Flow-level stability and performance of channel-aware priority-based schedulers

Samuli Aalto and Pasi Lassila

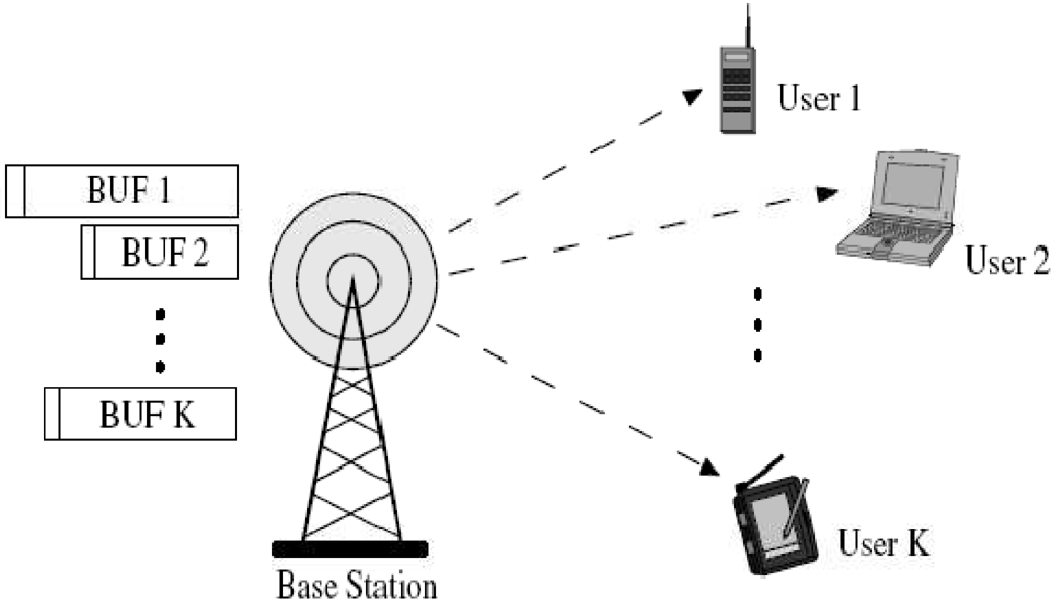
COMNET Department

TKK Helsinki University of Technology

Email: {Samuli.Aalto, Pasi.Lassila}@tkk.fi

Background (1)

– Flow-level scenario
  - Users around a BS
  - Users download files
  - Each file requires many time slots of service
  - Number of active users varies randomly



– HDR systems
  - BS transmits to one user in a time slot with full power
  - BS knows the channel quality
  - Decides on suitable coding to match channel state
  - Challenge: channel state varies randomly due to fading phenomena

Background (2)

- Consider system when time slot length → 0
- Assume that rate variations average out at flow time scale
  - System corresponds to M/G/1 queue
  - SRPT is optimal policy for minimizing mean flow delay
- Channel-aware scheduling
  - Base station knows instantaneous channel state of all active users
  - Can favor those users having instantaneously good channels
  - Analysis with a static number of users (w/wo packet-level dynamics)
    - Queue length-based policies shown to have many desirable properties (Stolyar 2005, Mandelbaum & Stolyar 2004,…)
  - Not much work on dynamic setting
    - Stability: seminal work by Borst 2005 and Jonckheere & Borst 2006
    - Minimizing mean delay very difficult and hardly anything is known

Overview

- We study so called priority-based channel–aware schedulers
  - Priority can be any strictly increasing function of instantaneous rate
  - Includes as special cases many proposed channel-aware schedulers
- Stability
  - Achieving maximum stability region is a robustness property
  - We give the general condition when necessary condition is also sufficient
  - When necessary condition is not sufficient, we give the sufficient condition for some special cases
- Performance
  - Simulation studies to gain insight on actual performance (including comparisons against α-fair policies)

