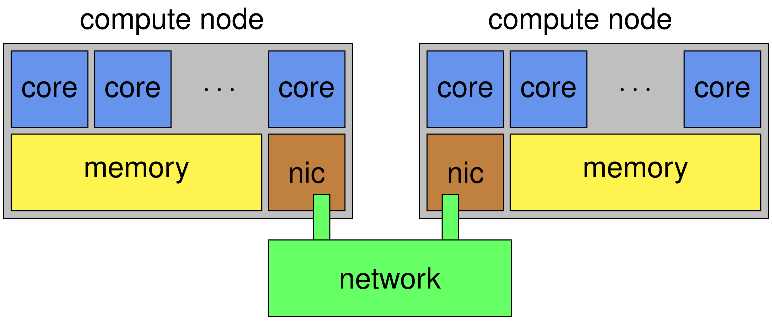
Basic concepts and assumptions



Hardware scheme

- *p* compute nodes
- *q* cores per compute node
- memory per compute node
- network interface per compute node
- nodes connected by a (fast) network

Hardware parameters

- p, number of nodes
- l_m, node performance [GFLOP/sec]
- b_{cl}, bandwidth between two nodes [GB/sec]
- q, number of cores per node
- l_c, core performance [GFLOP/sec]
- b_m, bandwidth between cores (caches) and memory [GB/sec]

Software parameters

- $\#op$, number of arithmetic operations per application problem
- $\#b$, number of bytes per application problem
- $\#x$, number of bytes exchanged between nodes per application problem

Performance

- t, computing time for a problem on a given system
- $l(q, p, \ldots) = \frac{\#op}{t}$, **performance** in terms of the parameters
- $\eta(q, p, \ldots) = \frac{l(q,p,\ldots)}{q\,p\,l_c}$, **efficiency** as a measure of how well the application uses its compute resources
- $S(q, p, \ldots) = \frac{l(q,p,\ldots)}{l_c}$, **speed-up** as a measure of how well the application scales with varying core and node numbers

efficiency and speed-up give insights

- what is the optimal number of cores and nodes for a given (or future) application on a given (or future) hardware ?
- use these optima as parameters for the batch system on a compute cluster to allocate the right number resources:

  determination of the right number of cores on clusters operated by sharing nodes between jobs

Multi-core node performance

computation time on one node

$$t \;\geq\; \frac{(\#op/q)}{l_c} + \frac{\#b}{b_m} = \frac{\#op}{q\,l_c}\left(1 + q \cdot \frac{l_c}{b_m}\frac{\#b}{\#op}\right) \qquad (1)$$

assume communication with the shared memory and the computation phases *do not overlap*

define $a = \frac{\#op}{\#b}$, $a^\star = \frac{l_c}{b_m}$

$$t \;\geq\; \frac{\#op}{q\,l_c}\left(1 + q\,\frac{a^\star}{a}\right)$$

a "software and problem demand", $a^\star$ "hardware capabilities" performance of a node $l_m = \#op/t$

$$l_m \le q\, l_c \frac{1}{1 + q \cdot \frac{a^\star}{a}} = q\, l_c \frac{\frac{a/a^\star}{q}}{1 + \frac{a/a^\star}{q}}$$

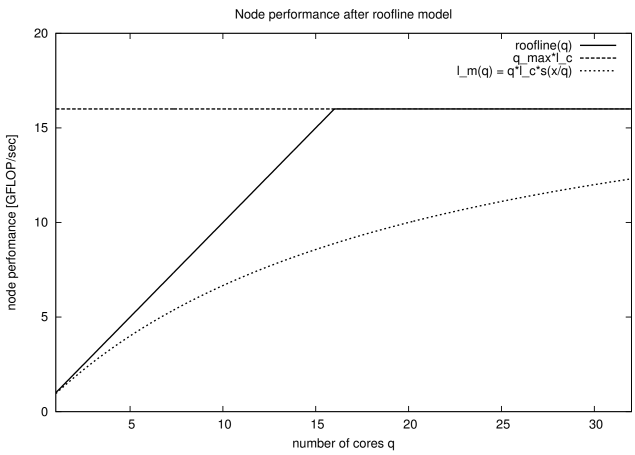
with dimensionless operational intensity $x = a/a^\star$

$$l_m(q) \le q\, l_c \frac{\frac{x}{q}}{1 + \frac{x}{q}}.$$

with Hockney $s(z) = \frac{z}{1+z}$, overlapping $s(z) = \min(1, z)$

$$l_m(q) \le q\, l_c\, s(\frac{x}{q}). \qquad (2)$$

Performance of multi-cores in the roofline model for $l_c = 0.5$ GFLOP/sec and operational intensity $x = 20$



Efficiency of multi-cores depending on x and q

$$\eta_m(x,q) = \frac{l_m(q)}{q\, l_m(1)} \quad \leq \quad \frac{1}{q}\,\frac{1+x}{1+\frac{x}{q}}$$

Amdahl's law: along x-axis, Gustafson's law: along y-axis