Toward Understanding Heterogeneity in Computing

Arnold L. Rosenberg
Ron C. Chiang

Department of Electrical and Computer Engineering
Colorado State University
Fort Collins, CO, USA
{rsnbrg, ron.chiang@colostate.edu}

Motivation

- Goal
  - to increase our understanding of heterogeneity in computing platforms

Motivation

- Goal
  - to increase our understanding of heterogeneity in computing platforms

- Heterogeneous computing platforms
  - different computing speeds

Motivation

- Goal
  - to increase our understanding of heterogeneity in computing platforms

- Heterogeneous computing platforms
  - different computing speeds
  - architecturally balanced

"Understanding" Heterogeneity

Suppose we have

- n+1 computers:
  - the server C_0
  - a "cluster" ***C*** comprising n computers, C_1, …, C_n
- Heterogeneity profile of ***C***
  - C_i can complete one unit of work in time ρ_i
  - $< \rho_1, \ldots, \rho_n >$
  - $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_n$

The Cluster-Exploitation Problem (CEP)

- C_0 must complete as many units of work as possible on cluster $\boldsymbol{C}$ within a given lifespan of L time units

The Cluster-Exploitation Problem (CEP)

- C_0 must complete as many units of work as possible on cluster $\mathbf{C}$ within a given lifespan of L time units

- *A worksharing protocol*
  - a schedule that solves the CEP

Architectural Parameters

Fixed communication cost

– setup time σ

– latency λ

negligible over a long lifespan

Architectural Parameters and Sample Values

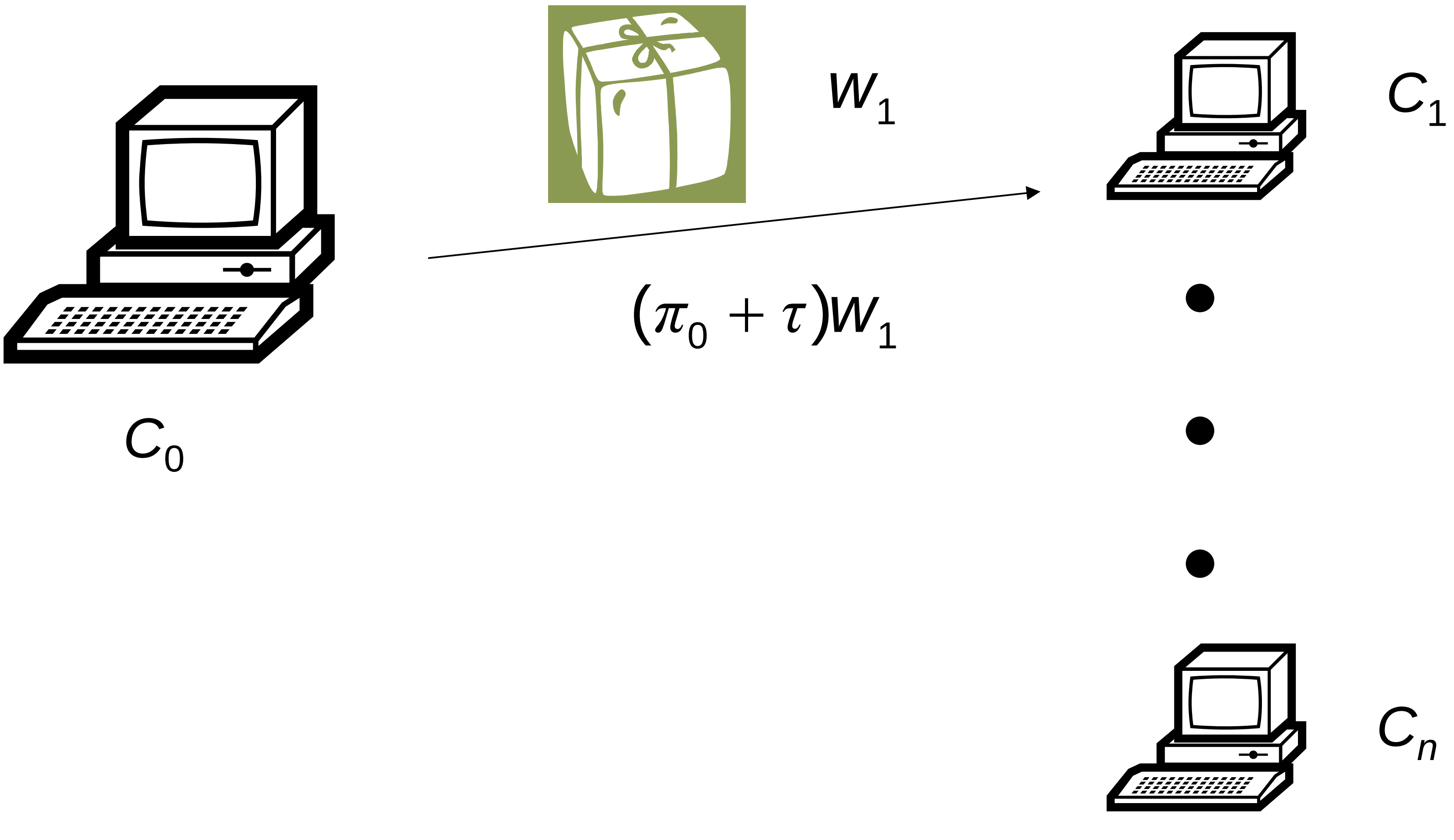
Common parameters:

- transmission rate τ (e.g. 1 μsec. / work unit)
- output-to-input length ratio δ (= 1)

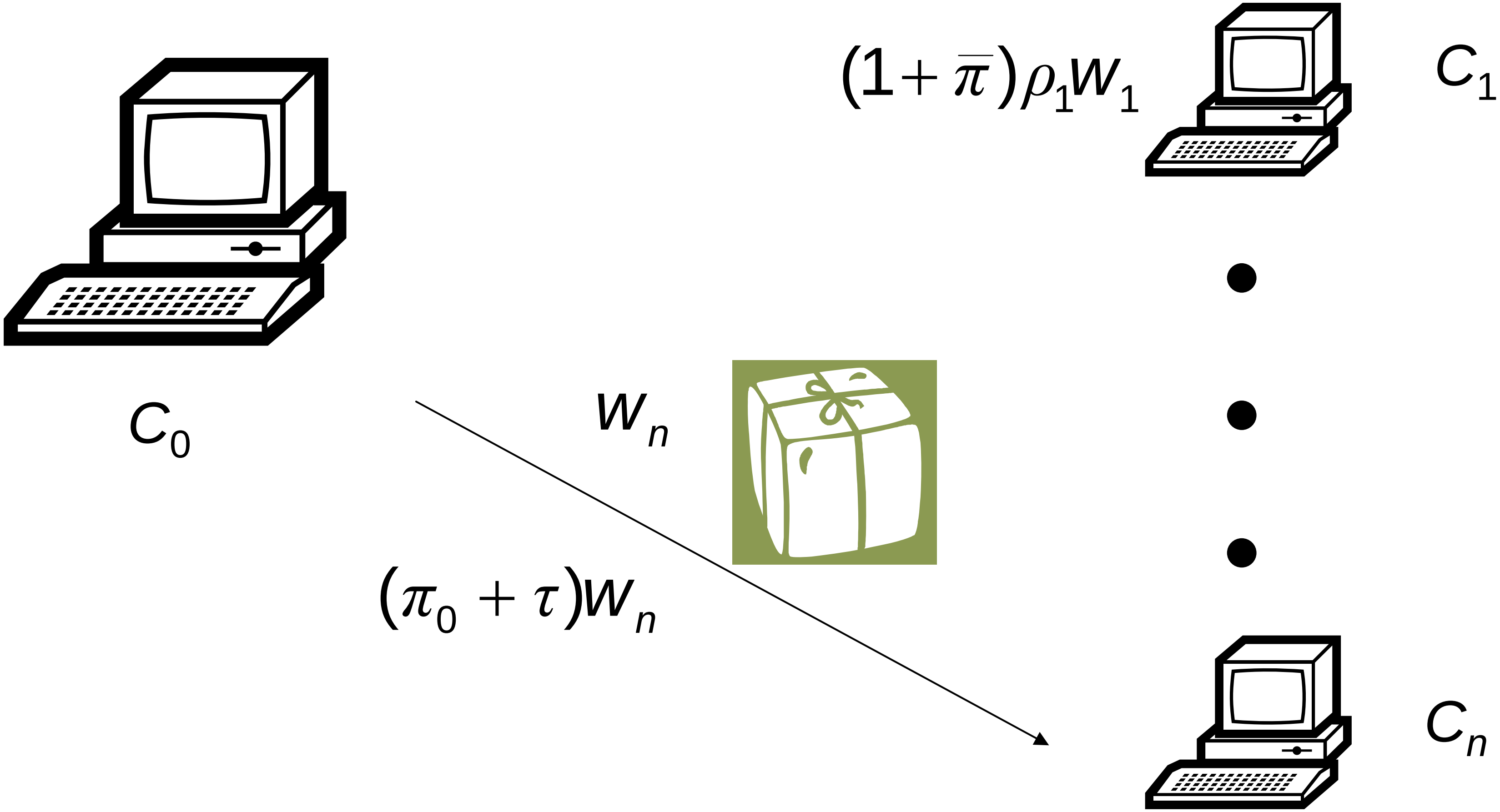
For computer *i*,

- packaging rate π_i (e.g. 10 μsec. / work unit)
- unpackaging rate $\bar{\pi}_i$ (e.g. 10 μsec. / work unit)
- workload w_i (work units)

