Paul Chelarescu

# Optimizing experiments' software stack management with Spack

# About

Paul Chelarescu

Computer science - The University of Manchester

Supervisors:

Javier CERVANTES VILLANUEVA

Patricia MENDEZ LORENZO

Graeme STEWART

# Spack

"Flexible package manager that supports multiple versions, configurations, platforms, and compilers" - https://spack.io/

# What is a package manager?

Installing, configuring, upgrading, removing <packages>

Extract archives, ensure integrity, group by function, avoiding "dependency hell"

# HPC software

Massively combinatorial

package implementations × package versions × platforms × compilers types × compiler versions

→thousands of combinations; manual deployment does not scale

Reproducibility is important

# Simple install syntax

Sensible defaults

```
$ spack install zlib
==> Installing zlib
==> Warning: Suspicious requests to set or unset 'LD_LIBRARY_PATH' found
==> Warning:           env.unset('LD_LIBRARY_PATH') at /var/build/pchelare/home-gcc6/spack/lib/spack/spack/build_environment.py:276
==> Warning:    --->    env.set('%s' % key, value) at /var/build/pchelare/home-gcc6/spack/lib/spack/spack/build_environment.py:299
==> Fetching http://zlib.net/fossils/zlib-1.2.11.tar.gz
################################################################## 100.0%
==> Staging archive: /var/build/pchelare/home-gcc6/spack/var/spack/stage/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv/zlib-1.2.11.tar.gz
==> Created stage in /var/build/pchelare/home-gcc6/spack/var/spack/stage/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv
==> No patches needed for zlib
==> Building zlib [Package]
==> Executing phase: 'install'
==> Successfully installed zlib
  Fetch: 1.29s.  Build: 3.85s.  Total: 5.13s.
[+] /var/build/pchelare/home-gcc6/spack/opt/spack/linux-scientificcernslc6-x86_64/gcc-6.2.0/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv
```

# Flexibility and specificity

Specify version, compiler type and version, variants and additional constraints

```
$ spack install zlib@1.2.11%gcc@6.2.0+optimize+pic+shared arch=linux-scientificccernslc6-x86_64
==> Installing zlib
==> Warning: Suspicious requests to set or unset 'LD_LIBRARY_PATH' found
==> Warning:         env.unset('LD_LIBRARY_PATH') at /var/build/pchelare/home-gcc6/spack/lib/spack/spack/build_environment.py:276
==> Warning:   --->    env.set('%s' % key, value) at /var/build/pchelare/home-gcc6/spack/lib/spack/spack/build_environment.py:299
==> Fetching http://zlib.net/fossils/zlib-1.2.11.tar.gz
################################################################## 100.0%
==> Staging archive: /var/build/pchelare/home-gcc6/spack/var/spack/stage/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv/zlib-1.2.11.tar.gz
==> Created stage in /var/build/pchelare/home-gcc6/spack/var/spack/stage/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv
==> No patches needed for zlib
==> Building zlib [Package]
==> Executing phase: 'install'
==> Successfully installed zlib
  Fetch: 1.44s.  Build: 8.68s.  Total: 10.12s.
[+] /var/build/pchelare/home-gcc6/spack/opt/spack/linux-scientificccernslc6-x86_64/gcc-6.2.0/zlib-1.2.11-bcfxtupznjkymum6qgum4tod2yk3n4gv
```

```
zlib@1.2.11%gcc@6.2.0+optimize+pic+shared arch=linux-scientificccernslc6-x86_64
```

