Scream: a software-efficient stream cipher

Shai Halevi Don Coppersmith Charanjit Jutla

IBM T. J. Watson Research Center
P.O. Box 704, Yorktown Heights, NY 10598, USA

{shaih,copper,csjutla}@watson.ibm.com

February 25, 2002

Abstract

We report on the design of Scream, a new software-efficient stream cipher, which was designed to be a "more secure SEAL". Following SEAL, the design of Scream resembles in many ways a block-cipher design. The new cipher is roughly as fast as SEAL, but we believe that it offers a significantly higher security level. In the process of designing this cipher, we re-visit the SEAL design paradigm, exhibiting some tradeoffs and limitations.

Key words: Stream ciphers, Block ciphers, Round functions, SEAL.

1 Introduction

A stream cipher (or pseudorandom generator) is an algorithm that takes a short random string, and expands it into a much longer string, that still "looks random" to adversaries with limited resources. The short input string is called the seed (or key) of the cipher, and the long output string is called the output stream (or key-stream). Stream ciphers can be used for shared-key encryption, by using the output stream as a one-time-pad. In this work we aim to design a secure stream cipher that has very fast implementations in software.

1.1 A more secure SEAL

The starting point of our work was the SEAL cipher. SEAL was designed in 1992 by Rogaway and Coppersmith [6], specifically for the purpose of obtaining a software efficient stream cipher. Nearly ten years after it was designed, SEAL is still the fastest steam cipher for software implementations on contemporary PC's, with "C" implementations running at 5 cycle/byte on common PC's (and 3.5 cycle/byte on some RISC workstations).

The design of SEAL shares many similarities with the design of common block ciphers. It is built around a repeating *round function*, which provides the "cryptographic strength" of the cipher. Roughly speaking, the main body of SEAL keeps a state which is made of three parts: an *evolving state*, some *round keys*, and a *mask table*. The output stream is generated in steps (or rounds). In each step, the round function is applied to the evolving state, using the round keys. The new evolving state is then masked by some of the entries in the mask table and this value is output as a part of the stream. The mask table is fixed, and some of the round keys are be changed every so often (but not every step).

In terms of security, SEAL is somewhat of a mixed story. SEAL is designed to generate up to 2^{48} bytes of output per seed. In 1997, Handschuh and Gilbert showed, however, that the output stream can be distinguished from random after seeing roughly 2^{34} bytes of output [4]. SEAL was slightly modified after that attack, and the resulting algorithm is known as SEAL 3.0. Recently, Fluhrer described an attack on SEAL 3.0, that can distinguish the output stream from random after about 2^{44} output bytes [3]. Hence, it seems prudent to avoid using the same seed for more than about 2^{40} bytes of output.

The goal of the current work was to come up with a "more secure SEAL". As part of that, we studied the advantages, drawbacks, and tradeoffs of this style of design. More specifically, we tried to understand what makes a "good round function" for a stream cipher, and to what extent a "good round function" for a block cipher is also good as the basis for a stream cipher. We also studied the interaction between the properties of the round function and other parts of the cipher. Our design goals for the cipher were as follows:

- Higher security than SEAL: It should be possible to use the same seed for 2^{64} bytes of output. More precisely, an attacker that sees a total of 2^{64} bytes of output (possibly, using several IV's of its choice), would be forced to spend an infeasible amount of time (or space) in order to distinguish the cipher from a truly random function. A reasonable measure of "infeasibility" is, say, 2^{80} space and 2^{96} time, so we tried to get the security of the cipher comfortably above these values.[1]
- Comparable speed to SEAL, i.e., about 5 cycles per byte on common PC's.
- We want to allow a full 128-bit input nonces (vs. 32-bit nonce in SEAL).
- Other, secondary, goals were to use smaller tables (SEAL uses 4KB of secret tables), get faster initialization (SEAL needs about 200 applications of SHA to initialize the tables), and maybe make the cipher more amenable to implementation in other environments (e.g., hardware, smartcard, etc.) We also tried to make the cipher fast on both 32-bit and 64-bit architectures.

1.2 The end result(s)

In this report we describe three variants of our cipher. The first variant, which we call Scream-0, should perhaps be viewed as a "toy cipher". Although it may be secure enough for some applications, it does not live up to our security goals. In the full version of this report we describe a "low-diffusion attack" that works in time 2^{79} and space 2^{50}, and distinguishes Scream-0 from random after seeing about 2^{44} bytes of the output stream.

We then describe Scream, which is the same as Scream-0, except that it replaces the fixed S-boxes of Scream-0 by key-dependent S-boxes. Scream has very fast software implementations, but to get this speed one has to use secret tables roughly as large as those of SEAL (mainly, in order to store

[1]This security level is arguably lower than, say, AES. This seems to be the price that one has to pay for the increased speed. We note that the "obvious solution" of using Rijndael with less rounds, fails to achieve the desired security/speed tradeoff.

the S-boxes). On our Pentium-III machine, an optimized "C" implementation of Scream runs at 4.9 cycle/byte, slightly faster than SEAL. On a 32-bit PowerPC, the same implementation runs at 3.4 cycle/byte, again slightly faster than SEAL. This optimized implementation of Scream uses about 2.5 KB of secret tables. Scream also offers some space/time tradeoffs. (In principle, one could implement Scream with less than 400 bytes of memory, but using so little space would imply a slowdown of at least two orders of magnitude, compared to the speed-optimized implementation.) In terms of security, if the attacker is limited to only 2^{64} bytes of text, we do not know of any attack that is faster than exhaustively searching for the 128-bit key. On the other hand, we believe that it it possible to devise a linear attack to distinguish Scream from random, with maybe 2^{80} bytes of text.

At the end of this report we describe another variant, called Scream-F (for Fixed S-box), that does not use secret S-boxes, but is slower than Scream (and also somewhat "less elegant"). An optimized "C" implementation of Scream-F runs at 5.6 cycle/byte on our Pentium-III, which is 12% slower than SEAL. On our PowerPC, this implementation runs at 3.8 cycle/byte, 10% slower than SEAL. This implementation of Scream-F uses 560 bytes of secret state. We believe that the security of Scream-F is roughly equivalent to that of Scream.

1.3 Organization

In Section 2 below we first describe Scream-0 and then Scream. In Section 3 we discuss implementation issues and provide some performance measurements. In Section 4 we briefly discuss the cryptanalysis of Scream-0. (A more detailed analysis can be found in the full version.) Finally, in Section 5, we describe the cipher Scream-F. In the appendix we give the constants that are used in Scream, and also provide some "test vectors".

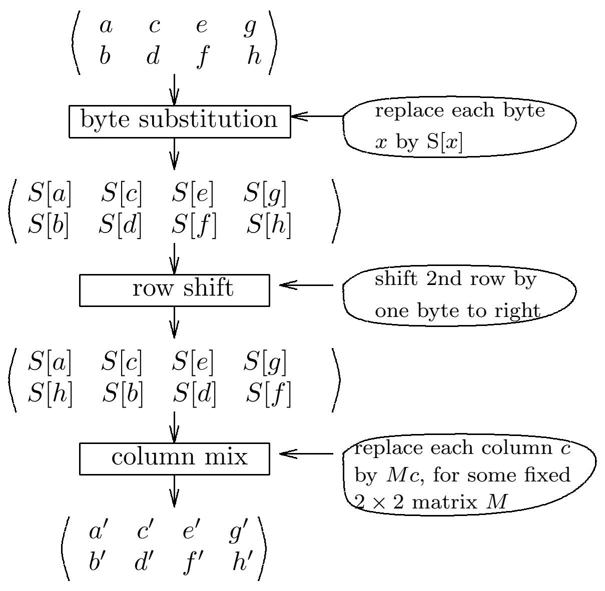
2 The design of Scream

We begin with the description of Scream-0. As with SEAL, this cipher too is built around a "round function" that provides the cryptographic strength. Early in our design, we tried to use an "off the shelf" round function as the basis for the new cipher. Specifically, we considered using the Rijndael round function [2], which forms the basis of the new AES. However, as we discuss in the full paper, the "wide trail strategy" that underlies the design of the Rijndael round function is not a very good match for this type of design. We therefore designed our own round function.

At the heart of our round function is a scaled-down version of the Rijndael function, that operates on 64-bit blocks. The input block is viewed as a 2×4 matrix of bytes. First, each byte is sent through an S-box, $S[\cdot]$, then the second row in the matrix is shifted cyclically by one byte to the right, and finally each column is multiplied by a fixed 2×2 invertible matrix M. Below we call this function the "half round function", and denote it by $G_{S,M}(x)$. A pictorial description of $G_{S,M}$ can be found in Figure 1.