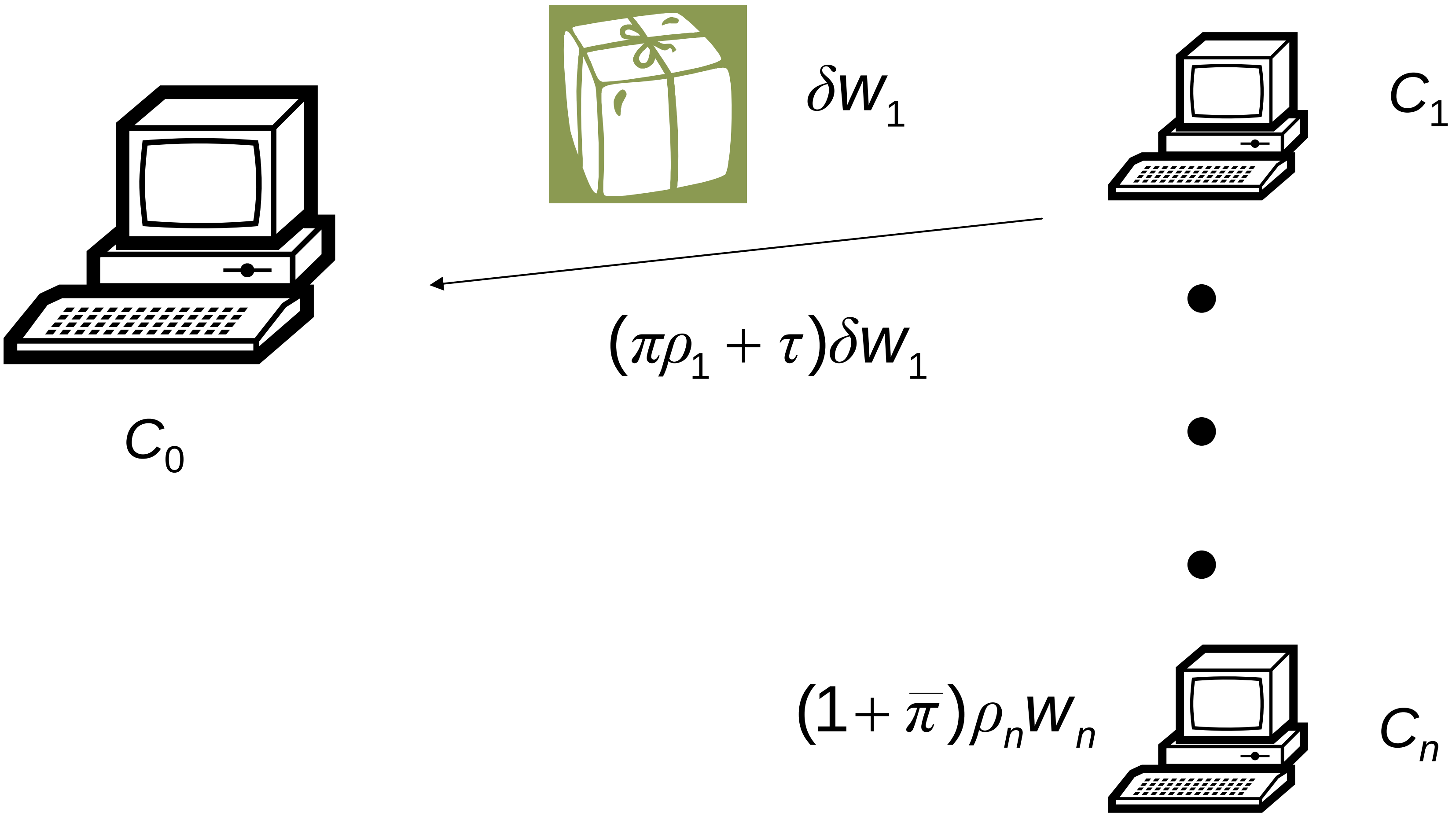
Worksharing Protocols



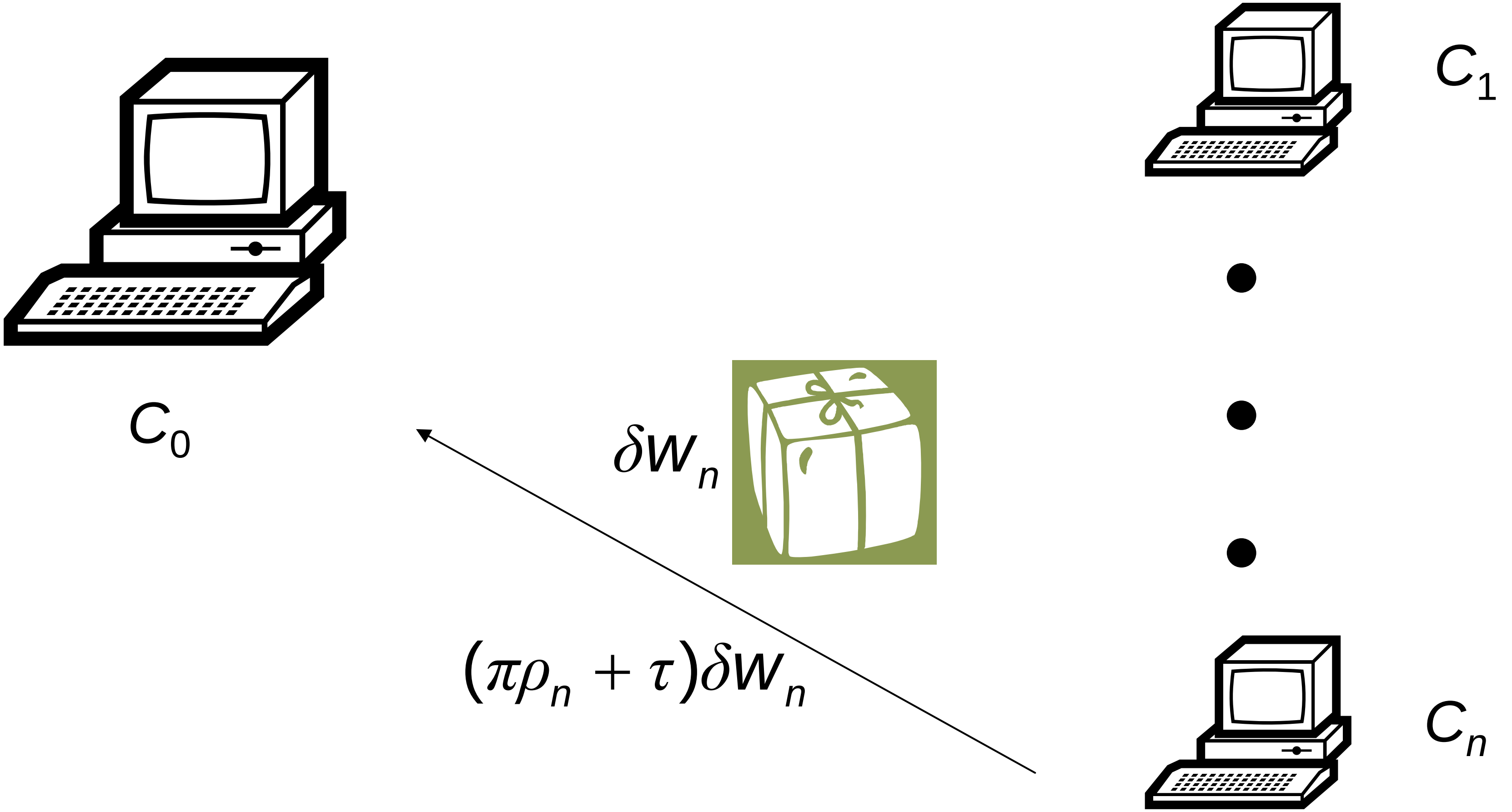
Worksharing Protocols



Worksharing Protocols



Worksharing Protocols



The FIFO Protocol

C_0	sends work to C_1	sends work to C_2	sends work to C_3
	$(\pi_0 + \tau)w_1$	$(\pi_0 + \tau)w_2$	$(\pi_0 + \tau)w_3$
C_1	waits	processes	results
		$(1+\bar{\pi})\rho_1 w_1$	$(\pi\rho_1 + \tau)\delta w_1$
C_2	waits	processes	results
		$(1+\bar{\pi})\rho_2 w_2$	$(\pi\rho_2 + \tau)\delta w_2$
C_3	waits	processes	results
		$(1+\bar{\pi})\rho_3 w_3$	$(\pi\rho_3 + \tau)\delta w_3$

(NOT TO SCALE)

The FIFO Protocol is Optimal

- Theorem [Adler-Gong-Rosenberg]

  Over any sufficiently long lifespan *L*, for any heterogeneous cluster ***C*** — *no matter what its heterogeneity profile*:

  – FIFO worksharing protocols provide optimal solutions to the cluster-exploitation problem

  – ***C*** is equally productive under every FIFO protocol, i.e., *under all startup orderings*

The Work-Production of FIFO

Let

$$X = \sum_{i=1}^{n} \frac{1}{(\pi_0 + \tau) + (1 + \overline{\pi} + \pi\delta)\rho_i} \prod_{j=1}^{i-1} \left(1 - \frac{\pi_0 + \tau - \tau\delta}{(\pi_0 + \tau) + (1 + \overline{\pi} + \pi\delta)\rho_j} \right)$$

The Work-Production of FIFO

Let

$$X = \sum_{i=1}^{n} \frac{1}{(\pi_0 + \tau) + (1 + \bar{\pi} + \pi\delta)\rho_i} \prod_{j=1}^{i-1} \left(1 - \frac{\pi_0 + \tau - \tau\delta}{(\pi_0 + \tau) + (1 + \bar{\pi} + \pi\delta)\rho_j} \right)$$

Then,

$$W = \frac{1}{\tau\delta + 1/X} \cdot L$$

The Work-Production of FIFO

Let

$$X = \sum_{i=1}^{n} \frac{1}{(\pi_0 + \tau) + (1 + \overline{\pi} + \pi\delta)\rho_i} \prod_{j=1}^{i-1} \left(1 - \frac{\pi_0 + \tau - \tau\delta}{(\pi_0 + \tau) + (1 + \overline{\pi} + \pi\delta)\rho_j} \right)$$

≈ ≈

To simplify, let $A = \pi_0 + \tau$ and $B = 1 + \overline{\pi} + \pi\delta$,

$$X = \sum_{i=1}^{n} \frac{1}{A + B\rho_i} \prod_{j=1}^{i-1} \left(\frac{B\rho_j + \tau\delta}{A + B\rho_j} \right)$$

On Comparing Heterogeneity Profiles

- For any cluster ***C*** with heterogeneity profile $P = \langle \rho_1, \ldots, \rho_n \rangle$

On Comparing Heterogeneity Profiles

- For any cluster ***C*** with heterogeneity profile

  $P = \langle \rho_1 , \ldots , \rho_n \rangle$

- ***C***'s *homogeneous-equivalent computing rate* (*HECR*) is

$$\rho_c = \max_{\rho} \left\{ X(P^{(\rho)}) \geq X(P) \right\}$$

  where $P^{(\rho)} = \langle \rho , \ldots , \rho \rangle$

Heterogeneity Profiles

Profile 1: $\rho_i = \dfrac{n-i+1}{n}$, which spreads evenly in a range

when $n = 8$, $\left\langle \dfrac{8}{8}, \dfrac{7}{8}, \dfrac{6}{8}, \ldots, \dfrac{1}{8} \right\rangle$

	Number of Computers		
	8	16	32
HECR	0.362	0.297	0.251

Recall: faster cluster has smaller HECR value

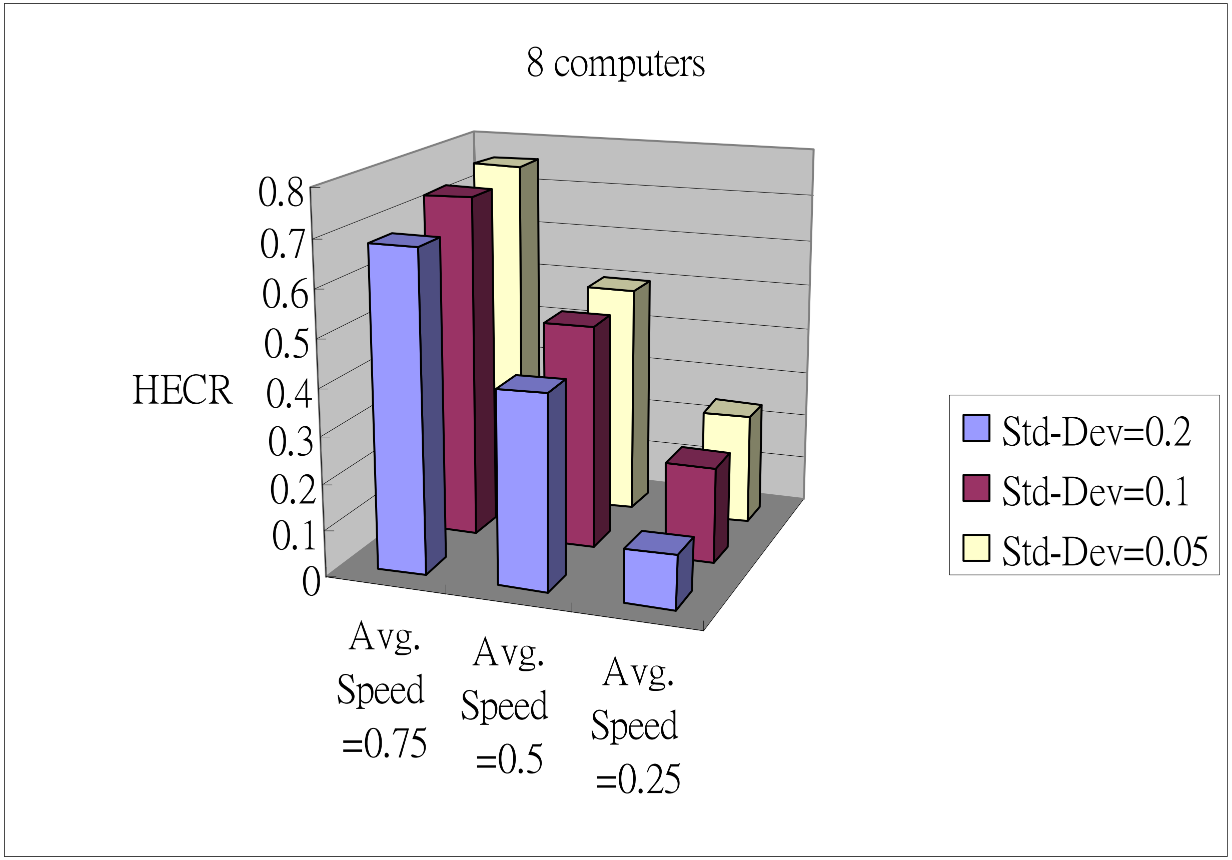
Heterogeneity Profiles

Profile 2 : $\rho_i = \frac{1}{i}$

when $n = 8, \left\langle \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{8} \right\rangle$

	Number of Computers		
	8	16	32
HECR	0.216	0.116	0.061

Avg. Speed vs. Std-Dev of Speed



Randomly generate 100 profiles for each combination

Avg. Speed vs. Std-Dev of Speed

8 computers' HECR		Std-Dev		
		0.2	0.1	0.05
Avg. Speed	0.75	0.681	0.735	0.759
	0.5	0.411	0.482	0.501
	0.25	0.113	0.208	0.239

The probability that these two groups have the same mean is 2×10^{-10}

Avg. Speed vs. Std-Dev of Speed

8 computers' HECR		Std-Dev		
		0.2	0.1	0.05
Avg. Speed	0.75	0.681	0.735	0.759
	0.5	0.411	0.482	0.501
	0.25	0.113	0.208	0.239

Trials with 16, 32 computers show similar pattern

Speeding Up Clusters Optimally under FIFO Protocols

- Which one computer should you speed up, if you can speed up only one?