# Spec syntax is recursive

```
fccdevel

Concretized
-------------------------------
fccdevel@1.0%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^acts-core@0.05.03%gcc@6.2.0 build_type=RelWithDebInfo patches=691b104fa8dae2b9c29218ed5c98f55eef8575140f823f35d2406eb8021aed9f arch=linux-scientificcernslc6-x86_64
        ^boost@1.66.0%gcc@6.2.0+atomic+chrono~clanglibcpp cxxstd=default +date_time~debug+exception+filesystem+graph~icu+iostreams+locale+log+math~mpi+multithreaded~numpy patches=2ab6c72d03dec6a4ae20220a9dfd5c8c572c5294252155b85c6874d97c323199 +program_options~python+random+regex+serialization+shared+signals~singlethreaded+system~taggedlayout+test+thread+timer~versionedlayout+wave arch=linux-scientificcernslc6-x86_64
        ^cmake@3.7.0%gcc@6.2.0~doc+ncurses+openssl+ownlibs patches=dd3a40d4d92f6b2158b87d6fb354c277947c776424aa03f6dc8096cf3135f5d0 ~qt arch=linux-scientificcernslc6-x86_64
        ^dd4hep@01-05%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
        ^doxygen@1.8.11%gcc@6.2.0 build_type=RelWithDebInfo ~graphviz arch=linux-scientificcernslc6-x86_64
        ^eigen@3.2.9%gcc@6.2.0 build_type=RelWithDebInfo +fftw+metis+mpfr+scotch+suitesparse arch=linux-scientificcernslc6-x86_64
        ^graphviz@2.28.0%gcc@6.2.0~expat~ghostscript~go~gtkplus~gts~guile~io~java~libgd~lua~ocaml~pangocairo~perl~php~python~qt~r~ruby~sharp~tcl arch=linux-scientificcernslc6-x86_64
        ^root@6.12.06%gcc@6.2.0 build_type=RelWithDebInfo ~graphviz arch=linux-scientificcernslc6-x86_64
    ^delphes@3.4.2pre05%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
        ^tcl@8.6.8%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
            ^zlib@1.2.11%gcc@6.2.0+optimize+pic+shared arch=linux-scientificcernslc6-x86_64
    ^fastjet@3.3.0%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^fcc-edm@0.5.2%gcc@6.2.0 build_type=Release patches=1b3509ea1724ae5c78f67ecdf207c946a18f2455441d6e6b335d1e8f8b2d4cca arch=linux-scientificcernslc6-x86_64
        ^dag@0.1%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
        ^podio@0.8%gcc@6.2.0 build_type=Release arch=linux-scientificcernslc6-x86_64
            ^py-pyyaml@3.11%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
                ^python@2.7.13%gcc@6.2.0+dbm~optimizations patches=123082ab3483ded78e86d7c809e98a804b3465b4683c96bd79a2fd799f572244 +pic+pythoncmd+shared~tk~ucs4 arch=linux-scientificcernslc6-x86_64
    ^fcc-physics@0.2.1%gcc@6.2.0 build_type=Release arch=linux-scientificcernslc6-x86_64
        ^hepmc@2.06.09%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
        ^pythia8@235%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^gaudi@v29r2%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
        ^aida@3.2.1%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
        ^clhep@2.4.0.1%gcc@6.2.0 build_type=RelWithDebInfo +cxx11~cxx14 arch=linux-scientificcernslc6-x86_64
        ^cppunit@1.12.1_p1%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
        ^gperftools@2.5%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
        ^heppdt@2.06.01%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
        ^intel-tbb@2018_U1%gcc@6.2.0 cxxstd=default patches=ca08c28bdb15582c30777f9303d1986e4c09b3d514776494f3fbf5f19381bfda +shared+tm arch=linux-scientificcernslc6-x86_64
        ^py-qmtest@2.4.1%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
        ^xerces-c@3.1.3%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^geant4@10.04%gcc@6.2.0 build_type=RelWithDebInfo +cxx11~cxx14~motif~opengl~qt~vecgeom~x11 arch=linux-scientificcernslc6-x86_64
    ^heppy@develop%gcc@6.2.0 patches=ea347b3ae1ec00c7215c52c1e04ff6ca69e3c7713dbb063aab4c4fe03d873004 arch=linux-scientificcernslc6-x86_64
        ^py-gitpython@2.1.8-0%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^papas@1.2.0%gcc@6.2.0 build_type=Release patches=b1a3e0842c2f0d788f595c56f12c054efb7b049d99a26edefc80de8b08187de1 arch=linux-scientificcernslc6-x86_64
    ^pkg-config@0.29.2%gcc@6.2.0+internal_glib arch=linux-scientificcernslc6-x86_64
    ^tricktrack@1.0.5%gcc@6.2.0 build_type=RelWithDebInfo arch=linux-scientificcernslc6-x86_64
```

