Dissimilarity Measures for Clustering Space Mission Architectures

Cody Kinneer

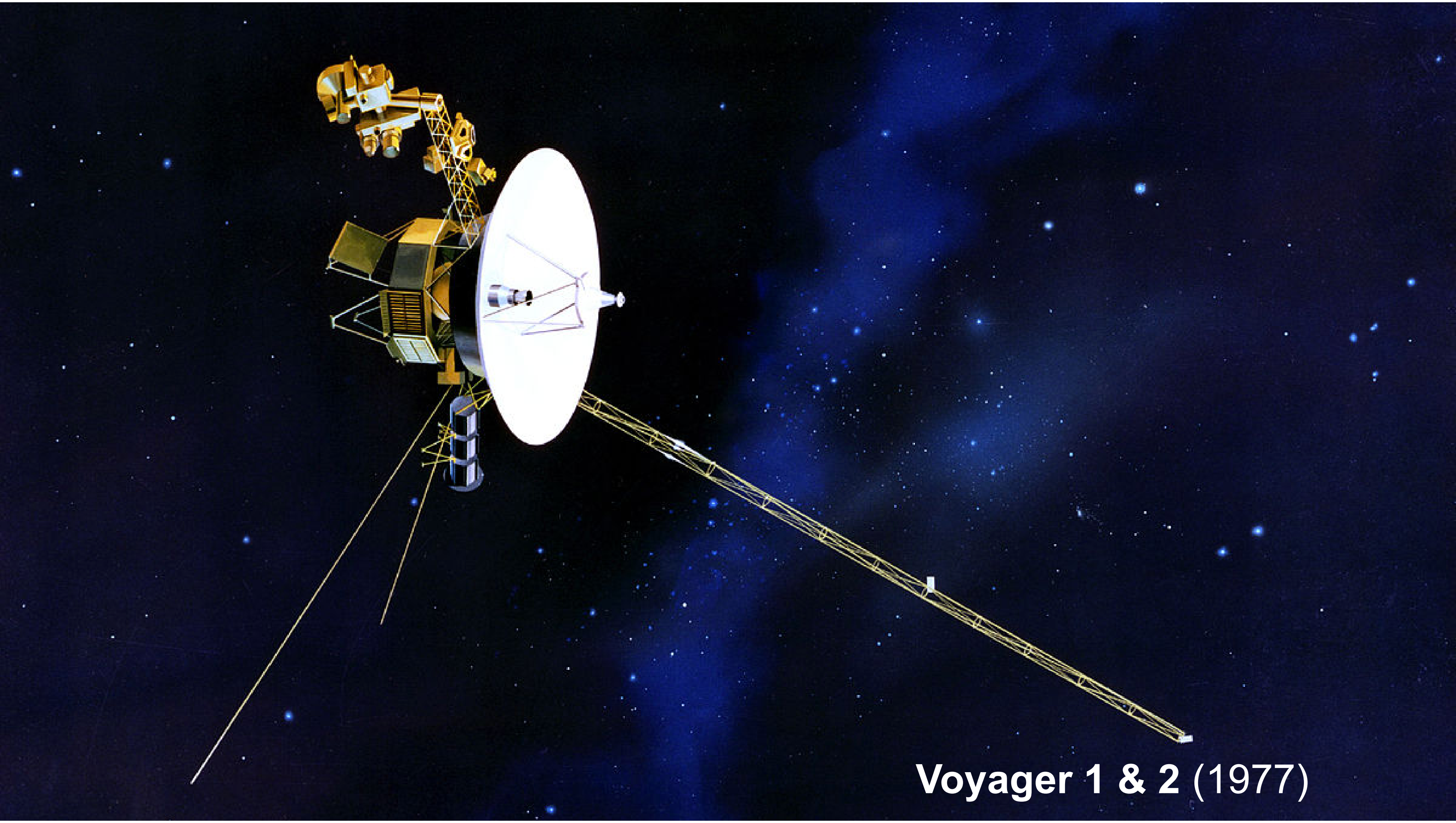
Institute for Software Research, Carnegie Mellon University

Sebastian J. I. Herzig

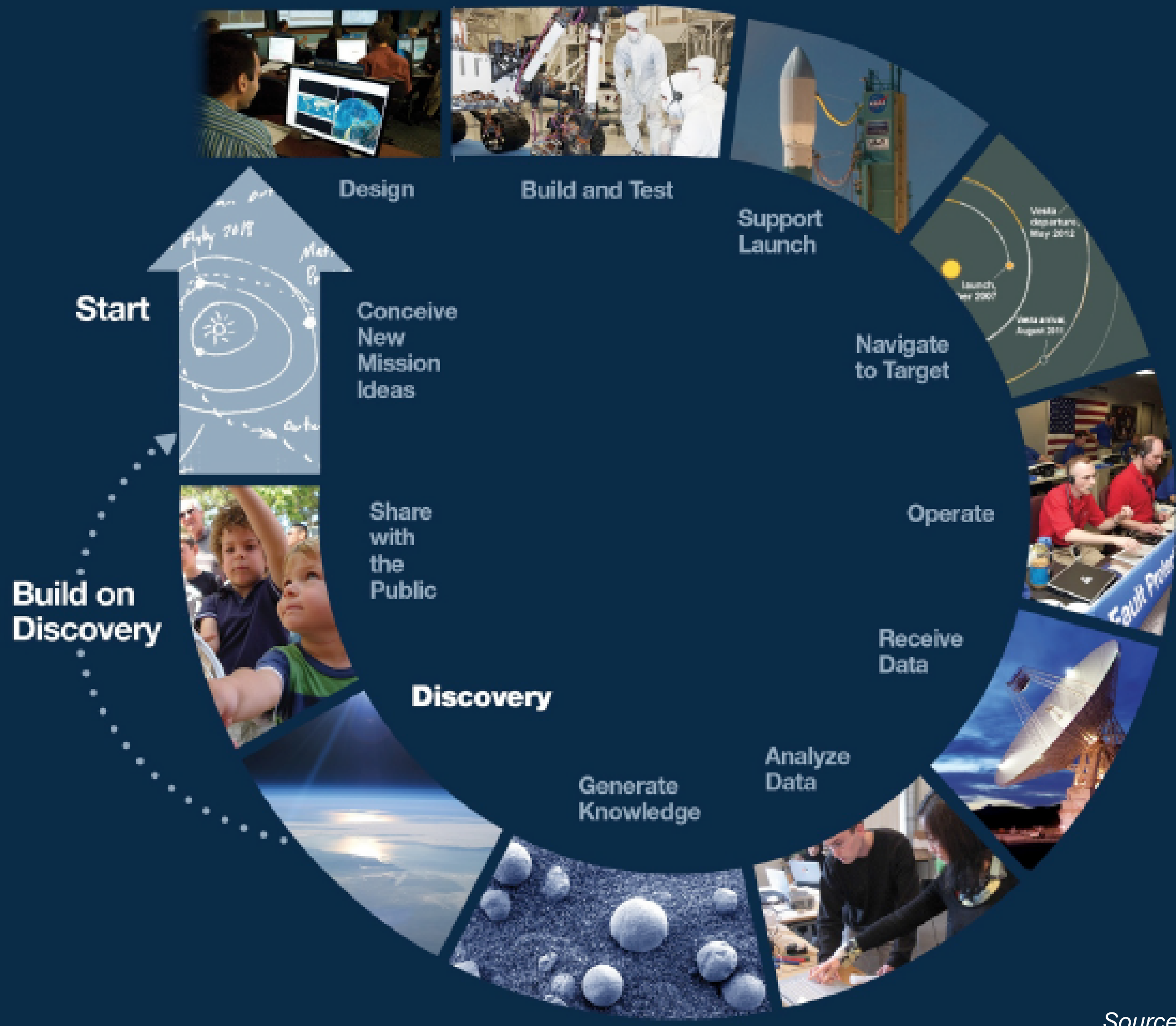
Jet Propulsion Laboratory, California Institute of Technology

