University of Zurich[UZH]

High Performance Computing Lecture 2

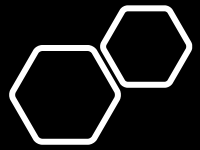
Douglas Potter
Jozef Bucko
Noah Kubli

Last Week

- What is HPC?
- What is a Supercomputer?
- Connecting to Piz Daint & Eiger
- SSH & X11
- Editors
- Unix / Linux

Today's Topics

- Revision Control Systems
- How compilers work
- Piz Daint "Environment"
- Batch Queues
- Processes & Threads
- Compiling and running code



Online Resources

http://goalkicker.com/

- Git Notes for Professionals
- Linux Notes for Professionals
- Bash Notes for Professionals
- C Notes for Professionals

GNU Make

- GNU Autoconf
- CMake

CSCS User Portal

- SLURM Jobscript Generator

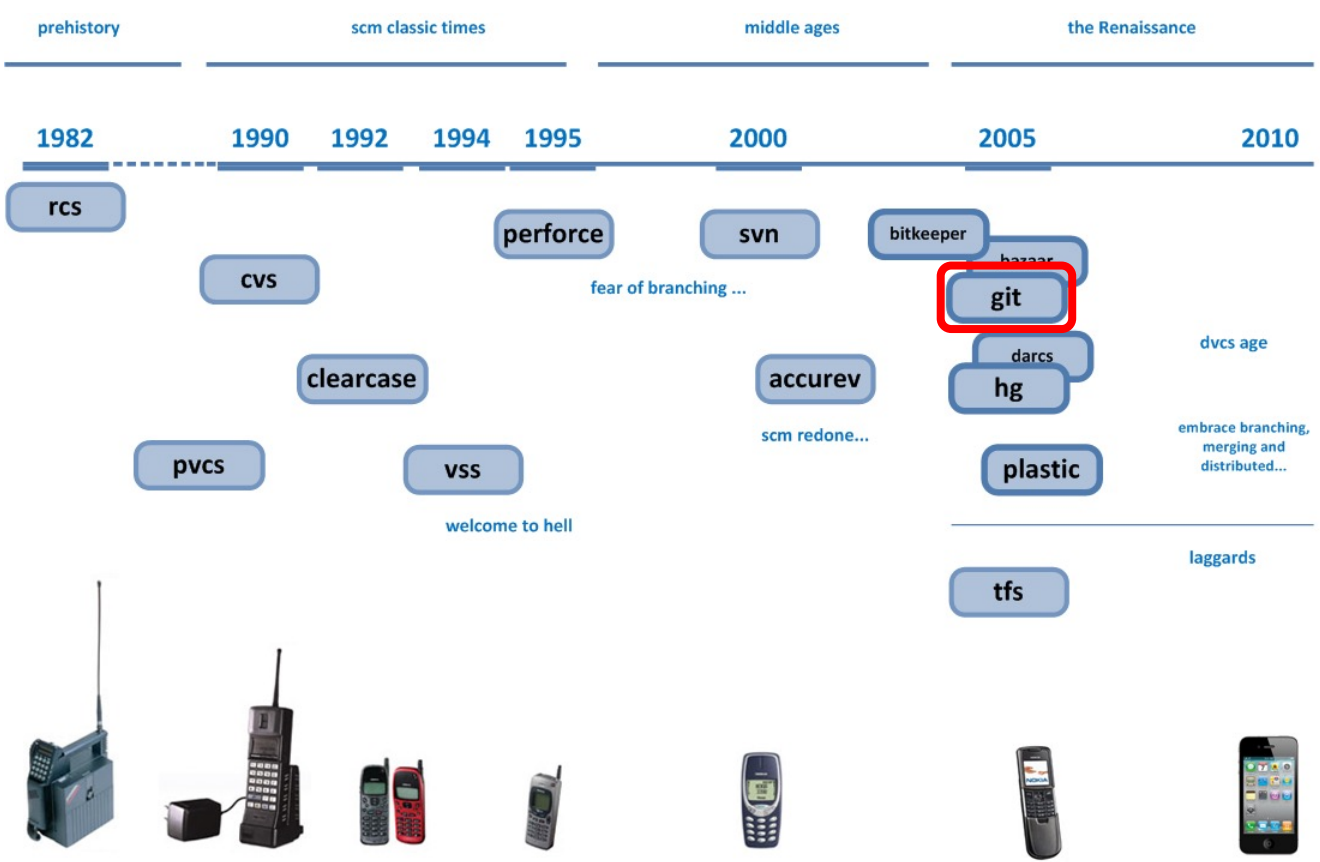
GitHub

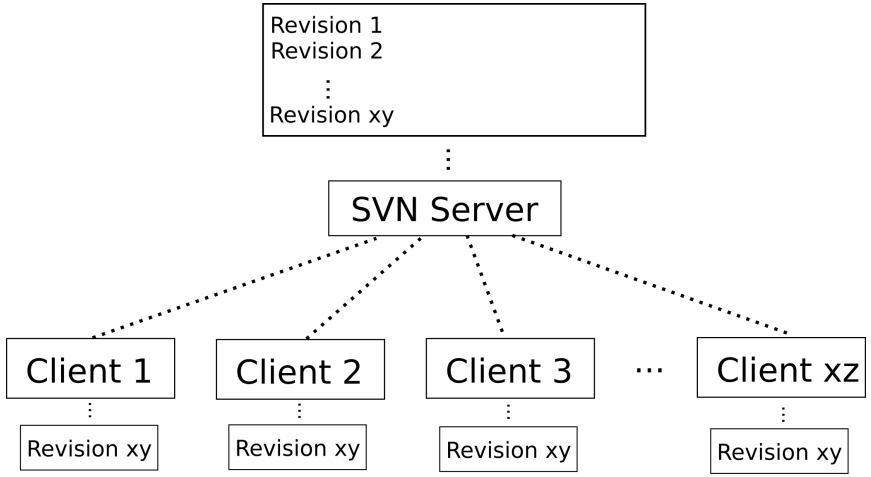
Source Code Management

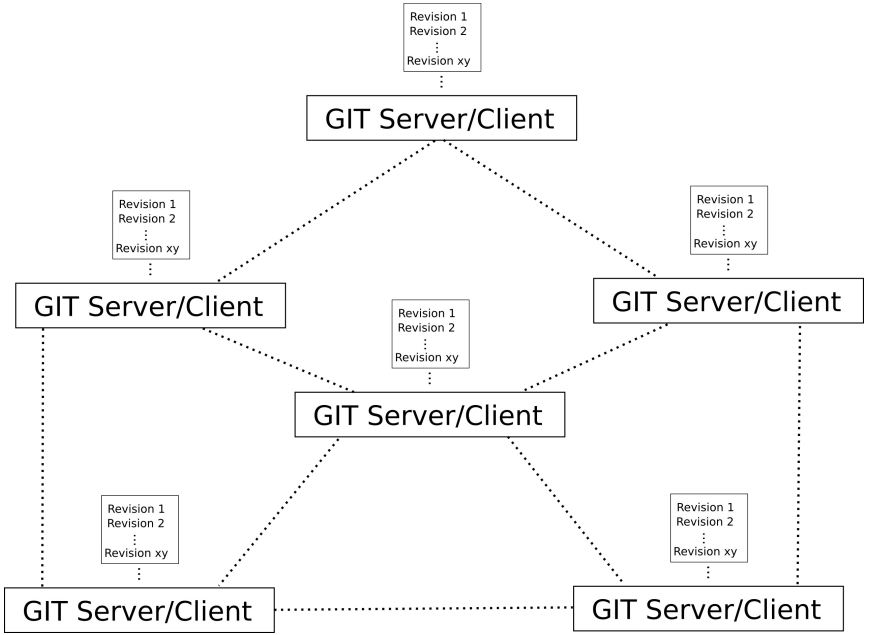
What/Why of Revision Control

- You change something and now it’s broken
- You forgot what you changed
- You start something and want to go back
- You want to know exactly what was changed
- You want to know who changed something
- You want to “undo” an old change
- You “accidentally” something ☺
- You want to work in parallel (different versions)
- You want to work as part of a team

SCM History


prehistory
scm classic times
middle ages
the Renaissance
1982
1990
1992
1994
1995
2000
2005
2010
rcs
cvs
perforce
svn
bitkeeper
bazaar
git
fear of branching ...
darcs
hg
dvcs age
clearcase
accurev
scm redone...
embrace branching, merging and distributed...
pvcs
vss
plastic
welcome to hell
laggards
tfs

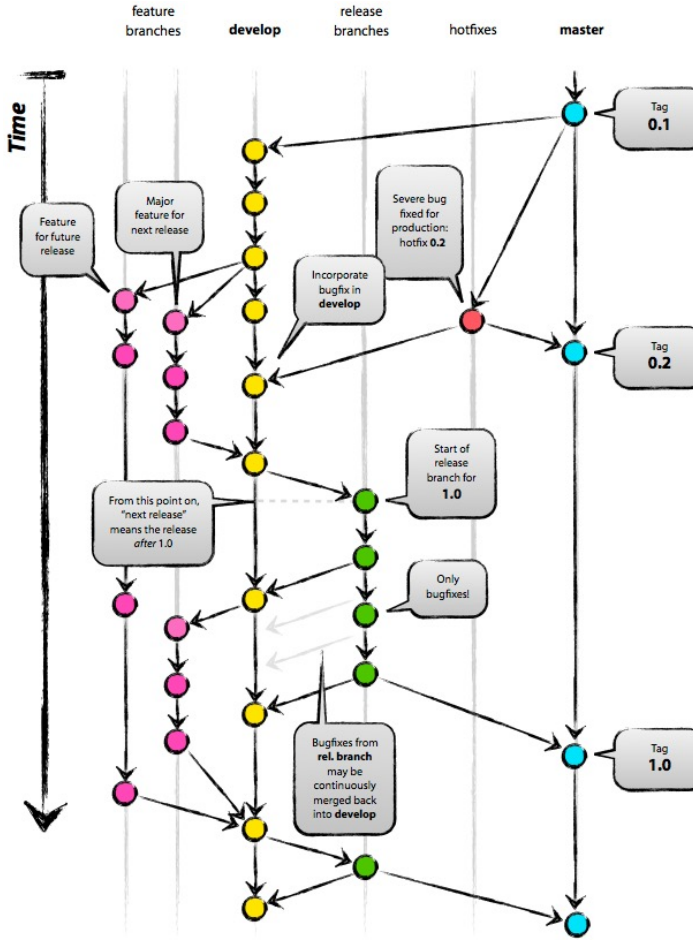




Git versus Subversion

Branches

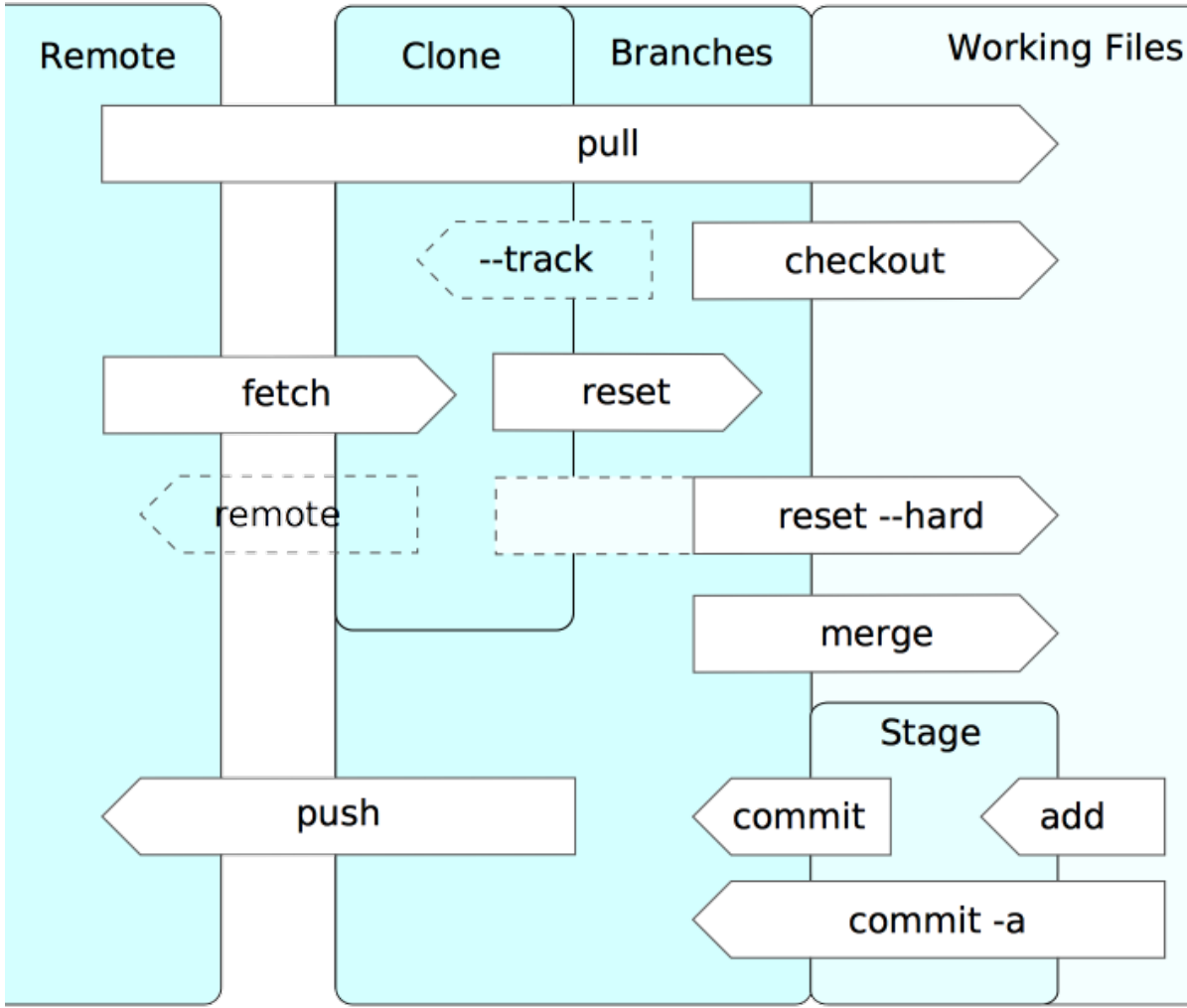
- git-flow (both an extension and a philosophy)
- “master” branch – always stable working code
- “develop” branch – work in progress (waiting release)
- “feature” branches – new features you are working on
- “release” branches – versions that end up in “master”



