Cloud Computing Uncertainty:
opportunities and challenges

Andrei Tchernykh

CICESE Research Center,
Ensenada, Baja California, México
chernykh@cicese.mx
http://usuario.cicese.mx/~chernykh/

PPAM 2017

12th International Conference on Parallel Processinng and Applied Mathematics

Lyblin, Poland. September 12, 2017

Towards Understanding Uncertainty in Cloud Computing Resource Provisioning

Andrei Tchernykh	CICESE Research Center, Mexico
Uwe Schwiegelshohn	University of Dortmund, Germany
El-ghazali Talbi	University of Lille, France
Vassil Alexandrov	Barcelona Supercomputing Centre, Spain
Michael Babenko	North Caucasus Federal University, Russia

ICCS-SPU 2015. Procedia Computer Science, Elsevier, 2015

Journal of Computational Science. Elsevier, 2016

Cloud Computing

- Illusion of infinite computing resources on demand
- Pay as you go
- Shared resources (dynamic performance changing)
- Massive, diverse, incomplete, heterogeneous data
- Big data
- Hybrid infrastructure
- Scalability and flexibility (dynamic elasticity)
- Privacy, security and availability concerns
- Virtualization, loosely coupling application to the infrastructure
- Resource provisioning time variation
- Inaccuracy of application runtimes
- Variation in data transmission
- Workload uncertainty

Cloud Computing Uncertainty

Long-term uncertainty

- It is poorly understood
- Inadvertent factors can influence its behavior

Retrospective uncertainty

- Lack of information about its behavior in the past

Decision uncertainty

- Impossibility of predicting the exact results of decisions

Stochastic uncertainty

- There is no reliable statistical information
- Stochastic nature requires verification
- Inherent randomness, intrinsic variability

Cloud Computing Uncertainty

Constraint uncertainty

- There is no ambiguity when choosing solutions
- There exists a possibly infinite number of Pareto optimal solutions

Participant uncertainty

- Conflict of main stakeholders: cloud providers, users and administrators.
- Own preferences, incomplete, inaccurate information about the motives and behavior of opposing parties

Goal uncertainty

- Inability to select one goal
- Conflicts in building multi objective optimization model
- Competing interests

Condition uncertainty

- Failure or a complete lack of information about the conditions under which decisions are made

Cloud Computing Uncertainty

Parameter uncertainty

- Partial or complete ignorance of the conditions
- Exact values are unknown and cannot be controlled
- Cannot be exactly inferred by statistical methods.

Uncertainty of user perception

- Privacy concerns
- Loss control over personal information
- Privacy and information security

User submission behavior

- Correlation between response time of a job and the time until the next job is submitted
- Lock-in effects

User Satisfaction

- React on a short-term,
- Abandonment rate
- Leaving the service after a start-up delay of 2000ms losing 5.8% of users for each additional second.

Cloud Computing Uncertainty

Irrereduced uncertainty

- Cannot be reduced by additional data

Aleatory Uncertainty

- Industrial processes
- Inherent randomness
- intrinsic variability

Epistemic Uncertainty

- Lack of knowledge about appropriate values to use (just belief)

Reducible uncertainty

- Can be reduced through increased understanding (research)

Challenges

The question is:

How to deliver scalable and robust behavior
under uncertainties
and
specific constraints, such as budgets, QoS, SLA,
energy costs; etc.

Research directions

Cloud Computing with Different Service Levels

Non-clairvoyant knowledge-free scheduling

Adaptive Admissible Allocation

Modeling applications with communications and uncertainty

Towards Secure data storage

- Uncertainty of storage system
- Multi-Cloud environment

Scheduling with Uncertainty

- User Run Time Estimates
- Game Scheduling
- Runtime Uncertainty

Uncertainty in urban public transport

Resource Contention

Adaptive Energy-Aware Resource Allocation

Adaptive VoIP Service for Cloud Infrastructure

Modeling applications with communication uncertainty

Dzmitry Kliazovich
Johnatan E. Pecero
Pascal Bouvry

University of Luxembourg, Luxembourg

Andrei Tchernykh

CICESE Research Center, Mexico

Samee U. Khan

North Dakota State University, U.S.A

Albert Y. Zomaya

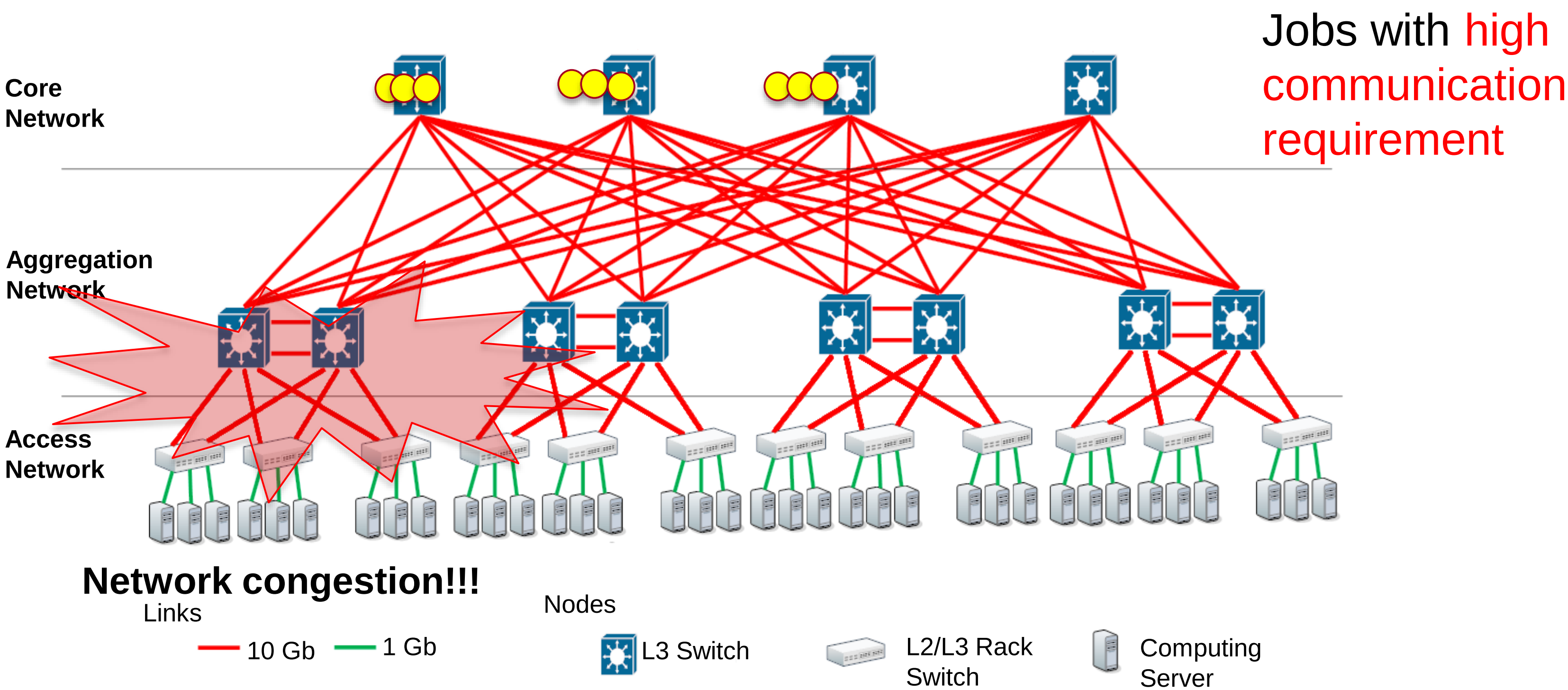
University of Sydney, Australia

- **IEEE CLOUD 2013** - IEEE 6th International Conference on Cloud Computing.
- *Journal of Grid Computing* , Springer, 2015

Consolidation in Data Centers

Most of energy saving is due to consolidation procedures.

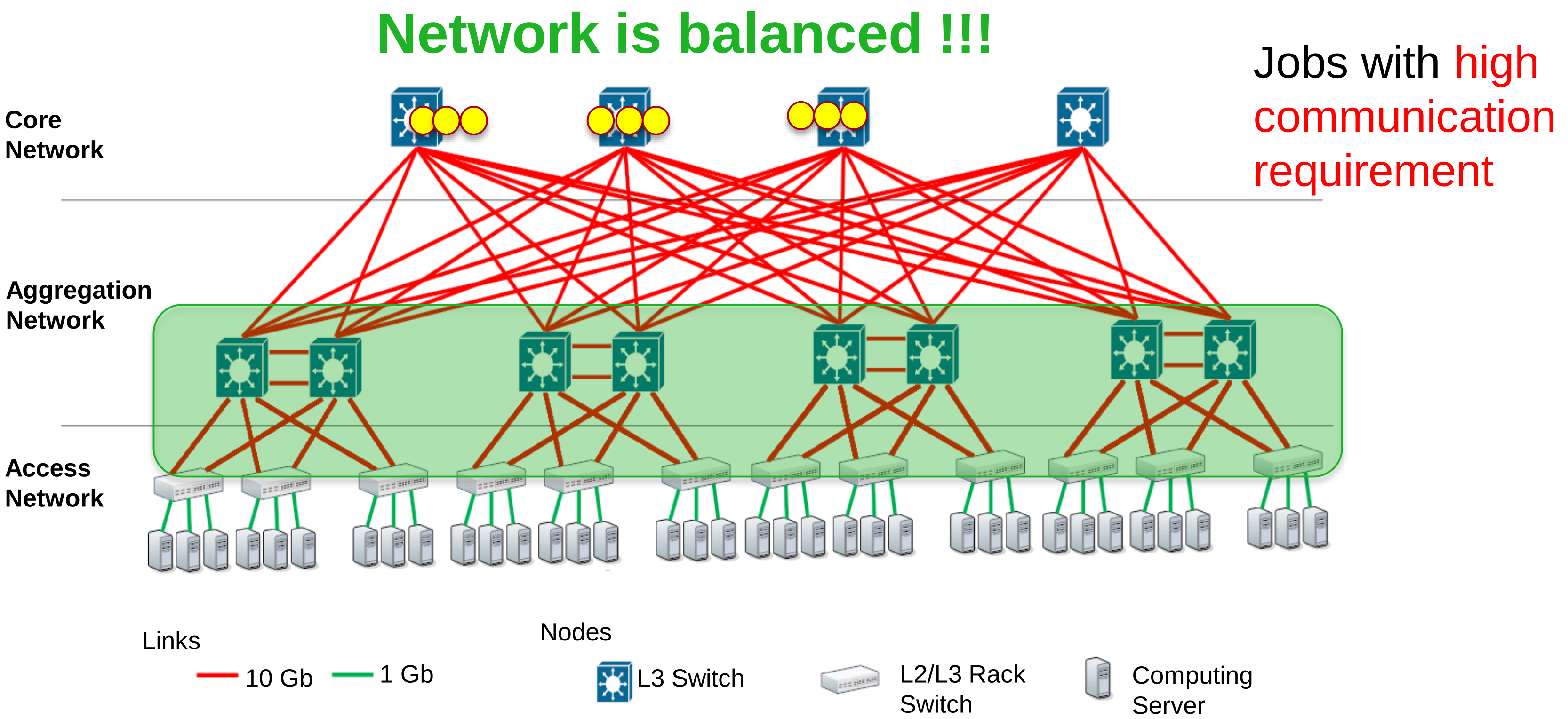
Increase number of server that can be put into “sleep” mode.



by Dzmitry Kliazovich

Consolidation with data balancing

Scheduler should tradeoff workload concentration with load balancing of network traffic



by Dzmitry Kliazovich

Modeling Applications

How to model applications with communication processes?

