Digital sufficiency behaviors to deal with intermittent energy sources in data center

GreenDays2024 @ Toulouse

Jolyne Gatt, Maël Madon, Georges Da Costa

March 27, 2024
Introduction: rebound effect

What is the average annual PUE for your largest data center? (n=669)

- **2007**: 2.50
- **2011**: 1.98
- **2014**: 1.65
- **2018**: 1.58
- **2019**: 1.67
- **2020**: 1.59
- **2021**: 1.57
- **2022**: 1.55

**Source**: Uptime Institute Global Data Center Survey 2022

**Global trends in digital and energy indicators, 2015-2022**

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2022</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet users</td>
<td>3 billion</td>
<td>5.3 billion</td>
<td>+78%</td>
</tr>
<tr>
<td>Internet traffic</td>
<td>0.6 ZB</td>
<td>4.4 ZB</td>
<td>+600%</td>
</tr>
<tr>
<td>Data centre workloads</td>
<td>180 million</td>
<td>800 million</td>
<td>+340%</td>
</tr>
<tr>
<td>Data centre energy use (excluding crypto)</td>
<td>200 TWh</td>
<td>240-340 TWh</td>
<td>+20-70%</td>
</tr>
<tr>
<td>Crypto mining energy use</td>
<td>4 TWh</td>
<td>100-150 TWh</td>
<td>+2300-3500%</td>
</tr>
<tr>
<td>Data transmission network energy use</td>
<td>220 TWh</td>
<td>260-360 TWh</td>
<td>+18-64%</td>
</tr>
</tbody>
</table>

**Source**: International Energy Agency
Efficiency is not enough: sufficiency

Sufficiency policies (IPCC, 2022)
A set of measures and daily practices that avoid demand for energy, materials, land and water while delivering human well-being for all within planetary boundaries.

What would “sufficiency” mean for data centers?
→ voluntary limitation, empower and involve the user
Data center model

Users submit jobs to the Scheduler, which uses various scheduling algorithms such as FCFS or easy-bf. The Scheduler interacts with the IT platform, which consists of multiple servers, possibly mono- or multicore. The IT platform consumes electricity.
Sufficiency behaviors

- job final state = $n \times see\_you\_later + b$
- $b \in \{ \text{rigid, degrad (space), degrad (temp), renounce, reconfig} \}$
Renewable energy production

- Solar panels:

![Graph showing renewable energy production and demand over time]

- Power (percentage vs max production)
  - IT demand
  - Renewable power
  - Underproduction
  - Overproduction

Jolyne Gatt, Maël Madon, Georges Da Costa
3-state energy model

- 3-color state for energy production:
  - **green state**: everything is fine (production $\geq 100\%$ max conso)
  - **yellow state**: some disturbance (production $\geq 50\%$ max conso)
  - **red state**: system critical (production $< 50\%$ max conso).
3-state energy model

- **3-color state for energy production:**
  - **green state:** everything is fine (production ≥ 100% max conso)
  - **yellow state:** some disturbance (production ≥ 50% max conso)
  - **red state:** system critical (production < 50% max conso).
Energy-aware behaviors

- **job to submit**
  - no effort, $p=1-\alpha$
  - effort, $p=\alpha$

- **which energy state?**
  - green: how many cores requested?
    - 1 core
    - >1 core
  - yellow: how many cores requested?
    - 1 core
    - >1 core
  - red: how many cores requested?
    - 1 core
    - >1 core

- **user choice**
  - rigid
  - degrad (temp)
  - degrad (space)
  - reconfig
  - see you later
  - renounce

---

Jolyne Gatt, Maël Madon, Georges Da Costa
Sufficiency behaviors for data center users
8 / 14
Energy-aware behaviors

- choice of behavior at random depending on the state

- rig
- see you later
- degrad (space)
- degrad (temp)
- reconfig
- renounce
- no effort, \( p = 1 - \alpha \)
- effort, \( p = \alpha \)

Which energy state?

Job to submit

- no effort, \( p = 1 - \alpha \)

User choice

- rigid
- degrad (temp)
- degrad (space)
- reconfig
- see you later
- renounce

- how many cores requested?

1 core

>1 core

- how many cores requested?

