上海科技大学
ShanghaiTech University

CS289: **Mobile Manipulation Fall 2026**

Sören Schwertfeger

ShanghaiTech University

STAR CENTER

Outline

- Admin
- Planning
- OMPL
- MoveIt

ADMIN

Paper Presentation

- Choose one paper from ICRA or IROS which is relevant to your project!
  - ICRA: https://ieeexplore.ieee.org/xpl/conhome/1000639/all-proceedings (any year!)
  - IROS: https://ieeexplore.ieee.org/xpl/conhome/1000393/all-proceedings (any year!)
  - Only full papers (6 or more pages) are allowed; no workshop papers

- **Send TA an email with your paper – no double papers are allowed – first come first serve!**

- Present the paper as if it were your own work!
- Front page: Name of the Paper; Full citation of the paper; **Your name in Pinyin; Your email address**
- Last slide: **ONE** slide about how this paper is relevant to your project.
- Your presentation has to be professional – not cute…
- **No videos/ no animations allowed!**
- **Put slide numbers in your presentation!**
- Submit pdf or ppt to the paper repository till Monday, Oct 26 23:59 to repo! Late submissions (or if you come with the ppt/ pdf to the presentation time) will receive a **flat 33% loss of points**!
- Presentations in 2 (or 3) slots – during lecture slots Oct 27 & 29.
- 10 minutes presentation plus 1 minute project relevance plus 2 minutes questions
  - Do not rush your presentation! Better present less items more slowly!
  - 10 minute presentation => 8 – max. 12 slides
  - Maybe have a slide towards the end that you can skip if you run out of time.
  - Give a test presentation to your friends beforehand!
- Finish early for practicing – don't learn by heart.

Grading of the Presentation

- 10 %: Your basic understanding/ knowledge about the paper you present
- 20 %: Presentation timing (plus or minus one minute is ok) – no rushing – good speed!
- 10 %: Correct written English in presentation:
  - No complete sentences, no grammatical or spelling mistakes
- 10 %: Good structure of presentation:
  - Depends on the type of paper, how much time you have, how long you need to present the main achievement.
  - For example: outline, introduction/ motivation, **problem statement**, state of the art, **approach**, experiments, **results**, **conclusion**, outlook
- 20 %: Clarity of written presentation
- 10 %: Good presentation style:
  - Interact with audience: look at the whole room (not just your slides, notes, or the back of the room)
  - Present the paper – do not read (or repeat the learned) speech from a prepared text
  - Use the presentation as visual aid – not as your tele-prompter to read from
  - Move your body – do not stand frozen at one place
- 10 %: Answering the questions
  - Questions have to be asked and answered in English – Chinese can be used for clarification
- 10 %: Asking questions to other students!
- **Not** scored: Your English skill

Project Selection so far

osmAG-moma	Ge Yang	Ziyu Fan	Zhengyi Fan
Small Humanoid Robot Crawling	Yibo He	Yi Huang	Qinyuan Li
Fetch Pick and Place	Shiyi Wang	Guolin Wang	
Noetix Humanoid Robot	Zhongli Yin	Yawen Wang	
HumanDex (Tianji)	Weihan Xu	Shenrui Wu	
DexManipulation (Tianji)	Wenkai Zhu	Junlei Zhu	Ran Ji

3 persons project selection outstanding – maybe talk to me after the lecture…

Project Proposal (1)

- **Title:** Find a nice, catchy title for your project
- **Abstract**: A short abstract/ summary what the project is about
- **Introduction**: general description & Motivation
- **State of the Art**: Literature & open-source-ROS packages
- Per team member:
  - present and cite three papers with just three or four sentences
  - present in more detail one further paper relevant to your project. Describe it with at least 1/3rd of a page.
  - present in detail one open source ROS package relevant to your project. At least 1/3rd of a page
  - => about one page per team member – => 3 pages for 3 persons team

Project Proposal (2)

- **System Description**
- **System Evaluation**: Describe how you want to test your system.
  - Experiments & how to measure their success
- **Work Plan**: Define some mile stones.
  - Possible phases: Algorithm design, implementation, testing, evaluation, documentation – some of those things can also happen in a loop (iteration).
  - Deliverables of Project:
    - Proposal (this document)
    - Final demo
    - Final Report
    - Website
- **Conclusions**: Short summary and conclusions

Project Proposal (3)

- Important dates:
- **Oct 8, 23:59: due date for the proposal**
- **Jan 08, 23:59 : due date for the final report**
  Jan 10: due date for the final demo & website
- **You can make the demo/ report also between December 1st and 10th!**

- Parts of proposal go into the final project report.
- Please don't forget to take **pictures and videos** when testing your system!

- In English! Using LaTeX!
- Put sources and PDF in git.

- **Additional task:** In glit/ gitlab: "Readme.txt" with:
  - Team Name and Members; email addresses
  - Documentation and how to's regarding your project.

- Next Lecture: Oct 8 ☺
- Write your project proposal during the missing lecture slot – it counts as Project time.

PICK & PLACE

Kinematic Problems for Manipulation

- Reliably position the tip - go from one position to another position
- Don’t hit anything, avoid obstacles
- Make smooth motions
  - at reasonable speeds and
  - at reasonable accelerations
- Adjust to changing conditions -
  - i.e. when something is picked up *respond to the change in weight*

Planning Problem

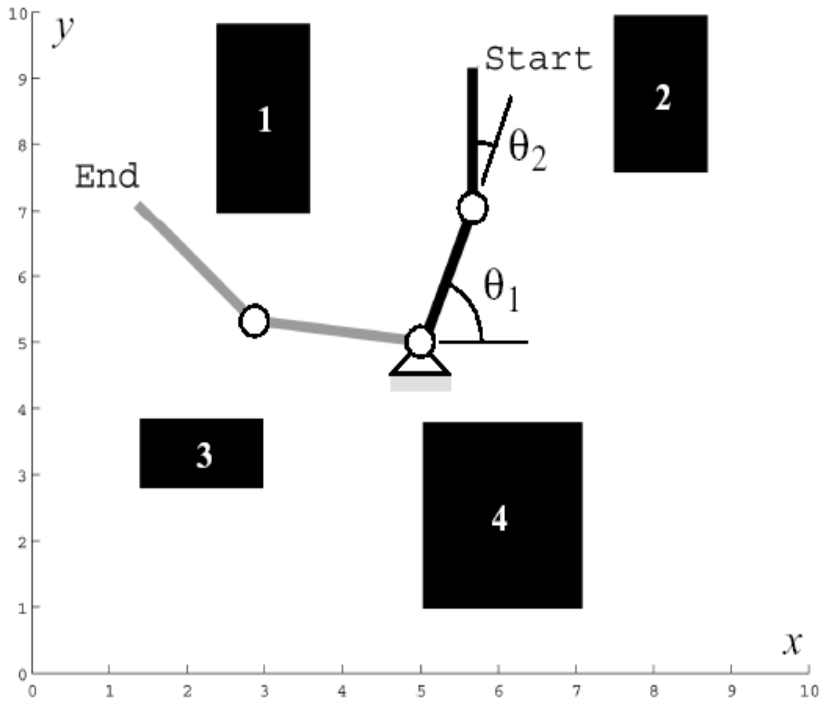
- (Arm) Pose: Set of join values
- (Arm) Trajectory:
  - Given a start pose and an end pose
  - A list of intermediate poses
  - That should be reached one after the other
  - With associated (desired) velocities and accelerations (maxima)
    - Without time (without velocity and acceleration): path! So:
    - Path: poses; Trajectory: poses with speeds (and maybe accelerations)
- Constrains:
  - Don’t collide with yourself
  - Don’t collide with anything else (except: fingers with the object to manipulate!!!)
  - Additional possible constrains:
    - Maximum joint velocities or accelerations
    - Keep global orientation of a joint (often end-effector) within certain boundaries

Planning Problem cont.

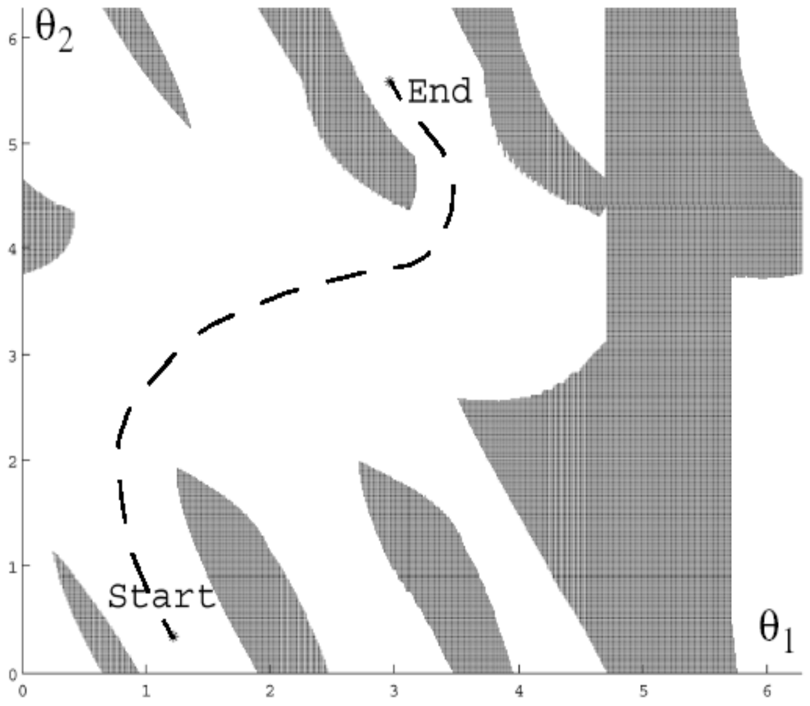
- Often the goal specified in Cartesian space (not joint space)
- => use IK to get joint space
- => often multiple (even infinitely many) solutions
  - Which one select for planning?
  - Plan for several solutions and select best!?