Speeding Up Clusters Optimally under FIFO Protocols

- Which one computer should you speed up, if you can speed up only one?
- We study two variants of this question

Speeding Up Clusters Optimally under FIFO Protocols

For convenience,

- let cluster ***C*** have heterogeneity profile $P = <\rho_1, \ldots, \rho_n>$,
  where $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_n$
- let i and $j > i$ be two computer indices

Fixed and Proportional Speed-up

- *Fixed-speedup* scenario
- by a fixed amount $\varphi < \rho_n$

$$P^{(i)} = \left\langle \rho_1, \ldots, \rho_{i-1}, \rho_i - \varphi, \rho_{i+1}, \ldots, \rho_{j-1}, \rho_j, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

$$P^{(j)} = \left\langle \rho_1, \ldots, \rho_{i-1}, \rho_i, \rho_{i+1}, \ldots, \rho_{j-1}, \rho_j - \varphi, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

Fixed and Proportional Speed-up

- *Fixed-speedup* scenario (by a fixed amount $\varphi < \rho_n$)

$$P^{(i)} = \left\langle \rho_1, \ldots, \rho_{i-1}, \rho_i - \varphi, \rho_{i+1}, \ldots, \rho_{j-1}, \rho_j, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

$$P^{(j)} = \left\langle \rho_1, \ldots, \rho_{i-1}, \rho_i, \rho_{i+1}, \ldots, \rho_{j-1}, \rho_j - \varphi, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

- *Proportional-speedup* scenario
- by a relative amount $\psi < 1$

$$P^{[i]} = \left\langle \rho_1, \ldots, \rho_{i-1}, \psi\rho_i, \rho_{i+1}, \ldots, \rho_{j-1}, \rho_j, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

$$P^{[j]} = \left\langle \rho_1, \ldots, \rho_{i-1}, \rho_i, \rho_{i+1}, \ldots, \rho_{j-1}, \psi\rho_j, \rho_{j+1}, \ldots, \rho_n \right\rangle$$

Proposition for Fixed-Speedup

- Under the fixed-speedup scenario, the most advantageous single computer to speed up is ***C***'s fastest computer

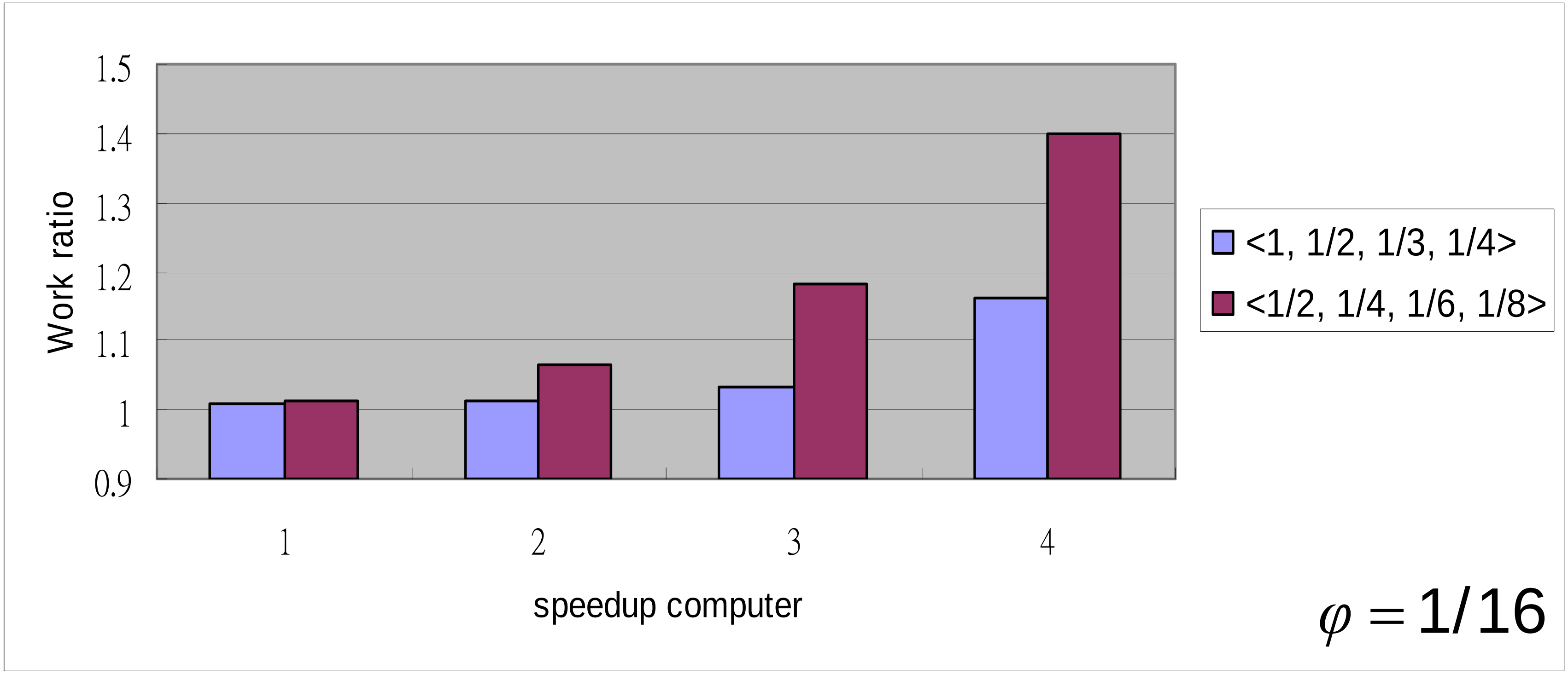
Terms for following figures

- Recall: work production $W = \frac{1}{\tau\delta + 1/X} \cdot L$

- Work ratio
  – the ratio of work production after speedup to work production before speedup

- Speedup computer
  – the single computer that is sped up

Fixed-Speedup Scenario



Proposition for Proportional-Speedup

(Recall : $\text{A} = \pi_0 + \tau$, $\text{B} = 1 + \bar{\pi} + \pi\delta$, and $\rho_i > \rho_j$)

- If $\psi\rho_i\rho_j > A\tau\delta / B^2$
  - – speeding up C_j (faster) is better

- If $\psi\rho_i\rho_j < A\tau\delta / B^2$
  - – speeding up C_i (slower) is better

Proposition for Proportional-Speedup

(Recall : A $= \pi_0 + \tau$, B $= 1 + \bar{\pi} + \pi\delta$, and $\rho_i > \rho_j$)

- If $\psi\rho_i\rho_j > A\tau\delta / B^2 = 1.0 \times 10^{-5}$
  - speeding up C_j (faster) is better
- If $\psi\rho_i\rho_j < A\tau\delta / B^2 = 1.0 \times 10^{-5}$
  - speeding up C_i (slower) is better

Parameter	Rate
A	11 μsecond / work unit
B with coarse (1 sec / task) tasks	1.000011 second / work unit

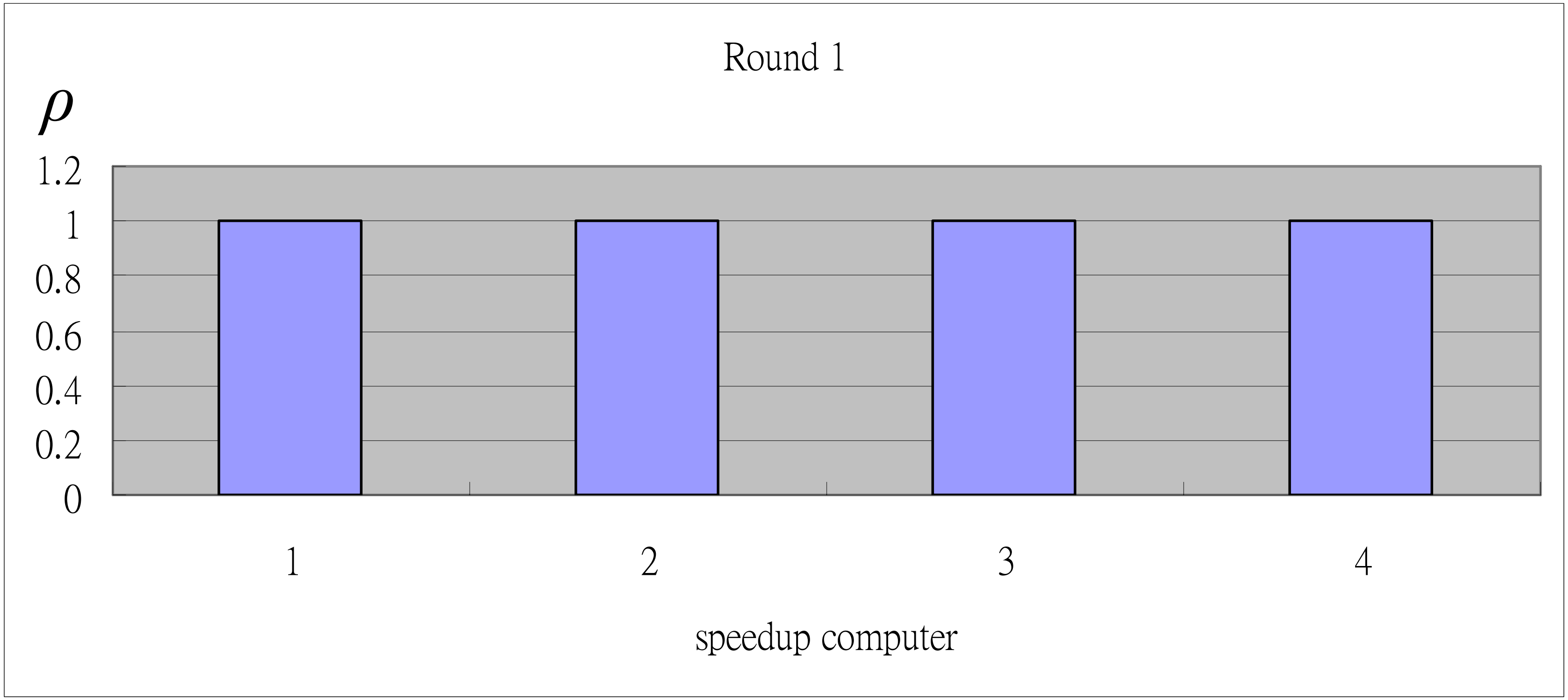
Proposition for Proportional-Speedup

(Recall : $\text{A} = \pi_0 + \tau$, $\text{B} = 1 + \bar{\pi} + \pi\delta$, and $\rho_i > \rho_j$)

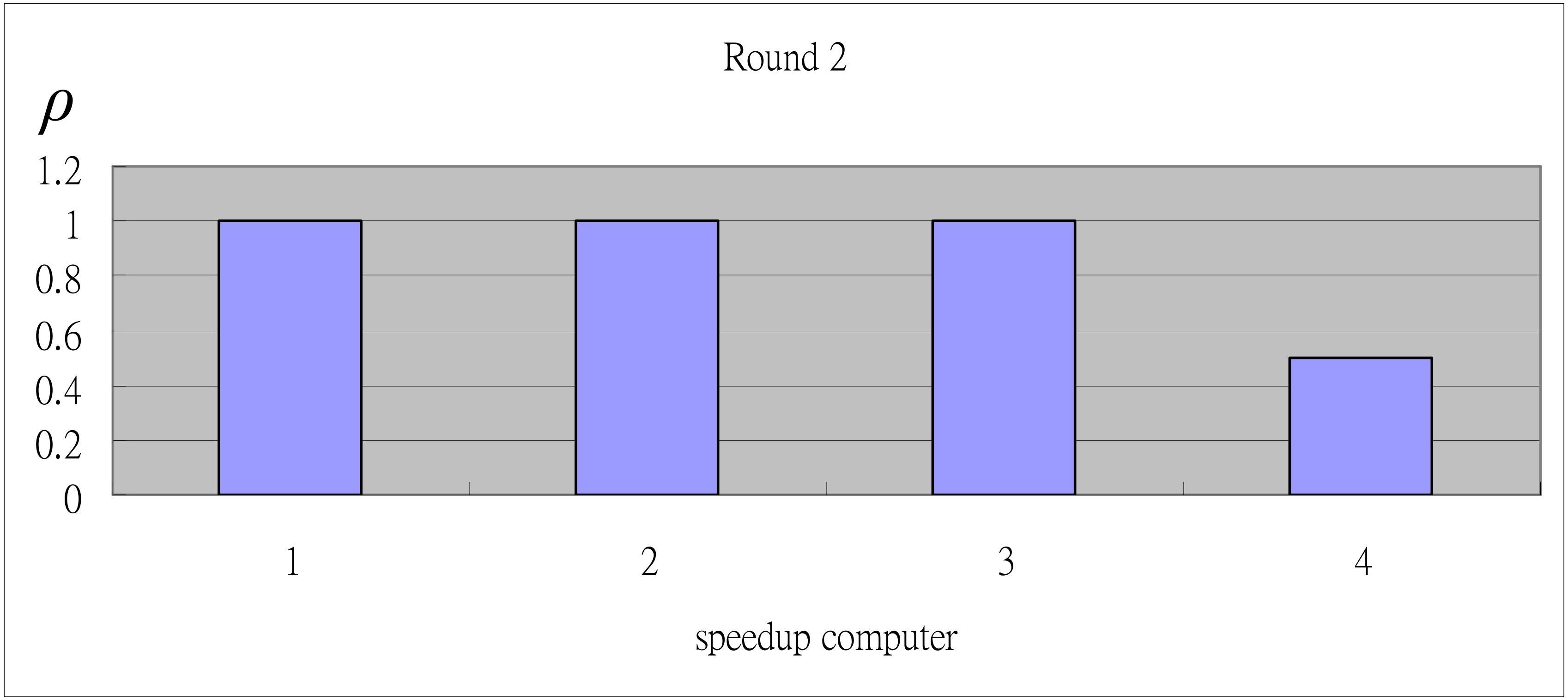
- If $\psi\rho_i\rho_j > A\tau\delta / B^2 = 1.0 \times 10^{-5}$
  - speeding up C_j (faster) is better
- If $\psi\rho_i\rho_j < A\tau\delta / B^2 = 1.0 \times 10^{-5}$
  - speeding up C_i (slower) is better

That is, it is more advantageous to speed up the faster one unless either both computers are already "very fast" or the speedup factor is "very large."

