

The Deconstruction of Dyninst





Building new tools, exposing new features

Deconstruction Principles

- ➤ Abstract
 - Divide the complex problems of binary analysis and instrumentation into simple, well-defined pieces
 - Generalize prior special-purpose solutions
- > Extensible
 - Allow users to add new functionality
 - Export the results of analyses for use by other tools
- > Portable
 - Hide platform-specific details
- Promote sharing and reuse
 - Small, well-defined components are easy to adopt
 - Encourage competing alternative approaches

Status

- > SymtabAPI supports PE/PDB, ELF/DWARF
- ➤ InstructionAPI supports x86, x86-64, PowerPC
- ParseAPI supports x86, x86-64, PowerPC
- StackwalkAPI supports Linux, Windows, BlueGene, FreeBSD
- ProcControlAPI supports Linux, Windows, BlueGene, FreeBSD
- ➤ DataflowAPI beta supports x86, x86-64, PowerPC

ProcControlAPI

- Controls processes: start, stop, spawn, kill
- Monitors processes: fork/exec, library load/unload, signals
- Modifies processes: poke code/data into address space

PatchAPI

- Specifies where to instrument a binary via instrumentation point abstraction
- Splices new code into a binary

StackwalkAPI

DyninstAPI

- Generates call stack traces in both 1st-party and 3rd-party modes
- Understands frameless functions, signal handlers, and more
- Extensible to new frame layouts, such as instrumentation

DataflowAPI

- Collection of dataflow analyses
- Includes stack depth, liveness, slicing, and symbolic evaluation

ParseAPI

- Performs control flow analysis
- Builds control flow graph (CFG) and call graph for other components to use

Instruction Semantics

- Adds semantic information to the InstructionAPI representation
- Provides a foundation for constant propagation, partial evaluation, execution simulation

Code Generator

 Converts architectureindependent abstract syntax tree (AST) representation to machine language

SymtabAPI

- Reads and updates symbol tables, debug information, dynamic linkage information, exception information, and type information
- Supports multiple file formats across multiple platforms

InstructionAPI

- Decodes machine instructions to an abstract representation
- Represents operand address calculations
- Provides register liveness and control flow target information