18 October 2018 – ACM/IEEE MODELS Conference, Copenhagen, Denmark

Robotic Space Exploration

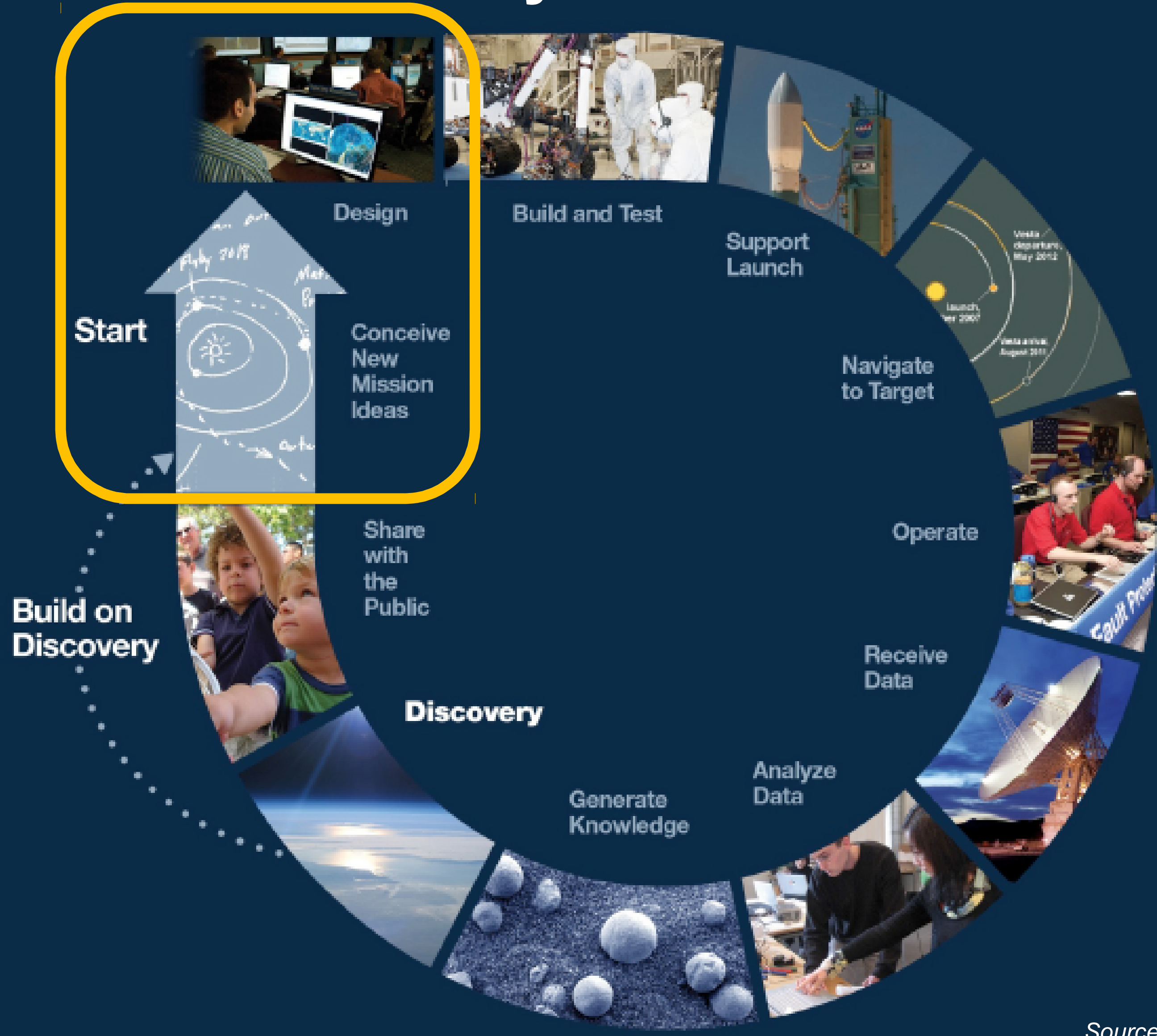


The JPL Product Lifecycle



Source: Nichols & Lin, 2014

The JPL Product Lifecycle

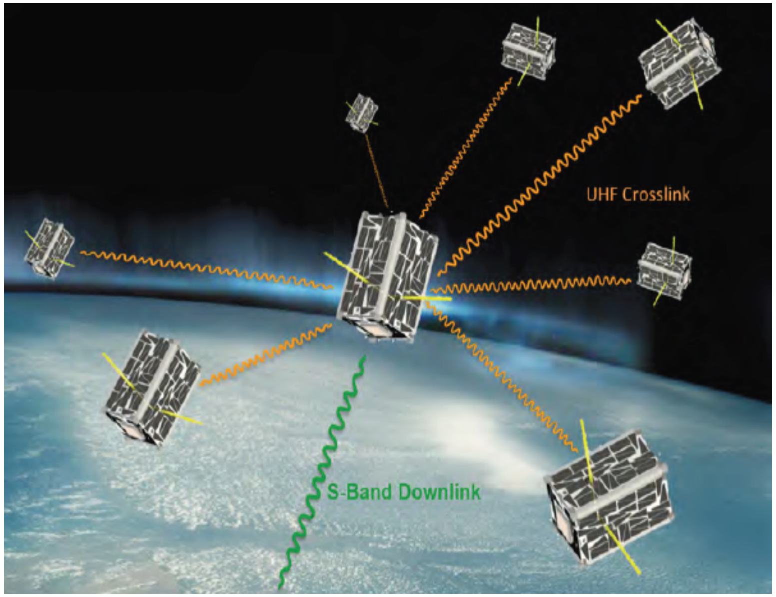


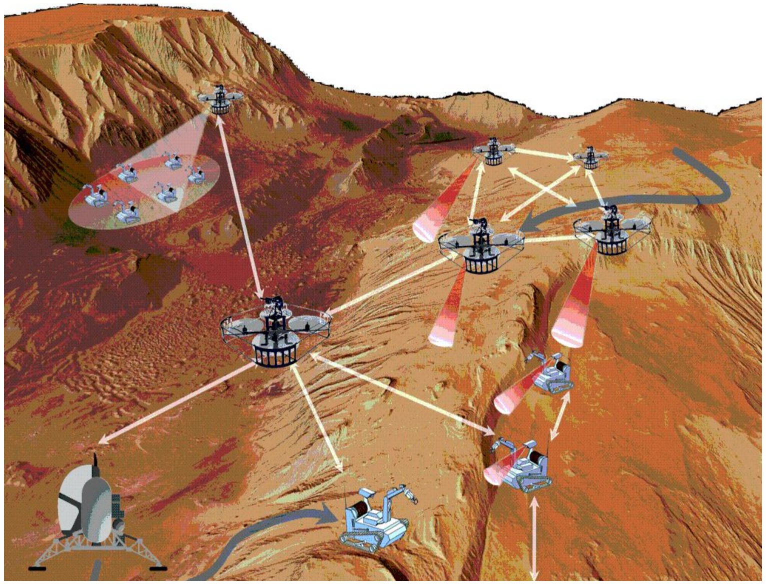
Source: Nichols & Lin, 2014

Networked Constellations of Spacecraft

- Small spacecraft enable innovative low-cost multi-asset missions

- Goal of initiative is to develop new technologies that support novel mission concept proposals





Motivating Case Study

Spacecraft-Based Radio Interferometry

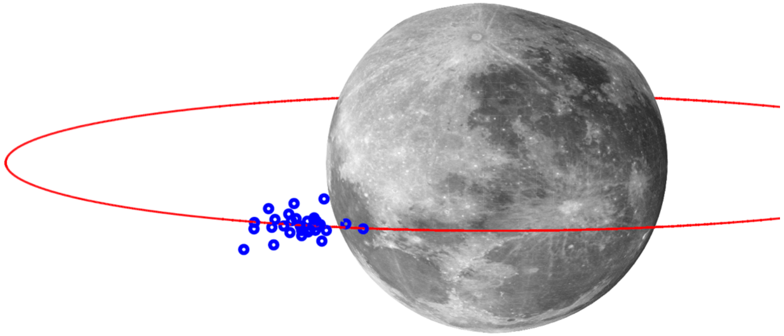
Source: http://www.passmyexams.co.uk/GCSE/physics/images/radio-telescopes-outdoors.jpg

Radio interferometers:

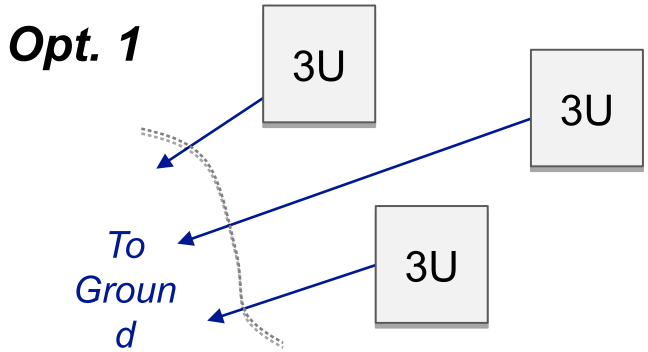
- Radio telescopes consisting of multiple antennas
- Achieve the same angular resolution as that of a single telescope with the same aperture
- Typically ground-based

Want to do this in space:

- Frequencies < 30Mhz blocked by ionosphere
- Cluster of spacecraft (3 – 50) functioning as telescopes in LLO
- CubeSats or SmallSats are promising enablers for this

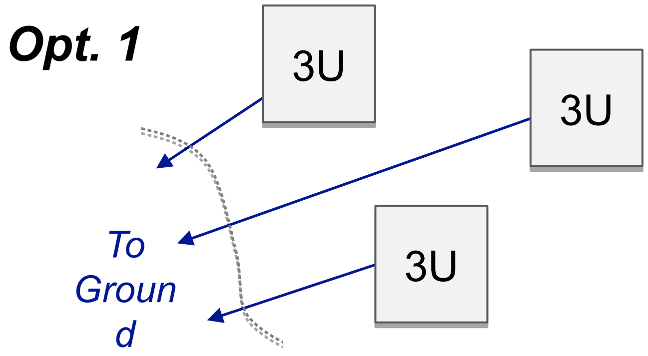


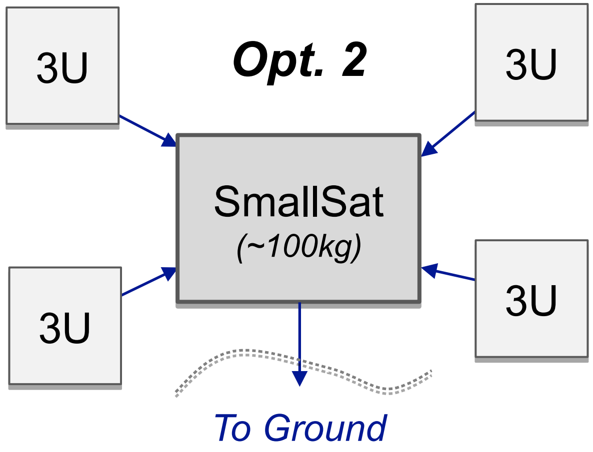
Which Architecture is Optimal?



Challenge: transmit very large data volume from LLO to Earth

Which Architecture is Optimal?

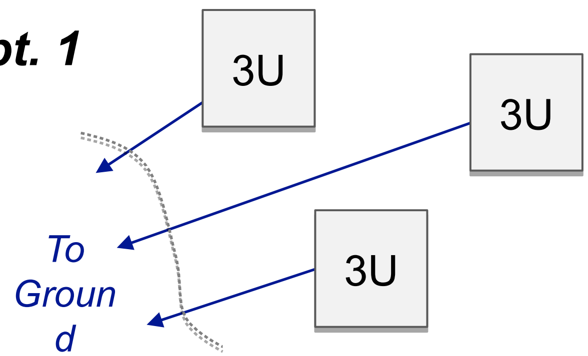




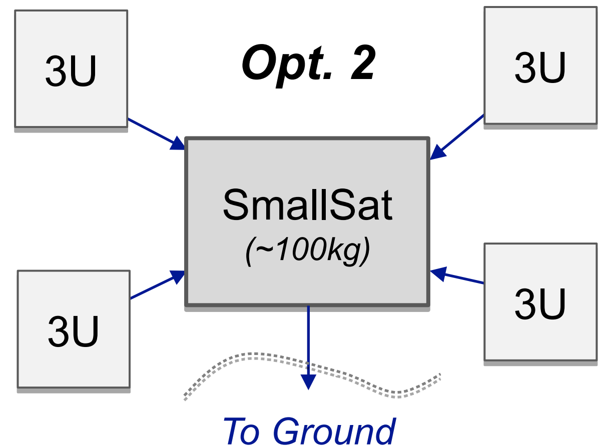
Challenge: transmit very large data volume from LLO to Earth

Which Architecture is Optimal?

