^{\infty} \left(1 + \frac{1}{4\left(t_i + \frac{1}{2}\right)^2}\right)^{1/2^{i+1}}$$

- Diameter

  For $k \leq n^2 + 1$, the diameter D is bounded below by $D \geq 2^{k-1}$.

  In [Guo 08], it is shown that $D \leq 2^{k+1} - 1$.

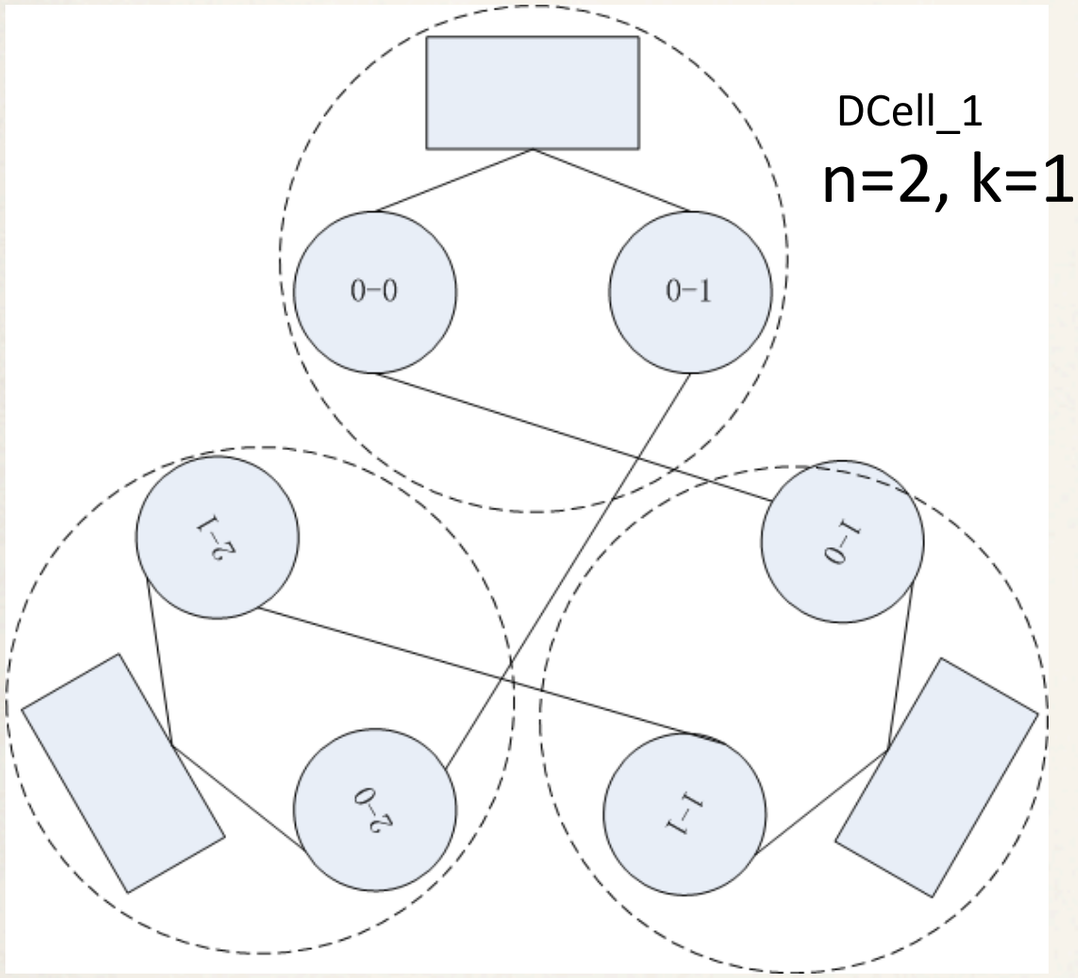
  So, for realistic values of *k* and *n*, these bounds determine the diameter to within a factor of 4.

