Architecting System Performance; Resource Management

by *Gerrit Muller*   TNO-ESI, University of South-Eastern Norway]

e-mail: `gaudisite@gmail.com`

`www.gaudisite.nl`
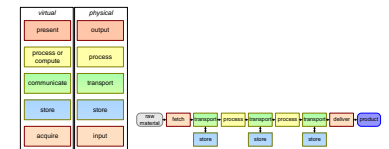
Abstract

The management of the resources largely determines system performance. This document discusses concepts related to resource management, such as caching, concurrency, and scheduling.
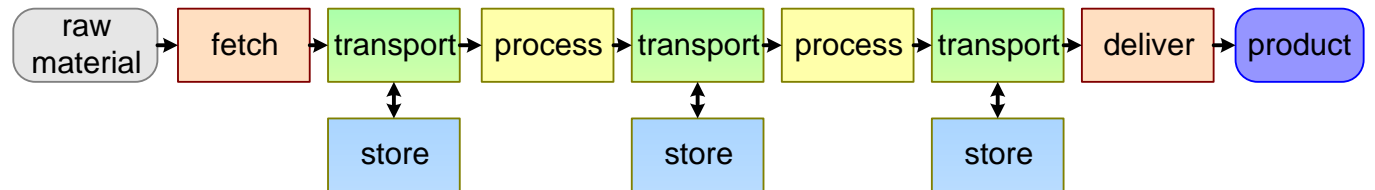
August 21, 2020
status: preliminary
draft
version: 0.1

Generic Resource Model

virtual	physical
present	output
process or compute	process
communicate	transport
store	store
acquire	input

raw material → fetch → transport → process → transport → process → transport → deliver → product

transport ↕ store

transport ↕ store

transport ↕ store

Design Considerations for Resource Management

Performance depends on resource utilization and management.

The design of the logistics, how does EMI[1] flow through the resources, is critical.

Critical design aspects are:

- concurrency (parallelism, pipelining)

- granularity of EMI

- scheduling (allocation of resources)

[1]Energy Material Information

Granularity as Key Design Choice

unit of buffering == or <> unit of synchronization == or <> unit of processing == or <> unit of I/O