Model and assumptions

- Traffic assumptions
  - Consider K classes of flows (or users)
  - Classes correspond to flows with different channel properties
  - Flows arrive according to a Poisson process with rate λ_k
  - Flow sizes, X_i, are i.i.d. with mean $\bar{x} = E[X_i]$

- Channel assumptions
  - Class-k rates vary independently according to a stationary process $R_k(t)$
  - Base station knows instantaneous flow rates and channel statistics
  - Discrete set of possible rates $\{r_1,\ldots,r_J\}$ (in HDR systems $J = 11$)
  - Each class has its own set R_k with $F_k(r) = P\{R_k \le r\}$
  - $\bar{r}_k = E[R_k]$, $r_k^* = \max \mathcal{R}_k$, $r_k^{**} = \max(\mathcal{R}_k \setminus \{r_k^*\})$

- Traffic load: $\rho_k^* = \dfrac{\lambda_k \bar{x}}{r_k^*}$ and $\rho^* = \rho_1^* + \ldots + \rho_K^*$

Rate-based priority scheduler

- Let $h_k(r)$ denote any strictly increasing function of instantaneous rate r
- Priority of flow i in class k

$$P_i(t) = h_k(R_i(t))$$

- Priority-based scheduler selects user i^* at time t for which

$$P_{i^*}(t) = \max_{i \in \mathcal{N}(t)} P_i(t)$$

- The set of possible priorities of class-k flows is discrete

$$\mathcal{P}_k = \{h_k(r) : r \in \mathcal{R}_k\}$$

- Priority class = all flows with same priority
  - Within a priority class there may be flows from different classes
  - We allow ties to occur between user classes

Special cases of rate-based priority scheduler

- Linear weight-based strategies so that $h_k(r) = w_k\, r$
  - Absolute rate priority: $w_k = 1$
  - Relative rate priority: $w_k = 1/\bar{r}_k$
  - Proportional rate priority: $w_k = 1/r_k^*$

- Non-linear weight
  - CDF-based priority: $h_k(r) = F_k(r)^{1/w_k}$

- Tie breaking
  - Randomized tie breaking ⇒ MR, RB, PB, CS (with $w_k = 1$)
  - Possible to use, e.g., information about remaining size Y_i

Stability under necessary condition (1)

– Necessary condition for channel aware-schedulers (B&J 2006)

$$\rho^* = \rho_1^* + \ldots + \rho_K^* \leq 1$$

– Utility-based policies

- Utility U depends on flow throughput $T_i(t)$ and the policy selects flow with

$$i^* = \arg \max_{i \in \mathcal{N}(t)} R_i(t)\, U'(T_i(t))$$

- For fixed nof flows, asymptotically maximizes $\sum_i U(T_i(t))$
- Utility-based policies are stable under (B&J 2006)

$$\rho^* < 1$$

- When is the above sufficient also for priority-based policies?

Stability under necessary condition (2)

- Some notation
  - Highest priority in class *k*, $p_k^* = h_k(r_k^*) = \max \mathcal{P}_k$
  - 2nd highest priority in class *k*, $p_k^{**} = h_k(r_k^{**}) = \max(\mathcal{P}_k \setminus \{p_k^*\})$

- **Theorem 1**: If $p_k^* > p_l^{**}$ for all $k \neq l$, then the rate-based priority policy is stable when

$$\rho^* < 1 \qquad (1)$$

  - Proof based on showing that $\rho^* \geq 1$ when the system is unstable

- Theorem 1 implies that at stability limit the scheduler always serves a flow with its class specific maximum rate.

Comments on Theorem 1

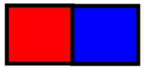
– Ties may occur between priorities of different classes

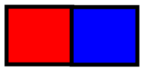
– Tie-breaking rule can be any work-conserving policy

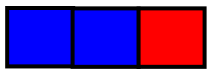
– Corollaries:
  - *Corollary 1*: Proportional rate policy is stable under (1).
  - *Corollary 2*: CDF-based priority policy is stable under (1).
  - *Corollary 3*: If $r_k^* = r_J$ for all classes, absolute rate priority policy is stable under (1).
  - *Corollary 4*: If $r_k^* / \bar{r}_k > r_l^{**} / \bar{r}_l$ for all $k \neq l$, then any relative rate priority policy is stable under (1).