Two known approaches

– **CU-DAG** Communication-unaware model

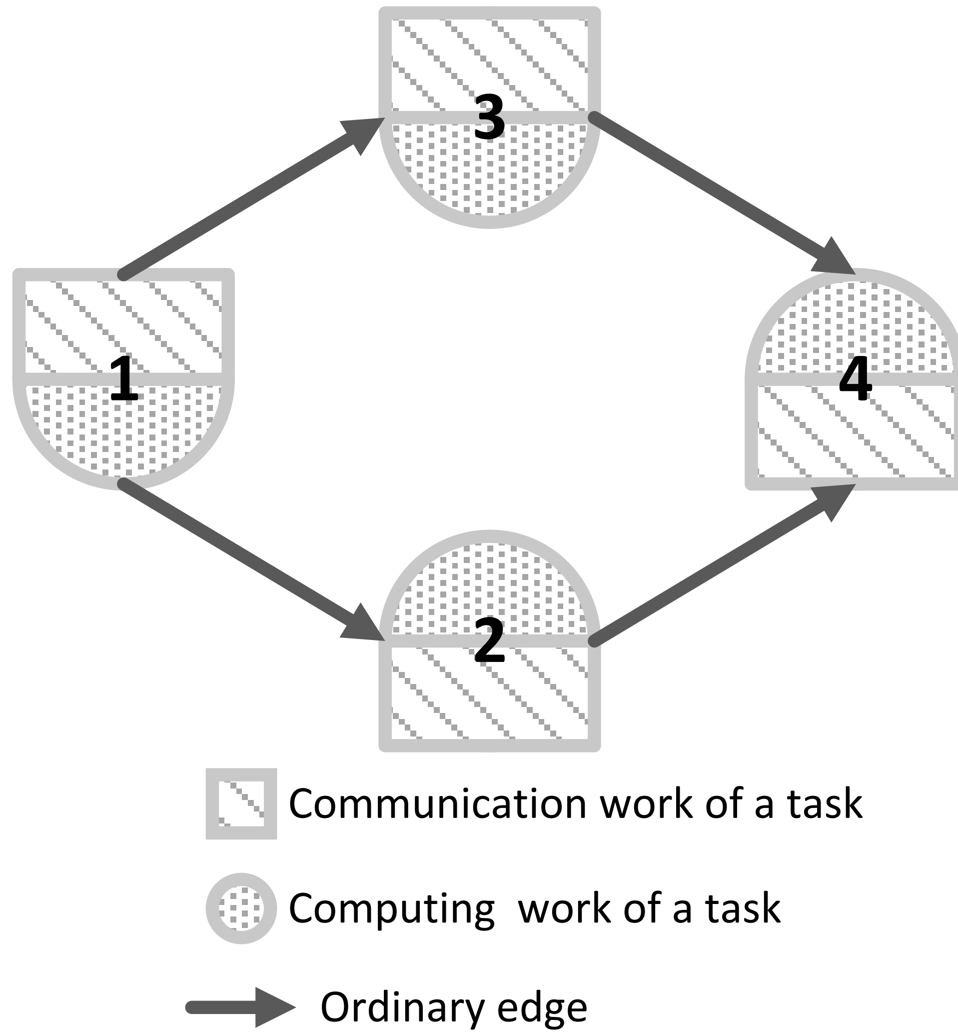
– **EB-DAG** Edges-based model

New approach

- **CA-DAG** **Communication-aware model**

Communication-unaware model

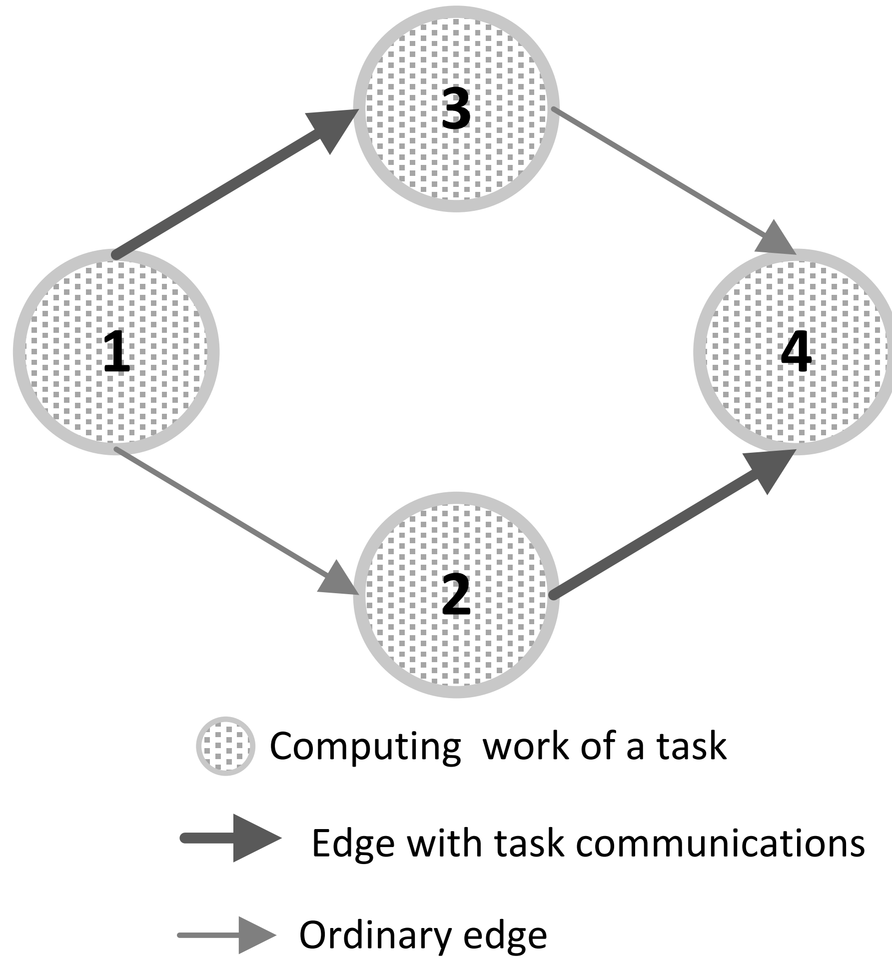
- vertex represents both computing and communication
- Edges: dependencies



- Main drawback
  - Difficult to make separate scheduling decisions

Edge-based model

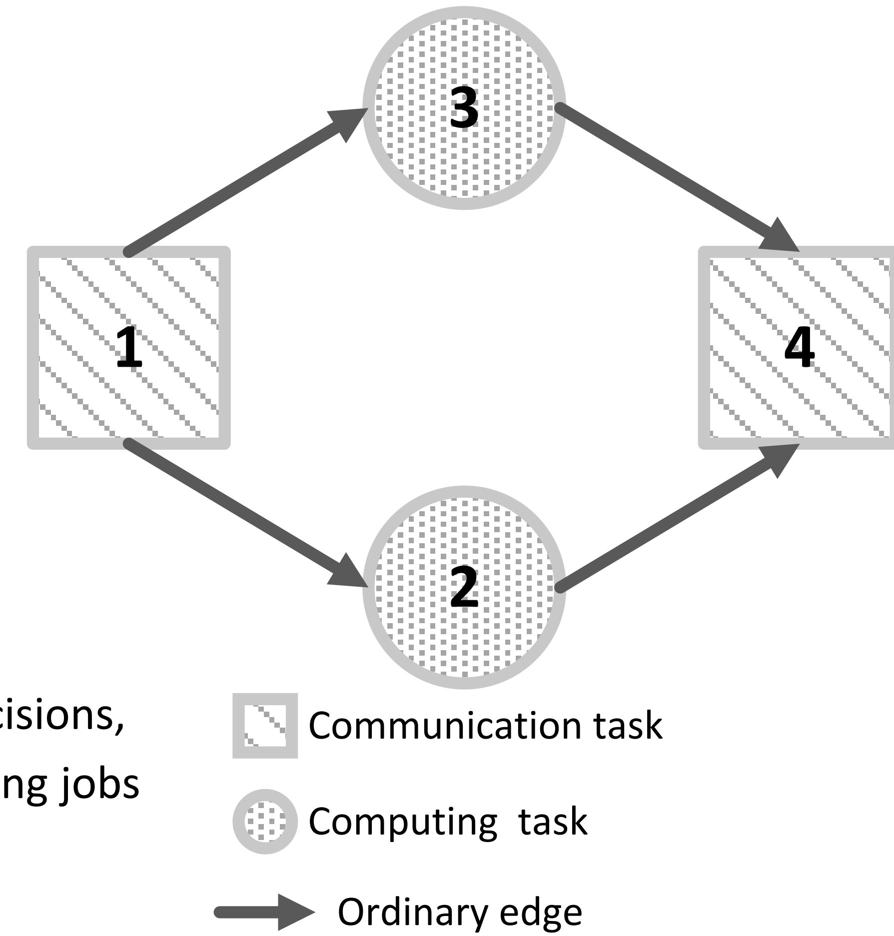
– Vertex represents computing
– Edges represent communication



- Main drawback
  – Two computing tasks cannot have the same data transfer to input
  – singe edge cannot lead to two different vertices

CA-DAG: Communication-Aware model

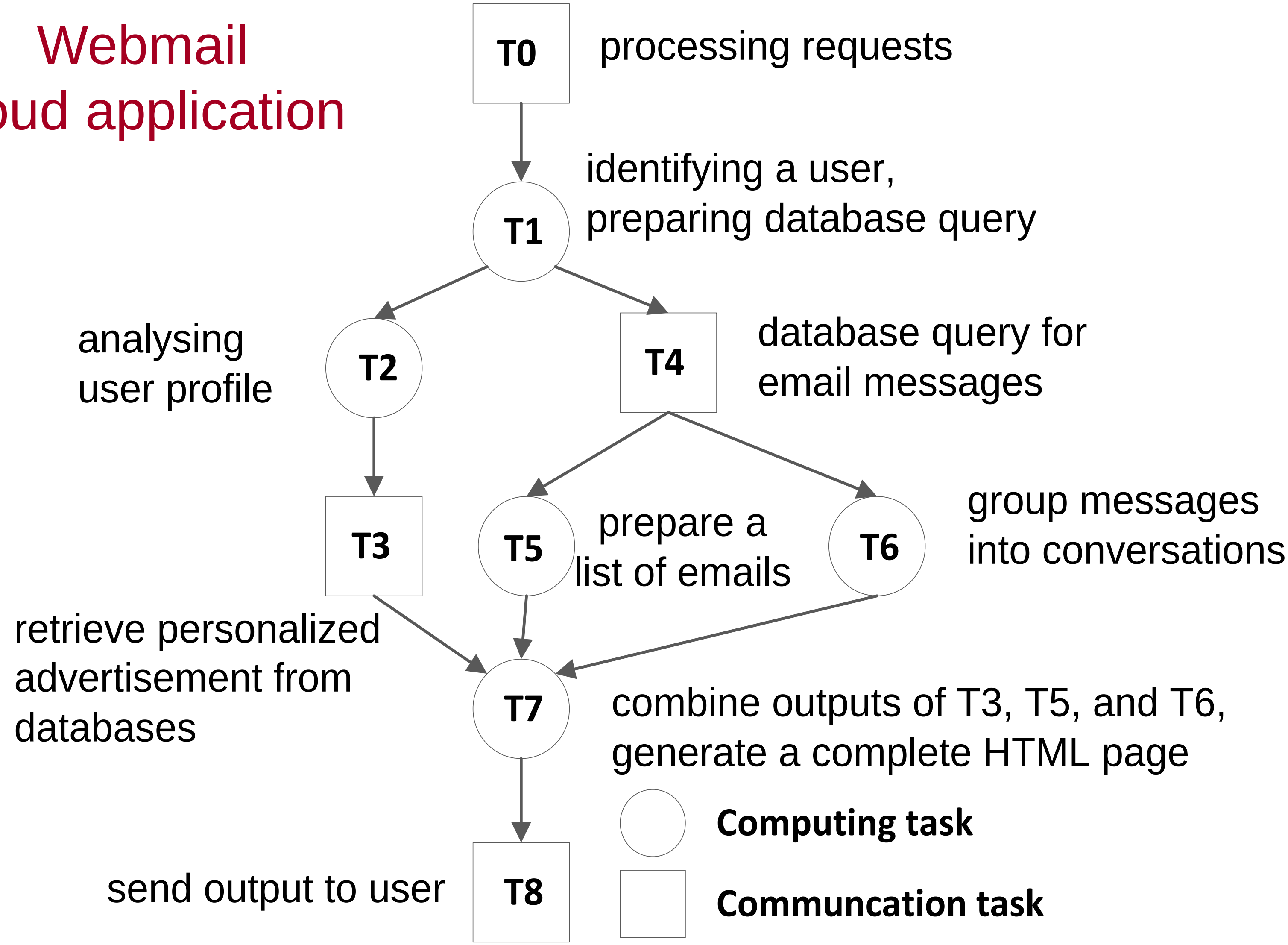
- Two types of vertices:
  - one for computing
  - one for communications
- Edges define dependences between tasks and order of execution