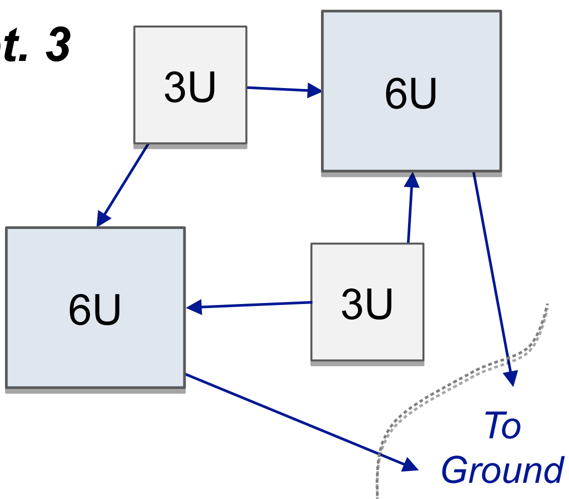
Opt. 1



Opt. 2

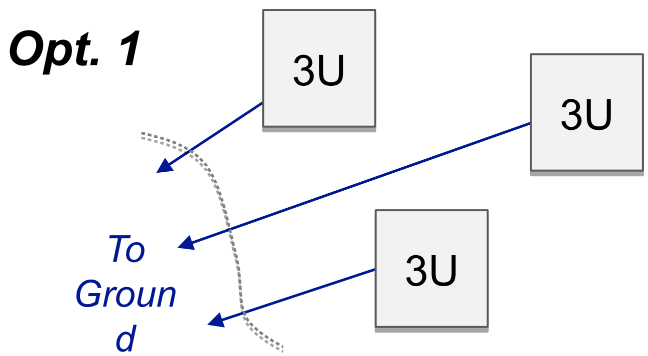


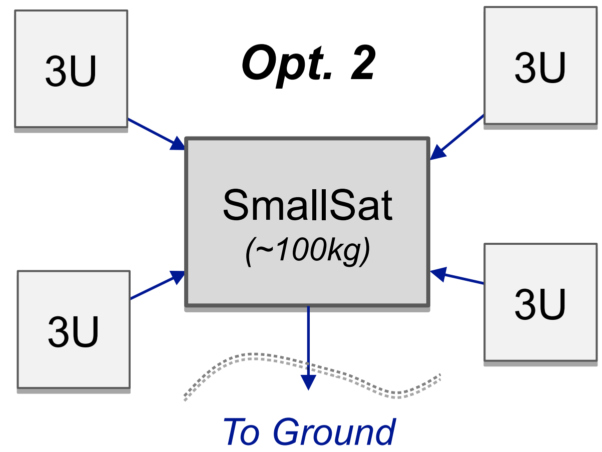
Opt. 3

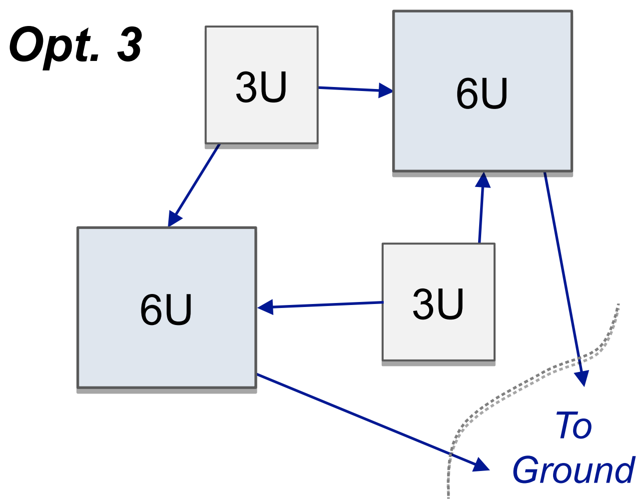


Challenge: transmit very large data volume from LLO to Earth

Which Architecture is Optimal?





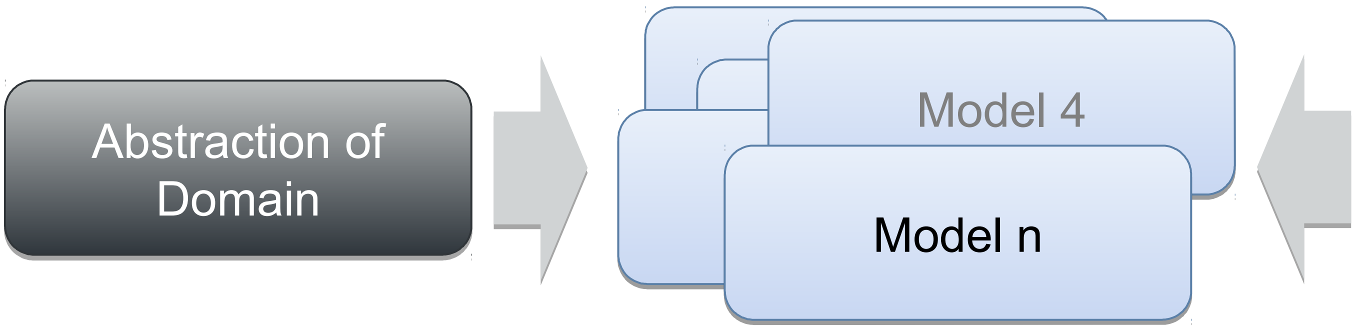


Challenge: transmit very large data volume from LLO to Earth

- How many spacecraft?
- Are all equipped with interferometry payload? Are some just relays?
- Who communicates with Earth?
- What frequency bands? Multi-hop?
- …
- Optimal w.r.t. cost? Science value?

Mission Architecture Trade Space Exploration

Mechanized Exploration



Which interferometry missions are optimal with respect to cost & scientific benefit?

“A constellation mission consists of at least 2 spacecraft and at most 100”

“A spacecraft can, but does not have to contain the interferometry payload”

“Operation of the interferometry payload operation requires power”

Solution Generation
Models in domain

“Constellation mission A with 3 spacecraft, one of which has a payload and solar cells”

Problem Description
Which models in the domain are we looking for?

Mission Architecture Trade Space Exploration

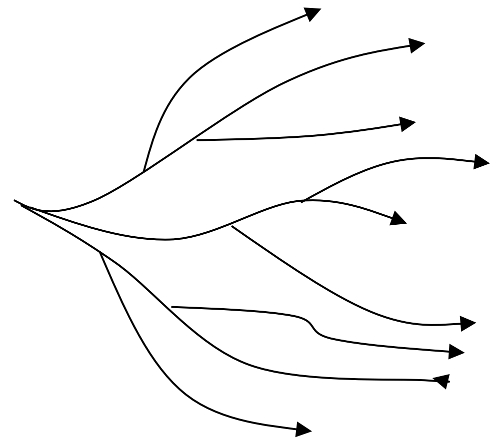
Mechanized Exploration

Abstraction of Domain

"A constellation mission consists of at least 2 spacecraft and at most 100"

"A spacecraft can, but does not have to contain the interferometry payload"

"Operation of the interferometry payload operation requires power"



Solution Search

Models in domain

"Constellation mission A with 3 spacecraft, one of which has a payload and solar cells"

Which interferometry missions are optimal with respect to cost & scientific benefit?

Problem Description

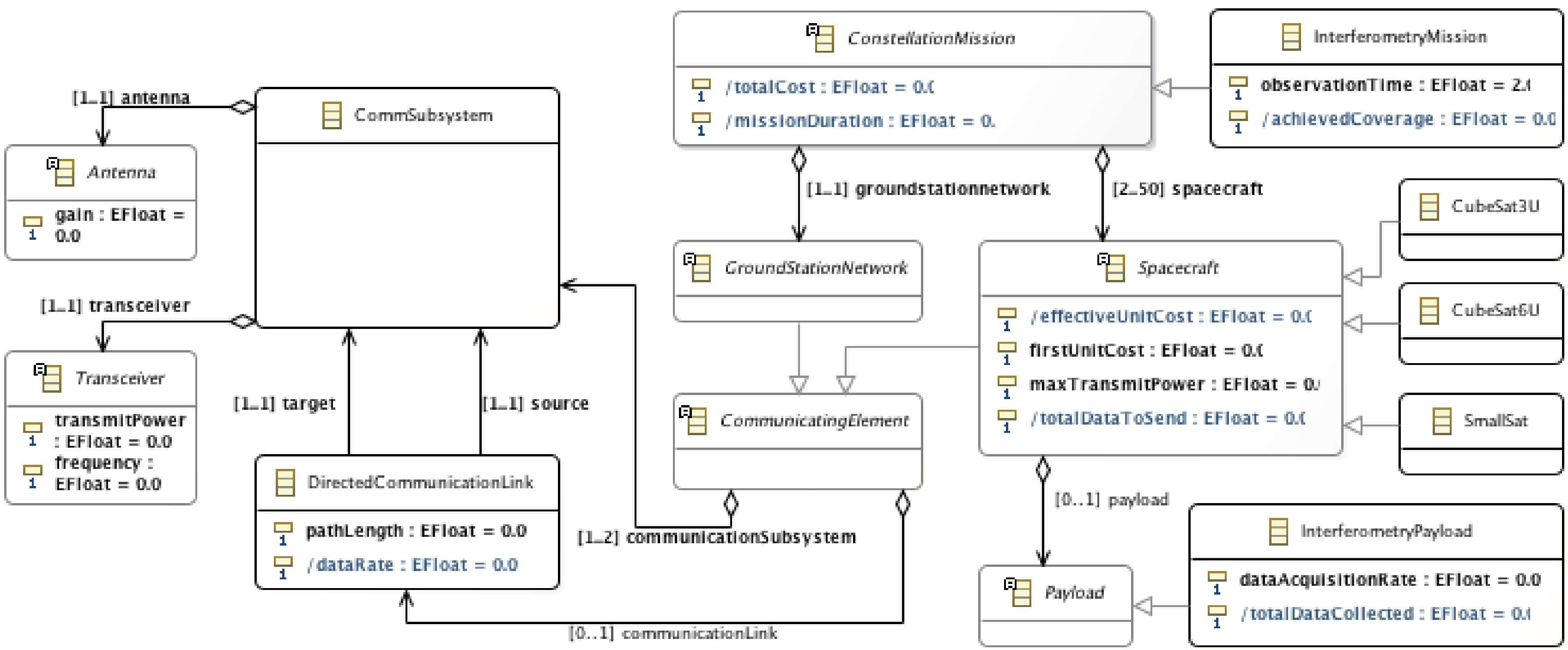
Which models in the domain are we looking for?