Work Space (Map) → Configuration Space

- State or configuration q can be described with k values q_i



a)



b)

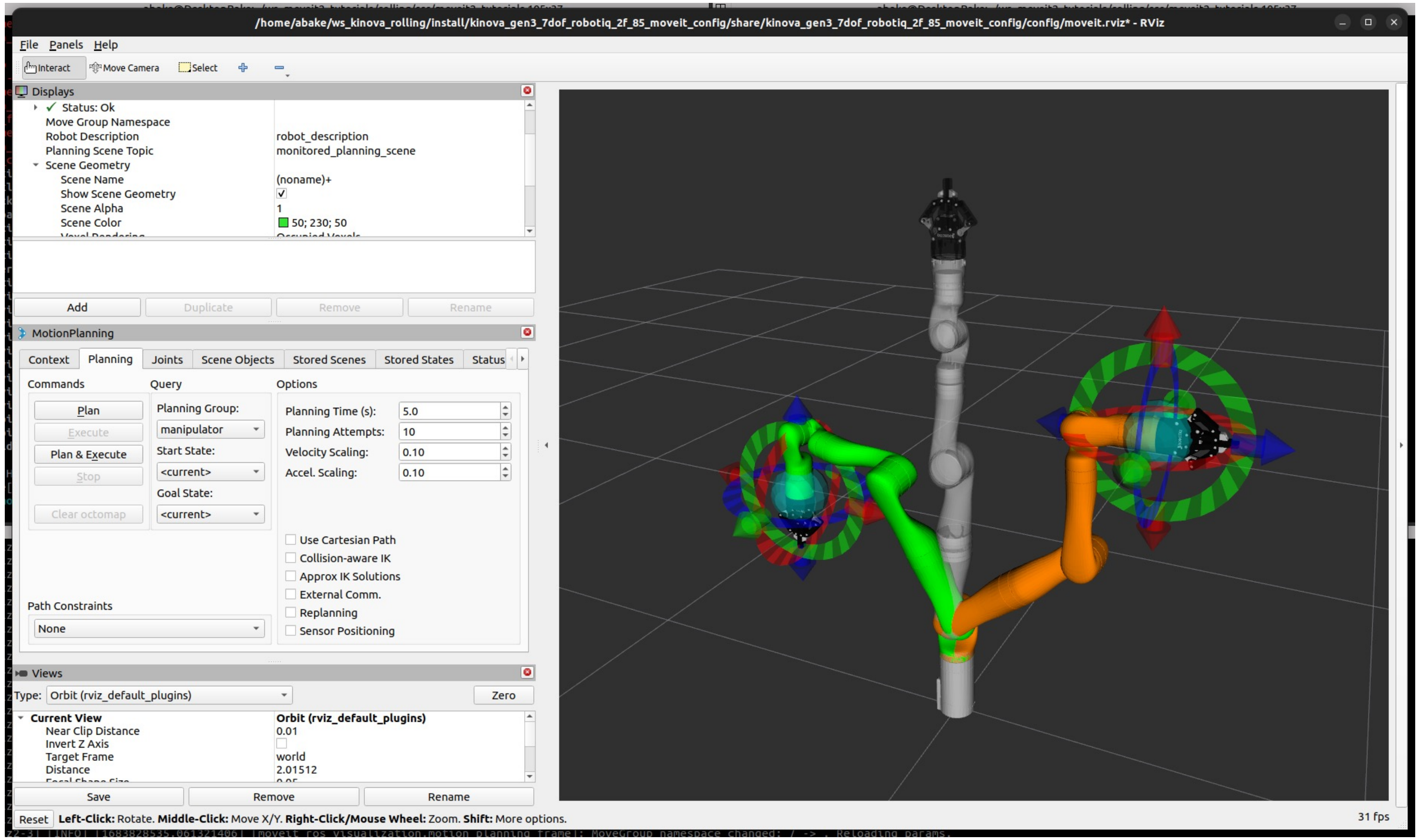
Work Space
Dimension depends on map
Dimension – typically 2D or 3D

Configuration Space:
the dimension of this
space is equal to the Degrees of Freedom (DoF)
of the robot

MOVEIT

MoveIt!

