BioSolveIT
expect actives!

C-S-D Deployment Options

- Three Ways to Run C-S-D
- Chose Your Environment

Requirements

Minimum Hardware (HPSee)

Memory (RAM)
- 32 GB minimum
- ~2 GB per core

Disk Space
- ≥ 10GB per C-S-D
- 500 GB minimum

Cores
- 32 cores minimum (→ ~5-6 days per CSD run)

More Hardware = Faster Processing

- 32 cores → ~5-6 days per CSD run
- 128 cores → ~1-2 days per CSD run
- 1024 cores → ~1-2 hrs per CSD run

SeeSAR
- 16 GB RAM

Operating Systems (HPSee)

Linux
- Docker Engine

MacOS
- currently not supported

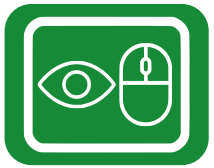
Windows
- Windows 10 and later
- WSL2
- Docker Desktop

Server

Client

Fully supported
- Linux
- MacOS
- Windows

Three Ways to Operate C-S-D



GUI

SeeSAR interface

Manually driven by the user from within SeeSAR. HPSee orchestrates remote compute behind the scenes.

BEST FOR Medicinal Chemists; interactive, visual work

- Interactive
- Visual control
- HPSee required

API

e.g., Python, Jupyter Notebook

Accessible upon request

Scripted access to the HPSee API. HPSee still orchestrates the remote heavy-duty compute.