University of Zurich[UZH]

Basic Git Operations

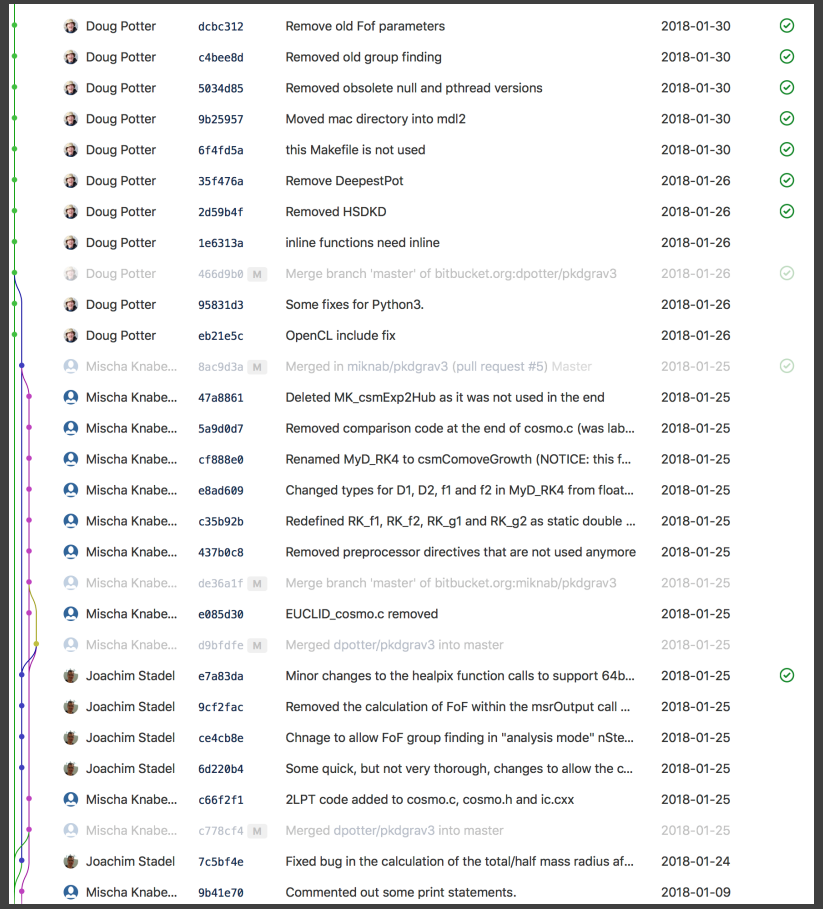
- “Working Files” are what you are probably used to.
- Clone/Branches are a local copy of the complete history.
- Remote
  - github.com
  - bitbucket.org
  - gitlab.uzh.ch
  - Or any other computer (or directory even!)



Almost Linear Development

Author	Commit	Message	Date	Builds
Doug Potter	4697bf1	Add openpa COPYRIGHT file	2018-02-14	✓
Joachim Stadel	1dfc65e	Went one step too far in the loop since D1[i+1] was set!	2018-02-13	✓
Doug Potter	1a4269e	MSC macros only under MSC	2018-02-13	✓
Doug Potter	58292c7	Now should compile under native Windows	2018-02-13	✓
Doug Potter	c222b31	More native windows compilation work (still in progress)	2018-02-13	✓
Doug Potter	191fb00	You cannot use a variable for size of local variables	2018-02-13	✓
Doug Potter	ab78c44	Why was mpi included here?	2018-02-13	✓
Doug Potter	71014b5 M	merge	2018-02-12	✓
Doug Potter	edff8a7	Some updates to support native windows	2018-02-12	
Doug Potter	dfe9661	Start of integrating base CUDA into MDL	2018-02-08	✓
Doug Potter	5d62d1c	Set sensible defaults when present	2018-02-07	✓
Doug Potter	0062ef1	Simple Windows instructions	2018-02-07	✓
Doug Potter	f1c8375	Add pkdgrav3 to the installation	2018-02-07	
Doug Potter	fc06662	Compilation on Windows	2018-02-07	✓
Doug Potter	c262c4f	A little to overzealous with the separation (for now)	2018-02-06	✓
Doug Potter	a6b5eeb	Separate generic CUDA routines better	2018-02-06	✓
Doug Potter	69fb03b	compilation fixes for Cray under cmake	2018-02-06	✓
Doug Potter	76acd63	Fix string literal problems	2018-02-06	
Doug Potter	4b8ce39	cmake should now work (perhaps everywhere)	2018-02-06	✓
Doug Potter	607ba6d	Preliminary cmake support (not yet completely ready)	2018-02-06	✓
Doug Potter	0fe6d50	Remove null/pthread build targets (there weren't used)	2018-02-05	✓
Doug Potter	63917ff	Features/macros cleanup	2018-02-05	✓
Doug Potter	1159819	Removed COOLING and Fortran support	2018-02-05	✓
Doug Potter	ab9106f	CHANGESOFT is not used	2018-02-05	✓
Doug Potter	82e60f0	Removed global group assignment code. This will be done diff...	2018-02-02	✓
Doug Potter	c60199c	Removed nLowTot/nHighTot from the pst; they were never used.	2018-02-02	✓
Doug Potter	f1d372f	Make argv processing optional by passing NULL	2018-02-01	✓
Doug Potter	a967fb1	mdl2 can now be compiled stand-alone with cmake	2018-01-31	✓

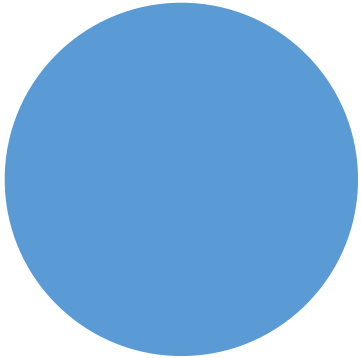
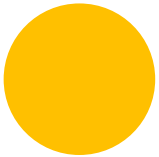
Parallel Development



Doug Potter	dcbc312	Remove old Fof parameters	2018-01-30
Doug Potter	c4bee8d	Removed old group finding	2018-01-30
Doug Potter	5034d85	Removed obsolete null and pthread versions	2018-01-30
Doug Potter	9b25957	Moved mac directory into mdl2	2018-01-30
Doug Potter	6f4fd5a	this Makefile is not used	2018-01-30
Doug Potter	35f476a	Remove DeepestPot	2018-01-26
Doug Potter	2d59b4f	Removed HSDKD	2018-01-26
Doug Potter	1e6313a	inline functions need inline	2018-01-26
Doug Potter	466d9b0 M	Merge branch 'master' of bitbucket.org:dpotter/pkdgrav3	2018-01-26
Doug Potter	95831d3	Some fixes for Python3.	2018-01-26
Doug Potter	eb21e5c	OpenCL include fix	2018-01-26
Mischa Knabe...	8ac9d3a M	Merged in miknab/pkdgrav3 (pull request #5) Master	2018-01-25
Mischa Knabe...	47a8861	Deleted MK_csmExp2Hub as it was not used in the end	2018-01-25
Mischa Knabe...	5a9d0d7	Removed comparison code at the end of cosmo.c (was lab...	2018-01-25
Mischa Knabe...	cf888e0	Renamed MyD_RK4 to csmComoveGrowth (NOTICE: this f...	2018-01-25
Mischa Knabe...	e8ad609	Changed types for D1, D2, f1 and f2 in MyD_RK4 from float...	2018-01-25
Mischa Knabe...	c35b92b	Redefined RK_f1, RK_f2, RK_g1 and RK_g2 as static double ...	2018-01-25
Mischa Knabe...	437b0c8	Removed preprocessor directives that are not used anymore	2018-01-25
Mischa Knabe...	de36a1f M	Merge branch 'master' of bitbucket.org:miknab/pkdgrav3	2018-01-25
Mischa Knabe...	e085d30	EUCLID_cosmo.c removed	2018-01-25
Mischa Knabe...	d9bfdfe M	Merged dpotter/pkdgrav3 into master	2018-01-25
Joachim Stadel	e7a83da	Minor changes to the healpix function calls to support 64b...	2018-01-25
Joachim Stadel	9cf2fac	Removed the calculation of FoF within the msrOutput call ...	2018-01-25
Joachim Stadel	ce4cb8e	Chnage to allow FoF group finding in "analysis mode" nSte...	2018-01-25
Joachim Stadel	6d220b4	Some quick, but not very thorough, changes to allow the c...	2018-01-25
Mischa Knabe...	c66f2f1	2LPT code added to cosmo.c, cosmo.h and ic.cxx	2018-01-25
Mischa Knabe...	c778cf4 M	Merged dpotter/pkdgrav3 into master	2018-01-25
Joachim Stadel	7c5bf4e	Fixed bug in the calculation of the total/half mass radius af...	2018-01-24
Mischa Knabe...	9b41e70	Commented out some print statements.	2018-01-09

GIT Example

```
dpotter@daint103:~> cat .gitconfig              (Set this first!)
[user]
        email = douglas.potter@uzh.ch
        name = Doug Potter
dpotter@daint103:~> cd cpi
dpotter@daint103:~/cpi> ls
cpi_mpi.c  cpi_openmp.c  Makefile
dpotter@daint103:~/cpi> git init                (Done only once)
Initialized empty Git repository in /users/dpotter/cpi/.git/
dpotter@daint103:~/cpi> git status          (Output abbreviated)
Untracked files:
        Makefile
        cpi_mpi.c
        cpi_openmp.c
dpotter@daint103:~/cpi> git add Makefile cpi_mpi.c cpi_openmp.c
dpotter@daint103:~/cpi> git commit -m "Initial import"
[master (root-commit) b0f821d] Initial import
 3 files changed, 145 insertions(+)
 create mode 100644 Makefile
 create mode 100644 cpi_mpi.c
 create mode 100644 cpi_openmp.c
```

Compilers | C & Fortran

High Performance Computing

“Hello World” in C

```c
#include <stdio.h>

int main(int argc,char *argv[]) {
    printf("Hello world\n");
    return 0;
    }
```

“Hello World” in “Machine Code”

```
.data
S0:     .text “Hello world\n”

.code
main:   mova S0,r0
        call printf
        movi #0,r0
        ret
```

```c
int main(int argc,char *argv[]) {
    printf("Hello world\n");
    return 0;
    }
```