- Software to Control & Plan with robot arms



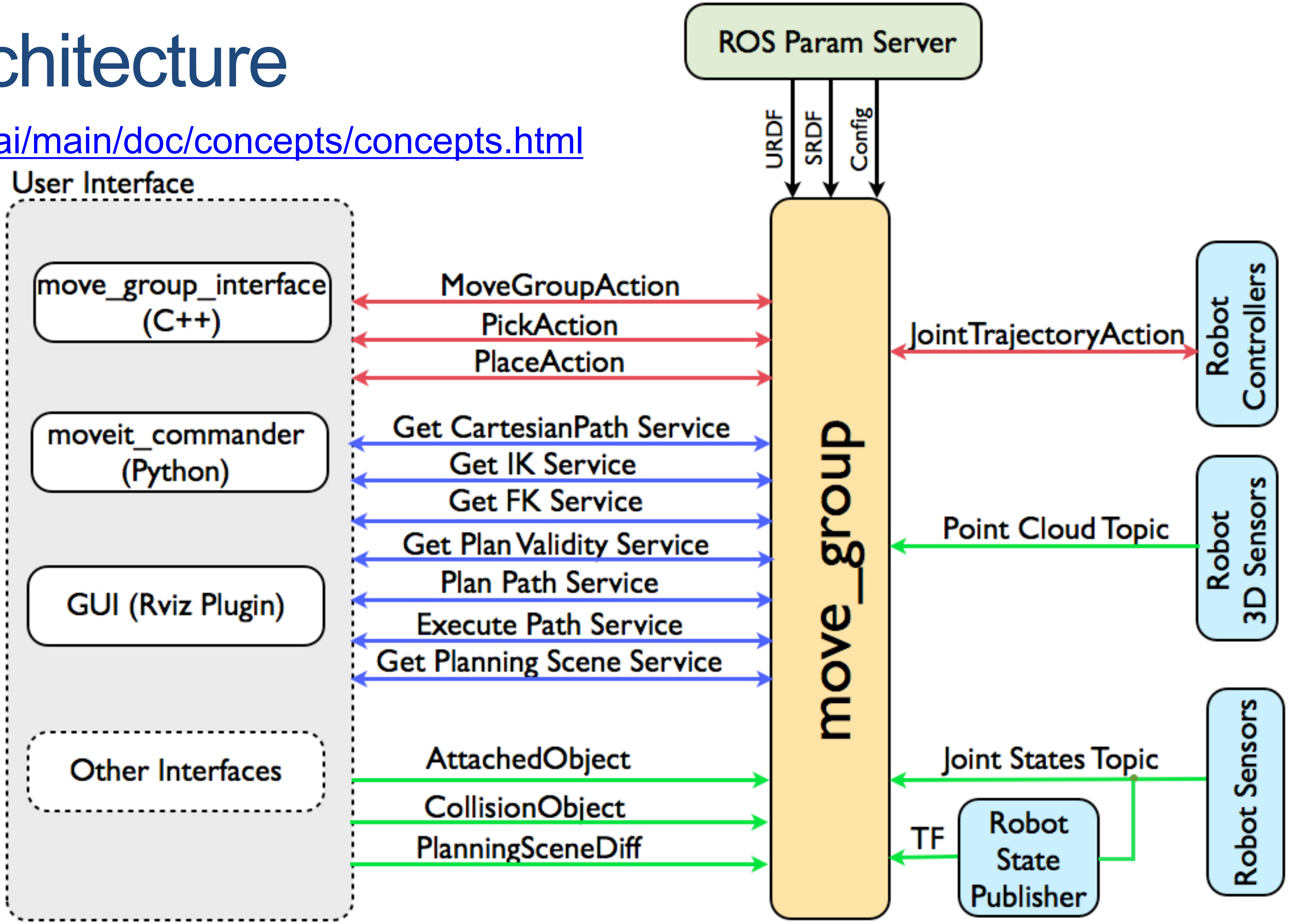
- Part of ROS

- https://moveit.picknik.ai/

- First let's explore how to tell MoveIt! what kind of arm you have…

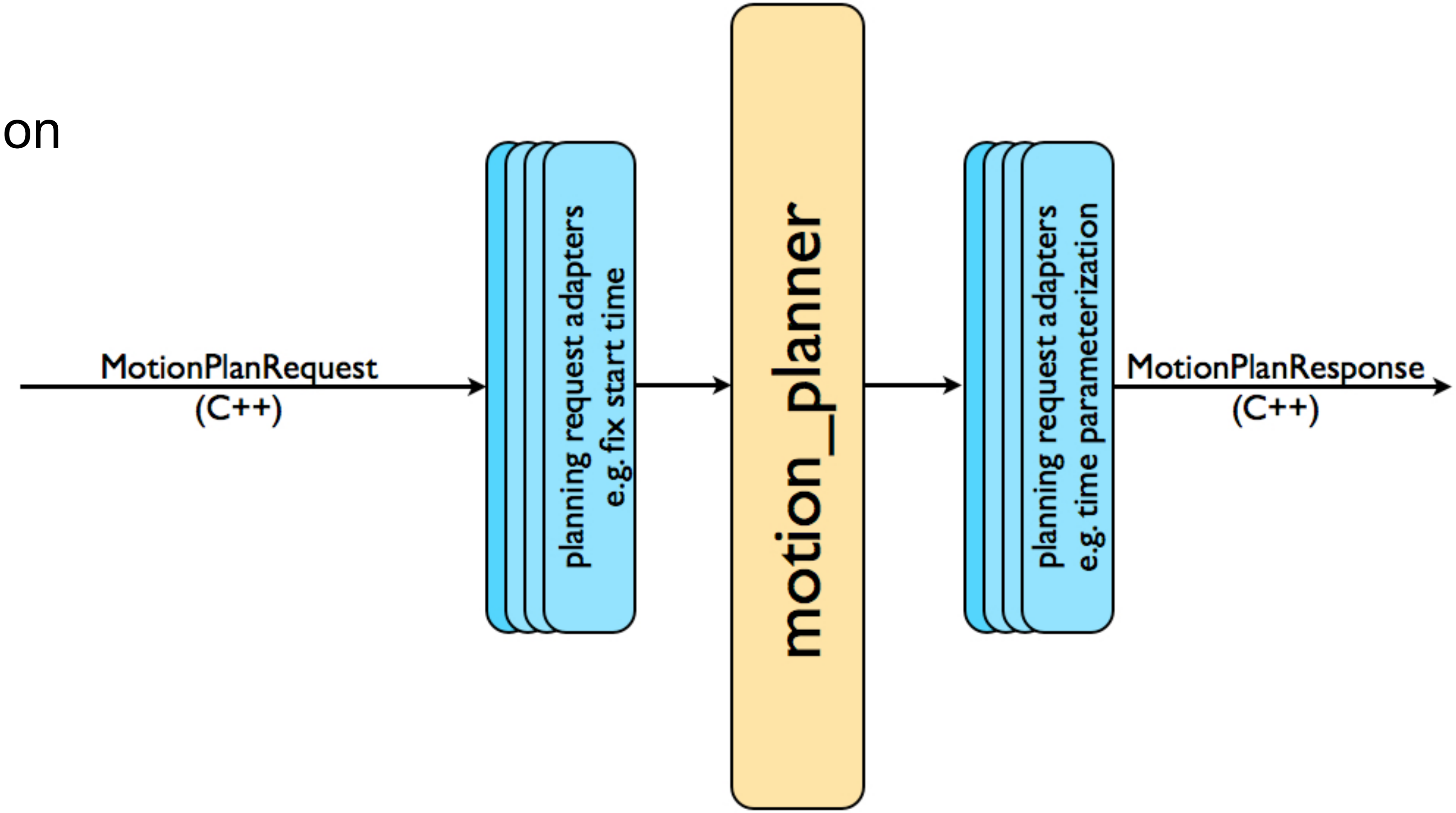
System Architecture

https://moveit.picknik.ai/main/doc/concepts/concepts.html



Motion Planning

- Mainly:
- OMPL (Open Motion Planning Library)

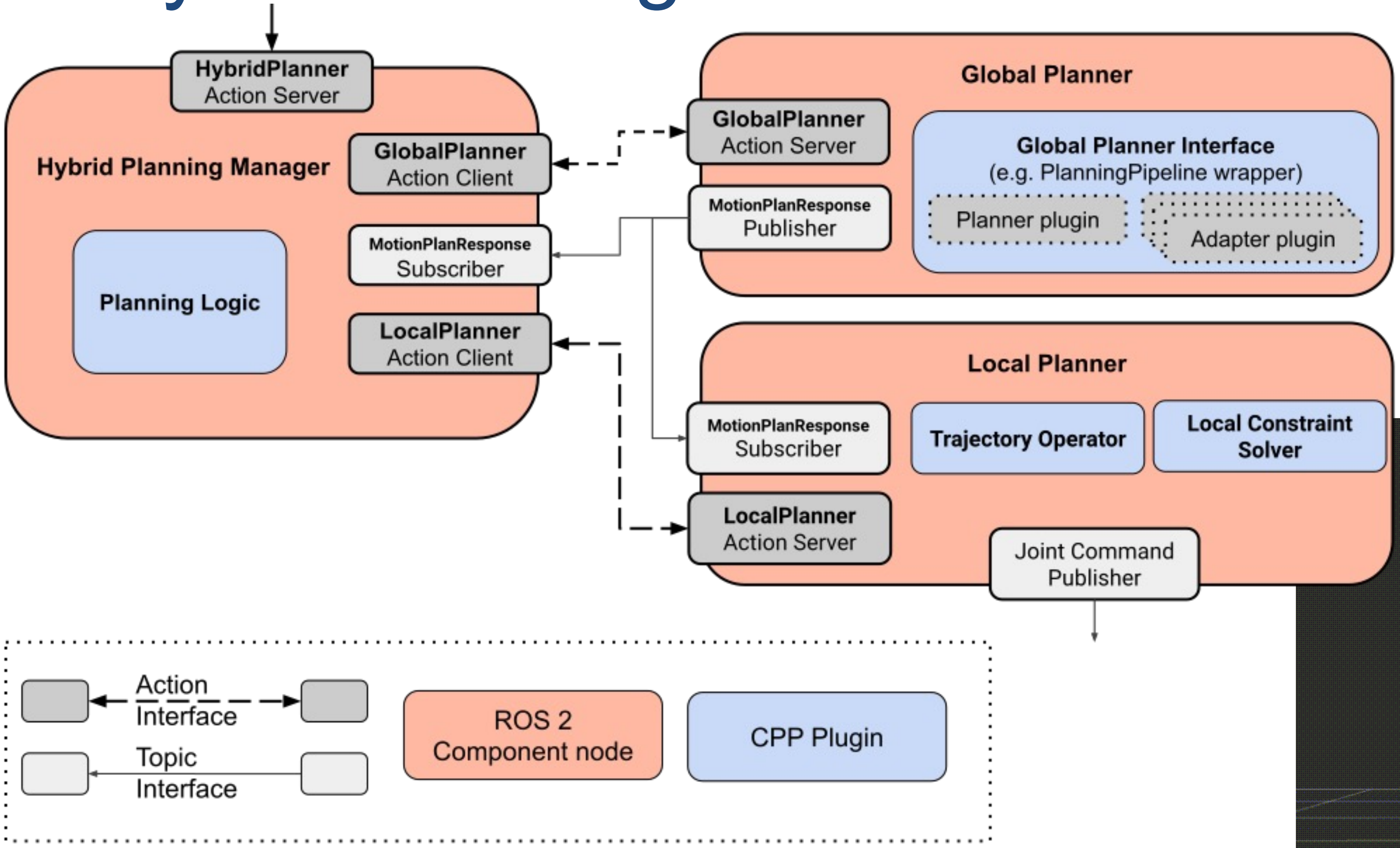


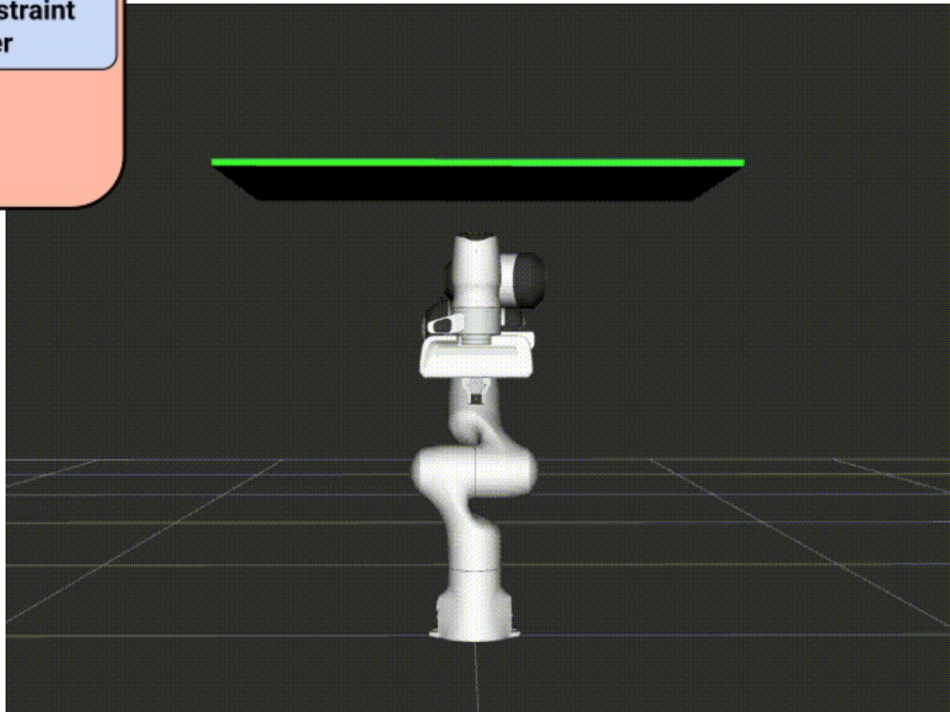
Hybrid Planning
Global and Local

- Online motion planning: The global planner creates an initial global solution and continuously optimizes it. Simultaneously, the local planner executes the reference trajectory and blends updated trajectory segments into it.
- Reactive Motion: The global planner is used to fix invalidated solutions (replanning) while the local planner slows down or halts before collisions
- Adaptive Motion: The local planner is used to adapt a global solution to dynamic conditions like keeping steady tool contact with an uneven surface