Illustration of Theorem 1 with $K = 2$

- $R_1 = R_2 = \{1,2,4,8\}$, $p_1 = \{0.1,0.2,0.3,0.4\}$, $p_2 = \{0.9,0.05,0.03,0.02\}$

- Absolute rate
  - $P_1 = \{8, 4, 2, 1\}$, $P_2 = \{8, 4, 2, 1\}$
  - highest indexes :  = stable

- Proportional rate
  - $P_1 = \{1, 0.5, 0.25, 0.125\}$, $P_2 = \{1, 0.5, 0.25, 0.125\}$
  - highest indexes :  = stable

- Relative rate
  - $P_1 = \{0.20, 0.41, 0.82, 1.63\}$, $P_2 = \{0.78, 1.56, 3.13, 6.25\}$
  - highest indexes :  = unstable

Sufficient conditions with randomized tie breaking (1)

– Assume that ties within a priority class are broken at random

– **Theorem 2**: If $p_k^* \geq p_l^{**}$ for all $k \neq l$, then the rate-based priority policy where ties are broken at random is stable when

$$\rho^* < 1$$

Sufficient conditions with randomized tie breaking (2)

– Consider the special case with $K = 2$ and randomized tie breaking

– **Theorem 3**: If $p_1^* < p_2^{**}$, then the rate-based priority policy is stable under the condition

$$\rho_1^* + P\{h_2(\tilde{M}_2) > p_1^*\} < 1$$

- If $p_1^* < p_2^{**}$, class 1 becomes unstable
- At stability limit, class 1 is served at r_1^* when scheduled
- Class 2 flows can "beat" class 1 flows whenever for some flow i in class 2 we have that $P_i(t) > p_1^* \Rightarrow$ loss in efficiency
- $P\{h_2(\tilde{M}_2) > p_1^*\}$ represents proportion of time that class-2 flows are served in a hypothetical reference system where $N_1(t) \to$ ◎
- Determined by an M/G/1-PS queue with state-dependent service rates

Impact of continuous rate distributions

– Consider the case where the rate distribution becomes continuous

– Proportional rate and CDF-based schedulers are still stable under (1)
  - Follows from Corollaries 1 and 2

– Absolute rate priority policy is stable under (1) if rate distributions $F_k(r)$ of all classes have same support.

– Relative rate priority policies are not stable under (1).

