The making of Rijndael

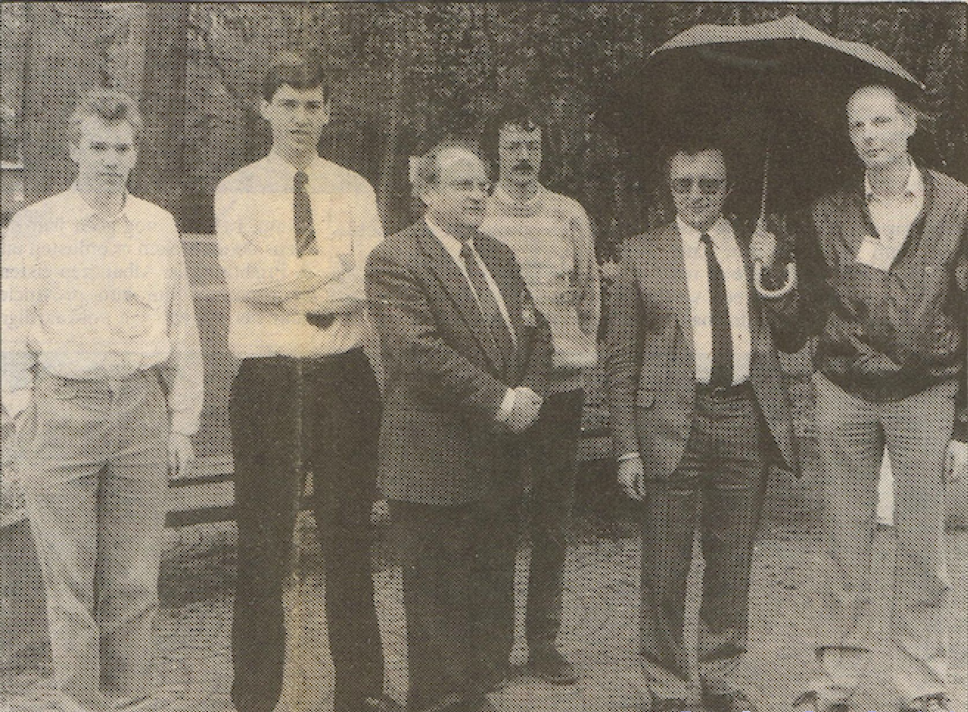
Joan Daemen[1] and Vincent Rijmen[2]

[1]Radboud University [2]KU Leuven

Histocrypt 2019,
June 23-26, Mons, Belgique

Dark ages of symmetric cryptography

- Before the discovery of differential and linear cryptanalysis
- Research in stream ciphers dominated by
  - linear feedback shift register (LFSR) bases schemes
  - study of properties of sequences, e.g., linear complexity
- Research in block ciphers dominated by
  - Data Encryption Standard (DES)
  - design rationale not published
  - study of properties of (vector) boolean functions



How to build a block cipher, back then

- Claude Shannon's concepts:
  - confusion: mainly associated with non-linearity
  - diffusion: mixing of bits
- Property-preserving paradigm
  - *security of a block cipher lies in its S-boxes*
  - build strong cipher with right properties
  - ...from S-boxes with same properties
- Non-linearity
  - distance to linear functions
  - *bent* and *almost perfect non-linear* (APN) functions
- Diffusion
  - avalanche effect
  - strict avalanche criterion (SAC)

Discovery of differential and linear cryptanalysis

- Biham and Shamir 1990: differential cryptanalysis (DC)
  - exploits high-probability difference propagation
  - to guess a partial key used in remaining rounds
  - propagation along *trail* Q with probability $\mathrm{DP}(Q)$
- Matsui 1992: linear cryptanalysis (LC)
  - exploits high correlations over all but a few rounds
  - to guess a partial key used in remaining rounds
- Statistical attacks that require many input-output pairs
- Many flavours, variants and combinations exist
- LC/DC resistance is foundation of block cipher design