Our round function, denoted $F(x)$, uses two different instances of the "half-round" function, G_{S_1,M_1} and G_{S_2,M_2}, where S_1, S_2 are two different S-boxes, and M_1, M_2 are two different matrices. The S-boxes S_1, S_2 in Scream-0 are derived from the Rijndael S-box, by setting $S_1[x] = S[x]$, and $S_2[x] = S[x \oplus 00010101]$, where $S[\cdot]$ is the Rijndael S-box. The constant 00010101 (decimal 21) was chosen so that S_2 will not have a fixed-point or an inverse fixed-point.[2] The matrices M_1, M_2 were

[2]An inverse fixed-point is some x such that $S[x] = \bar{x}$.



$\begin{pmatrix} a & c & e & g \\ b & d & f & h \end{pmatrix}$
byte substitution
replace each byte x by S[x]
$\begin{pmatrix} S[a] & S[c] & S[e] & S[g] \\ S[b] & S[d] & S[f] & S[h] \end{pmatrix}$
row shift
shift 2nd row by one byte to right
$\begin{pmatrix} S[a] & S[c] & S[e] & S[g] \\ S[h] & S[b] & S[d] & S[f] \end{pmatrix}$
column mix
replace each column c by Mc, for some fixed 2×2 matrix M
$\begin{pmatrix} a' & c' & e' & g' \\ b' & d' & f' & h' \end{pmatrix}$

Figure 1: The "half round" function $G_{S,M}$

chosen so that they are invertible, and so that neither of M_1, M_2 and $M_2^{-1}M_1$ contains any zeros. Specifically, we use

$$M_1 = \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 1 & x+1 \\ x+1 & 1 \end{pmatrix}$$

where $1, x, x+1$ are elements of the field $GF(2^8)$, which is represented as $\mathbb{Z}_2[x]/(x^8+x^7+x^6+x+1)$.

The function F is a mix of a Feistel ladder and an SP-network. A pseudocode of F is provided below, and a pictorial description can be found in Figure 2.

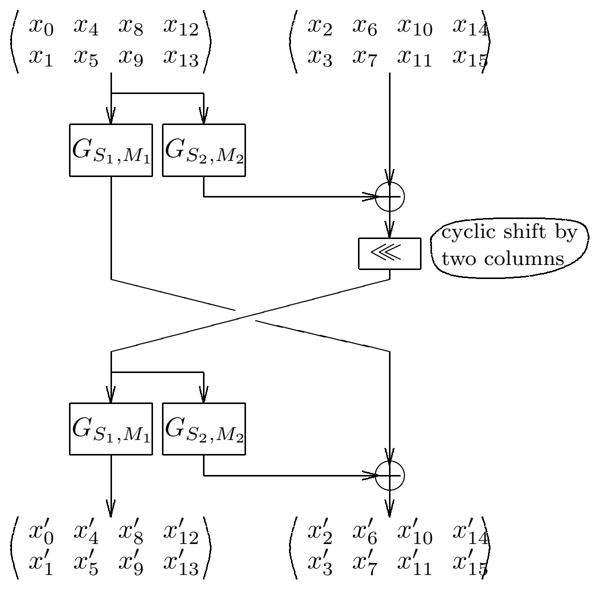
```
Function $F(x)$:
  1. Partition $x$ into two $2 \times 4$ matrices
       $A := \begin{pmatrix} x_0 & x_4 & x_8 & x_{12} \\ x_1 & x_5 & x_9 & x_{13} \end{pmatrix}$    $B := \begin{pmatrix} x_2 & x_6 & x_{10} & x_{14} \\ x_3 & x_7 & x_{11} & x_{15} \end{pmatrix}$

  2. $B := B \oplus G_{S_2,M_2}(A)$                // use $A$ to modify $A, B$
  3. $A := G_{S_1,M_1}(A)$

  4. $B := \begin{pmatrix} B_{0,2} & B_{0,3} & B_{0,0} & B_{0,1} \\ B_{1,2} & B_{1,3} & B_{1,0} & B_{1,1} \end{pmatrix}$    // rotate $B$ by two columns

  5. Swap $A \leftrightarrow B$
  6. $B := B \oplus G_{S_2,M_2}(A)$                // use $A$ to modify $A, B$
  7. $A := G_{S_1,M_1}(A)$

  8. Collect the 16 bytes in $A, B$ back into $x$
       $x' := (A_{0,0}\ A_{1,0}\ B_{0,0}\ B_{1,0}\ A_{0,1}\ A_{1,1}\ B_{0,1}\ B_{1,1}\ A_{0,2}\ A_{1,2}\ B_{0,2}\ B_{1,2}\ A_{0,3}\ A_{1,3}\ B_{0,3}\ B_{1,3})$
```



$\begin{pmatrix} x_0 & x_4 & x_8 & x_{12} \\ x_1 & x_5 & x_9 & x_{13} \end{pmatrix}$
$\begin{pmatrix} x_2 & x_6 & x_{10} & x_{14} \\ x_3 & x_7 & x_{11} & x_{15} \end{pmatrix}$
G_{S_1,M_1}
G_{S_2,M_2}
$\lll$
cyclic shift by two columns
$\begin{pmatrix} x'_0 & x'_4 & x'_8 & x'_{12} \\ x'_1 & x'_5 & x'_9 & x'_{13} \end{pmatrix}$
$\begin{pmatrix} x'_2 & x'_6 & x'_{10} & x'_{14} \\ x'_3 & x'_7 & x'_{11} & x'_{15} \end{pmatrix}$

Figure 2: The round function, F

The main loop of Scream-0. As with SEAL, the cipher Scream-0 maintains a state that consists of the "evolving state" x, some round keys y, z, and a "mask table" W. In Scream-0, x, y and z are 16-byte blocks, and the table W consists of 16 blocks, each of 16 bytes. In step i of Scream-0, the evolving state is modified by setting $x := F(x \oplus y) \oplus z$, and we then output $x \oplus W[i \bmod 16]$.

In Scream-0, both the mask table and the round keys are modified, albeit slowly, throughout the computation. Specifically, after every pass through the mask table (i.e., every 16 steps), we modify y, z and one entry in W, by passing them through the F function. The entries of W are modified in order: after the j'th pass through the table we modify the entry $W[j \bmod 16]$. Moreover, instead of keeping both y, z completely fixed for 16 rounds, we rotate y by a few bytes after each use. The rotation amounts were chosen so that the rotation would be "almost for free" on 32-bit and 64-bit machines. This simple measure provides some protection against "low-diffusion attacks" and linear analysis. A pseudocode of the body of Scream-0 is described in Figure 3.

Key- and nonce-setup. The key- and nonce-setup procedures of Scream-0 are quite straightforward: We just use the round function F to derive all the quantities that we need. The key-setup routine fills the table W with some initial values. These values are later modified during the nonce-setup routine, and they also double as the equivalent of a "key schedule" for the nonce-setup routine. A pseudocode for these two routines is provided in Figures 4 and 5.

2.1 The ciphers Scream

The cipher Scream is the same as Scream-0, except that we derive the S-boxes $S_1[\cdot], S_2[\cdot]$ from the Rijndael S-box $S[\cdot]$ in a key-dependent fashion. We replace line 0a in Figure 4 by the following

0a. set $S_1[x] := S[\ldots S[S[x + \text{seed}_0] + \text{seed}_1] \ldots + \text{seed}_{15}]$ for all x

(Notice that + denotes integer addition mod 256, rather then exclusive-or.) In terms of speed (in software), Scream-S is just as fast as Scream-0, except for the key-setup. However, it has a

```
The main loop of Scream:
State: $x,\ y,\ z$ – three 16-byte blocks
       $W$       – a table of 16 16-byte blocks
       $i_w$     – an index into W (initially $i_w = 0$)
1.  repeat (until you get enough output bytes)
2.      for $i = 0$ to 15                    // generate the next 16 output blocks
3.          $x := F(x \oplus y)$                      // modify the "evolving state" $x$
4.          $x := x \ \oplus \ z$
5.          output $x \oplus W[i \bmod 16]$
6.          if $i = 0$ or 2 mod 4                      // rotate $y$
7.              rotate $y$ by 8 bytes, $y := y_{8..15,0..7}$
8.          else if $i = 1$ mod 4
9.              rotate each half of $y$ by 4 bytes, $y := y_{4..7,0..3,12..15,8..11}$
10.         else if $i < 15$                           // no point in rotating when $i = 15$
11.             rotate each half of $y$ by three bytes to the right, $y := y_{5..7,0..4,13..15,8..12}$
12.         end-if
13.     end-for
14.     $y := F(y \oplus z)$                       // modify $y,\ z$, and $W[i_w]$
15.     $z := F(z \oplus y)$
16.     $W[i_w] := F(W[i_w])$
17.     $i_w := i_w + 1 \bmod 16$
18. end-repeat
```

Figure 3: The main body of Scream and Scream-0

```
Key-setup:
Input:   seed     – a 16-byte block
State:   $a, b$     – temporary variables, each a 16-byte block
Output:  $W0$       – a table of sixteen 16-byte blocks

0a. set $S_1[x] := S[x]$ for all $x$                 // $S[\cdot]$ is the Rijndael S-box
0b. set $S_2[x] := S_1[x \oplus 00010101]$ for all $x$

1.  $a :=$ seed
2.  $b := F(a \oplus pi)$          // $pi$ is a constants: the first 16 bytes in the binary expansion of $\pi$
3.  for $i = 0$ to 15
4.      $a := F^4(a) \oplus b$     // four applications of the function $F$
5.      $W0[i] := a$
6.  end-for
```

