On the compromise between burstiness and frequency of events

Alain Jean-Marie Yvan Calas Tigist Alemu

INRIA/LIRMM CNRS U. of Montpellier 2
ID-IMAG U. Grenoble
ajm@lirmm.fr http://www.lirmm.fr/~ajm

Presented at
Performance 2005, 7 october 2005
with minor corrections

Contents

The Problem (1)

Consider a process of events with two types: good and lost.

n consecutive events are called a block.

Time may be continuous

but the model will be in discrete-time and ignore actual time intervals between events.

The Problem (1)

Consider a process of events with two types: good and lost.

n consecutive events are called a block.

Time may be continuous

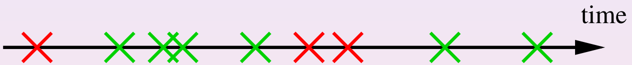
but the model will be in discrete-time and ignore actual time intervals between events.

The Problem (1)

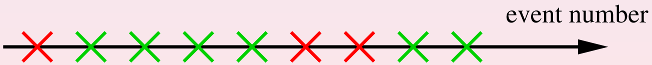
Consider a process of events with two types: good and lost.

n consecutive events are called a block.

Time may be continuous



but the model will be in discrete-time and ignore actual time intervals between events.



The Problem (2)

Metric of interest: given h and n,

$$P(\text{ the block is ``lost'' }) \quad = \quad P(\, > h \text{ losses among } n \text{ events})$$

Usual objective: find the smallest h (redundancy) such that

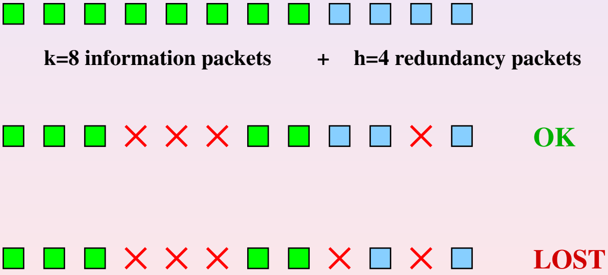
$$P(\, > h \text{ losses among } n + h \text{ events}) \quad < \quad \varepsilon\ .$$

Today's objective: compare two situations

- same event loss probability
- different "burstiness" patterns

Motivation #1: Forward Error Correction

Forward error correction at the packet level: able to repair up to h lost packets, using h packets of redundancy.



Motivation #1 (ctd)

Different queue management schemes at routers produce different loss patterns.

Assuming the loss rate is the same: is it better

- to have losses regularly spaced,
- or have losses clustered?

Additional motivation

Reliability/real time systems:

- n tasks to be executed within a time frame
  each one may fail
  execute $m = n + k$ of them
  "k-out-of-m"

Bandwith reduction in a slotted network:

- frames of n slots $\rightarrow$ frames of h slots
  no buffer
  probability of overflow?

Pedestrian crossing...

Well known facts...

Several facts are well known:

- Variability worsen things (Folk result)
  $\Longrightarrow$ the situation with the "most regular arrivals" should be better
- The independence assumption is optimistic
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  $\Longrightarrow$ the situation with the "most independent arrivals" should be better

Investigate the issue with a focus on the bursts of losses.

Well known facts...

Several facts are well known:

- Variability worsen things (Folk result)
  $\Longrightarrow$ the situation with the "most regular arrivals" should be better
- The independence assumption is optimistic
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  $\Longrightarrow$ the situation with the "most independent arrivals" should be better

Investigate the issue with a focus on the bursts of losses.

Well known facts...

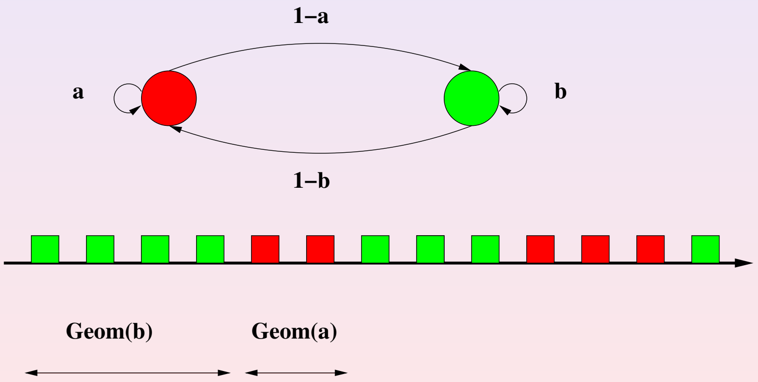
Several facts are well known:

- Variability worsen things (Folk result)
  $\Longrightarrow$ the situation with the "most regular arrivals" should be better
- The independence assumption is optimistic
  If the loss events are independent, the block loss probabilities are (much) smaller than if they are correlated
  $\Longrightarrow$ the situation with the "most independent arrivals" should be better

Investigate the issue with a focus on the bursts of losses.