1 core

>1 core
Experimental setup

- **Software**: Batsim + Batmen

- **IT workload**: filtered version of MetaCentrum from Parallel Workload Archive
  - June 1 to November 11, 2014 (4.5 months)
  - 650000 jobs and 500 users

- **Energy production data**: 145 m² solar panels
  - weather data Toulouse 2019 from Renewable Ninja (days aligned with IT)

- **IT platform**: 42 18-core machines

- **Scheduler**: bin-packing scheduler which shutdown machine when idle.
Experimental campaign

- $\alpha = \text{probability of modifying a job in red / yellow}$
- 6 scenari:
  - full rigid ($\alpha = 0$)
  - low effort ($\alpha = 0.25$)
  - medium effort ($\alpha = 0.5$)
  - big effort ($\alpha = 0.75$)
  - max effort ($\alpha = 1$)
  - full renounce/degrad/reconfig in red
- each scenario run 30 times to minimize the effect of randomness
How much does user effort impact energy consumption?

→ if 50% jobs modified in red/yellow (medium effort), underproduction reduced by 8%

→ if 100% jobs modified in red/yellow (max effort), underproduction reduced by 18%

Energy savings linear with effort

\[ Y = -37.701X + 8253 \]

r-squared: 0.957
How much does user effort impact energy consumption?

→ if 50% jobs modified in red/yellow (medium effort), underproduction reduced by 8%

→ if 100% jobs modified in red/yellow (max effort), underproduction reduced by 18%
Results

How much does user effort impact energy consumption?

→ if 50% jobs modified in red/yellow (medium effort), underproduction reduced by 8%
→ if 100% jobs modified in red/yellow (max effort), underproduction reduced by 18%

Energy savings linear with effort
Results: ratio energy/effort

- low effort
- low effort (yellow)
- medium effort
- medium effort (yellow)
- big effort
- big effort (yellow)
- max effort
- max effort (yellow)

Gains in underproduction per modified job (in Wh/job)

Marginal gains increase with $\alpha$: "the more people who make an effort, the greater the impact of a user's additional effort."
Results: ratio energy/effort

→ marginal gains increase with $\alpha$: “the more people who make an effort, the greater the impact of a user’s additional effort”

→ gains with yellow windows of the same scale than with red
Conclusion

- 3-state energy model and user behaviors to adapt job to energy consumption

Possible improvements:
- Thresholds on instantaneous available energy
- Collaboration with the scheduler
- More realistic replay method
- Social science studies (willingness to adopt behaviors, impact of eco-feedback)

First step towards studying sufficiency and not efficiency.

Simulation campaign reproducible.

Article in review: J. Gatt, Mael Madon, and Georges Da Costa, "Digital sufficiency behaviors to deal with intermittent energy sources in data center."
Conclusion

- 3-state energy model and user behaviors to adapt job to energy consumption
- Possible improvements:
  - thresholds on instantaneous available energy
  - collaboration with the scheduler
  - more realistic replay method
  - social science studies (willingness to adopt behaviors, impact of eco-feedback)
Conclusion

- 3-state energy model and user behaviors to adapt job to energy consumption
- Possible improvements:
  - thresholds on instantaneous available energy
  - collaboration with the scheduler
  - more realistic replay method
  - social science studies (willingness to adopt behaviors, impact of eco-feedback)
- First step towards studying **sufficiency** and not **efficiency**
Conclusion

- 3-state energy model and user behaviors to adapt job to energy consumption
- Possible improvements:
  - thresholds on instantaneous available energy
  - collaboration with the scheduler
  - more realistic replay method
  - social science studies (willingness to adopt behaviors, impact of eco-feedback)
- First step towards studying **sufficiency** and not **efficiency**
- Simulation campaign **reproducible**
Conclusions

- 3-state energy model and user behaviors to adapt job to energy consumption

Possible improvements:
- Thresholds on instantaneous available energy
- Collaboration with the scheduler
- More realistic replay method
- Social science studies (willingness to adopt behaviors, impact of eco-feedback)

First step towards studying sufficiency and not efficiency

Simulation campaign reproducible

Article in review:
- J. Gatt, M. Madon, and G. Da Costa, “Digital sufficiency behaviors to deal with intermittent energy sources in data center.”
Come to my **PhD defense: April 30, 14:00** (link to come) !!

Do not hesitate to contact me :-)

- [www.irit.fr/~Mael.Madon](http://www.irit.fr/~Mael.Madon)
- mael.madon@irit.fr