In practice, too many possible solutions to generate & compare all

View as a search problem

Application to Case Study

Domain model in Ecore + OCL (Excerpt)

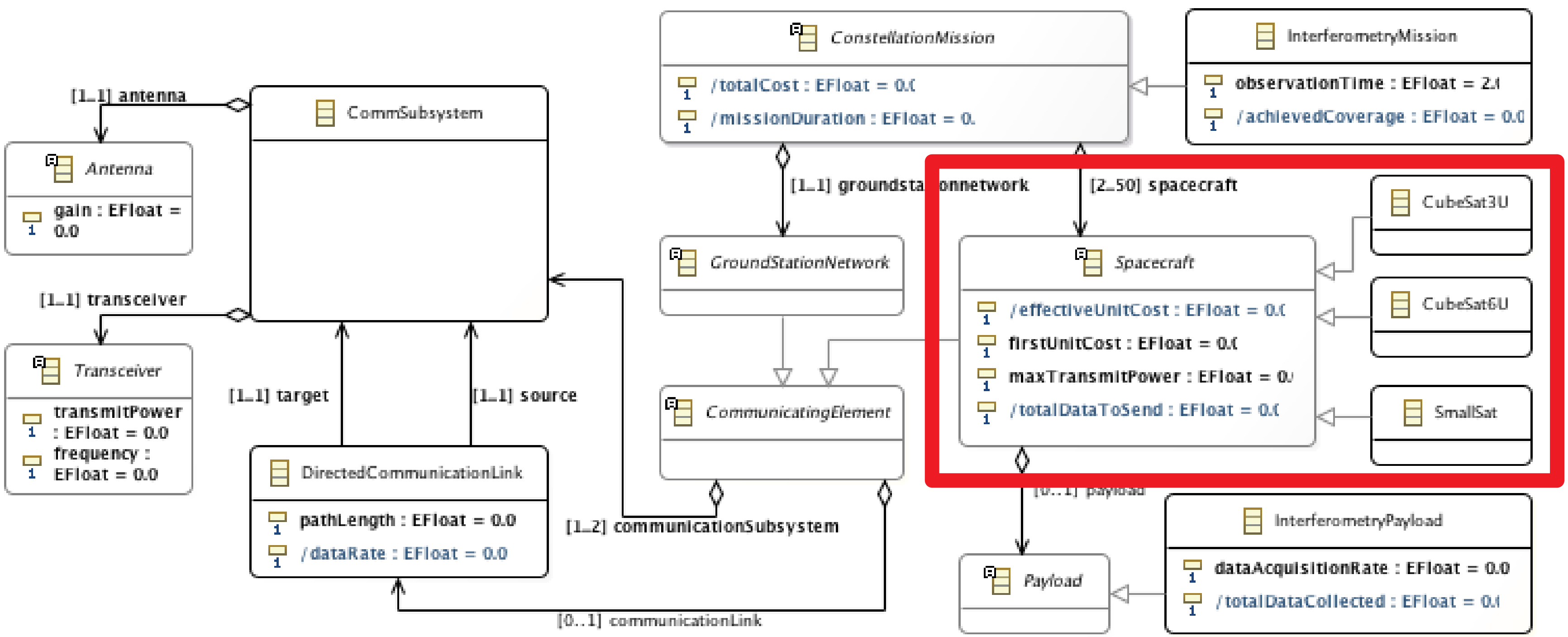


20 concepts, 9 associations, 15 attributes / parameters

> 48[10] possible models

Application to Case Study

Domain model in Ecore + OCL (Excerpt)

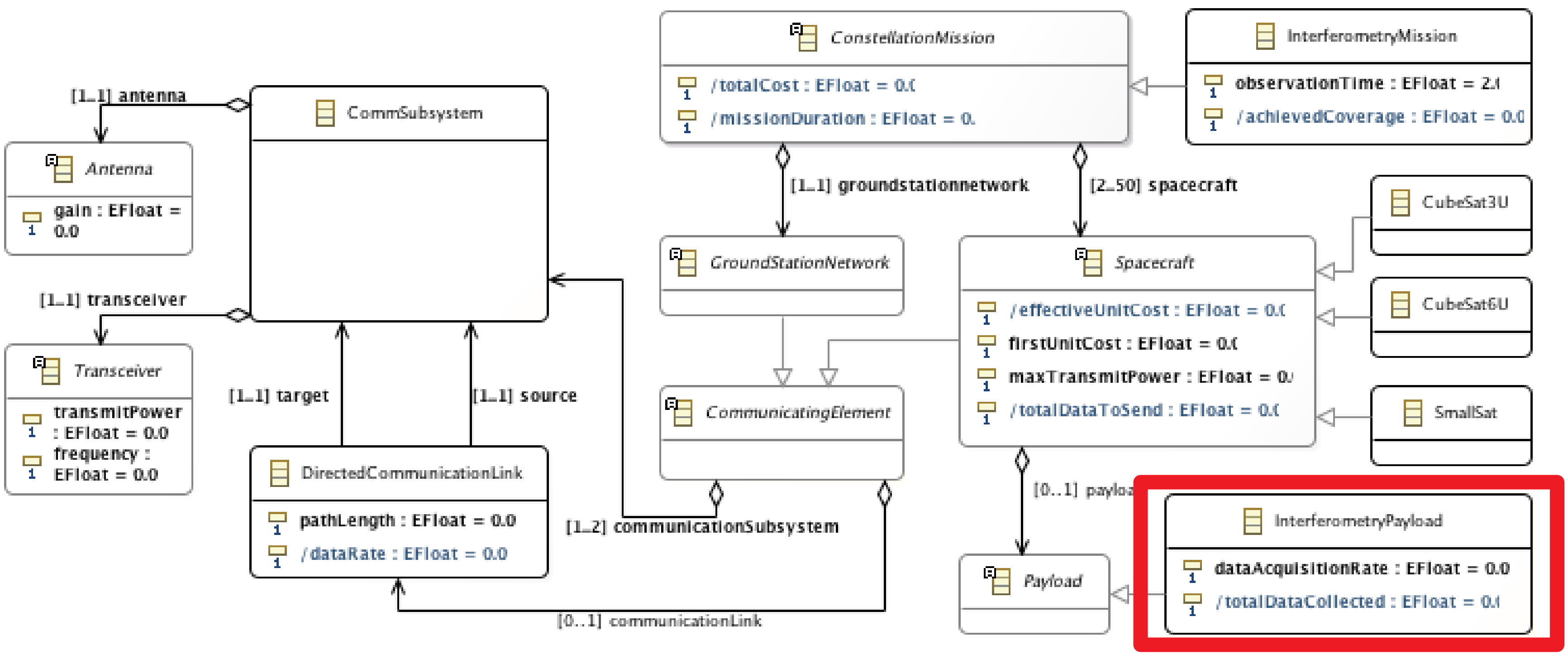


20 concepts, 9 associations, 15 attributes / parameters

> 48^{10} possible models

Application to Case Study

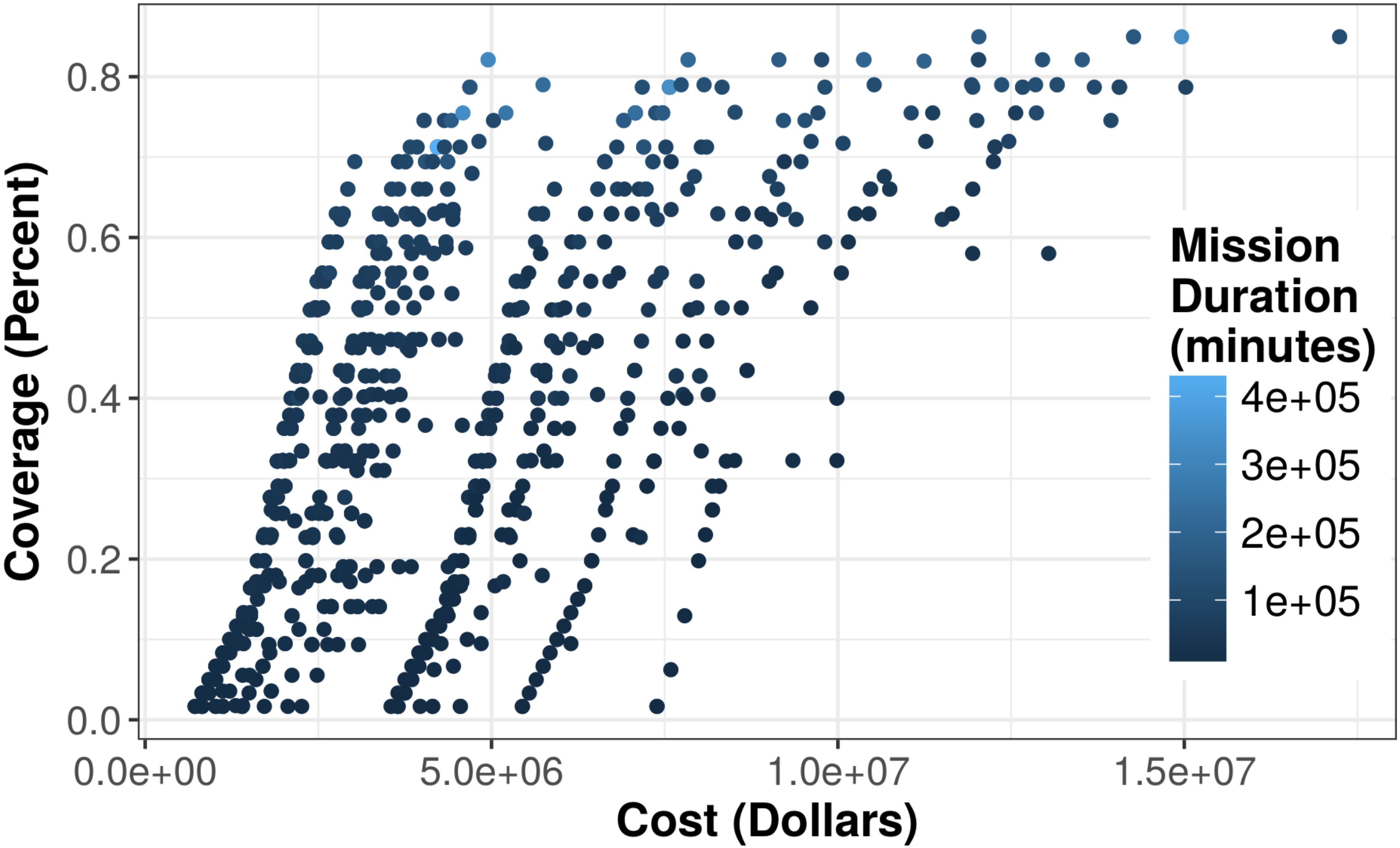
Domain model in Ecore + OCL (Excerpt)