Figure 4: The key-setup of Scream-0

Nonce-setup:

Input:	nonce	– a 16-byte block
State:	$W0$	– a table of sixteen 16-byte blocks
	a, b	– temporary variables, each a 16-byte block
Output:	x, y, z	– three 16-byte blocks
	W	– a table of sixteen 16-byte blocks

```
 1. $z := F^2(\text{nonce} \oplus W0[1])$        // two applications of the function $F$
 2. $y := F^2(z \oplus W0[3])$
 3. $a := F^2(y \oplus W0[5])$
 4. $x := F(a \oplus W0[7])$                   // only one application of $F$
 5. $b := x$
 6. for $i = 0$ to 7                       // set $W$ as a modification of $W0$
 7.     $b := F(b \oplus W0[2i])$
 8.     $W[\,2i\,] \quad := W0[\,2i\,] \oplus a$
 9.     $W[2i+1] := W0[2i+1] \oplus b$
10. end-for
```

Figure 5: The nonce-setup of Scream and Scream-0

much larger secret state (a speed-optimized software implementation of Scream-S uses additional 2Kbyte of secret tables). We note that we still have $S_2[x] = S_1[x \oplus 00010101]$, so a space-efficient implementation need only store S_1.

3 Implementation and performance

Software implementation of the F function. A fast software implementation of the F function uses tricks similar to Rijndael: Namely, we can implement the two "half round" functions G_{S_1,M_1}, G_{S_2,M_2} together, using just eight lookup operations into two tables, each consisting of 256 four-byte words. Let the eight-byte input to G_{S_1,M_1}, G_{S_2,M_2} be denoted $(x_0, x_1,\ x_4, x_5,\ x_8, x_9,\ x_{12}, x_{13})$, the output of G_{S_1,M_1} be denoted $(u_0, u_1,\ u_4, u_5,\ u_8, u_9,\ u_{12}, u_{13})$, and the output of G_{S_2,M_2} be denoted $(u_2, u_3,\ u_6, u_7,\ u_{10}, u_{11},\ u_{14}, u_{15})$. Then we can write:

$$\begin{array}{rcl} u_0 &=& M_1(0,0)\cdot S1[x_0] \quad\oplus\quad M_1(0,1)\cdot S1[x_{13}] \\ u_1 &=& M_1(1,0)\cdot S1[x_0] \quad\oplus\quad M_1(1,1)\cdot S1[x_{13}] \\ u_2 &=& M_2(0,0)\cdot S2[x_0] \quad\oplus\quad M_2(0,1)\cdot S2[x_{13}] \\ u_3 &=& M_2(1,0)\cdot S2[x_0] \quad\oplus\quad M_2(1,1)\cdot S2[x_{13}] \end{array}$$

(where $M(i,j)$ is the entry in row i, column j of matrix M, indexing starts from zero). Similar expressions can be written for the other bytes of u. Therefore, if we set the tables T_0, T_1 as

$$T_0(x) = \left\langle M_1(0,0)\cdot S1[x] \;\middle|\; M_1(1,0)\cdot S1[x] \;\middle|\; M_2(0,0)\cdot S2[x] \;\middle|\; M_2(1,0)\cdot S2[x] \right\rangle$$
$$T_1(x) = \left\langle M_1(0,1)\cdot S1[x] \;\middle|\; M_1(1,1)\cdot S1[x] \;\middle|\; M_2(0,1)\cdot S2[x] \;\middle|\; M_2(1,1)\cdot S2[x] \right\rangle$$

Then we can compute $u_{0..3} := T_0[x_0] \oplus T_1[x_{13}]$, $u_{4..7} := T_0[x_4] \oplus T_1[x_1]$, $u_{8..11} := T_0[x_8] \oplus T_1[x_5]$, and $u_{12..15} := T_0[x_{12}] \oplus T_1[x_9]$. A "reasonably optimized" implementation of the round function F (on a 32-bit machine) may work as follows:

```
Function $F(x_0, x_1, x_2, x_3)$:                 // each $x_i$ is a four-byte word
Temporary storage: $u_0, u_1, u_2, u_3$, each a four-byte word
 1. $u_0 := T_0[\text{byte0}(x_0)] \oplus T_1[\text{byte1}(x_3)]$          // first "half round"
 2. $u_1 := T_0[\text{byte0}(x_1)] \oplus T_1[\text{byte1}(x_0)]$
 3. $u_2 := T_0[\text{byte0}(x_2)] \oplus T_1[\text{byte1}(x_1)]$
 4. $u_3 := T_0[\text{byte0}(x_3)] \oplus T_1[\text{byte1}(x_2)]$
 5. $[\text{byte2}(u_0) \mid \text{byte3}(u_0)] := [\text{byte2}(u_0) \mid \text{byte3}(u_0)] \oplus [\text{byte2}(x_0) \mid \text{byte3}(x_0)]$
 6. $[\text{byte2}(u_1) \mid \text{byte3}(u_1)] := [\text{byte2}(u_1) \mid \text{byte3}(u_1)] \oplus [\text{byte2}(x_1) \mid \text{byte3}(x_1)]$
 7. $[\text{byte2}(u_2) \mid \text{byte3}(u_2)] := [\text{byte2}(u_2) \mid \text{byte3}(u_2)] \oplus [\text{byte2}(x_2) \mid \text{byte3}(x_2)]$
 8. $[\text{byte2}(u_3) \mid \text{byte3}(u_3)] := [\text{byte2}(u_3) \mid \text{byte3}(u_3)] \oplus [\text{byte2}(x_3) \mid \text{byte3}(x_3)]$

 9. $u_0 := u_0 \lll 2$ bytes                        // swap the two halves
10. $u_1 := u_1 \lll 2$ bytes
11. $u_2 := u_2 \lll 2$ bytes
12. $u_3 := u_3 \lll 2$ bytes

13. $x_0 := T_0[\text{byte0}(u_2)] \oplus T_1[\text{byte1}(u_2)]$          // second "half round"
14. $x_1 := T_0[\text{byte0}(u_3)] \oplus T_1[\text{byte1}(u_3)]$
15. $x_2 := T_0[\text{byte0}(u_0)] \oplus T_1[\text{byte1}(u_0)]$
16. $x_3 := T_0[\text{byte0}(u_1)] \oplus T_1[\text{byte1}(u_1)]$
17. $[\text{byte2}(x_0) \mid \text{byte3}(x_0)] := [\text{byte2}(x_0) \mid \text{byte3}(x_0)] \oplus [\text{byte2}(u_0) \mid \text{byte3}(u_0)]$
18. $[\text{byte2}(x_1) \mid \text{byte3}(x_1)] := [\text{byte2}(x_1) \mid \text{byte3}(x_1)] \oplus [\text{byte2}(u_1) \mid \text{byte3}(u_1)]$
19. $[\text{byte2}(x_2) \mid \text{byte3}(x_2)] := [\text{byte2}(x_2) \mid \text{byte3}(x_2)] \oplus [\text{byte2}(u_2) \mid \text{byte3}(u_2)]$
20. $[\text{byte2}(x_3) \mid \text{byte3}(x_3)] := [\text{byte2}(x_3) \mid \text{byte3}(x_3)] \oplus [\text{byte2}(u_3) \mid \text{byte3}(u_3)]$

21. output $(x_0, x_1, x_2, x_3)$
```

We note the need for explicit swapping of the two halves above (lines 9-12). The reason for that is that the tables T_0, T_1 are arranged so that the part corresponding to G_{S_1,M_1} is in the first two bytes of each entry, and the part of G_{S_2,M_2} is in the last two bytes. The code above can be optimized further, combining the rotation in these lines with the masking, which is implicit in lines 5-8, 17-20. Hence, the rotation becomes essentially "for free".

This structure provides a space/time tradeoff similar to Rijndael: Since the matrices M_1, M_2 are symmetric, one can obtain $T_2(x)$ from $T_1(x)$ using a few shift operations. Hence, it is possible to store only one table, at the expense of some slowdown in performance. This tradeoff is particularly important for Scream, where the tables T_0, T_1 are key-dependent.

The nonce-setup routine. The nonce-setup routine was designed so that the first output block can be computed as soon as possible. Although all the entries of the table W have to be modified during the nonce-setup, an application that does not use all of them can modify only as many as it needs. Hence an application that only outputs a few blocks per input nonce, does not have to complete the entire nonce-setup. Alternatively, an application can execute the nonce-setup together with the first "chunk" of 16 steps, modifying each mask of W just before this mask is needed.

