Architecting System Performance; Resource Management

by *Gerrit Muller* TNO-ESI, University College of South East Norway
e-mail: gaudisite@gmail.com
www.gaudisite.nl

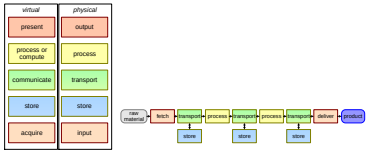
Abstract

The management of the resources largely determines system performance. This document discusses concepts related to resource management, such as caching, concurrency, and scheduling.

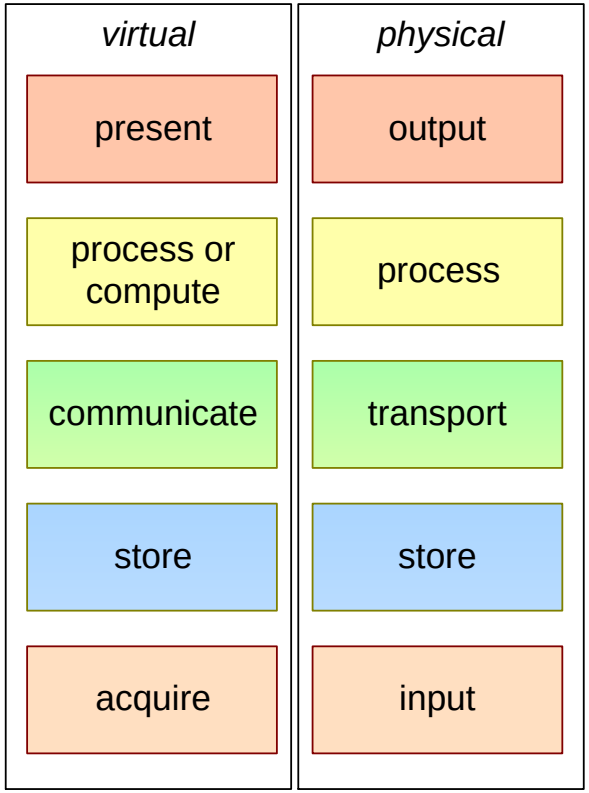
Distribution

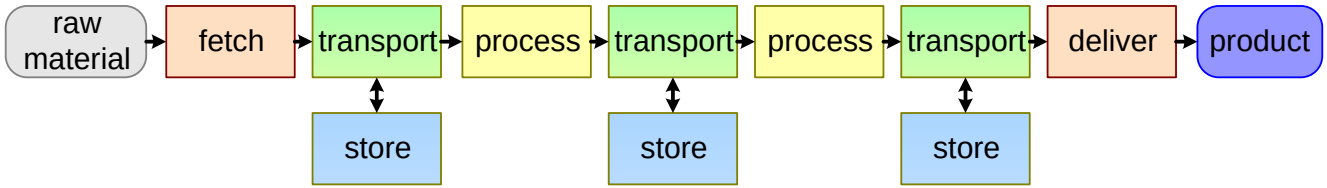
This article or presentation is written as part of the Gaudí project. The Gaudí project philosophy is to improve by obtaining frequent feedback. Frequent feedback is pursued by an open creation process. This document is published as intermediate or nearly mature version to get feedback. Further distribution is allowed as long as the document remains complete and unchanged.

October 3, 2016
status: preliminary draft
version: 0.1



Generic Resource Model





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.

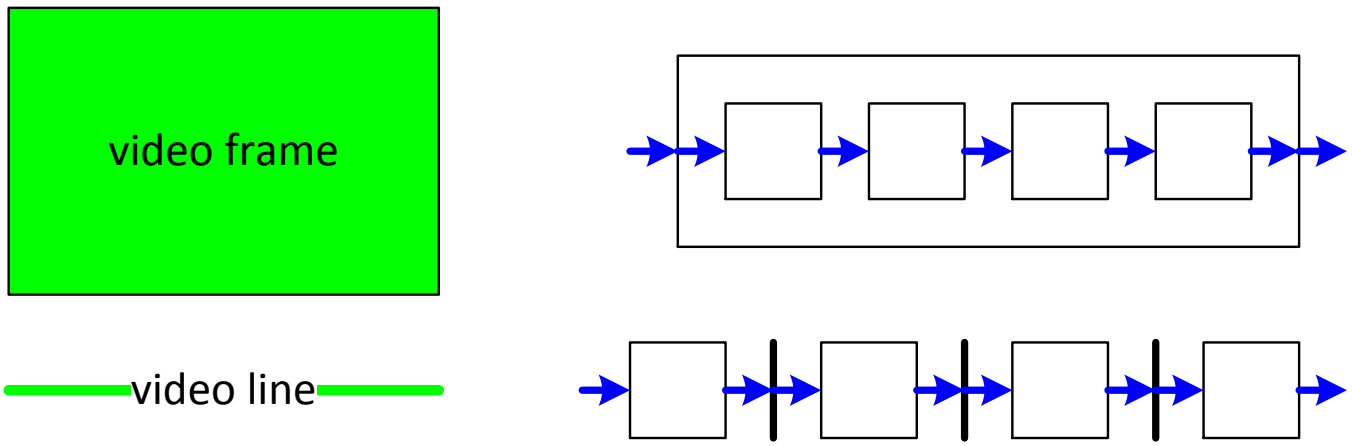
Critical design aspects are:

