Radboud University

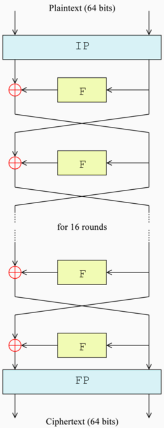
The origins of Rijndael (AES)

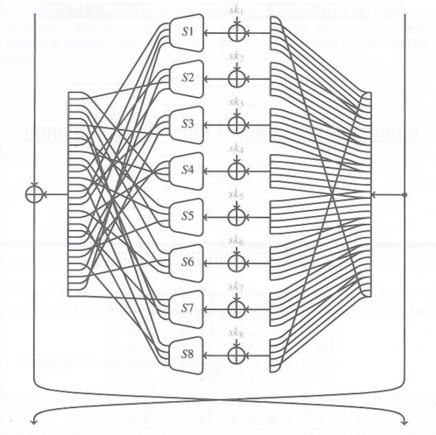
Joan Daemen, Radboud University, NL

AES workshop

Monday April 13, 2026, Gausta

ESCADA

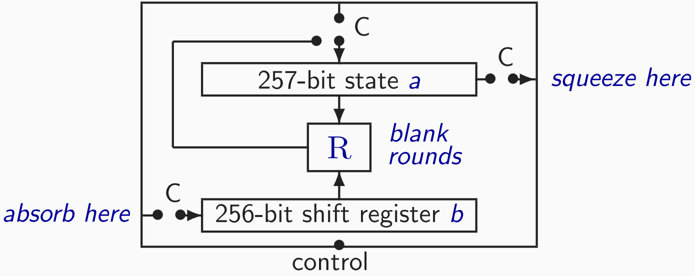




Copied from The Block Cipher Companion, Lars Knudsen and Matt Robshaw

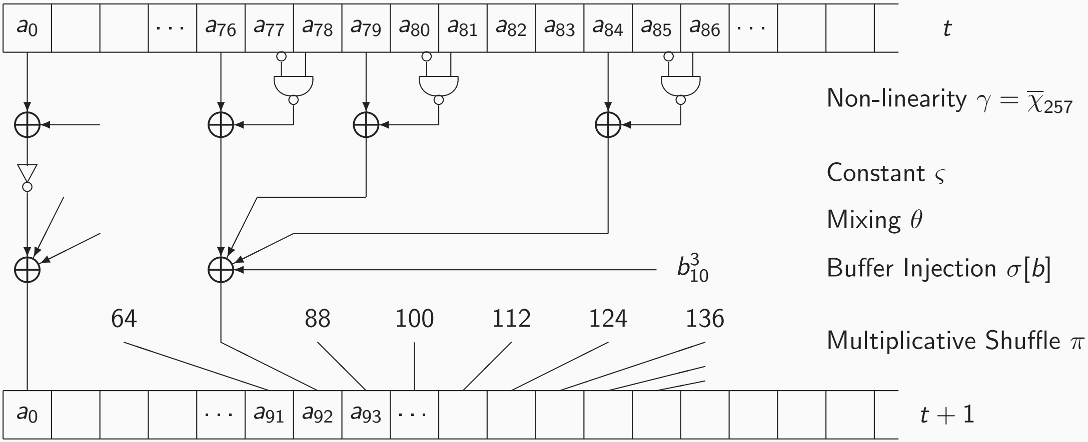
No dedicated mixing layer; Magic (Diffusion & Confusion) resides in the 8 *S-boxes*

SUBTERRANEAN [Daemen et al., 1992]: a stream/hash module



Substream:	short $K; D$	$\rightarrow$	long Z
Subhash:	long M	$\rightarrow$	short T

The round function of SUBTERRANEAN: lots of shift-invariance



Great but "Block ciphers that are efficient in software, that what you should work on!"

Intermezzo: Branch number $\mathcal{B}$ [Daemen et al., 1993]

- Mixing layer θ: $b_i \leftarrow a_i + a_{i+3} + a_{i+8}$
- Fan-in: number of terms in right-hand side
  - Measures computational effort: fan-in d requires $d-1$ XORs per bit
  - Measures diffusion power: 1 bit spreads to (or depends on) 3 bits
- Consider θ: $b_i \leftarrow a_i + a_{i+1} + a_{i+2} + a_{i+3} + a_{i+4}$ with fan-in 5
- Can be implemented with 2 XORs per bit
  - $c_{2i} \leftarrow a_{2i} + a_{2i+1}$
  - $d_{2i} \leftarrow c_{2i} + c_{2i+2}$
  - $b_{2i} \leftarrow d_{2i} + a_{2i+4}$
  - $b_{2i-1} \leftarrow a_{2i-1} + d_{2i}$
- Input difference 11 gives output difference 100001
- Better measure for the diffusion power: the branch number $\mathcal{B}$