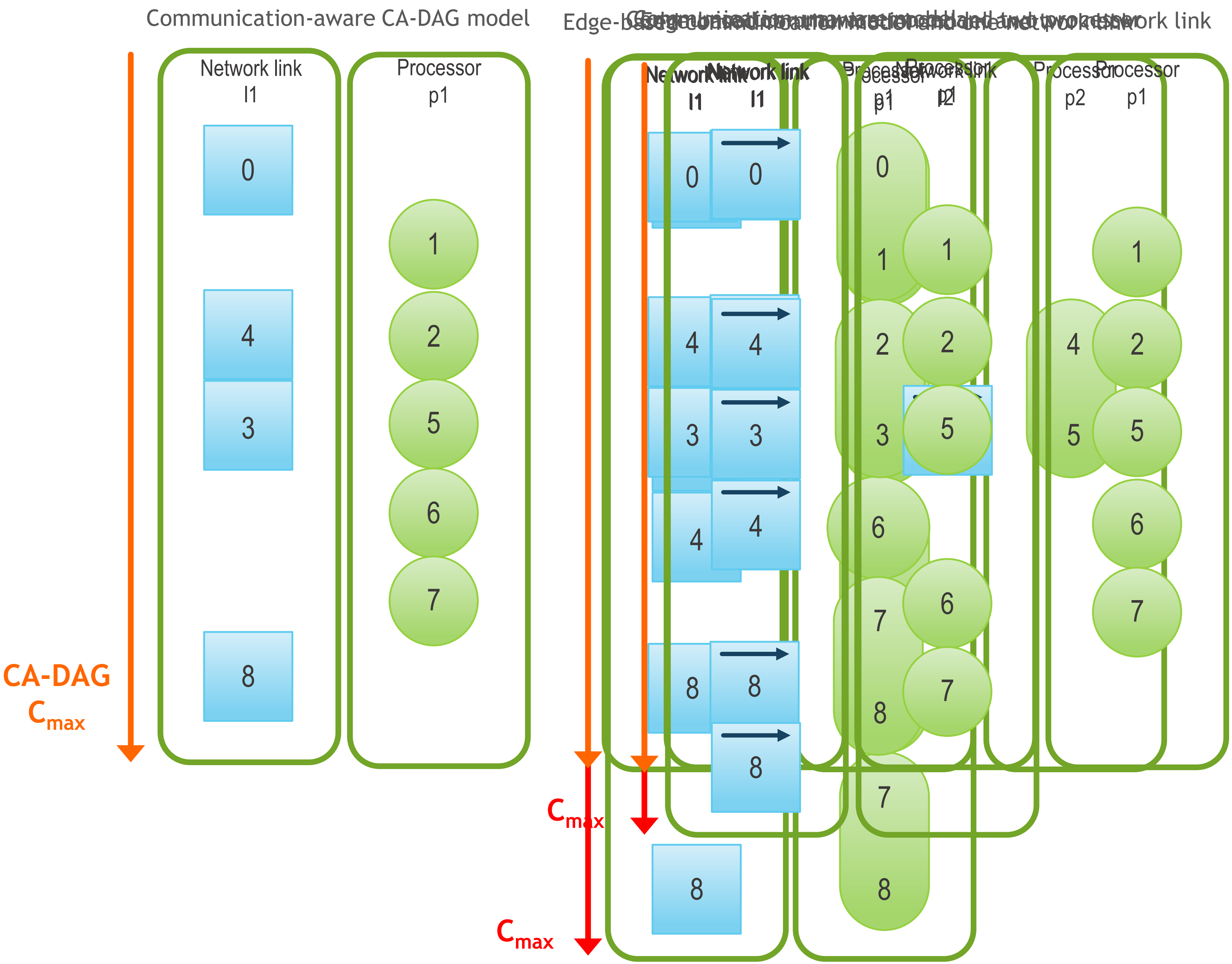


- Main advantage
  - Allows separate resource allocation decisions,
  - assigning processors to handle computing jobs
  - network resources for information transmissions

CA-DAG: Communication-Aware DAG

Webmail cloud application

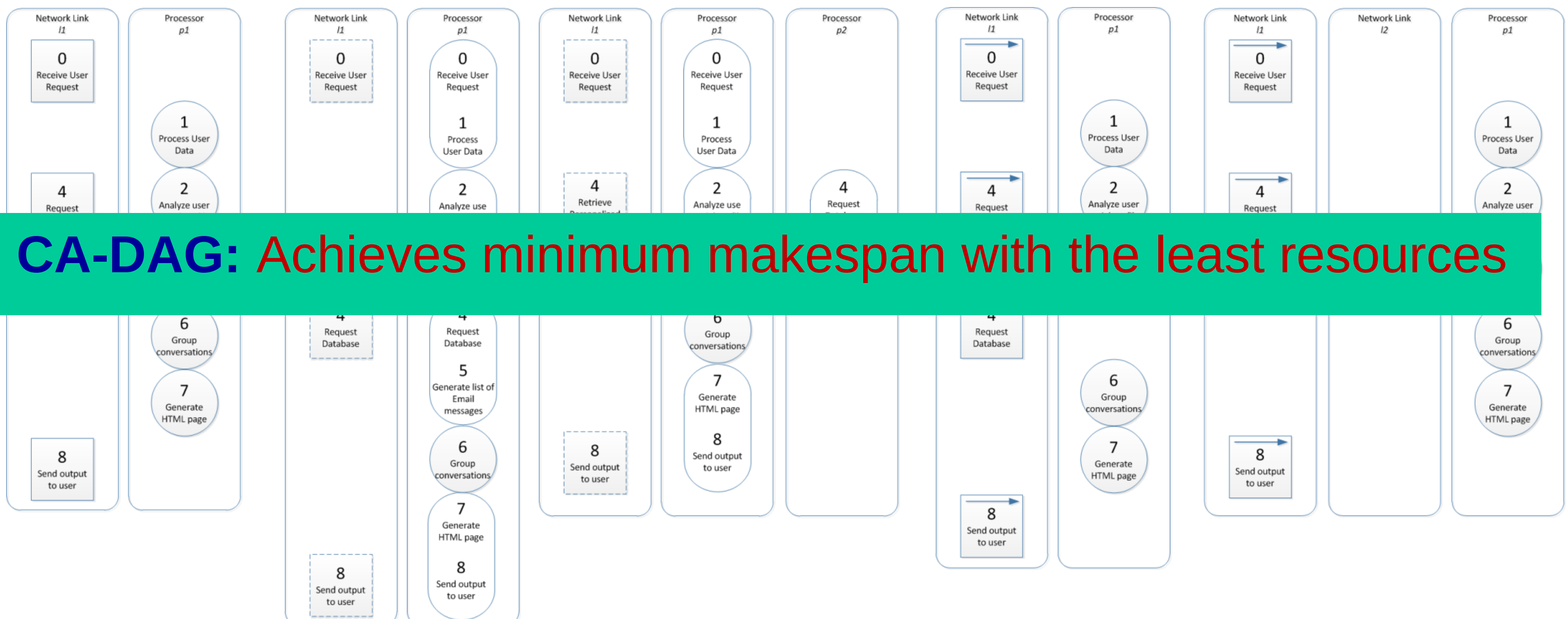




Communication-aware CA-DAG model
Edge-b
network link
Network link
l1
Processor
p1
p2
0
1
2
3
4
5
6
7
8
CA-DAG
C_{max}
C_{max}
C_{max}

Comparison of schedules

CA-DAG model **Communication-unaware model** **Edges-based model**

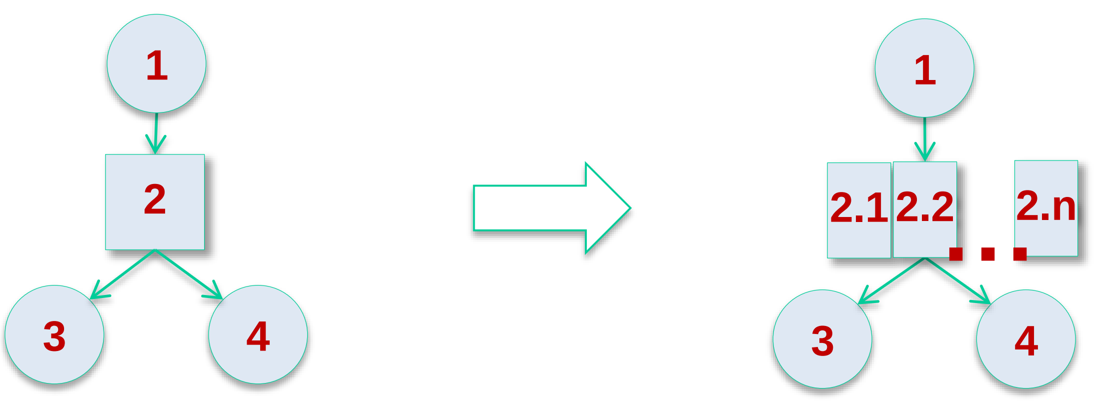


CA-DAG: Achieves minimum makespan with the least resources

# of Processors	# of Network links	Communication-unaware model	Edges-based model	**Proposed CA-DAG model**
1	1	9	8	**7**
1	2	9	7	**7**
2	1	7	8	**7**

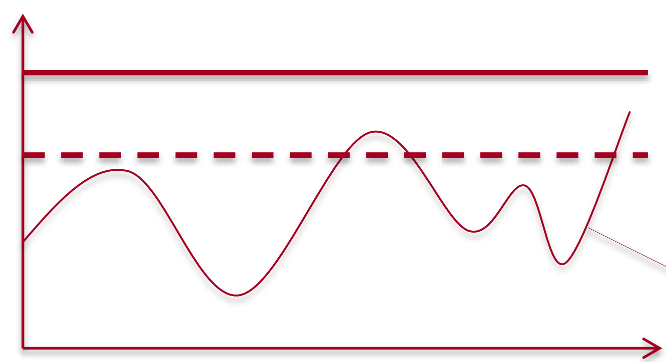
Task Parallelization

- Each communication task/vertex can be divided into n different independent communication tasks that can be executed in parallel



Bandwidth uncertainty

- Mapping of DAG to communication system with uncertainty is not efficient
- CA-DAG can use
  - Available connections and bandwidth
  - Parallel transmission