Mathematical experiment # 1: The Gilbert model

Assumption: losses occur according to the state of a (two-state) Markov chain.



The Gilbert model (2)

Gilbert as a Markov-Additive process:

$$L_{m+1} \;=\; L_m \;+\; \mathbf{1}_{\{X_m = \textcolor{red}{\bullet}\}} \;.$$

$$\begin{aligned} E(\textcolor{blue}{z}^{L_n}) &= \pi_0 \; M(\textcolor{blue}{z})^n \; \mathbf{1} \\ &= (\pi_{\textcolor{red}{\bullet}}, \pi_{\textcolor{green}{\bullet}}) \;\times\; \begin{pmatrix} a\textcolor{blue}{z} & (1-a)\textcolor{blue}{z} \\ 1-b & b \end{pmatrix}^n \;\times\; \begin{pmatrix} 1 \\ 1 \end{pmatrix} \;, \end{aligned}$$

where

$$\pi_{\textcolor{red}{\bullet}} \;=\; \frac{1-b}{2-a-b} \qquad \pi_{\textcolor{green}{\bullet}} \;=\; \frac{1-a}{2-a-b} \;.$$

Gilbert model (3)

Loss Run Length (LRL):

$$\text{LRL} = \frac{1}{1-a}$$

Good Run Length (GRL):

$$\text{GRL} = \frac{1}{1-b}$$

Stationary loss probability:

$$p = \pi_{\bullet} = \frac{\text{LRL}}{\text{LRL} + \text{GRL}} .$$

Problem: with a fixed LRL (or a), the range of p is

$$\left[0, \frac{\text{LRL}}{\text{LRL}+1}\right) .$$

Gilbert model (3)

Loss Run Length (LRL):

$$\text{LRL} = \frac{1}{1-a}$$

Good Run Length (GRL):

$$\text{GRL} = \frac{1}{1-b}$$

Stationary loss probability:

$$p = \pi_{\bullet} = \frac{\text{LRL}}{\text{LRL} + \text{GRL}} .$$

Problem: with a fixed LRL (or a), the range of p is

$$\left[0, \frac{\text{LRL}}{\text{LRL}+1}\right) .$$

Skewed Gilbert Model

Solution: make Good Runs Geometrically distributed on $\{0, 1, \ldots\}$ instead of $\{1, 2, \ldots\}$.

$\Longrightarrow$ another Gilbert process with matrix:

$$\begin{pmatrix} 1 - b(1-a) & b(1-a) \\ 1-b & b \end{pmatrix} ,$$

and

$$\text{LRL} = \frac{1}{1-a} \qquad \text{GRL} = \frac{b}{1-b} \qquad p = \frac{1-b}{1-ab} .$$

Now the range of p is $[0, 1]$!

Skewed Gilbert Model

Solution: make Good Runs Geometrically distributed on $\{0, 1, \ldots\}$ instead of $\{1, 2, \ldots\}$.

$\Longrightarrow$ another Gilbert process with matrix:

$$\begin{pmatrix} 1 - b(1-a) & b(1-a) \\ 1 - b & b \end{pmatrix},$$

and

$$\text{LRL} = \frac{1}{1-a} \qquad \text{GRL} = \frac{b}{1-b} \qquad p = \frac{1-b}{1-ab}.$$

Now the range of p is $[0, 1]$!

Skewed Gilbert Model

Solution: make Good Runs Geometrically distributed on $\{0, 1, \ldots\}$ instead of $\{1, 2, \ldots\}$.

$\Longrightarrow$ another Gilbert process with matrix:

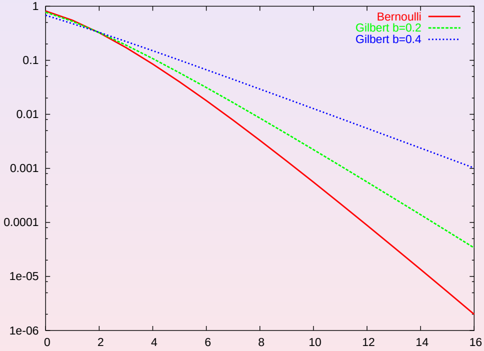
$$\begin{pmatrix} 1 - b(1-a) & b(1-a) \\ 1 - b & b \end{pmatrix} ,$$

and

$$\text{LRL} = \frac{1}{1-a} \qquad \text{GRL} = \frac{b}{1-b} \qquad p = \frac{1-b}{1-ab} .$$

Now the range of p is $[0, 1]$!