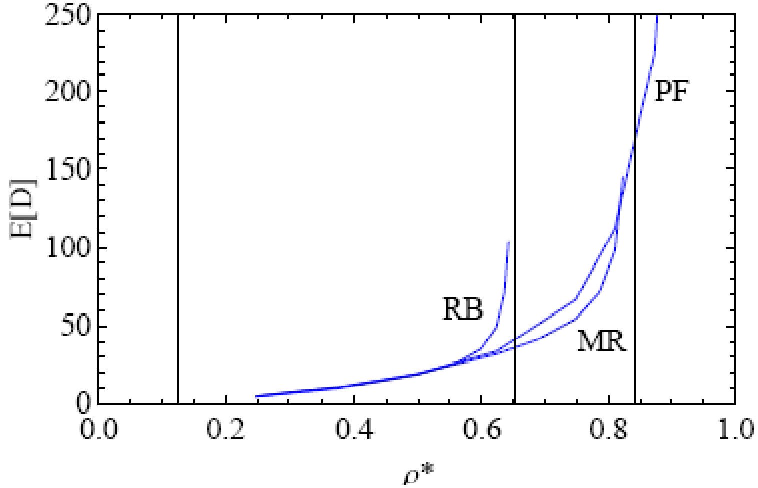
Numerical examples

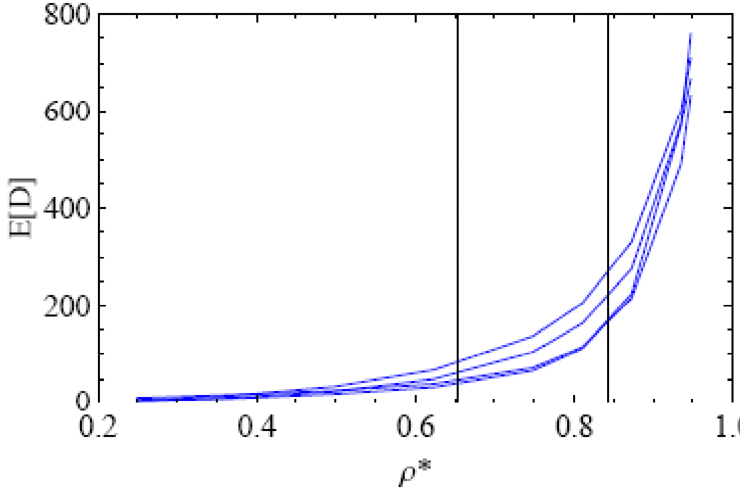
- Study performance of
  - Priority based: MR, RB, PB and CS
  - Utility based: PF (Proportional Fair, $\alpha = 1$) and PD (Potential Delay, $\alpha = 2$)
  - Also, investigate impact of using SRPT-like tie-breaking rules
- Parameters
  - 2 classes, flows arrive according to Poisson process with $\lambda_1 = \lambda_2 = 0.5$
  - HDR rates, i.e., $J = 11$
  - Class 1 flows can achieve 7 lowest rates
  - Class 2 flows can achieve all 11 rates
  - Rate distributions obey truncated geometric distr. with $q_1 = 1$, $q_2 = 0.5$

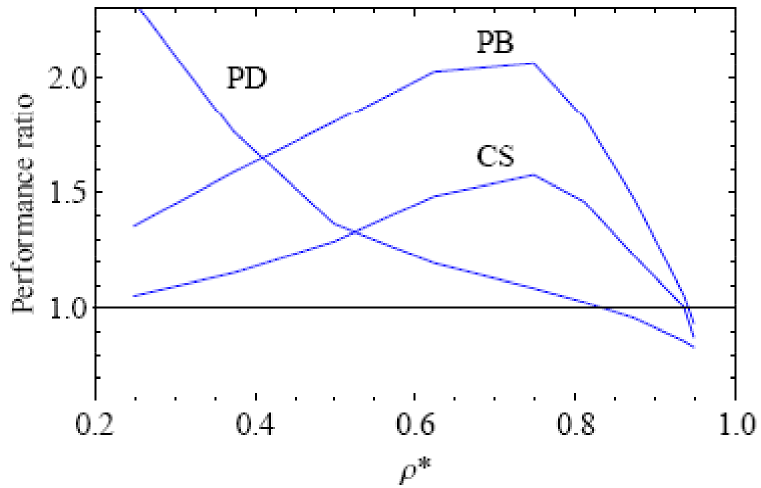
$$P\{R_1 = r_j\} = \frac{q_1^j}{\sum_{n=1}^{j_1} q_1^n}, \quad P\{R_2 = r_j\} = \frac{q_2^j}{\sum_{n=1}^{J} q_2^n}$$

