Vicente Cavanna
VLSI Architect
19 March, 2001

STORAGE
Networking

iSCSI Digests – CRC or Checksum?

Agilent Technologies

Purpose

STORAGE Networking

- Describe error detection characteristics of CRCs and Checksums.
- Describe types of errors each method is best suited for.
- Describe what type of errors iSCSI should protect from
- Make recommendation

Premise for superiority of CRC

STORAGE Networking

- Commonly held belief of superiority of CRC based on premise that most likely error patterns consists of few erroneous bits.
- Such a bias towards small errors results in weakness in detecting other error patterns.
- There are causes of errors that result in patterns with large number of errors being more likely or even equiprobable.
- If all error patterns are equally probable one in 2^32 will appear to be valid data to both CRC and Checksum.

Probability of various patterns

STORAGE Networking

- When SNR (signal to noise ratio) is the cause of errors, as in a communications link, patterns with few bits in error are more probable.
- When high amplitude noise disturbance is the cause of errors, patterns with large number of errors during the disturbance can be more probable.

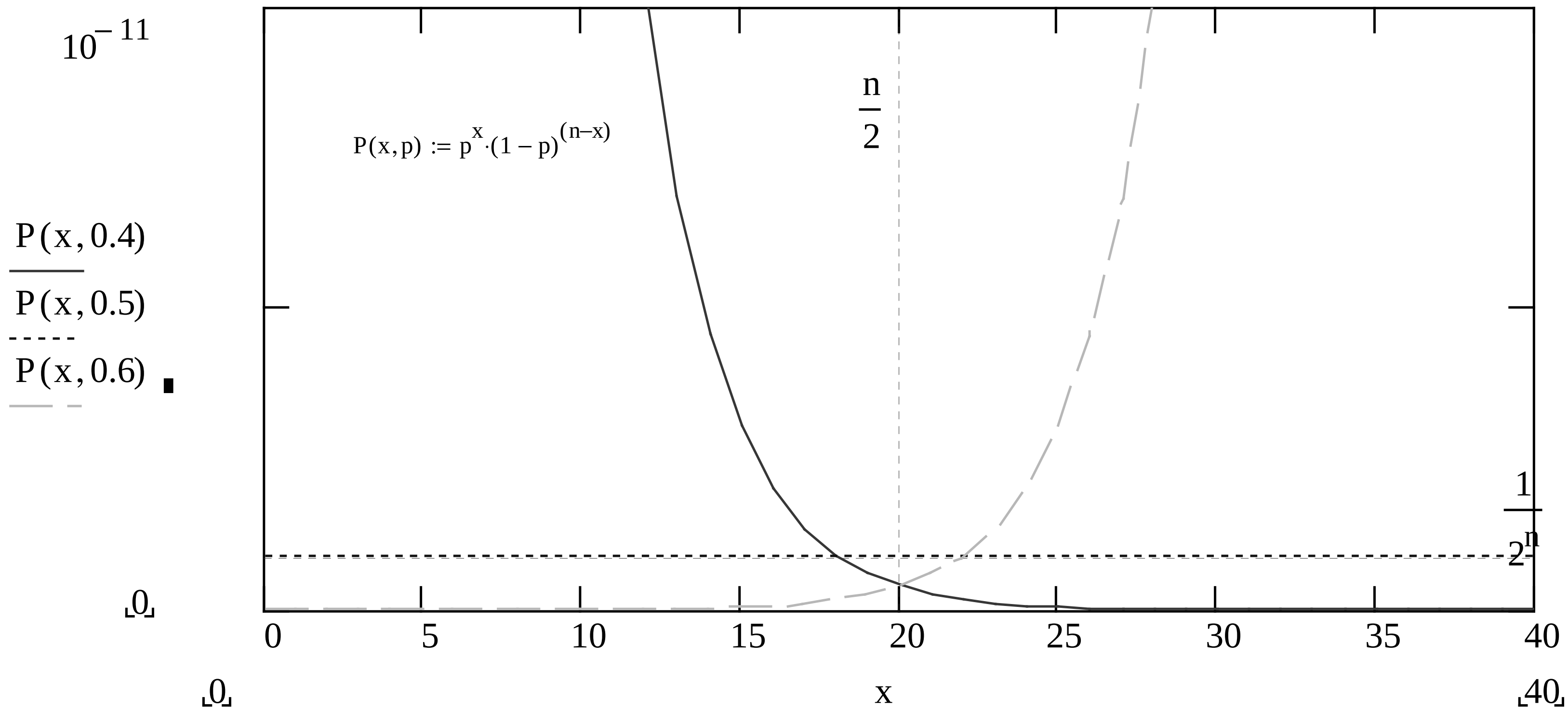
Disturbance can cause large local BER

STORAGE Networking

- When disturbance spanning multiple bits clamps signal to either level.
- And bits are independent
- And individual bits are 1 or 0 with equal probability. BER during the disturbance will be 0.5
- If data is biased, BER can be higher than 0.5

Probability of Any One Given n-bit Pattern with x errors

STORAGE Networking



Probability Distribution of X, the number of errors in n bits.