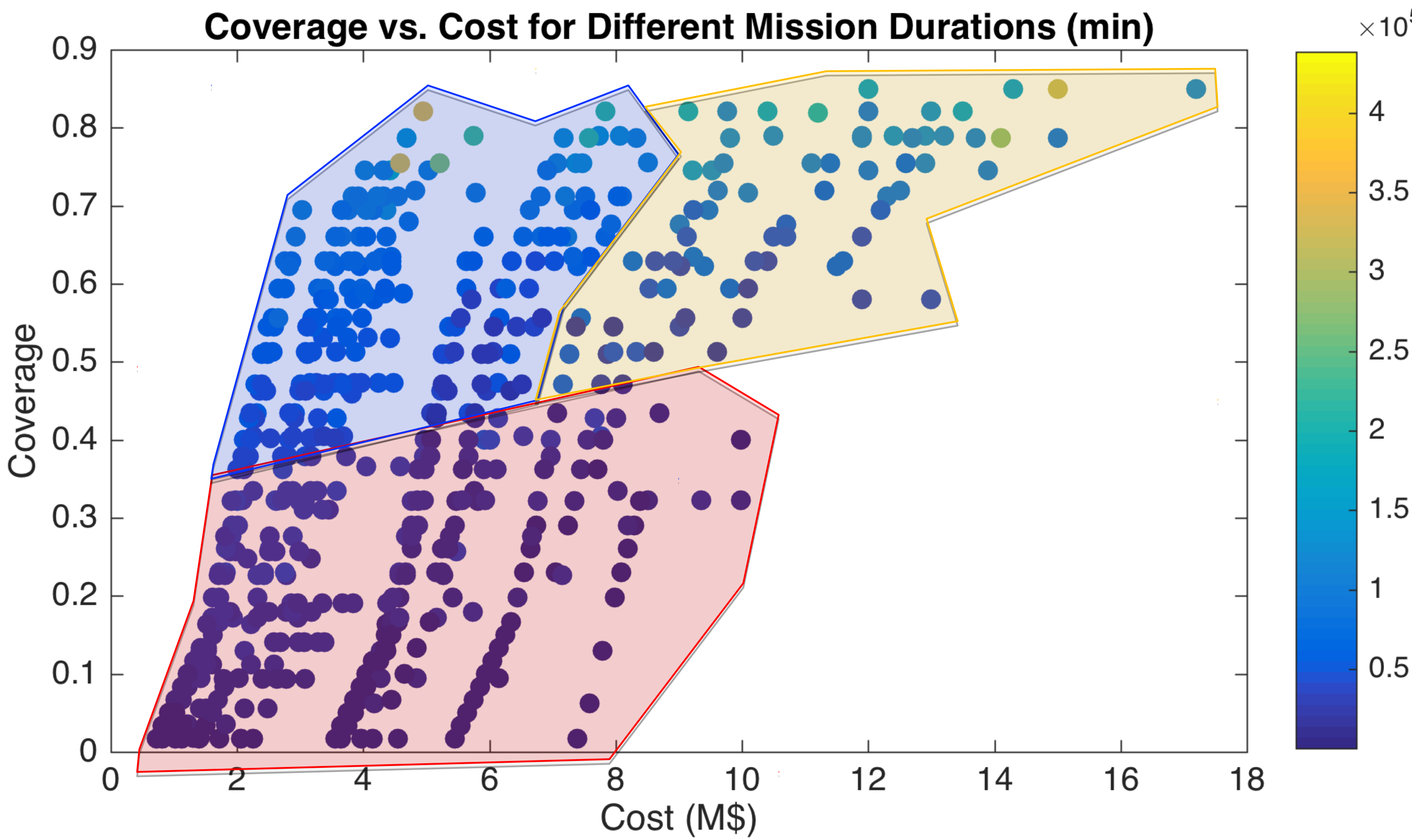
20 concepts, 9 associations, 15 attributes / parameters

> 48^{10} possible models

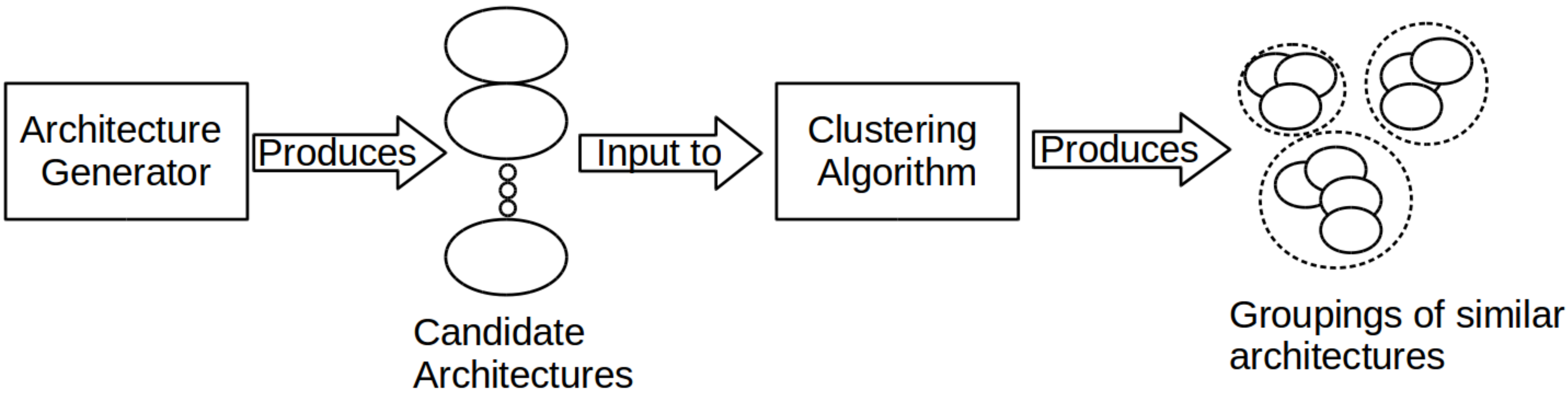
Problem: Too Many Architectures!