Block ciphers with SUBTERRANEAN philosophy: 3-WAY and BASEKING

- Circulant symmetry and . . .
  - suitable for software and hardware
  - encryption and decryption very similar
  - round constants but no key schedule
  - only bitwise Boolean operations and cyclic shifts
- Structure
  - 4-layer round function alternated with round key addition

    θ circulant mixlayer

    π_1 first (ShiftRows-like) bit shuffle

    γ non-linear S-box layer

    π_2 second (ShiftRows-like) bit shuffle

  - extra θ at the end

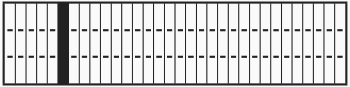
The γ S-box: $\bar{\chi}_3$

x	000	001	010	100	110	101	011	111
y	111	010	100	001	011	110	101	000

Computation (per bit): 1 XOR + 1 AND + 1 NOT or 1 2-NAND + 2 3-NAND

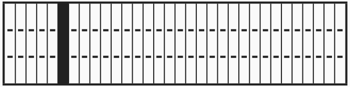
0 $3k-1$

3 rows:



0 tot $k-1$
k tot $2k-1$
$2k$ tot $3k-1$

$\Downarrow \gamma$



Propagation properties of χ_3

DP-table, aka scaled DDT

	0	1	2	3	4	5	6	7
0	4							
1		1		1		1		1
2			1	1			1	1
3		1	1			1	1	
4					1	1	1	1
5		1		1	1		1	
6			1	1	1	1		
7		1	1		1			1

LP-table, elements are $4C^2(u, v)$

	0	1	2	3	4	5	6	7
0	4							
1		1		1		1		1
2			1	1			1	1
3		1	1			1	1	
4					1	1	1	1
5		1		1	1		1	
6			1	1	1	1		
7		1	1		1			1

Masks and differences propagate in the same way!

Inverse mapping in $\mathbb{F}_8$ with degree 2

Power functions of $\mathbb{F}_{2^n}$ are popular in the BFA community, in particular:

- Gold functions:
  $y \leftarrow x^e$ with $\mathrm{HW}(e) = 2$
- Inverse function (Kaisa):
  $y \leftarrow x^{2^n-2} (= x^{-1}$ for $x \neq 0)$

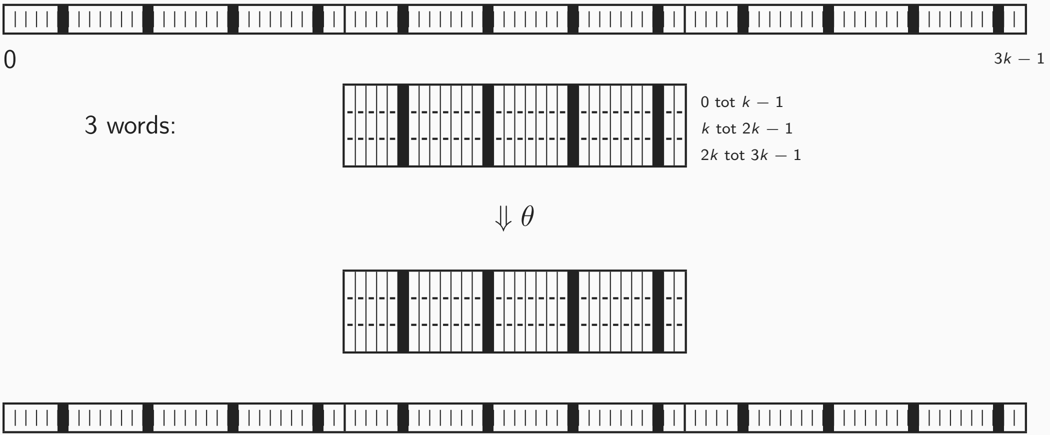
χ_3: $y \leftarrow x^6$ in $\mathbb{F}_{2^3}$ with a normal basis

- Singular object with great symmetry
- There exists no other such S-box

DP/LP-table

		α^0	α	α^2	α^3	α^4	α^5	α^6
		7	1	2	5	4	6	3
α^0	7	1	1	1		1		
α	1	1	1		1			1
α^2	2	1		1			1	1
α^3	5		1			1	1	1
α^4	4	1			1	1	1	
α^5	6			1	1	1		1
α^6	3		1	1	1		1	