BEST FOR Semi- to fully automated workflows

- Automated
- Workflow integration
- HPSee required

CLI

Command-line tools

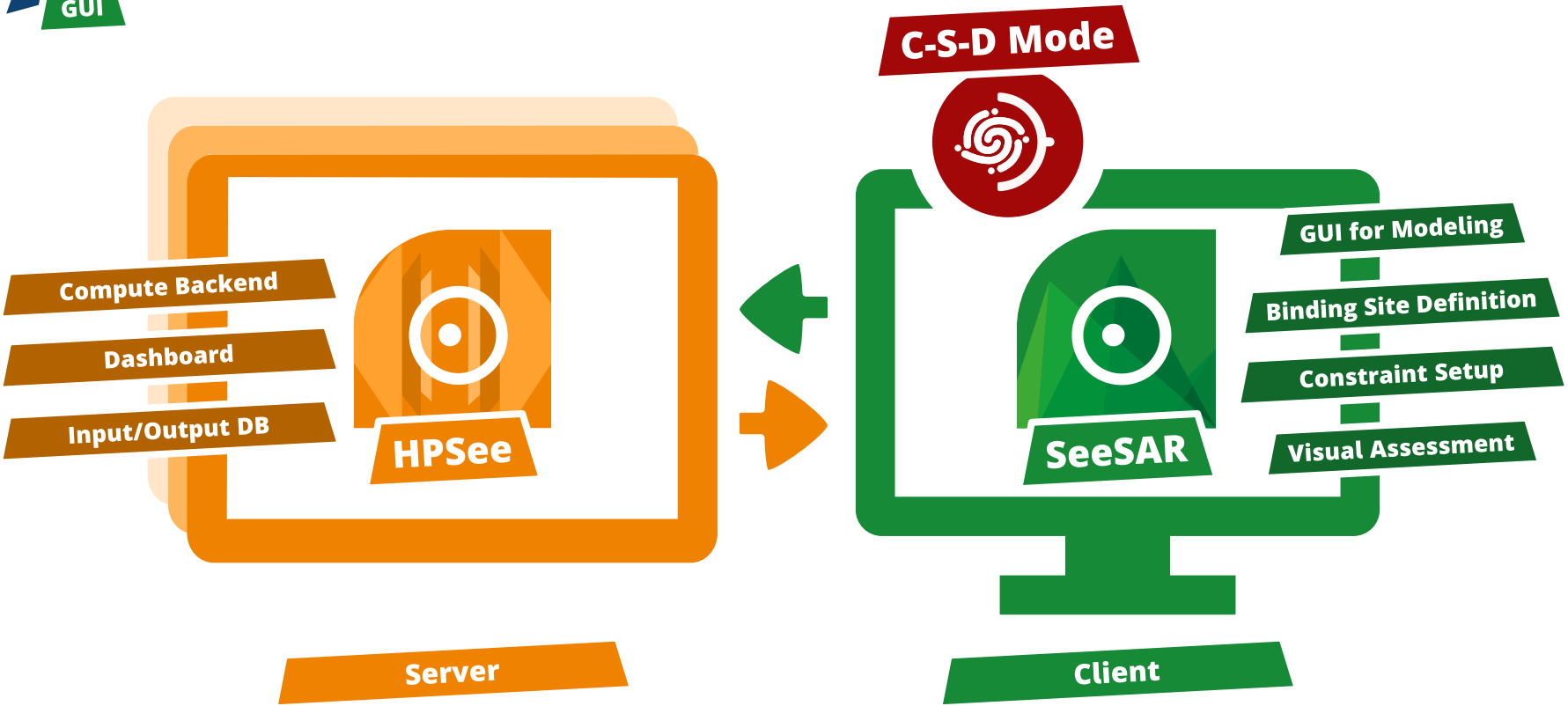
In development

Scripted access via standalone executables. Full flexibility, full responsibility for execution.

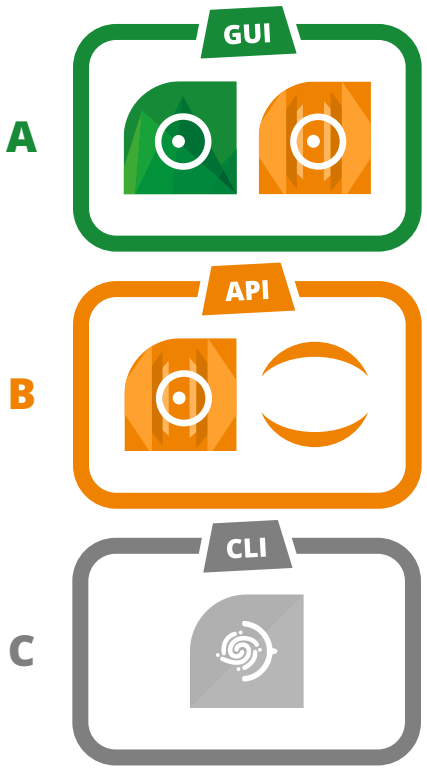
BEST FOR Computational experts needing full control