Meanwhile Joan @ COSIC: wide trail strategy

- Probability of a difference propagation trail: $\mathrm{DP}(Q)$ is the product of those of its active S-boxes: $\prod_i \mathrm{DP}(\mathrm{Sbox}_i)$
- DC of DES: few active S-boxes per round
- decrease $\mathrm{DP}(Q)$: S-boxes with low maximum DP
  - $\mathrm{DP}_{\mathrm{max}}(\mathrm{Sbox}) \geq 2^{1-b}$ with b S-box width
  - so this implies big S-boxes
- Cost of S-boxes strongly increases with size
  - Software: lookup tables of size 2^b
  - Hardware: increase of combinatorial logic

Principle of the wide trail strategy [PhD Daemen, 1995]

- **Many** active S-boxes rather than *big* S-boxes
  - Assure that any trail has many active S-boxes
  - multiple active S-boxes per round
- Separate layers for nonlinearity, mixing and dispersion
  - nonlinear layer: e.g. S-boxes with some max. DP and LP
  - mixing layer: local spreading of differences (and correlations)
  - dispersion layer: moves nearby bits to remote positions

Branch number

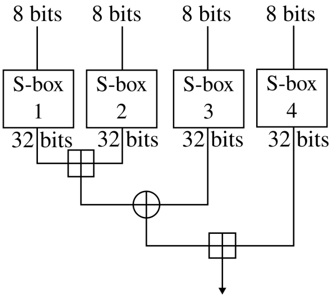
- Desired properties of diffusion layer:
  - avalanche: few active S-boxes at input $\rightarrow$ many at output
  - avalanche of inverse: few at output $\rightarrow$ many at input
- Branch number $\mathcal{B}$ of a diffusion layer
  - minimum number of active S-boxes at input and output
  - two types: linear and differential
  - relative to a state partition in bits, bytes, or ...

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: a core idea of Rijndael/AES takes shape

- Joan's last month in COSIC
- Blowfish [Schneier, 1993]
- F function:

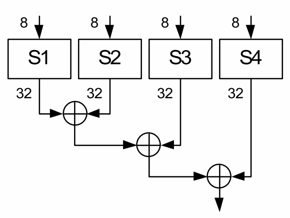


- 8-to-32-bit Sboxes
- Derived from key

- Great potential
  - Only 4 TLU and 3 add.
  - Very high diffusion
- Cryptanalysis contest
- Won by Serge Vaudenay
  - Exploiting local collisions
  - In S-box: weak keys
  - In F-function
  - [Vaudenay, 1996]
- But it can be fixed