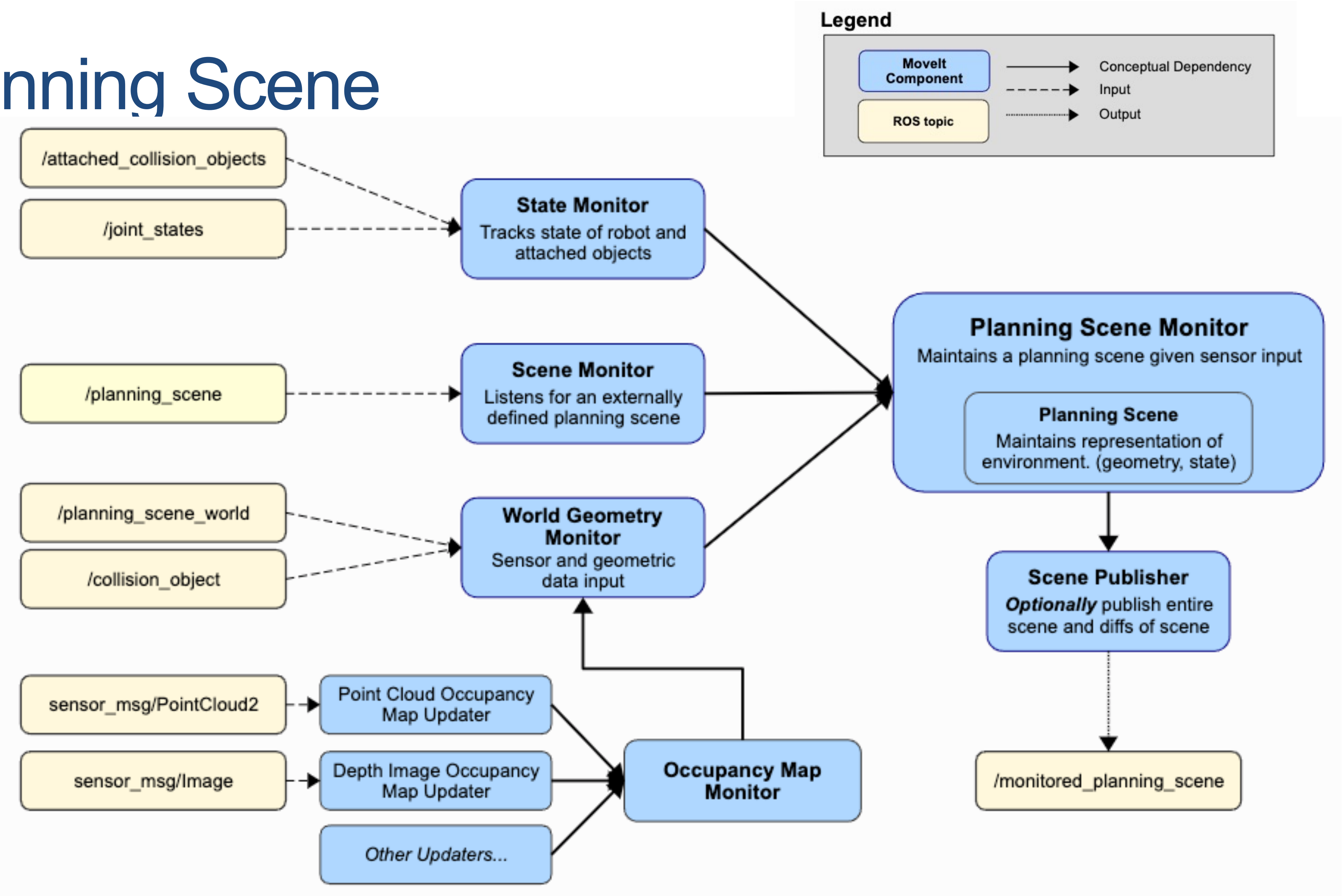
Global Planner	Local Planner
• Solve global solution trajectory • Optimize trajectory path (continuously)	• Follow global reference trajectory • Solve local problem constraints • May process sensor input • Optimize solution locally • Compute controller commands
• Complete • No restricted computation time • Not real-time safe • Not necessarily deterministic	• Can get stuck in local minima • Low computation time • Realtime-safe (depends on solver) • Deterministic
• OMPL planner • STOMP • TrajOpt • Cartesian motion planner • Pilz Industrial Motion Planner • MTC	• IK solver, Jacobian • Potential field planner • Trajectory optimization algorithm • Model Predictive Control (MPC) • Sensor-based Optimal Control

Hybrid Planning: Global and Local





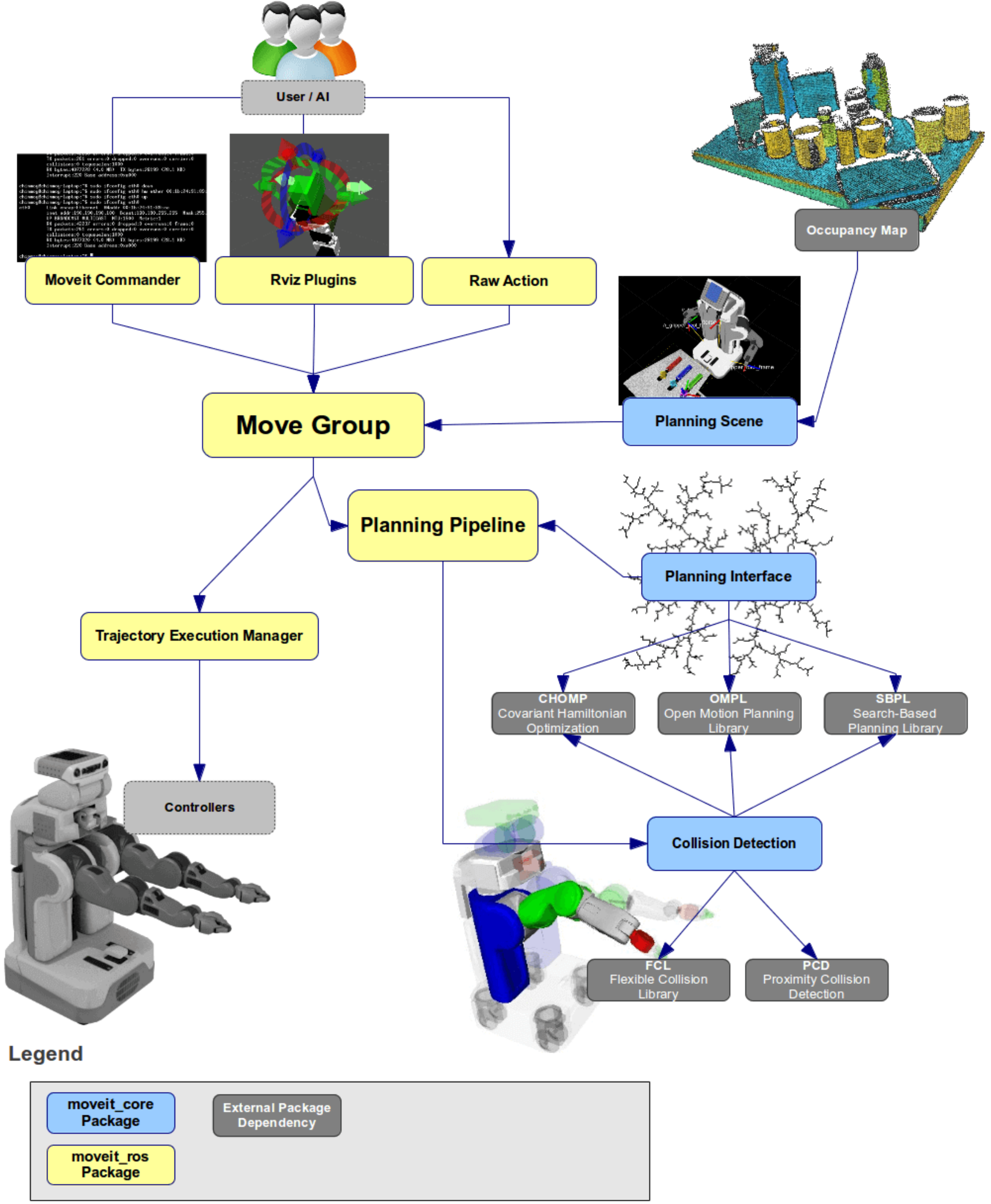
Planning Scene



MoveIt!

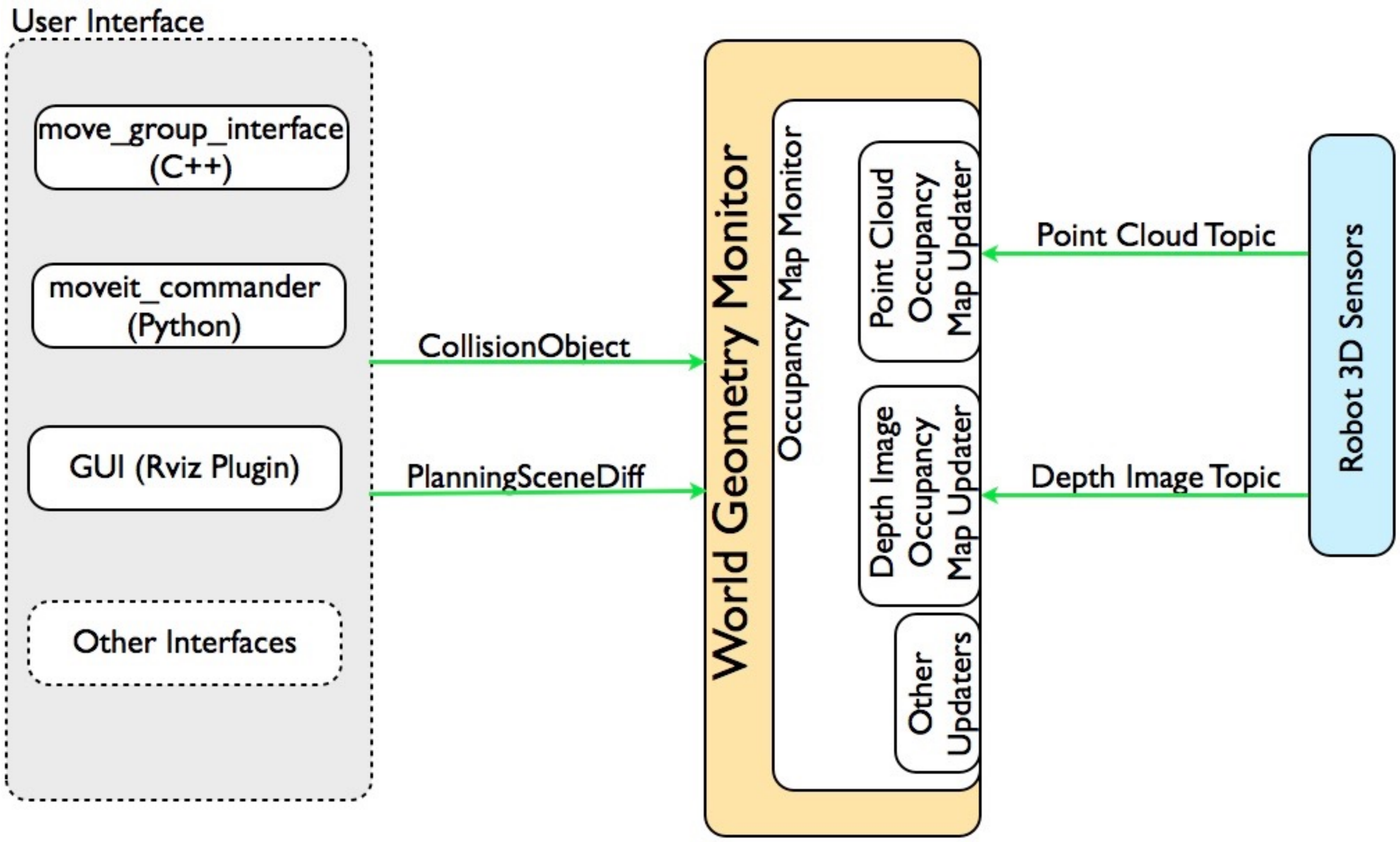
- Configuration:

```
my_robot_moveit_config
    config/
        kinematics.yaml
        joint_limits.yaml
        *_planning.yaml
        moveit_controllers.yaml
        moveit_cpp.yaml
        sensors_3d.yaml
        ...
    launch/
    .setup_assistant
    CMakeLists.txt
    package.xml
```