- Symmetry

  Only one nontrivial automorphism for α-DCell (when *n* is small)

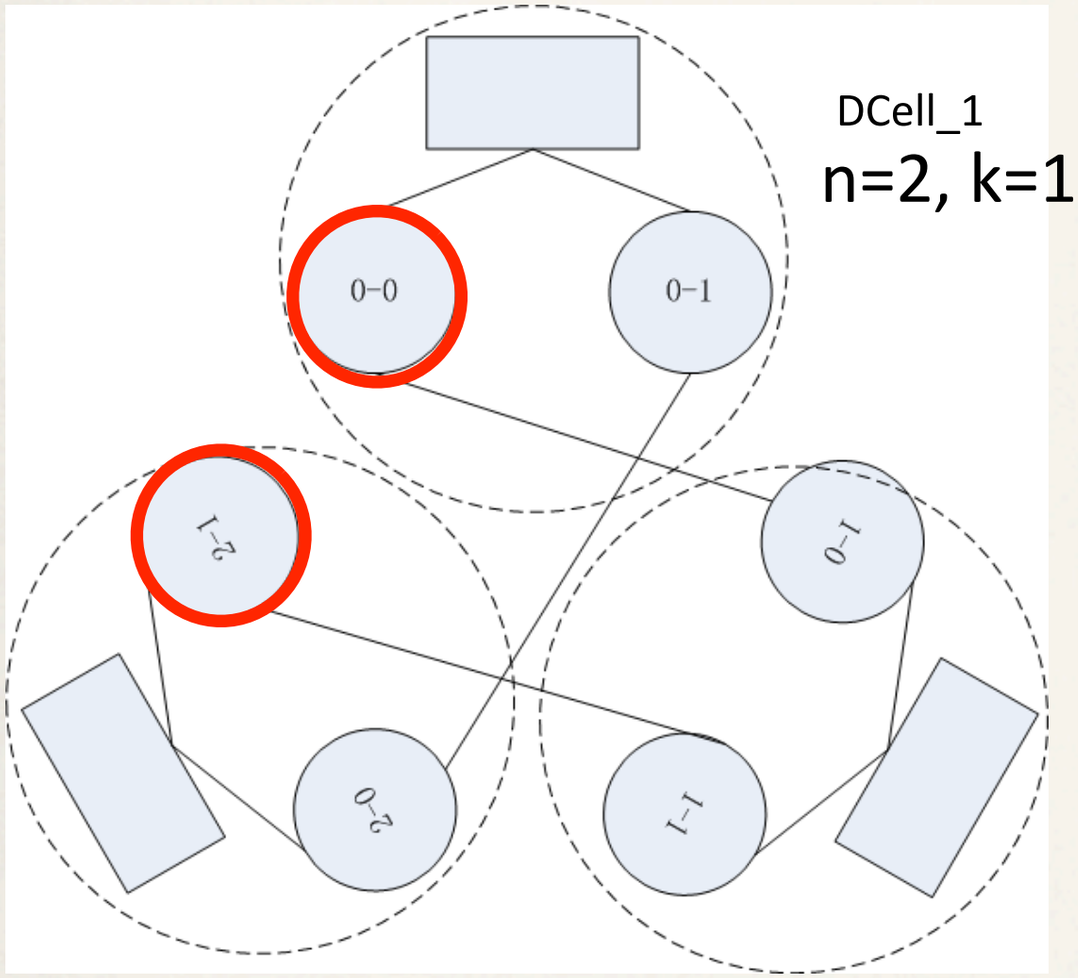
  For other connection rules, there is a cyclic subgroup of the automorphism group of order $t_{k-1} + 1$