Performance in software. We tested the software performance of Scream and Scream-F on two platforms, both with word-length of 32 bits: One platform is an IBM PC 300PL, with a 550MHz Pentium-III processor, running Linux and using the gcc compiler, version 3.0.3. The other platform is an RS/6000 43P-150 workstation, with a 375MHz 304e PowerPC processor, running AIX 4.3.3 and using the IBM C compiler (xlc) version 3.6.6. On both platforms, we measured peak throughput, and also timed the key-setup and nonce-setup routines. To measure peak throughput, we timed a procedure that produces 256MB of output (all with the same key and nonce). Specifically, the procedure makes one million calls to a function that outputs the next 256 bytes of the cipher. To eliminate the effect of cache misses, we used the same output buffer in all the calls. We list our test results in the table below.

Platform	Operation	Scream-F	Scream	SEAL
Pentium-III	throughput	5.6 cycle/byte	4.9 cycle/byte	5.0 cycle/byte
550 MHz	key-setup	3190 cycles	27500 cycles	
Linux, gcc	nonce-setup	1276 cycles	1276 cycles	
604e PowerPC	throughput	3.8 cycle/byte	3.4 cycle/byte	3.45 cycle/byte
375 MHz	key-setup	1950 cycles	16875 cycles	
AIX, xlc	nonce-setup	670 cycles	670 cycles	

Implementation in different environments. Being based on a Rijndael-like round function, Scream is amenable for implementations in many different environments. In particular, it should be quite easy to implement it in hardware, and the area/speed tradeoff in such implementation may be similar to Rijndael (except that Scream needs more memory for the mask table). Also, it should be quite straightforward to implement it for 8- and 16-bit processors (again, as long as the architecture has enough memory to store the internal state). Scream is clearly not suited for environments with extremely small memory, but it can be implemented with less than 400 bytes of memory (although such implementation would be quite slow).

4 Security Analysis

Below we examine some possible attacks on Scream-0 and Scream. The discussion below deals mostly with Scream-0. At the end we briefly discuss the effect of Scream's key-dependent S-boxes on these attacks. We examine two types of attacks, one based on linear approximations of the F function, and the other exploits the low diffusion provided by a single application of F. In both attacks, the goal of the attacker is to distinguish the output of the cipher from a truly random stream.[3]

4.1 Linear attacks

It is not hard to see that the F function has linear approximations that approximate only three of the 8-by-8 S-boxes. Since the S-boxes in Scream-0 are based on the Rijndael S-box, the best approximation of them has bias 2^{-3}, so we can probably get a linear approximation of the F function with bias 2^{-9}. Namely, there exists a linear function L such that $\Pr_x[L(x, F(x)) = 0] = 1/2 \pm 2^{-9}$.

[3]In a separate paper [1], we show that these two types of attacks can be viewed as two special cases of a generalized distinguishing attack.

In Appendix A, we show that there are no approximation of the F function with bias of more than 2^{-9}.

To use such approximation, we need to eliminate the linear masking, introduced by the y, z and the $W[i]$'s. Here we use the fact that each one of these masks is used sixteen times before it is modified. For each step of the cipher, the attacker sees a pair $(x \oplus y \oplus W[i],\ F(x) \oplus z \oplus W[i+1])$, where x is random. Applying the linear approximation L to this pair, we get the bit

$$\sigma \ = \ L(x, F(x)) \oplus L(y, z) \oplus L(W[i], W[i+1])$$

For simplicity, we ignore for the moment the rotation of the y block after each step. If we add two such σ's that use the same y and z blocks, we get $\tau = \sigma \oplus \sigma' = L(x, F(x)) \oplus L(x', F(x')) \oplus L(W[i], W[i+1]) \oplus L(W[j], W[j+1])$. The last bit does not depend on y, z anymore. We can repeat this process, adding two such τ's that use the same masks, we end up with a bit

$$\mu = \tau \oplus \tau' = L(x, F(x)) \oplus L(x', F(x')) \oplus L(x'', F(x'')) \oplus L(x''', F(x'''))$$

Since $L(x, F(x))$ has bias of 2^{-9}, the bit μ has bias of 2^{-36}, so after seeing about 2^{72} such bits, we can distinguish the cipher from random.

Since each of the masks is used sixteen times before it is modified, we have about $\binom{16}{2}$ choices for the pairs of σ's to add (still ignoring the rotation of y), and about $\binom{16}{2}$ choices for the pairs of τ's to add. Hence, 256 steps of the cipher gives us about $\binom{16}{2}^2 \approx 2^{14}$ bits μ. After seeing roughly $256 \cdot 2^{58} = 2^{66}$ steps of the cipher (i.e., 2^{70} bytes of output), we can to collect the needed 2^{72} samples of μ's to distinguish the cipher from random.

The rotation of y. The rotation of y makes it harder to devise attacks as above. To cancel both the y and the z blocks, one would have to use two different approximations with the same output bit pattern, but where the input bit patterns are rotated accordingly. We do not know if it possible to devise such approximation with "reasonably high" bias.

The secret S-boxes. The introduction of key-dependent S-boxes in Scream does not significantly alter the analysis from above. Since the S-boxes are key-dependent, an attacker cannot pick "the best approximations" for them, but on the other hand these S-boxes have better approximations than the Rijndael S-box. Thus, the attacker can use a random approximation, and it will likely to be roughly as good as the best approximation for the fixed S-boxes.

The nonce-setup procedure. The analysis from above assumed that the attacker only uses one nonce, and watches many output bytes from the resulting stream. In our attack model, however, the attacker is able to feed the cipher with many different nonces. To see why this may be an effective attack, consider what would happen if we eliminate the mask modification process (Lines 5–10) from the nonce-setup procedure. The attacker could then feed many different nonces, watching only the first few output blocks from each nonce. In this process, the masks are fixed, and therefore there is no need to cancel them out. The only thing that needs to be canceled out are the y, z blocks, and the attacker can do that by approximating only two steps for each nonce. This could potentially yield an approximation with bias as high as 2^{-18}, so the attacker only needs about 2^{36} different nonces before it can distinguish the cipher from random. [4]

[4] As noted above, the achievable bias is likely to be smaller, due to the rotations of the y block.

The simplest fix is to modify all the masks (in an “uncorrelated” way) during the nonce-setup. However, doing that is rather expensive. We therefore used an “optimization trick”, where we modified the odd entries by adding to them different values, and modified all the even entries by adding to them the same value. The reason that this helps, is that an approximation of a single step includes two masks, one even and one odd. Thus, we still need to cancel out the odd masks, which means that we still need to add at least four approximations.

The only way to avoid using masks from odd steps, is to use an approximation of two consecutive F functions, and this is likely to have small bias. Moreover, to be able to cancel the value that was added to the even masks, and also the y and z blocks, and to do it using just two steps, one must use approximations of the F function, where

(a) the same bitwise pattern is used on both the input and the outputs of the function; and

(b) this bitwise pattern is periodic.

We were not able to find such an approximation that uses less than all the S-boxes.

4.2 Low-diffusion attacks

A low-diffusion attack exploits the fact that not every byte of $F(x)$ is influenced by every byte of x (and vise versa). For example, there are output bytes that only depend on six input bytes. In fact, in Appendix B we show that knowing two bytes of x and one byte of (linear combination of bytes in) $F(x)$, we can compute another byte of (linear combination of bytes in) $F(x)$. Namely, we have a (non-degenerate) linear function L with output length of four bytes, so that we can write $L(X, F(x))_3 = g(L(X, F(x))_{0..2})$, where g is an known deterministic function (with three bytes of input and one byte of output).

As for the linear attacks, here too we need to eliminate the linear masking, introduced by the y, z and the $W[i]$'s. This is done in very much the same way. Again, we ignore for now the rotation of the block y. For each step of the cipher the attacker sees the four bytes $L(x \oplus y \oplus W[i], F(x) \oplus z \oplus W[i+1])$. We eliminate the dependence on y, z by adding two such quantities that use the same y, z blocks. This gives a four-byte quantity $L(x, F(x)) \oplus L(x', F(x')) \oplus L(W[i], W[i+1]) \oplus L(W[j], W[j+1])$. Adding two of those with the same i, j, we then obtain the four byte quantity

$$L(x, F(x)) \oplus L(x', F(x')) \oplus L(x'', F(x'')) \oplus L(x''', F(x'''))$$

We can write this last quantity in terms of the function g, as a pair $(s \oplus t \oplus u \oplus v,\ g(s) \oplus g(t) \oplus g(u) \oplus g(v)\ )$, where each of s, t, u, v is three-byte long, and the $g(\star)$'s are one-byte long. In a separate paper [1], we analyze the statistical properties of such expressions, and calculate the number of samples that needs to be seen to distinguish them from random.

The rotation of y. Again, the rotation of y makes it harder to devise attacks as above. In Appendix B we show, however, that we can still use a low-diffusion attack on the F function, in which guessing six bytes of $(x, F(x))$ yields the value of four other bytes. Applying tools from our paper [1] to this relation, we can compute that the amount of output text that is needed to distinguish the cipher from random along the lines above, is merely 2^{43} bytes. However, the procedure for distinguishing is quite expensive. The most efficient way that we know how to use these 2^{43} bytes would require roughly 2^{50} space and 2^{80} time.