3D Perception & Collision Checking

- Octomap



Simple C++

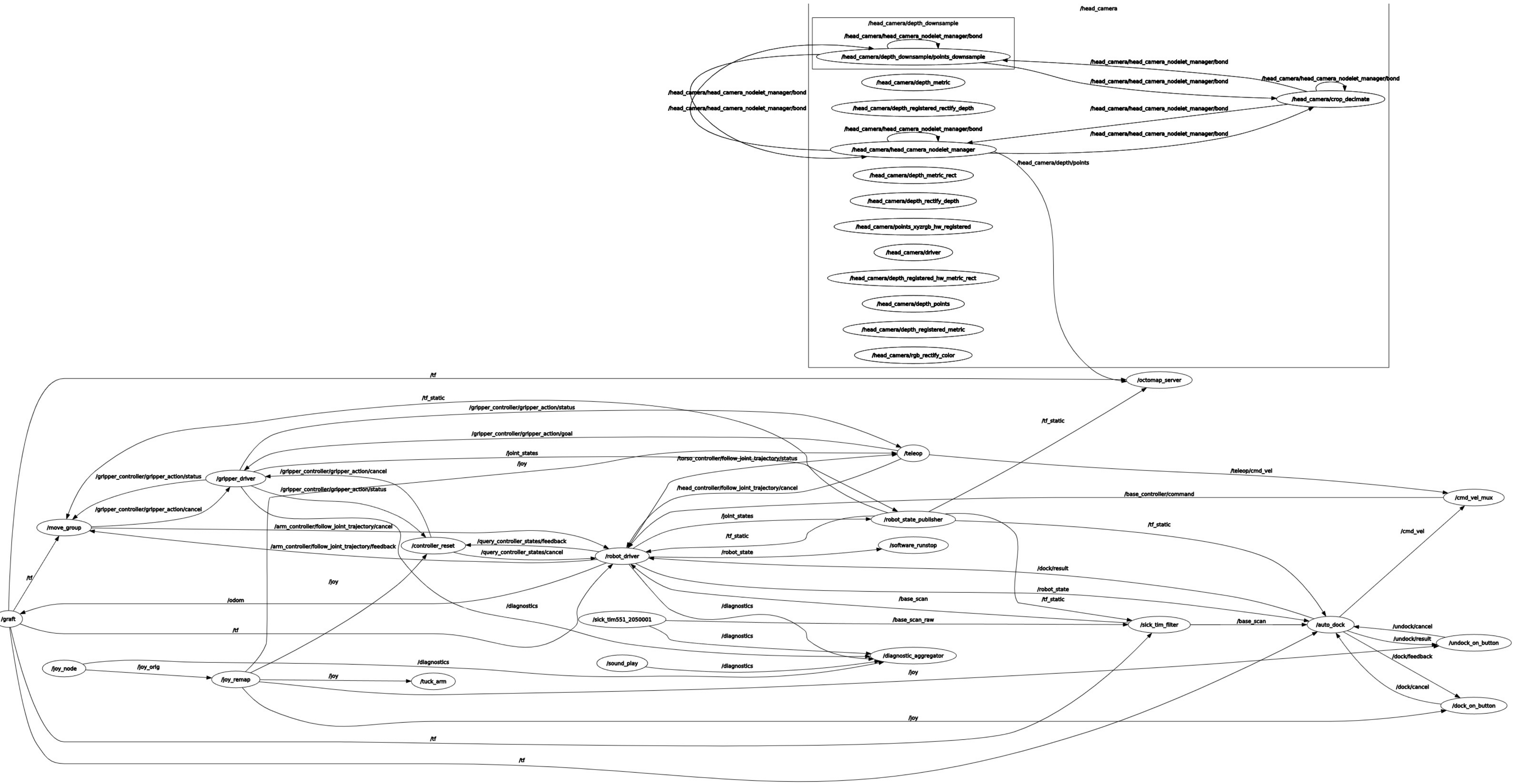
```cpp
// Create the MoveIt MoveGroup Interface
using moveit::planning_interface::MoveGroupInterface;
auto move_group_interface = MoveGroupInterface(node, "manipulator");

// Set a target Pose
auto const target_pose = []{
  geometry_msgs::msg::Pose msg;
  msg.orientation.w = 1.0;
  msg.position.x = 0.28;
  msg.position.y = -0.2;
  msg.position.z = 0.5;
  return msg;
}();
move_group_interface.setPoseTarget(target_pose);

// Create a plan to that target pose
auto const [success, plan] = [&move_group_interface]{
  moveit::planning_interface::MoveGroupInterface::Plan msg;
  auto const ok = static_cast<bool>(move_group_interface.plan(msg));
  return std::make_pair(ok, msg);
}();

// Execute the plan
if(success) {
  move_group_interface.execute(plan);
} else {
  RCLCPP_ERROR(logger, "Planning failed!");
}
```

EXAMPLE

Fetch Robot; ROS 1

- Examples from running robot

- Without localization, SLAM, mapping navigation, task or mission planning


/head_camera
/head_camera/depth_downsample
/head_camera/head_camera_nodelet_manager/bond
/head_camera/depth_downsample/points_downsample
/head_camera/depth_metric
/head_camera/depth_registered_rectify_depth
/head_camera/head_camera_nodelet_manager
/head_camera/crop_decimate
/head_camera/depth/points
/head_camera/depth_metric_rect
/head_camera/depth_rectify_depth
/head_camera/points_xyzrgb_hw_registered
/head_camera/driver
/head_camera/depth_registered_hw_metric_rect
/head_camera/depth_points
/head_camera/depth_registered_metric
/head_camera/rgb_rectify_color
/tf
/octomap_server
/tf_static
/gripper_controller/gripper_action/status
/gripper_controller/gripper_action/goal
/teleop
/joint_states
/torso_controller/follow_joint_trajectory/status
/joy
/gripper_controller/gripper_action/cancel
/gripper_driver
/teleop/cmd_vel
/head_controller/follow_joint_trajectory/cancel
/base_controller/command
/cmd_vel_mux
/move_group
/robot_state_publisher
/arm_controller/follow_joint_trajectory/cancel
/controller_reset
/query_controller_states/feedback
/query_controller_states/cancel
/software_runstop
/arm_controller/follow_joint_trajectory/feedback
/robot_driver
/robot_state
/cmd_vel
/dock/result
/odom
/diagnostics
/base_scan
/graft
/sick_tim551_2050001
/base_scan_raw
/sick_tim_filter
/auto_dock
/undock/cancel
/undock/result
/undock_on_button
/dock/feedback
/diagnostic_aggregator
/sound_play
/joy_node
/joy_orig
/joy_remap
/tuck_arm
/dock/cancel
/dock_on_button

Nodes

- 35 different nodes running
- Without any navigation for the mobile base! (at least +15)

```
/auto_dock
/cmd_vel_mux
/controller_reset
/diagnostic_aggregator
/dock_on_button
/graft
/gripper_driver
/head_camera/crop_decimate
/head_camera/depth_downsample/points_downsample
/head_camera/depth_metric
/head_camera/depth_metric_rect
/head_camera/depth_points
/head_camera/depth_rectify_depth
/head_camera/depth_registered_hw_metric_rect
/head_camera/depth_registered_metric
/head_camera/depth_registered_rectify_depth
/head_camera/driver
/head_camera/head_camera_nodelet_manager
/head_camera/points_xyzrgb_hw_registered
/head_camera/rgb_rectify_color
/joy_node
/joy_remap
/move_group
/octomap_server
/robot_driver
/robot_state_publisher
/rosout
/rviz_1713876364443101616
/sick_tim551_2050001
/sick_tim_filter
/software_runstop
/sound_play
/teleop
/tuck_arm
/undock_on_button
```