DCellRouting [Guo 08] : Flow Distribution



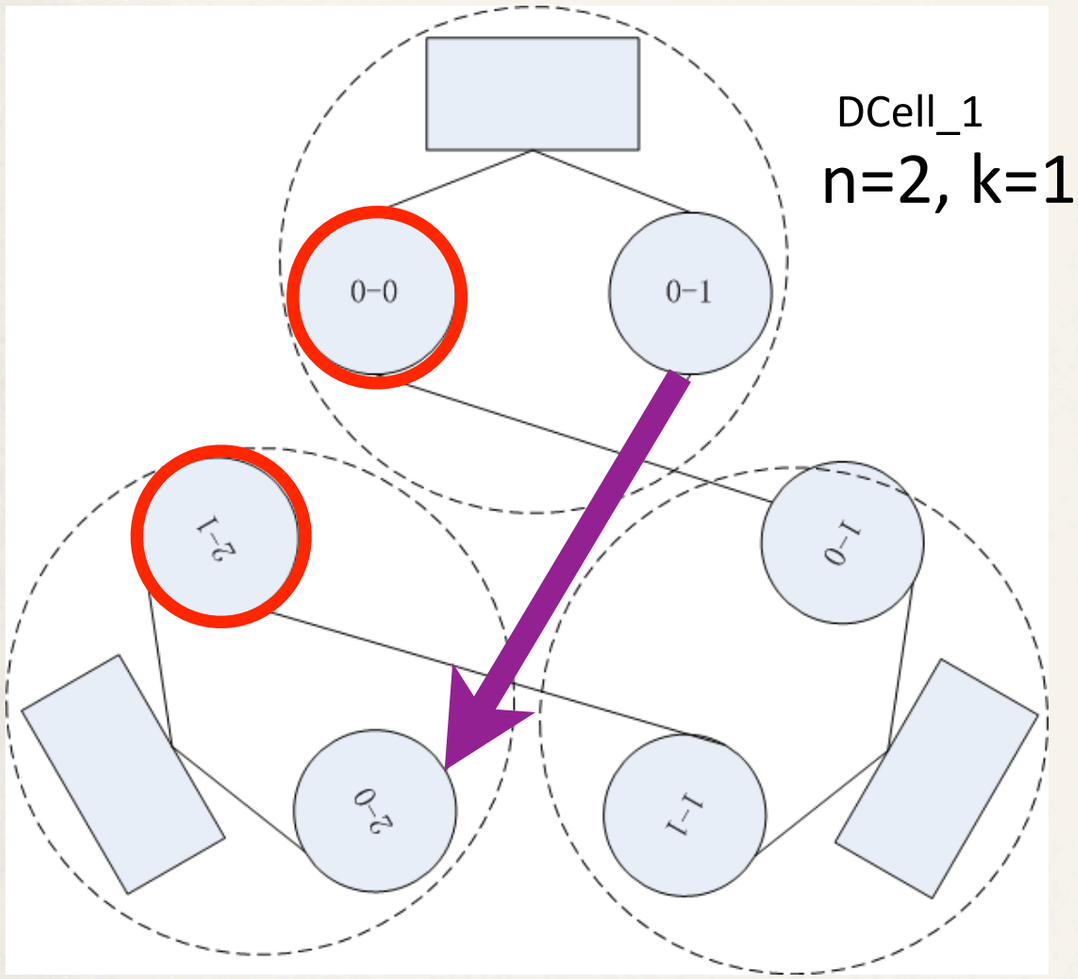
- Let F_L denote the number of flows carried by a L-level link, in all-to-all communication using the DCellRouting algorithm.