- Standalone
- Maximum flexibility
- No HPSee

Technical Setup

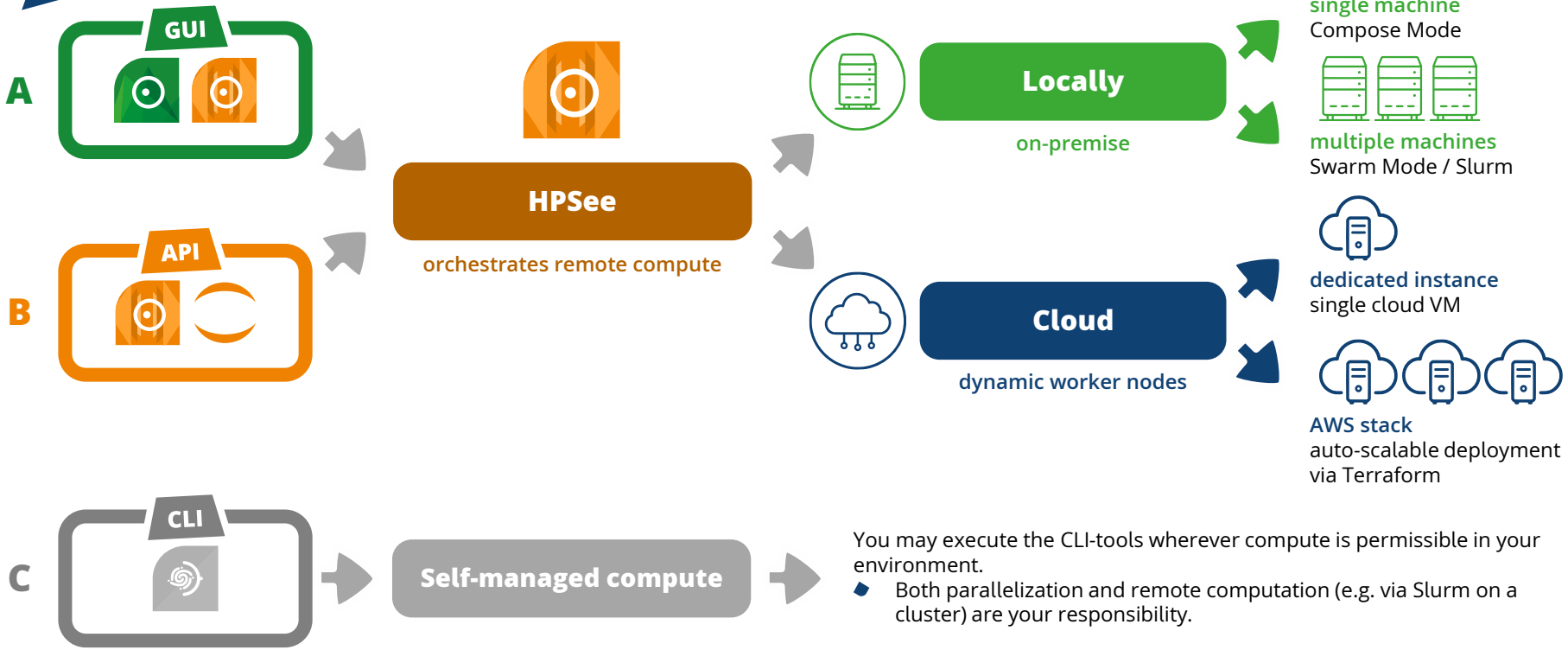


The Right Setup for Every Scenario



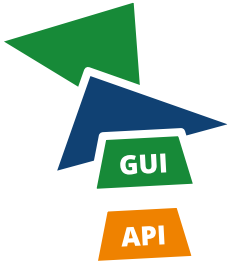
	Strengths	Good-to-Know
A	• Easy for medchems • Compute happens behind the scenes • Tailor binding site (e.g., pharmacophores) • User is in full control of every step • Visual inspection/manual selection	• HPSee setup required • Requires IT expert maintenance • No fully automated workflow
B	• Easy to use for computational chemists • Compute happens behind the scenes • Semi- to fully automated workflows • Integrate with your internal workflows • Full process control retained	• HPSee setup required • Binding site tailoring & pharmacophores require manual SeeSAR setup • Only basic example scripts included for filtering & selection
C	• No HPSee required • Most flexible deployment mode • May be combined with external tools if synthon name and USMILES stay unchanged	• User handles large-scale execution • No built-in parallelization or remote computation • Requires advanced compute expertise • Only basic example scripts included for filtering & selection

Where the Computation Happens



You may execute the CLI-tools wherever compute is permissible in your environment.

- Both parallelization and remote computation (e.g. via Slurm on a cluster) are your responsibility.



On-Premise Computation

Note: Hardware must be 100% dedicated to HPSee

HPSee will not throttle itself for other jobs on the same machine. A fixed share such as 50% may be reserved **only during setup** but cannot be changed dynamically afterward.

HPSee also cannot yet use machines managed by external batch-queuing systems (Slurm or PBS). Until batch-queue support is available, all hardware assigned to HPSee must remain dedicated. BioSolveIT is working on compatibility.

Cloud Computation

dedicated instance
single cloud VM

AWS stack (under development)
auto-scalable deployment via Terraform

Always-On Core
A minimum footprint **runs permanently**.
The HPSee API and the database (storing all input/output).

On-Demand Spot Worker Nodes (under development)
AQS Spot workers automatically scan with demand (within predefined limits) and shut down when no longer required.
A highly cost-efficient cloud option for accelerating compute-intensive workloads.