Adaptive Resource Allocation
- CPU intensive
- Communication intensive jobs

Andrei Tchernykh
Manuel Simon Combarro
CICESE Research Center, Mexico

Dzmitry Kliazovich
University of Luxembourg, Luxembourg

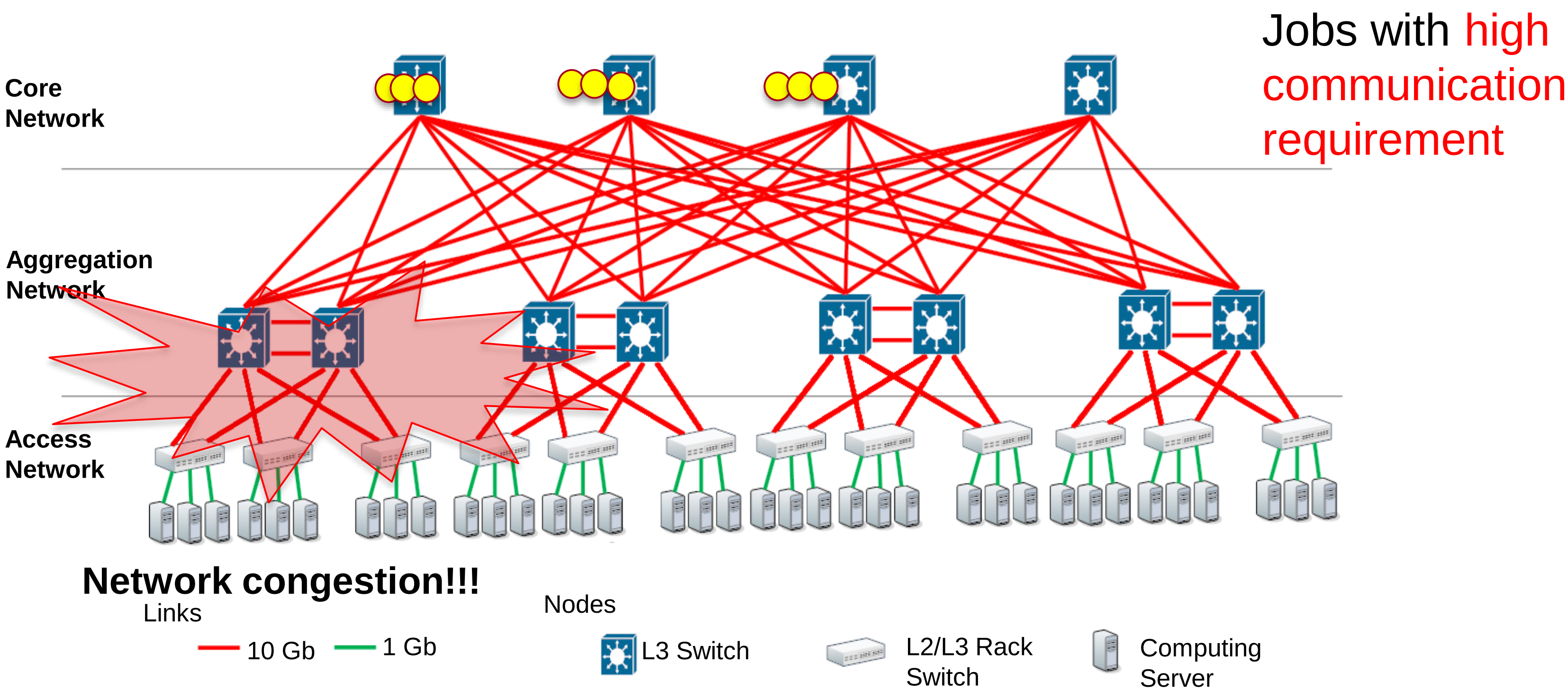
Alexander Yu. Drozdov
Moscow Institute of Physics and Technology

Gleb Radchenko
South Ural State University, Russia

Consolidation in Data Centers

Most of energy saving is due to consolidation procedures.

Increase number of server that can be put into "sleep" mode.

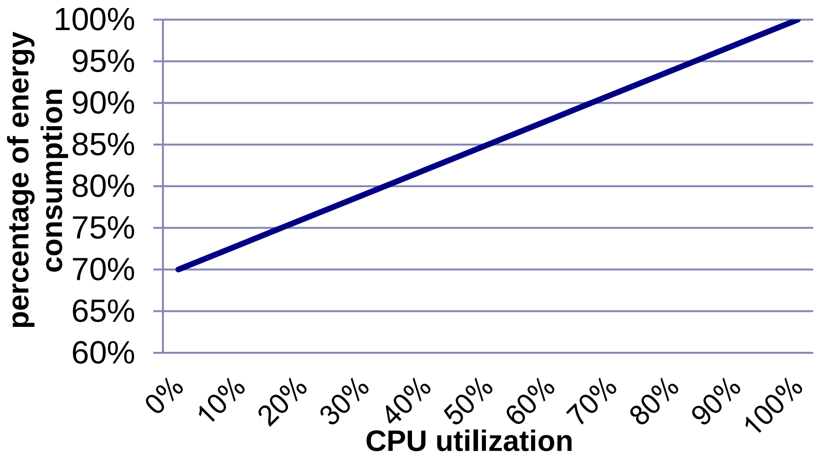


by Dzmitry Kliazovich

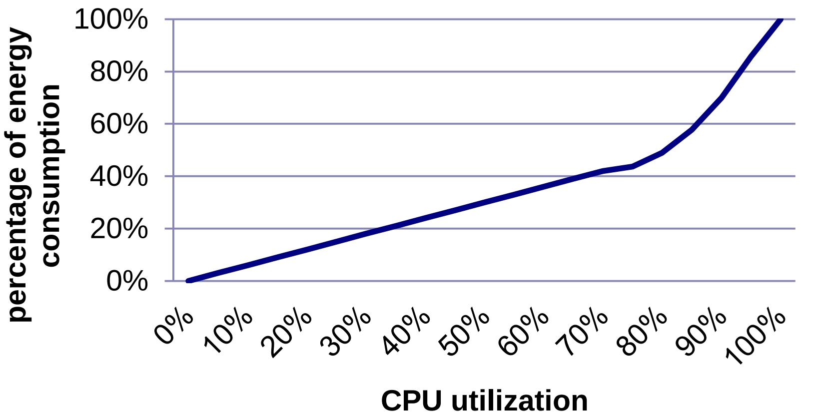
Model

- Machines $\mathrm{m} = (l_i^p(t), l_m^{cp}(t), W_m^{cp}(t))$.
  - l_i^p average paths utilization at time t
  - $l_m^{cp}(t)$ utilization at time t
  - $W_m^{cp}(t)$ power consumption at time t
    - $F_m^{cp}\left(l_m^{cp}(t)\right)$ function that represents power consumption vs utilization
    - $W_m^{cp}(t) = F_m^{cp}\left(l_m^{cp}(t)\right)$
- Jobs $j = (r_j, l_j^{cp}, l_j^{cm})$
  - r_j release time
  - l_j^{cp} computation requirements (MIPS)
  - l_j^{cm} communication requirements (Mbps)

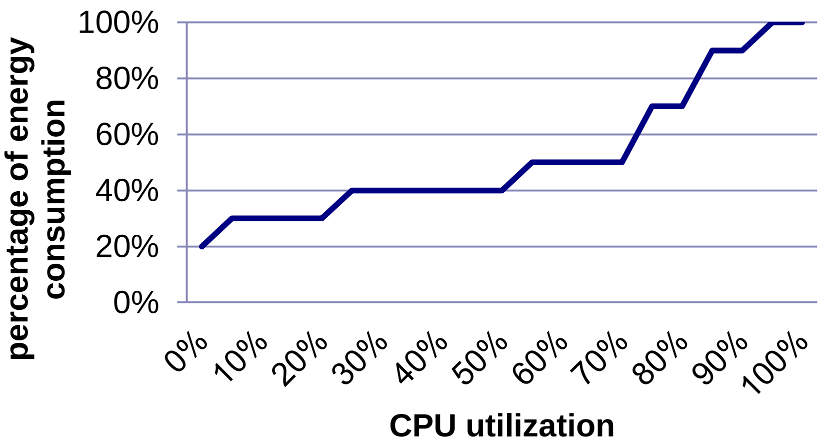
Existed Energy models



A. Beloglazov, et.al "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing" 2012.



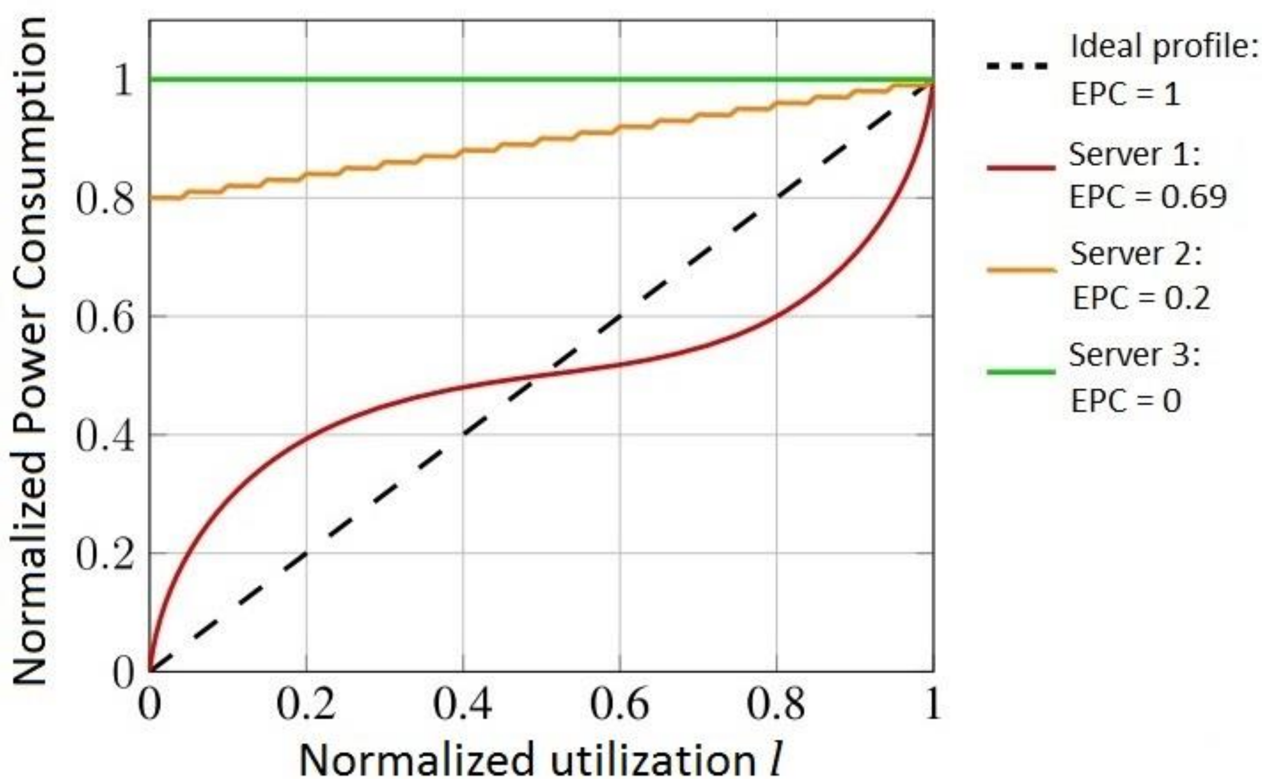
Y. Gao, et. al "An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems," 2013.



C.-H. Hsu, et. al, "Optimizing Energy Consumption with Task Consolidation in Clouds," 2014.

Energy model

Energy Proportionality Coefficient (EPC)



$\boldsymbol{EPC} = \mathbf{1}$ equal increase of utilization, equal increase of power consumption

$\boldsymbol{EPC} = \mathbf{0}$ for each increase of utilization, power consumption remains constant

Score Function

- $f_i = \alpha f_i^{cp} + (1-\alpha) f_i^{cm}$

  Network Balance ← α → Energy save (0 … 1)