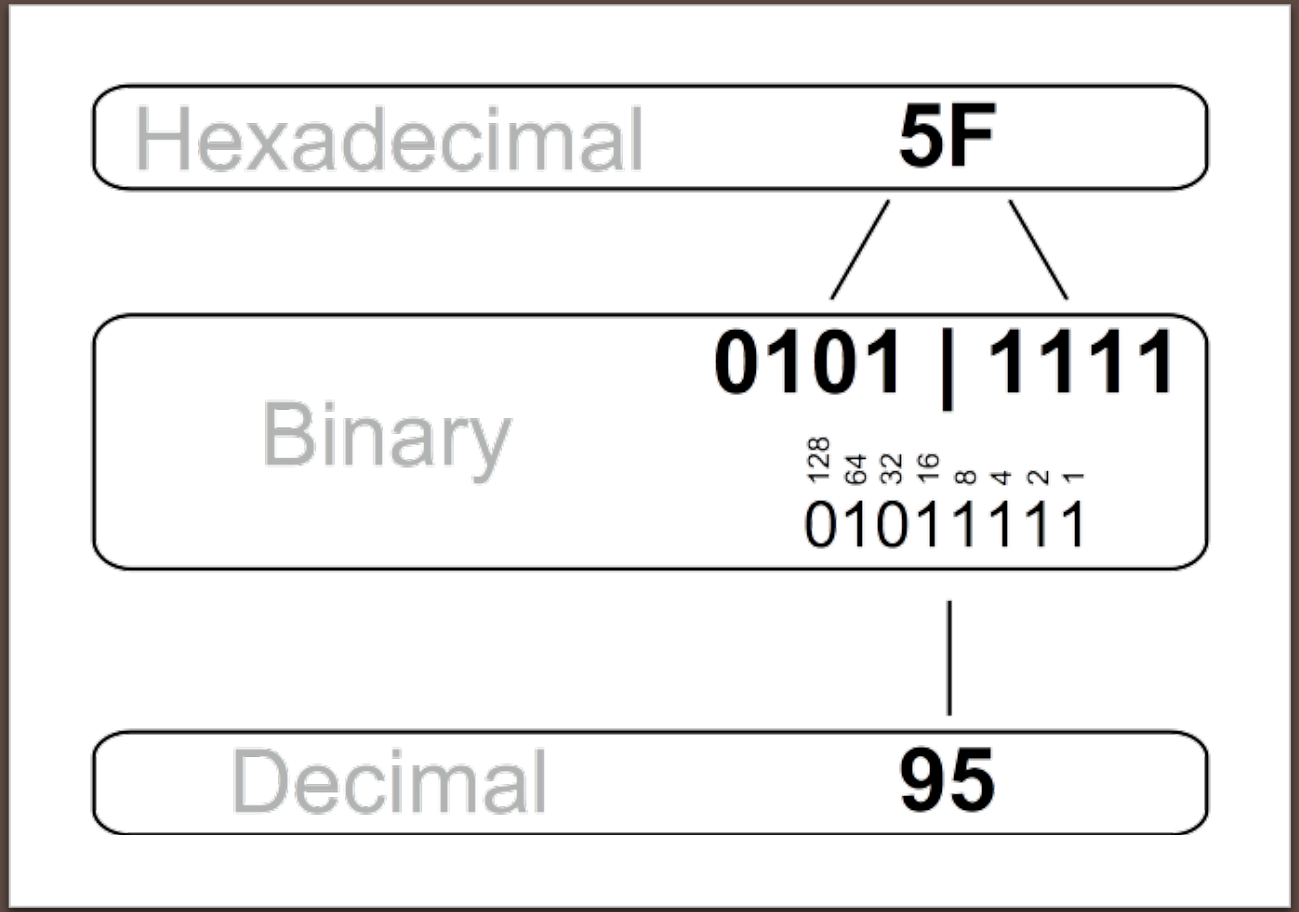
Real Assembler Output

```
        .section    __TEXT,__text,regular,pure_instructions
        .macosx_version_min 10, 13
        .globl      _main
        .p2align    4, 0x90
_main:                                  ## @main
        .cfi_startproc
## BB#0:
        pushq       %rbp
Lcfi0:
        .cfi_def_cfa_offset 16
Lcfi1:
        .cfi_offset %rbp, -16
        movq        %rsp, %rbp
Lcfi2:
        .cfi_def_cfa_register %rbp
        leaq        L_str(%rip), %rdi
        callq       _puts
        xorl        %eax, %eax
        popq        %rbp
        retq
        .cfi_endproc

        .section    __TEXT,__cstring,cstring_literals
L_str:                                  ## @str
        .asciz      "Hello world"


.subsections_via_symbols
```

```c
int main(int argc,char *argv[]) {
    printf("Hello world\n");
    return 0;
    }
```

“man puts”

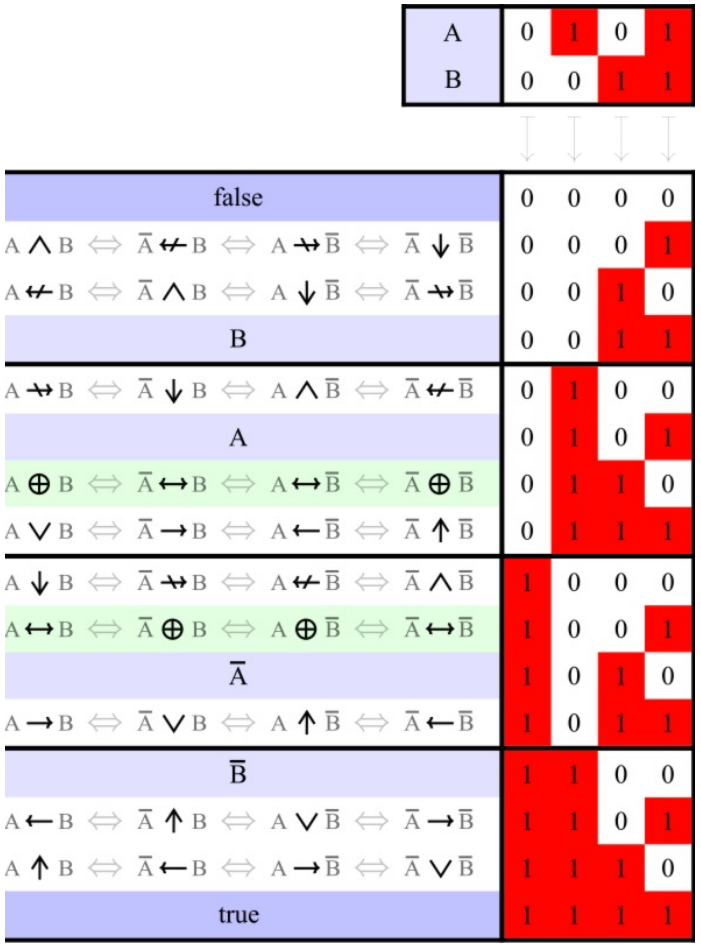
```
PUTS(3)       BSD Library Functions Manual       PUTS(3)

NAME
     puts -- output a line to a stream

LIBRARY
     Standard C Library (libc, -lc)

SYNOPSIS
     #include <stdio.h>
     int     puts(const char *s);

DESCRIPTION
     The function puts() writes the string s, and a
     terminating newline character to the stream stdout.
```


Hexadecimal
5F
Binary
0101 | 1111
128 64 32 16 8 4 2 1
01011111
Decimal
95

Binary & Hex

Boolean Logic

```
leaq L_str(%rip), %rdi
callq    _puts
xorl %eax, %eax
popq %rbp
retq
```



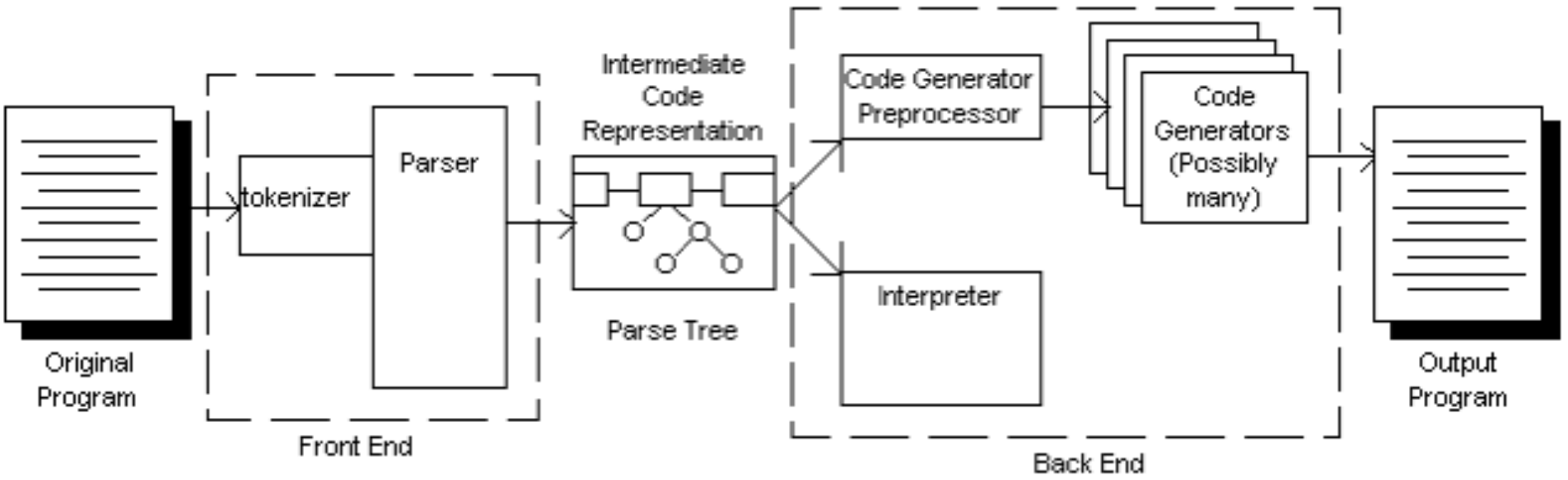
Simple Arithmetic Function

```
float f(float a,float b,
        float c,float d,
        float e){
  return (a-b)*c + d/e;
  }
```

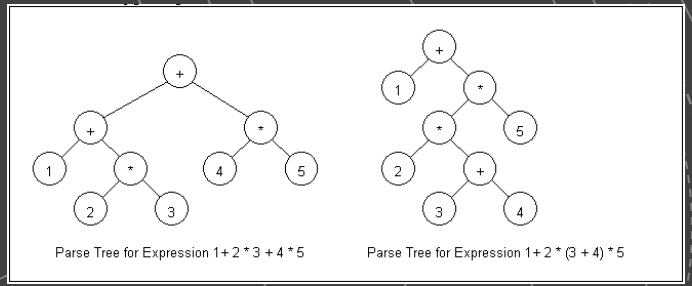
```
Disassembly of section .text:
   0:  55             push   %rbp
   1:  48 89 e5       mov    %rsp,%rbp
   4:  f3 0f 5c c1    subss  %xmm1,%xmm0
   8:  f3 0f 59 c2    mulss  %xmm2,%xmm0
   c:  f3 0f 5e dc    divss  %xmm4,%xmm3
  10:  f3 0f 58 c3    addss  %xmm3,%xmm0
  14:  5d             pop    %rbp
  15:  c3             retq
```

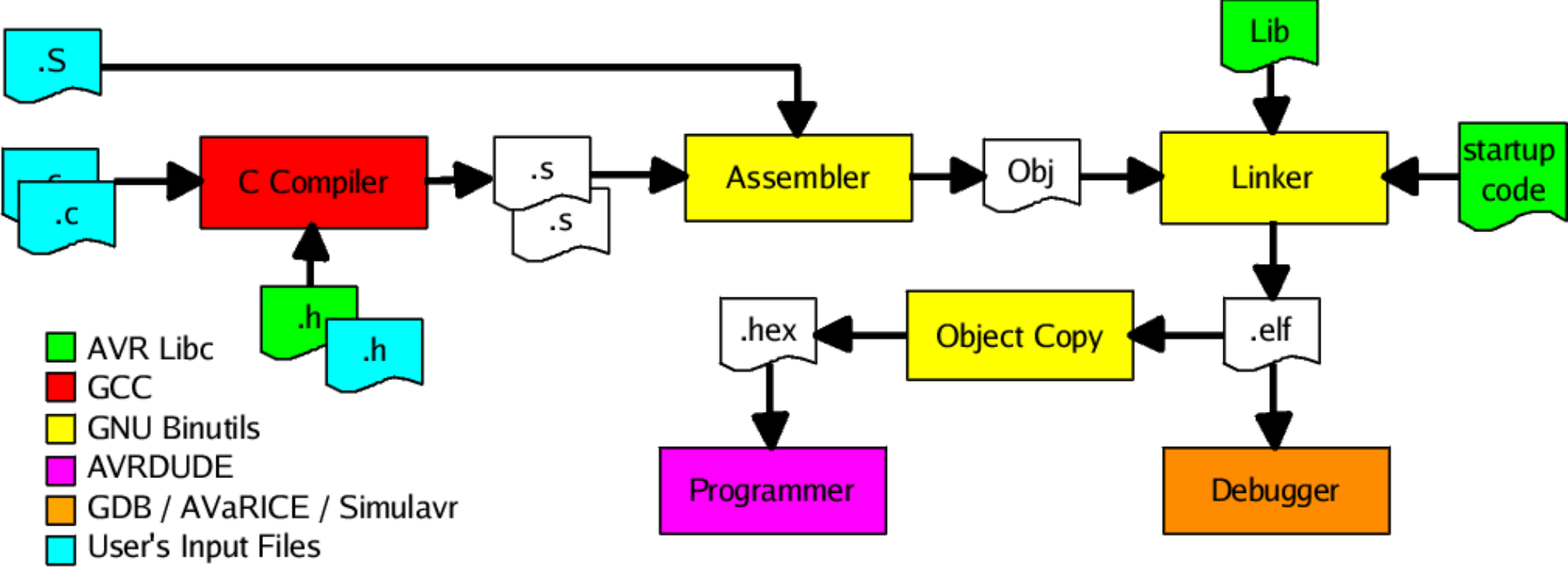
```
_f:
  pushq %rbp
  movq  %rsp, %rbp
  subss %xmm1, %xmm0
  mulss %xmm2, %xmm0
  divss %xmm4, %xmm3
  addss %xmm3, %xmm0
  popq  %rbp
  retq
```

gcc -O3 -S -c -o test.s test.c



Anatomy of a Compiler





Compilation Process

“Modules” on Piz Daint

```
dpotter@ela3:~> ssh daint
Last login: Thu Feb 22 08:21:39 2018 from 148.187.1.7
================================================================
      IMPORTANT REMINDER FOR USERS of CSCS facilities

      help@cscs.ch - +41 91 610 82 10 - http://user.cscs.ch
================================================================

Please load 'daint-gpu' module for using the GPU/Haswell nodes
or
load 'daint-mc' module for the Multicore/Broadwell nodes

For more info, please refer to the User Portal:
http://user.cscs.ch/getting_started/running_jobs/piz_daint

dpotter@daint105:~> module load daint-mc
dpotter@daint105:~> module swap PrgEnv-cray PrgEnv-gnu
```