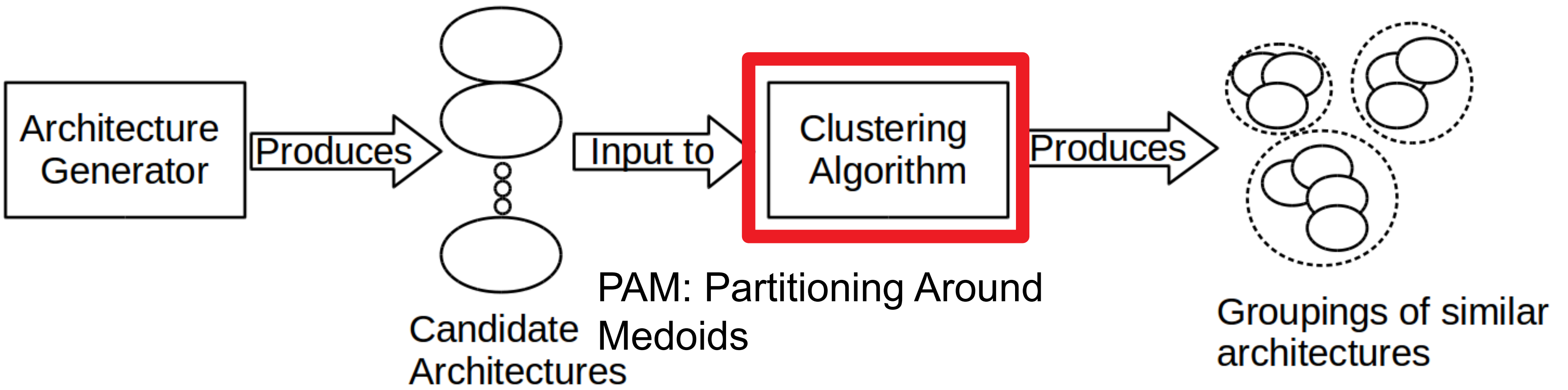
Idea: Clustering Similar Architectures



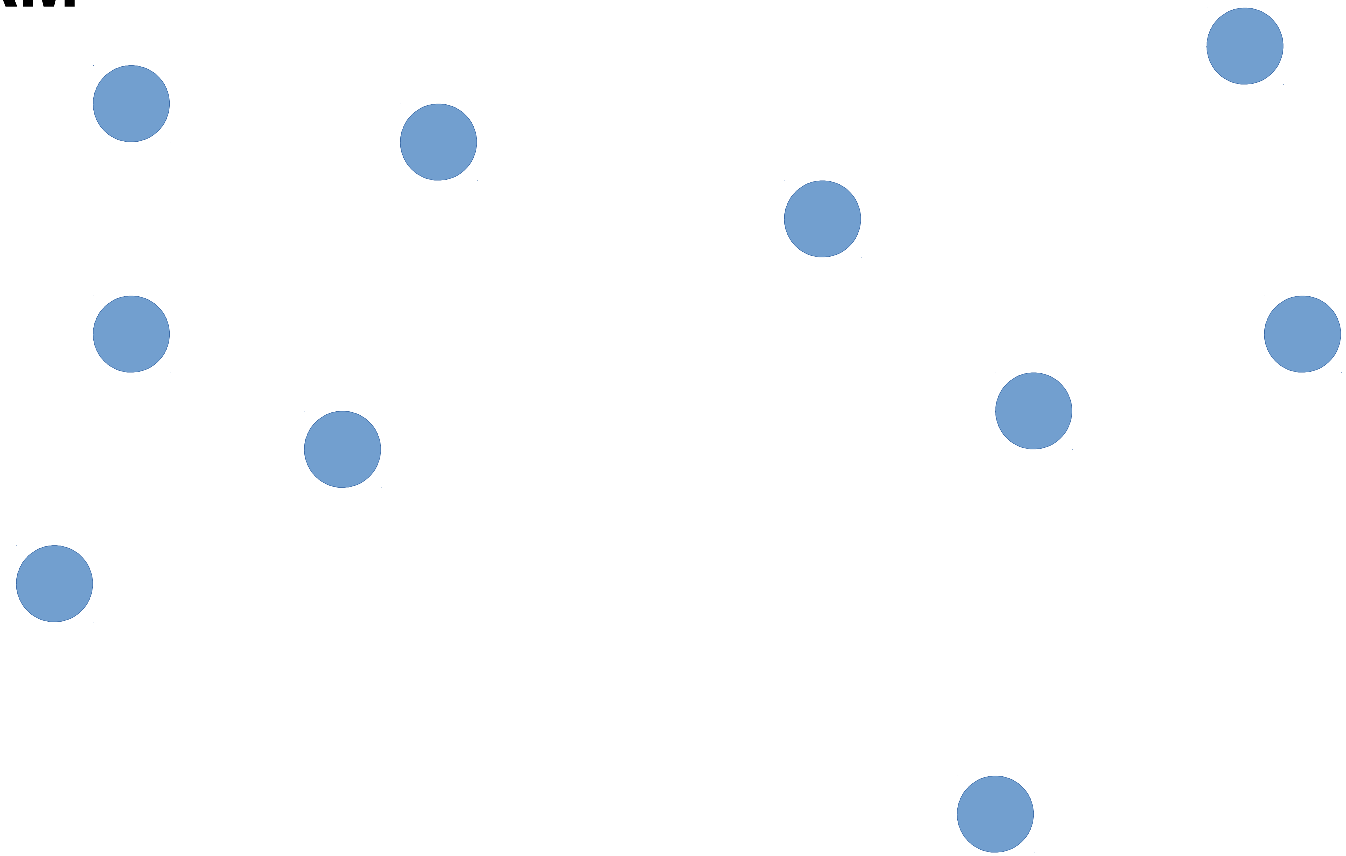
Overview of Approach



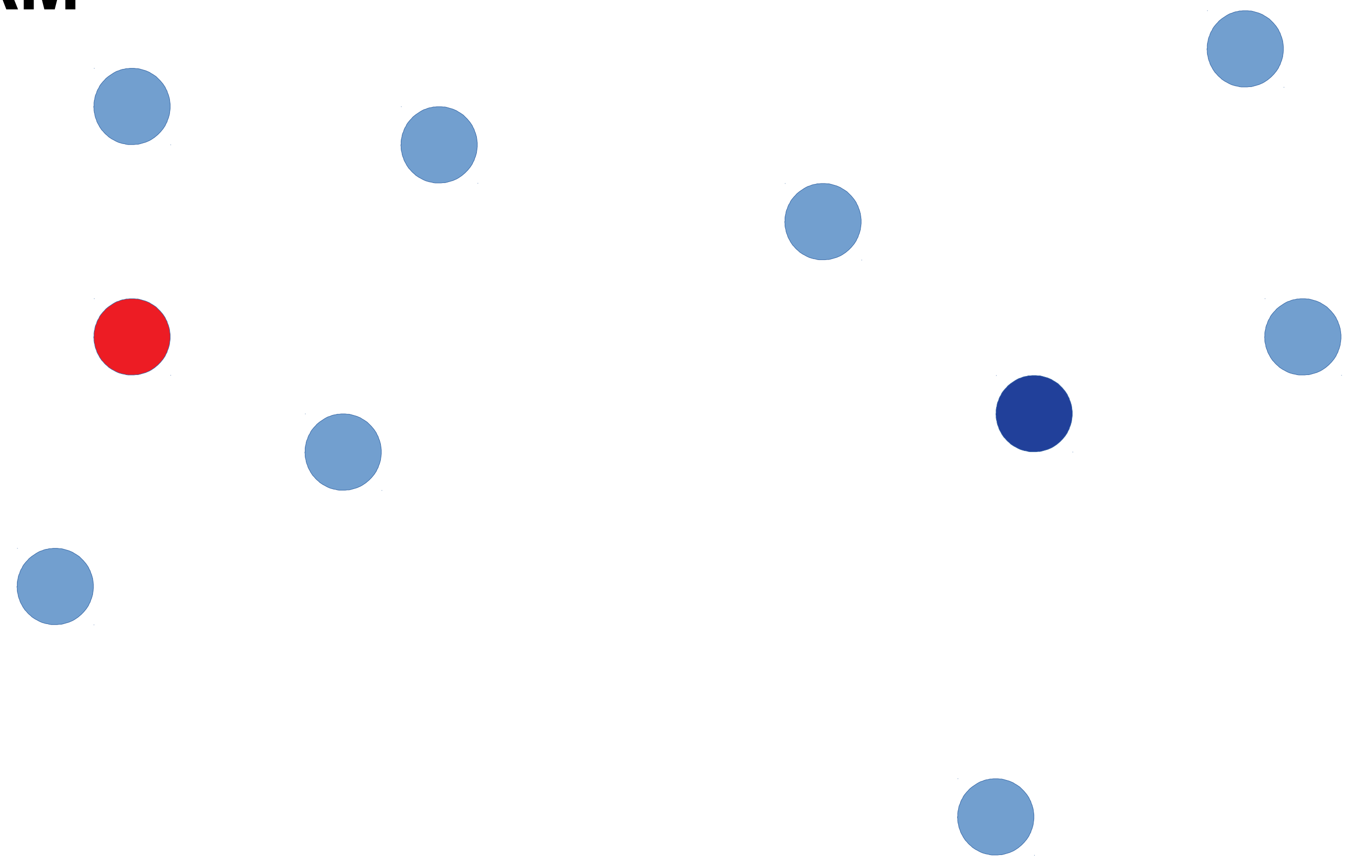
Overview of Approach



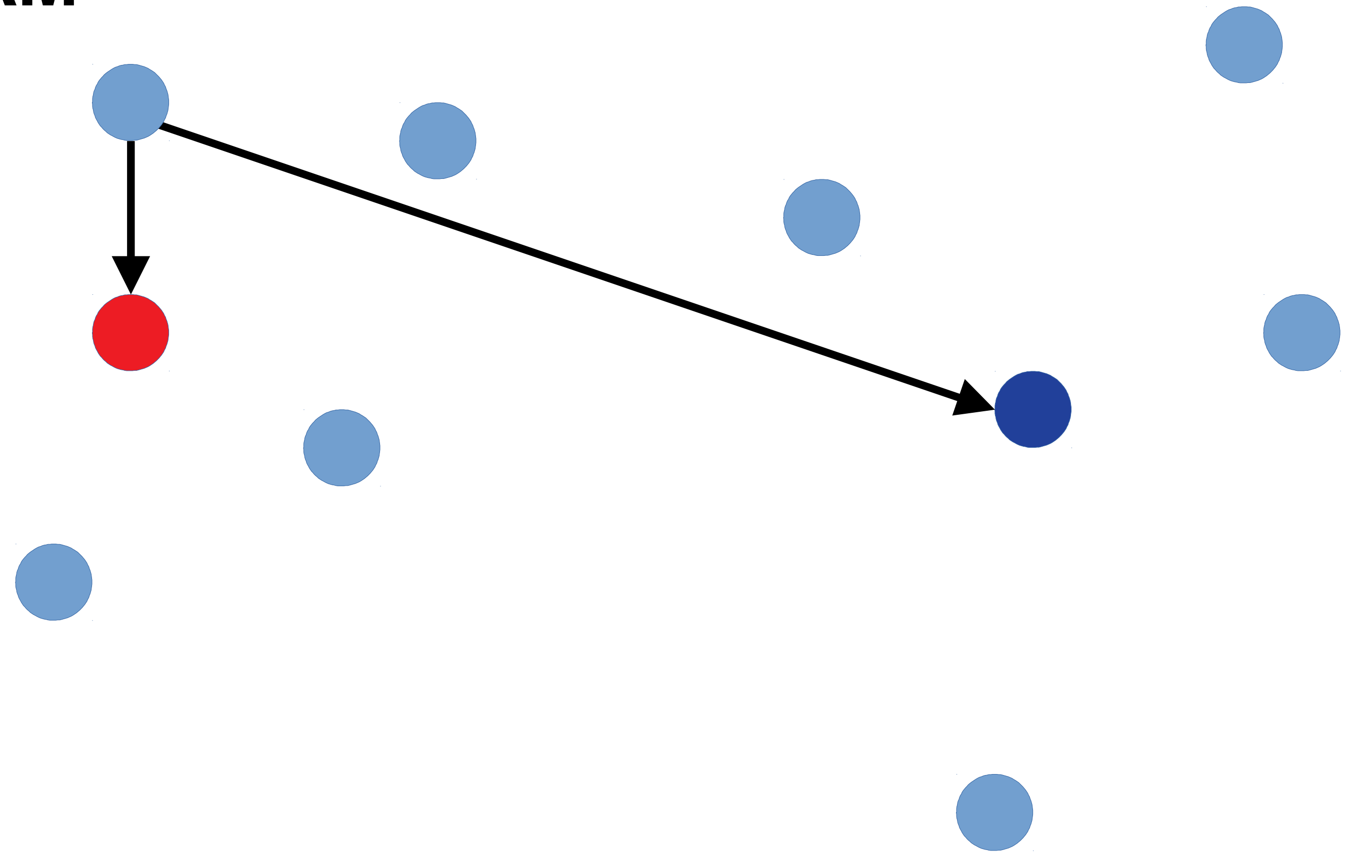
PAM



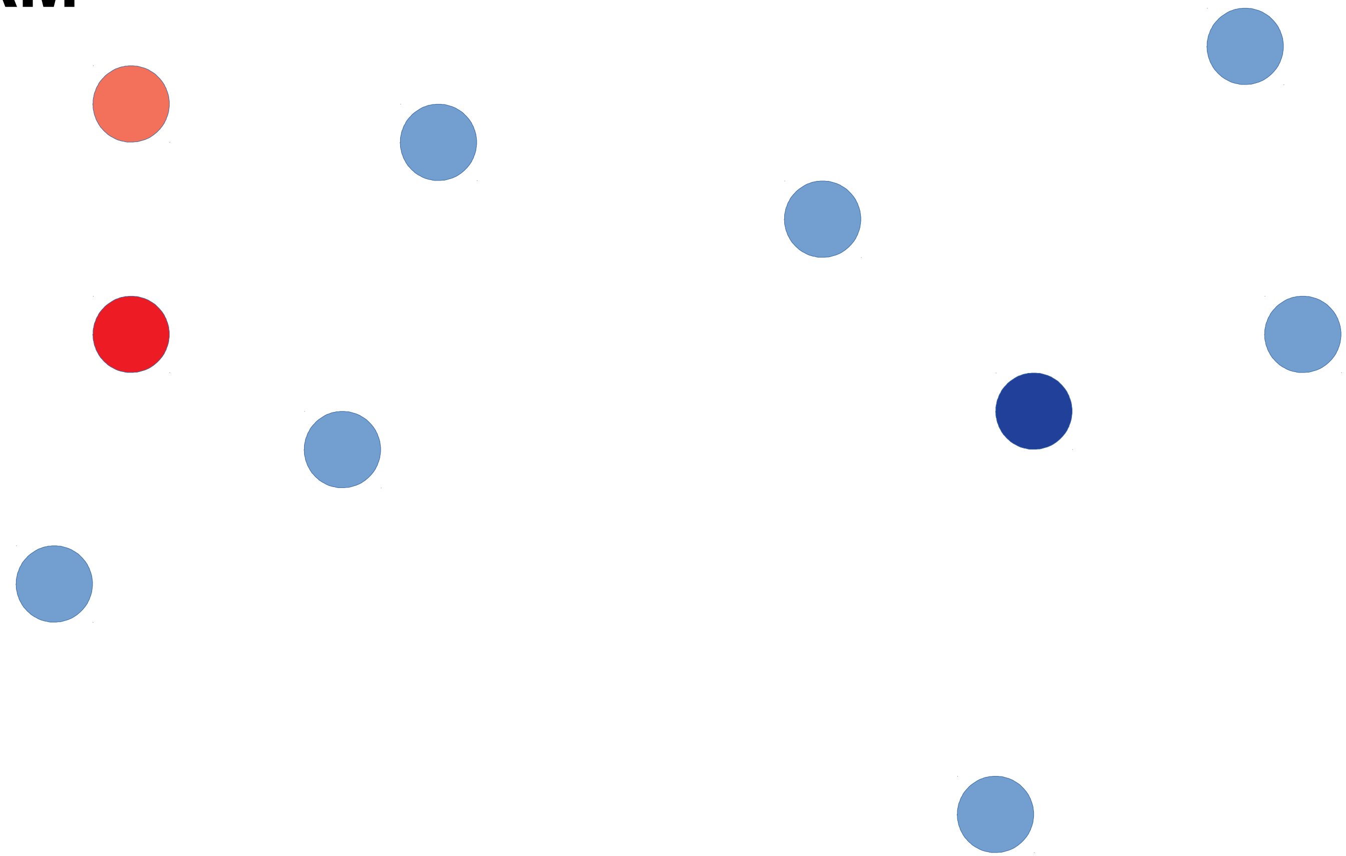
PAM



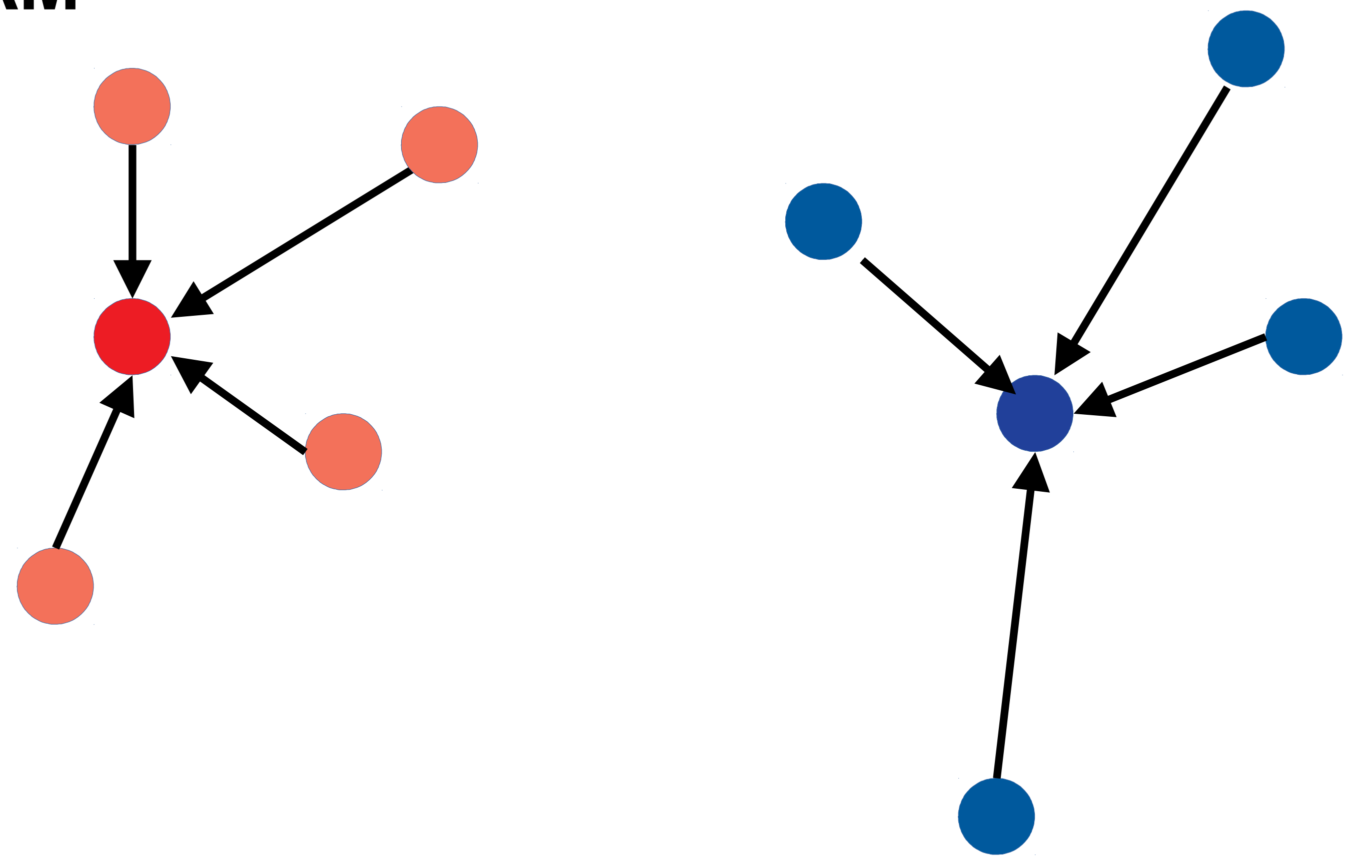