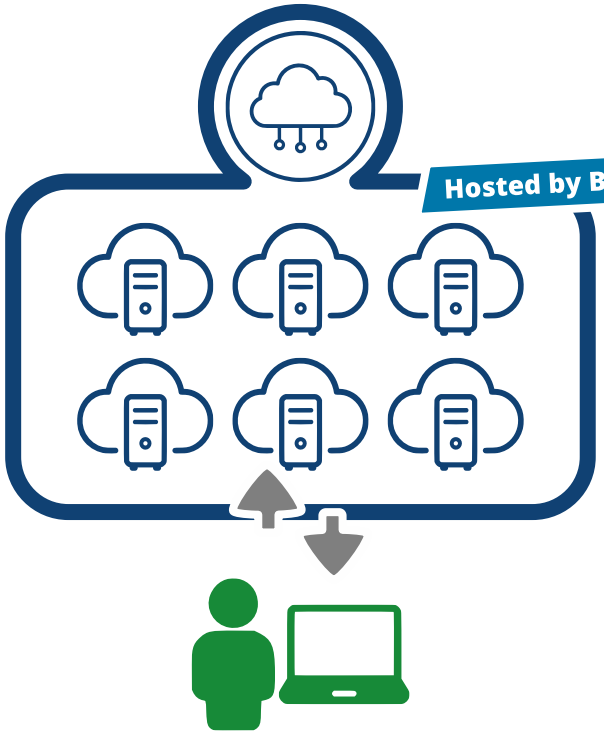
Autoscaling validated with up to 200 Amazon EC2 virtual machines running in parallel.

- C-S-D runs can be completed in a few hours
- Scaling workers directly cuts total compute time.

Hosted Evaluation Environment

GUI

API



BioSolveIT offers a dedicated, **fully pre-configured AWS sub-account** for users, who would like to evaluate C-S-D.

Designed for Easy Evaluation

- Fully configured before access
- Chemical Spaces of interest are preloaded
- No local infrastructure required
- Ideal for users with little cloud experience

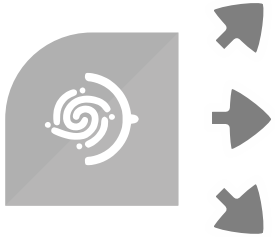
Cost & Security Considerations

- Actual AWS usage is billed separately.
- Safeguards according to AWS standards

Dock Trillions+ Your Way: C-S-D Command-Line Tools

Versatile Sandbox Engine

C-S-D CLI enables to screen Chemical Spaces within any in-house infrastructure with full control over resources, processing, and data.



C-S-D CLI Tools

- SynthonPrepper
  Extracts and prepares synthons
- SynthonDocker
  Docks and scores synthons
- SynthonExtender
  Extends given synthons based on the chemistry in the Space

Screening trillions and more in your ecosystem:

- Runs where your compute and data already live.

