Heavy Tails: Performance Models and Scheduling Disciplines

Part II – Workload Asymptotics for Generalized Processor Sharing Systems

R. Núñez-Queija

CWI & TU/e

Organization

1. Background & motivation

2. Generalized Processor Sharing (GPS)

3. Performance evaluation

4. Model description

5. Workload asymptotics in various scenarios

6. Discussion & conclusion

7. References

Background & motivation

- Support variety of services
- Voice and video vs data applications
- Need for differentiated QoS
- Packet scheduling: IntServ/DiffServ
- Possible intermediate scenario
  - ◇ Fine-grained scheduling at network edge (in particular wireless access and application servers)
  - ◇ Coarse-level or no scheduling in network core

Generalized Processor Sharing (GPS)

Bandwidth sharing in proportion to weight factors

Two crucial properties

- Minimum-rate guarantees, providing flow isolation and preventing starvation effects
- Work conservation, achieving statistical multiplexing gains and thus ensuring efficient bandwidth utilization

GPS includes strict-priority scheduling as special case

Performance evaluation

Focus on evaluation of performance for given weights

Inverse problem: how to set weights to meet given performance target

Performance evaluation

Focus on evaluation of performance for given weights

Inverse problem: how to set weights to meet given performance target

Exact analysis extremely difficult, motivating derivation of bounds and asymptotics

GPS = Coupled-processors model

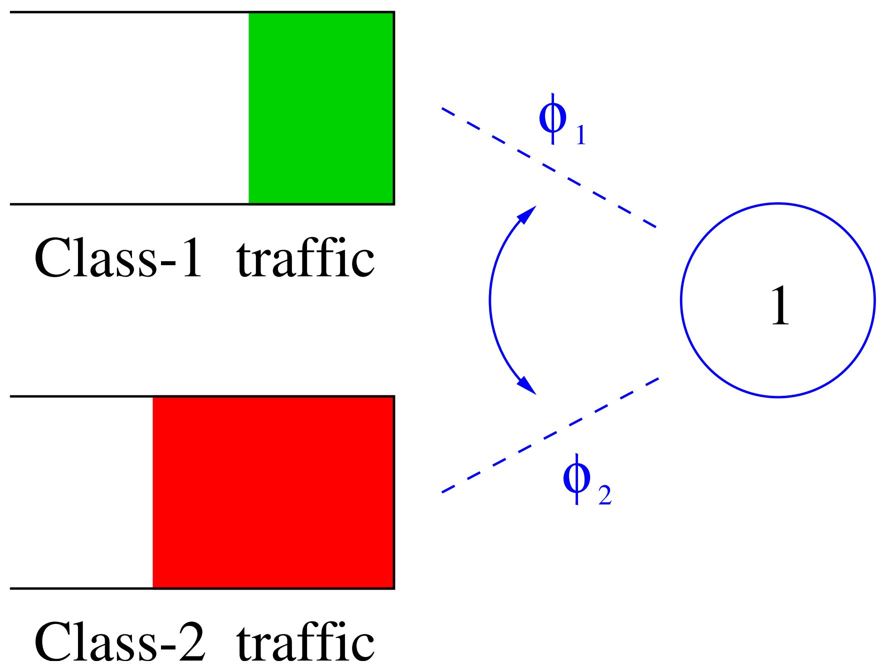
- Fayolle & Iasnogorodski (1979) consider exponential service times and reduce analysis of joint queue length distribution to Riemann-Hilbert problem

- Cohen & Boxma (1983) extend analysis to general service times and obtain joint workload distribution as solution to boundary-value problem

- Focus on exact large-buffer asymptotics for combination of heavy-tailed and light-tailed traffic

Model description

Two classes sharing link of unit rate



Class i **is assigned weight** $\phi_i \geq 0$**, with** $\phi_1 + \phi_2 = 1$

If both classes are backlogged, then class i receives service at rate ϕ_i

Excess capacity is re-allocated to the other class

If both classes are backlogged, then class i receives service at rate ϕ_i

Excess capacity is re-allocated to the other class

$\rho_i =$ **traffic intensity of class i**

$\mathrm{V}_i^{GPS} =$ **stationary workload of class i**

Traffic assumptions

Class 1 has 'light-tailed' characteristics, e.g.,

- G/G/1 input with 'exponentially-bounded' service times
- Markov-modulated fluid input

Traffic assumptions

Class 1 has 'light-tailed' characteristics, e.g.,

- G/G/1 input with 'exponentially-bounded' service times
- Markov-modulated fluid input