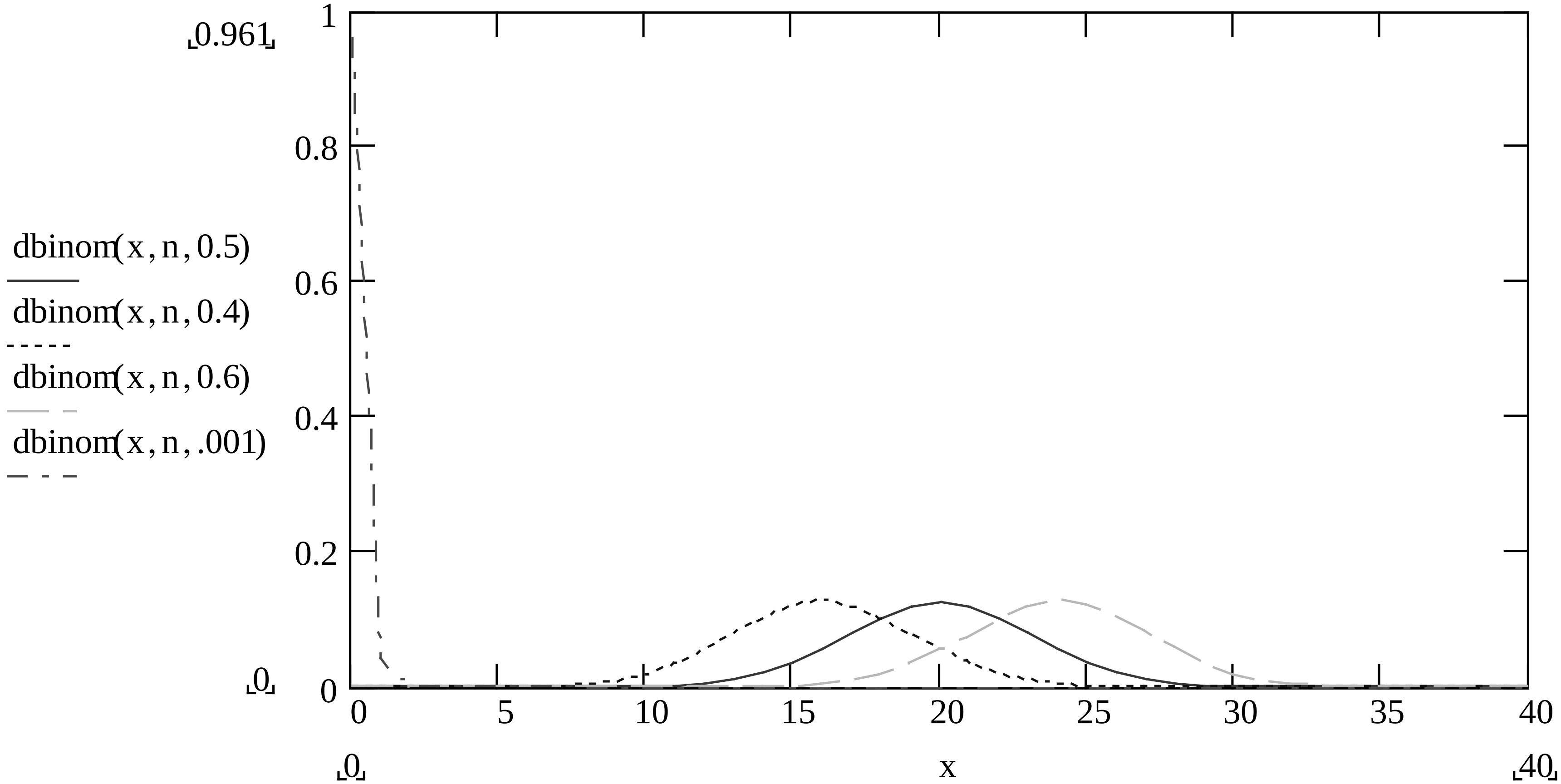
STORAGE Networking

- Probability of x errors in n bits is $p^{x} \cdot (1-p)^{(n-x)}$
- Where p is the probability of error in each trial (i.e. the BER).
- Number of ways in which the errors can be distributed among the n bits is $\frac{n!}{x! \cdot (n-x)!}$
- Probability of X=x is the sum of the probabilities of all sequences containing x errors and n-x non-errors in some order. The Binomial Distribution.

$$P(x) := \left(\frac{n!}{x! \cdot (n-x)!}\right) \cdot \left[p^{x} \cdot (1-p)^{(n-x)} \right]$$

Probability Distribution of the number of errors, X in 40 bits

STORAGE Networking



Which patterns are more probable depends on BER

STORAGE Networking

- If BER < 0.5 patterns with less than half the bits in error are more probable. See plot for BER=0.4
- If BER > 0.5 patterns with more than half the bits in error are more probable. See plot for BER=0.6
- If BER = 0.5 all patterns are equally probable. See plot for BER =0.5.
- If BER <<0.5 distribution centered around very few errors. See plot for BER=0.001