The mixlayer θ, applied to 12-bit units



M_θ, the circulant matrix of the mixlayer θ

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \end{pmatrix} = \begin{pmatrix}
1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 & 1 & 1 \\
1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 & 1 \\
1 & 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 \\
1 & 1 & 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot \\
\cdot & 1 & 1 & 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 \\
1 & \cdot & 1 & 1 & 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 \\
1 & 1 & \cdot & 1 & 1 & 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & 1 & 1 & \cdot & 1 & 1 & 1 & 1 & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & 1 & 1 & \cdot & 1 & 1 & 1 & 1 & \cdot & 1 & \cdot \\
\cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 & 1 & 1 & 1 & \cdot & 1 \\
1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 & 1 & 1 & 1 & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & 1 & 1 & 1 & 1
\end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \end{pmatrix}$$

- Computation (per bit): 3.5 XOR, depth 4 or 4.25 XOR, depth 3
- $\mathrm{M}_\theta^{-1} = \mathrm{M}_\theta^{\mathrm{T}}$: masks and differences propagate the same way!

Diffusion properties of θ

HW(x)\ HW($\theta(x)$)	1	2	3	4	5	6	7	8	9	10	11
1	-	-	-	-	-	-	12	-	-	-	-
2	-	-	-	-	-	60	-	-	-	6	-
3	-	-	-	-	180	-	-	-	40	-	-
4	-	-	-	255	-	-	-	240	-	-	-
5	-	-	180	-	-	-	600	-	-	-	12
6	-	60	-	-	-	804	-	-	-	60	-
7	12	-	-	-	600	-	-	-	180	-	-
8	-	-	-	240	-	-	-	255	-	-	-
9	-	-	40	-	-	-	180	-	-	-	-
10	-	6	-	-	-	60	-	-	-	-	-
11	-	-	-	-	12	-	-	-	-	-	-

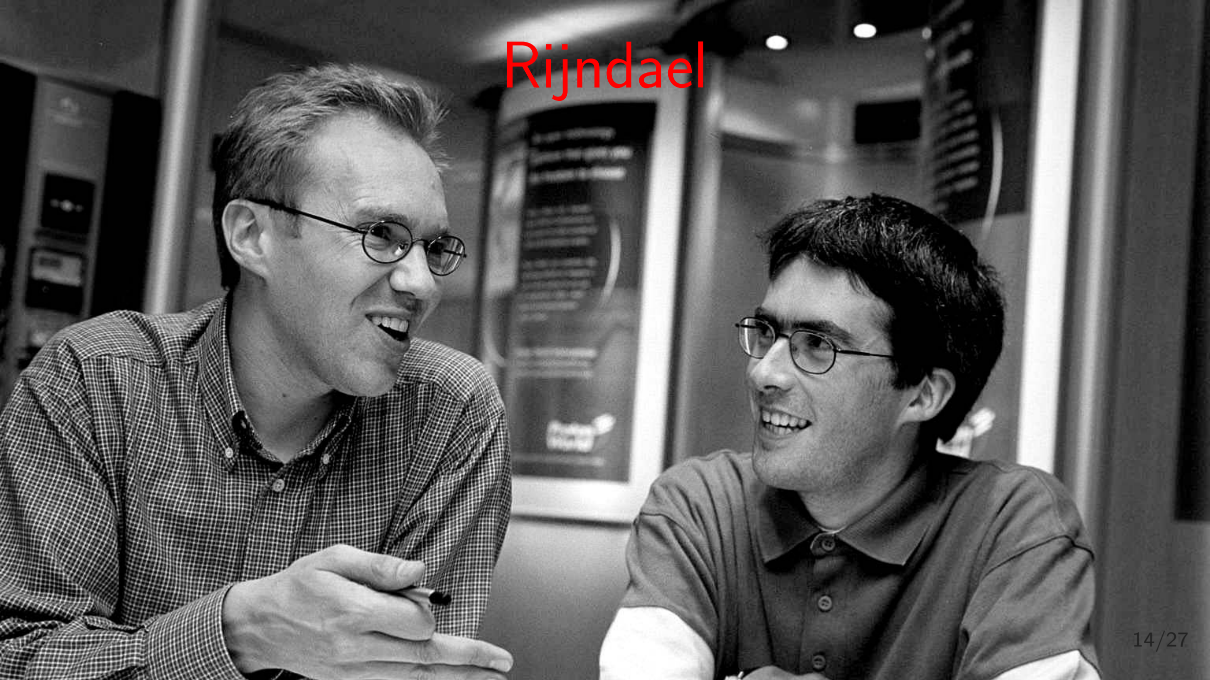
- Branch number $\mathcal{B} = 8$
- Corresponds to $[24, 12, 8]$-code: extended binary Golay code
- Singular object with amazing symmetry