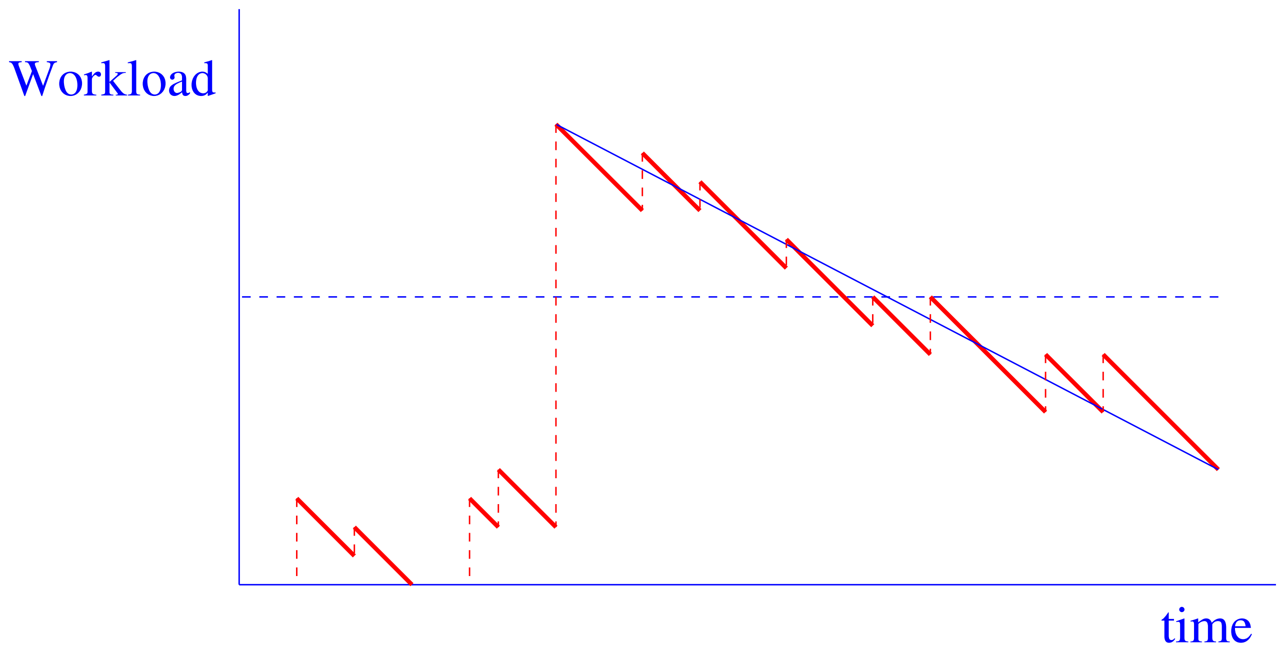
Class 2 has 'heavy-tailed' characteristics, e.g.,

- Instantaneous 'heavy-tailed' bursts B_2
- On-Off process with 'heavy-tailed' On-periods A_2 with fraction On-time p_2, peak rate r_2

Theorem [Cohen (1973), Pakes (1975)]

If B_i^r is subexponential, and $\rho_i < c$, then

$$\mathbb{P}\{\mathbf{V}_i^c > x\} \sim \frac{\rho_i}{c - \rho_i}\mathbb{P}\{\mathrm{B}_i^r > x\} \qquad \textbf{as } x \to \infty$$



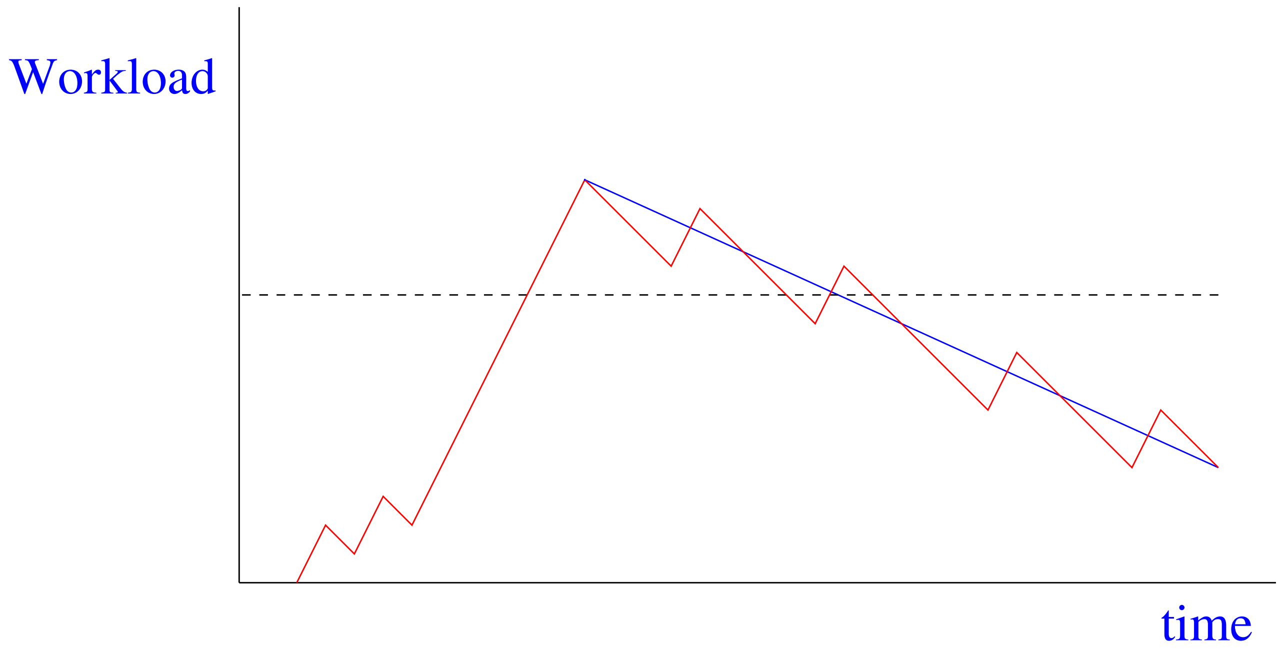
Catastrophe scenario:

Due to SINGLE extremely large burst

Theorem [Jelenković & Lazar (1999)]