Compiler Suites

Cray Compiler (PrgEnv-cray)	Available on Cray Computers
GNU Compiler (PrgEnv-gnu)	Free for everyone
Intel Compiler (PrgEnv-intel)	Now free for everyone
Portland (PrgEnv-pgi)	Important for OpenACC (GPU)
Clang (Apple & Cray)	Open Source & GNU Compatible
Visual Studio (Windows)	Not usually High Performance

University of Zurich[UZH]

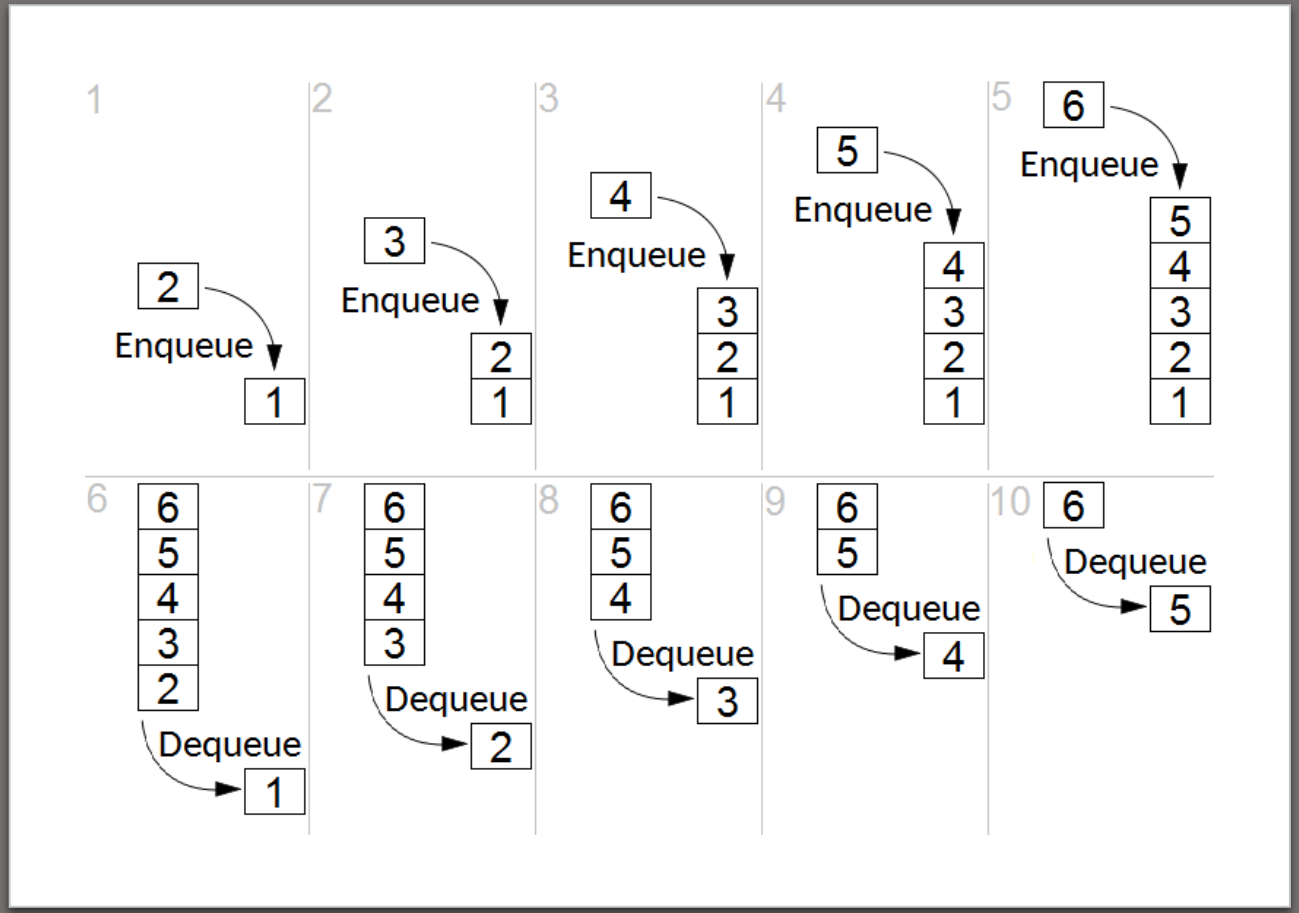
List Loaded Modules

```
dpotter@daint105:~> module list
Currently Loaded Modulefiles:
  1) modules/3.2.10.6                                 13) ugni/6.0.14-6.0.4.0_14.1__ge7db4a2.ari
  2) eproxy/2.0.16-6.0.4.1_3.1__g001b199.ari          14) pmi/5.0.12
  3) gcc/5.3.0                                        15) dmapp/7.1.1-6.0.4.0_46.2__gb8abda2.ari
  4) craype-broadwell                                 16) gni-headers/5.0.11-6.0.4.0_7.2__g7136988.ari
  5) craype-network-aries                             17) xpmem/2.2.2-6.0.4.0_3.1__g43b0535.ari
  6) craype/2.5.12                                    18) job/2.2.2-6.0.4.0_8.2__g3c644b5.ari
  7) cray-mpich/7.6.0                                 19) dvs/2.7_2.2.32-6.0.4.1_7.1__ged1923a
  8) slurm/17.02.9+git20180119.b04278-1               20) alps/6.4.1-6.0.4.0_7.2__g86d0f3d.ari
  9) xalt/daint-2016.11                               21) rca/2.2.11-6.0.4.0_13.2__g84de67a.ari
 10) daint-mc                                         22) atp/2.1.1
 11) cray-libsci/17.06.1                              23) perftools-base/6.5.1
 12) udreg/2.3.2-6.0.4.0_12.2__g2f9c3ee.ari           24) PrgEnv-gnu/6.0.4
dpotter@daint105:~> module avail
---------------------------- /opt/cray/pe/perftools/6.5.1/modulefiles -----------------------------
perftools             perftools-lite-events perftools-lite-hbm    perftools-nwpc
perftools-lite        perftools-lite-gpu    perftools-lite-loops

-------------------- /apps/daint/UES/jenkins/6.0.UP04/mc/easybuild/modules/all ---------------------
Amber/16-CrayGNU-17.08-parallel(default)            PLUMED/2.3.0-CrayGNU-17.08
Amber/16-CrayGNU-17.08-serial                       Paraver/4.6.4.rc1(default)
Boost/1.65.0-CrayGNU-17.08                          Perl/5.26.1-bare(default)
Boost/1.65.0-CrayGNU-17.08-python2                  PyExtensions/2.7-CrayGNU-17.08
Boost/1.65.0-CrayGNU-17.08-python3(default)         PyExtensions/3.5-CrayGNU-17.08(default)
… and many more

dpotter@daint105:~> module avail gcc
------------------------------ /opt/modulefiles -------------------------------
gcc/4.9.3          gcc/6.1.0          gcc/7.1.0
gcc/5.3.0(default) gcc/6.2.0
```

List Loaded Modules (Eiger)

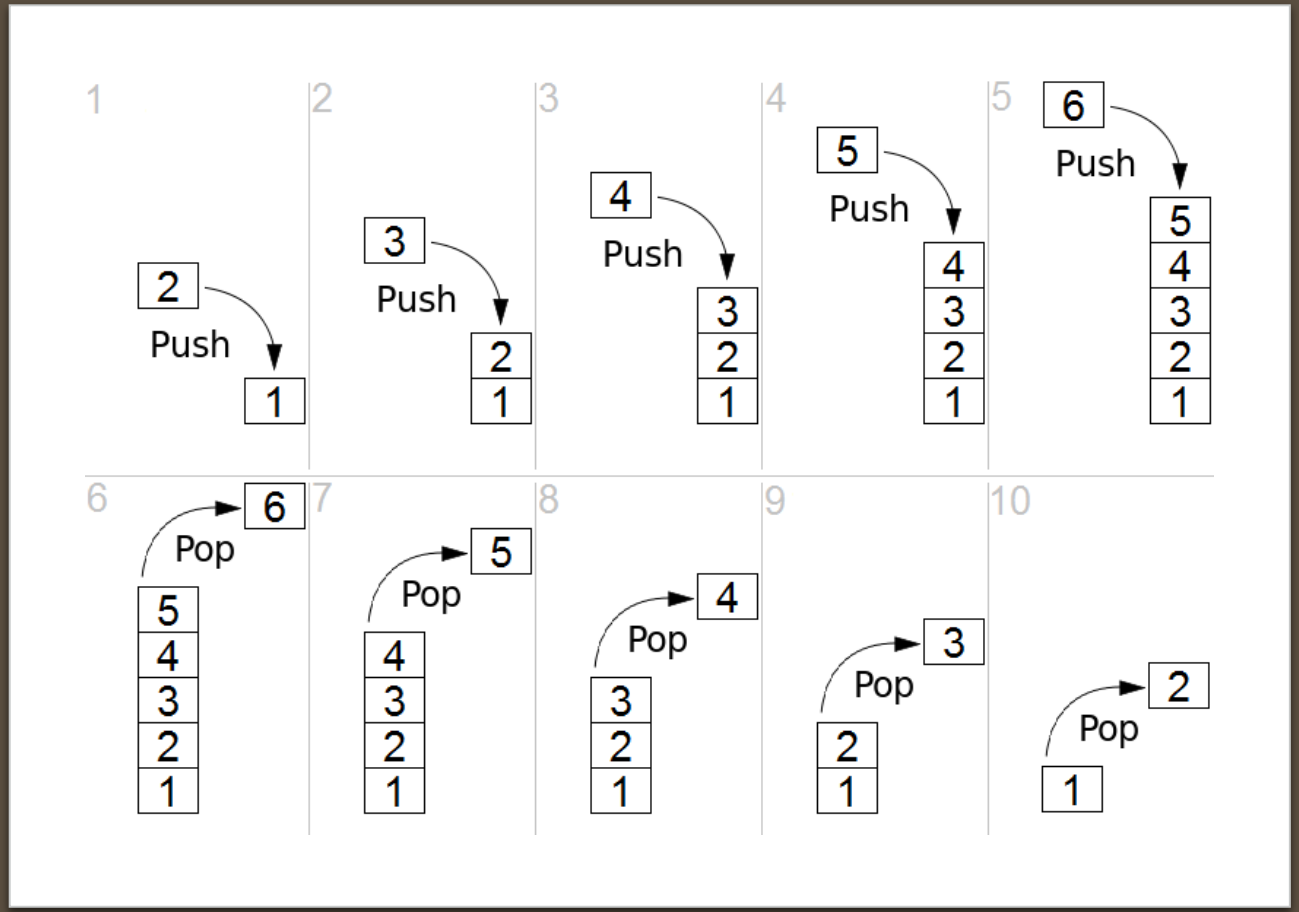
```
[eiger][dpotter@nid001765 ~]$ module list
Currently Loaded Modules:
  1) craype-x86-rome      4) perftools-base/21.12.0                7) craype/2.7.13      10) cray-libsci/21.08.1.2
  2) libfabric/1.15.0.0   5) xpmem/2.4.4-2.3_9.1__gff0e1d9.shasta  8) cray-dsmml/0.2.2   11) PrgEnv-cray/8.3.0
  3) craype-network-ofi   6) cce/13.0.0                            9) cray-mpich/8.1.12

[eiger][dpotter@nid001080 ~]$ cc --version
Cray clang version 13.0.0  (24b043d62639ddb4320c86db0b131600fdbc6ec6)
Target: x86_64-unknown-linux-gnu
Thread model: posix
InstalledDir: /opt/cray/pe/cce/13.0.0/cce-clang/x86_64/share/../bin

[eiger][dpotter@nid001765 ~]$ module swap PrgEnv-cray PrgEnv-gnu
Due to MODULEPATH changes, the following have been reloaded:
  1) cray-mpich/8.1.12

[eiger][dpotter@nid001765 ~]$ module list
Currently Loaded Modules:
  1) craype-x86-rome      4) perftools-base/21.12.0                7) craype/2.7.13      10) cray-libsci/21.08.1.2
  2) libfabric/1.15.0.0   5) xpmem/2.4.4-2.3_9.1__gff0e1d9.shasta  8) cray-dsmml/0.2.2   11) PrgEnv-gnu/8.3.0
  3) craype-network-ofi   6) gcc/11.2.0                            9) cray-mpich/8.1.12

[eiger][dpotter@nid001080 ~]$ cc --version
gcc (GCC) 11.2.0 20210728 (Cray Inc.)
Copyright (C) 2021 Free Software Foundation, Inc.
This is free software; see the source for copying conditions.  There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
```

Batch Queue Systems



"Queue"