- 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*



- pixel

fine grain:	*coarse grain:*
flexible	rigid
high overhead	low overhead

Size versus Performance Trade off

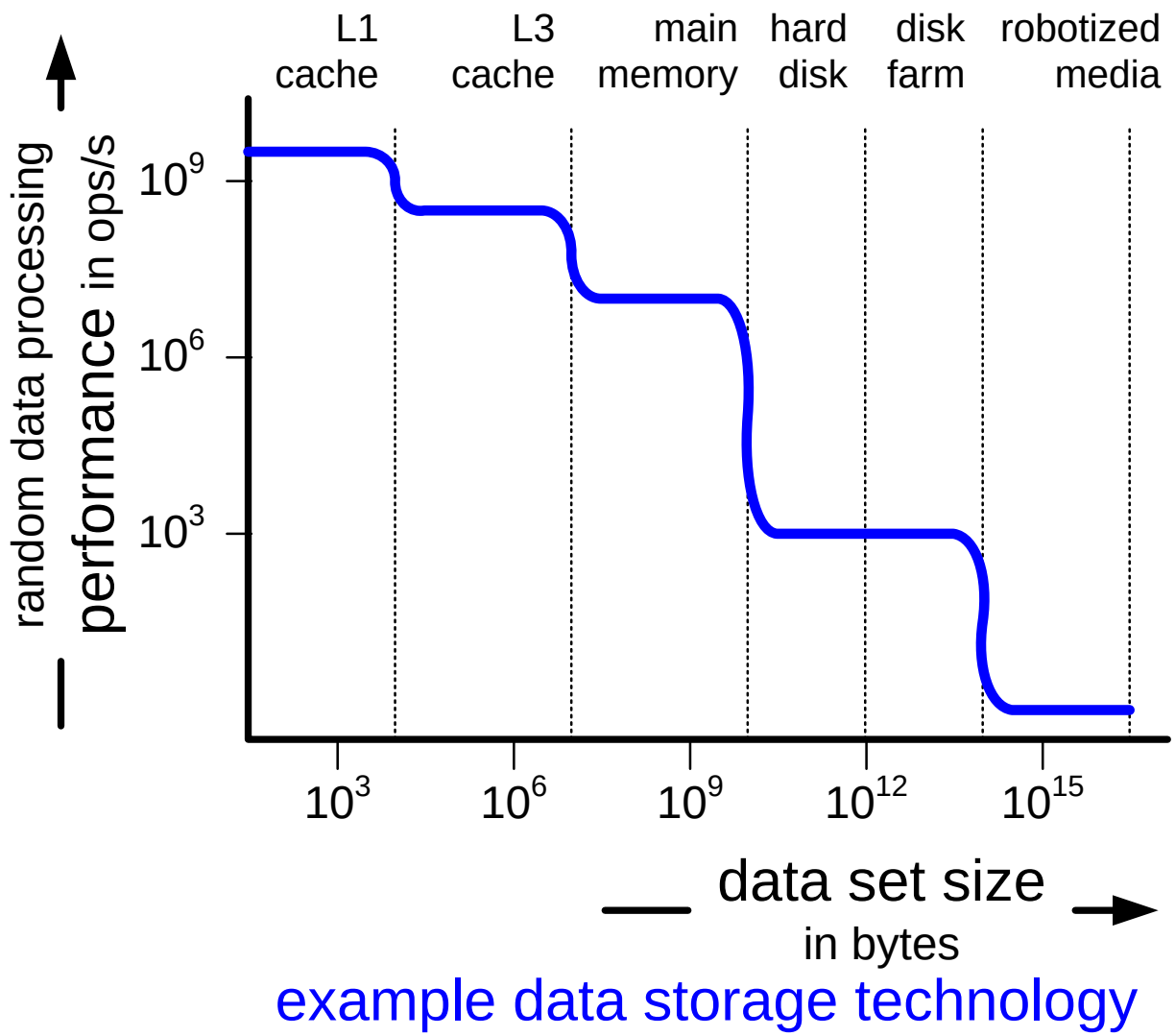
small capacity

fast technology
small
expensive

large capacity

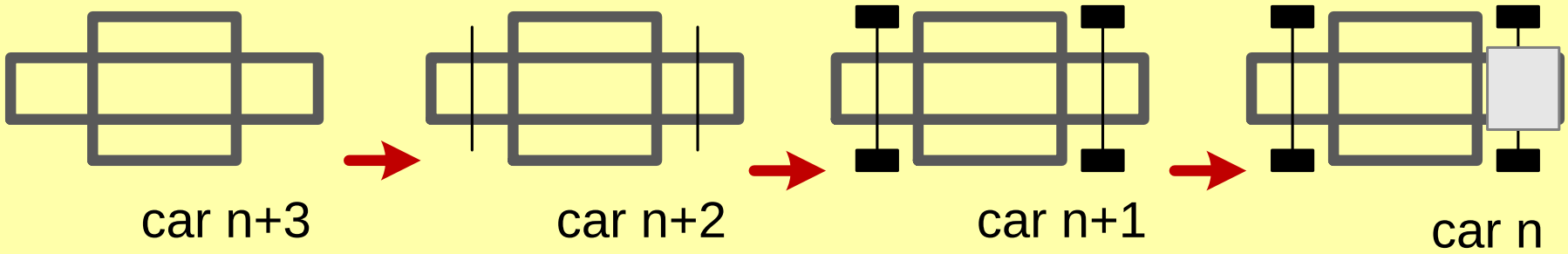
slow technology