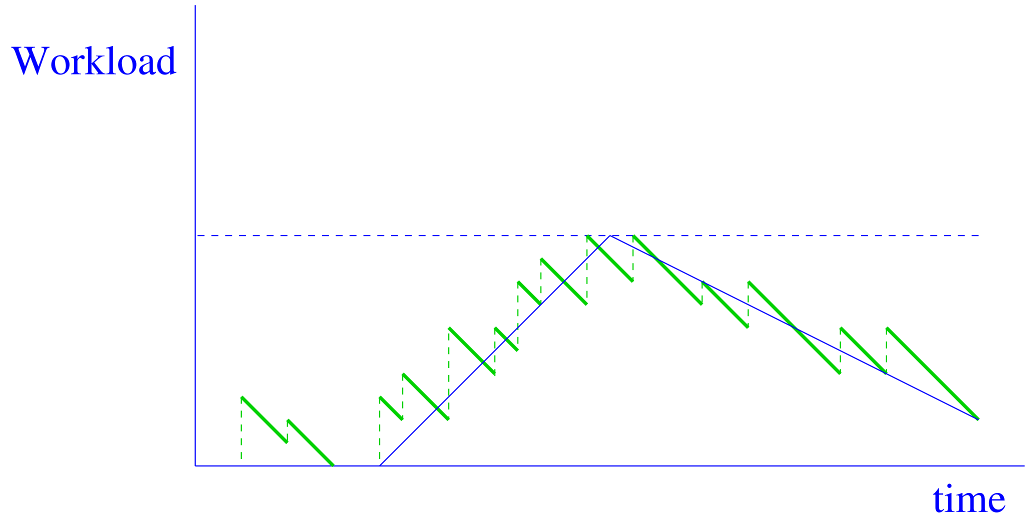
If A_i^r is subexponential, and $\rho_i < c < r_i$, then

$$\mathbb{P}\{\mathbf{V}_i^c > x\} \sim (1 - p_i)\frac{\rho_i}{c - \rho_i}\mathbb{P}\{\mathbf{A}_i^r > x/(r_i - c)\} \qquad \textbf{as } x \to \infty$$



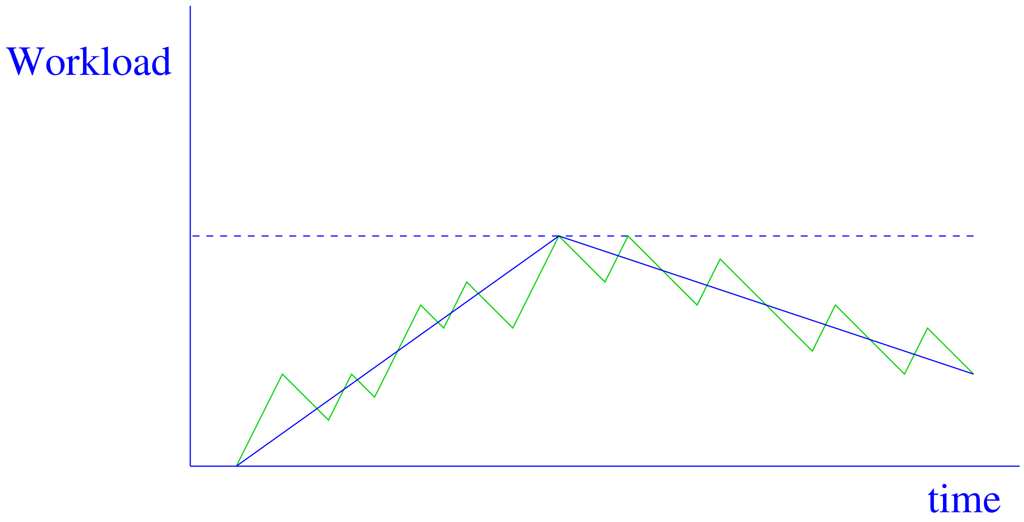
Due to SINGLE extremely long On-period

In contrast, class-1 builds up large workload level in gradual manner



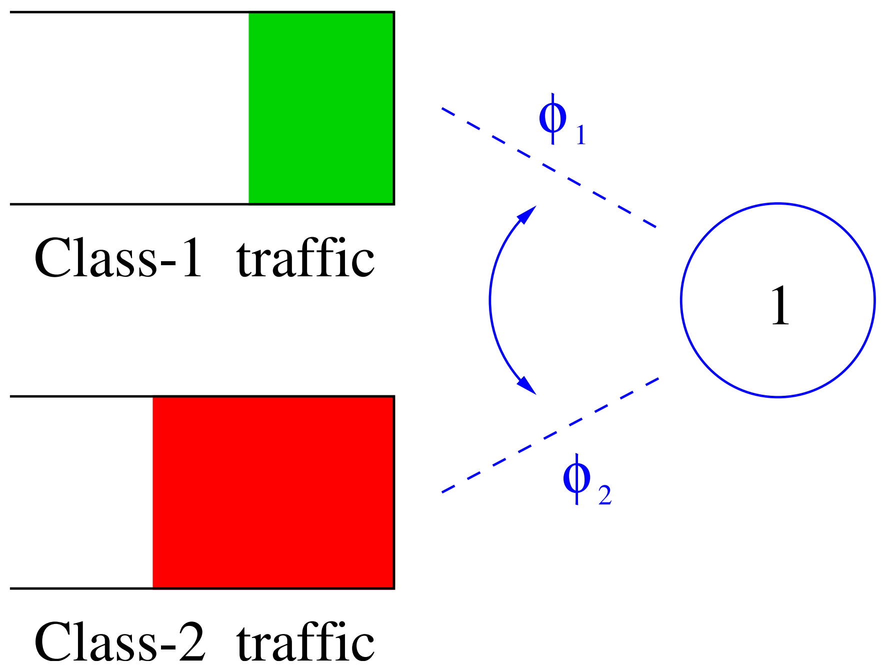
Conspiracy scenario:

Combination of MANY relatively large bursts and MANY relatively short interarrival times



Combination of MANY relatively long On-periods and MANY relatively short Off-periods