The secret S-boxes. At present, we do not know how to extend low-diffusion attacks such as above to deal with secret S-boxes. Although we can still write the same expression $L(X, F(x))_3 = g(L(X, F(x))_{0..2})$, the function g now depends on the key, so it is not known to the attacker. Although it is likely that some variant of these attacks can be devised for this case too, we strongly believe that such variants would require significantly more text than the 2^{64} bytes that we "allow" the attacker to see.

5 The cipher Scream-F

In Scream, we used key-dependent S-boxes to defend against "low-diffusion attacks". A different approach is to keep the S-box fixed, but to add to the main body of the cipher some "key dependent operation" before outputting each block. This approach was taken in Scream-F, where we added one round of Feistel ladder after the round function, using a key-dependent table. However, since the only key-dependent table that we have is the mask table W, we let W double also as an "S-box". Specifically, we add the following lines 3a-3e between lines 3 and 4 in the main-loop routine from Figure 3.

```
3a.    view the table $W$ as an array of 64 4-byte words $\hat{W}[0..63]$
3b.    $x_{0..3} := x_{0..3} \oplus \hat{W}[1 + (x_4 \wedge 00111110)]$
3c.    $x_{4..7} := x_{4..7} \oplus \hat{W}[x_8 \wedge 00111110]$
3d.    $x_{8..11} := x_{8..11} \oplus \hat{W}[1 + (x_{12} \wedge 00111110)]$
3e.    $x_{12..15} := x_{12..15} \oplus \hat{W}[x_0 \wedge 00111110]$
```

We note that the operation $x_i \wedge 00111110$ in these lines returns an even number between 0 and 62, so we only use odd entries of W to modify $x_{0..3}$ and $x_{8..11}$, and even entries to modify $x_{4..7}$ and $x_{12..15}$. The reason is that to form the output block, the words $x_{0..3}, x_{8..11}$ will be masked with even entries of W, and the words $x_{4..7}, x_{12..15}$ will be masked by odd entries. The odd/even indexing is meant to avoid the possibility that these masks cancel with the entries that were used in the Feistel operation.[5]

5.1 Conclusions

We presented Scream, a new stream cipher with the same design style as SEAL. The new cipher is roughly as fast as SEAL, but we believe that it is more secure. It has some practical advantages over SEAL, in flexibility of implementation, and also in the fact that it can take a full 128-bit nonce (vs. 32 bits in SEAL). In the process of designing Scream, we studied the advantages and pitfalls of the SEAL design style. We hope that the experience from this work would be beneficial also for future ciphers that uses this style of design.

Acknowledgments

This design grew out of a study group in IBM, T.J. Watson during the summer and fall of 2000. Other than the authors, the study group also included Ran Canetti, Rosario Gennaro,

[5]It is still possible that two words, say $x_{0..3}$ and $x_{4..7}$, are masked with the same mask, but it seems less harmful.

Nick Howgrave-Graham, Tal Rabin and J.R. Rao. The motivation for this work was partly due to the NESSIE "call for cryptographic primitives" (although we missed their deadline by more than one year).

References

[1] D. Coppersmith, S. Halevi, and C. Jutla. Cryptanalysis of stream ciphers with linear masking. Available from the ePrint archive, at `http://eprint.iacr.org/2002/020/`, 2002.

[2] J. Daemen and V. Rijmen. AES proposal: Rijndael. Available on-line from NIST at `http://csrc.nist.gov/encryption/aes/rijndael/`, 1998.

[3] S. Fluhrer. Cryptanalysis of the SEAL 3.0 pseudorandom function family. In *Proceedings of the Fast Software Encryption Workshop (FSE'01)*, 2001.

[4] H. Handschuh and H. Gilbert. χ^2 cryptanalysis of the SEAL encryption algorithm. In *Proceedings of the 4th Workshop on Fast Software Encryption*, volume 1267 of *Lecture Notes in Computer Science*, pages 1–12. Springer-Verlag, 1997.

[5] K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptography, Eurocrypt'93*, volume 765 of *Lecture Notes in Computer Science*, pages 55–64. Springer-Verlag, 1993.

[6] P. Rogaway and D. Coppersmith. A software optimized encryption algorithm. *Journal of Cryptology*, 11(4):273–287, 1998.

A Linear approximations of the F function of Scream-0

A detailed description of the round function F is provided again in Figure 6. Throughout the next two sections, we refer to the notation that are used in that figure.

A.1 The S-boxes

Since every S-box lookup in F uses the same input for the two S-boxes S_1, S_2, one should view these two S-boxes as one box, with one byte input and two byte output. As both S_1, S_2 are permutations, any approximation of this 8×16 S-box (with non-zero bias) must look at at least two of the three bytes $(x, S_1[x], S_2[x])$. Below we say that an approximation is a *two-value approximation* if it uses only two of these values, and it is a *three-value approximation* if it uses all three.

Proposition 1 *For the S-boxes S_1, S_2 of Scream-0:*
(a) Any two-value approximation of $x, S_1[x], S_2[x]$ has bias of at most 2^{-3}.
(b) Any three-value approximation of $x, S_1[x], S_2[x]$ has bias of at most 2^{-2}.

Proof: These facts can be confirmed by an exhaustive search. Below we provide an analytical proof for (a): Recall that the S-boxes S_1, S_2 in Scream-0 are defined as $S_1[x] = A(\frac{1}{x}) \oplus b$, $S_2[x] = A(\frac{1}{x\oplus\delta}) \oplus b$, where A is an invertible boolean matrix, the $(1/x)$ operation is defined in $GF(2^8)$,

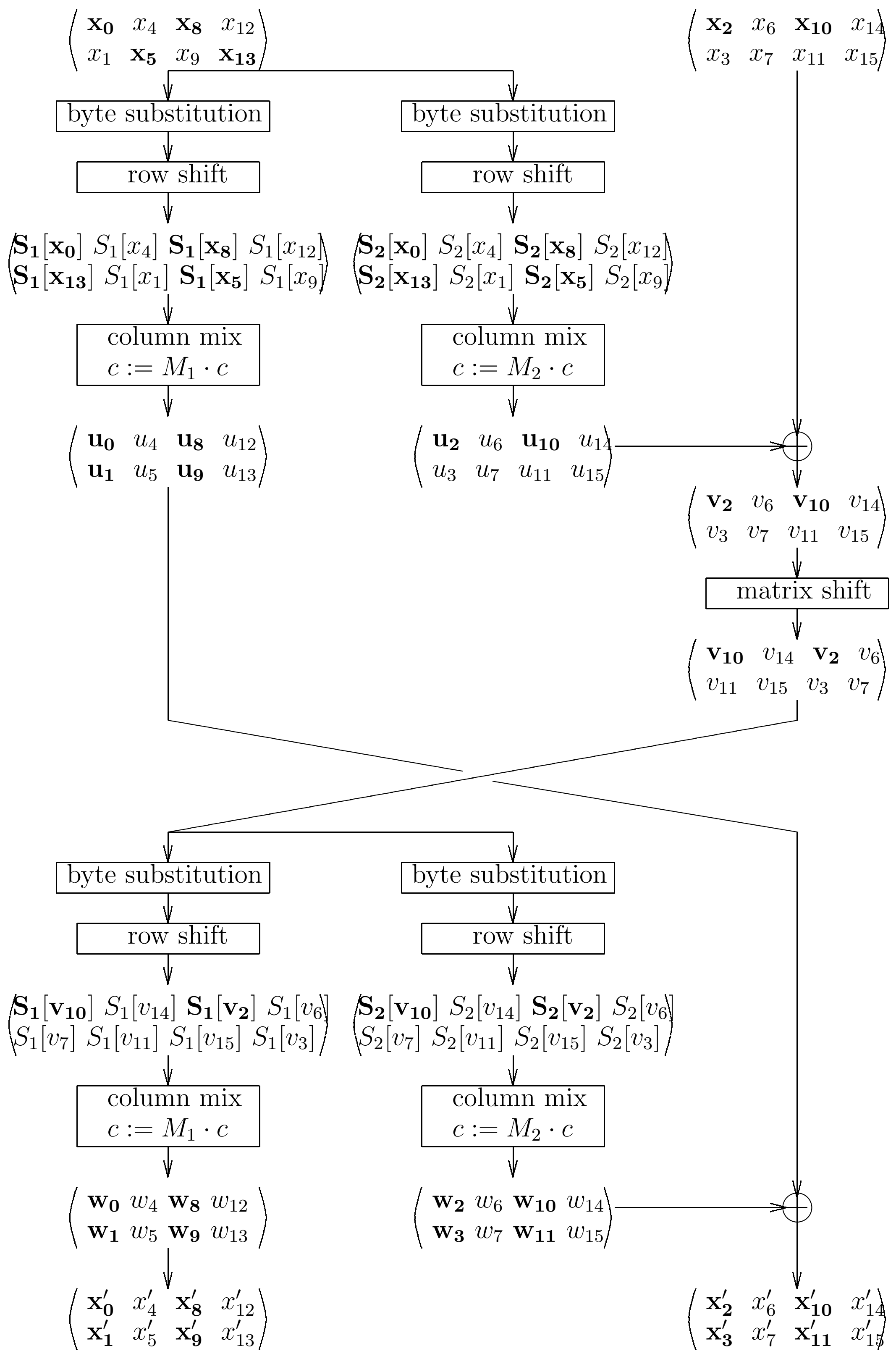


Figure 6: Details of the round function F. (The bytes in boldface are used in the low-diffusion attack from Appendix B.)