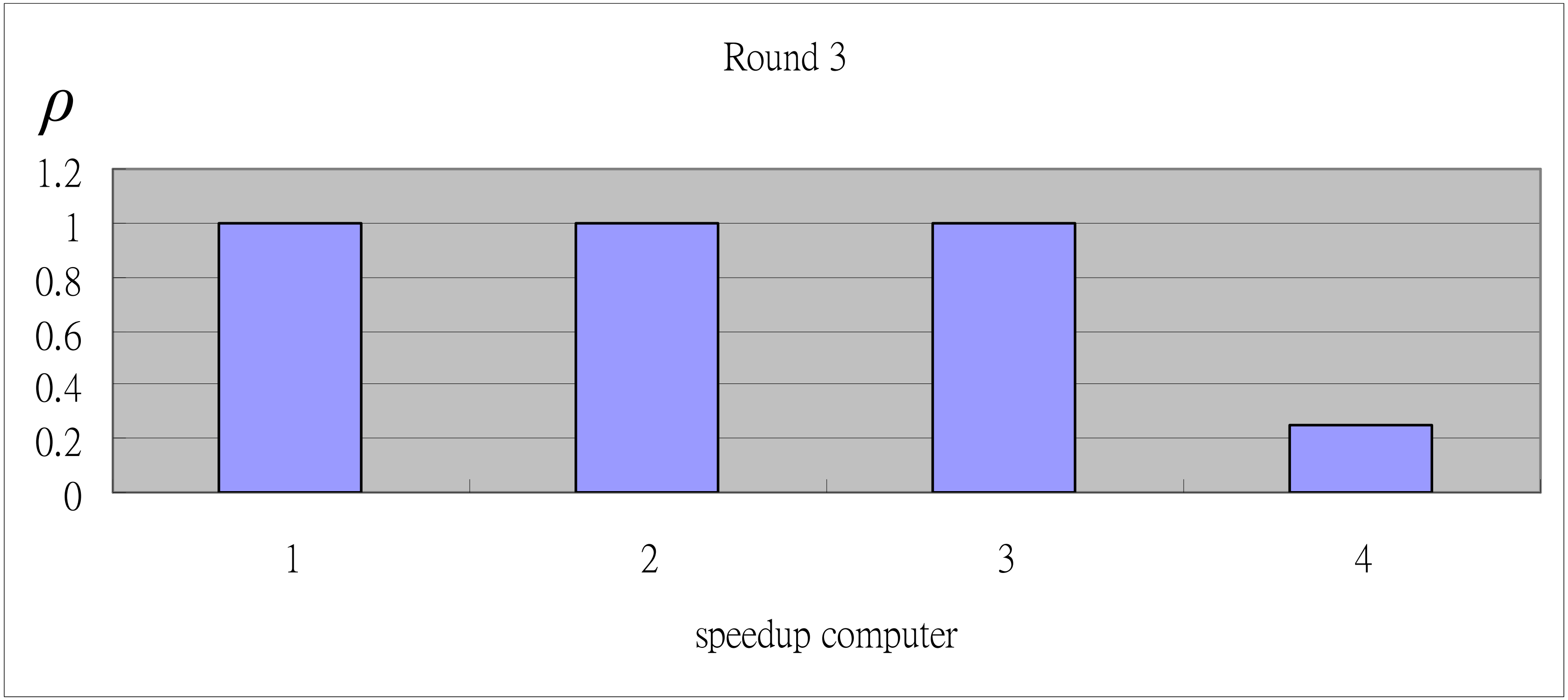
Proportional-Speedup in Action



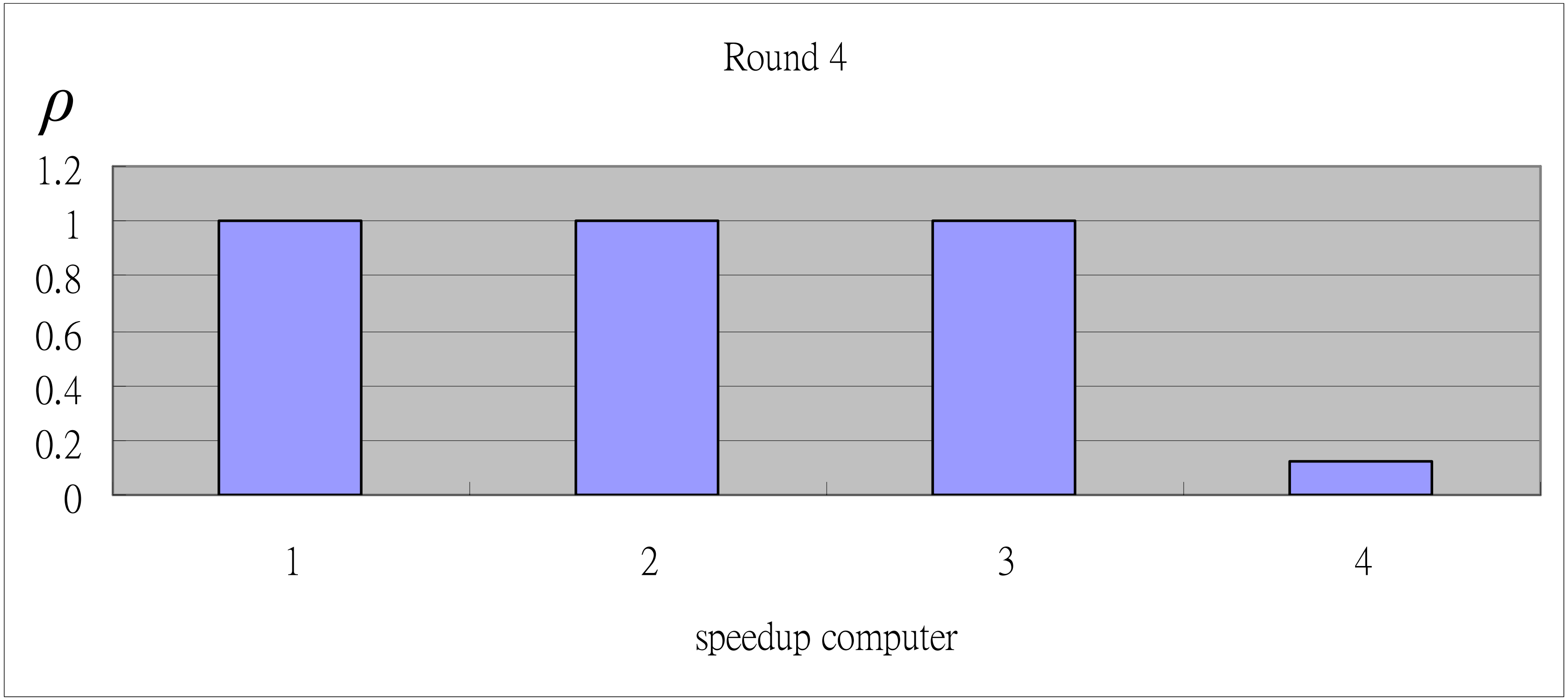
Proportional-Speedup in Action



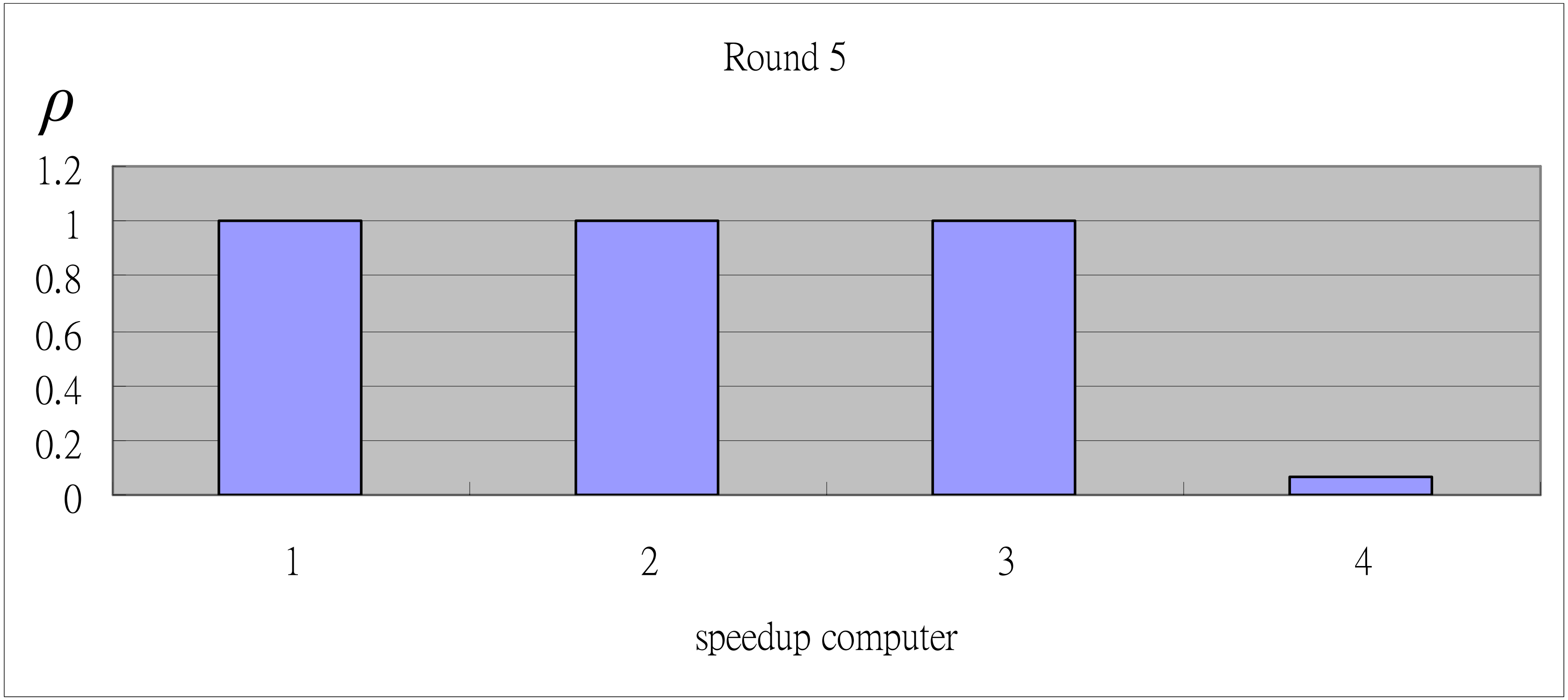
Proportional-Speedup in Action



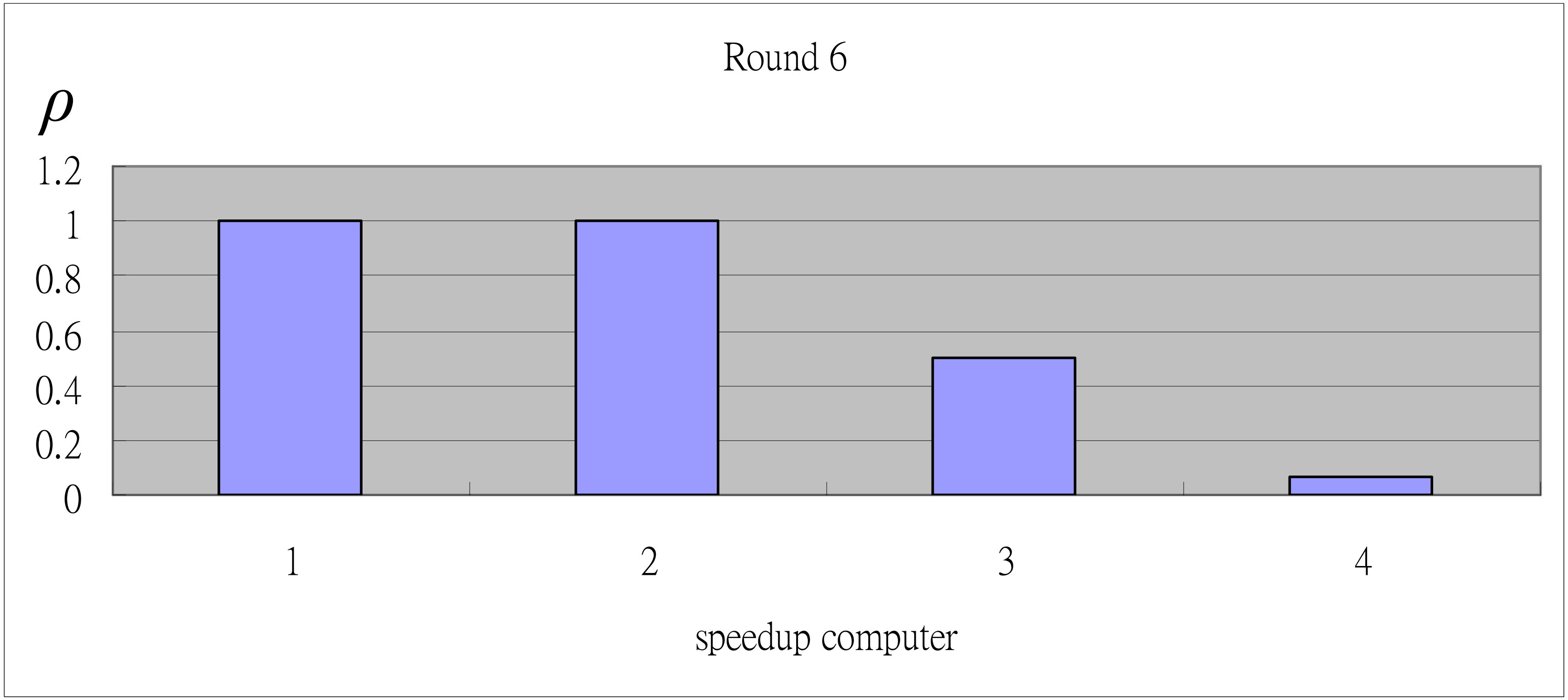
Proportional-Speedup in Action



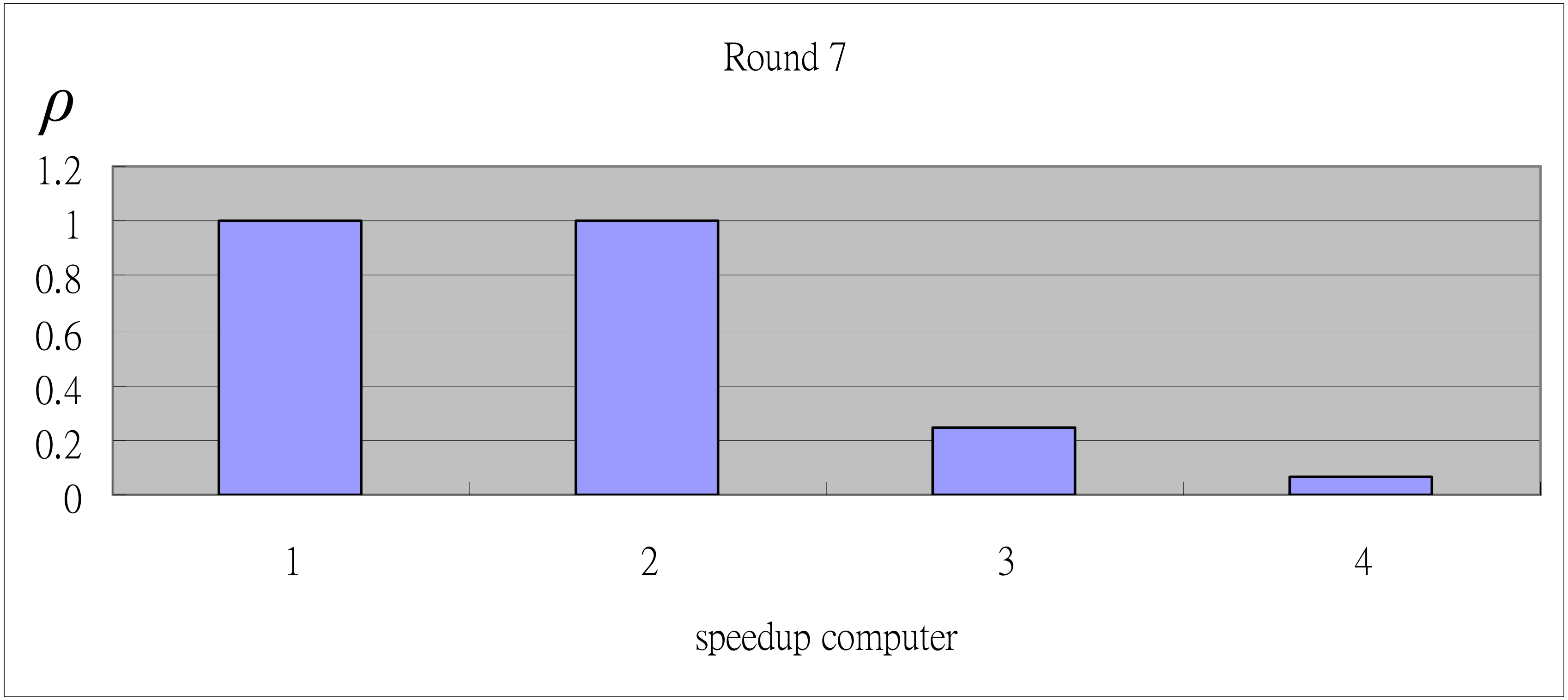