video frame

——video line——

• pixel
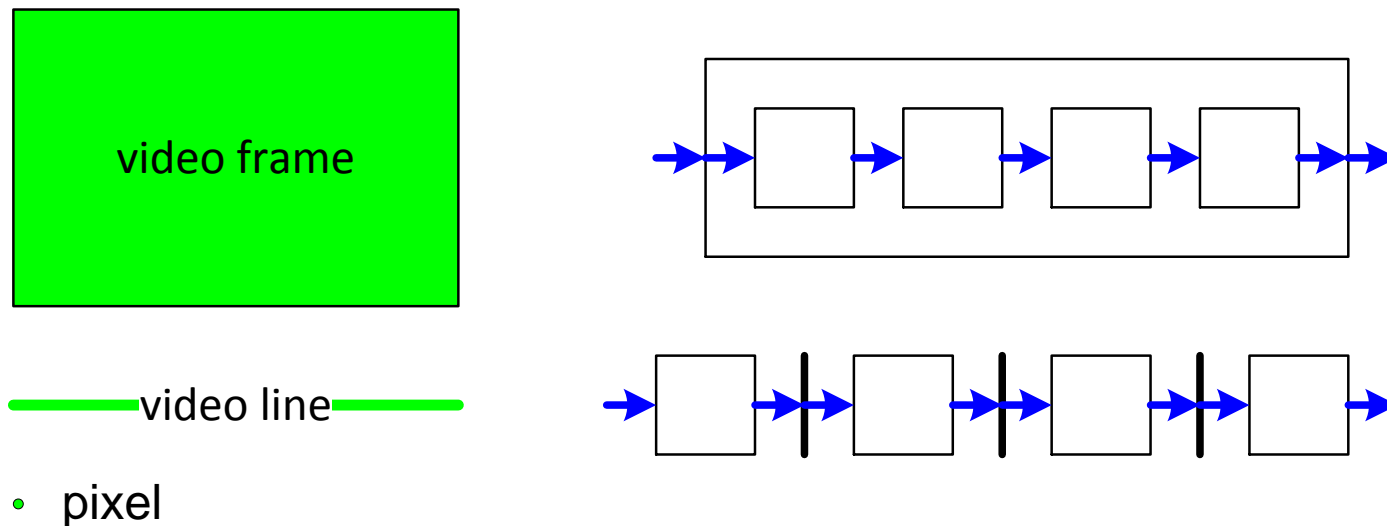
fine grain:
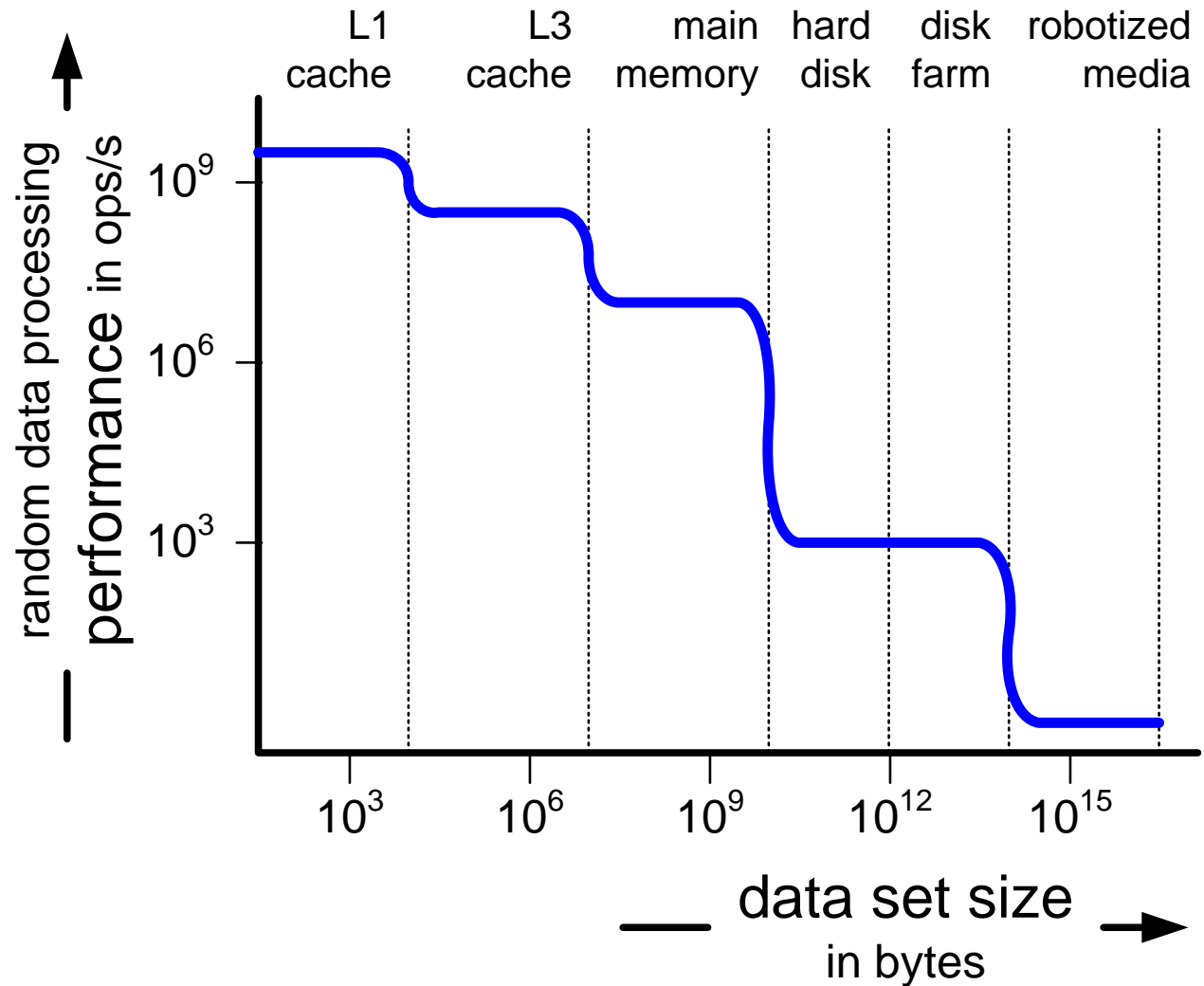flexible
high overhead