# Spec syntax is recursive



```
fccdevel@1.0%gcc@6.2.0 arch=linux-scientificcernslc6-x86_64
    ^acts-core@0.05.03%gcc@6.2.0 build_type=RelWithDebInfo patches=6
linux-scientificcernslc6-x86_64
        ^boost@1.66.0%gcc@6.2.0+atomic+chrono~clanglibcpp cxxstd=def
+log+math~mpi+multithreaded~numpy patches=2ab6c72d03dec6a4ae20220a9d
+regex+serialization+shared+signals~singlethreaded+system~taggedlayo
c6-x86_64
        ^cmake@3.7.0%gcc@6.2.0~doc+ncurses+openssl+ownlibs patches=d
rch=linux-scientificcernslc6-x86_64
        ^dd4hep@01-05%gcc@6.2.0 build_type=RelWithDebInfo arch=linux
        ^doxygen@1.8.11%gcc@6.2.0 build_type=RelWithDebInfo ~graphvi
        ^eigen@3.2.9%gcc@6.2.0 build_type=RelWithDebInfo +fftw+metis
        ^graphviz@2.28.0%gcc@6.2.0~expat~ghostscript~go~gtkplus~gts~
arp~tcl arch=linux-scientificcernslc6-x86_64
        ^root@6.12.06%gcc@6.2.0 build_type=RelWithDebInfo ~graphviz
```

# Variants

Options for packages, switch on/off

```
==> Error: An unsatisfiable variant constraint has been detected for spec:

    pcre@8.38%gcc@7.3.0 cppflags="-std=c++17" ~jit+utf arch=linux-scientificcernslc6-x86_64


while trying to concretize the partial spec:

    git@2.18.0%gcc@7.3.0 cppflags="-std=c++17" ~tcltk arch=linux-scientificcernslc6-x86_64
```
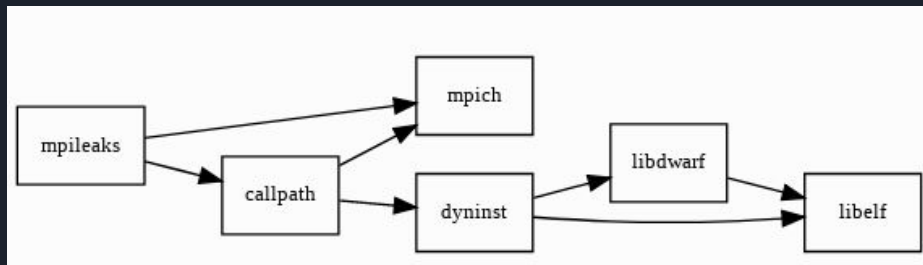
# Dependency graphs

Directed Acyclic Graph (DAG)