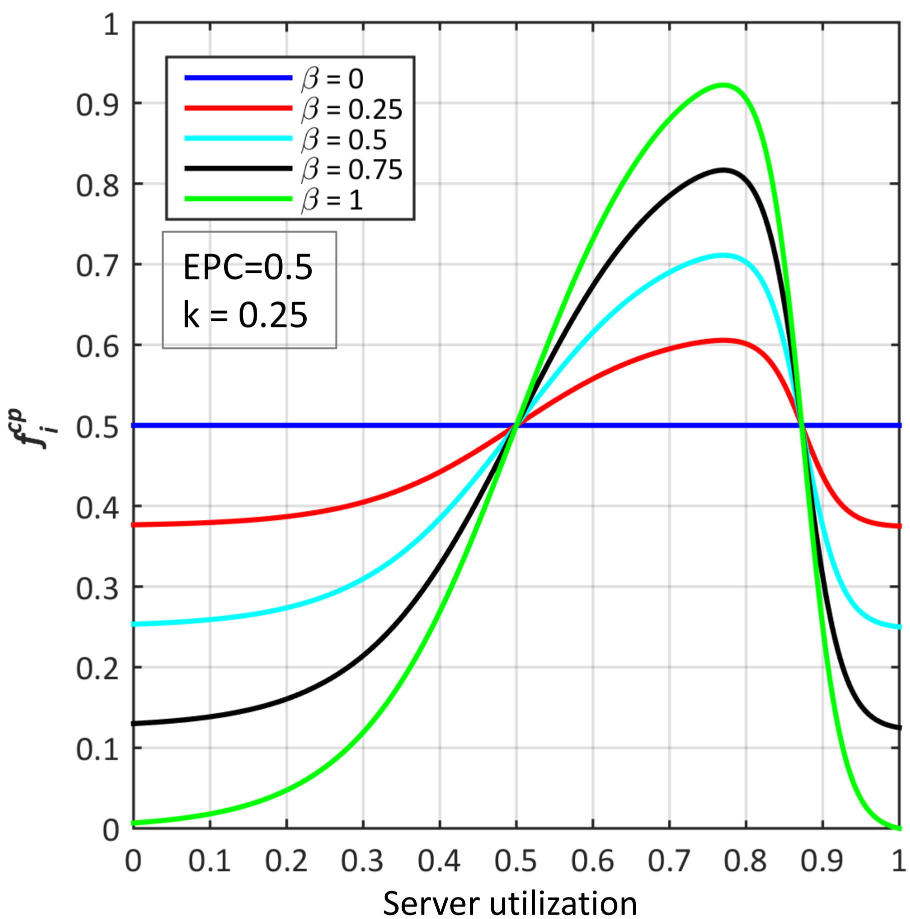
  - $f_i^{cp} = \beta \overline{f_i} + (1-\beta)\mathrm{EPC}_i^{cp}$
    - $\overline{f}_i$ [1] function of server load $l_i^{cp}(t)$

  Efficient Server ← β → Server Utilization (0 … 1)

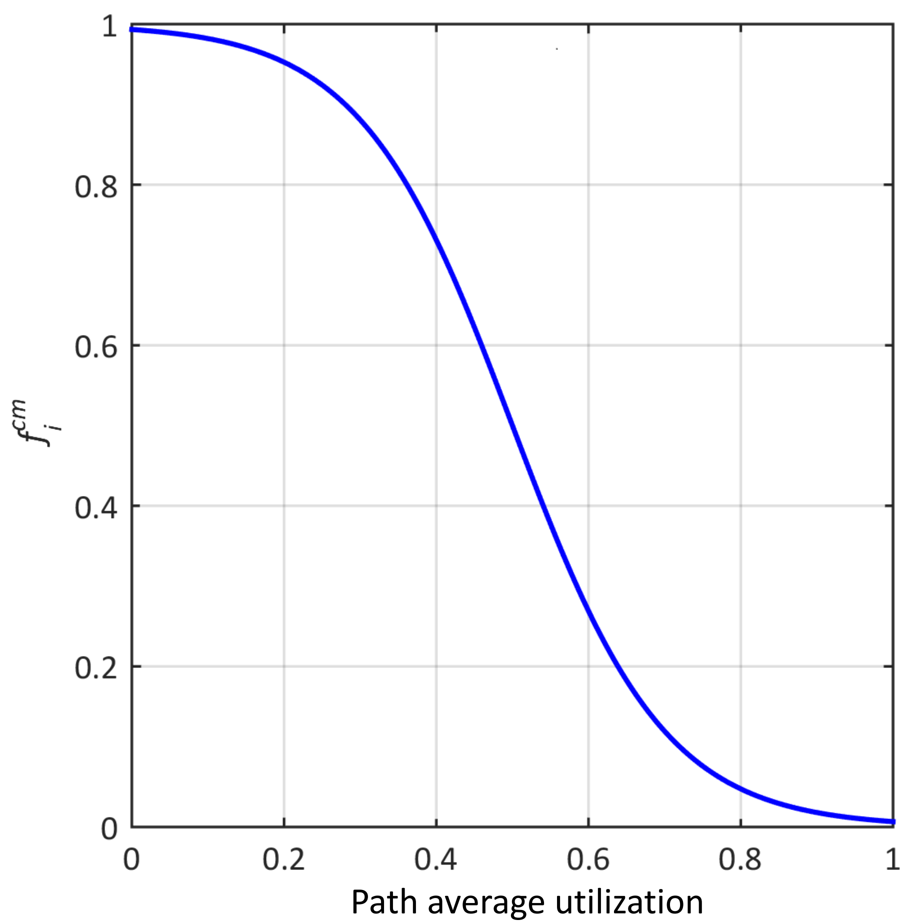
- Allocate jobs to the suitable server i with the highest f_i
- α and β can be tuned or adapted

