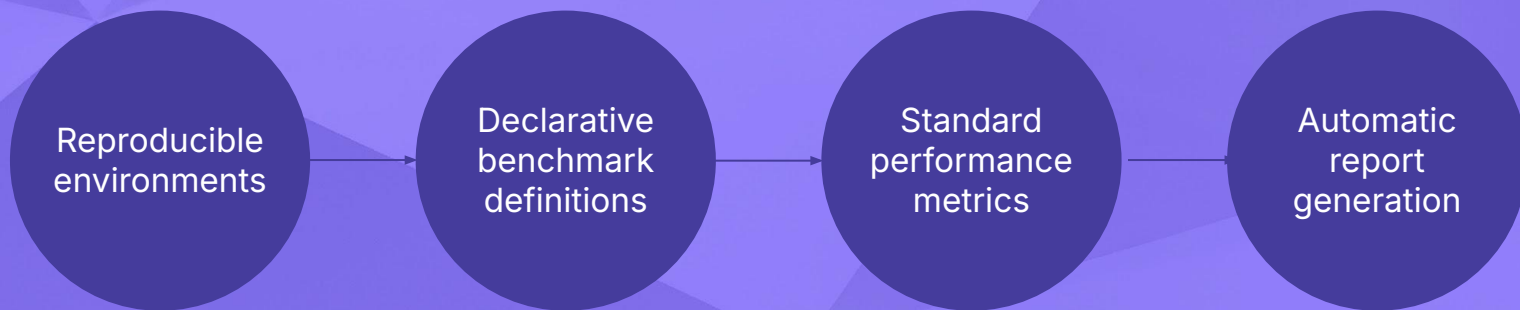# lambkin
benchmarking toolbox for localization

Localization **A**nd **M**apping **B**enchmar**KIN**g toolkit

Localization And Mapping BenchmarKINg Toolkit

Reproducible environments → Declarative benchmark definitions → Standard performance metrics → Automatic report generation

architecture
dependencies and orchestration

Architecture

Reproducible environments

Earthly

Docker

Architecture

Standard performance metrics

Earthly

evo

timemory

Docker

ROS 2

Architecture

Declarative benchmark definitions

Robot Framework

Earthly

evo

timemory

Docker

ROS

Architecture

Automatic report generation

Robot Framework

Earthly

evo

timemory

latex

Docker

ROS

Architecture

User defined benchmarks

lambkin-shepherd   lambkin-clerk
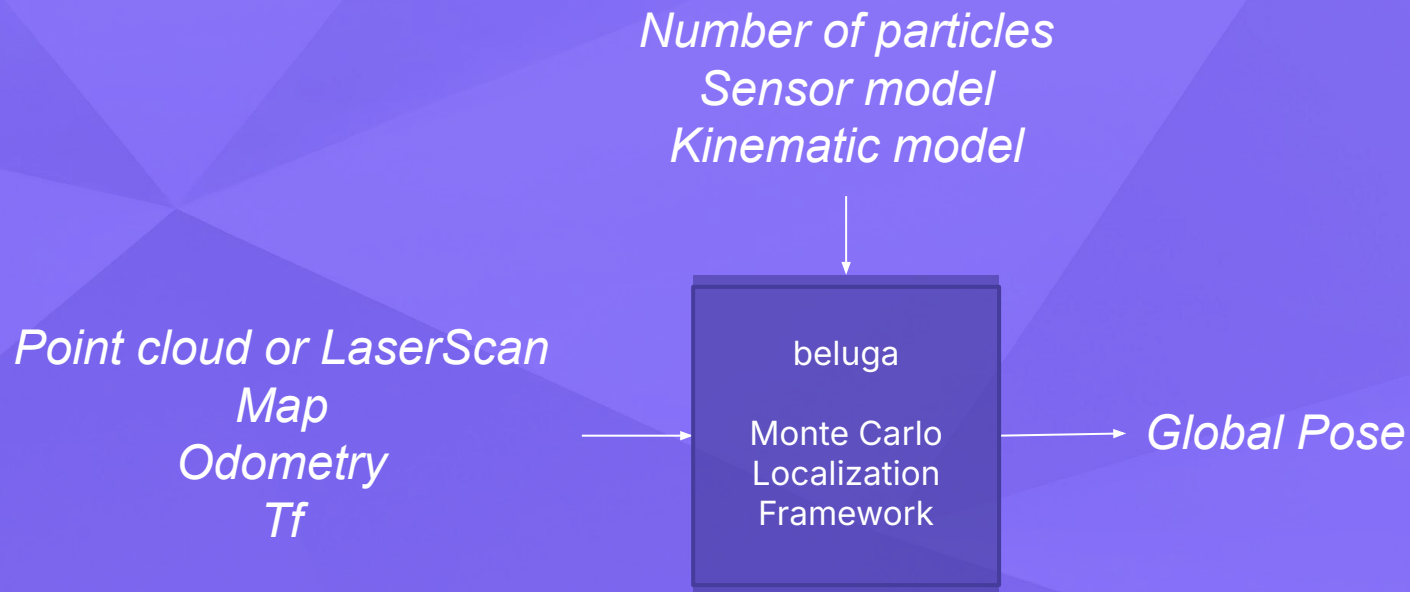
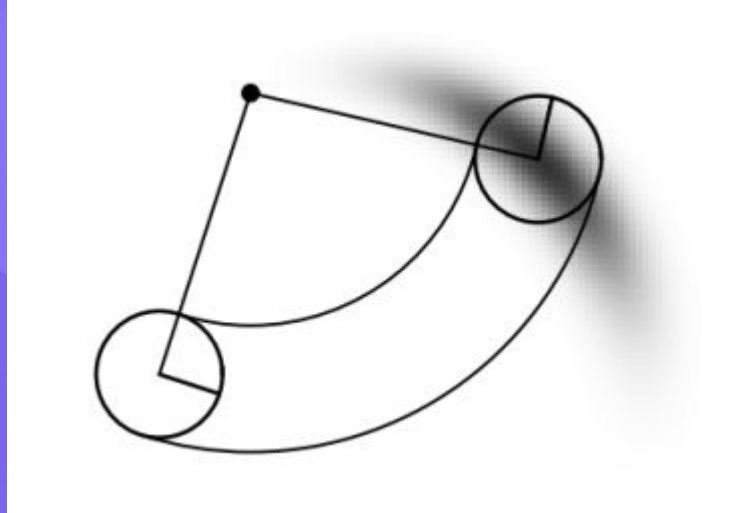| Robot Framework | Earthly | evo | timemory | latex |