Resulting block ciphers

- Two block ciphers:
  - 3-Way [Daemen et al., 1993]: 96-bit block and key, in 3 32-bit rows
  - BaseKing [Daemen 1995]: 192-bit block and key, in 12 16-bit rows
- The first wide-trail block ciphers
- Supersymmetry
  - one-to-one mapping between differential and linear trails
  - propagation $\leftarrow$ same as $\rightarrow$ with bit order reversed
- Implementation
  - (almost) same circuit for cipher and its inverse
  - (almost) same code for cipher and its inverse
  - difference lies in round constant generation and round key addition
- (first) *bit-slice* block ciphers: code with only Boolean instructions and cyclic shifts

. . . I was very enthusiastic . . . but nobody else

Rijndael

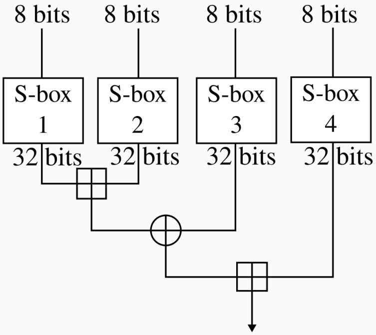


A Fortunate Event

- Summer 1993:
    - COSIC gets some classified contract work
    - Security evaluation of a proprietary cipher
- Profs. Vandewalle and Govaerts decide to put on it:
    - Joan Daemen
    - Vincent Rijmen
- Result of contract work:
    - Some new types of cryptanalysis
    - Classified, unfortunately
    - Later *re-invented* and published
    - . . . by someone else

March 1995: looking at Blowfish [Schneier, 1993]

- My last month in COSIC
- Blowfish cryptanalysis contest
- 64-bit Feistel cipher with F function:

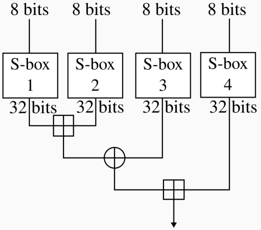


- 8-to-32-bit S-boxes, derived from key

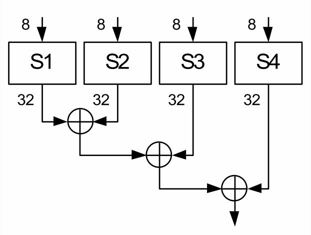
- Great potential
  - only 4 TLU and 3 additions
  - very high diffusion
- Contest won by Serge Vaudenay
  - Exploiting local collisions
  - In S-box: weak keys
  - In F-function
  - [Vaudenay, 1996]
- But I thought I could fix it

Fixing the Blowfish F function in 4 steps

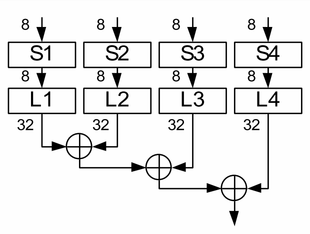
0: Blowfish



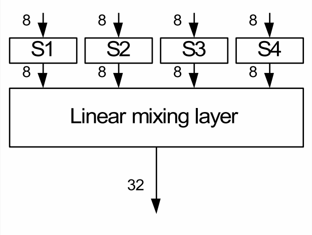
1: additions → XORs



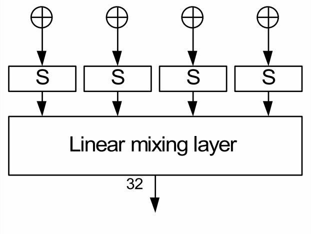
2: split linear non-linear



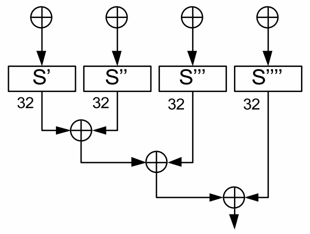
3: integrate L-boxes



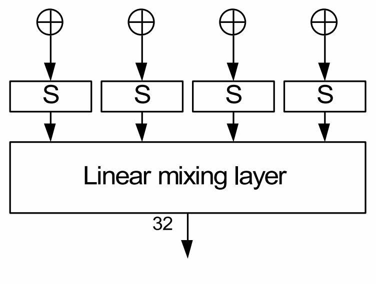
4: split key S-box



5: TLU still works



March 1995: a core idea of Rijndael/AES takes shape



- Mixing $\circ$ S-box $\circ$ key
- Each invertible
- 4 TLU and 4 XORs

- No need for Feistel
  - 64-bit block: 8 bytes wide
  - 128-bit block: 16 bytes wide
- S-boxes
  - just take a single one
  - with minimal max DP and LP
- Linear mixing layer
  - Maximum $\mathcal{B} : n + 1$
  - $n = 8, 16$: seemed possible
- Challenge: finding decent S-box and mixing layer