Proportional-Speedup in Action



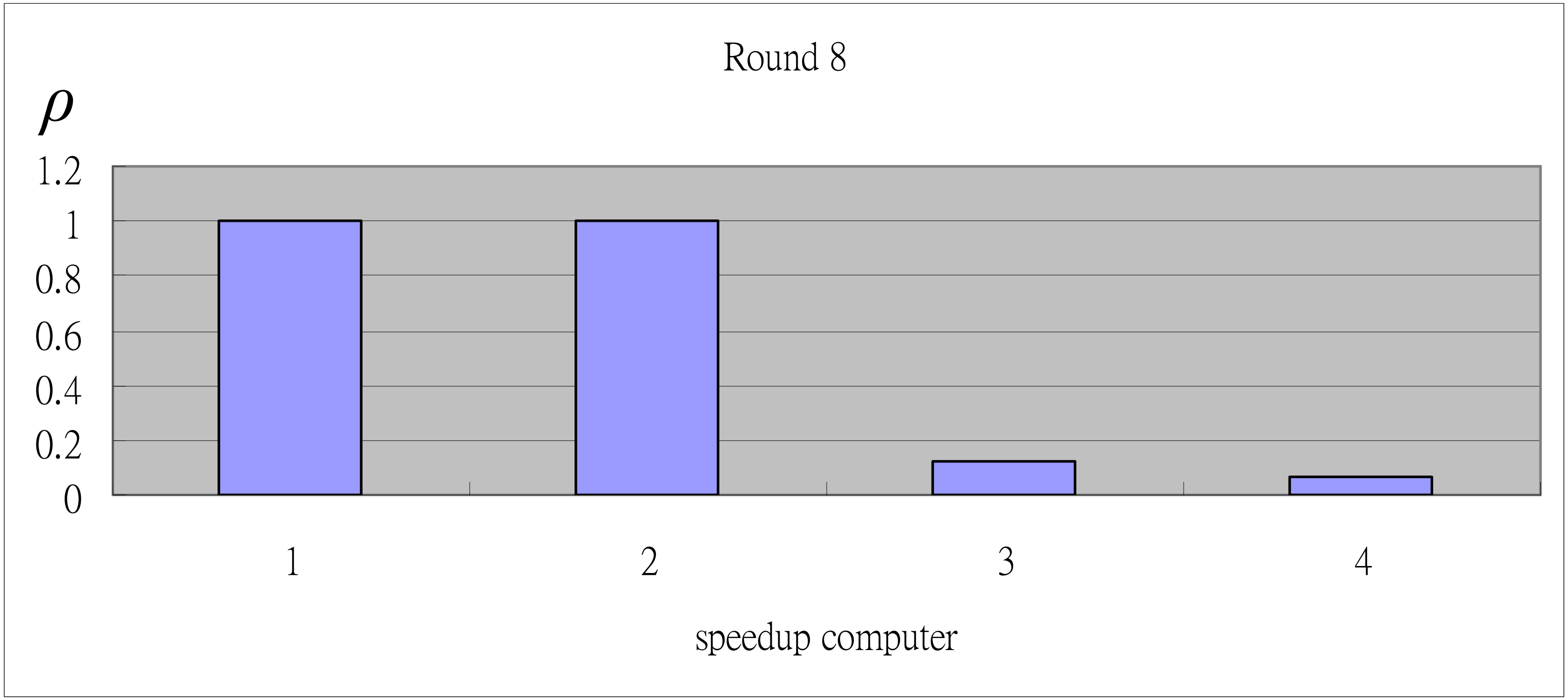
Proportional-Speedup in Action



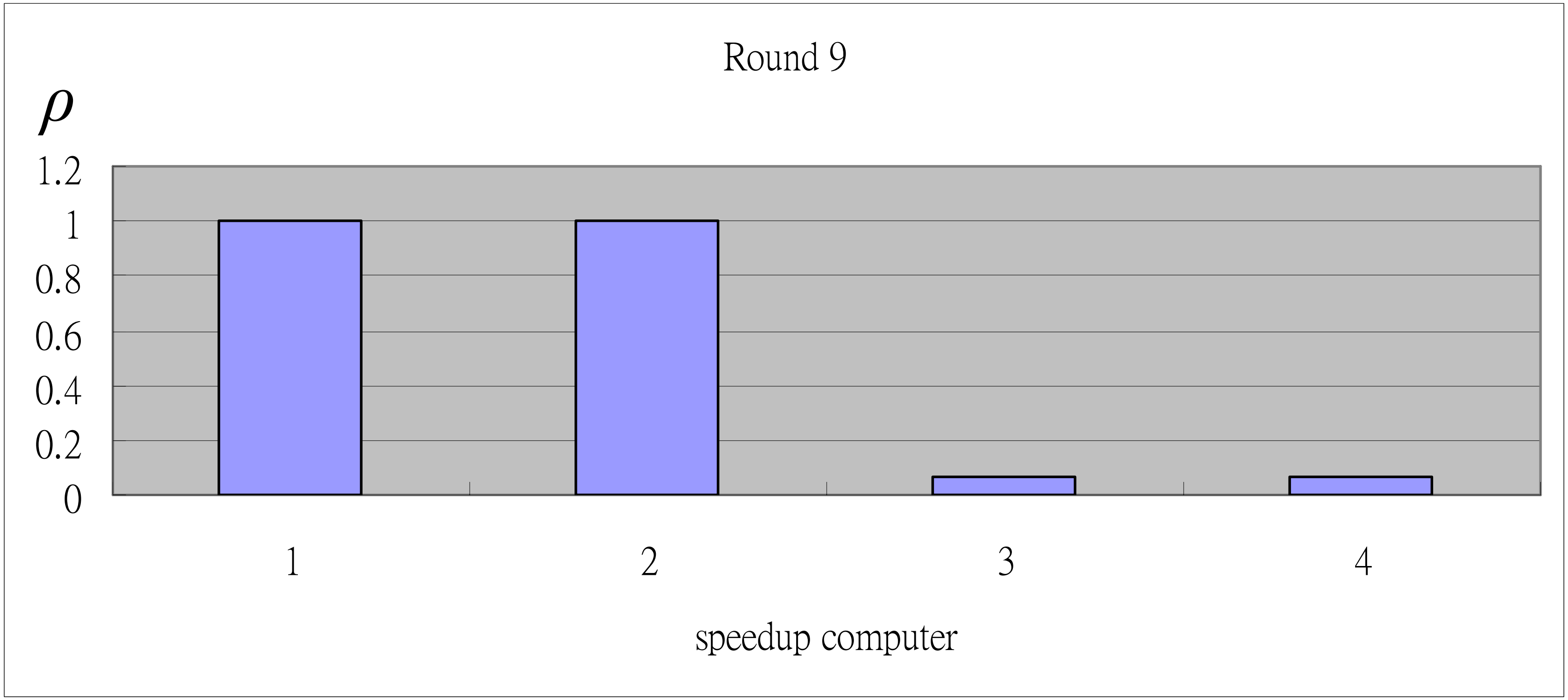
Proportional-Speedup in Action



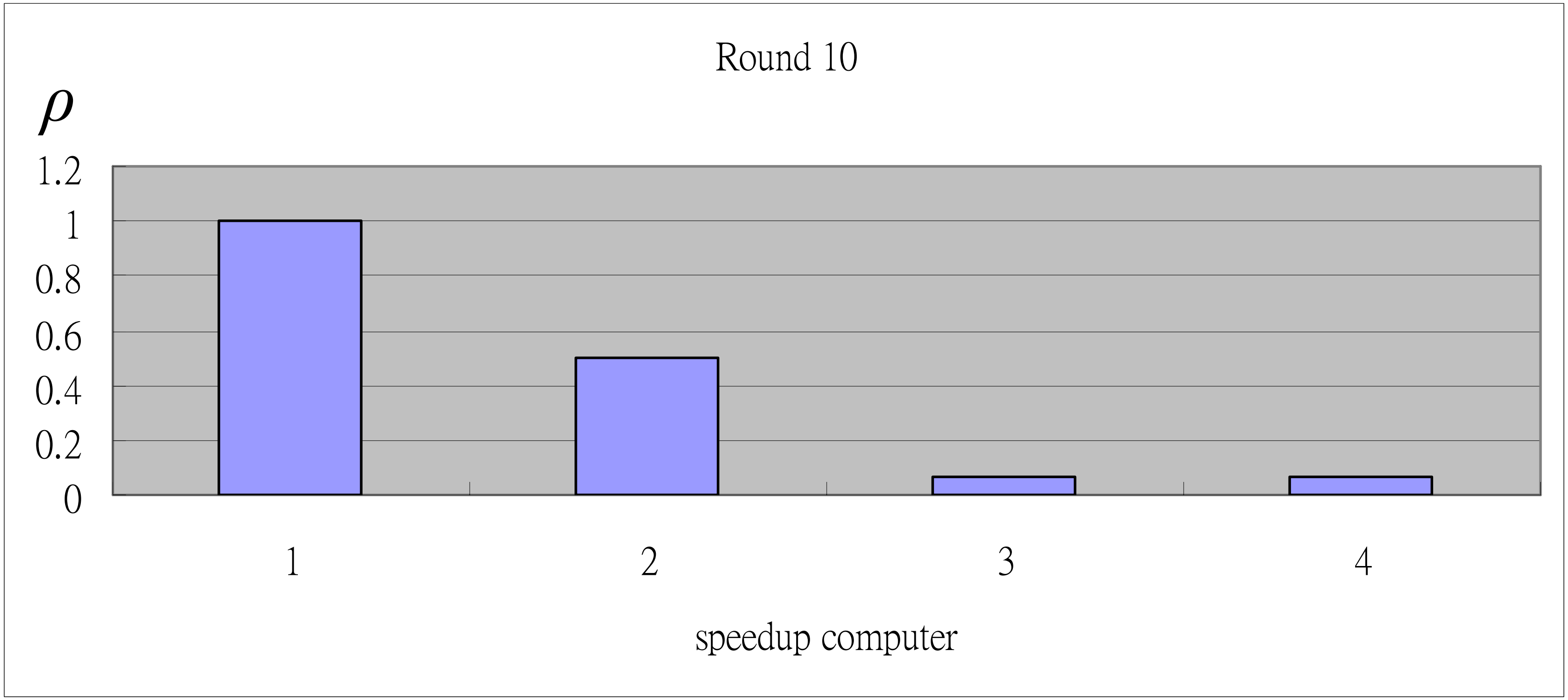
Proportional-Speedup in Action



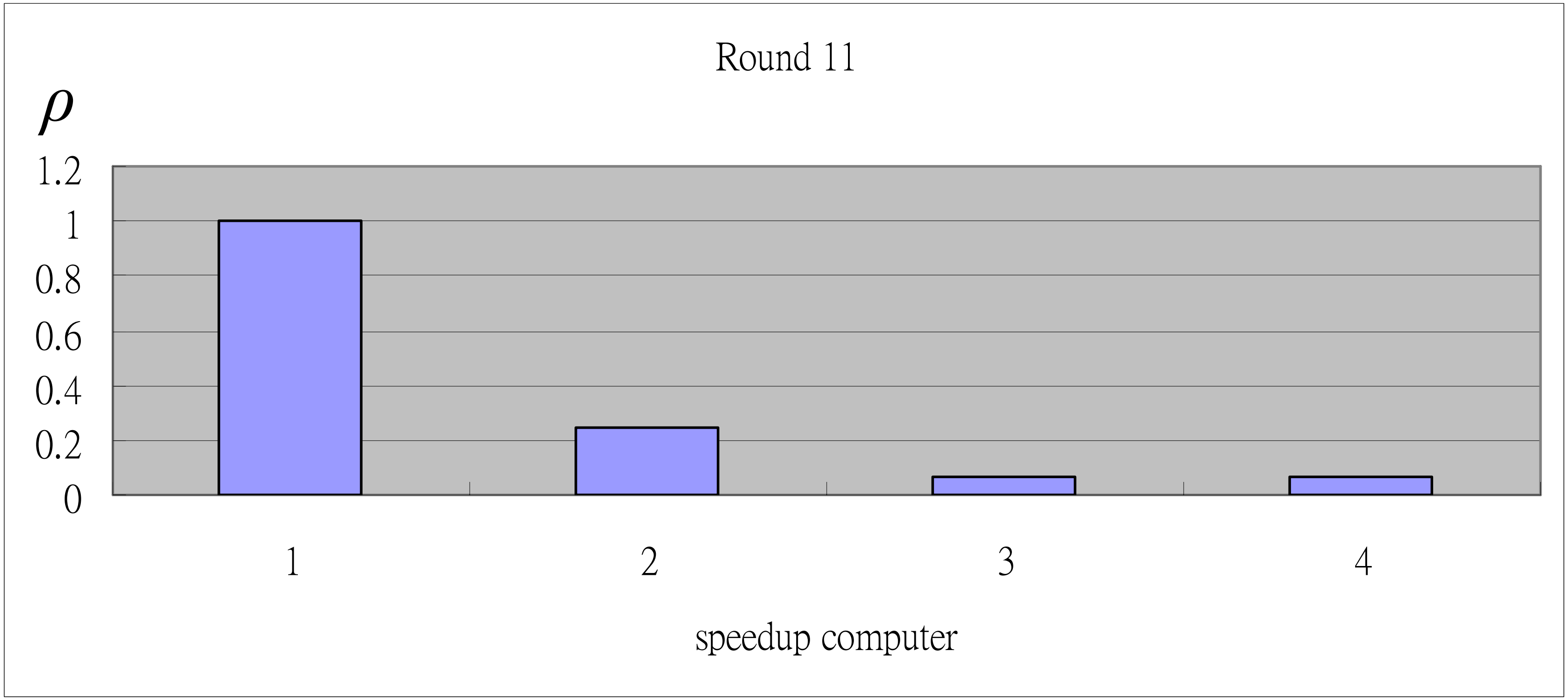
Proportional-Speedup in Action