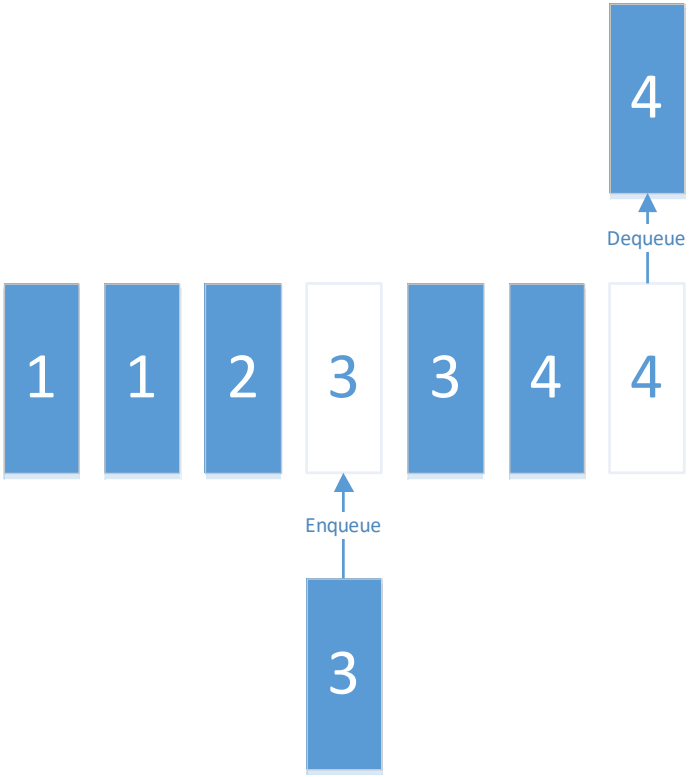
First In, First Out (FIFO)



"Stack"

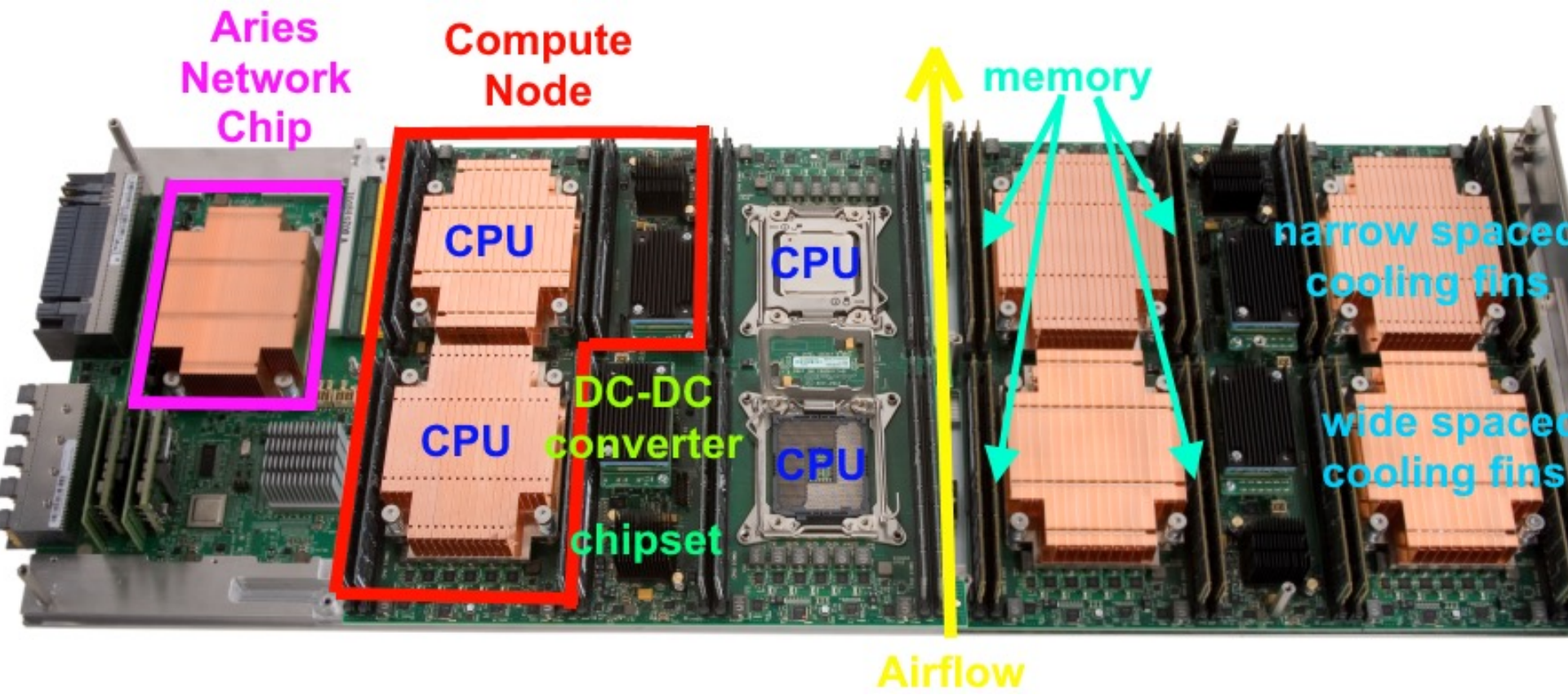
Last In, First Out (LIFO)

Priority Queue



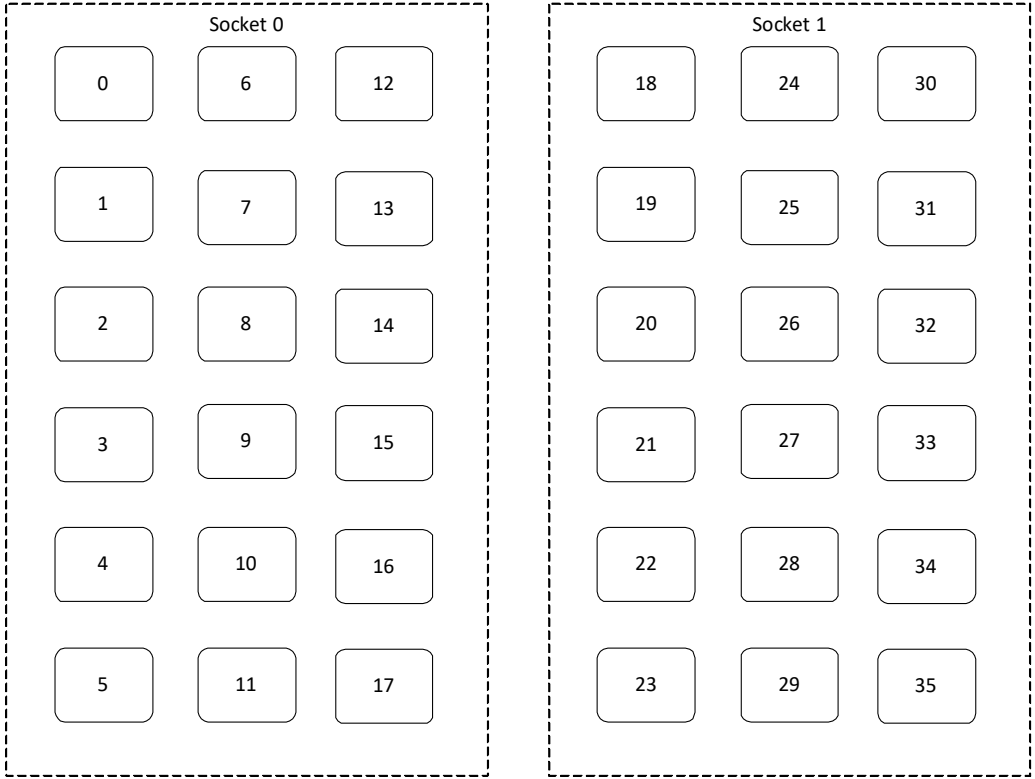
Generate Job Script

```
#!/bin/bash -l
#SBATCH --account=uzh8
#SBATCH --job-name=hpc_test
#SBATCH --time=01:00:00
#SBATCH --nodes=1
#SBATCH --ntasks-per-core=1
#SBATCH --ntasks-per-node=36
#SBATCH --cpus-per-task=1
#SBATCH --partition=normal
#SBATCH --constraint=mc

export OMP_NUM_THREADS=$SLURM_CPUS_PER_TASK

srun hostname
```

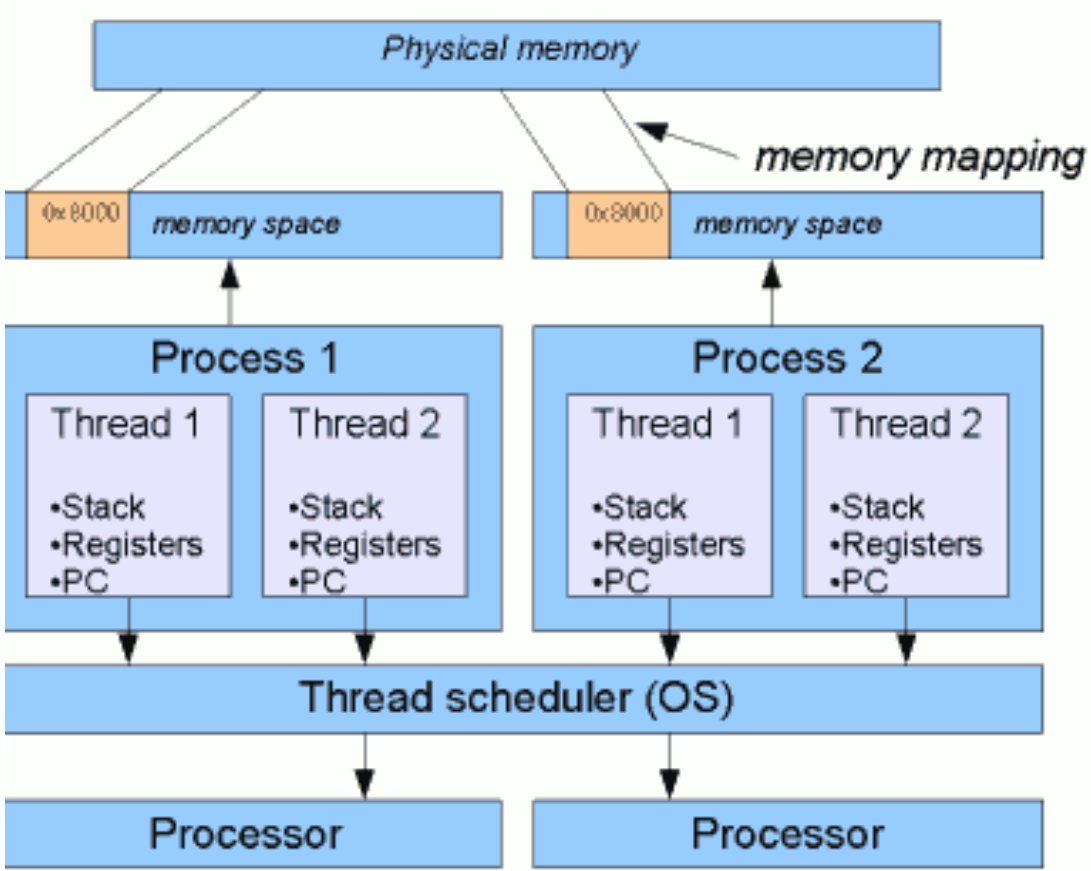


Cray XC40 Blade (daint-mc)