Consistent versions across DAG

Full control over the combinatorial build space

# Tasks

Add a package to the built-in repository

Improve build time using ccache

Upgrade gcc from version 6 to 7

Reuse previous installations from cvmfs using chained builds

Integrate build logs with cdash

# Packages are python classes

```python
class ActsCore(CMakePackage):
    """A Common Tracking Software (ACTS) Project provides a
    experiment-independent set of track reconstruction tools"""

    homepage = "https://gitlab.cern.ch/acts/acts-core"
    url      = "https://gitlab.cern.ch/acts/acts-core/repository/v0.05.02/archive.tar.gz"

    version('0.05.03', '872272ff18b38a01fc3f7b5f33be9d01')
    version('0.05.02', 'c824e925145bbd316b892ebe0c1eddc3')

    depends_on('cmake@3.5:', type='build')
    depends_on('boost@1.62:')
    depends_on('eigen@3.2.9:')
    depends_on('root@6.08.00:')
    depends_on('dd4hep@1.02:')

    # Optional for documentation
    depends_on('doxygen@1.8.11:')
    depends_on('graphviz@2.26.00:')

    conflicts("%gcc@:6.1")

    patch('cmake.patch', when='@0.05.03')

    def url_for_version(self, version):
        url = "https://gitlab.cern.ch/acts/acts-core/repository"
        return "{0}/v{1}/archive.tar.gz".format(url, version)

    def cmake_args(self):
        spec = self.spec
        args = [
            "-DBUILD_DD4HEP_PLUGIN=ON",
            "-DBUILD_MATERIAL_PLUGIN=on",
            "-DBUILD_TGEO_PLUGIN=ON",
            "-DEIGEN_INCLUDE_DIR=%s" % spec['eigen'].prefix + "/include/eigen3"
        ]
        if self.compiler.cxx14_flag:
            args.extend(["-DCMAKE_CXX_FLAGS=-std=c++14"])
        elif self.compiler.cxx17_flag:
            args.extend(["-DCMAKE_CXX_FLAGS=-std=c++17"])
        return args
```

# Packages are python classes

```python
class Pcre(AutotoolsPackage):
    """The PCRE package contains Perl Compatible Regular Expression
    libraries. These are useful for implementing regular expression
    pattern matching using the same syntax and semantics as Perl 5."""

    homepage = "http://www.pcre.org"
    url      = "https://ftp.pcre.org/pub/pcre/pcre-8.42.tar.bz2"

    version('8.42', '085b6aa253e0f91cae70b3cdbe8c1ac2')
    version('8.41', 'c160d22723b1670447341b08c58981c1')
    version('8.40', '41a842bf7dcecd6634219336e2167d1d')
    version('8.39', 'e3fca7650a0556a2647821679d81f585')
    version('8.38', '00aabbfe56d5a48b270f999b508c5ad2')

    patch('intel.patch', when='@8.38')

    variant('jit', default=False,
            description='Enable JIT support.')

    variant('utf', default=True,
            description='Enable support for UTF-8/16/32, '
            'incompatible with EBCDIC.')

    def configure_args(self):
        args = []

        if '+jit' in self.spec:
            args.append('--enable-jit')

        if '+utf' in self.spec:
            args.append('--enable-utf')
            args.append('--enable-unicode-properties')

        return args
```

# Packages are python classes

```python
class Minuit(AutotoolsPackage):
    """MINUIT is a physics analysis tool for function minimization."""

    homepage = "https://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/home.html"
    url      = "http://www.cern.ch/mathlibs/sw/5_34_14/Minuit2/Minuit2-5.34.14.tar.gz"
    list_url = "https://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/release/download.html"

    version('5.34.14', '7fc00378a2ed1f731b719d4837d62d6a')
    version('5.28.00', '536a1d29e5cc9bd4499d17d665021370')
    version('5.27.02', 'b54673f2b9b62a8ff4e6937a2ed8fda5')
    version('5.24.00', '9a915e56bee1e8986a719aa80e7b05d6')
    version('5.22.00', '2cbc34907bfe202c7a779e9713355846')
    version('5.21.06', 'b431ed129abb6c5020fd58d53cb8d27c')
    version('5.20.00', '3083d31e3764de45d477d082d60f2c29')
    version('5.18.00', 'a8764e7213fe811e56d5b6e5f3a91f5e')
    version('5.16.00', '6ea5feca06fca365d324bcfe16db7f08')
    version('5.14.00', 'b7452867b01c76cd115c696336c202d2')
    version('5.12.00', '36726b8c6fcddf4f0837c900461a1d3c')
    version('5.10.00', 'dfc7afc0add70deaca105ff549f5a786')
    version('5.08.00', '1cc8da07c4a247c877f39acf8d76ba02')
    version('1.7.9', '10fd518fc778317fdadbc4ef6f7ce8e4')
    version('1.7.6', '6a9a8341557de154274caff810686364')
    version('1.7.1', 'd202a1cf58662e9833f2967b4dc8808e')
    version('1.6.3', '955f560d0fb17bd7f081eddd7080fad6')
    version('1.6.0', '6992d70fc8fded50be49b6b358b58507')
    version('1.5.2', '31a0698febe59edd70aa001c4d7a56f8')
    version('1.5.0', 'bc502c66af071fcdc0a2ae45a8740c75')

    def url_for_version(self, version):
        if version > Version('5.0.0'):
            url = "http://www.cern.ch/mathlibs/sw/{0}/Minuit2/Minuit2-{1}.tar.gz"
            return url.format(version.underscored, version)
        else:
            url = "http://seal.web.cern.ch/seal/minuit/releases/Minuit-{0}.tar.gz"
            return url.format(version.underscored)

    patch('sprintf.cxx.patch', when='@5.08.00:5.18.00')
    patch('sprintf.patch', when='@:1.7.9')
    patch('LASymMatrix.h.patch', when='@:1.7.6')
```