Score Function

Computation component f_i^{cp}

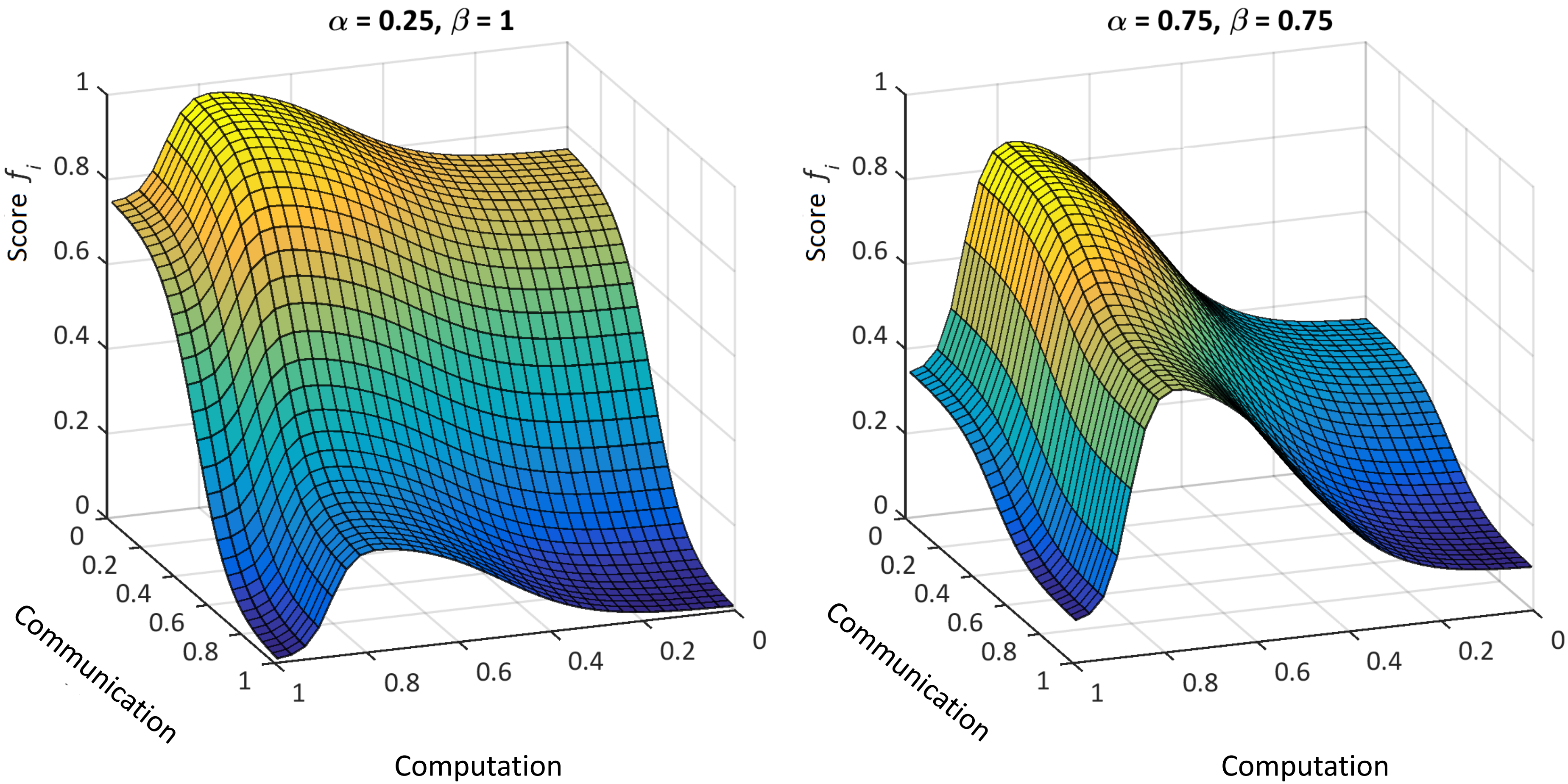


Communication component f_i^{cm}



Score Function

$\text{EPC} = 0.5$



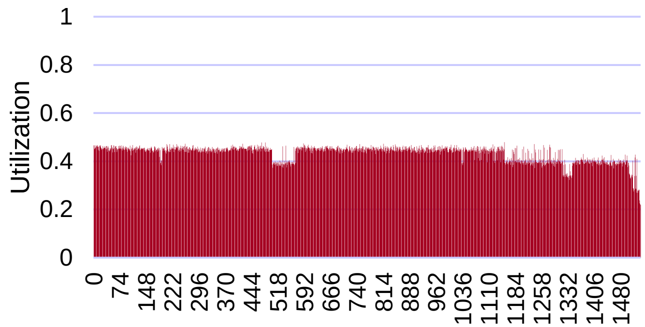
Schedulers

- Static-ACCURATE (S-ACCURATE)
  - α, β
    - Tuned
    - Before execution

- Adaptive-ACCURATE (A-ACCURATE)
  - α, β
    - Adaptive
    - During execution

S-ACCURATE

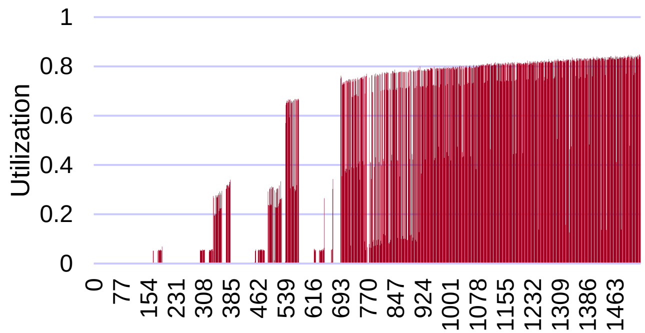
SERVERS



Config. 0.25-1

Energy 5220 Wh

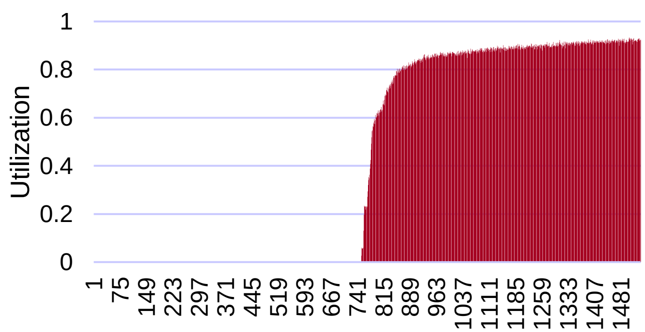
SLA violation rate 0



Config. 0.75-0.75

Energy 4455 Wh

SLA violation rate 0



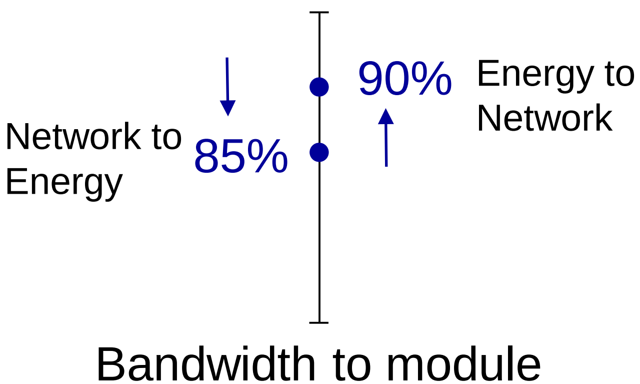
Config. 1-0

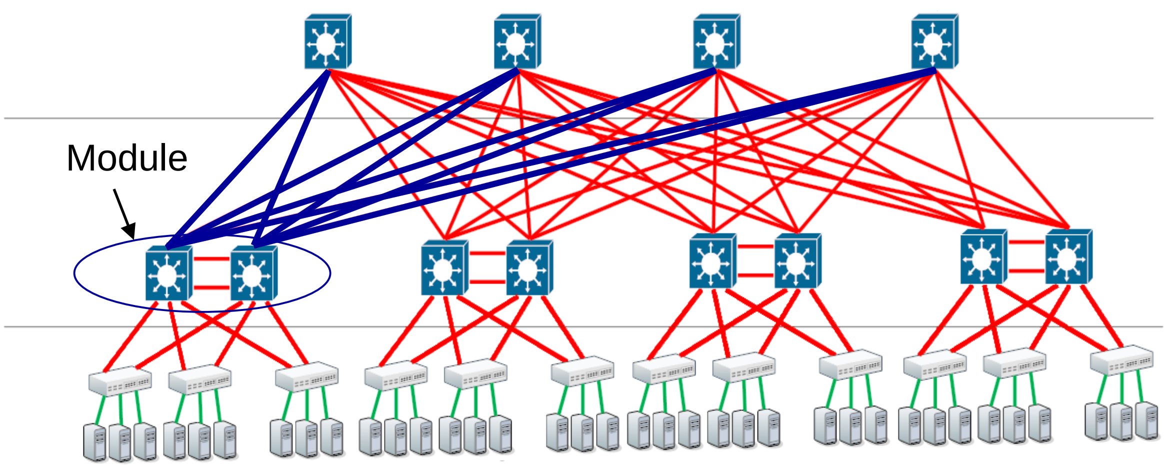
Energy 4204 Wh

SLA violation rate 0.31

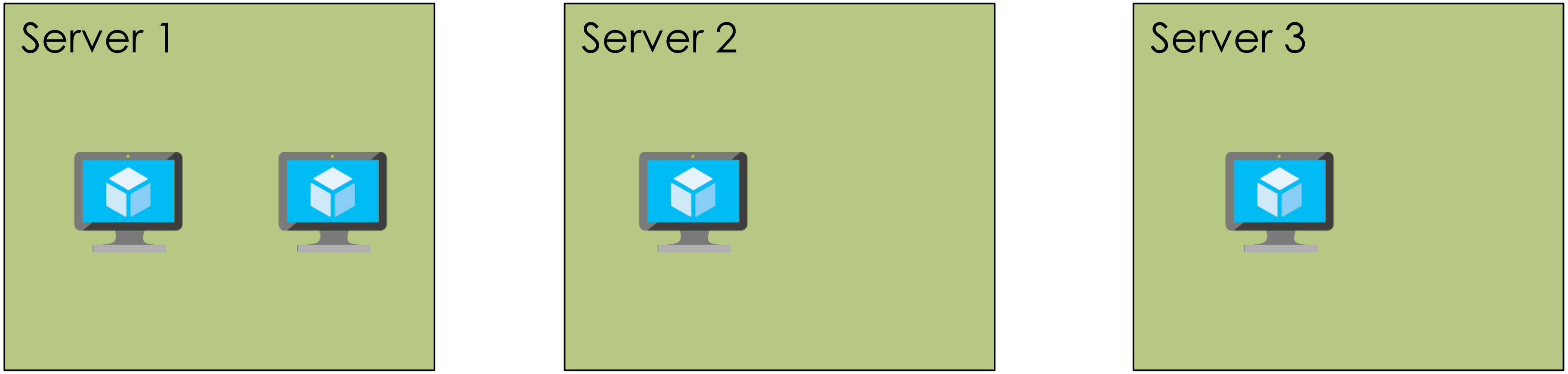
A-ACCURATE

- 2 Configurations
  - Save energy 1-0 (HPC)
  - Balanced network load: 0.25-1 (DIW)
- Adaptation criteria
  - Amax-ACCURATE (Am-ACCURATE) . Module with max bandwidth
  - Aaverage-ACCURATE (Aa-ACCURATE). Average bandwidth

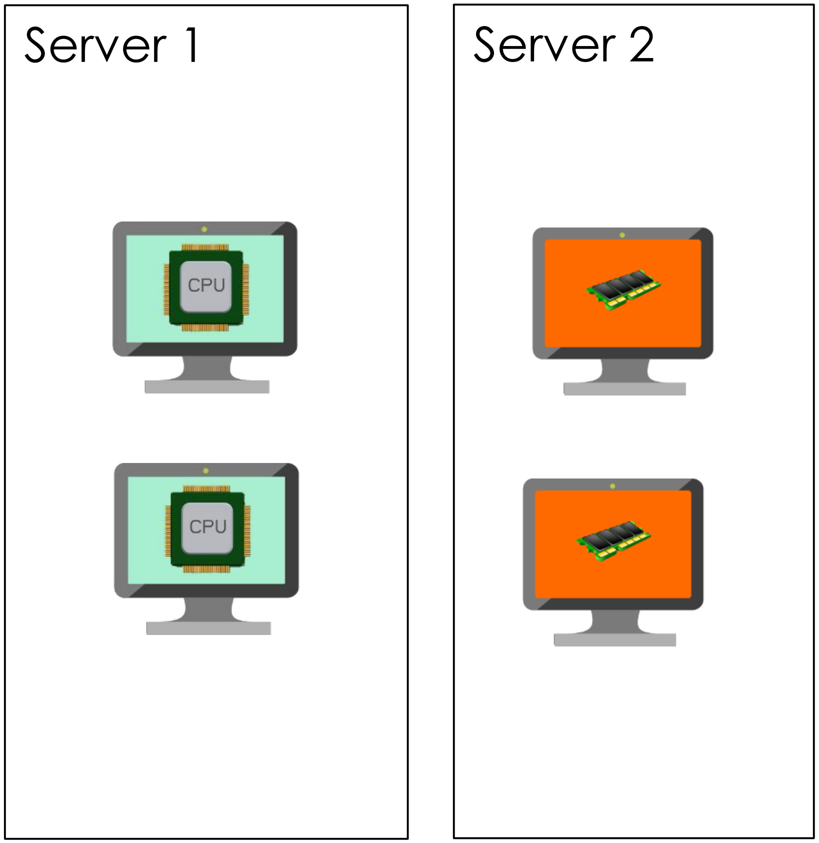


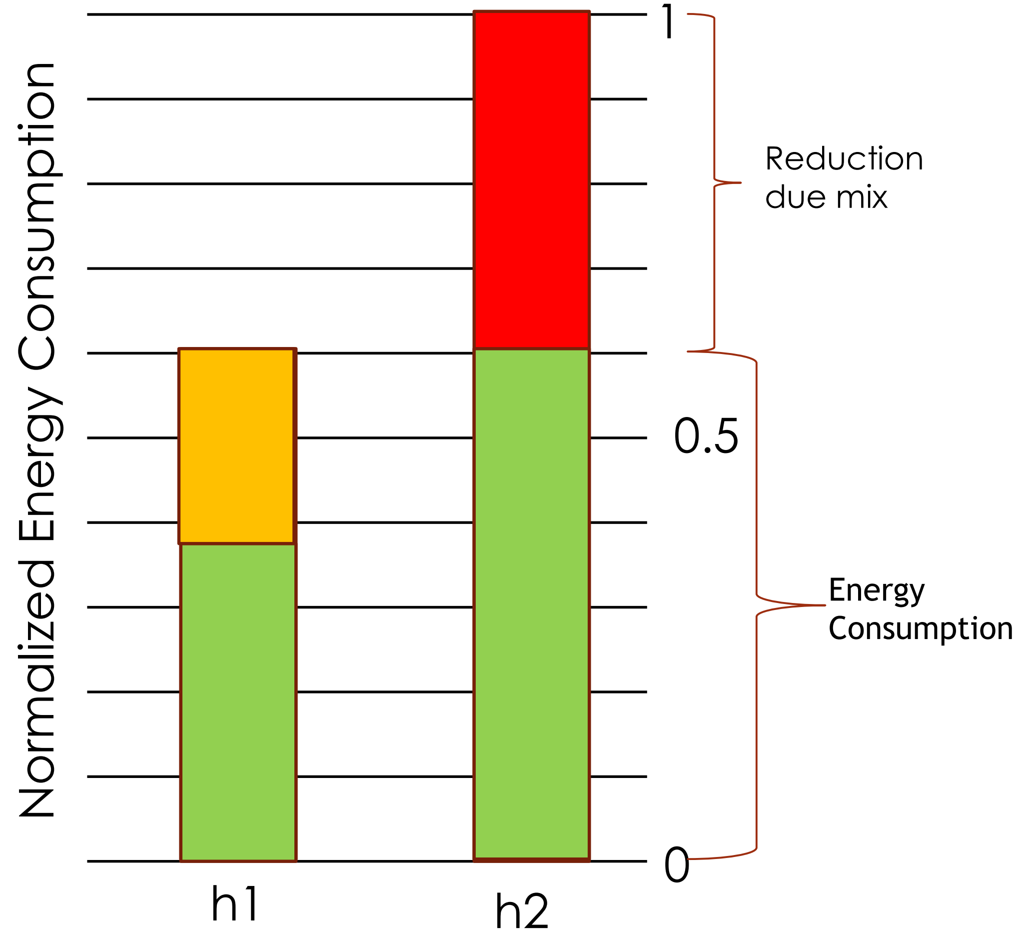


Consolidation



Consolidation with concentration





Concentration factor

Power contribution of each application in the processor take into account the combination of job types

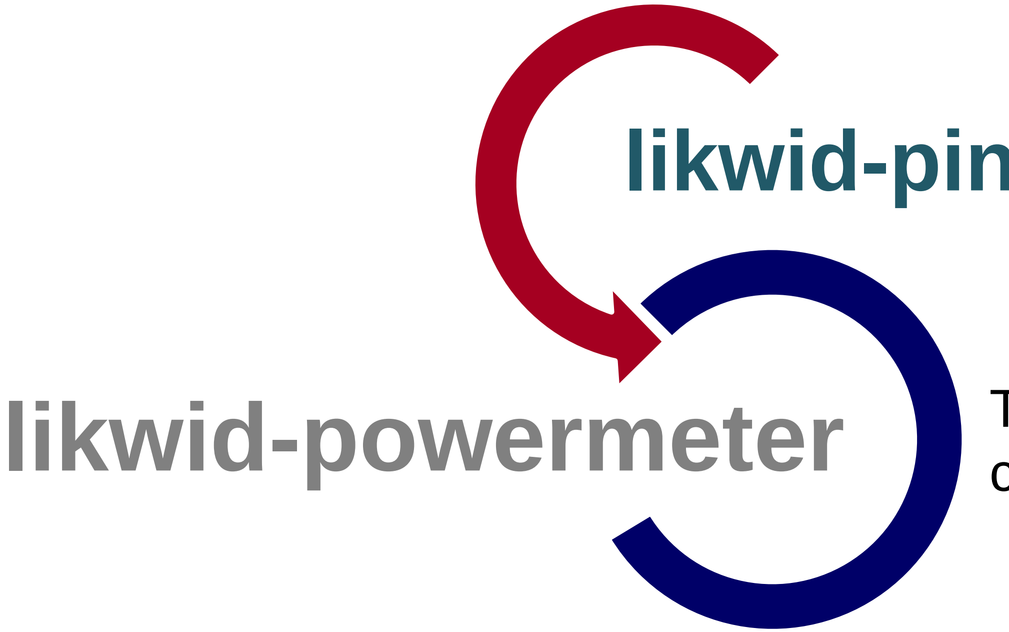
$$e(t) = o(t)\left(e_{idle} + e_{used}(t)\right)$$

$$e_{used}(t) = \ (e_{max} - e_{idle}) * F(t) * g\left(\ \alpha_{CI}(t)\right)$$

Power contribution of each application in the processor separately

$$g\left(\ \alpha_{CI}(t)\right) = 1$$

LIKWID: Tools



Tool for accessing RAPL counters on Intel processor

Other tools

- likwid-features
- likwid-topology
- likwid-perfCtr
- likwid-mpirun
- likwid-bench

PDU

A Power Distribution Unit (PDU) is a device with multiple outlets designed to distribute electric power and a digital load meter for local current monitoring to enable load balancing