Comparison Bernoulli/Gilbert



Loss probability of a block of size $n = h + 16$, depending on h.

Comparison experiments (1)

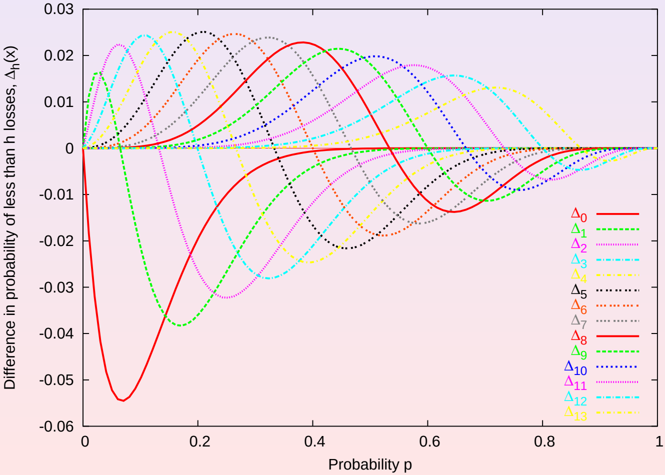
Experiment: consider two cases 1 and 2.

- Fix tho Loss Run lengths: $\text{LRL}_1 < \text{LRL}_2$ ($a_1 < a_2$),
- fix a block length k and a "redundancy" quantity h
- vary the Loss Probability p
- plot the difference:

$$\begin{aligned}\Delta_h(p) &= P(\text{ block saved in case 1 }) \\ &\quad - P(\text{ block saved in case 2 })\end{aligned}$$

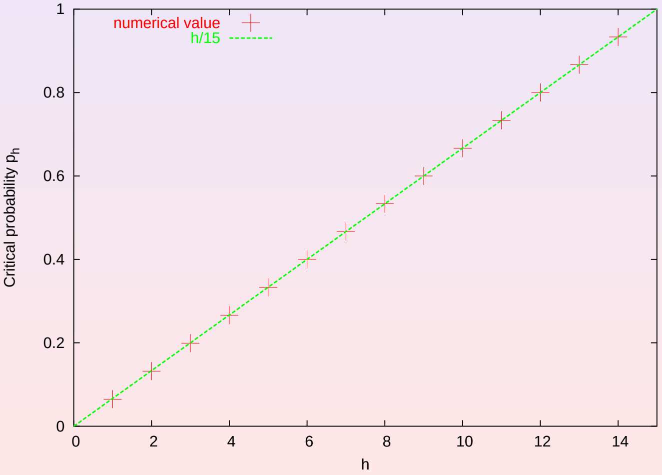
Comparison experiments (2)

h grows from 0 (left, red) to 13 (right, yellow).



Comparison experiments (3)

Let p_h be the value at which $\Delta_h(p_h) = 0$.



Work in progress

Empirical finding: when n is large,

$$x_h \sim \frac{h}{n-1} .$$

How to prove it?
If the loss rate is $p = h/n$,

$$P(\leq h \text{ losses}) = [z^h]\frac{1}{1-z}\begin{pmatrix}(1-c)z & cz \\ 1-b & b\end{pmatrix}^n$$

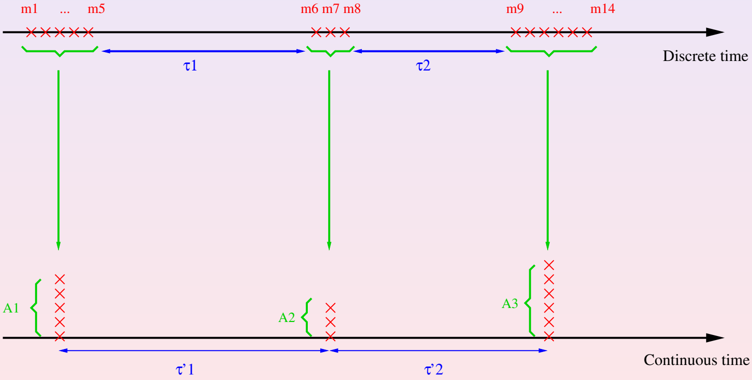
with

$$c = (1-b)\frac{n-h}{n} .$$

Work in progress...