PAM



PAM

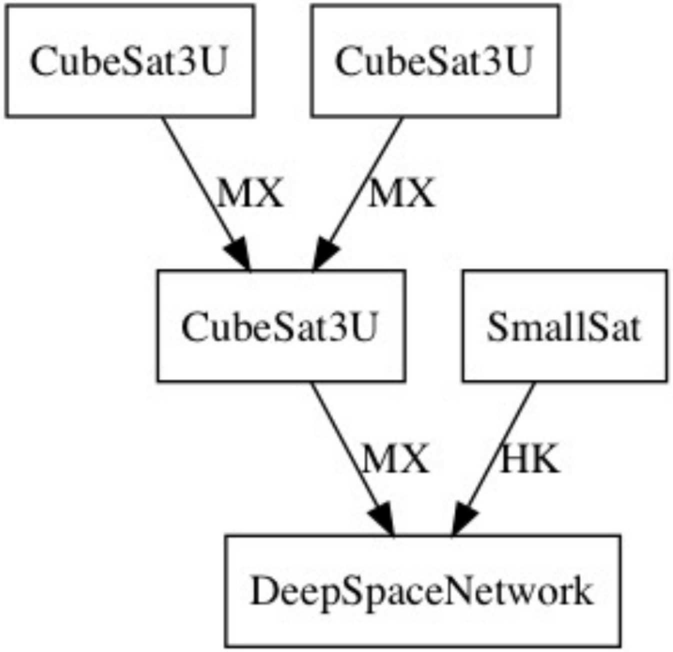


PAM

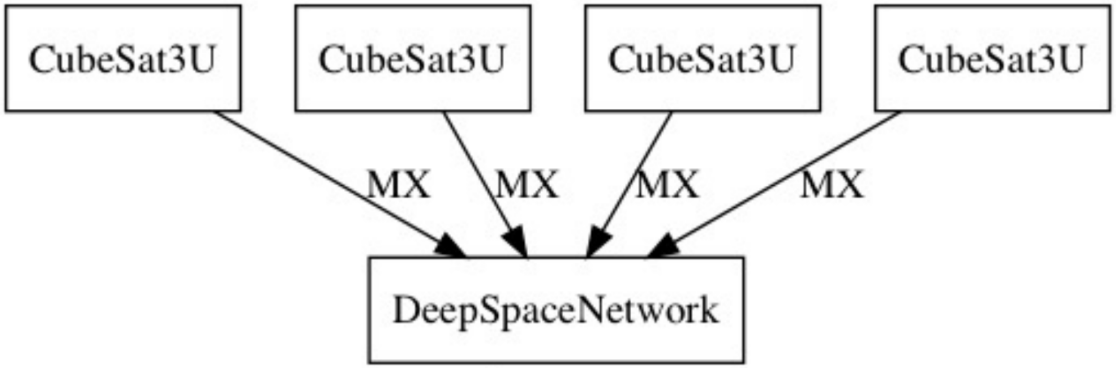


Distance Measure?

Distance Measure?

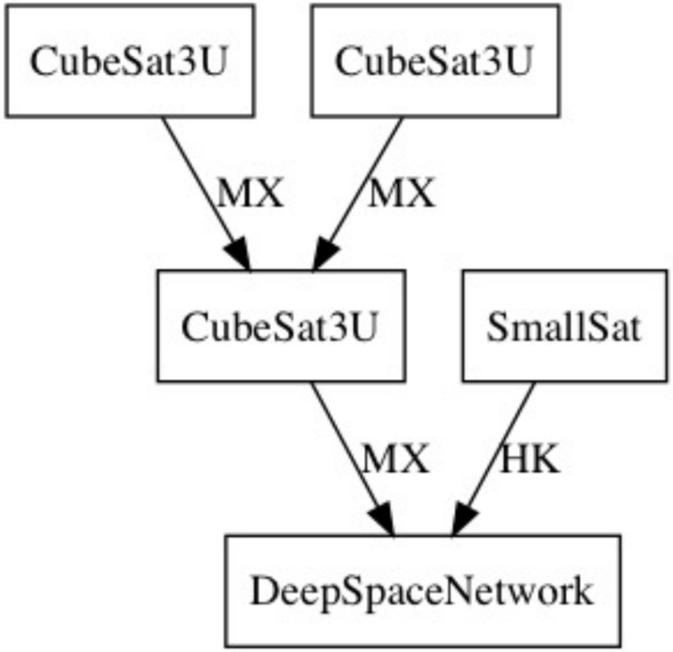


ID: 702 C:0.36 $4.97 MD: 22.97 OT: 8.0

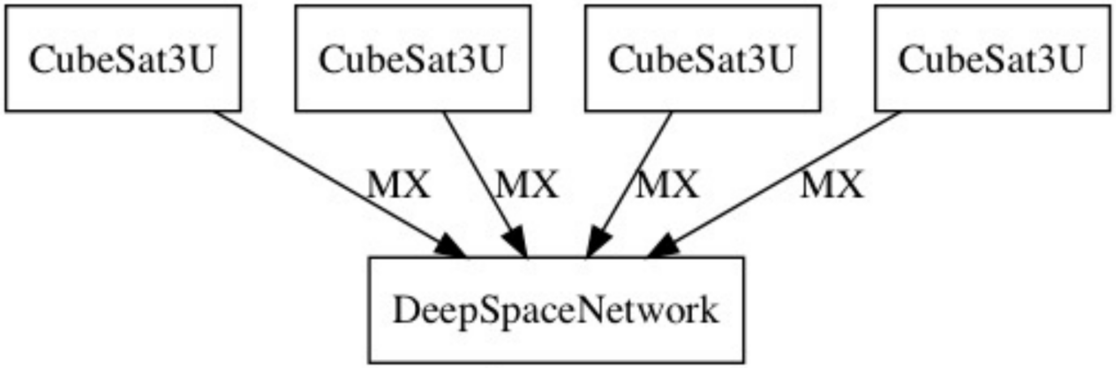


ID: 162 C:0.28 $1.91 MD: 17.82 OT: 6.0

Distance Measure?