Autumn of 1995: SHARK

- I could not not use these ideas but had very little time for it
- I decided to contact Vincent to do it together
  - later Bart Preneel, Antoon Bosselaers and Erik De Win joined
  - result: paper on SHARK [SHARK, FSE 1996]
- 3-layers: key addition, *b*-bit S-boxes and *n*-wise mixing
- Mixing layer with maximum branch number
  - Link with maximum distance separable (MDS) codes
  - # active S-boxes per two rounds $\geq \mathcal{B} = n + 1$
- Concretely in SHARK:
  - $b = 8, n = 8$ so block length is 64
  - S-box: multiplicative inverse in $\mathbb{F}_{2^8}$ [Nyberg, 1994]

The trouble with SHARK

- In general: 8-wise and 16-wise MDS layers are expensive
  - software: n look-up tables with 2^8 entries each n bytes
  - hardware: # gates per bit grows quickly as a function of n

1996: Square

- I had an idea: just add a *shuffle* layer
  - as in Subterranean, 3-Way and BaseKing
  - promising pencil-and-paper exercises
  - # active S-boxes per 4 rounds always large!
- I contacted Vincent again to work this out
  - this led to Square [Square, FSE 1997]
  - later Lars Knudsen joined

The Square approach

- Add an optimal shuffle layer moving bytes around
  - moves bytes in MDS block to all different MDS blocks
  - we proved: # active S-boxes per four rounds $\geq \mathcal{B}^2$
- SQUARE concretely
  - 16 bytes in 4 by 4 square
  - mixing layer: circulant matrix over $\mathbb{F}_{2^8}$ with *small* elements
  - shuffle: taking transpose of the square
  - S-box: same one as later in Rijndael
    - added affine layer
    - to counter *interpolation attacks* [Jacobsen, Knudsen, '97]
    - a 4-round trail has at least 25 active S-boxes
  - lightweight recursive key schedule that is **linear**

Lars Knudsen's Square attack

- Our working version of SQUARE had only 6 rounds because:
  - DP of 4-round differential trails $\leq 2^{-150}$
  - LP of 4-round linear trails $\leq 2^{-150}$
- Lars' Square attack
  - input sets: constant in some and complete in other bytes
  - properties decay only slowly through steps of the round
  - 4-round distinguisher, breaking full 6 rounds
  - lesson learnt: interpret trail bounds with caution
- How we fixed it:
  - increase number of rounds to 8
  - ask Lars as co-author and include attack in paper

Winter 1996-1997: BKSQ

- Need for 96-bit block cipher for hash-based signature scheme at Banksys
- I contacted Vincent again
- This resulted in SQUARE variant BKSQ [Cardis 1998]:
  - 12-byte block length instead of 16-byte
  - MDS operating separately on 3-byte columns
  - shuffle: transpose replaced by `ShiftRows`-like
  - linear recursive key schedule
- External evaluation by two independent parties:
  - both produced a report: no weaknesses found
  - but reports had concerns with linear key schedule
- Lessons learnt:
  - `ShiftRows` shuffle allows varying block size
  - linear key schedule raises eyebrows

The AES competition

- January 1997: NIST announces the AES initiative
  - replacement of DES
  - open call for block cipher proposals
  - ... and for analysis, comparisons, etc.
  - draft call requires several block and key lengths
- We had already most ingredients in SQUARE and BKSQ while the other teams still had to start

Rijndael

- 25 block ciphers with block and key lengths $\in \{128, 160, 192, 224, 256\}$
- Efficient in software: 1 table-lookup and XOR per byte per round
- Big S-box not optimal for hardware but symmetry gives much flexibility:
  - all rounds are identical
  - ...except for the round keys and MixColumns in last round
  - parallel and symmetric
- Key schedule
  - expansion of cipher key to round key sequence
  - recursive procedure that can be done in-place
- NIST only standardized block length 128

Conclusion

- Rijndael was a triomf and AES is everywhere, even as instructions on many CPUs

  . . .
- But I always liked much more the SUBTERRANEAN and 3-WAY/BASEKING approach
  - NOEKEON was an attempt at a 3-WAY-like 128-bit block cipher
  - KECCAK and even ASCON build further on the approach
- I think building standard crypto around block ciphers was a mistake
  - unkeyed crypto: sponge/XOF is a better alternative
  - keyed crypto: deck (and duplex) functions

Thanks for your attention!