# Customizable prefixes

Hash of DAG is part of prefix

As many installations are needed

Database is in json format

```
"xg5bwjh6umkzp2b6xtnikjqztlybi4yn": {
 "path": "/cvmfs/fcc.cern.ch/sw/releases/externals/93.1.0/x8
6_64-slc6-gcc62-opt/linux-scientificcernslc6-x86_64/gcc-6.2
.0/fccdevel-1.0-xg5bwjh6umkzp2b6xtnikjqztlybi4yn",
 "explicit": true,
 "spec": {
  "fccdevel": {
   "version": "1.0",
   "arch": {
    "platform": "linux",
    "platform_os": "scientificcernslc6",
    "target": "x86_64"
   },
   "compiler": {
    "name": "gcc",
    "version": "6.2.0"
   },
   "namespace": "fcc",
   "parameters": {
    "cflags": [],
    "cppflags": [],
    "cxxflags": [],
    "fflags": [],
    "ldflags": [],
    "ldlibs": []
   },
   "dependencies": {
    "acts-core": {
     "hash": "25innat6rdwrnbigc7btwovr4yosfy6e",
     "type": [
      "build",
      "link"
     ]
    },
```

# Ccache

Compiler Cache - Speed up recompilation

Now has simple integration with Spack

Provided by LCG releases or installed by Spack

```
config:
  ccache: true
```

# Comparison

Without ccache:

```
$ time spack install fccdevel %gcc@7.3.0

real    52m4.625s
user    229m22.115s
sys     43m42.191s
```

With ccache:
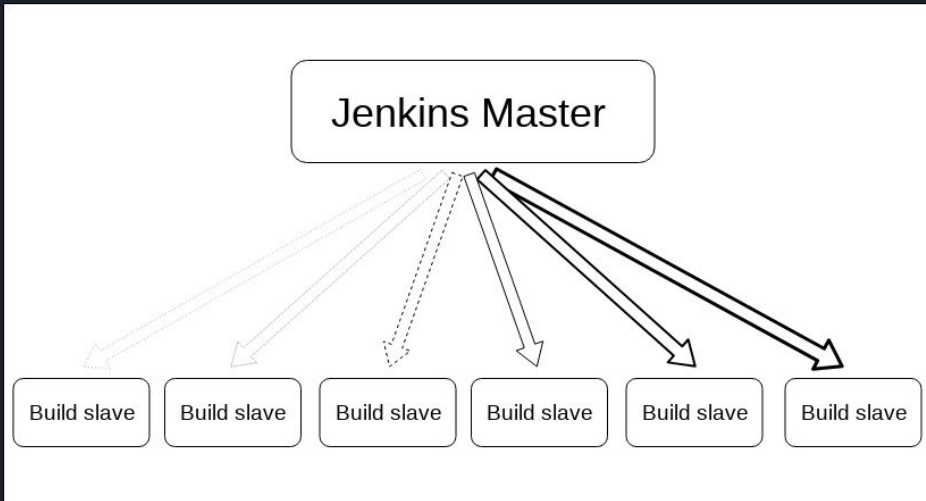
```
$ time spack install fccdevel %gcc@7.3.0

real    37m35.251s
user    58m54.019s
sys     33m18.348s
```