| Docker | ROS |

use case
beluga amcl

Localization system with beluga

Number of particles
Sensor model
Kinematic model

↓

Point cloud or LaserScan
Map
Odometry
Tf

→

beluga

Monte Carlo
Localization
Framework

→ *Global Pose*

Localization system with beluga

benchmark
properties
syntax
reports

EKUMEN

benchmarks
Tests - Variations - Iterations - Metrics

```
*** Settings ***
Documentation        Nominal Beluga AMCL vs Nav2 AMCL benchmark using 2D datasets.

Resource             lambkin/shepherd/robot/resources/all.resource

Suite Setup          Setup Beluga vs Nav2 benchmark suite
Suite Teardown       Teardown Beluga vs Nav2 benchmark suite
Test Template        Run Beluga vs Nav2 benchmark case for each ${dataset} ${laser_model}
```

benchmarks
Tests - **Variations** - Iterations - Metrics


```
*** Variables ***
@{LASER_MODELS}          beam   likelihood_field


*** Test Cases ***           DATASET                  LASER_MODEL
Hallway Localization         hallway_localization     ${{LASER_MODELS}}
Hallway Return               hallway_return           ${{LASER_MODELS}}
```

benchmarks
Tests - Variations - **Iterations** - Metrics

```
Beluga vs Nav2 benchmark case
    Extends ROS 2 system benchmark case
    Extends ROS 2 2D SLAM system benchmark case
    # Setup benchmark inputs
    ${dataset_path} = Set Variable  ${EXECDIR}/beluga-datasets/${dataset}
    Uses ${dataset_path}/ROS2/${dataset}_bag as input to ROS 2 system
    ${package_share_path} = Find ROS 2 Package  beluga_vs_nav2  share=yes
    ${qos_override_path} =  Join Path  ${package_share_path}  config  qos_override.yml
    Configures QoS overrides from ${qos_override_path} for input to ROS 2 system
    # Setup benchmark rig
    Uses beluga_vs_nav2.launch in beluga_vs_nav2 ROS package as rig
    Sets map_filename launch argument to ${dataset_path}/map.yaml
    Sets laser_model_type launch argument to ${laser_model}
    Sets use_sim_time launch argument to true
    # Setup benchmark analysis
    Tracks /nav2_amcl/pose /beluga_amcl/pose trajectories
    Uses ${dataset_path}/groundtruth.tum as trajectory groundtruth
    Performs trajectory corrections  align=yes  t_max_diff=${0.1}

    Uses 10 iterations
```

benchmarks
Tests - Variations - Iterations - **Metrics**

```
Beluga vs Nav2 benchmark suite
    Extends ROS 2 system benchmark suite
    Extends ROS 2 2D SLAM system benchmark suite
    Generates latexpdf report from nominal_report template in beluga_vs_nav2 ROS 2 package
```

benchmarks

Nav2 Likelihood Sensor Implementations Evaluation
Release 0.1.0

Ekumen Inc.

Jan 30, 2025

CHAPTER
ONE

DATASET: WILLOW GARAGE

1.1 Dataset information

These results are based on a modified version of the TorWIC SLAM dataset, which was repurposed to serve as a 2D localization test scenario. The original dataset is no longer available online, but it was described in the following reference:

"An Object-Based Semantic World Model for Long-Term Change Detection and Semantic Querying.", by Julian Mason and Bhaskara Marthi, IROS 2012.