A Compound Poisson Model (1)

Simplification: move to continuous time



A Compound Poisson Model (1)

Process of loss:

- groups of losses occur according to a Poisson process with rate λ,
- groups have random sizes with identical distribution and mean a.

Global loss rate: $p = \lambda \times a$

Distribution of the number of losses:

$$\sum_k z^k P(k \text{ losses in } [0,t)) \; = \; e^{\lambda(A(z)-1)} \; .$$

Comparison experiments (1)

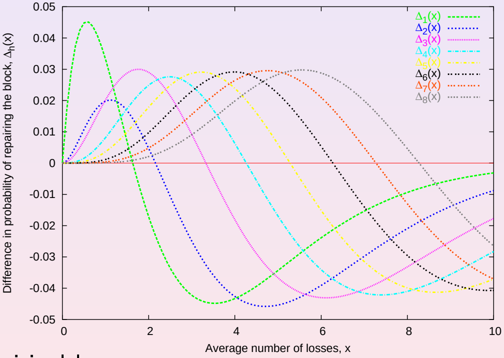
Comparison of two cases:

- Small bursts case: losses of
  1 with proba 0.9,
  2 with proba 0.1
- Larger bursts case: losses of
  1 with proba 0.6,
  2 with proba 0.4
- Same average packet loss number $x = p \times T$

$$\Delta_h(x) = P(\text{ block saved with small bursts}) - P(\text{ block saved with larger bursts})$$

Comparison experiments (2)

Difference $\Delta_h(x)$ as the average number of losses x grows



Again an empirical law

$$x_h \sim h + C .$$

Analysis of limits

Analysis of extreme cases: consider the probability of success of a block

$$P(N_T \leq h) \;=\; \sum_{n=0}^{h} \frac{x^n}{(\mathbb{E}A)^n n!} e^{-xT/\mathbb{E}A} P(A_1 + \ldots + A_n \leq h)\;.$$

i/ Assume that

$$\frac{\mathbb{P}(A^{(1)} > h)}{\mathbb{E}A^{(1)}} \;<\; \frac{\mathbb{P}(A^{(2)} > h)}{\mathbb{E}A^{(2)}}\;.$$

Then $\Delta_h(x) > 0$ when $x \to 0$.

ii/ Assume that $m^{(1)} < m^{(2)}$. Then $\Delta_h(x) < 0$ when $x \to \infty$.

Asymptotic Analysis (1)

Consider the quantity:

$$d_h(y) \;=\; P(\,\leq h \text{ losses in } h+y \text{ time units})\;.$$

Then we find:

$$d_h(y) \;=\; \frac{1}{2} + \frac{1}{\sqrt{2\pi h}}\sqrt{\frac{\mu_1}{\mu_2}}\left(\frac{1}{2} + \frac{\mu_3}{2\mu_2} - y\right) + o(h^{-1/2})\,,$$

where $\mu_1 = EA$, $\mu_2 = E(A^2)$, $\mu_3 = E(A^3)$.

Asymptotic Analysis (1)

Consider the quantity:

$$d_h(y) \;=\; P(\,\le h \text{ losses in } h + y \text{ time units})\;.$$

Then we find:

$$d_h(y) \;=\; \frac{1}{2} \;+\; \frac{1}{\sqrt{2\pi h}}\,\sqrt{\frac{\mu_1}{\mu_2}}\,\left(\frac{1}{2} \;+\; \frac{\mu_3}{2\mu_2} \;-\; y\right) \;+\; o(h^{-1/2})\;,$$

where $\mu_1 = EA$, $\mu_2 = E(A^2)$, $\mu_3 = E(A^3)$.

Asymptotic analysis (2)

Accordingly: for all real y, we have:

$$\Delta_h(h+y) = \frac{1}{\sqrt{2\pi h}}\left(C_0 - C_1 y\right) + o(h^{-1/2}),$$

where:

$$C_0 = \sqrt{\frac{\mu_1^{(1)}}{\mu_2^{(1)}}}\left(\frac{1}{2} + \frac{\mu_3^{(1)}}{6\mu_2^{(1)}}\right) - \sqrt{\frac{\mu_1^{(2)}}{\mu_2^{(2)}}}\left(\frac{1}{2} + \frac{\mu_3^{(2)}}{6\mu_2^{(2)}}\right)$$

$$C_1 = \sqrt{\frac{\mu_1^{(1)}}{\mu_2^{(1)}}} - \sqrt{\frac{\mu_1^{(2)}}{\mu_2^{(2)}}}.$$

Asymptotic analysis (3)

Finally, we have indeed:

$$\Delta_h(h + y_h) \; = \; 0 \quad \Longrightarrow \quad y_h \; \sim \; \frac{C_1}{C_0} \; ,$$

and therefore

$$\boxed{x_h \;\; \sim \;\; \frac{C_1}{C_0} \; + \; h \; .}$$

FEC and the Queue Management

Packet queues inside network routers are handled by a Queue Management scheme.