Mechanisms in iSCSI favoring few errors

STORAGE Networking

- Memory – random errors
- Gradual degradation of hardware
- Crosstalk
- Serial links within middle boxes

Mechanisms in iSCSI favoring many errors

STORAGE Networking

- Software/firmware errors
- FIFO control errors
- Bus errors
- Large datapath errors

Performance for few errors

STORAGE Networking

	Number of Single bit errors				Number of Bursts	
	1	2	3	beyond	1	2
CRC32	All, if P(x) has two or more terms[i]	All, if P(x) has a factor with 3 or more terms and the record length is less than the period of P(x)[ii]	All provided P(x) has a factor x^c+1, regardless of record length	All odd number of single-bit errors provided P(x) has a factor x^c+1, regardless of record length[iii]. Most other errors.	All up to 32 bits in size provided P(x) has an x^0 term. Most larger single bursts [iv]	All, provided P(x) has a factor x^c+1 and sum of burst lengths does not exceed c+1 and record length does not exceed period of P(x) [v]
Adler32	All	All, if record length in bytes is less than the modulus, 65521. Otherwise will miss a small subset with just the right spacing	Some errors get through	Similar to CRC except no complete coverage for odd number of bits in error	All, up to 16 bits because data is handled 8 bits at a time	All up to 8 bits if record length in bytes is less than the modulus, provided bursts are byte aligned.
Fletcher32	All	All, if record length in words (2 bytes) is less than the modulo, 65535. Otherwise will miss a small subset with just the right spacing.	Some errors get through	Similar to CRC except no complete coverage for odd number of bits in error	All, up to 32 bits except when a 16 bit pattern of all 0s changes to all 1s or vice-versa[vi]	Most up to 16 bits if record length in words (2 bytes) is less than the modulo, 65535 provided bursts are byte aligned [vii]

Performance for few errors

STORAGE Networking

[i] P(x) is the CRC polynomial and is of degree n

[ii] The period of P(x) is typically 2^n-1 or 2^(n-1)-1

[iii] The x^c+1 factor makes P(x) reducible. 1+x is a commonly used factor since it allows the remaining factor, which is usually primitive, to be of highest order which maximizes the record length covered.

[iv] The only 33 bit wide burst that is undetected is the polynomial itself

[v] A polynomial may detect triple single bit errors and double bursts even if it does not have x^c+1 as a factor but the record length will be smaller. For example the Ethernet polynomial does not have x^c+1 as a factor but will detect all double bursts of size 9 if the record length is less than 13000 bits and all triple bit errors in a record less than 12144 bits long.

[vi] How likely it is to have an error mechanism that converts one word form all zeroes to all ones without touching other words?

[vii] Some double bursts with spacing a multiple of some of the factors of the modulus (3,5,17,257) are missed.

Performance for few errors

STORAGE Networking

- If small number of errors dominate then CRC can easily provide more complete coverage.

Performance for many errors

STORAGE Networking

- For given number of check bits better performance for few errors implies poorer performance when errors are not few.
- Difficult to characterize in detail.

Properties of CCITT-CRC32 for single bit errors

STORAGE Networking

- The poly: $1+x^4+x^{31}+x^{32}$
- Detects all odd number of single errors regardless of record length.
- Detects all double bit errors provided the record length is less than $2^{31}-1$.
- Therefore Hamming distance is at least 4 up to record length of $2^{31}-1$.

Properties of CCITT-CRC32 for error bursts

STORAGE Networking

- Detects all single bursts of size 32.
- Detects all single bursts of size 33 except the for its own pattern.
- Detects 99.9999999767% of all larger single bursts if all patterns are equiprobable.
- Double burst protection not explored.

Cost of Poly complexity

STORAGE Networking

- I compared implementations that divide by the CRC Polynomial taking 32 bits of the input at one time.
- CCITT-CRC32 has 4 terms. Next state equations consist of
  - 5 equations with almost 32 terms; 2 equations with 6 terms; 2 equations with 5 terms the remaining 23 equations with 4 terms.
- CRC32Q has 12 terms. Next-state equations consist of
  - Most of the 32 equations have close to 32 terms.

Recommendation

STORAGE Networking

- CRC due to its better coverage for small number of errors than Checksums.
- CCITT-CRC32 due to its relatively low implementation cost without compromising much in error detection

If time permits ....

STORAGE Networking

- I have detected confusion about the difference between an irreducible or prime polynomial and a primitive polynomial.

Irreducible or Prime Polynomial

STORAGE Networking

- Not divisible by any poly of a degree greater than 0 but less than n.
- Its period is not 2^n-1 unless primitive.
- Has multiple sequences of length equal to the period.
- Has one sequence of length one - the zero sequence.

Primitive Polynomial

STORAGE Networking

- a primitive polynomial is an irreducible polynomial whose period is 2^n-1, i.e. maximum possible length.
- Primitive polynomials are not necessarily most desireable. It depends on what error detection properties are more important.
- We have seen that a factor of 1+x is quite desireable as it provides odd parity.