Workload asymptotics in various scenarios



- $\rho_1 < \phi_1$, $\rho_2 < \phi_2$

- $\rho_1 > \phi_1$, $\rho_2 < \phi_2$

- $\rho_1 < \phi_1$, $\rho_2 > \phi_2$

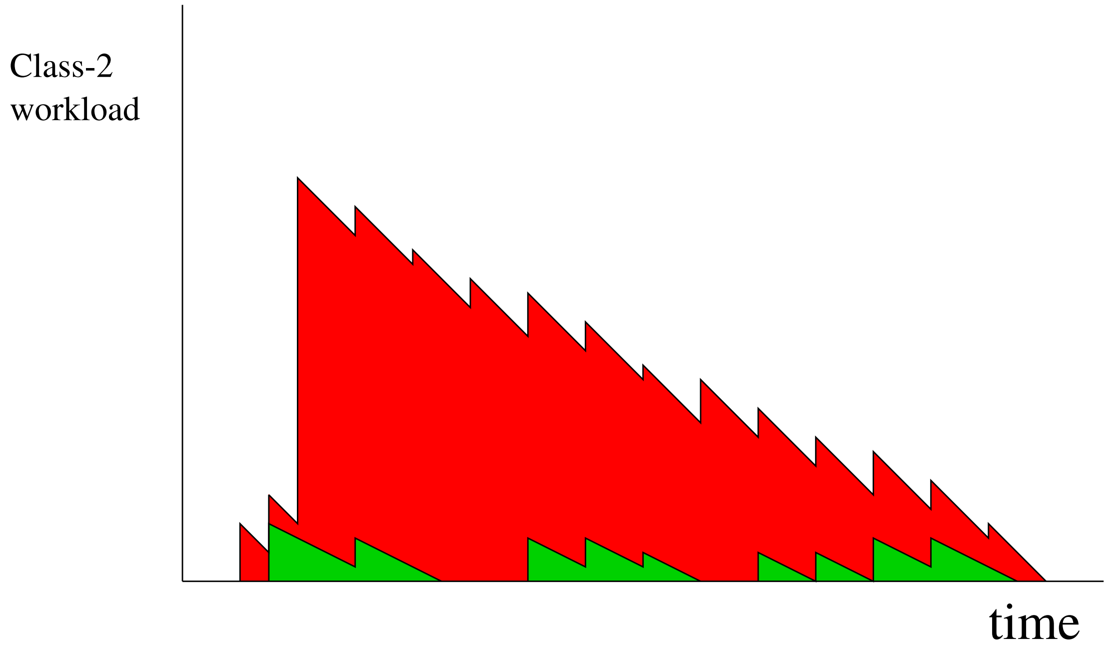
Workload asymptotics in various scenarios

Class-2 workload behavior

Case I: $\rho_1 < \phi_1$, $\rho_2 < \phi_2$

Catastrophe scenario:

- **Class 2 generates large burst (or long On-period)**
- **Class 1 generates traffic at rate** $\rho_1 < \phi_1$
- **Class 2 is effectively served at rate** $1 - \rho_1$


Class-2
workload
time

Theorem

If A_2^r or B_2^r is regularly varying, $\rho_1 < \phi_1$ and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \text{as } x \to \infty$$

Reduced-load equivalence (RLE)

Similar behavior has been shown for total workload in queues fed by mixture of heavy-tailed and light-tailed input [Agrawal, Nain & Makowski (1999), Zwart, B & Mandjes (2001)]

Note: here *independent* of class-1 traffic characteristics

Sample path lower bound

$$V_i^{GPS}(t) \geq V_i^{1-\rho_{-i}+\delta}(t) - \underbrace{U_{-i}^{\rho_{-i}-\delta}(t) - \sum_{j \neq i} V_j^{\phi_j}(t)}_{\text{"small correction terms"}}$$

Proof

Sample path wise,

$$
\begin{aligned}
V_i^{GPS}(t) &= V^{GPS}(t) - \sum_{j \neq i} V_j^{GPS}(t) \\
&\overset{Min-rate \;\; guarantee}{\geq} V^{GPS}(t) - \sum_{j \neq i} V_j^{\phi_j}(t) \\
&\overset{Work-conservation}{=} \sup_{0 \leq s \leq t} \{A(s,t) - (t-s)\} - \sum_{j \neq i} V_j^{\phi_j}(t) \\
&\geq \sup_{0 \leq s \leq t} \{A_i(s,t) - (1-\theta)(t-s)\} \\
&- \sup_{0 \leq s \leq t} \{\theta(t-s) - A_{-i}(s,t)\} - \sum_{j \neq i} V_j^{\phi_j}(t) \\
&= V_i^{1-\theta}(t) - U_{-i}^{\theta}(t) - \sum_{j \neq i} V_j^{\phi_j}(t)
\end{aligned}
$$

Then take $\theta = \rho_{-i} - \delta$

Sample path upper bound

$$V_i^{GPS}(t) \leq \min\{V_i^{\phi_i}(t), V_i^{1-\rho_{-i}-\delta}(t) + \underbrace{V_{-i}^{\rho_{-i}+\delta}(t)}_{\text{“correction term”}}\}$$

Proof

Sample path wise,

$$
\begin{aligned}
V_i^{GPS}(t) &\leq V^{GPS}(t) \\
&\overset{Work-conservation}{=} \sup_{0\leq s\leq t}\{A(s,t)-(t-s)\} \\
&\leq \sup_{0\leq s\leq t}\{A_i(s,t)-(1-\theta)(t-s)\} \\
&+ \sup_{0\leq s\leq t}\{A_{-i}(s,t)-\theta(t-s)\} \\
&= V_i^{1-\theta}(t)+V_{-i}^{\theta}(t)
\end{aligned}
$$