DCellRouting [Guo 08] : Flow Distribution



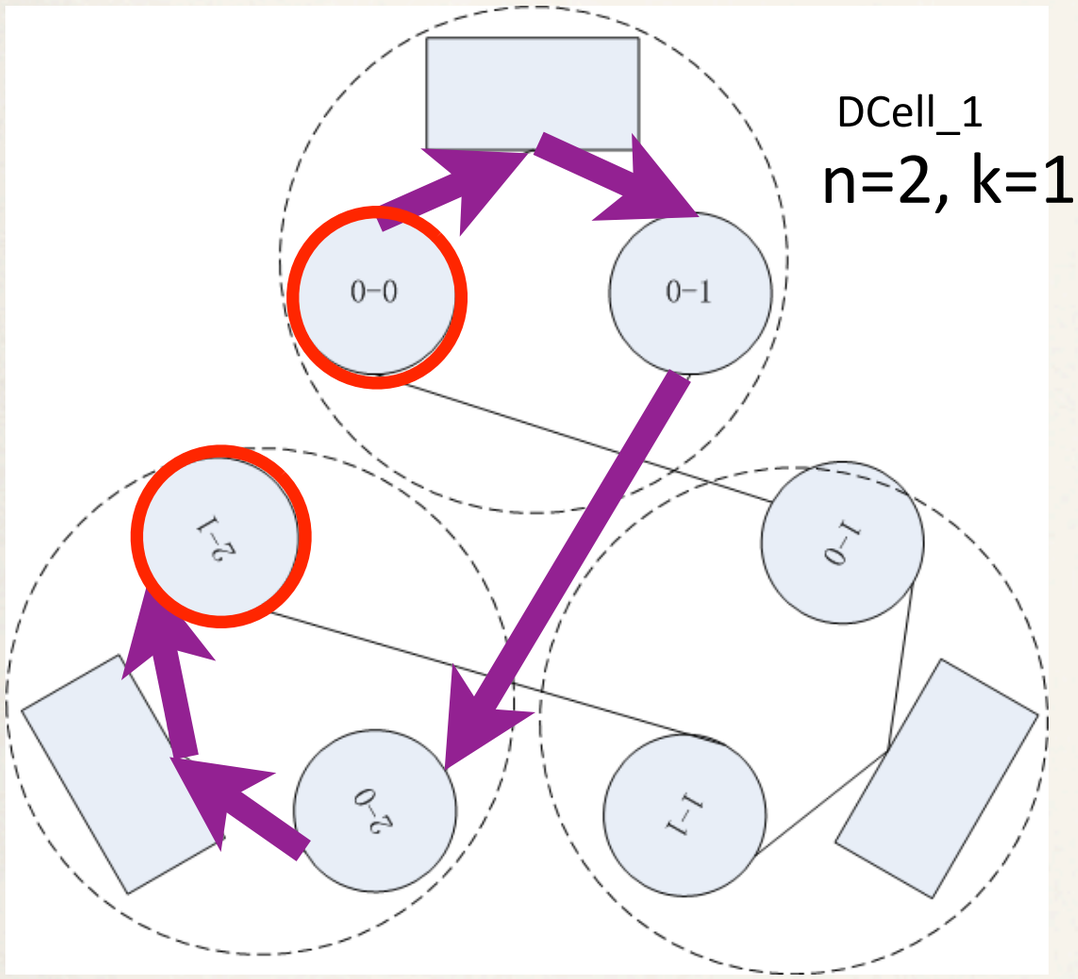
- Let F_L denote the number of flows carried by a L-level link, in all-to-all communication using the DCellRouting algorithm.