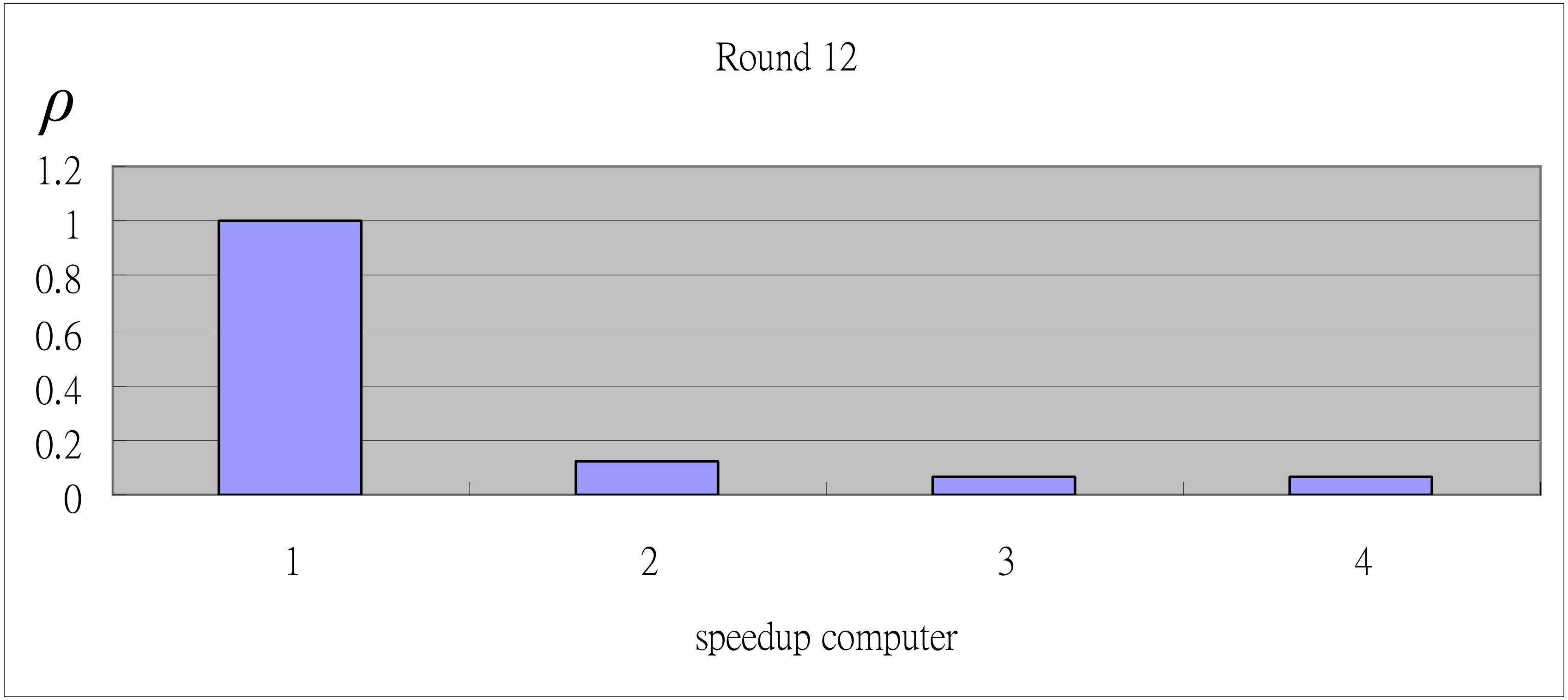
Proportional-Speedup in Action



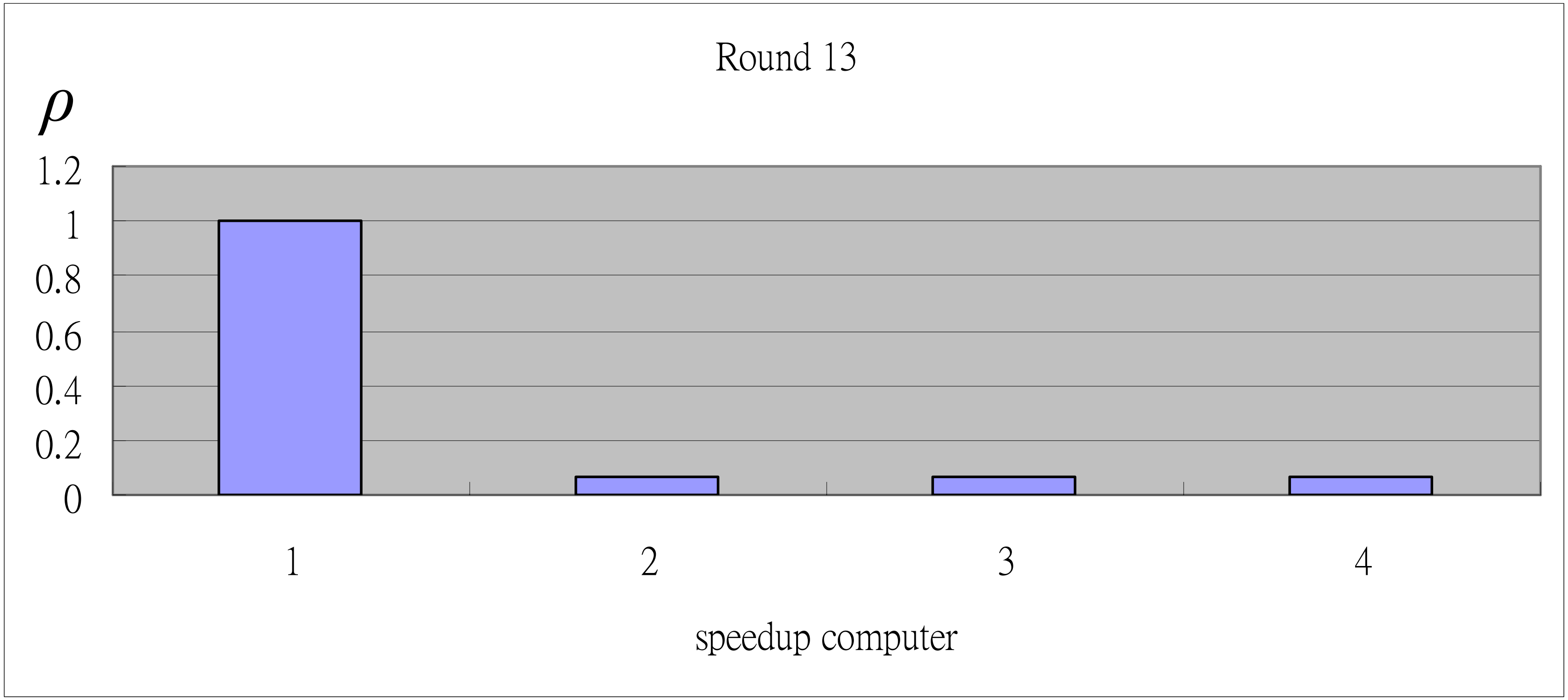
Proportional-Speedup in Action



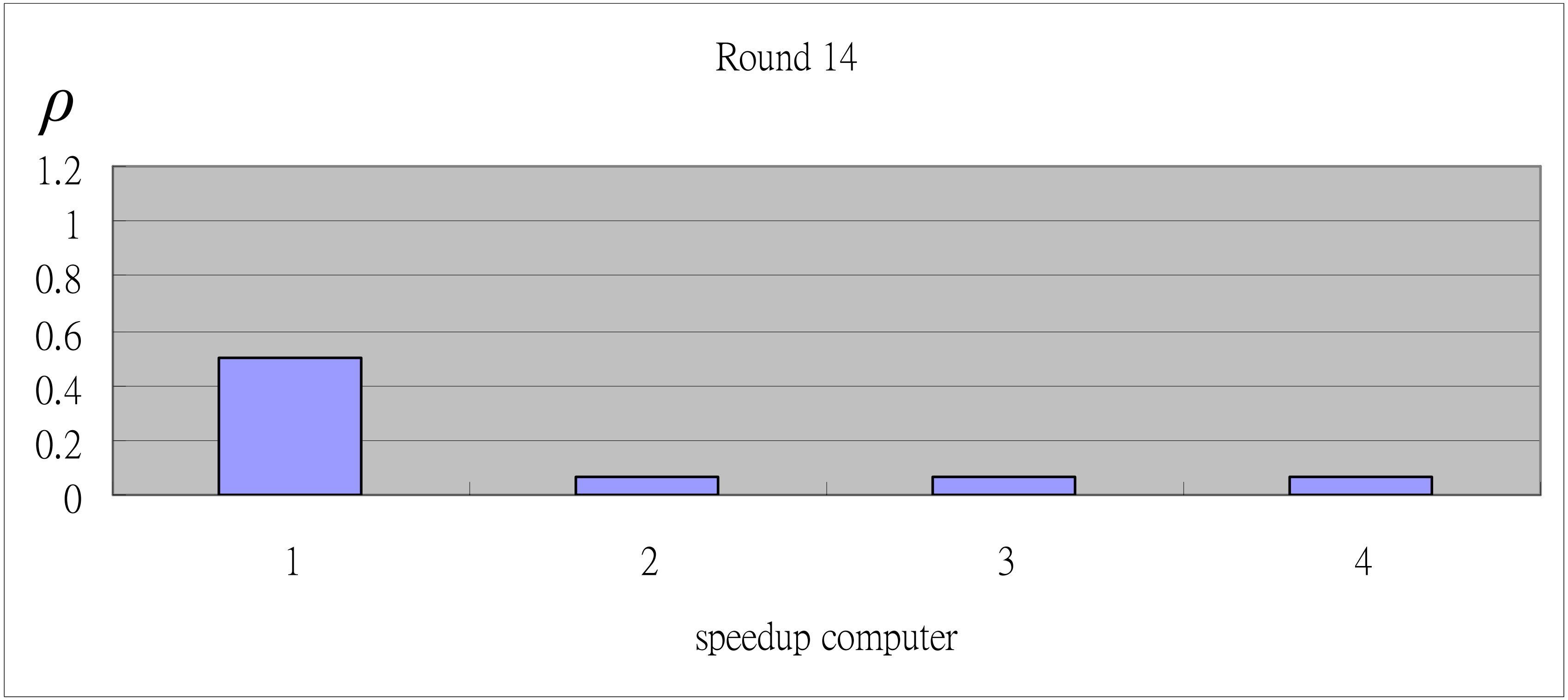
Proportional-Speedup in Action



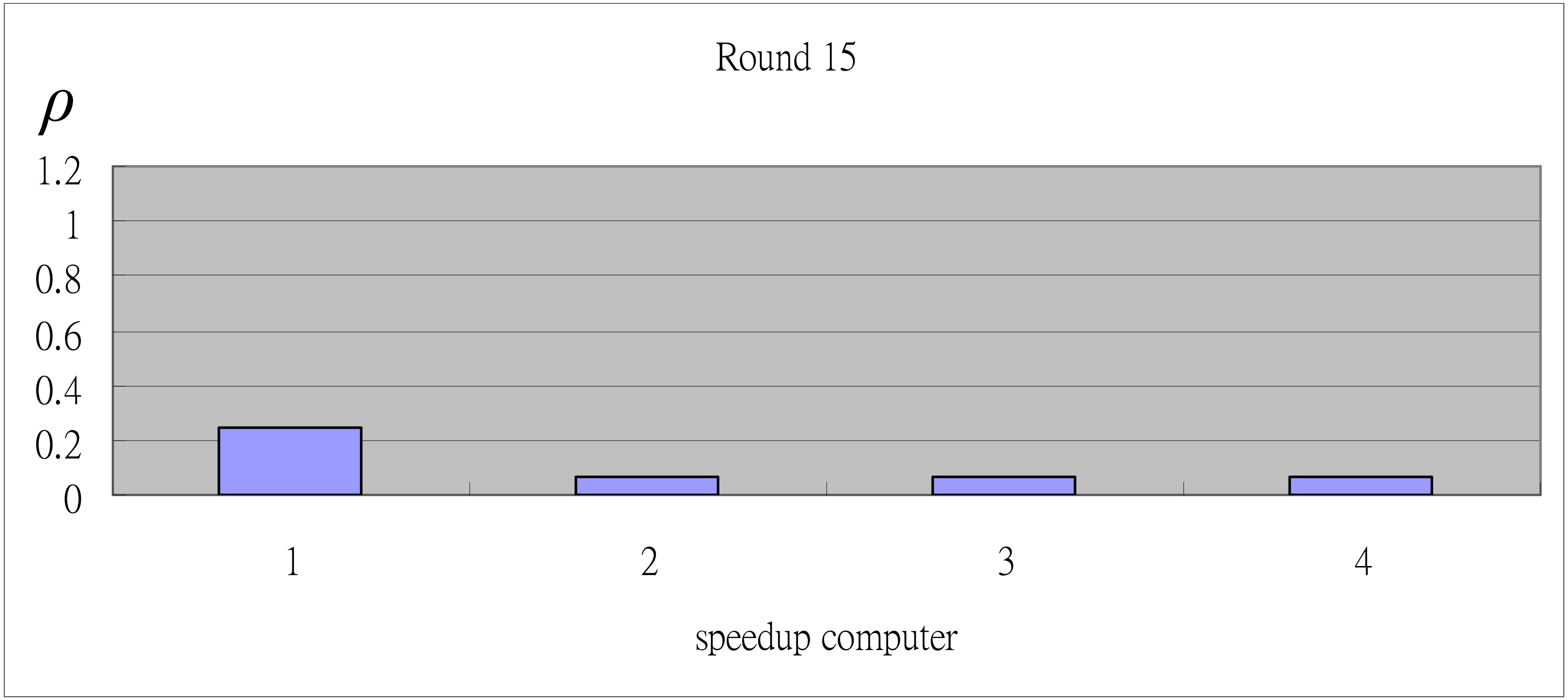
Proportional-Speedup in Action