and b, δ are two bytes, with $\delta \neq 0$. As the operation $Az \oplus b$ is just an affine transformation, it is sufficient to prove the assertion for the "underlying S-boxes"

$$S'_1[x] = \frac{1}{x} \qquad\qquad S'_2[x] = \frac{1}{x \oplus \delta}$$

It is known that the best linear approximation of the function $f(x) = 1/x$ in $GF(2^n)$ has bias $2^{1-n/2}$ [5]. Hence, every approximation that only uses $(x, S'_1[x])$ (or only uses $(x, S'_2[x])$) must have bias at most 2^{-3}. It is left to show that an approximation that only uses the values $(S'_1[x], S'_2[x])$ cannot do any better. For any two masks ρ_1, ρ_2 and any bit σ, denote

$$X_{\rho_1,\rho_2,\sigma} = \{x \ : \ \rho_1 \cdot (1/x) \oplus \rho_2 \cdot (1/(x \oplus \delta)) = \sigma\}$$

We need to show that for any ρ_1, ρ_2, σ, we have $|X_{\rho_1,\rho_2,\sigma}| \leq \frac{256}{2}(1 + 2^{-3}) = 144$.

We now change the variable, setting $y = 1 \oplus (\delta/x)$ (all the operations in $GF(2^8)$). With this assignment, we have $1/x = (1 \oplus y)/\delta$ and $1/(x \oplus \delta) = (1 \oplus (1/y))/\delta$. The transformation from x to y is a bijection, so we have

$$|X_{\rho_1,\rho_2,\sigma}| \ = \ \left| \left\{ y \ : \ \rho_1 \cdot \left(\frac{1 \oplus y}{\delta} \right) \ \oplus \ \rho_2 \cdot \left(\frac{1 \oplus (1/y)}{\delta} \right) = \sigma \right\} \right|$$

Finally, since δ is fixed, then division by δ in $GF(2^8)$ is a linear operation, and therefore it commutes with the inner-product operation. Hence, we have

$$|X_{\rho_1,\rho_2,\sigma}| = \left| \{y \ : \ (\rho_1/\delta) \cdot y \ \oplus \ (\rho_2/\delta) \cdot (1/y) = \sigma \oplus ((\rho_1/\delta) \cdot 1) \oplus ((\rho_2/\delta) \cdot 1)\} \right| \leq \ 144$$

where the last inequality follows since this is the bias of some linear approximation of the function $f(y) = 1/y$, and we already know that that function cannot be approximated with bias of more than 2^{-3}. $\square$

A.2 The F function

It is easy to see that the F function has approximations that only look at three S-boxes. For example (using the notations from Figure 6), one can use the 2-value approximations $(x_0, S_2[x_0])$, $(x_{13}, S_2[x_{13}])$ and $(v_2, S_1[v_2])$. Since there is a linear relation between $S_2[x_0], S_2[x_{13}], x_2$ and v_2, and another linear relation between $S_1[v_2], x'_8, x'_9$, we obtain an approximation of input bytes x_0, x_{13}, x_2 and output bytes x'_8, x'_9. Since the approximations of the S-boxes have bias at most 2^{-3}, this approximation of F has bias at most 2^{-9}. This, however, is the best possible approximation of the F function, as we prove below.

Proposition 2 *Any approximation of the F function has bias of at most* 2^{-9}.

Proof: First some terminology: As usual, we view an approximation of the function F as a sum (modulo 2) of approximations involving the intermediate components. Since the only non-linear components are the S-boxes, we identify an approximation of F with an assignment of masks to the inputs and outputs of *all* the S-boxes in Figure 6. For example, we may talk about "the mask of $S_1[x_0]$" in some approximation. We say that an approximation *uses* some specific byte, if it assigns a non-zero mask to that byte, and it uses some S-box if it uses some of its input and outputs.

We now present some useful observations. Consider, for example, the mask of the bytes $S_2[x_0]$, $S_2[x_{13}]$, $S_2[v_{10}]$ and $S_2[v_7]$. Note that these bytes (and the output bytes x'_2, x'_3) are related by

$$\begin{pmatrix} S_2[v_{10}] \\ S_2[v_7] \end{pmatrix} = M_2^{-1} \begin{pmatrix} x'_2 \\ x'_3 \end{pmatrix} \oplus M_2^{-1} M_1 \begin{pmatrix} S_1[x_0] \\ S_1[x_{13}] \end{pmatrix}$$

Therefore, in any approximation (with non-zero bias), the masks of these bytes must also be related. Namely, if we have an approximation with the masks $\rho_1, \rho_2, \rho_3, \rho_4$ assigned respectively to $S_2[x_0]$, $S_2[x_{13}]$,$S_2[v_{10}]$,$S_2[v_7]$, then it must be the case that

$$\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = M_2^{-1} M_1 \begin{pmatrix} \rho_3 \\ \rho_4 \end{pmatrix} \tag{1}$$

In particular, the values of ρ_1, ρ_2 determine the values of ρ_3, ρ_4, and vice versa. Moreover, since $M_2^{-1} M_1$ is an MDS matrix, then if one of these masks is non-zero, at least three of them must be non-zero. Similarly, the masks assigned to the bytes $S_2[x_0], S_2[x_{13}], v_2$ and v_3 are related by

$$\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = M_2 \begin{pmatrix} \rho_3 \\ \rho_4 \end{pmatrix} \tag{2}$$

Again, since M_2 is MDS, then if one of these masks is non-zero, at least three of them must be non-zero. In Table 1 we list eight sets of bytes for which the masks must be related in the same manner. The analysis above says that for any column in Table 1, either all the bytes are assigned the mask zero, or at least three are assigned a non-zero mask. The rest of the proof follows by some observations on the structure of that table.

Bytes related by Eq. (1)

$S_1[x_0]$	$S_1[x_4]$	$S_1[x_7]$	$S_1[x_{12}]$
$S_1[x_{13}]$	$S_1[x_1]$	$S_1[x_5]$	$S_1[x_9]$
$S_2[v_{10}]$	$S_2[v_{14}]$	$S_2[v_2]$	$S_2[v_6]$
$S_2[v_7]$	$S_2[v_{11}]$	$S_2[v_{15}]$	$S_2[v_3]$

Bytes related by Eq. (2)

$S_2[x_0]$	$S_2[x_4]$	$S_2[x_7]$	$S_2[x_{12}]$
$S_2[x_{13}]$	$S_2[x_1]$	$S_2[x_5]$	$S_2[x_9]$
v_2	v_6	v_{10}	v_{14}
v_3	v_7	v_{11}	v_{15}

Table 1: Some intermediate bytes of the F function

Looking at Table 1, we see that it contains two of the three bytes from every S-box in the F function, and that these two bytes never appear in the same column. Therefore, every two-value approximation of an S-box in F must assign non-zero mask to at least one byte in Table 1, and every three-value approximation of an S-box must assign non-zero mask to at least two byte in two different columns.

Below we say that two columns in Table 1 *intersect*, if they contain bytes from the same S-box. We say that the columns *intersect at the top* if the upper halves of these columns intersect, and that they *intersect at the bottom* if the lower halves intersect. Due to the "matrix shift" operation in the F function, there are no two columns in Table 1 that intersect both at the top and at the bottom. That is, if the top halves of two columns intersect, then their bottom halves do not, and vice versa.

We now can prove Proposition 2. First, every approximation must assign non-zero value to at least one byte in some column the table. But this implies that at least three bytes in that column has to be assigned non-zero values. Hence, every approximation must use bytes from at least three different S-boxes.

Next, if an approximation of F uses any three-value approximation of an S-box, then it must assign non-zero values to bytes in at least two different columns of Table 1. Since in each of these columns there must be at least three bytes with non-zero masks (and since these columns cannot intersect both at the top and at the bottom), it follows that the approximation uses bytes from at least four different S-boxes.