DCellRouting [Guo 08] : Flow Distribution



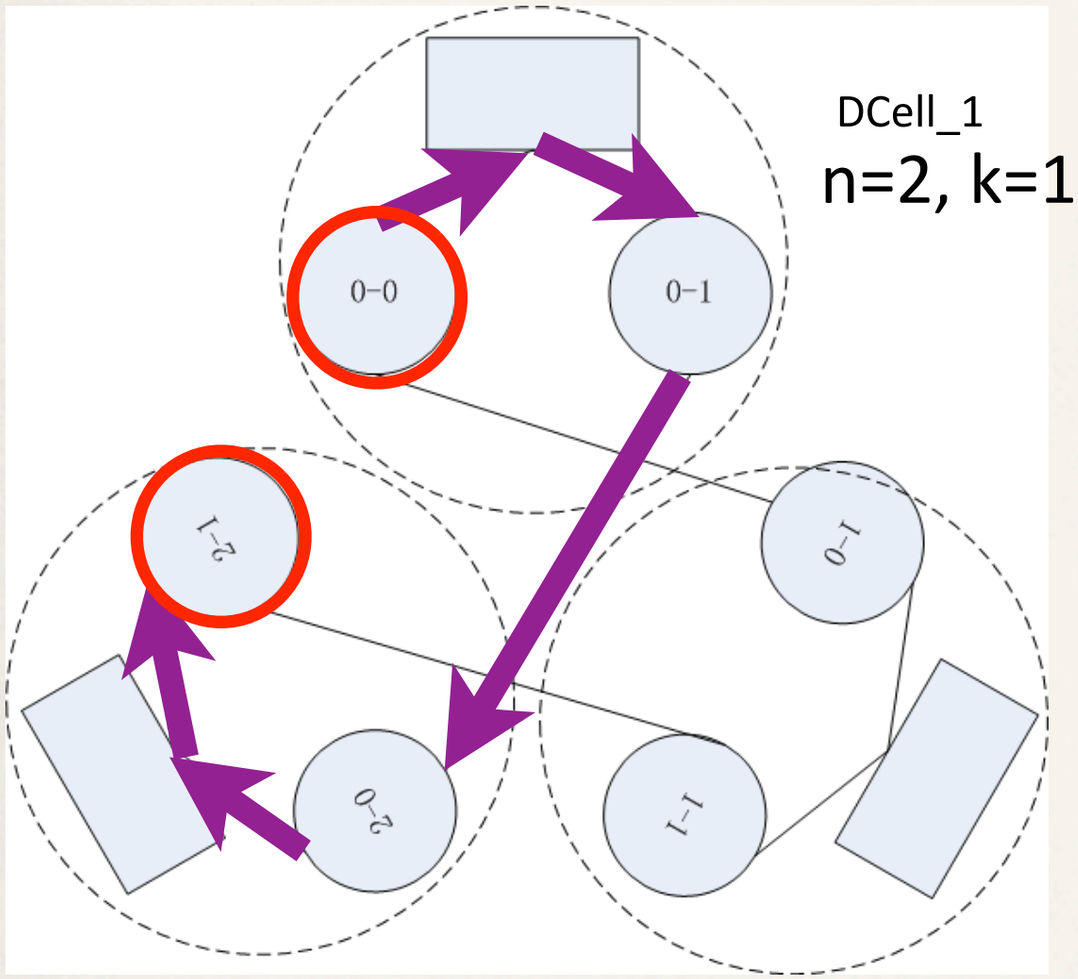
- Let F_L denote the number of flows carried by a L-level link, in all-to-all communication using the DCellRouting algorithm.

DCellRouting [Guo 08] : Flow Distribution



- Let F_L denote the number of flows carried by a L-level link, in all-to-all communication using the DCellRouting algorithm.