  - With these parameters $p_1^* < p_2^{**}$ for both MR and RB

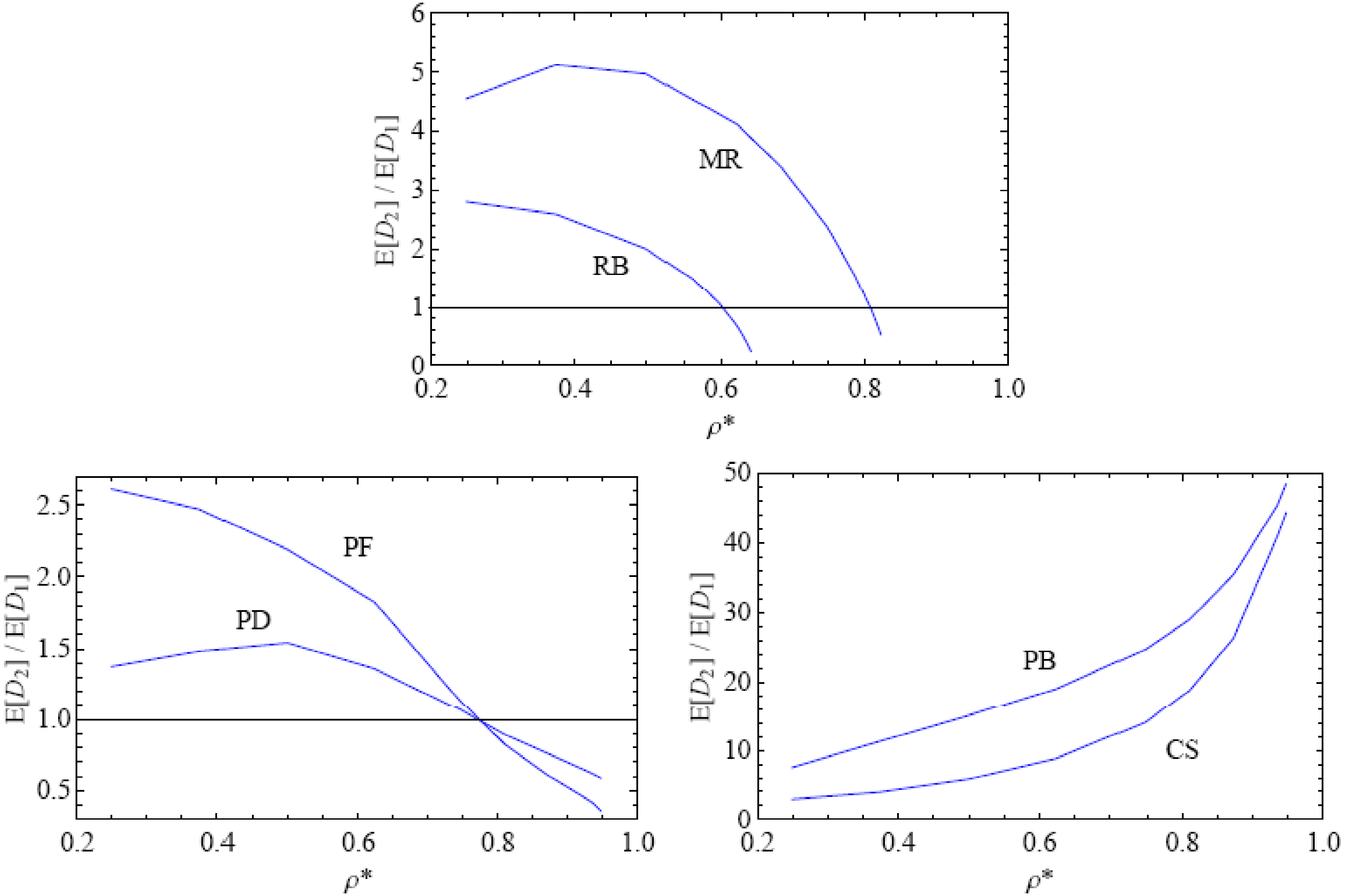
Overall performance (mean delay)