Two common ones:

Tail Drop Drops packets if and only if the buffer is full
$\Longrightarrow$ tends to produce bursts of losses

RED Drops packets at random preventively
$\Longrightarrow$ tends to produce isolated losses

Two loss paterns: which one works better with FEC?

Application of the model

Alain Jean-Marie, Yvan Calas, Tigist Alemu

Admitting that the smaller bursts (RED) work better when

$$x \le h + C$$

for some constant C.

Equivalently, RED better if:

$$\begin{array}{rcl} \text{small block} & k \le & \frac{1-p}{p} h + \frac{C}{p} \\ \text{large redund. ratio} & \frac{h}{k} \ge & \frac{p}{1-p} - \frac{C}{1-p}\frac{1}{k} \\ \text{small loss rate} & p \le & \frac{h+C}{h+k}. \end{array}$$

Application of the model

Admitting that the smaller bursts (RED) work better when

$$p(k+h) \;\leq\; h \;+\; C$$

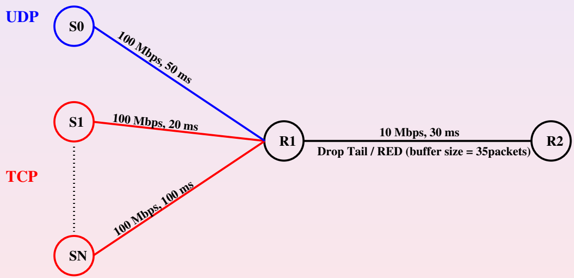
for some constant C.
Equivalently, RED better if:

small block	$k \leq \dfrac{1-p}{p}\, h + \dfrac{C}{p}$
large redund. ratio	$\dfrac{h}{k} \geq \dfrac{p}{1-p} - \dfrac{C}{1-p}\dfrac{1}{k}$
small loss rate	$p \leq \dfrac{h+C}{h+k}$.

Experimental setup

Simulations with the `ns-2` program.

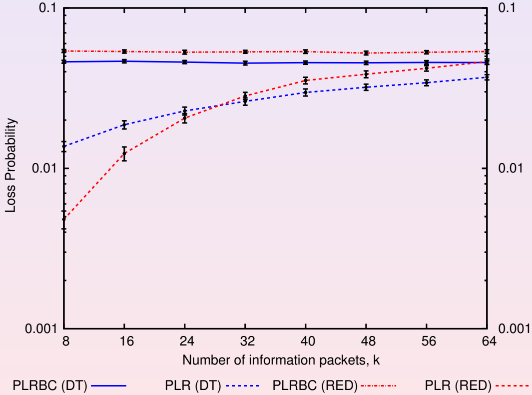
- Source of packets with the UDP protocol, 5-10% of the BW
- Background traffic of TCP flows, saturating the BW.



Statistics collected about Packet Loss Rate Before Correction and after correction.

Results of Simulations

Loss rates, $k = 16$ packets per block $+$ $h = 2$ FEC packets.



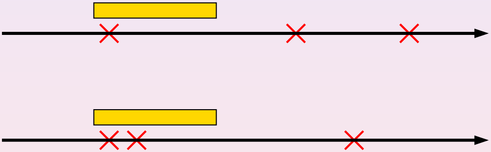
RED does not always win...

An Explanation

There is a compromise between loss “burstiness” and loss rate.
Assume blocks protected with $h = 1$ packet.

An Explanation

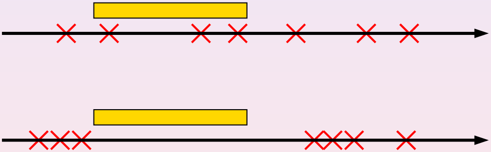
There is a compromise between loss “burstiness” and loss rate.
Assume blocks protected with $h = 1$ packet.



Low loss rate/small blocks

An Explanation

There is a compromise between loss "burstiness" and loss rate.
Assume blocks protected with $h = 1$ packet.



High loss rate/large blocks