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

- Mixing ∘ S-box
- Both invertible

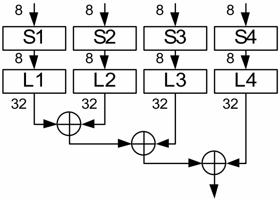


- 4 TLU and 4 XORs

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

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

- Mixing ∘ S-box
- Both invertible

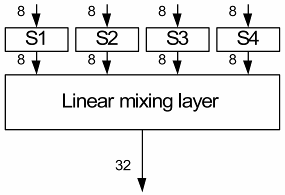


- 4 TLU and 4 XORs

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

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

- Mixing ∘ S-box
- Both invertible

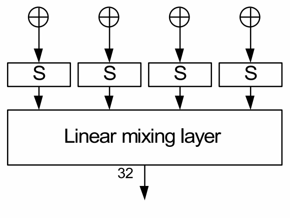


- 4 TLU and 4 XORs

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

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

- Mixing ∘ S-box
- Both invertible

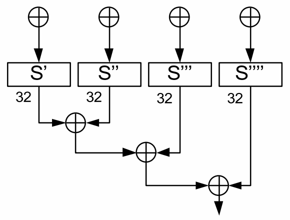


- 4 TLU and 4 XORs

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

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

- Mixing ∘ S-box
- Both invertible



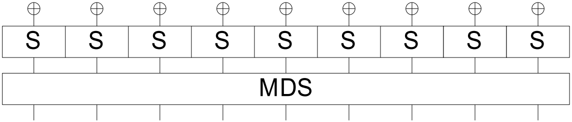
- 4 TLU and 4 XORs

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

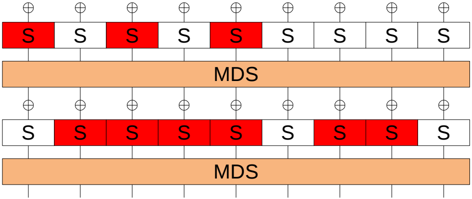
Autumn of 1995: SHARK

- Joan contacts Vincent to work out these ideas
  - 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 GF(2^8) [Nyberg, 1994]

SHARK principle illustrated



SHARK principle illustrated



The trouble with SHARK

- In general: n-wise MDS layer is expensive
  - software: n look-up tables with 2^8 entries of size $8n$
  - hardware: # gates per bit grows quickly as a function of n
  - instead of an expensive S-box,
  - ...we now have an expensive MDS matrix

1996: Square

- Idea: add a *dispersion* layer
  - like in earlier designs, e.g., Subterranean and 3-Way
  - promising pencil-and-paper exercises
  - # active S-boxes per 4 rounds always large!
- Joan contacts Vincent again to work this out
  - this led to Square [Square, FSE 1997]
  - later Lars Knudsen joined

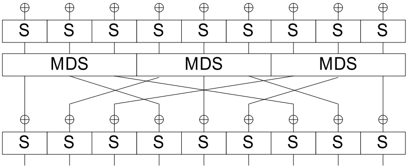
The Square approach

- Add a dispersion layer moving bytes around
- Use an optimal dispersion layer
  - 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
  - same mixing layer as Rijndael: circulant matrix
  - dispersion: taking transpose of square
  - S-box: same one as later in Rijndael
    - added affine layer
    - to counter *interpolation attacks* [Jacobsen, Knudsen, '97]
  - lightweight linear recursive key schedule

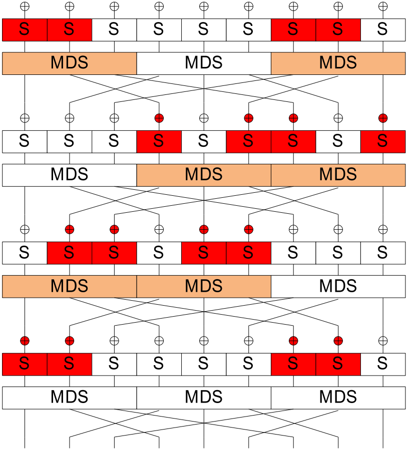
Lars' Square attack

- Our working version 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

The Square approach illustrated



The Square approach illustrated



Winter 1996-1997: BKSQ

- Need for 96-bit block cipher for Lamport-like signatures
- Joan contacts Vincent again
- This resulted in SQUARE variant BKSQ [Cardis 1998]:
  - 12-byte blocks instead of 16-byte
  - MDS operating separately on 3-byte columns
  - dispersion: 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` dispersion allows varying block size
  - linear key schedule raises eyebrows

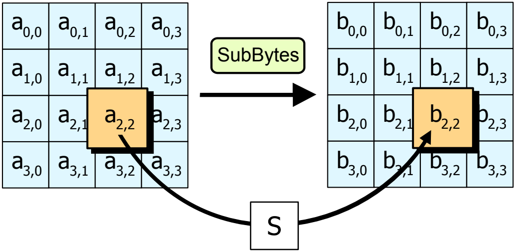
The start of the AES competition

- 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
- Remained to do:
  - specify a non-linear key schedule
  - prepare the documentation
  - prepare reference code
  - ...still more work than expected