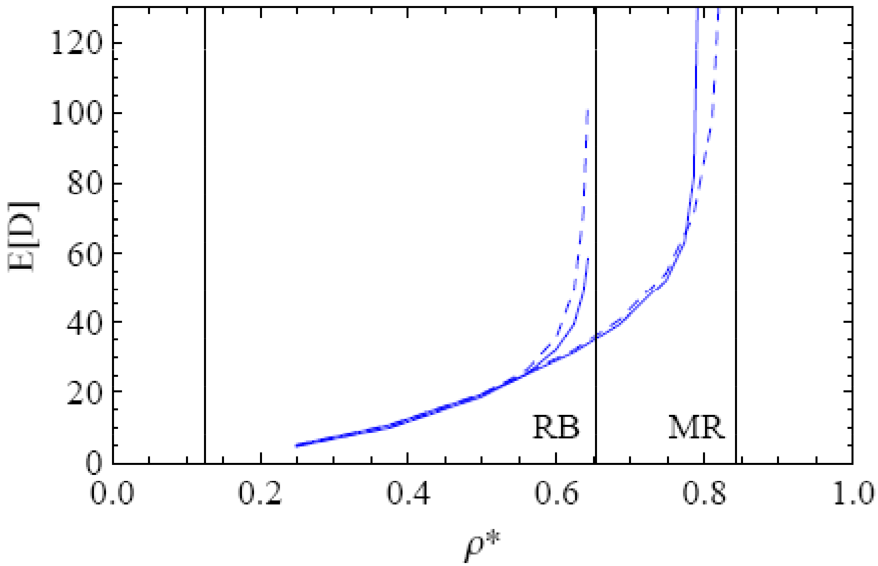


Fairness

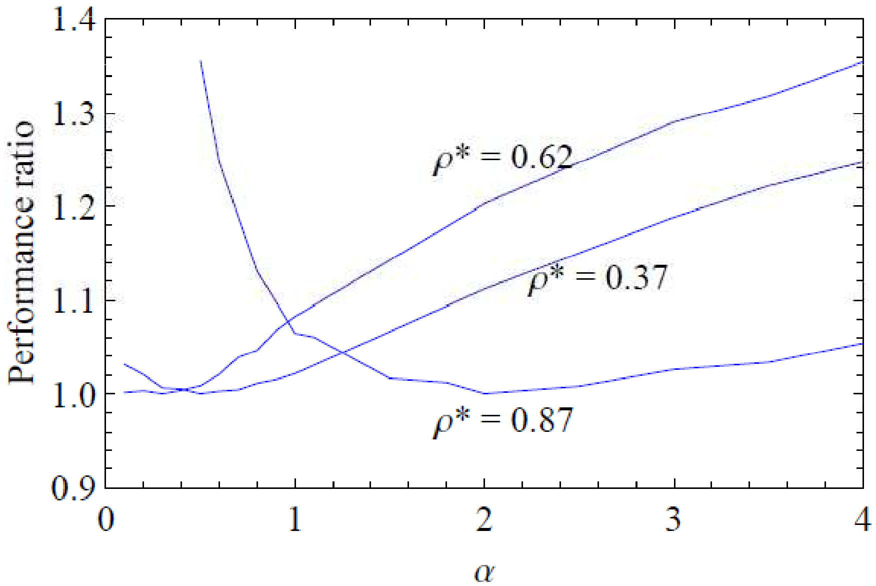

6
5
4
3
2
1
0
0.2
0.4
0.6
0.8
1.0
ρ^*
$E[D_2] / E[D_1]$
MR
RB
2.5
2.0
1.5
1.0
0.5
PF
PD
50
40
30
20
10
PB
CS

Other performance comparisons

Impact of SRPT-like tie breaking



Optimizing α-parameter



Conclusions

– Stability
  - Conditions under which necessary condition is sufficient for general rate-based priority policies
  - Stricter sufficient conditions for some special cases

– Performance
  - MR and RB offer quite good performance, but may become unstable
  - PB and CS policies are very unfair (although stable)
  - PF performs very well over a large region of loads (good overall)
  - PD can outperform PF at very high loads
  - SRPT-like tie-breaking heuristics do not work at the time-slot level
  - To minimize mean delay, flow-level information can be used to tune the packet level schedulers (cf., tests with impact of α)