coarse grain:
rigid
low overhead

USN ESI

Size versus Performance Trade off
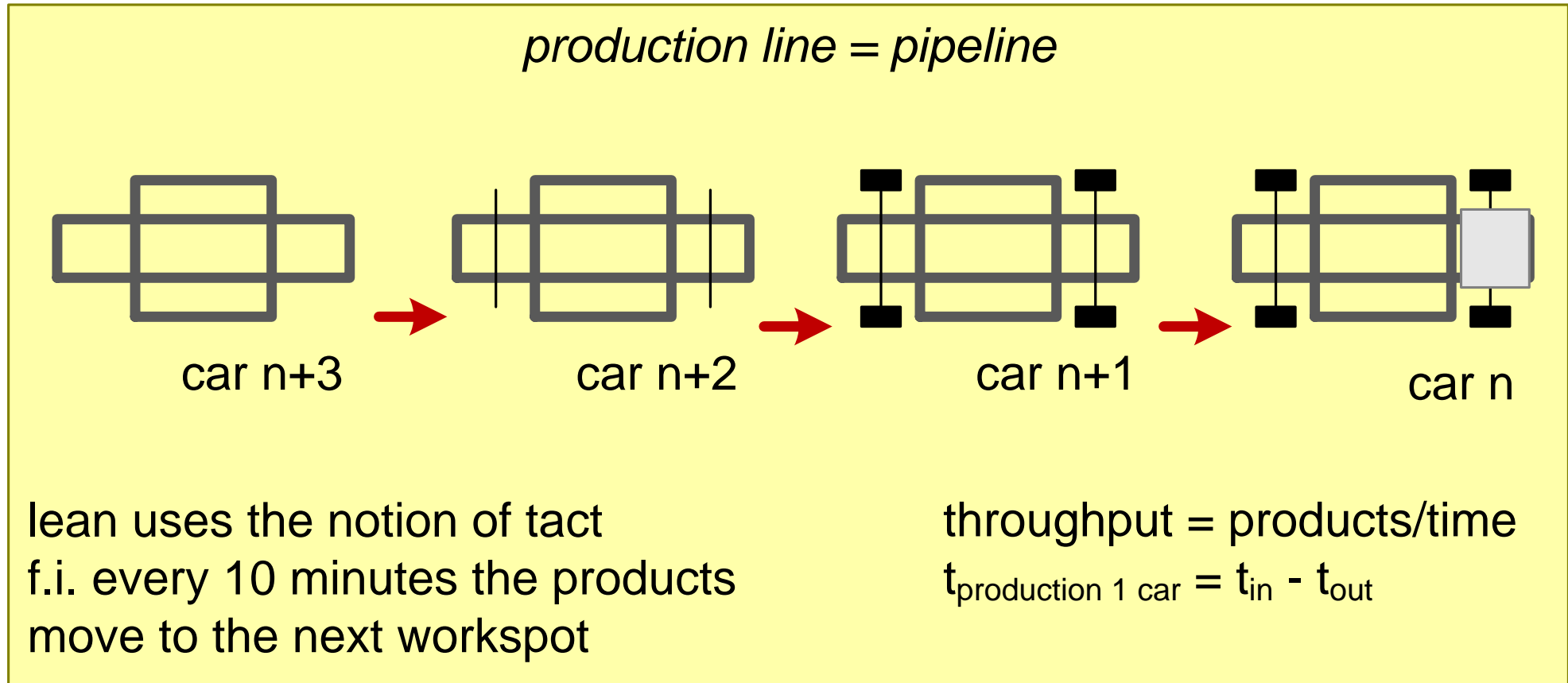
small capacity