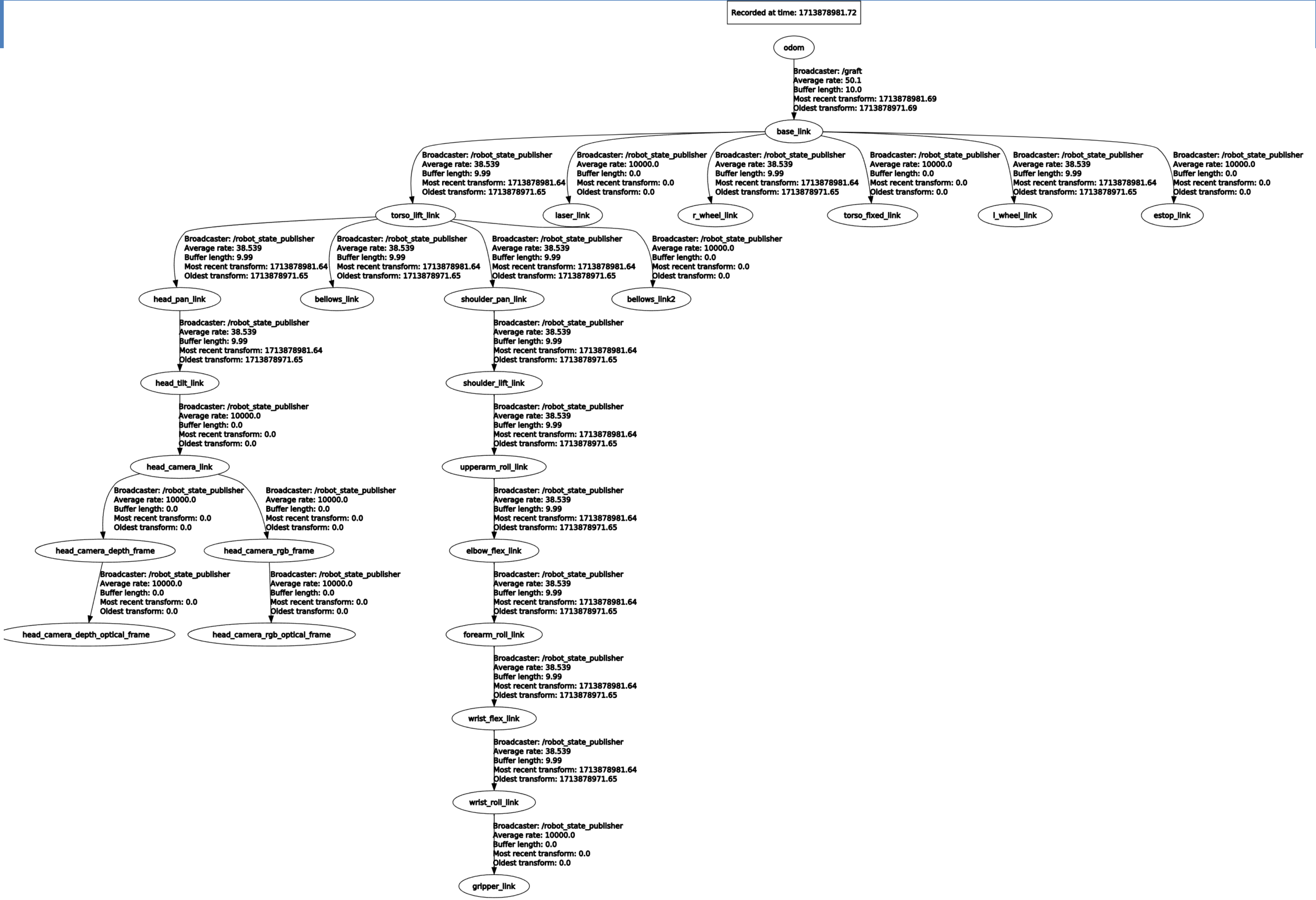
Topics

- 311 different topics
- w/o navigation!

- Each action has topics:
  cancel; feedback; goal; result; status

```
/arm_controller/cartesian_twist/command
/arm_controller/follow_joint_trajectory/cancel
/arm_controller/follow_joint_trajectory/feedback
/arm_controller/follow_joint_trajectory/goal
/arm_controller/follow_joint_trajectory/result
/arm_controller/follow_joint_trajectory/status
/arm_with_torso_controller/follow_joint_trajectory/cancel
/arm_with_torso_controller/follow_joint_trajectory/feedback
/arm_with_torso_controller/follow_joint_trajectory/goal
/arm_with_torso_controller/follow_joint_trajectory/result
/arm_with_torso_controller/follow_joint_trajectory/status
/attached_collision_object
/base_controller/command
/base_scan
/base_scan_no_self_filter
/base_scan_raw
/base_scan_tagged
/battery_state
/charge_lockout/cancel
/charge_lockout/feedback
/charge_lockout/goal
/charge_lockout/result
/charge_lockout/status
/clicked_point
/cmd_vel
/cmd_vel_mux/selected
/collision_object
/diagnostics
/diagnostics_agg
/diagnostics_toplevel_state
/dock/cancel
/dock/feedback
/dock/goal
/dock/result
```

tf tree



Recorded at time: 1713878981.72
odom
Broadcaster: /graft
Average rate: 50.1
Buffer length: 10.0
Most recent transform: 1713878981.69
Oldest transform: 1713878971.69
base_link
Broadcaster: /robot_state_publisher
Average rate: 38.539
Buffer length: 9.99
Most recent transform: 1713878981.64
Oldest transform: 1713878971.65
Broadcaster: /robot_state_publisher
Average rate: 10000.0
Buffer length: 0.0
Most recent transform: 0.0
Oldest transform: 0.0
torso_lift_link
laser_link
r_wheel_link
torso_fixed_link
l_wheel_link
estop_link
head_pan_link
bellows_link
shoulder_pan_link
bellows_link2
head_tilt_link
shoulder_lift_link
head_camera_link
upperarm_roll_link
head_camera_depth_frame
head_camera_rgb_frame
elbow_flex_link
head_camera_depth_optical_frame
head_camera_rgb_optical_frame
forearm_roll_link
wrist_flex_link
wrist_roll_link
gripper_link

SENSORS FOR COLLISION CHECKING

Sensor

- Requirements:
  - Range information
  - Dense data
  - Ideally: also have RGB data to combine computer vision with 3D range sensing
- Main sensor approaches:
  - LiDAR
  - Time of flight cameras
  - Structured light
  - Stereo cameras

Intel RealSense L515

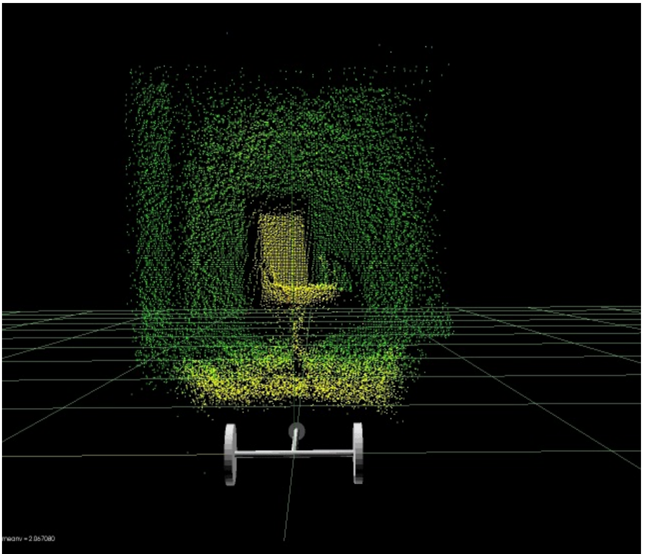
- 9m distance
- Depth: 1024 x 786 pixel @ 30Hz => 23mill pts per second
- RGB: 1920 × 1080 @ 30Hz
- Depth FOV: 70° × 55° (±2°)
- Weight: 100g
- With IMU

- Solid state laser with RGB camera

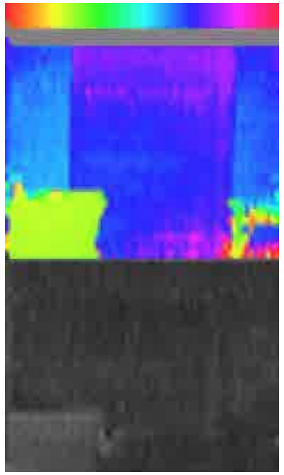
- Sensitive to ambient infrared light => Does NOT work outdoors!

3D Range Sensor: Time Of Flight (TOF) camera

- 3D information with high data rate (100 Hz)
- Compact and easy to manage
- High, non-uniform measurement noise
- High outlier rate at jump edges
- Wrap-around error (phase ranging)
- Sensitive to ambient infrared light => Does NOT work outdoors!

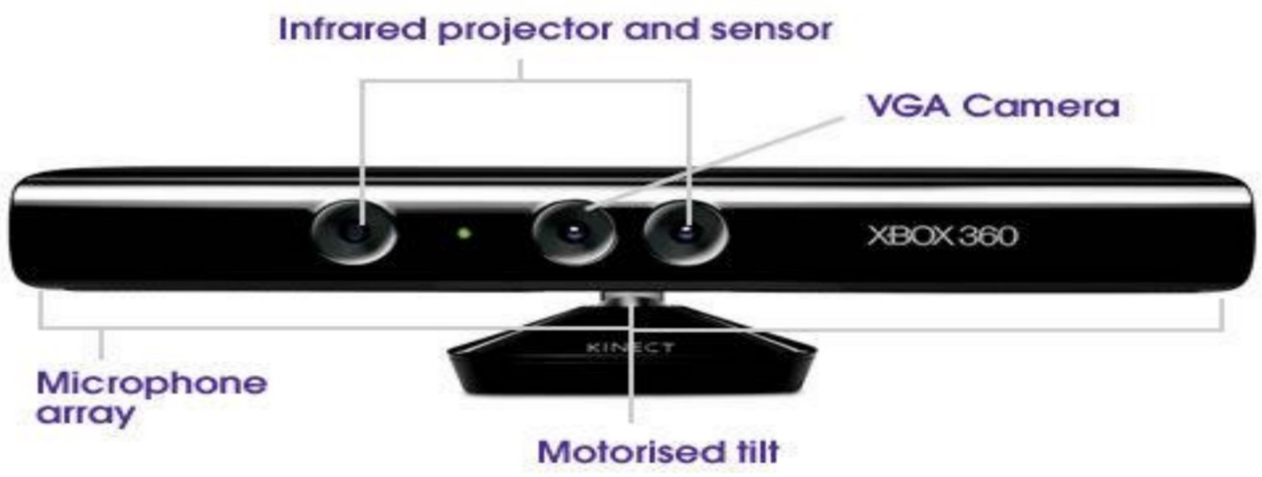


Swiss Ranger 3000 (produced by MESA)



- Kinect 2
- Resolution 1920x1080 pixels
- Field of view: 70 deg (H), 60 deg (V)
- Claimed accuracy: 1 mm
- Claimed max range: 6 meters