large
low cost

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



Pipeline pattern

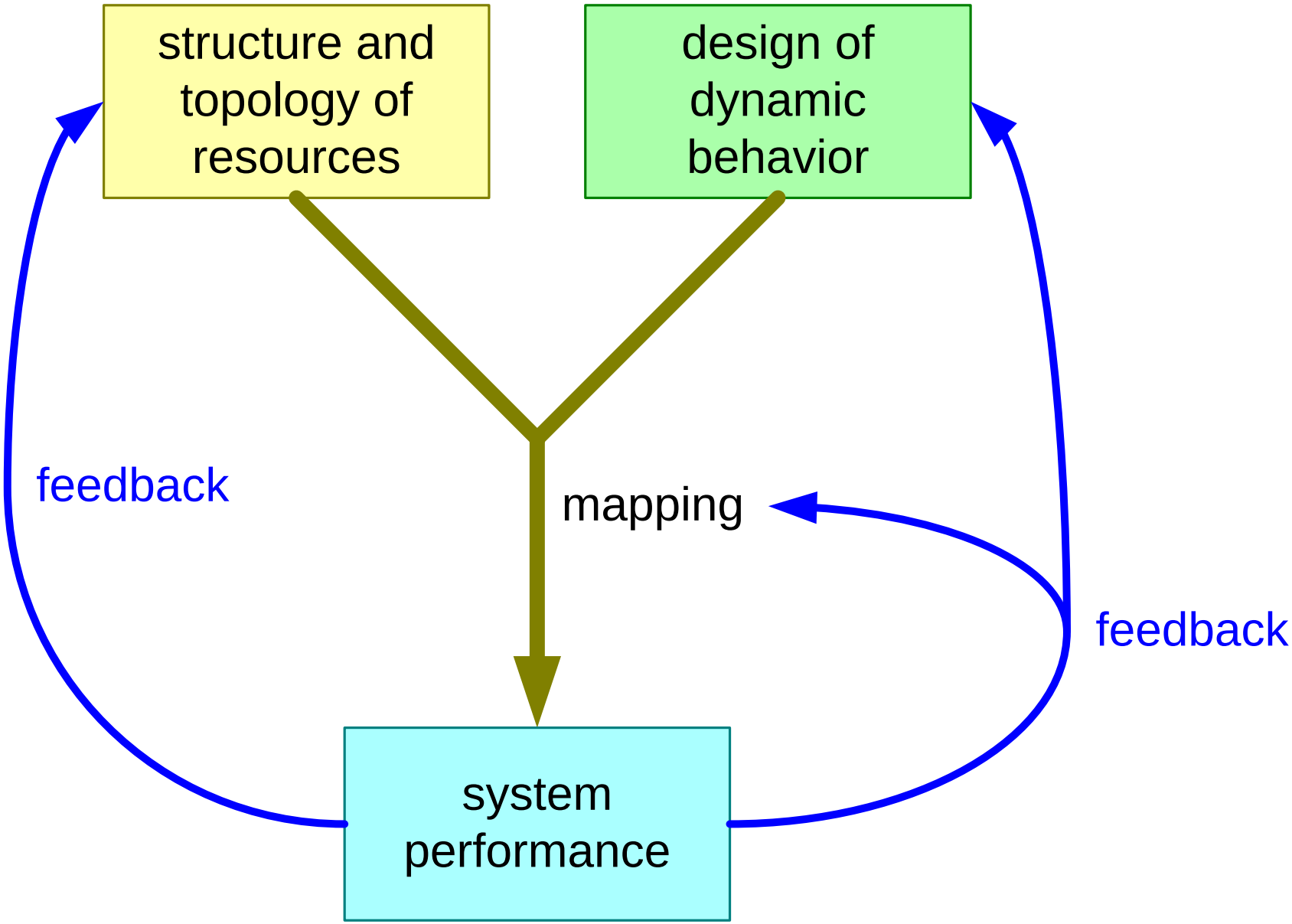
production line = pipeline



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

throughput = products/time
$t_{production\ 1\ car} = t_{in} - t_{out}$

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