fast technology
small
expensive

large capacity

slow technology
large
low cost

staircase effect:
performance and
size are non-linear
with thresholds

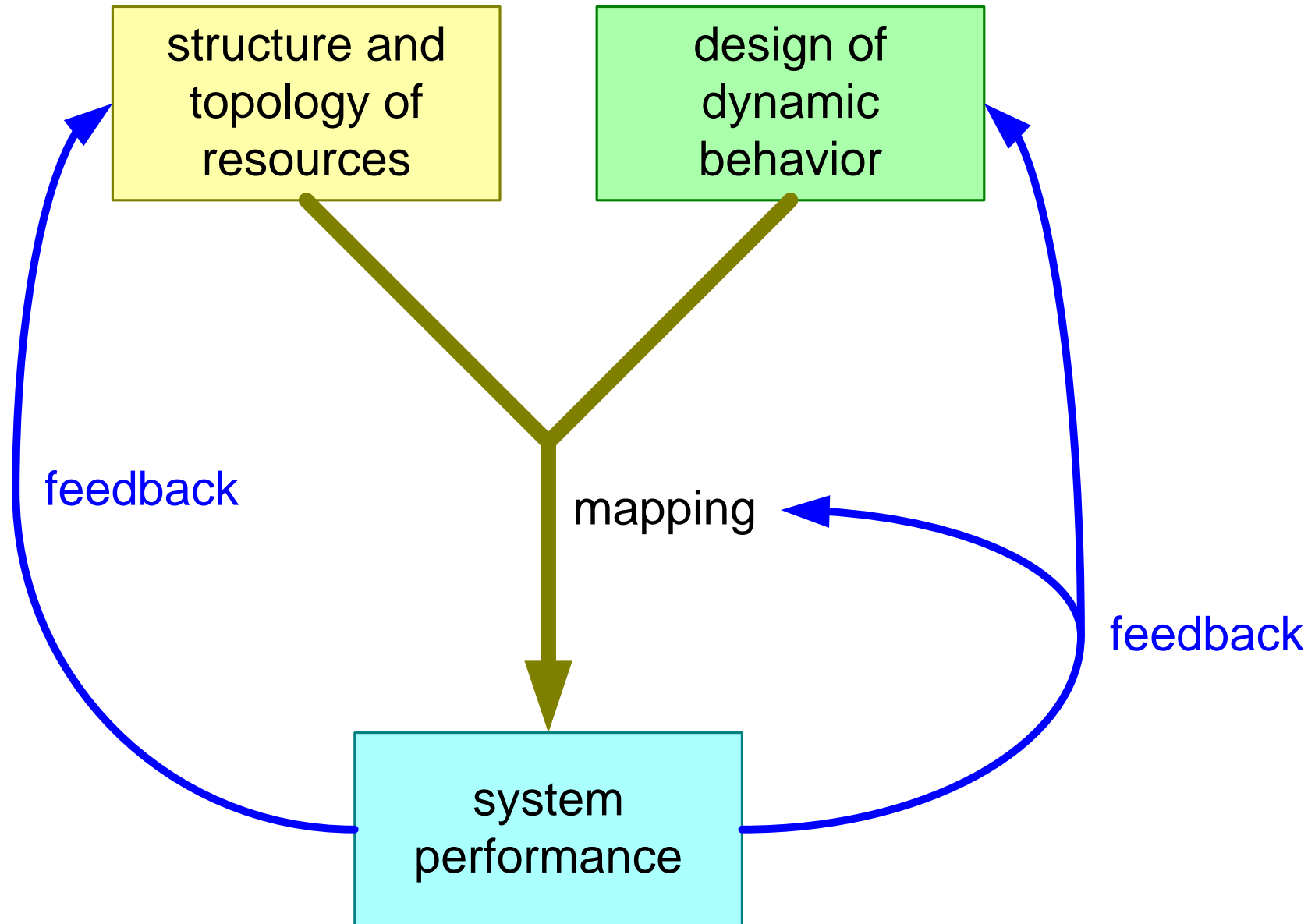random data processing performance in ops/s

L1
cache

L3
cache

main
memory

hard
disk

disk
farm

robotized
media

10^9

10^6

10^3

10^3 10^6 10^9 10^{12} 10^{15}

___ data set size
in bytes

example data storage technology

USN   ESI

Pipeline pattern

production line = pipeline

car n+3   →   car n+2   →   car n+1   →   car n

lean uses the notion of tact
f.i. every 10 minutes the products
move to the next workspot

throughput = products/time

$t_{\text{production 1 car}} = t_{in} - t_{out}$

USN   ESI

Y-chart Pattern

Performance Pitfalls and Resource Management

Overhead (control, handling)

Starvation (underrun)

Saturation/stagnation (overrun)

Variation (duration, quality)

Serialization

Interference with other work

Unnecessary conversions or adaptations

USN   ESI