Processor Cores Layout

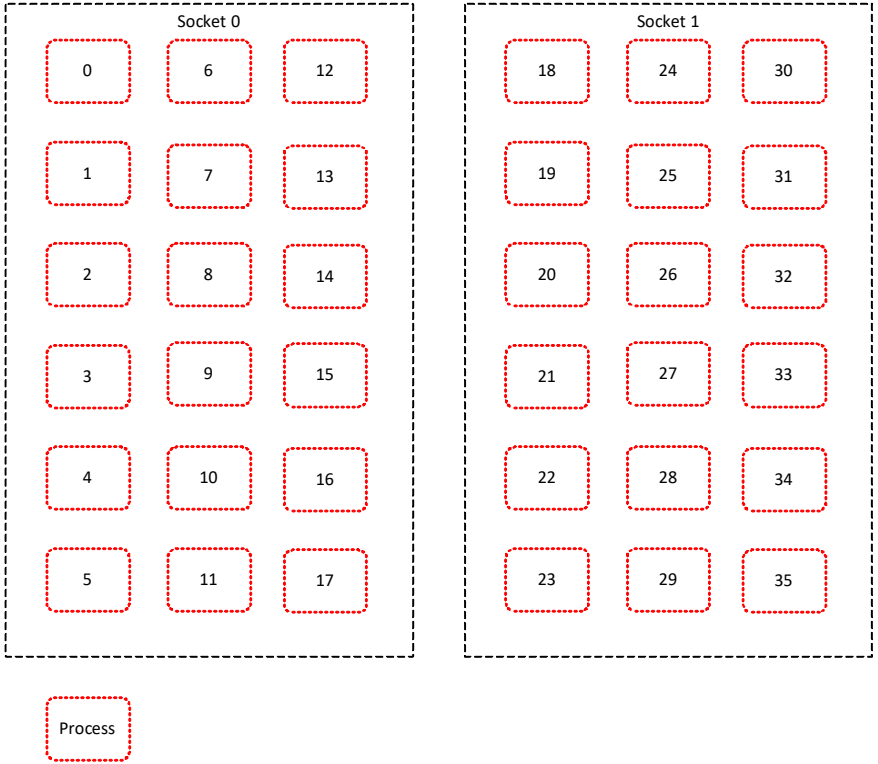


Processes & Threads



MPI Layout

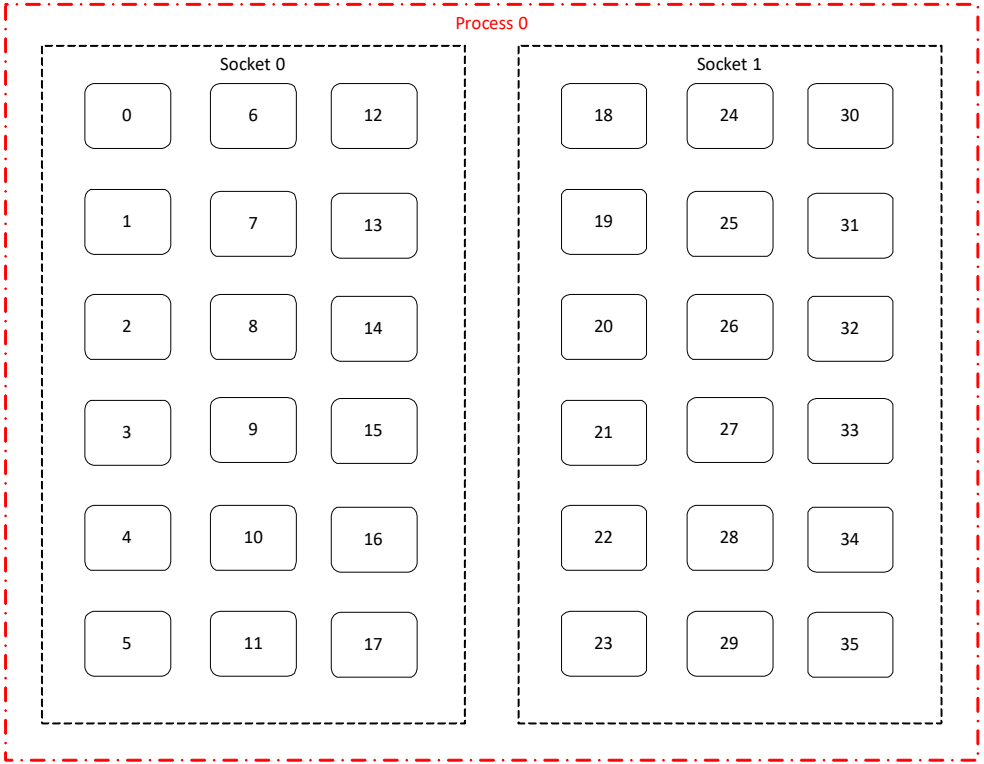
- --ntasks-per-core=1
- --ntasks-per-node=36
- --cpus-per-task=1



University of Zurich[UZH]

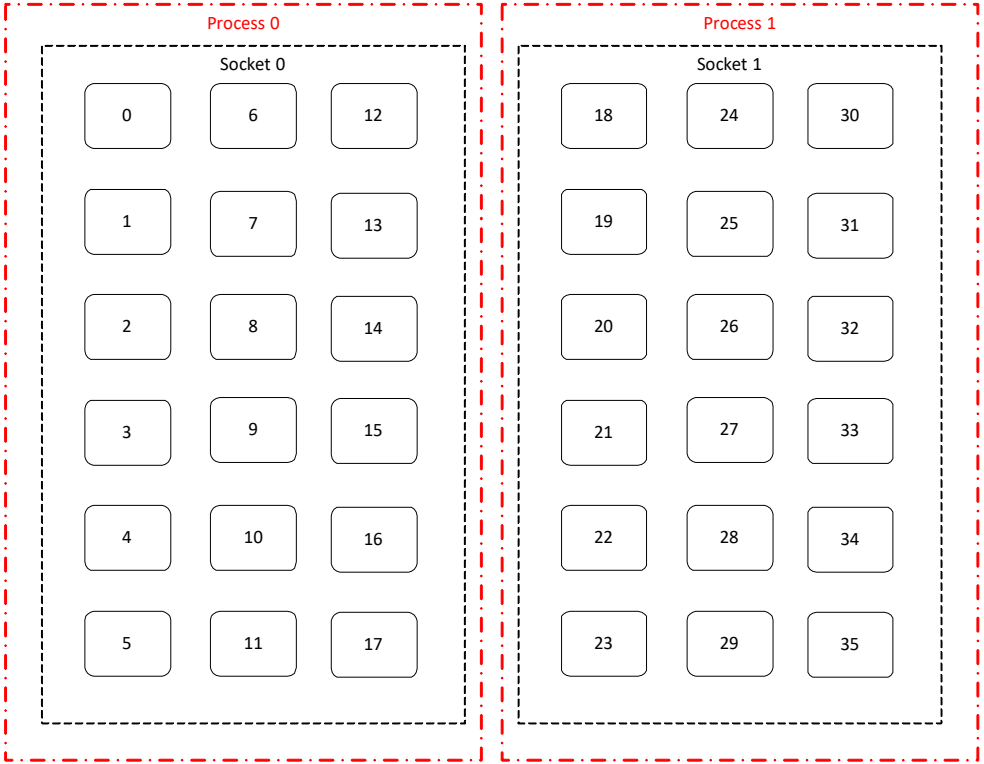
OpenMP Threads Layout

- --ntasks-per-core=1
- --ntasks-per-node=1
- --cpus-per-task=36



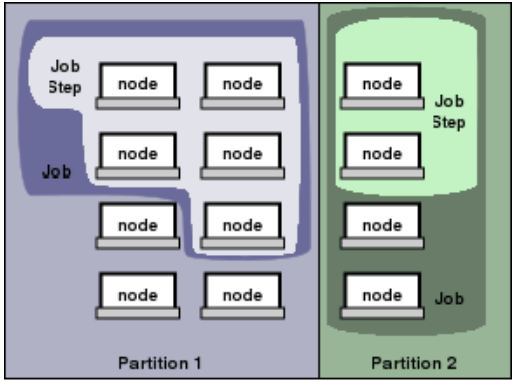
Hybrid Layout

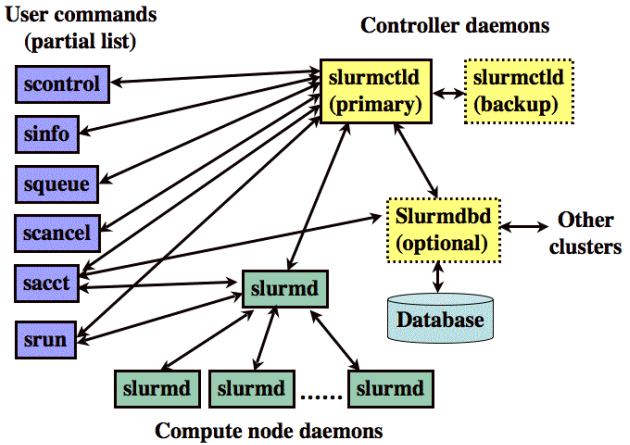
- --ntasks-per-core=1
- --ntasks-per-node=2
- --cpus-per-task=18

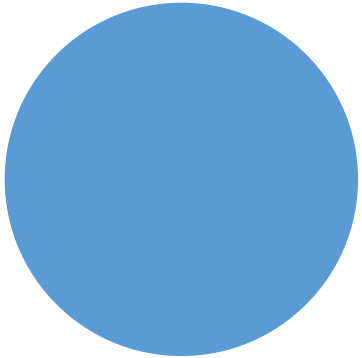
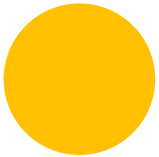


SLURM Job Submission

- ”sbatch” to submit a job
- “squeue” to see jobs
  - Pending jobs: -t pending
  - Running jobs: -t running
  - Your jobs: -u username
- “sinfo” to see partition information
  - Normal partition: -p normal
  - Debug partition: -p debug
- “srun” to run a job step
  - Always use “srun” in scripts
  - You can use it for interactive sessions
    - Not recommended if avoidable
- “scancel” to cancel a job
- “accounting” for usage
  - CHF 0.376 per “node hour”