Finally, if an approximation of F uses *only* three-value approximations of S-boxes, then it must assign non-zero values to bytes in more than two columns, and therefore use bytes from more than four different S-boxes (in fact, at least six S-boxes). Using the bounds from Proposition 1 on the bias of S-box approximations, we conclude that in either case, an approximation of the F function cannot have bias of more than 2^{-9}. □

B Low-diffusion attack on the F function

Low-diffusion attacks exploit the limited diffusion provided by just one application of the F function. The goal is to guess just a few of the input/output bytes of the function, and get at least one byte of consistency check. It is not hard to see that we can get some relations by guessing only four (linear combinations of the) input and output bytes. For example, it can be seen from Figure 6, that there is a relation involving input bytes x_0, x_{13}, a one-byte linear combination of x'_0, x'_1 (namely $S_1[v_{10}]$), and a one-byte linear combination of x'_2, x'_3: If we guess $S_1[v_{10}]$, then we also know $S_2[v_{10}]$, which is some linear combination of $w_2 = u_0 \oplus x'_2$ and $w_3 = u_1 \oplus x'_3$. If we also guess x_0 and x_{13}, we can compute u_0 and u_1, thereby deducing the linear combination of x'_2 and x'_3. [6] It is not hard to see that guessing any three of these bytes, we can compute the fourth byte. Another example is a relation between $S_1[v_{10}]$ and the three input bytes x_5, x_8, x_{10}.

These "guess three, get one free" relations, however, cannot be used directly to mount an attack, because of the rotations of the y block. To mount an attack, we need to cancel out both the rotated y blocks, and the fixed z block (and also the fixed masks $W[i]$). This means that we must find two relations that use the same (linear combinations of) output bytes, but rotated version of the (linear combinations of) input bytes.

The best strategy that we found for doing that, is to use two copies of the "guess three, get one free" relations from above. For example, the attacker can guess the six bytes x_0, x_{13}, $S_1[v_{10}]$ and x_8, x_5, $S_1[v_2]$ and use the fact that every other step, the y block is rotated by eight bytes. An added bonus (for the attacker) is that now both types of the relations from above can be used. Namely, guessing these six bytes, one can derive four other bytes: a one-byte linear combination of x'_2, x'_3, a one-byte linear combination of x'_{10}, x'_{11}, and the bytes x_{10} and x_2. Notice that the input bytes in these relations are arranged in pairs, (x_0, x_8), (x_5, x_{13}), and (x_2, x_{10}), so we still use the same six bytes of y as masks, even after y is rotated by eight positions.

To make the notations below a little less horrible, we denote $x'_a = S_1[v_2]$, $x'_b = S_1[v_{10}]$, also also denote the one-byte linear combination of x'_2, x'_3 by x'_c, and the one-byte linear combination of x'_{10}, x'_{11} by x'_d. The important thing to remember is that x'_a, x'_b, x'_c, x'_d are all one-byte linear

[6] We note that guessing just one of x_0, x_{13} yields some linear combination of u_0, u_1, but since $M_2^{-1}M_1$ is MDS, we cannot get the "right" combination.

combinations of some bytes of x'. The relation that we use can now be described as

$$\begin{pmatrix} f_1(x_0) \\ f_1(x_8) \\ f_2(x_0) \\ f_2(x_8) \end{pmatrix} + \begin{pmatrix} g_1(x_{13}) \\ g_1(x_5) \\ g_2(x_{13}) \\ g_2(x_5) \end{pmatrix} + \begin{pmatrix} h_1(x'_a) \\ h_1(x'_b) \\ h_2(x'_b) \\ h_2(x'_a) \end{pmatrix} = \begin{pmatrix} x_2 \\ x_{10} \\ x'_c \\ x'_d \end{pmatrix} \tag{3}$$

where the f_i, g_i, h_i functions are known permutations. (For example, $f_1(\alpha) = S_2[\alpha]$, and $h_1(\alpha) = S_2[S_1^{-1}[\alpha]]$.)
To cancel out the y, z blocks and the masks $W[i]$, the attacker considers a "2×2 matrix of steps", i, $i+1$, $i+16$, $i+17$ for even i. Since i is even, then the y block is rotated by eight bytes between steps $i, i+1$ and between steps $i+16$, $i+17$. Also, the masks from W that are used in the steps $i-1, i, i+1$ are the same as in steps $i+15, i+16, i+17$. For example, below we consider steps 2, 3, 18 and 19 in the cipher. We use the following notations:

- We denote by $x(i)$ the input block to the F function in step i, and by $x'(i)$ the corresponding output block. As before, we use subscript to denote bytes withing a block. For example, $x(2)_{10}$ or $x'(18)_c$.
- We denote the y, z blocks at step 2 by $y1, z1$, and the y, z blocks at step 18 by $y2, z2$. Note that in step 3 we have $z = z1, y = y1_{8..15,0..7}$ and similarly in step 19 we have $z = z2, y = y2_{8..15,0..7}$.
- We denote the mask in steps 1,17 by $W[1]$, the mask in steps 2,18 by $W[2]$, and the mask in steps 3,19 by $W[3]$.

The attacker, watching the output stream, can see the sums of those bytes. For example, it can get the byte $x(2)_2 \oplus y1_2 \oplus W[1]_2$ from the output block in step 1, and the bytes $x'(2)_c \oplus z1_c \oplus W[2]_c$ and $x(3)_2 \oplus y1_{10} \oplus W[2]_2$ from the output block from step 2. In the sums that we get from these six steps, each byte of $y1, y2, z1, z2, W[1], W[2], W[3]$ appears exactly twice. The attacker can then sum-up these bytes, to eliminate the y's, z's and $W's$. For example, we have the following:

from step 2 we have	$x'(2)_c \oplus z1_c \oplus W[2]_c$
from step 3 we have	$x'(3)_c \oplus z1_c \oplus W[3]_c$
from step 18 we have	$x'(18)_c \oplus z2_c \oplus W[2]_c$
from step 19 we have	$x'(19)_c \oplus z2_c \oplus W[3]_c$
summing them, we get	$x'(2)_c \oplus x'(3)_c \oplus x'(18)_c \oplus x'(19)_c$

Another example, relating to input bytes is as follows:

from step 1 we have	$x(2)_0 \oplus y1_0 \oplus W[1]_0$
from step 2 we have	$x(3)_8 \oplus y1_0 \oplus W[2]_8$
from step 17 we have	$x(18)_0 \oplus y2_0 \oplus W[1]_0$
from step 18 we have	$x(19)_8 \oplus y2_0 \oplus W[2]_8$
summing them, we get	$x(2)_0 \oplus x(3)_8 \oplus x(18)_0 \oplus x(19)_8$

The attacker collect ten such sums, one for each byte in Eq. (3). We now claim that the distribution over these ten bytes is significantly different than the uniform distribution. To see that, we show that we can write these ten sums as

$$\langle s \oplus t \oplus u \oplus v,\ F_1(s) \oplus F_2(t) \oplus F_1(u) \oplus F_2(v) \rangle \tag{4}$$

for some function $F_1, F_2 : \{0,1\}^{48} \rightarrow \{0,1\}^{32}$, both known to the adversary. Indeed, it is easy to check that this is exactly what we get when we set

$$\begin{aligned}
s_{0..5} &= \langle x(2)_0,\ x(2)_8,\ x(2)_{13},\ x(2)_5,\ x'(2)_a,\ x'(2)_b \rangle \\
t_{0..5} &= \langle x(3)_8,\ x(3)_0,\ x(3)_5,\ x(3)_{13},\ x'(3)_a,\ x'(3)_b \rangle \\
u_{0..5} &= \langle x(18)_0, x(18)_8, x(18)_{13}, x(18)_5, x'(18)_a, x'(18)_b \rangle \\
v_{0..5} &= \langle x(19)_8, x(19)_0, x(19)_5, x(19)_{13}, x'(19)_a, x'(19)_b \rangle
\end{aligned}$$

and

$$F_1(a_{0..5}) \stackrel{\text{def}}{=} \begin{matrix} f_1(a_0) \\ f_1(a_1) \\ f_2(a_0) \\ f_2(a_1) \end{matrix} + \begin{matrix} g_1(a_2) \\ g_1(a_3) \\ g_2(a_2) \\ g_2(a_3) \end{matrix} + \begin{matrix} h_1(a_4) \\ h_1(a_5) \\ h_2(a_5) \\ h_2(a_4) \end{matrix} \qquad F_2(a_{0..5}) \stackrel{\text{def}}{=} \begin{matrix} f_1(a_0) \\ f_1(a_1) \\ f_2(a_1) \\ f_2(a_0) \end{matrix} + \begin{matrix} g_1(a_2) \\ g_1(a_3) \\ g_2(a_3) \\ g_2(a_2) \end{matrix} + \begin{matrix} h_1(a_5) \\ h_1(a_4) \\ h_2(a_5) \\ h_2(a_4) \end{matrix}$$

Moreover, each of the functions F_i is a sum, $F_i(a_{0..5}) = F_i^1(a_{0,1}) + F_i^1(a_{2,3}) + F_i^1(a_{4,5})$.