Also,

$$V_i^{GPS}(t) \overset{Min-rate\ \ guarantee}{\leq} V_i^{\phi_i}(t)$$

Then take $\theta = \rho_{-i}+\delta$

Want to show

If A_2^r **or** B_2^r **is regularly varying,** $\rho_1 < \phi_1$ **and** $\rho_2 < \phi_2$**, then**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \text{as } x \to \infty$$

Want to show

If A_2^r **or** B_2^r **is regularly varying,** $\rho_1 < \phi_1$ **and** $\rho_2 < \phi_2$**, then**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \textbf{as } x \to \infty$$

Proof (sketch)

From sample path lower bound, for any $\delta > 0$ **and** y**,**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \geq \mathbb{P}\{\mathbf{V}_2{}^{1-\rho_1+\delta} > x + y\}\mathbb{P}\{\mathbf{U}_1^{\rho_1-\delta} + \mathbf{V}_1^{\phi_1} < y\}$$

Want to show

If A_2^r **or** B_2^r **is regularly varying,** $\rho_1 < \phi_1$ **and** $\rho_2 < \phi_2$**, then**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \text{as } x \to \infty$$

Proof (sketch)

From sample path lower bound, for any $\delta > 0$ **and** y**,**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \geq \mathbb{P}\{\mathbf{V}_2{}^{1-\rho_1+\delta} > x + y\}\mathbb{P}\{\mathbf{U}_1^{\rho_1-\delta} + \mathbf{V}_1^{\phi_1} < y\}$$

From sample path upper bound, for any $\delta > 0$ **and** y**,**

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \leq \mathbb{P}\{\mathbf{V}_2{}^{1-\rho_1-\delta} > x - y\} + \mathbb{P}\{\mathbf{V}_2^{\phi_2} > x\}\mathbb{P}\{\mathbf{V}_1^{\rho_1+\delta} > y\}$$

Want to show

If A_2^r or B_2^r is regularly varying, $\rho_1 < \phi_1$ and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \textbf{as } x \to \infty$$

Proof (sketch)

From sample path lower bound, for any $\delta > 0$ and y,

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \geq \mathbb{P}\{\mathbf{V}_2{}^{1-\rho_1+\delta} > x + y\}\mathbb{P}\{\mathbf{U}_1^{\rho_1-\delta} + \mathbf{V}_1^{\phi_1} < y\}$$

From sample path upper bound, for any $\delta > 0$ and y,

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \leq \mathbb{P}\{\mathbf{V}_2{}^{1-\rho_1-\delta} > x - y\} + \mathbb{P}\{\mathbf{V}_2^{\phi_2} > x\}\mathbb{P}\{\mathbf{V}_1^{\rho_1+\delta} > y\}$$

Show that, for $y \to \infty$, $\delta \downarrow 0$, both bounds behave as

$$\mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\}$$

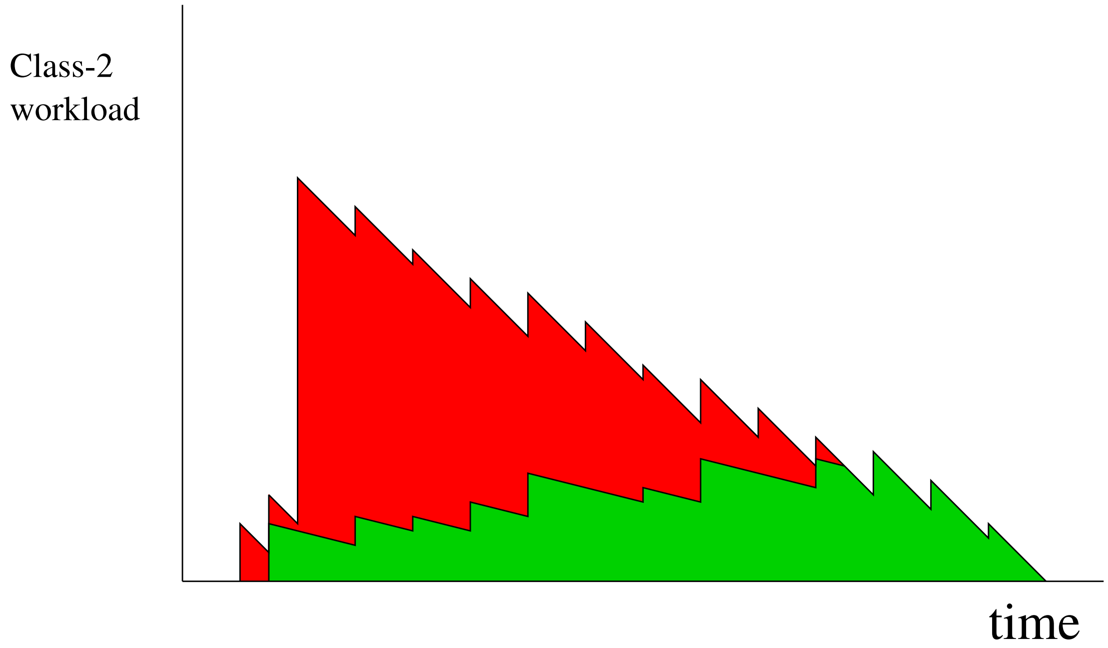
Requires that A_2^r or B_2^r is regularly varying