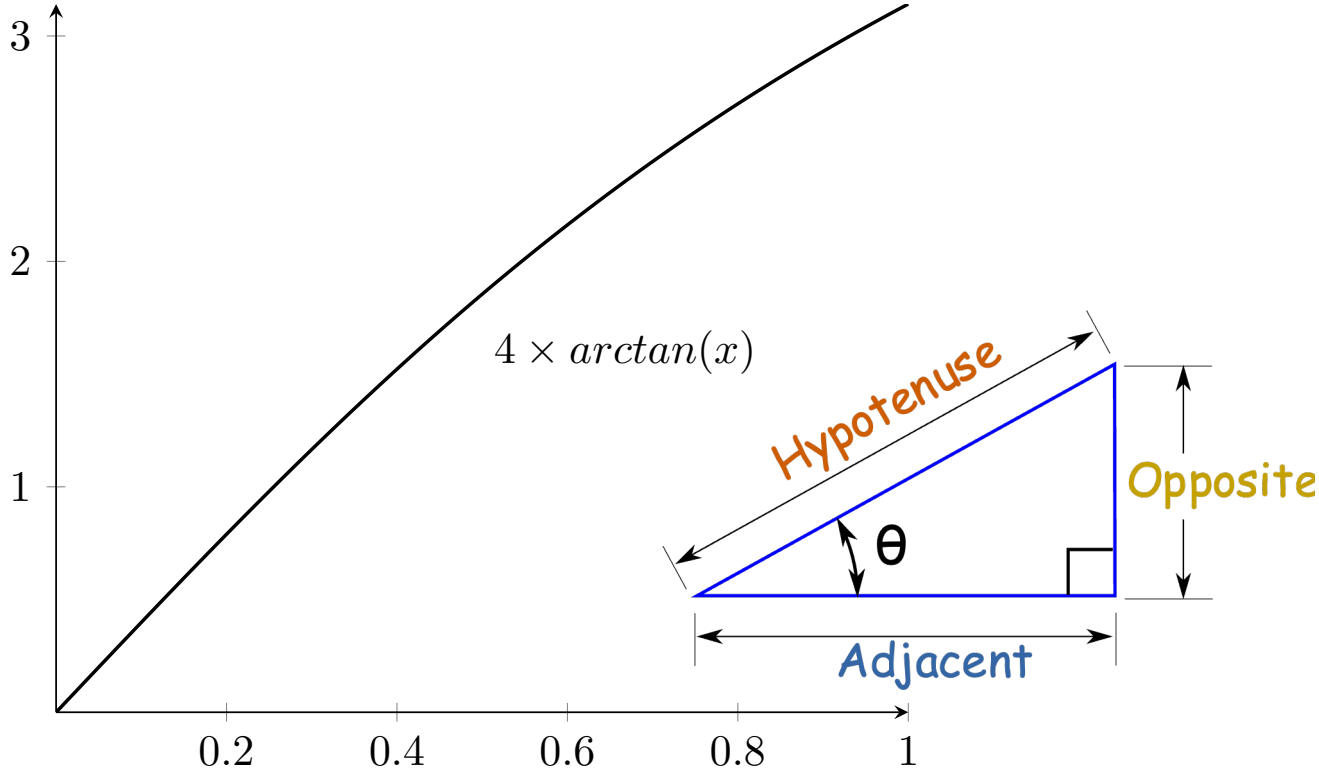


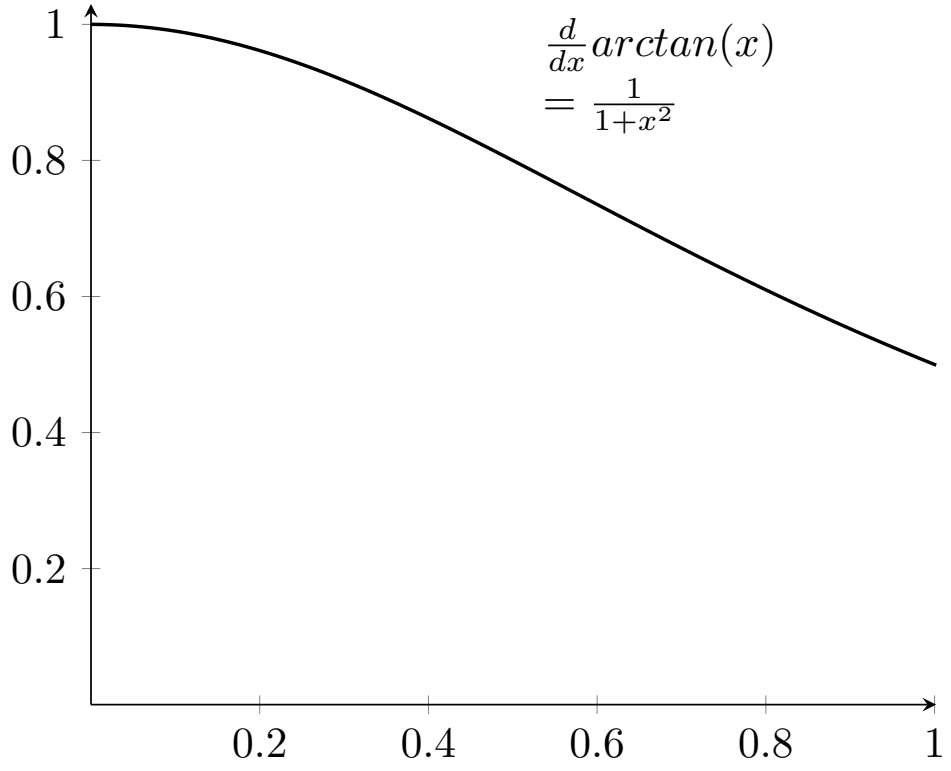
Sample Program

Calculate pi

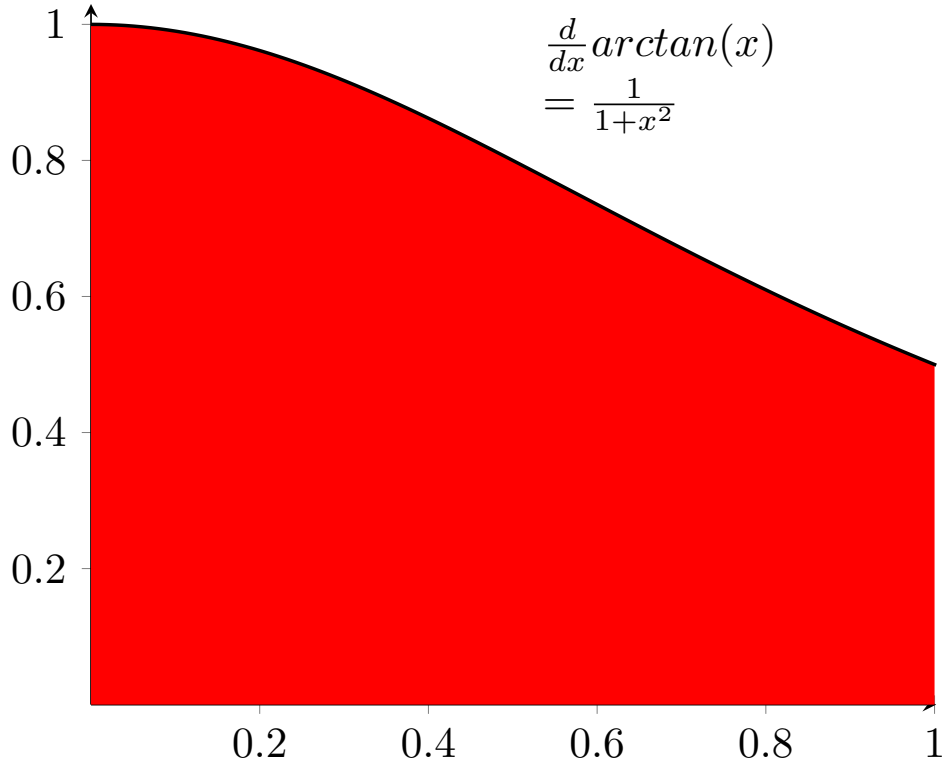
Calculate pi using arctan



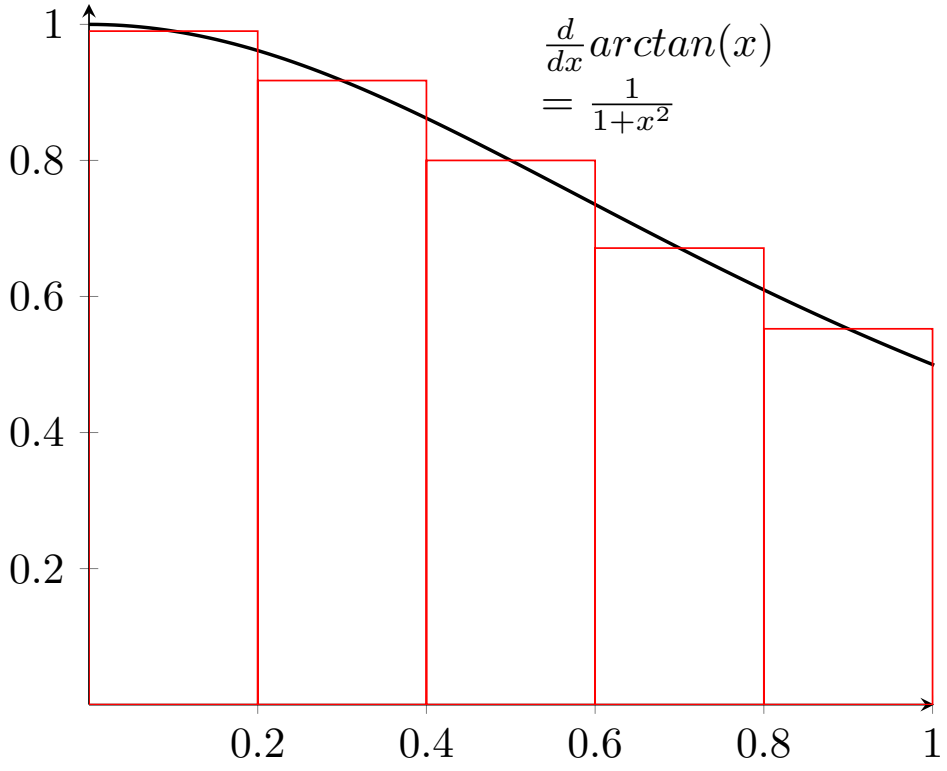
Derivative is well known



Integrate to get $\frac{\pi}{4}$

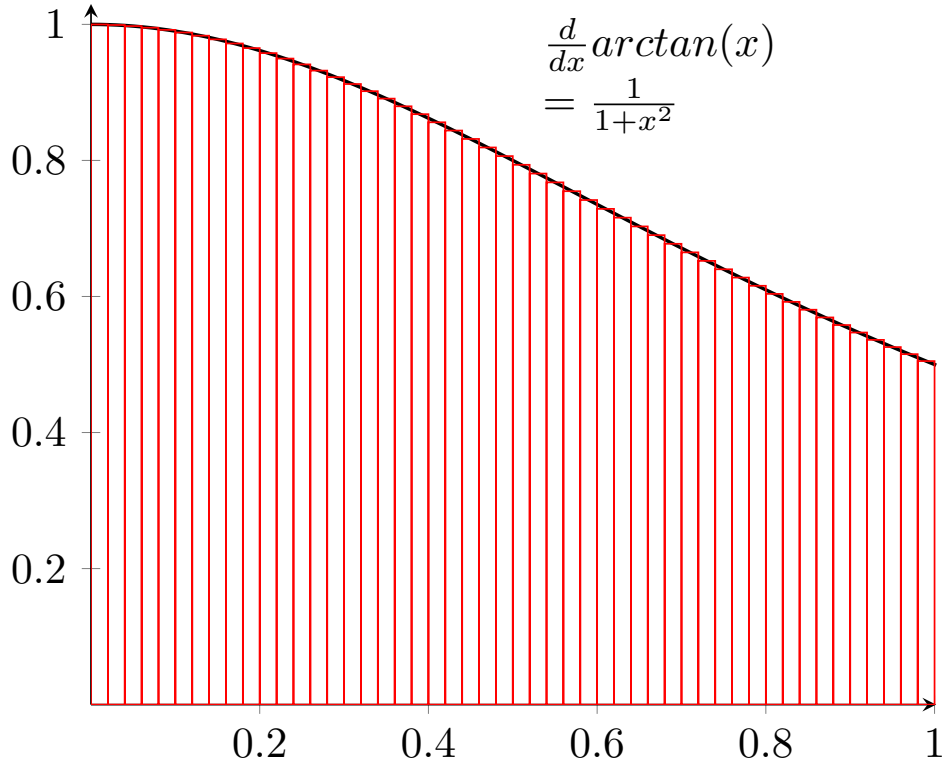


Integrate numerically



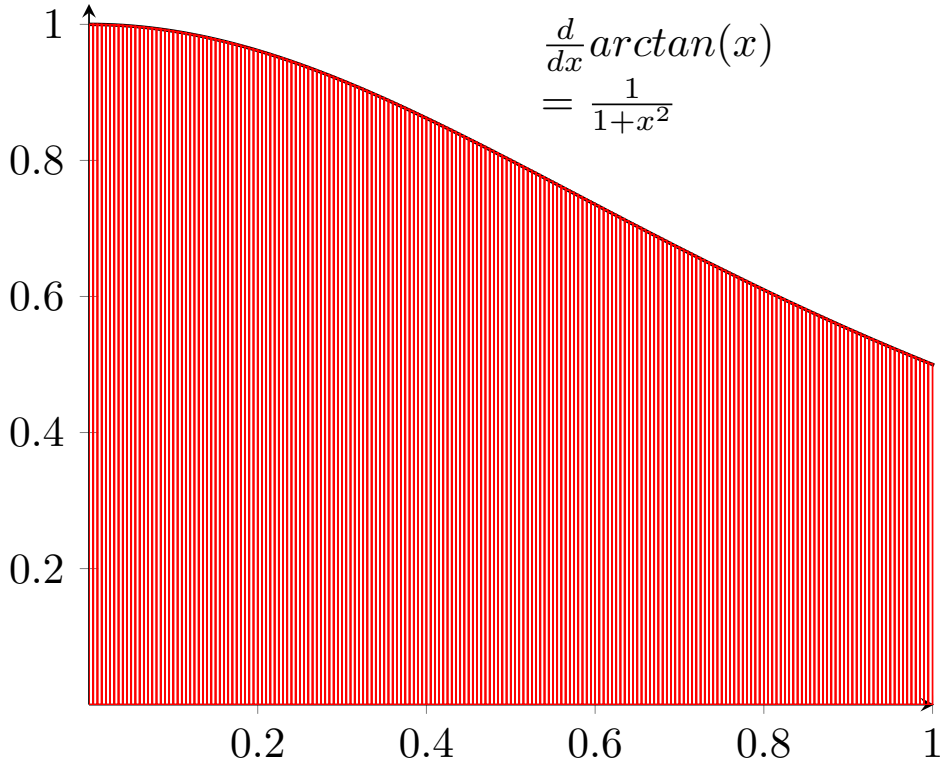
= 3.144925

Improve Accuracy



= 3.141625

We haven't even started!



= 3.141595

Integration in Python

```
N=5
dx=1.0/N
s = 0
for x in range(0, N):
  s += dx * 1 / ( 1 + ((x+0.5)*dx)**2 )
print(s*4)

dhcp-94-191:cpi$ python integrate.py
3.1449258640033277
```

Timing.

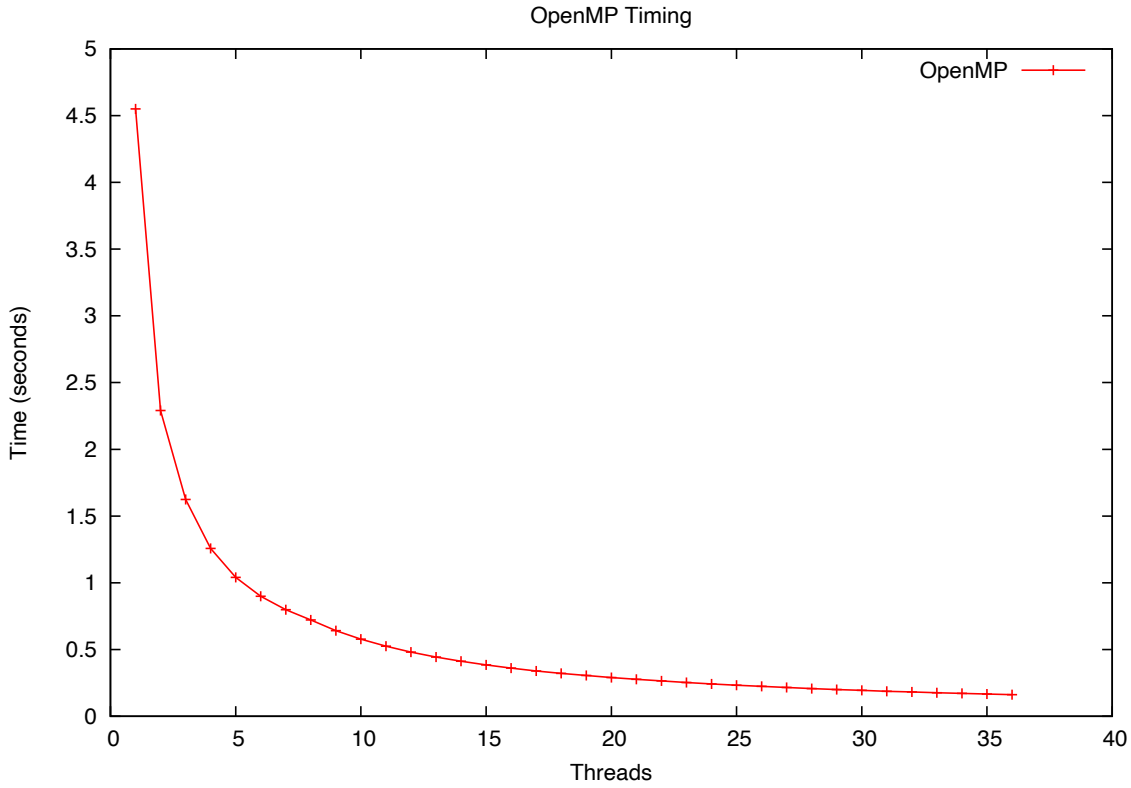
Is this good?