PrimeSense Cameras

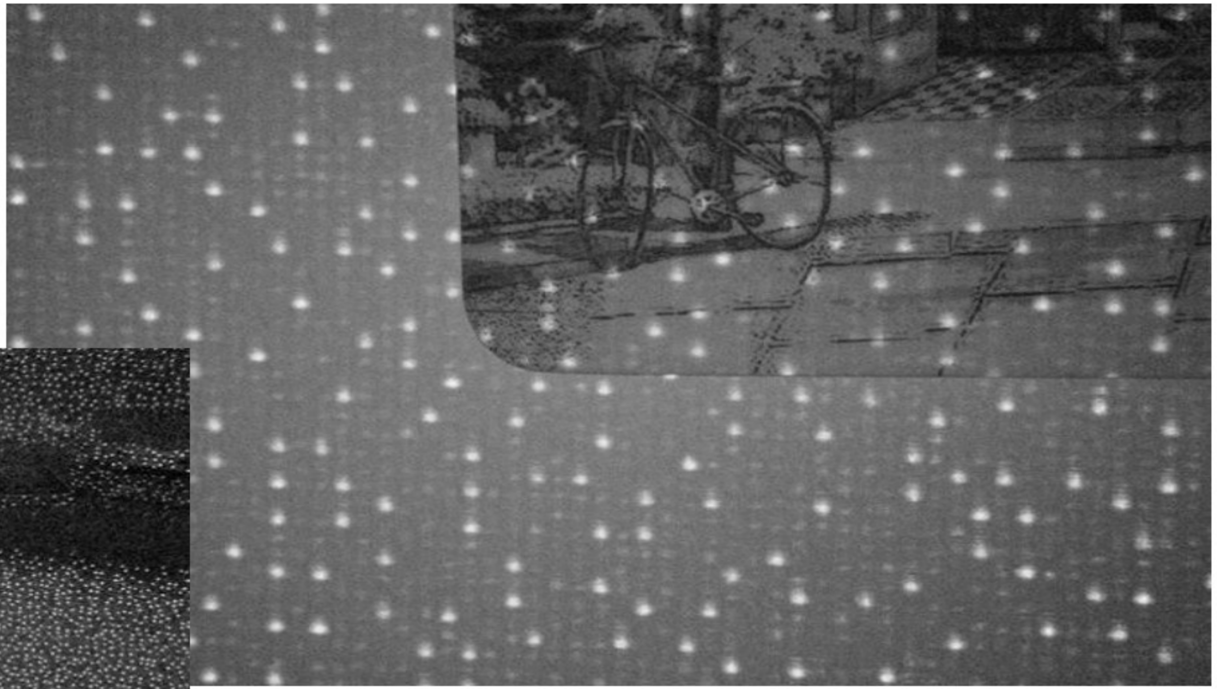


- Devices: Microsoft Kinect and Asus Xtion
- Developed by Israeli company PrimeSense in 2010

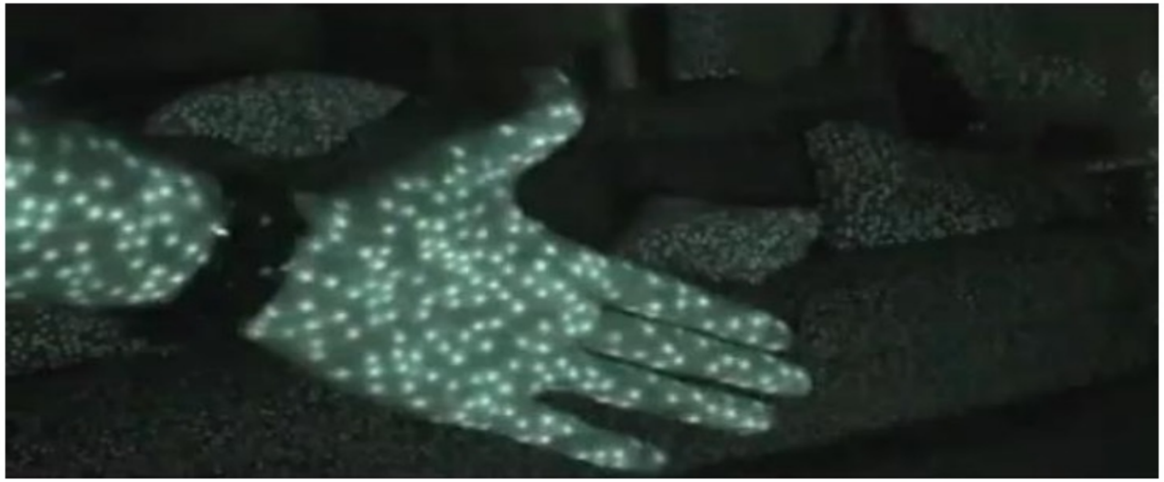
- Components:
  - IR camera (640 x 480 pixel)
  - IR Laser projector
  - RGB camera (640 x 480 or 1280 x 1024)
  - Field of View (FoV):
    - 57.5 degrees horizontally,
    - 43.5 degrees vertically

IR Pattern

Sensitive to infrared light
=>
Does NOT work outdoors!



Depth Map

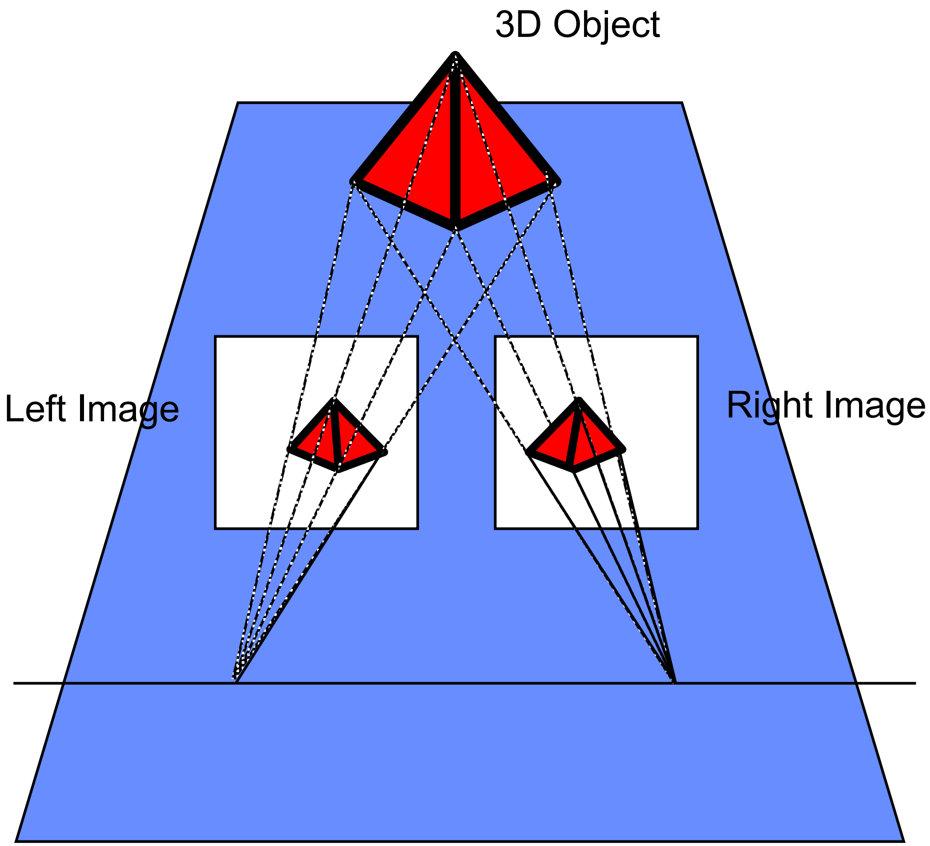


Microsoft Kinect: Depth Computation

- **Depth from Stereo**
  - The Kinect uses an infrared projector and an infrared sensor; it does not use its RGB camera for depth computation
  - The technique of analyzing a known pattern is structured light
  - The IR projector projects a pseudo-random pattern across the surface of the room.
  - The direction of each speckle of the patter is known (from pre calibration during manufacturing) and is hardcoded into the memory of the Kinect
  - By measuring the position of each speckle in the IR image, its depth can be computed

Stereo Cameras

- Theory will be covered in detail in Vision Lecture
- Estimate depth by using 2 cameras

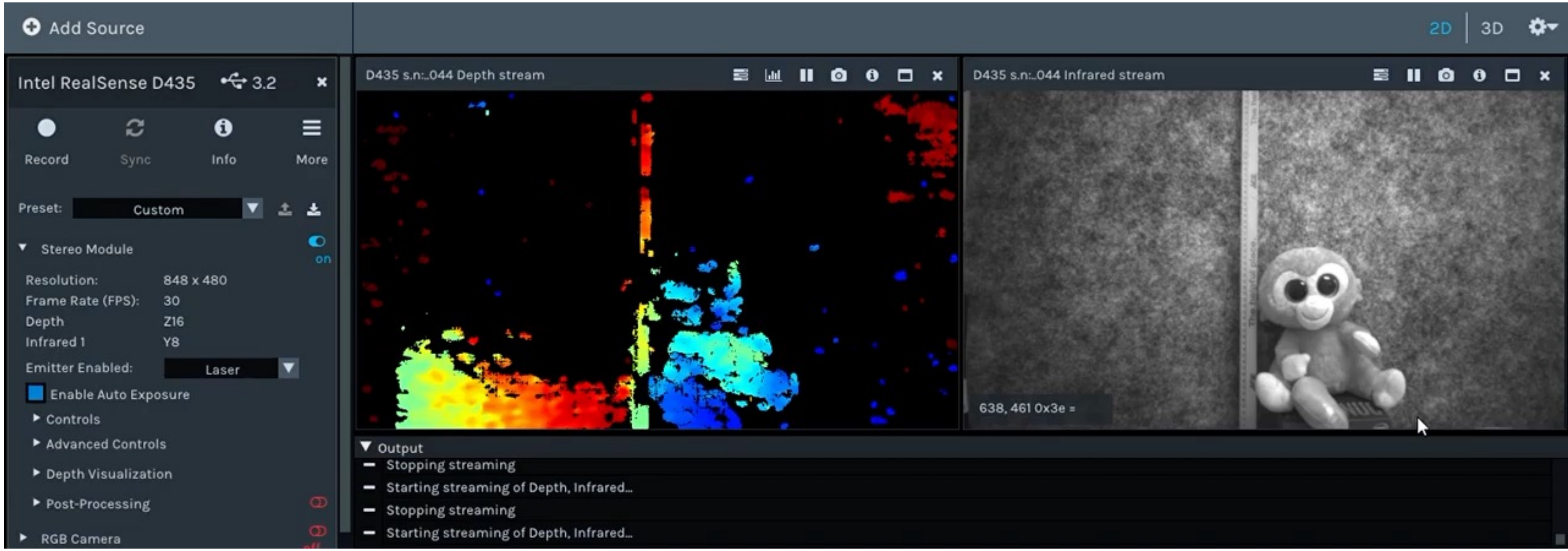


Stereo example: ZED camera

- Dual 4MP Camera @15Hz (lower resolution => higher fps)
- Up to 20m distance
- Passive Sensor
- Doesn’t work on single color surfaces (e.g. white wall)!

