MULTIPLE LANES

we can have array of 11 functional units or a combination 11 and pipelined units.

we can increase the peak throughput of a vector unit by adding more lanes

designers to trade off area, clock rate, voltage and energy without sacrificing performance.

VECTOR LENGTH REGISTERS

max. vector length (32) (mvl)

vector length register (vl)

```
for (i=0; i<n; i=i+1)
  y[i] = a * x[i] + y[i]
```

to add a vector length register (VL)
the VL controls the length of any vector operation including vector load and store.

when the vector is longer than the max. length, a technique called strip mining is used.

vector operation is done for a size <= MVL

PREDICATE REGISTERS

```
for (i=0; i<64; i=i+1)
    if (X[i] != 0)
        X[i] = X[i] - Y[i]
```

```
vsetdcfg   2*FP64      # enable 2 64-bit FP vector
                       # registers
vsetpcfgi  1           # enable 1 predicate register.
vld   v0, X5           # load vector X into V0
vld   v1, X6           # load vector Y into V1
fmv.d.x  f0, x0        # put FP zero into f0
vpne  p0, v0, f0       # Set p0(i) to 1 if v0(i) ≠ 0
vsub  v0, v0, v1       # subtract under vector.
vst   v0, x5           # store the result in X
vdisable               # disable vector registers.
vpdisable              # disable predicate registers.
```

MEMORY BANKS (storage) (holding area).

supplying bandwidth for vector load/store units.
most vector processor use memory bank which allow several independent accesses, rather than memory interleaving for following reasons.

1. Many vector computers support many loads or stores per clock cycle and the memory bank cycle time is usually several times larger than the processor cycle time.

2. most vector processor support the ability to load or store data words that are not sequential.

3. most vector pro computers support multiple processors sharing the same memory system, so each processor will be generating its own separate stream of addresses.

STRIDE

```
for (i=0; i<100; i++)
  for (j=0; j<100; j++)
    A[i][j]=0;
    for (k=0; k<100; k++)
      A[i][j]= A[i][j] + B[i][k] * D[k][j];
```

the distance separating elements to be gathered into a single vector register is called the stride.

GATHER AND SCATTER

```
for( i=0; i<n; i++)
  A[K[i]] = A[K[i]] + C[M[i]]
```

index vectors	X5	X6	X7	X28
	A	C	K	M

```
vsetdcfg    4*FP64      # 4 64 bit vector registers
vld         V0, X7      # load K[]
vldx        V1, X5, V0  # load A[K[]]
vld         V2, X28     # load M[]
vldi        V3, X6, V2  # load C[M[]]
vadd        V1, V1, V3  # add them
vstx        V1, X5, V0  # store A[K[]]
vdisable                # disable vector registers.
```

try vect asm.

gfx and computing gpus — fermi kepler pascal.

john nickolls. and david kirk

NVIDIA whitepapers

IEEE micro.