

Floating point

Registri Floating-Point

32 registri

width dipende dal' implementazione

32 bit  $\subset$  64 bit  $\subset$  128 bit

| Name          | Register Number | Usage                            |
|---------------|-----------------|----------------------------------|
| <b>ft0-7</b>  | f0-7            | Temporary variables              |
| <b>fs0-1</b>  | f8-9            | Saved variables                  |
| <b>fa0-1</b>  | f10-11          | Function arguments/Return values |
| <b>fa2-7</b>  | f12-17          | Function arguments               |
| <b>fs2-11</b> | f18-27          | Saved variables                  |
| <b>ft8-11</b> | f28-31          | Temporary variables              |

Sotto-argomento 2

Eccezione

| Flag Mnemonic | Flag Meaning      |
|---------------|-------------------|
| NV            | Invalid Operation |
| DZ            | Divide by Zero    |
| OF            | Overflow          |
| UF            | Underflow         |
| NX            | Inexact           |

le istruzioni floating-point non triggerano  
nessua eccezioni

settano un flag che può essere letto nel

FRCSR

puo essere letto tramite

**fcsr** è un **CSR**

che lo restituisce su un registor intero

**fcsr** Floating-Point Control and Status Register

Not a number NaN

esistno due tipi

Signalling NaN

segnale una operazione invalida

Quiet NaN

si propaga nelle operazione

MSB della mantissa = 1

Estensioni

in riferimento allo standard IEEE 754

**RVF** single-precision 32-bit

1 sign bit

8 exponent bit

23 fraction bit

**RVD** double-precision 64-bit

1 sign bit

11 exponent bit

52 fraction bit

**RVQ** quad-precision 128-bit

1 sign bit

15 exponent bit

112 fraction bit

Eccezioni

Positive infinity  $+\infty$

sign bit = 0

exponent bit = tutti 1

fraction bit = tutti 0

Negative infinity  $-\infty$

sign bit = 1

exponent bit = tutti 1

fraction bit = tutti 0

Positive Zero  $-0$

sign bit = 0

exponent bit = tutti 0

fraction bit = tutti 0

Negative Zero  $-0$

sign bit = 1

exponent bit = tutti 0

fraction bit = tutti 0

Not a number NaN

sign bit = x

exponent bit = tutti 1

fraction bit = x

le x possono essere qualisasi bit