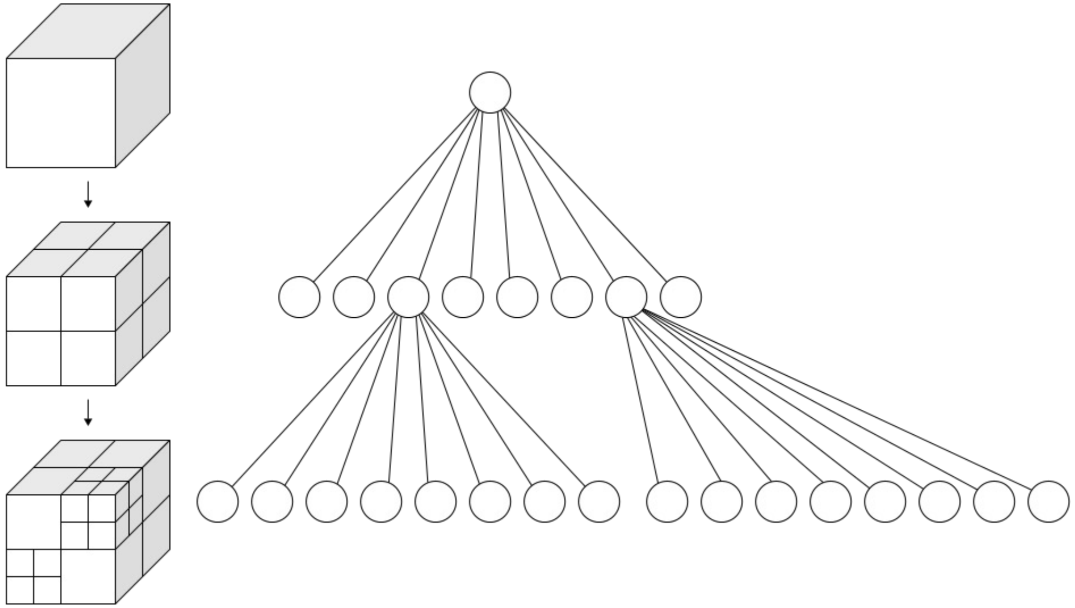
Intel RealSense D435

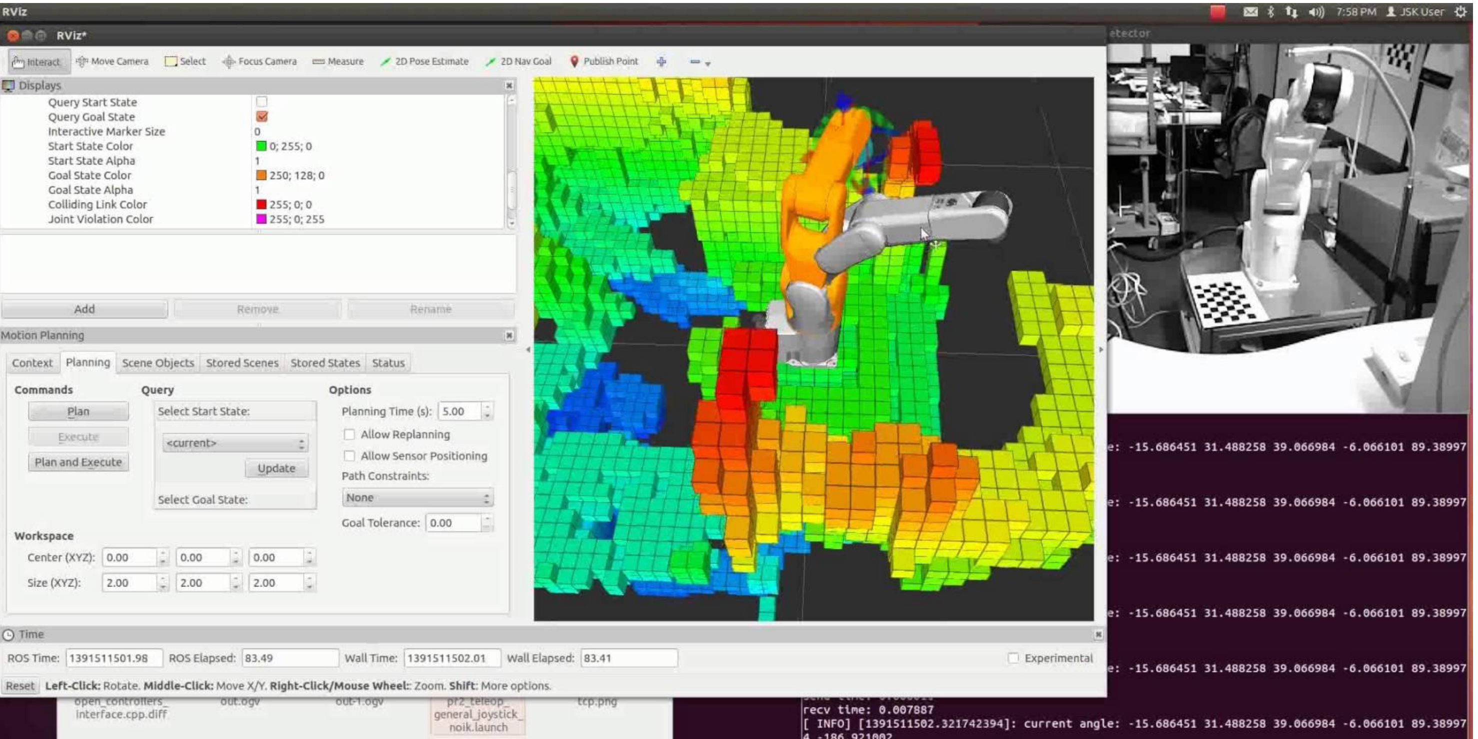
- Stereo Infrared – works indoors and outdoors
- Active pattern (e.g. for white wall) – only works indoors!
- Depth resolution: 1280 × 720
- Depth Field of View (FOV): 86° × 57° (±3°)
- RGB: 1920 × 1080
- Small and l
  ightweight
- With IMU



Octomap / Octree

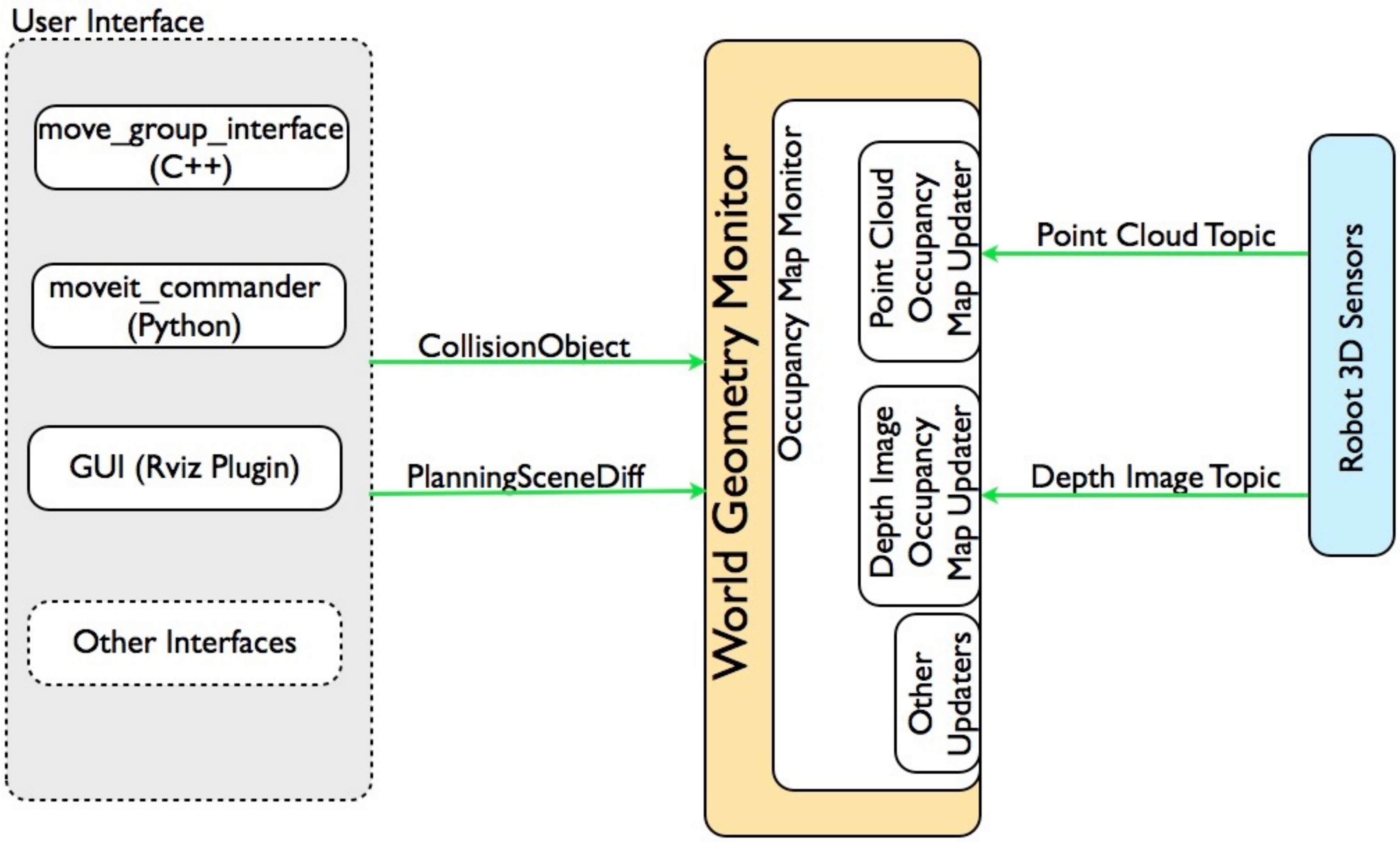
- Put 3d range sensor data in clever datastructure
- http://wiki.ros.org/octomap





3D Perception & Collision Checking

- Octomap



Mobile Manipulation