24 CPU Intel(R) Xeon(R) CPU E5-2630L  @ 2.00GHz
64 GB RAM

~30% reduction in real build time

# Jenkins integration

Last used node distribution algorithm

# GCC upgrade

Version 6.2 → 7.3

New configuration files

Patching CMake build files

Configuring std C++ flags in recipes

Investigate archive type incompatibility

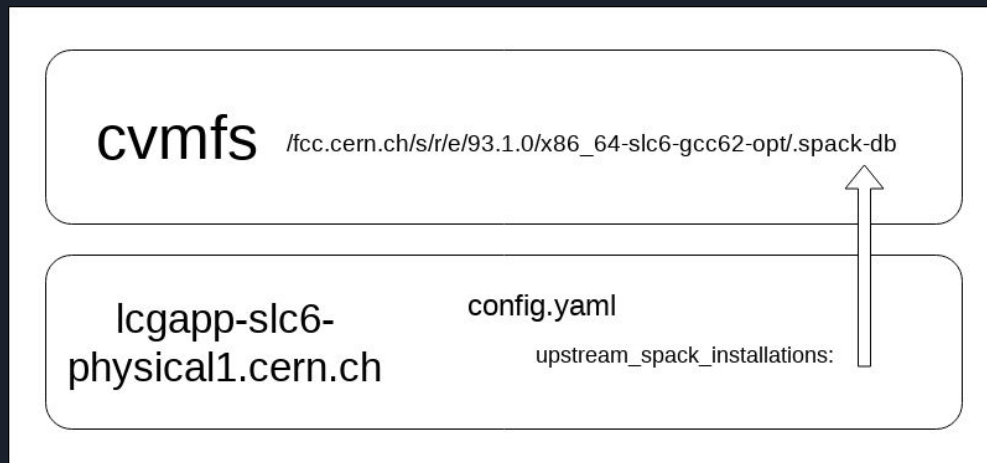Submit bug fixes to spack

Debug variant incompatibilities

Add new dependencies to fccdevel spec

# Chained installations

Reuse existing installed packages from other spack installations

Chain of spack install trees

# Only install what is needed

```
$ spack install fccdevel
==> acts-core is installed in an upstream Spack instance
==> tcl is installed in an upstream Spack instance
==> delphes is installed in an upstream Spack instance
==> dag is installed in an upstream Spack instance
==> py-pyyaml is installed in an upstream Spack instance
==> podio is installed in an upstream Spack instance
==> fcc-edm is installed in an upstream Spack instance
==> fcc-physics is installed in an upstream Spack instance
==> gaudi is installed in an upstream Spack instance
==> openssl is installed in an upstream Spack instance
==> curl is installed in an upstream Spack instance
==> bzip2 is installed in an upstream Spack instance
==> pkg-config is installed in an upstream Spack instance
==> ncurses is installed in an upstream Spack instance
==> tar is installed in an upstream Spack instance
==> gettext is installed in an upstream Spack instance
==> libiconv is installed in an upstream Spack instance
==> readline is installed in an upstream Spack instance
==> gdbm is installed in an upstream Spack instance
==> perl is installed in an upstream Spack instance
==> git is installed in an upstream Spack instance
==> Installing py-setuptools
  Fetch: 0.57s.  Build: 1.97s.  Total: 2.54s.
==> heppy is installed in an upstream Spack instance
==> Installing papas
  Fetch: 2.52s.  Build: 43.22s.  Total: 45.74s.
==> tricktrack is installed in an upstream Spack instance
==> Installing fccdevel
  Fetch: 1.76s.  Build: 2.02s.  Total: 3.79s.
```

# Significant reduction in build time

28 seconds vs 52 minutes, more than 100x speedup

```
$ spack install fccdevel

==> fccdevel is installed in an upstream Spack instance
```

# In progress

Migrating bash configuration scripts to python modules

# CDash

Future review - still in progress (incomplete)

Thank you