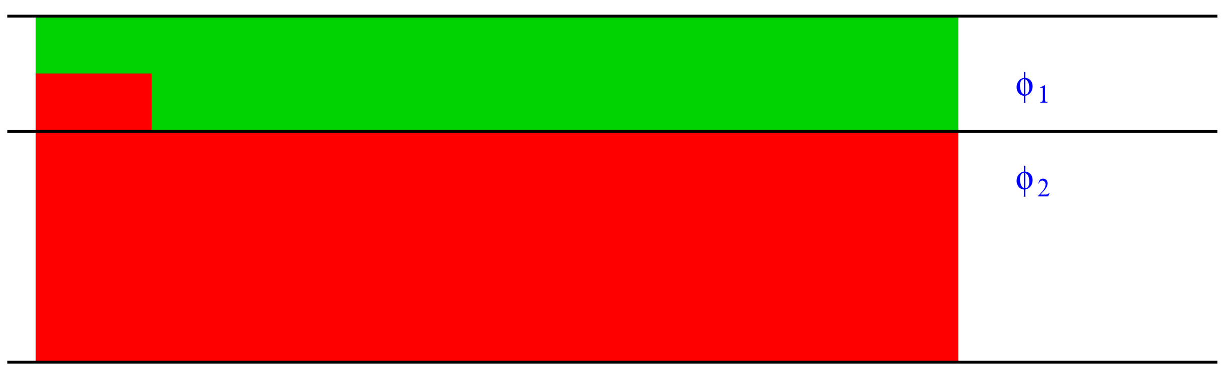
Class-2 workload behavior (cont'd)

Case II: $\rho_1 > \phi_1$, $\rho_2 < \phi_2$

Catastrophe scenario:

- **Class 2 generates large burst (or long On-period)**
- **Class 1 generates traffic at rate $\rho_1 > \phi_1$, but only receives service at rate ϕ_1**
- **Class 2 is effectively served at rate $\phi_2 = 1 - \phi_1$**


Class-2
workload
time


ϕ_1
ϕ_2

Theorem

If A_2^r or B_2^r is regularly varying, $\rho_1 > \phi_1$, and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{\phi_2} > x\} \qquad \text{as } x \to \infty$$

Reduced-weight equivalence (RWE):

Qualitatively similar to reduced-*load* equivalence in previous case

Note: *independent* **of class-1 traffic characteristics**

Class-2 workload behavior (cont'd)

Case III: $\rho_1 < \phi_1$, $\rho_2 > \phi_2$

Catastrophe scenario:

- **Class 2 generates large burst (or long On-period)**
- **Class 1 generates traffic at rate** $\rho_1 < \phi_1$
- **Class 2 is effectively served at rate** $1 - \rho_1$

Theorem

If A_2^r or B_2^r is regularly varying, $\rho_1 < \phi_1$, and $\rho_2 > \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_2^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_2^{1-\rho_1} > x\} \qquad \text{as } x \to \infty$$

Reduced-load equivalence (RLE):

Qualitatively similar as in previous two cases

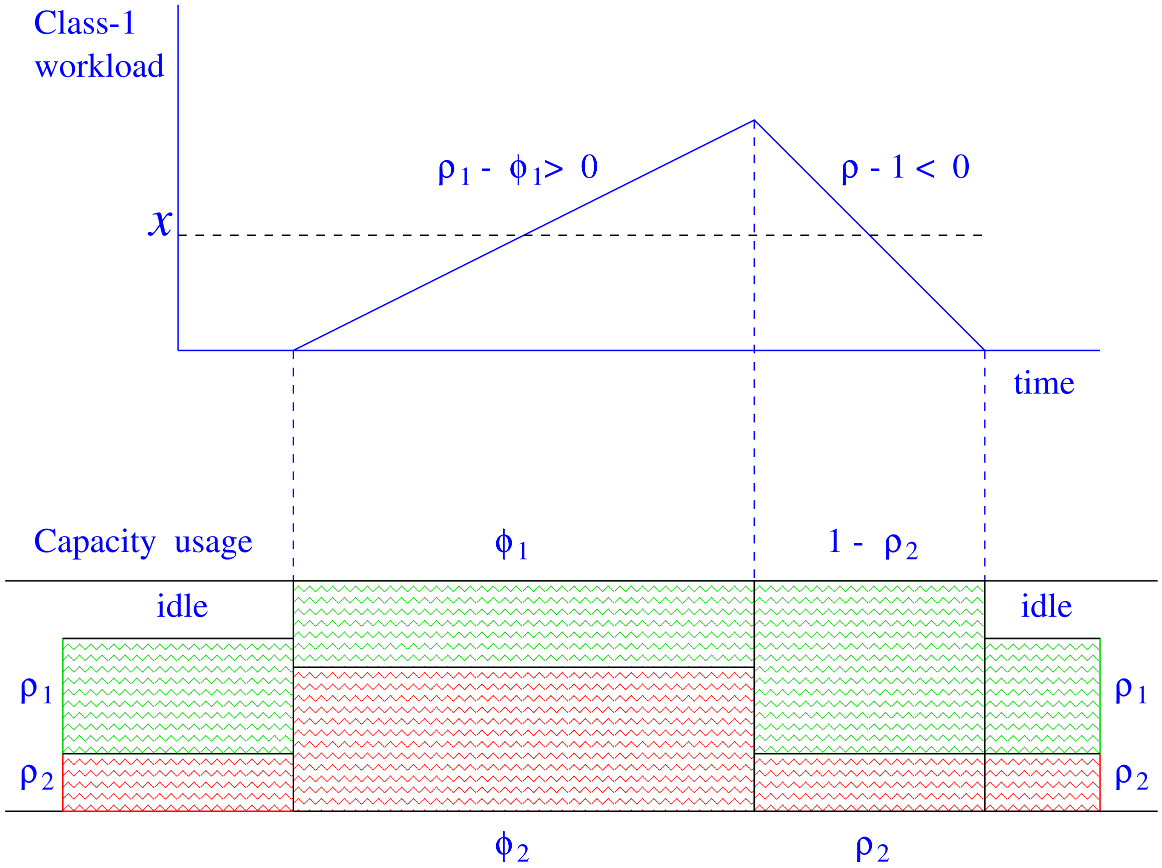
However, in contrast to previous two cases, now *it is* crucial that class-1 traffic is 'lighter'-tailed than class-2 traffic