In our paper [1] we analyze distributions such as the one in Eq. (4). In that paper, it is shown that for the case of two functions $F_i : \{0,1\}^m \rightarrow \{0,1\}^{m'}$, each a sum of three functions $F_i^j : \{0,1\}^{m/3} \rightarrow \{0,1\}^{m'}$, we expect the statistical distance between the distribution from Eq. (4) and the uniform distribution to be $\sqrt{16/\pi} \cdot 2^{3(m'-m)/2}$. Plugging in the values $m = 48$, $m' = 32$ we get $\sqrt{16/\pi} \cdot 2^{3\cdot(32-48)/2} \approx 2^{-20.5}$.

The attack. The attack can now be succinctly described as follows: Watching the output stream, the attacker collects sufficiently many samples from the distribution of Eq. (4), until it can distinguish this distribution from random. Since we have statistical distance $2^{-20.5}$, then we need roughly 2^{41} such samples.

It can be shown that from 256 steps of the cipher, the attacker can collect about 2^{10} samples. To see this, note that each pair of y, z blocks is used 16 times before it is modified, and we can partition these 16 times into eight pairs $(i, i+1)$ (with even i) and use each of these pairs in place of steps 2,3 above. Next, the attack above looks at three consecutive masks, $W[i-1], W[i], W[i+1]$. In Scream-0, we have 14 "batches" of 16 steps where all three masks are the same, before we modify one of them. We can choose any pair of these batches to cancel out the masks W. Namely, instead of steps 2,3,18,19 as in the example above, we can use any four steps of the form $i, i+1, i+16j, i+16j+1$, with even $i \leq 16$ and as long as the masks $W[i-1 \bmod 16]$, $W[i \bmod 16]$, $W[i+1 \bmod 16]$ remain the same in all these steps (which means that we must have $j \leq 14$). This gives us $8 \times \binom{14}{2} \approx 2^{10}$ smaples from each collection of 16 batches (or 256 steps). To get the 2^{41} samples that we need, we therefore need to see about 2^{31} collesctions, which is 2^{43} bytes of output.

The attack works by collecting these 2^{41} samples, and for each one computing the probability of that value according to the distribution from Eq. (4). In our paper [1], we show how one can efficiently compute this distribution, by pre-computing some large tables that can be used during the attack, using Walsh-Hadamard transforms. For the case of Scream-0, where each function F_i is a sum of three functions, each from $m/3$ to m' bits, the total time that we spend on the attack is roughly $m' \cdot 2^{m'}$ per sample, and the space requirement is $3 \cdot 2^{(m/3)+m'}$. In our case, we have $m = 48$ and $m' = 32$, so we need about 2^{50} space and $2^{42} \cdot 32 \cdot 2^{32} = 2^{79}$ time.

C Constants and Test Vectors

```
The Rijndael S-box, S[0..255] = [
  63 7c 77 7b f2 6b 6f c5 30 01 67 2b fe d7 ab 76 ca 82 c9 7d fa 59 47 f0
  ad d4 a2 af 9c a4 72 c0 b7 fd 93 26 36 3f f7 cc 34 a5 e5 f1 71 d8 31 15
  04 c7 23 c3 18 96 05 9a 07 12 80 e2 eb 27 b2 75 09 83 2c 1a 1b 6e 5a a0
  52 3b d6 b3 29 e3 2f 84 53 d1 00 ed 20 fc b1 5b 6a cb be 39 4a 4c 58 cf
  d0 ef aa fb 43 4d 33 85 45 f9 02 7f 50 3c 9f a8 51 a3 40 8f 92 9d 38 f5
  bc b6 da 21 10 ff f3 d2 cd 0c 13 ec 5f 97 44 17 c4 a7 7e 3d 64 5d 19 73
  60 81 4f dc 22 2a 90 88 46 ee b8 14 de 5e 0b db e0 32 3a 0a 49 06 24 5c
  c2 d3 ac 62 91 95 e4 79 e7 c8 37 6d 8d d5 4e a9 6c 56 f4 ea 65 7a ae 08
  ba 78 25 2e 1c a6 b4 c6 e8 dd 74 1f 4b bd 8b 8a 70 3e b5 66 48 03 f6 0e
  61 35 57 b9 86 c1 1d 9e e1 f8 98 11 69 d9 8e 94 9b 1e 87 e9 ce 55 28 df
  8c a1 89 0d bf e6 42 68 41 99 2d 0f b0 54 bb 16 ]

The constant pi (for key-setup)
    pi        = [24 3f 6a 88 85 a3 08 d3 13 19 8a 2e 03 70 73 44]


Test vectors for Scream-S
    *** key-setup test vectors ***
    key       = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
    W0[0]     = [b6 a5 0b bf f3 9b 9e 99 28 b0 35 18 7b 7d 9c 7b]
    W0[15]    = [83 32 53 22 db 10 00 31 49 3a a4 80 3a 41 8c b3]

    *** nonce-setup test vectors ***
    key       = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
    nonce     = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
    X         = [b4 b7 7e 35 6a 24 0c c8 a7 41 b8 c7 d7 29 68 82]
    Y         = [e4 f4 1d 3b fd 07 d4 3c cb df a9 bb 25 df 65 6c]
    Z         = [87 de 72 cd 96 5a 96 24 b4 eb 79 66 57 26 fd f9]
    W[0]      = [66 d4 35 4d 2c 90 5f 0e 7f cc 25 59 43 ba d2 22]
    W[15]     = [a8 0e b6 56 be aa 5d d2 8d ca fe 07 1b f9 9c 7a]

    *** stream test vectors ***
    key       = [01 23 45 67 89 ab cd ef fe dc ba 98 76 54 32 10]

    nonce     = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
    out[0]    = [74 8c 59 f2 0d 76 9e a8 7a 6d c1 87 46 e6 4a c0]
    out[1]    = [bd 3b 39 cd 12 18 43 0f 80 fa e0 1b 2e 60 f1 74]
    out[4]    = [15 21 8a 46 fb ee 26 54 98 8d 2b 80 8a 87 f4 5e]
    out[16]   = [cb 32 f4 d6 f7 ce 57 69 e2 a3 ac d8 37 e1 37 82]
    out[1023] = [97 ec 87 f0 a0 6c e7 0b 75 e6 12 25 50 1f 82 e3]

    nonce     = [01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01]
    out[0]    = [47 68 06 37 83 85 99 af d2 8f fb 2e dd fc 9d 2e]
    out[1]    = [7b d3 0b e4 7a a6 3b 5f 4f 5f 05 06 66 17 d5 a2]
    out[4]    = [98 aa 20 75 73 c7 fa fc 1c 4c 27 61 46 14 3c 1d]
    out[16]   = [b3 33 a4 8e 17 50 8e ab b2 68 fb 60 67 56 46 1e]
    out[1023] = [a5 41 b3 37 c6 bd 8a 4b 41 a1 40 5f ea c5 a3 f5]


Test vectors for Scream-F
```

```
*** key-setup test vectors ***
key       = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
W0[0]     = [be a0 cd 9a 5d f6 85 59 c0 3f a9 c5 53 fd ad e1]
W0[15]    = [eb 2e ab 45 26 ee 49 e1 34 db 97 87 62 d1 3b 25]

*** nonce-setup test vectors ***
key       = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
nonce     = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
X         = [d4 10 c5 bf bd 7b fd 81 37 4e e3 b0 c1 bf 8b a6]
Y         = [51 a6 7f 38 3d 0d 95 26 bf b5 b0 e8 26 b5 e4 93]
Z         = [53 50 b7 d6 87 3d df 8c 7f 9b 10 7c e0 92 d0 02]
W[0]      = [cb ad d5 c2 b0 85 af 77 6c d8 ef ce 7b 36 65 3a]
W[15]     = [19 14 5e 0a 4d 23 1c d5 f9 6f 85 8a 39 38 81 a1]

*** stream test vectors ***
key       = [01 23 45 67 89 ab cd ef fe dc ba 98 76 54 32 10]

nonce     = [00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00]
out[0]    = [39 ec 4a 06 45 4d c3 cd 96 dd ef 0c f0 c2 67 40]
out[1]    = [a0 ea 56 e7 e3 c8 f5 df 34 ea 35 ee 77 ed da 66]
out[4]    = [8a c8 93 af 83 ed 0a 53 6b e9 f4 7c b6 6d 21 67]
out[16]   = [e0 8c fe 31 34 a7 48 ca 14 10 f9 58 50 71 49 20]
out[1023] = [a4 e2 fc be 0a 47 53 9a 23 e0 79 25 5c be ea e7]

nonce     = [01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01]
out[0]    = [2e 70 fb 8c d5 d8 50 a8 94 38 0e 85 46 9d 33 fc]
out[1]    = [33 39 da 86 9c a1 f7 1b 3a d0 16 16 ea 42 24 1a]
out[4]    = [1a 79 cf 13 01 67 2c 52 25 13 8c c8 89 fb 50 72]
out[16]   = [c8 f2 3f ca 4e 0c 47 46 1a b3 7b 34 1b 57 c7 96]
out[1023] = [6e 63 21 c1 9b 49 08 57 84 87 14 ea 4f 08 4b 7d]
```