Threads	Seconds
1	4.550
2	2.291
3	1.624
4	1.258
5	1.041
6	0.8988
7	0.7989
8	0.7217
9	0.6415
10	0.5774
11	0.5249
12	0.4811
13	0.4441
14	0.4124
15	0.3849
16	0.3609
17	0.3397
18	0.3208

Threads	Seconds
19	0.3053
20	0.2897
21	0.2768
22	0.2644
23	0.2528
24	0.2423
25	0.2326
26	0.2237
27	0.2154
28	0.2077
29	0.2006
30	0.1939
31	0.1876
32	0.1818
33	0.1763
34	0.1711
35	0.1662
36	0.1616

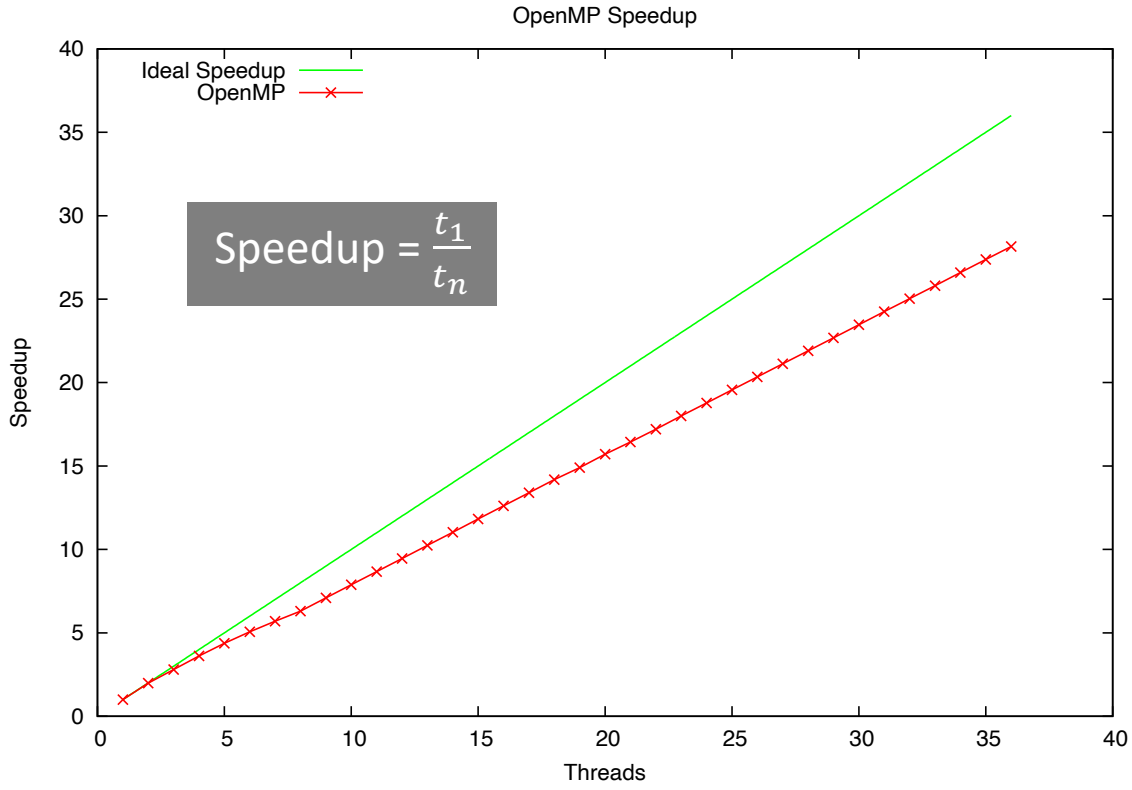
Timing Plot

```
set title 'OpenMP Timing'
set xlabel 'Threads’
det ylabel 'Time (seconds)'
set key top right
plot "cpi_openmp.dat" u 1:2 w lp lw 2 t 'OpenMP'
```

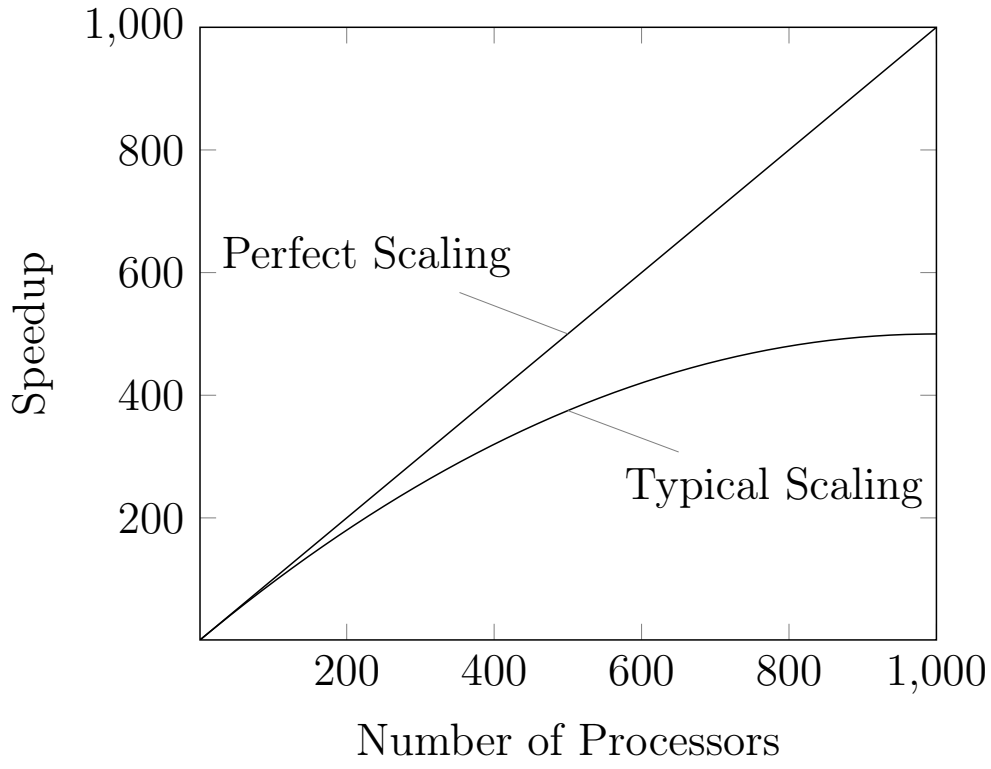


“Speedup” Plot

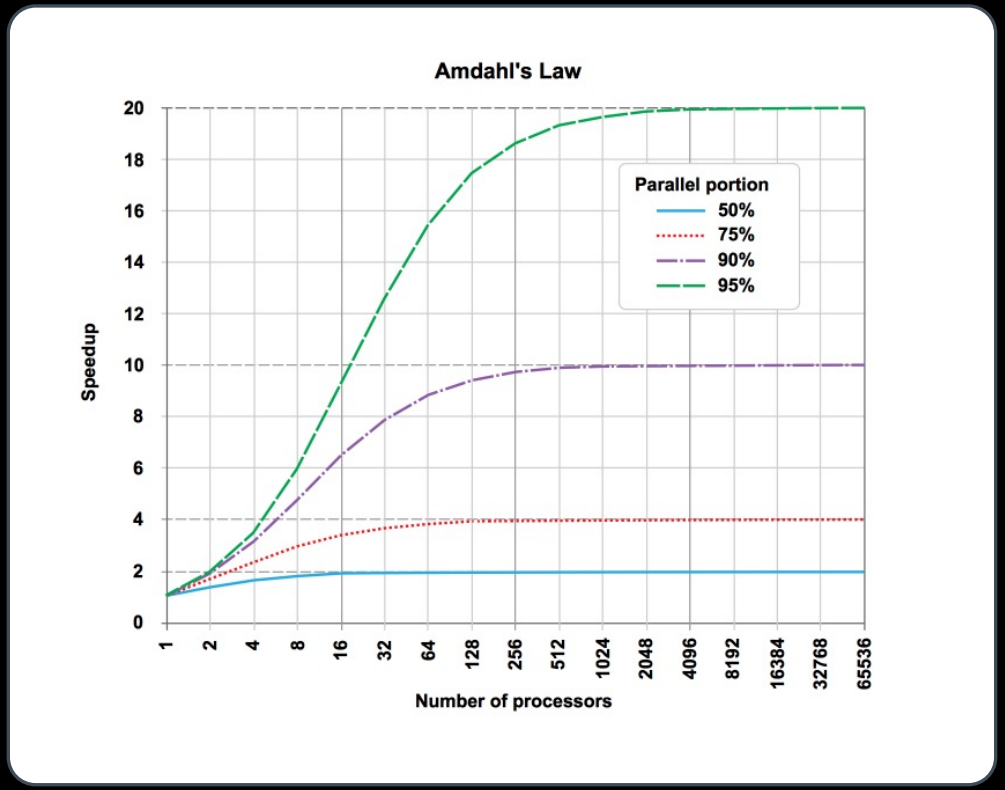
```
set title 'OpenMP Speedup'
set xlabel 'Threads'
set ylabel 'Speedup'
set key top left
plot x lc 2 lw 2 t 'Ideal Speedup’,\
   "cpi_openmp.dat" u 1:(4.55/$2)\
   w lp lc 1 lw 2 t 'OpenMP'
```



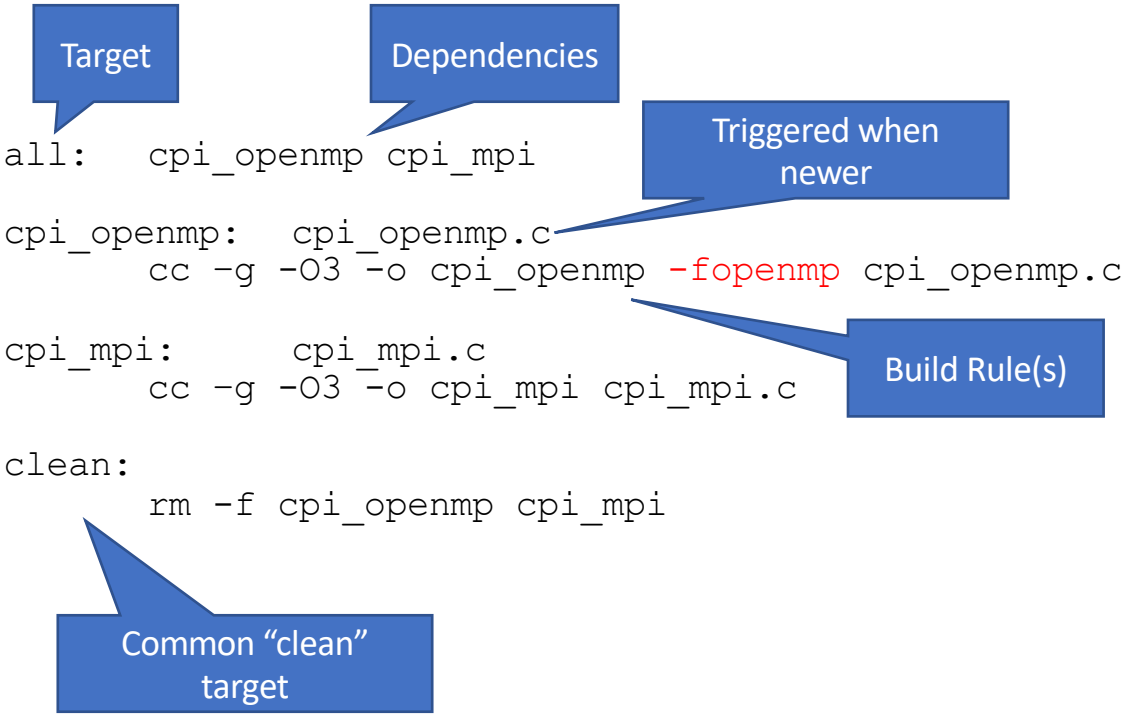
How does it scale?



Why won’t it scale?

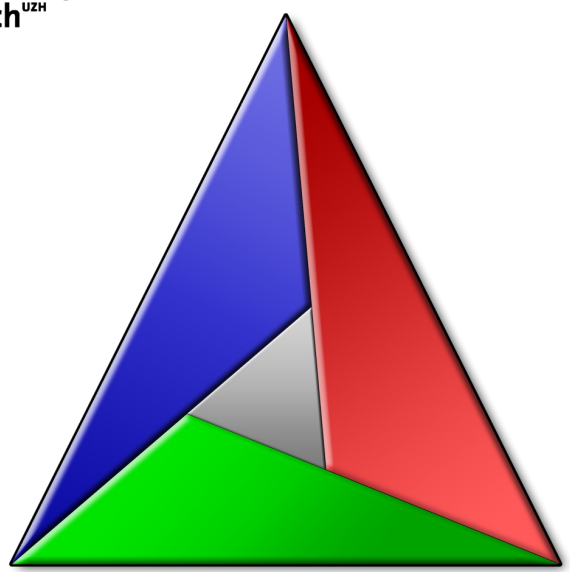


Makefile (for the C versions)




```
all:   cpi_openmp cpi_mpi

cpi_openmp:   cpi_openmp.c
	cc -g -O3 -o cpi_openmp -fopenmp cpi_openmp.c

cpi_mpi:      cpi_mpi.c
	cc -g -O3 -o cpi_mpi cpi_mpi.c

clean:
	rm -f cpi_openmp cpi_mpi
```

University of Zurich[UZH]



Advanced Make Systems

General Recipe

```
SRC=$HOME/source
DST=$HOME/install



cd $SRC
./configure\
  --prefix=$DST
make
make install
```

```
SRC=$HOME/source
DST=$HOME/install
mkdir $HOME/build
cd $HOME/build
cmake\
 -DCMAKE_INSTALL_PREFIX:PATH=$DST\
 $SRC
make
make install
```