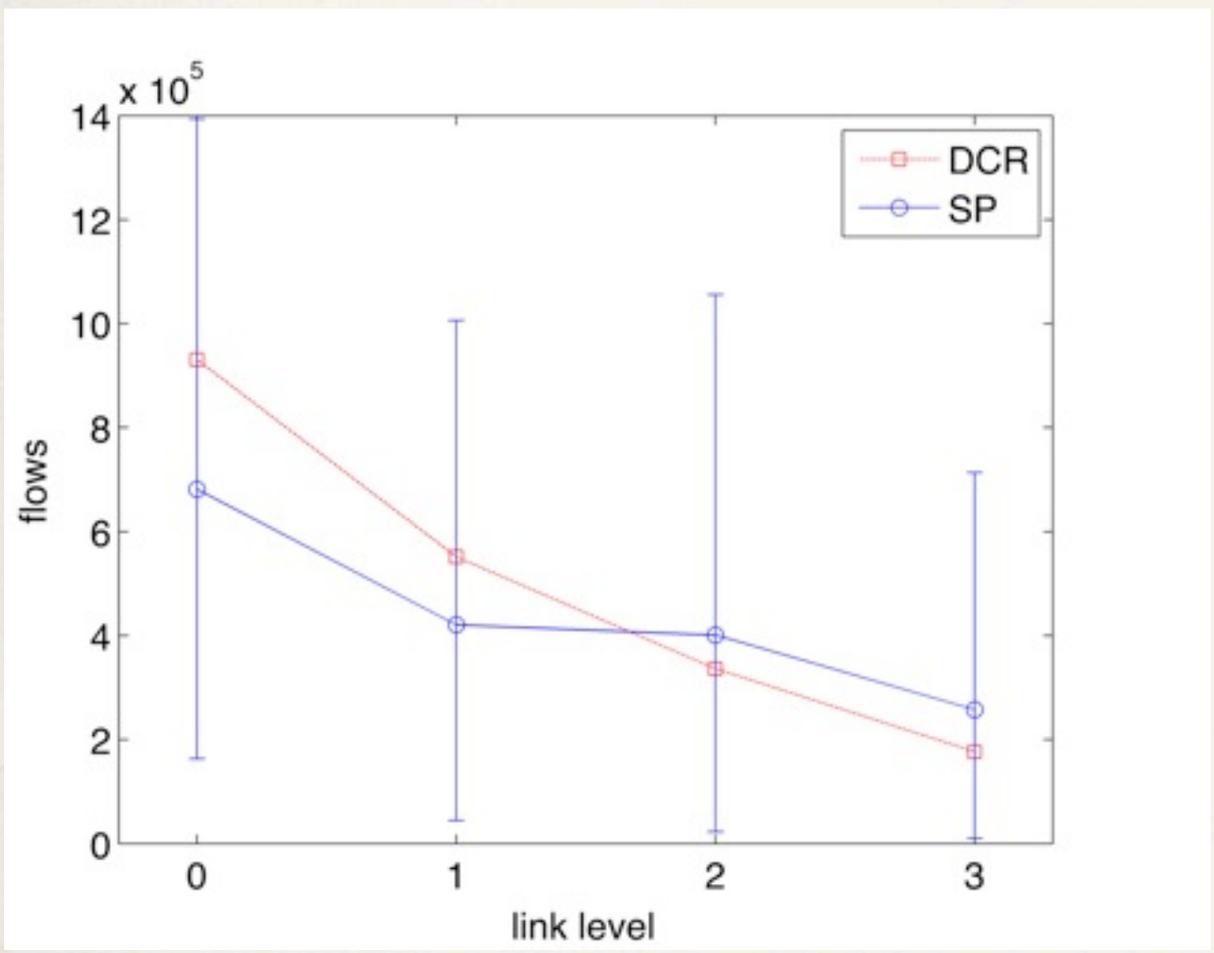
DCellRouting [Guo 08] : Flow Distribution