Class-1 workload

Case I: $\rho_1 > \phi_1$, $\rho_2 < \phi_2$

Catastrophe scenario:

- **Class 2 generates large burst (or long On-period)**
- **Leaves only service rate** $\phi_1 = 1 - \phi_2$ **for class 1**
- **Class 1 generates traffic at rate** $\rho_1 > \phi_1$
- **Class-1 workload builds up at rate** $\rho_1 - \phi_1 > 0$



Class-1
workload
$\rho_1 - \phi_1 > 0$
$\rho - 1 < 0$
x
time
Capacity usage
ϕ_1
$1 - \rho_2$
idle
idle
ρ_1
ρ_1
ρ_2
ρ_2
ϕ_2
ρ_2

Theorem

If B_2^r is regularly varying, $\rho_1 > \phi_1$ and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim \frac{\phi_2 - \rho_2}{\phi_2} \frac{\rho_2}{1 - \rho_1 - \rho_2} \mathbb{P}\{\mathbf{P}_2^r > \frac{x}{\rho_1 - \phi_1}\},$$

with P_2^r residual class-2 busy period when served at rate ϕ_2

Induced burstiness (IB):
Class-1 workload behaves as that of heavy-tailed On-Off process with as On-periods the class-2 busy periods, and inherits ill-behaved class-2 characteristics

Class-1 workload behavior (cont'd)

Case II: $\rho_1 < \phi_1$, $\rho_2 < \phi_2$

Class 1 remains stable even when class 2 is backlogged, so previous catastrophe scenario can no longer occur

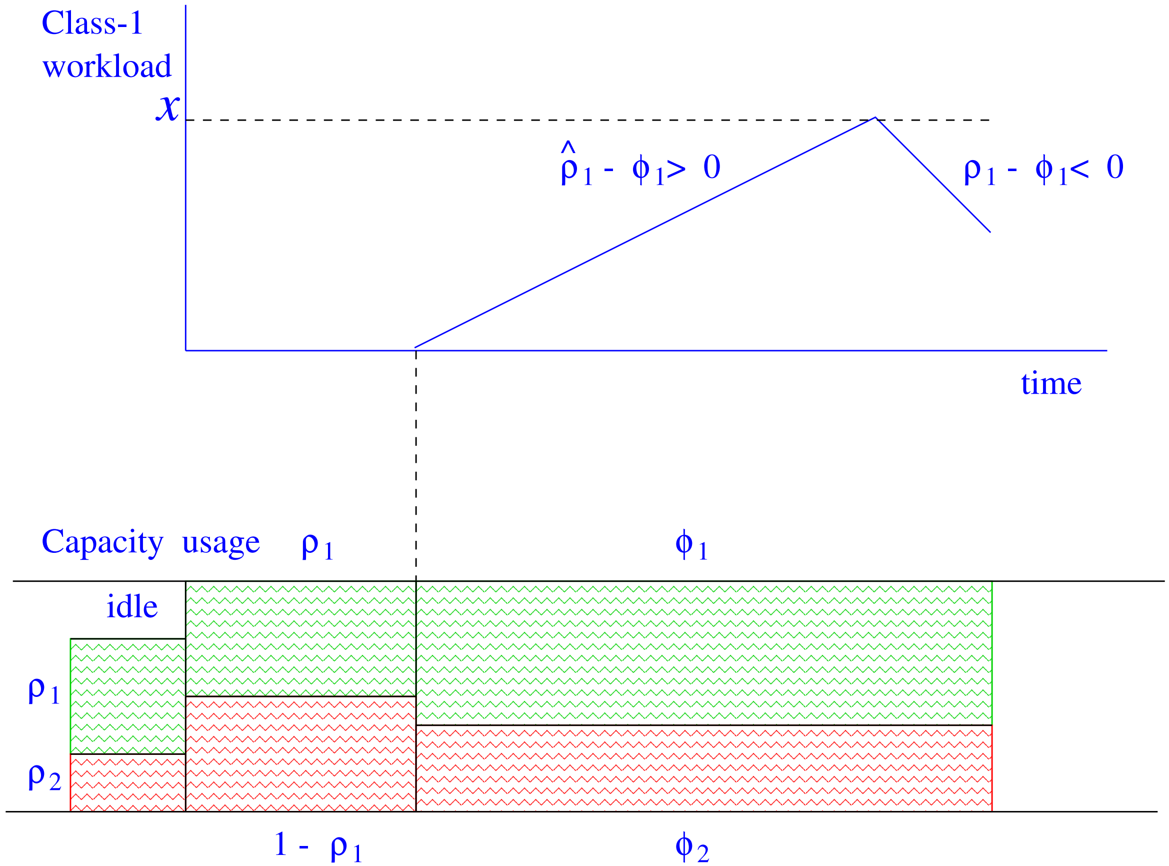
Class 1 too must show abnormal activity in order for large workload to build up

Recall class 1 in isolation builds up large workload in gradual manner by deviating from its normal traffic intensity for long period

Conspiracy scenario:

- **Class 1 shows similar abnormal activity as in isolation, raising its traffic intensity to** $\hat{\rho}_1 > \phi_1$ **for period** $\frac{x}{\hat{\rho}_1 - \phi_1}$

- **During that period, class 2 remains constantly back-logged, leaving service rate** $\phi_1 = 1 - \phi_2$ **for class 1**


Class-1
workload
x
$\hat{\rho}_1 - \phi_1 > 0$
$\rho_1 - \phi_1 < 0$
time
Capacity usage
ρ_1
ϕ_1
idle
ρ_1
ρ_2
$1 - \rho_1$
ϕ_2