The dataset was collected using a PR2 omnidirectional mobile robot developed by Willow Garage.

Our mobile robot is a Willow Garage PR2 with a headmounted Microsoft Kinect. When the robot is in motion, the Kinect is roughly 1.5 meters off the floor, and captures frames at 5Hz. The robot travels at 0.5 m/s. We rely on the PR2's other sensors and tools in ROS [19] to provide autonomous navigation and accurate localization, but only the current localization estimate and the Kinect's RGB-D point cloud are provided to our perceptual pipeline, detailed in Section V. As a result, the PR2 is not a prerequisite for our system; we only require a localized RGB-D camera. We do require that the sensor be fairly far from the floor, in order to perceive supporting surfaces and the objects on them.

The dataset provides 67 trajectories recorded over multiple days and time periods. All of the trajectories are in the same office-like environment, and each represents about 15 minutes of robot operation on average. The shortest trajectory is 8.5 minutes long, and the longest is 30 minutes long.

The following is a representative map from the set:

1.2 Transformations to the original dataset

To repurpose the TorWIC Mapping dataset for 2D localization evaluation, we made the following changes from the original dataset as distributed in here:

- The original ROS 1 bagfiles were converted to ROS 2 bagfiles.
- All topics not related to lidar-based 2D localization were removed to reduce file size.
- The map->odom_combined TF transformations were removed from the transform tree.
- The "/" prefix in all transform frame names was removed to avoid issues with the ROS 2 TF tree.
- A reference ground-truth and a matching occupancy map were generated for each bagfile using SLAM.

1

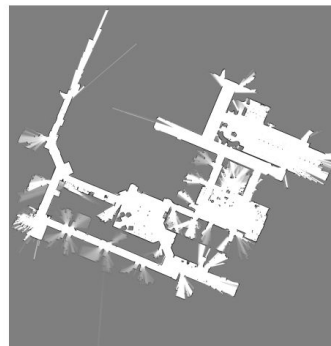Nav2 Likelihood Sensor Implementations Evaluation, Release 0.1.0

Fig. 1: Localization map of a representative example in the Willow Garage dataset.

1.2. Transformations to the original dataset   2

Nav2 Likelihood Sensor Implementations Evaluation, Release 0.1.0

1.3 2011-08-03-20-03-22

Bagfile metadata:

```
Bag size:    447.7 MiB
Storage id:
Duration:    1179.865735696s
Start:       Aug  4 2011 03:03:22.249058964 (1312427802.249058964)
End:         Aug  4 2011 03:23:02.114786654 (1312428182.114786654)
Messages:    90936
```

Evaluation results:

Table 1: APE metrics for 2011-08-03-20-03-22.

Sensor Model	Statistics				
	median	median c.i.	mean	mean c.i.	worst-case
Nav2 LF	0.048 m	0.047-0.050 m	0.059 m	0.057-0.060 m	0.294 m
Nav2 LFP	0.030 m	0.029-0.031 m	0.035 m	0.034-0.036 m	0.206 m
Nav2 LFP + BS	0.030 m	0.029-0.031 m	0.035 m	0.034-0.036 m	0.203 m

Table 2: Average CPU and peak Resident Set Size (RSS) values.

Implementation	Resource usage	
	cpu	rss
Nav2 LF	17.8%	67 MB
Nav2 LFP	18.5%	65 MB
Nav2 LFP + BS	18.2%	67 MB

1.4 2011-08-04-12-16-23

Bagfile metadata:

```
Bag size:    415.2 MiB
Storage id:
Duration:    1071.353156193s
Start:       Aug  4 2011 19:16:24.006902194 (1312485384.006902194)
End:         Aug  4 2011 19:34:15.360058387 (1312486455.360058387)
Messages:    83286
```

Evaluation results:

Table 3: APE metrics for 2011-08-04-12-16-23.

Sensor Model	Statistics				
	median	median c.i.	mean	mean c.i.	worst-case
Nav2 LF	0.054 m	0.052-0.057 m	0.063 m	0.061-0.065 m	0.246 m
Nav2 LFP	0.029 m	0.028-0.031 m	0.036 m	0.035-0.038 m	0.236 m
Nav2 LFP + BS	0.029 m	0.028-0.031 m	0.036 m	0.034-0.037 m	0.247 m

Table 4: Average CPU and peak Resident Set Size (RSS) values.

Implementation	Resource usage	
	cpu	rss
Nav2 LF	18.0%	63 MB
Nav2 LFP	18.6%	62 MB
Nav2 LFP + BS	18.2%	64 MB

1.3. 2011-08-03-20-03-22   3

Scan me ;)

Thank you!

Júlia Marsal Perendreu

Software Developer

julia.marsal@ekumenlabs.com

EKUMEN