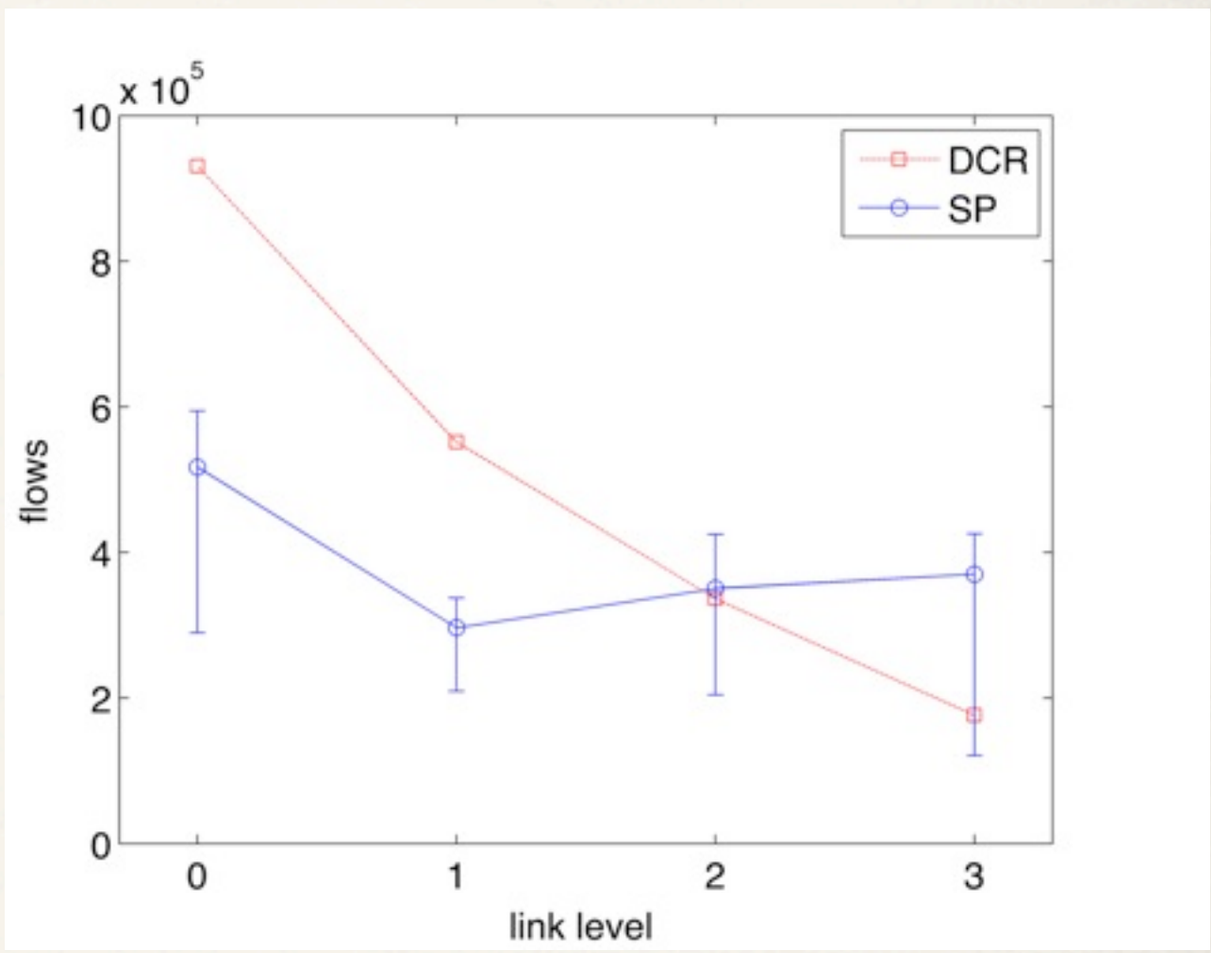
- Let F_L denote the number of flows carried by a L-level link, in all-to-all communication using the DCellRouting algorithm.

$$F_L = \begin{cases} t_{k-1}^2 & \text{for } L = k, \\ t_{L-1}^2 \prod_{j=L}^{k-1} (1 + 2t_j) & \text{for } 1 \le L \le k-1, \\ (n-1) \prod_{j=0}^{k-1} (1 + 2t_j) & \text{for } L = 0. \end{cases}$$