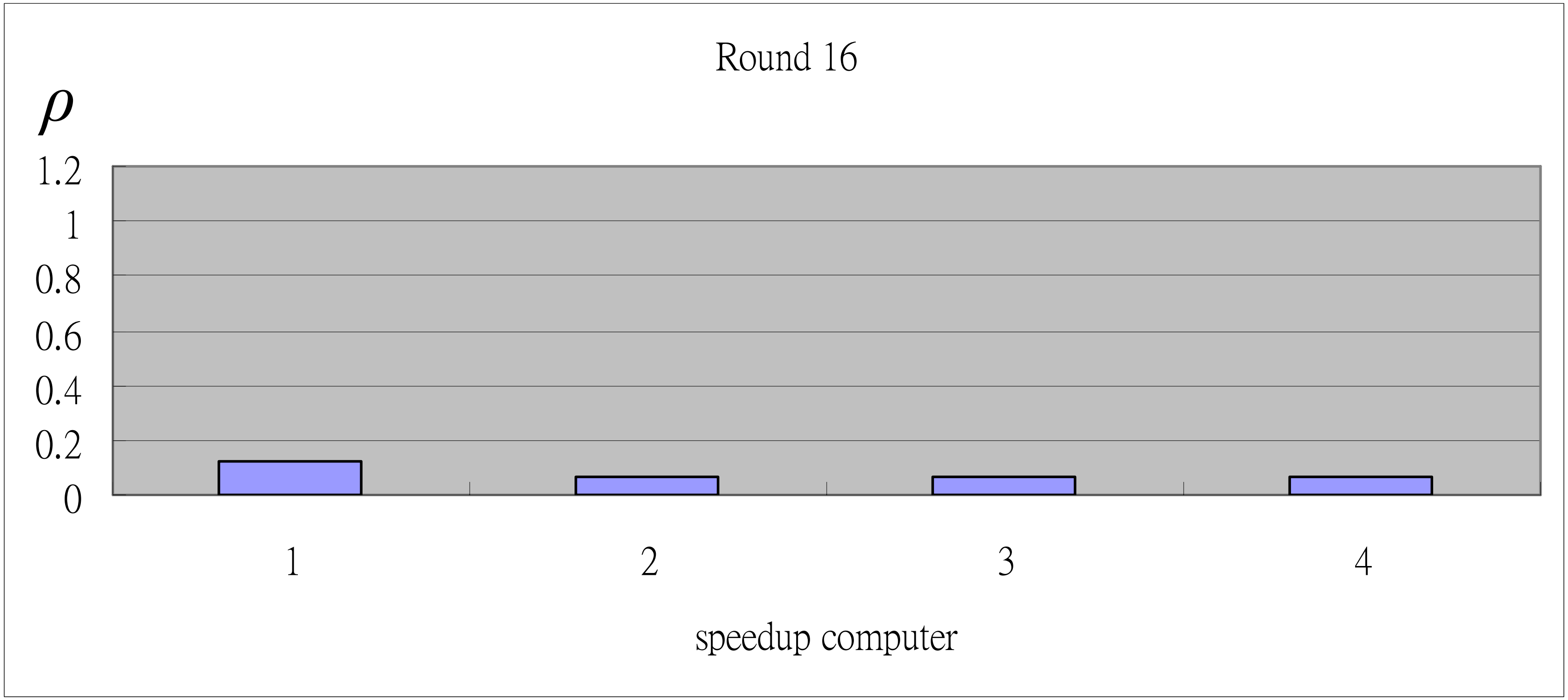
Proportional-Speedup in Action



Proportional-Speedup in Action



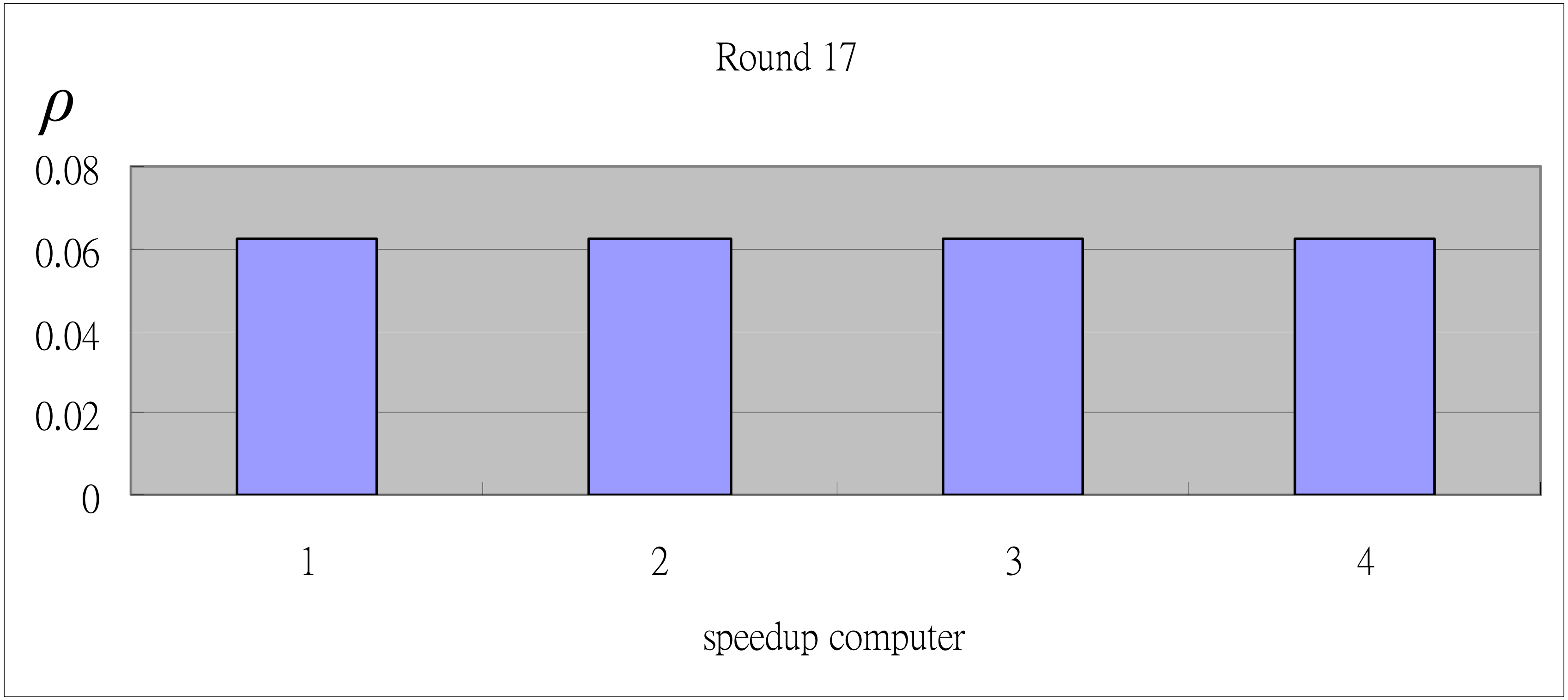
Proportional-Speedup in Action



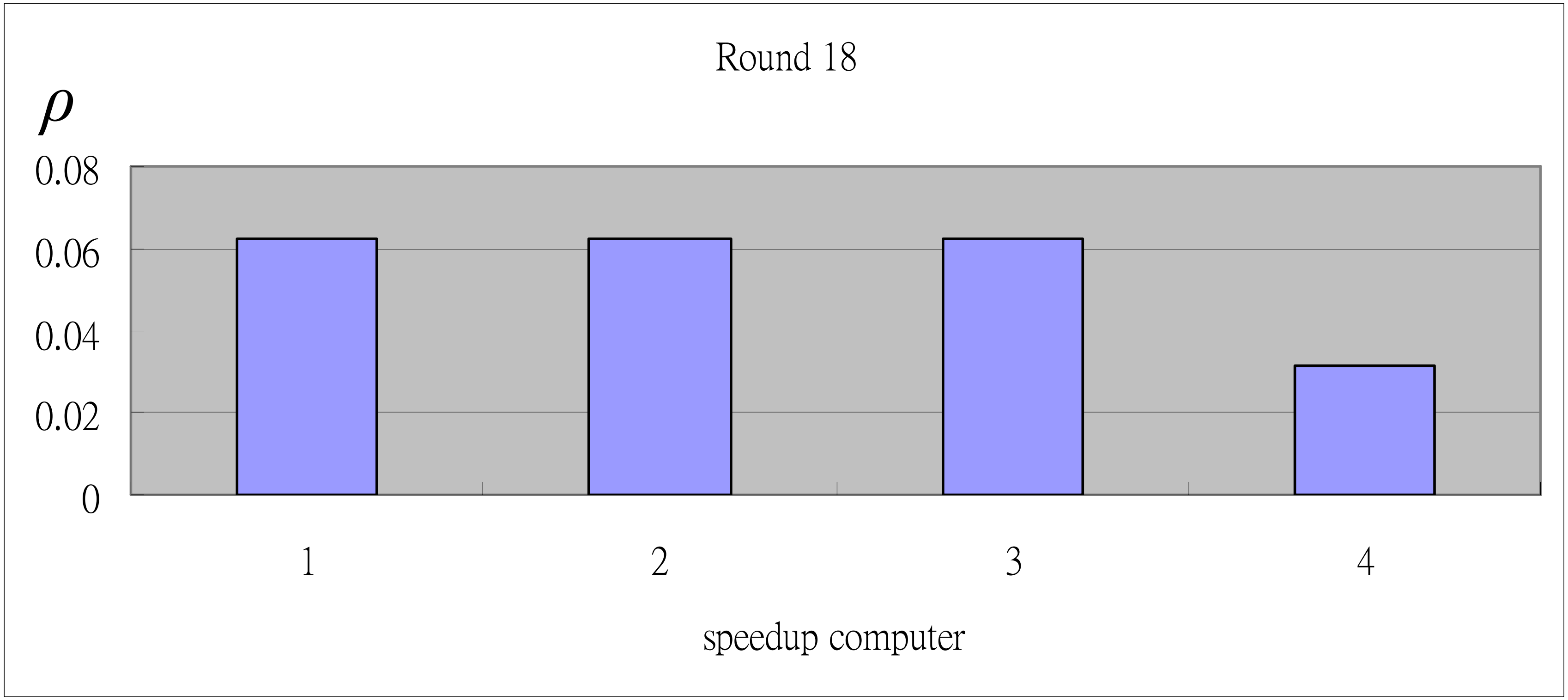
Proportional-Speedup in Action

- When all computers are very fast
  - It is more advantageous to speed up the slower one