Rijndael

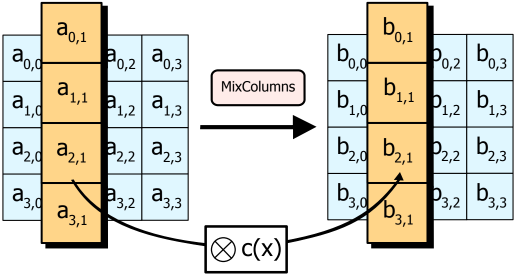
- Block cipher with block and key lengths $\in \{128, 160, 192, 224, 256\}$
- Simple round function with four steps
  - all rounds are identical
  - ...except for the round keys
  - parallel and symmetric
- Key schedule
  - expansion of cipher key to round key sequence
  - recursive procedure that can be done in-place

The non-linear layer: SubBytes



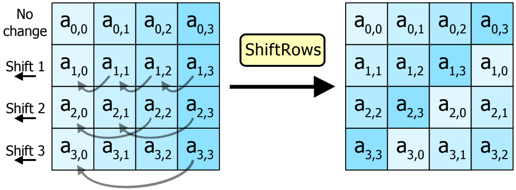
- Single S-box with two layers
- $y = x^{-1}$ in $GF(2^8)$, or more exactly $y = x^{254}$
    - max LP = max DP = 2^{-6} [Nyberg, Eurocrypt 1993]
- Affine mapping: multiplication by 8×8 matrix in GF(2)
    - to counter interpolation attacks [Jacobsen, Knudsen, FSE 1997]

The mixing layer: MixColumns



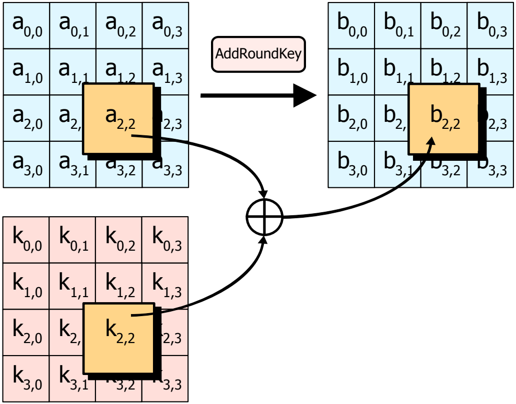
- Single MDS mapping applied to columns
- Multiplication by a 4×4 circulant matrix in GF(2^8)
  - Elements: 1, 1, x and $x + 1$
  - *circulant MDS matrix with the simplest elements*
  - Inverse has more complex elements

The dispersion layer: ShiftRows



- Each row is shifted by a different amount
- Different shift offsets for higher block lengths

Round key addition: AddRoundKey



Key schedule: 192-bit key, 128-bit block example

k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	$\cdots$

Round key 0	Round key 1	Round key 2	...

$$
\begin{aligned}
k_{6n} &= k_{6n-6} \oplus f(k_{6n-1}) \\
k_i &= k_{i-6} \oplus k_{i-1}, \quad i \neq 6n
\end{aligned}
$$

Rijndael: some distinguishing features

- Symmetric and (too) simple (to be secure)
- Inverse is different and slightly more expensive
- Table-lookup implementation:
  - 4 Kbytes of table
  - 1 table-lookup + 1 XOR per byte per round
  - inverse uses different tables
- No integer arithmetic

The Rijndael book

- Springer approached us for writing a book on Rijndael
  - more work than expected
  - very learnful experience
- New insights on LC and DC of key-alternating ciphers
  - linking linear trails, correlations and linear probability (LP)
  - clear and clean expressions
- Rijndael-GF
  - $GF(2^8)$ only: matrix in $GF(2)$ becomes *linearized polynomial*
  - linear cryptanalysis native in $GF(2^n)$

Conclusions

- Design process took years of elapsed time
- Ideas used from an even longer period
- But result seems to be tough: shape of AES 2019 AD:
  - theoretical security: small dents in armour
  - practical security: no threat
- ...and inspiring for both design (and attacks)
  - block ciphers and compression functions
  - stream ciphers
  - iterated permutations

Thanks for listening!