Flow Distribution: DCR vs. SP Routing



α-DCell



γ-DCell

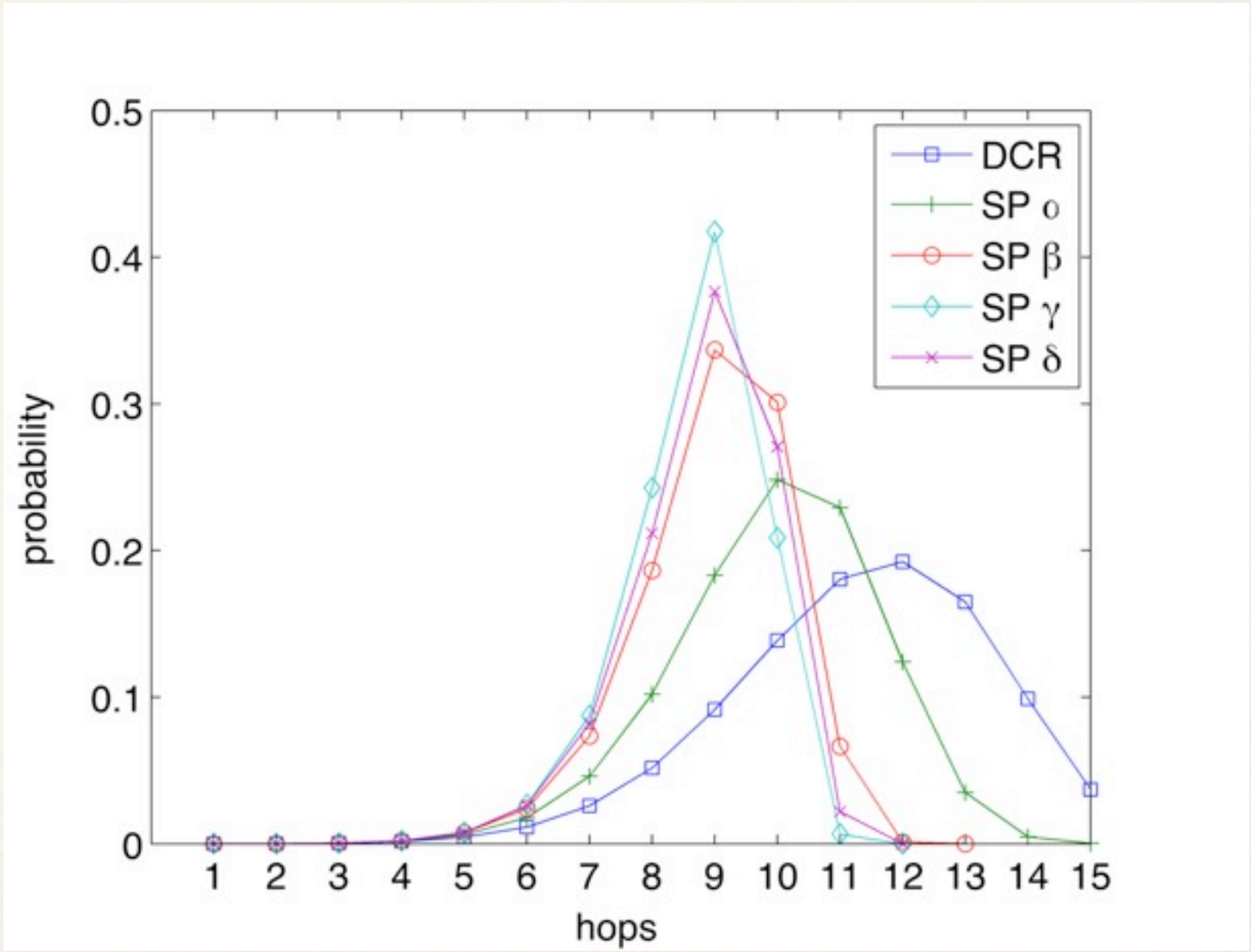
Distribution of flows by link level using all-to-all communication for n = 4 and k = 3. The error bars indicate the maximum and minimum values. The performance of β-DCell and δ-DCell is similar to that of γ-DCell.

Path-Length Distribution

n	k	α	β	γ	δ
2	2	7	6	6	6
4	2	7	7	7	7
2	3	15	10	10	10
4	3	15	13	12	12
2	4	27	17	15	16

Path-Length Distribution

n	k	α	β	γ	δ
2	2	7	6	6	6
4	2	7	7	7	7
2	3	15	10	10	10
4	3	15	13	12	12
2	4	27	17	15	16



Path-length distribution for n=4 and k=3.

Summary

- Generalized DCell
- Graph properties
  - Scalability, diameter and symmetry
- Routing algorithms
  - Path-length distribution and flow distribution
- Simulations
  - Smaller diameter and better load-balancing than the original DCell

Next Work

- We don’t know:
  - how the connection rules, are related with the properties of data center networks, e.g., low diameter and high symmetry.
  - how to exploit the freedom of the generalized DCell
  - whether there is a better routing algorithm
    - a practical, scalable and load-balancing shortest path algorithm

Thank you!

Questions ?