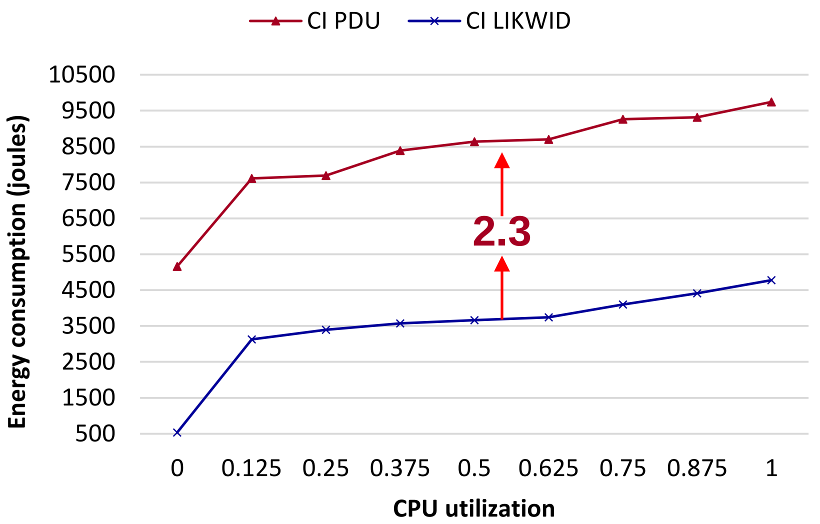
VMR-8HD20-1 Outlet Metered PDU Dual 20A 120V (8)5-15R

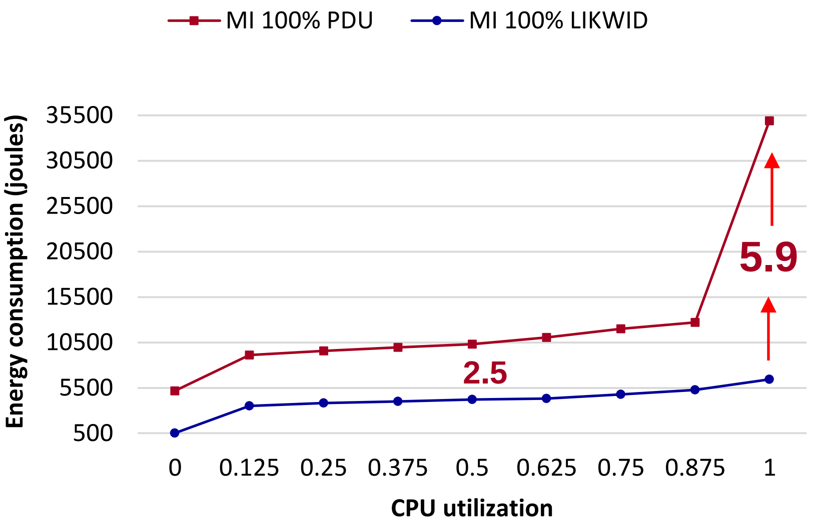
- Server Express x3650 M4
- Two Xeon IvyBridge processors E5-2650v2 95W 2.6GHz.
- Each processor has 8 Cores and two threads per core

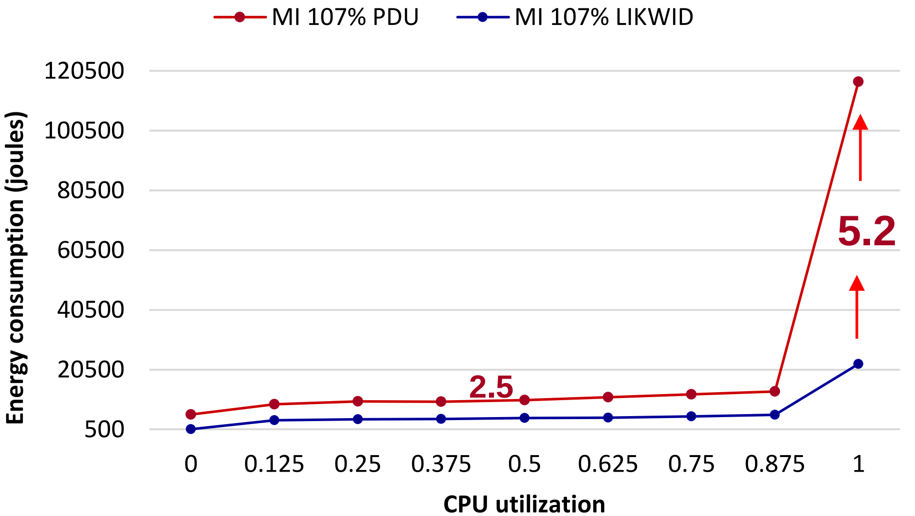
Benchmarks

Benchmark	CI	MI	NI	DI
LINPACK	●			
STREAM		●		
SysBench	●	●		●
iperf			●	
IOR				●
IOzone				●
NPB	●	●		●
Netperf			●	
SPEC	●	●		