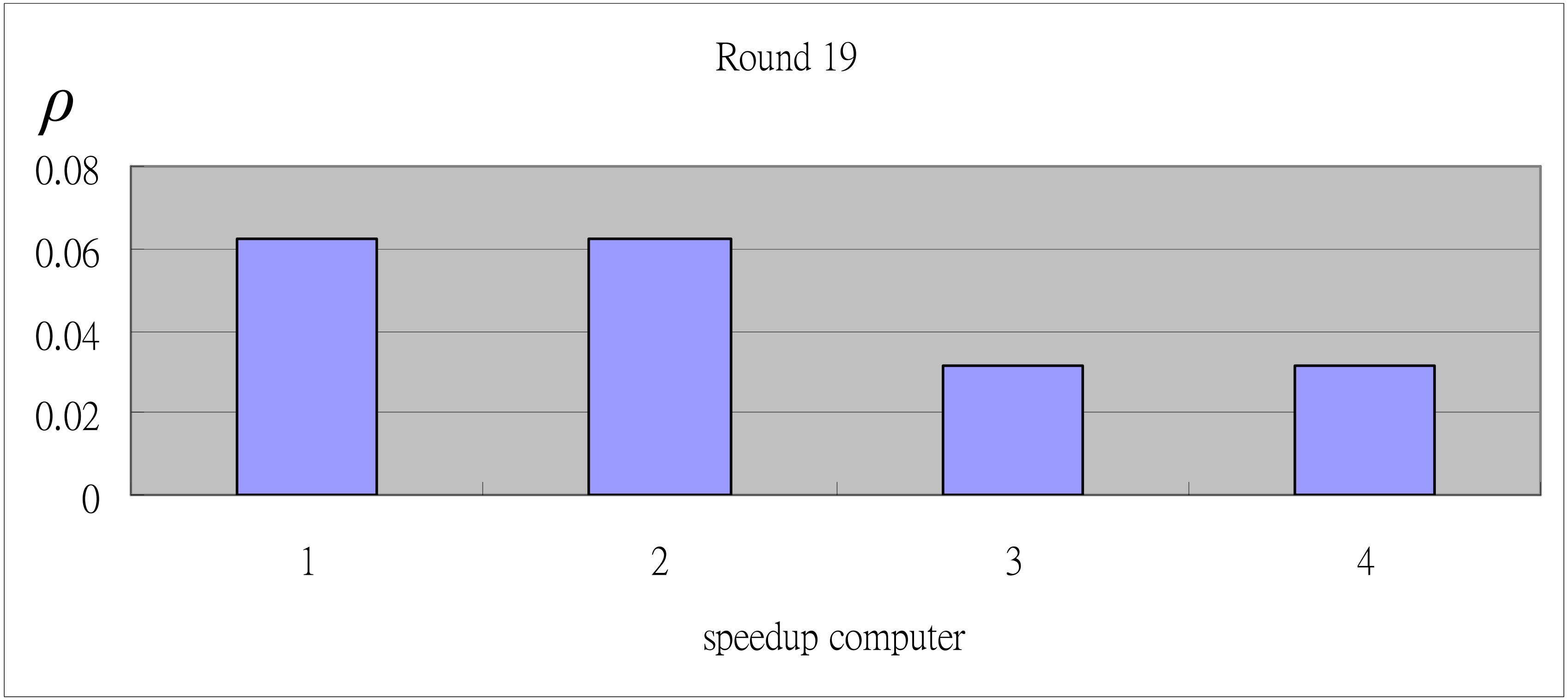
Proportional-Speedup in Action



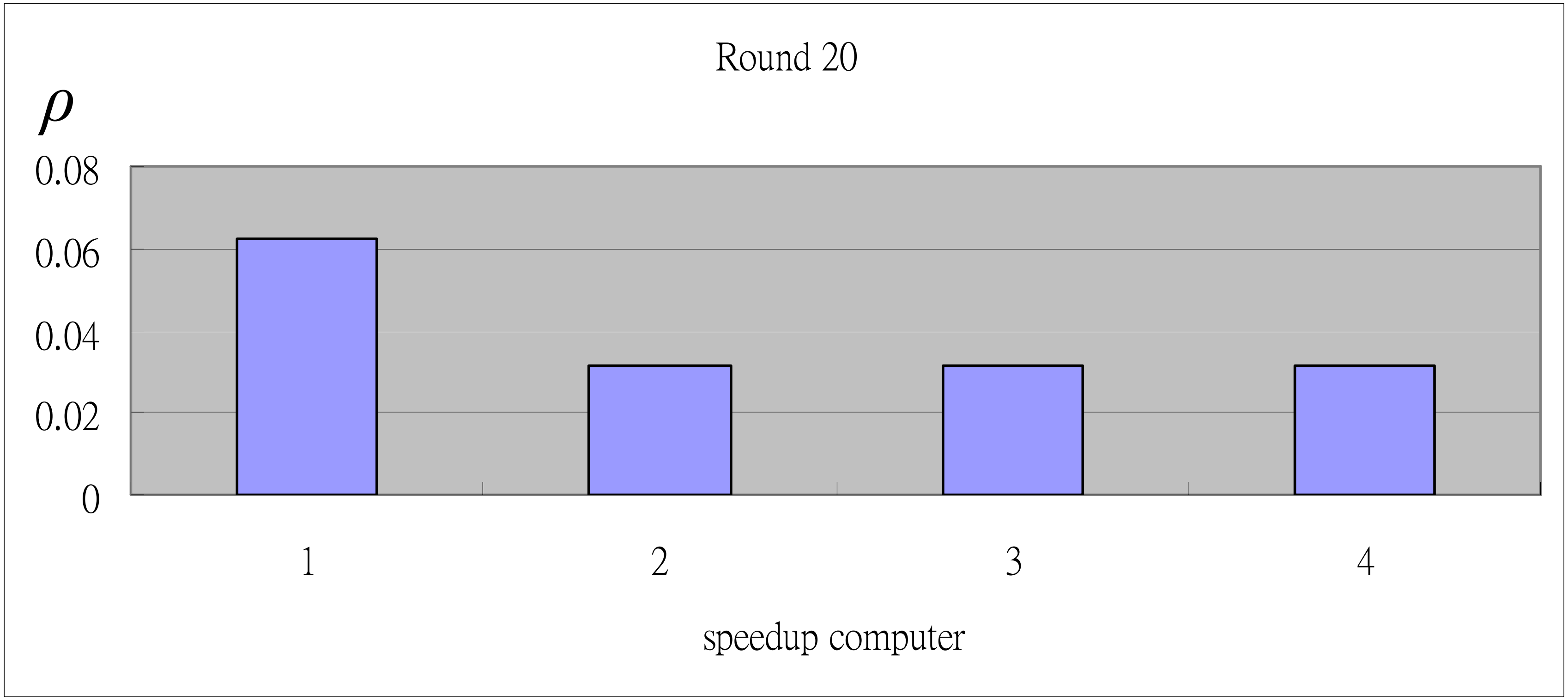
Proportional-Speedup in Action



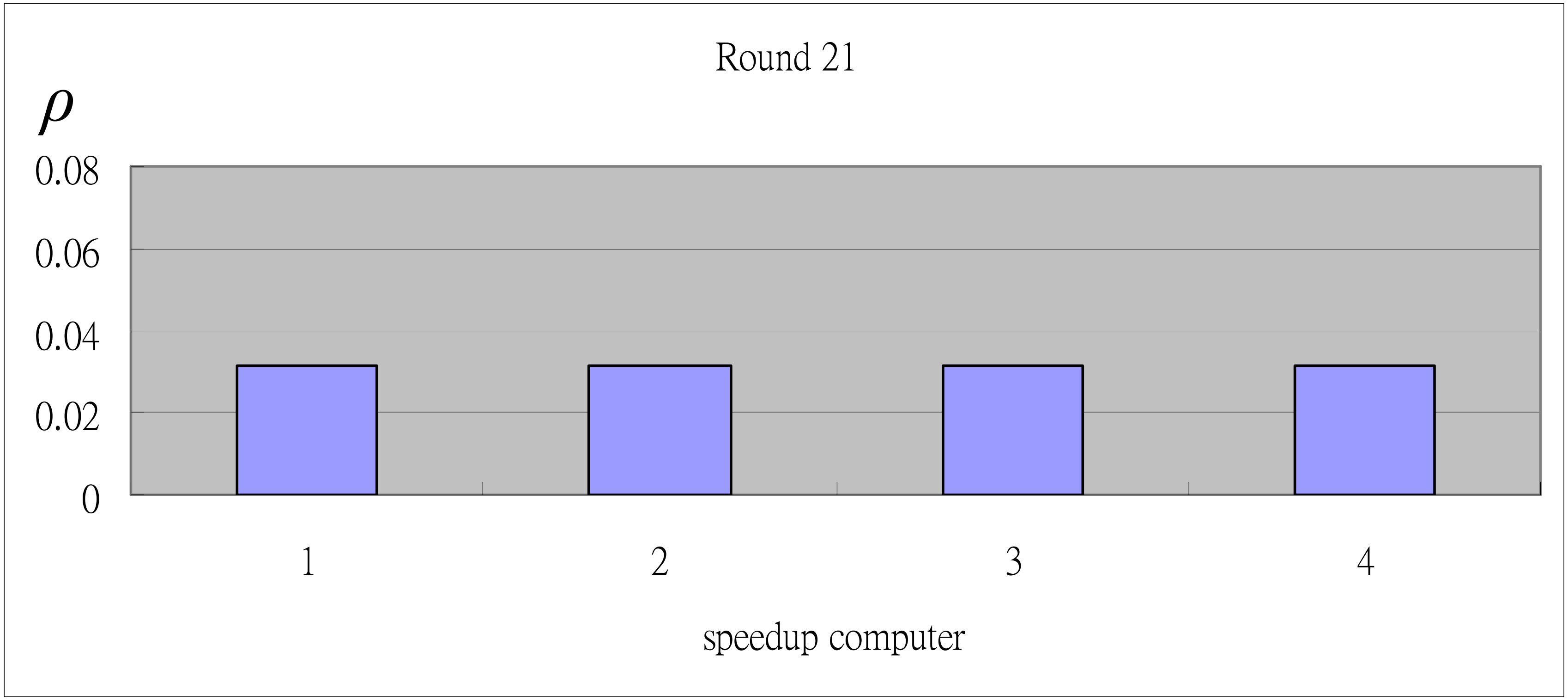
Proportional-Speedup in Action



Proportional-Speedup in Action



Proportional-Speedup in Action



Summary

- Two ways to measure computing power
  - the X function
  - the HECR value

Summary

- Two ways to measure computing power
  - the X function
  - the HECR value
- Standard deviation influences work production

Summary

- Two ways to measure computing power
  - the X function
  - the HECR value
- Standard deviation influences work production
- Speeding up a fast computer in a cluster is almost always more advantageous than speeding up a slower one

Thank you

Questions?

HECR values

	Number of Computers		
	8	16	32
Profile 1	0.362	0.297	0.251
Profile 2	0.216	0.116	0.061

Profile 1: $\rho_i = \frac{n - i + 1}{n}$ Profile 2: $\rho_i = \frac{1}{i}$

Recall: faster cluster has smaller HECR value

Avg. Speed vs. Std-Dev of Speed

16 computers' HECR		Std-Dev		
		0.2	0.1	0.05
Avg. Speed	0.75	0.671	0.723	0.768
	0.5	0.385	0.475	0.502
	0.25	0.110	0.194	0.239

Avg. Speed vs. Std-Dev of Speed

32 computers' HECR		Std-Dev		
		0.2	0.1	0.05
Avg. Speed	0.75	0.669	0.742	0.782
	0.5	0.380	0.478	0.502
	0.25	0.115	0.197	0.239