Theorem

If B_2^r is regularly varying, $\rho_1 < \phi_1$ and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_1^{\phi_1} > x\}\mathbb{P}\{\mathbf{T}_2 > \frac{x}{\hat{\rho}_1 - \phi_1}\},$$

with T_2 'drain' time of class 2 when served at rate ϕ_2 with initial workload $\mathrm{V}_2^{1-\rho_1}$

Reduced-weight equivalence (RWE):

Theorem

If B_2^r is regularly varying, $\rho_1 < \phi_1$ and $\rho_2 < \phi_2$, then

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_1^{\phi_1} > x\}\mathbb{P}\{\mathbf{T}_2 > \frac{x}{\hat{\rho}_1 - \phi_1}\},$$

with T_2 'drain' time of class 2 when served at rate ϕ_2 with initial workload $\mathrm{V}_2^{1-\rho_1}$

Reduced-weight equivalence (RWE):
but now *major* contribution from deviant class-2 behavior

Similar behavior has been shown for total workload in queues fed by mixture of heavy-tailed and light-tailed input [B & Zwart (2000)] and various related models [Boxma, Deng & Zwart (2002), Boxma & Kurkova (2000)]

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_1^{\phi_1} > x\}\mathbb{P}\{\mathbf{T}_2 > \frac{x}{\hat{\rho}_1 - \phi_1}\}$$

First term represents upper bound for class 1 based on minimum-rate guarantee ϕ_1,
and captures deviant behavior of class 1 itself

Second term reflects that class 2 must remain backlogged long enough for class-1 workload to build up,
and provides measure for gains from sharing surplus capacity with class 2

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim \mathbb{P}\{\mathbf{V}_1^{\phi_1} > x\}\mathbb{P}\{\mathbf{T}_2 > \frac{x}{\hat{\rho}_1 - \phi_1}\}$$

First term represents upper bound for class 1 based on minimum-rate guarantee ϕ_1,
and captures deviant behavior of class 1 itself

Second term reflects that class 2 must remain backlogged long enough for class-1 workload to build up,
and provides measure for gains from sharing surplus capacity with class 2

General decompositional form holds irrespective of detailed traffic characteristics of two classes

Instantaneous input

$$\mathbb{P}\{\mathbf{T}_2 > x\} \sim \frac{\rho_1}{1-\rho_1-\rho_2}\mathbb{P}\{\mathbf{B}_2^r > (\phi_2 - \rho_2)x\}$$

Class 2 must remain backlogged for period of length x

Normally generates traffic at rate ρ_2

Receives service at rate ϕ_2 while class-1 workload builds up

Instantaneous input (cont'd)

Class 2 needs to make up for 'deficit' amount $(\phi_2 - \rho_2)x$

Enjoys service at rate $1 - \rho_1$ before that

Most likely scenario: initial $V_2^{1-\rho_1}$ exceeds $(\phi_2 - \rho_2)x$ (due to earlier large burst), which occurs with probability

$$\mathbb{P}\{V_2^{1-\rho_1} > (\phi_2 - \rho_2)x\} \sim \frac{\rho_2}{1 - \rho_1 - \rho_2} \mathbb{P}\{B_2^r > (\phi_2 - \rho_2)x\}$$

Instantaneous input (cont'd)

Class 2 needs to make up for 'deficit' amount $(\phi_2 - \rho_2)x$

Enjoys service at rate $1 - \rho_1$ before that

Most likely scenario: initial $V_2^{1-\rho_1}$ exceeds $(\phi_2 - \rho_2)x$ (due to earlier large burst), which occurs with probability

$$\mathbb{P}\{V_2^{1-\rho_1} > (\phi_2 - \rho_2)x\} \sim \frac{\rho_2}{1 - \rho_1 - \rho_2} \mathbb{P}\{B_2^r > (\phi_2 - \rho_2)x\}$$

Fluid input

Similar yet slightly more involved scenario

Class-1 workload behavior (cont'd)

Case III: $\rho_1 < \phi_1$, $\rho_2 > \phi_2$

Now class 2 remains constantly backlogged with probability O(1) while class-1 workload builds up

$$\mathbb{P}\{\mathbf{V}_1^{GPS} > x\} \sim K_2 \mathbb{P}\{\mathbf{V}_1^{\phi_1} > x\} \qquad \text{as } x \to \infty$$

Constant K_2 is difficult to determine

Reduced-weight equivalence (RWE):
but now *minor* contribution from deviant class-2 behavior

Discussion & conclusion

Various scenarios for qualitative behavior

- **Reduced-load equivalence (RLE):**
  class receives total rate reduced by load of other class

- **Reduced-weight equivalence – no effort (RWE-0):**
  class gets total rate reduced by weight of other class;
  other class shows average behavior (prob. 1)

- **Reduced-weight equivalence – minor effort (RWE-1):**
  class gets total rate reduced by weight of other class;
  other class shows minor deviant behavior (prob. O(1))

- **Reduced-weight equivalence – major effort (RWE-2):** class gets total rate reduced by weight of other class; other class shows major deviant behavior (prob. o(1))

- **Induced burstiness (IB):** class affected by other class, and inherits ill-behaved traffic characteristics

Heavy Tails: Performance Models and Scheduling Disciplines

Part II – Workload Asymptotics for Generalized Processor Sharing Systems

References:

Borst, S.C., Boxma, O.J., Jelenković, P.R. (2003). Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. Queueing Systems 43, 273–306.

M.J.G. van Uitert. Generalized Processor Sharing Queues. PhD thesis. Eindhoven University of Technology, 2003.

http://www.cwi.nl/∼sindo