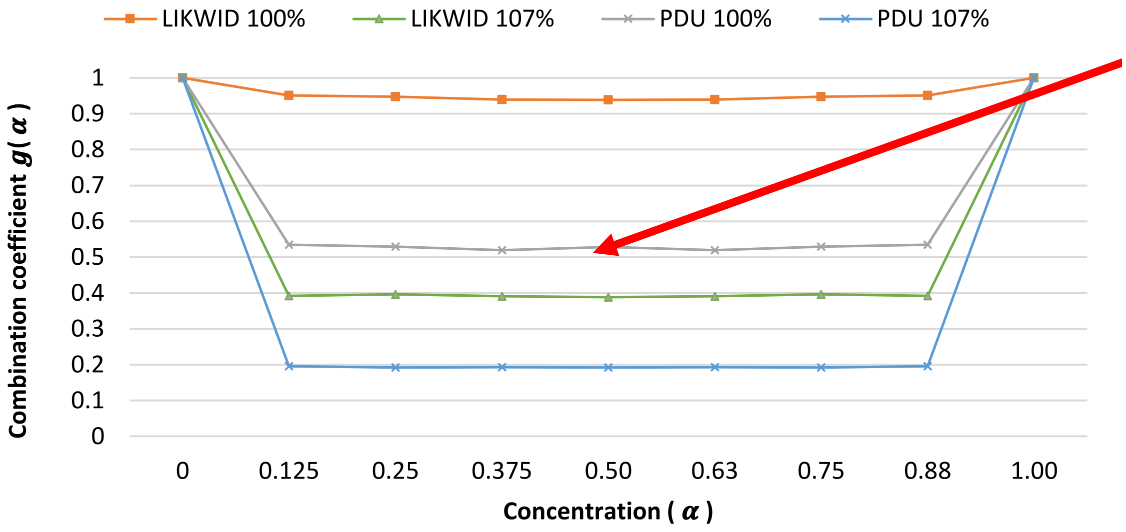
Energy consumption $F(t)$

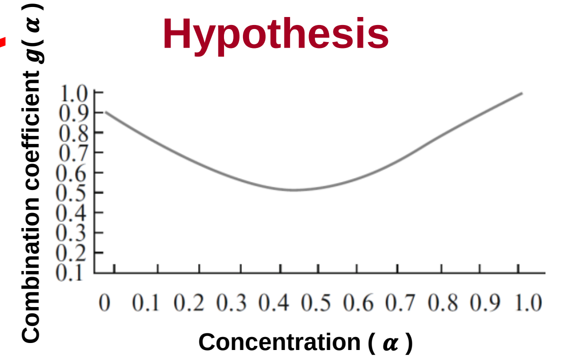


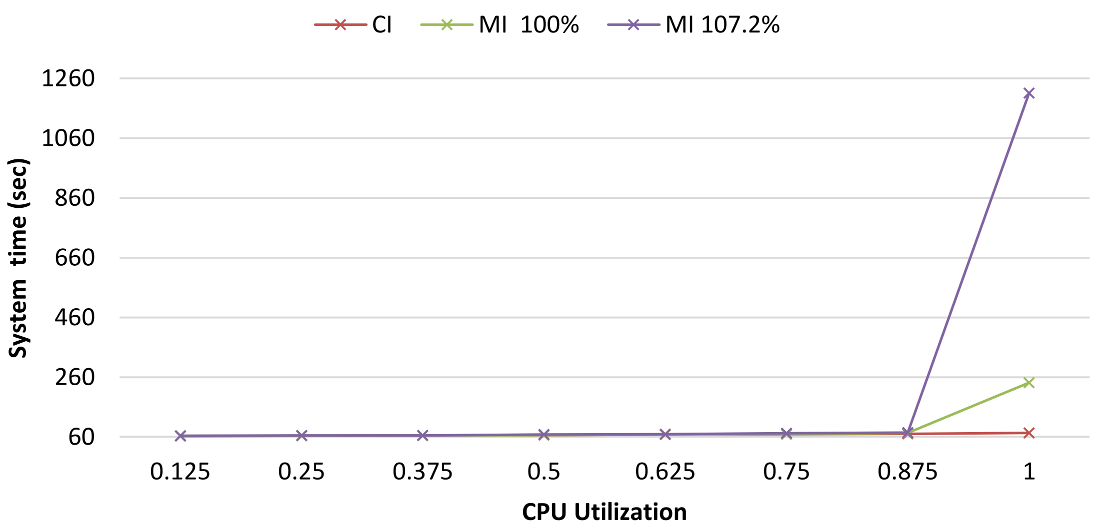




Concentration coefficient $g(\alpha_{CI}(t))$



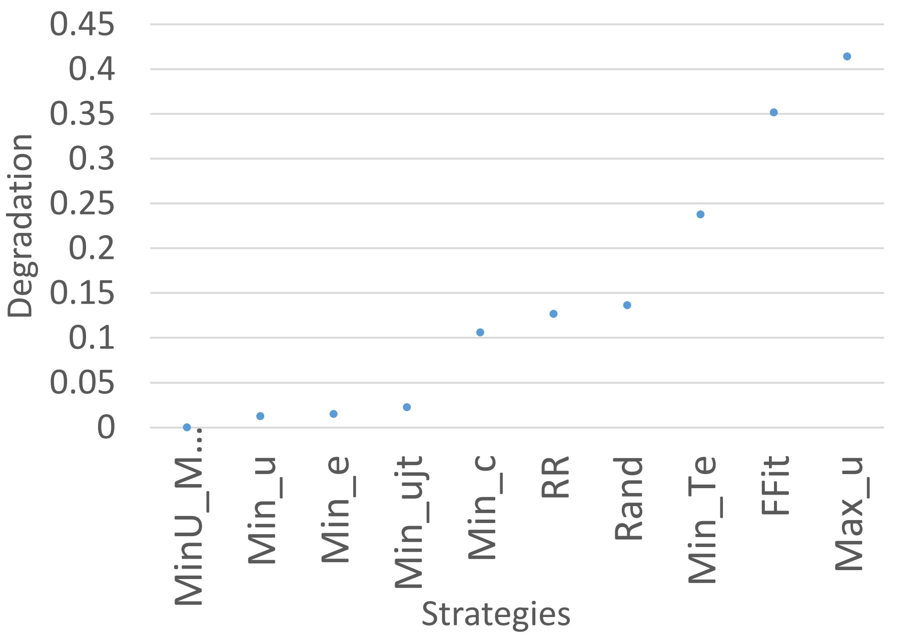




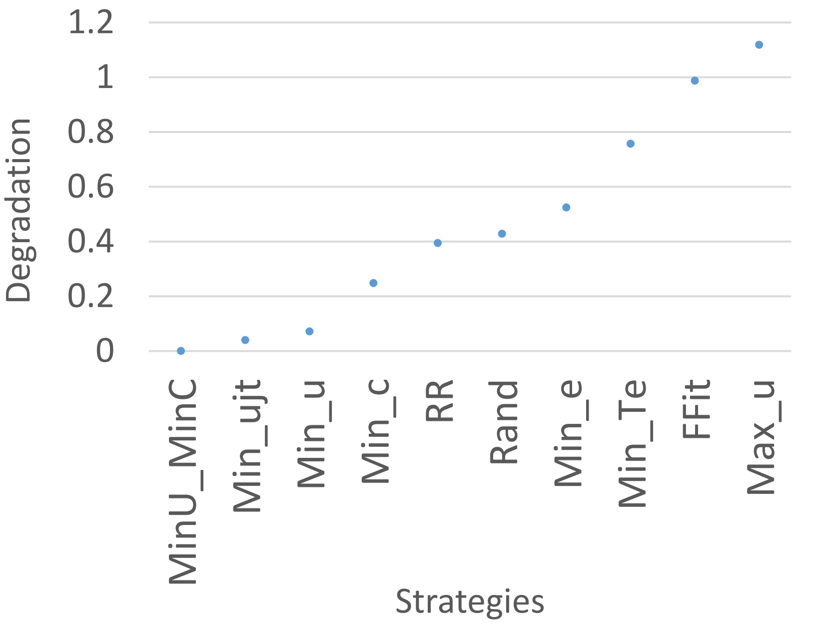
Job allocation strategies

Type	Strategy	Description
Knowledge Free	*Rand*	Allocates job j to a suitable machine randomly using a uniform distribution in the range $[1..m]$.
	FFit (First Fit)	Allocates job j to the first machine available and capable to execute it.
	RR (Round Robin)	Allocates job j to the machine available and capable to execute by Round Robin strategy
Energy-aware	*Min_Te* (Min-Total_energy)	Allocates job j to the machine with minimum total energy consumption at time r_j : $min_{i=1..m}\left(\sum_{t=1}^{r_j} e_i^{proc}(t)\right)$
	Min_e (Min-energy)	Allocates job j to the machine with minimum power consumption at time r_j : $min_{i=1..m}\left(e_i^{proc}(r_j)\right)$
Utilization Aware	*Min_u* (Min-utilization)	Allocates job j to the machine with minimum total utilization at time r_j $min_{i=1..m}(u_i^{proc})$
	Max_u (Max-utilization)	Allocates job j to the machine with maximum total utilization at time r_j $max_{i=1..m}(u_i^{proc})$
Job type	*MinU_MinC* (Min utilization and Min concentration)	Allocates job j to the machine in the subset of machines with minimum total utilization at time r_j $min_{i=1..m}(u_i^{proc})$ and minimum concentration of jobs of the same type.
	Min_ujt (Min-util_job_type)	Allocates job j to the machine with minimum utilization of jobs of the same type at time r_j
	Min_c (Min-concentration)	Allocates job j to the machine with minimum concentration of jobs of the same type at time r_j

Power consumption degradation analysis

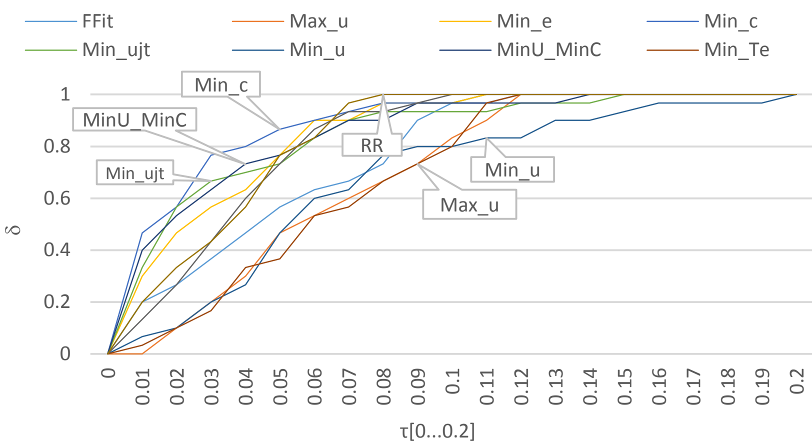


LIKWID

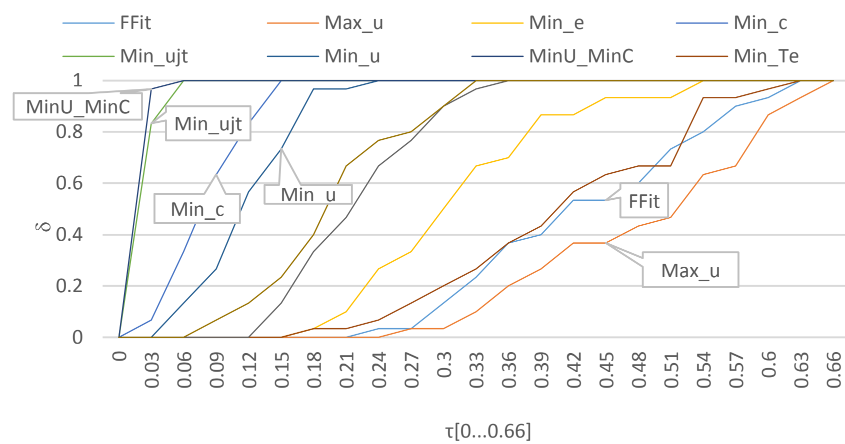


PDU

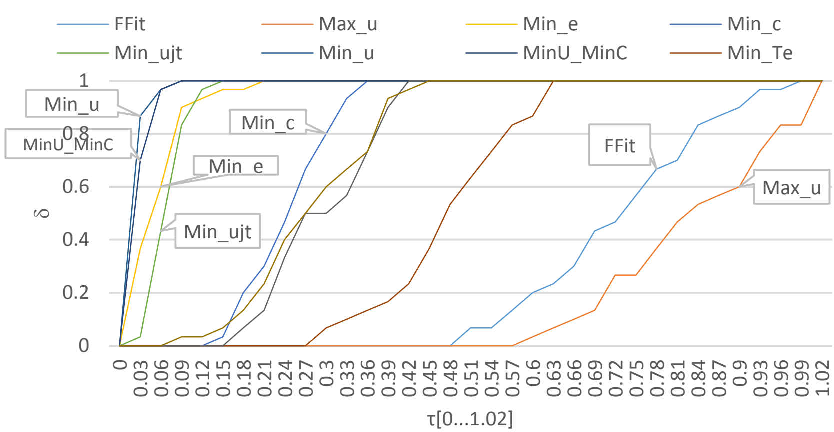
Performance profile of power consumption



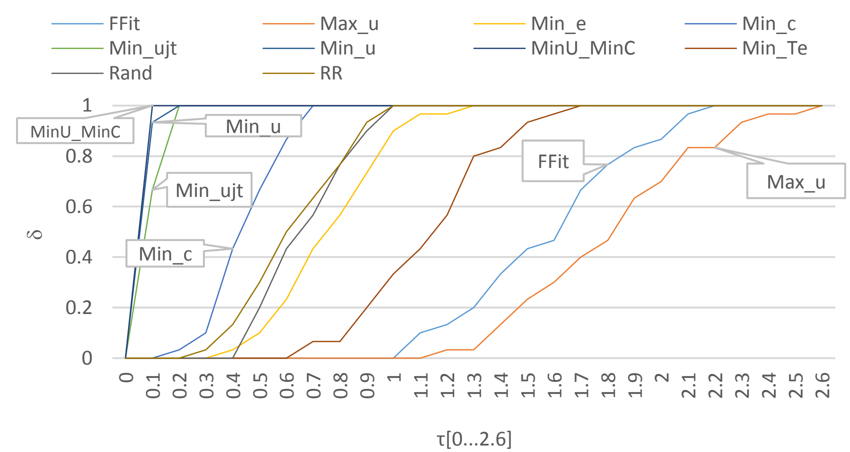
(a) LIKWID “A”



(b) PDU “A”

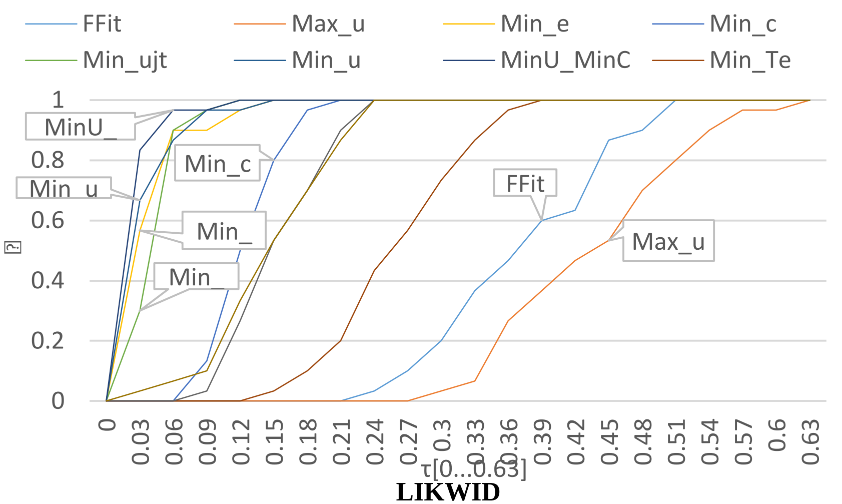


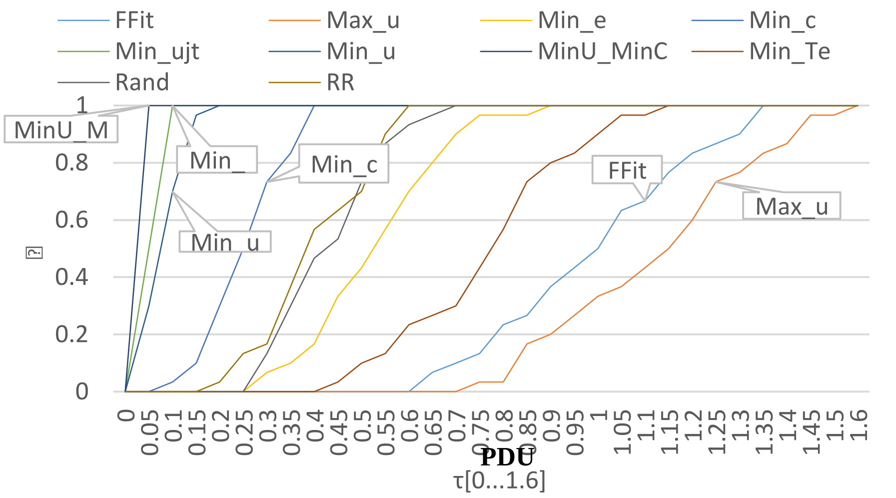
c)LIKWID “B”



(d) PDU “B”

Performance profile of power consumption





Mean performance profile of the power degradation of 10 strategies

Conclusions

To deal with uncertainty of communication

We propose

CA-DAG: Communication-Aware DAG model

- Allows separate resource allocation decisions
  - **computing task to** processors
  - **communication task to** network resources
- Task parallelization
- Multipath routing
- Adapt to bandwidth uncertainty

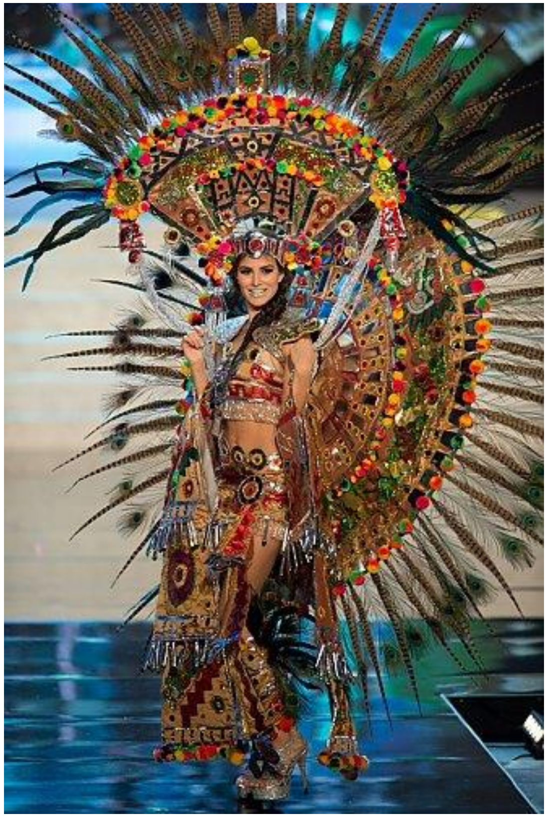
Adaptive Resource Allocation Strategy

- CPU intensive

- Communication intensive jobs

to cope with different objective preferences, workloads, and cloud roperties

Concentration policy for uncertainty of Resource Contention



Thanks for your attention!