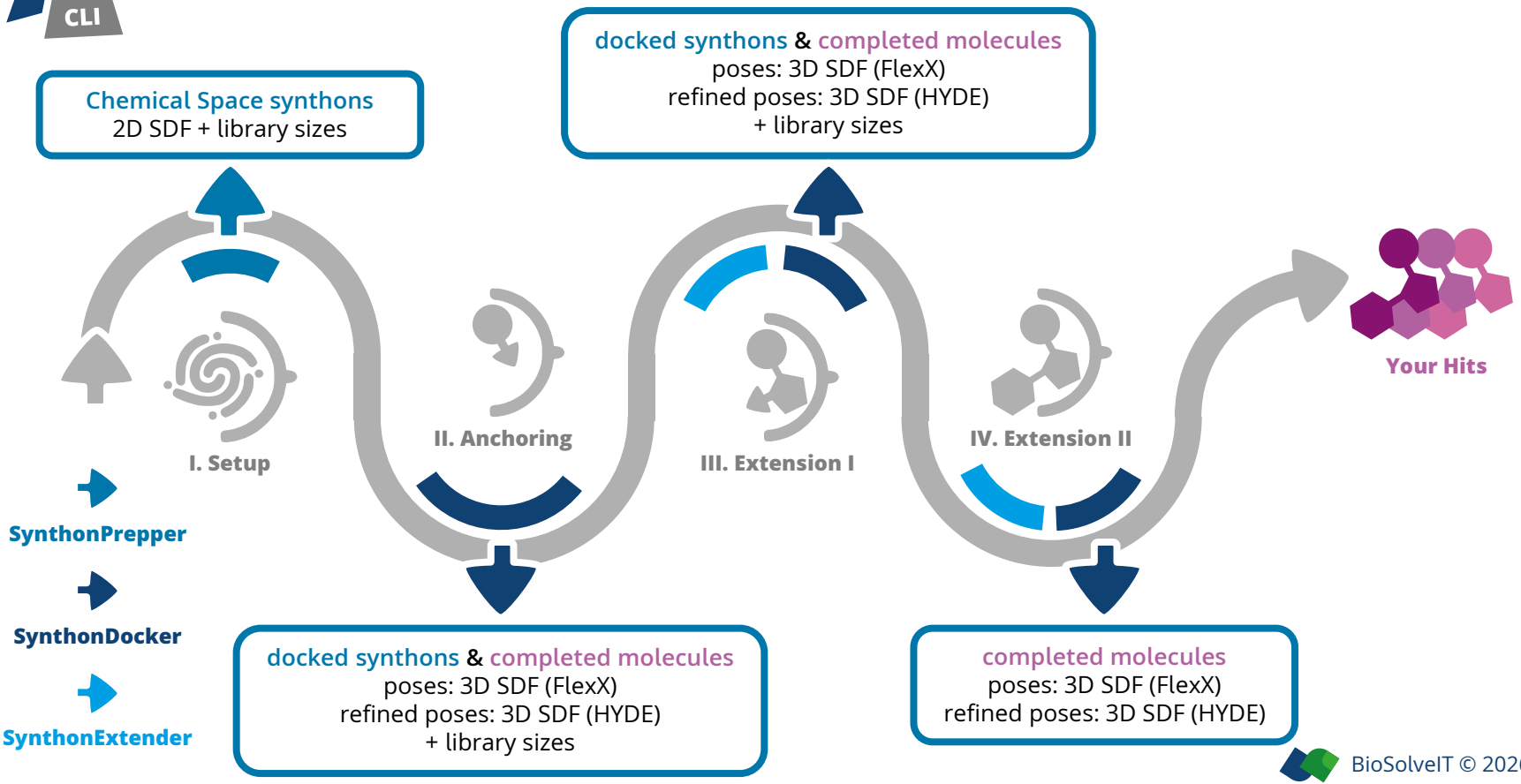
Augmentable with your tools:

- Connects to custom scripts, models, filters, and third-party software.

Workflow interaction on your terms:

- Export and re-introduce data at defined workflow stages.

What Data can be Accessed?

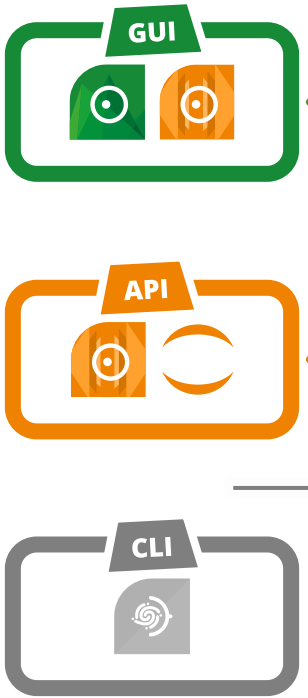


Choose Your Setup

Want an interactive, visual experience?

Want to automate and customize?

Want it running on your own infrastructure?



for dedicated hardware you control

Locally

for external compute resources

Cloud

- autoscaling possible
- BioSolveIT-hosted AWS evaluation environment offered