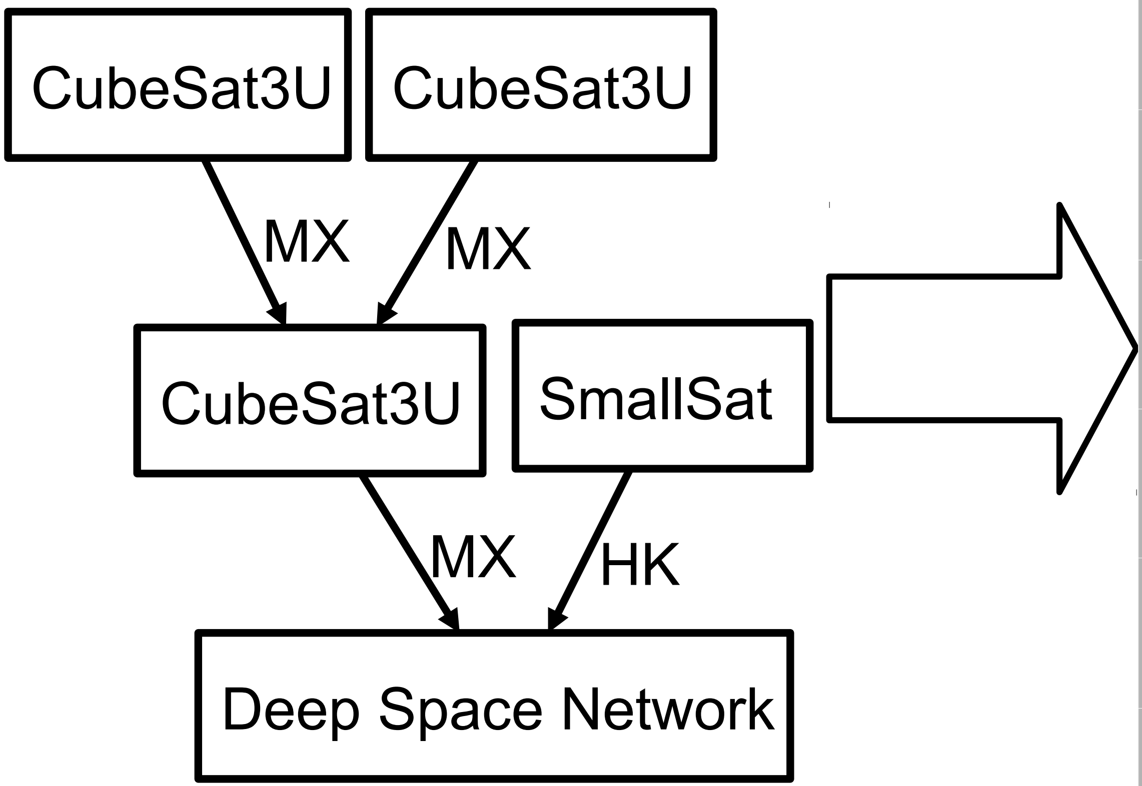
ID: 702 C:0.36 $4.97 MD: 22.97 OT: 8.0



ID: 162 C:0.28 $1.91 MD: 17.82 OT: 6.0

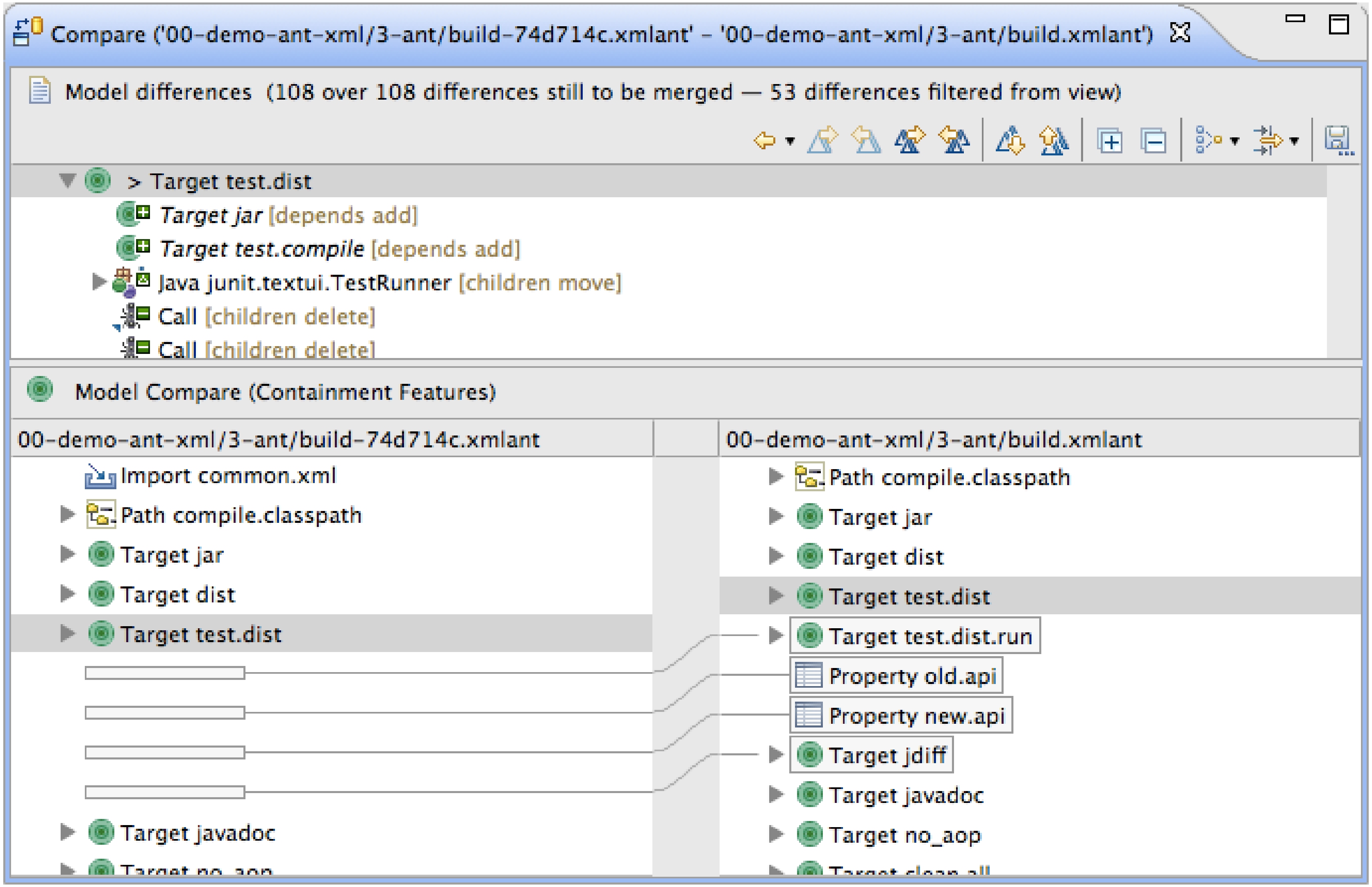
How to determine distance is non-trivial
We investigate three approaches

Feature Selection



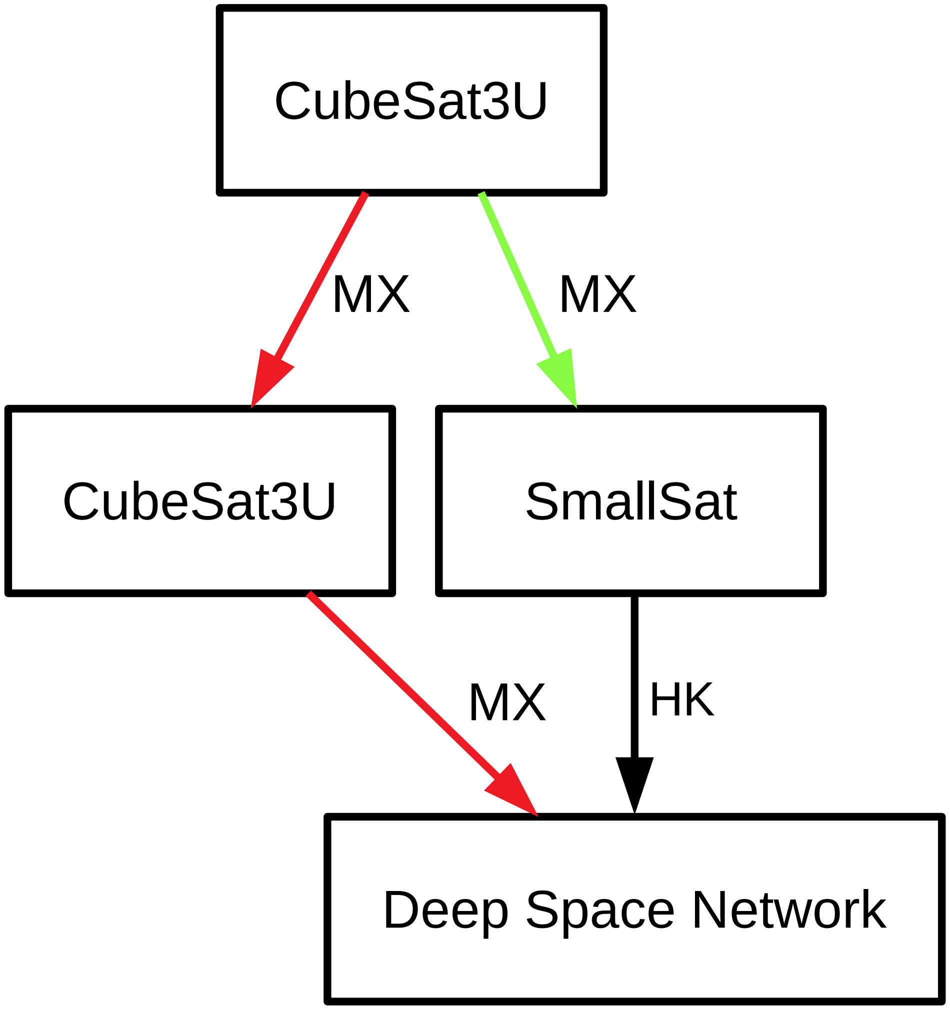
	Feature Vector
Number of Assets	4
Cost	4.97
Coverage	0.28
Mission Duration	22.97
...	...

EMF Compare

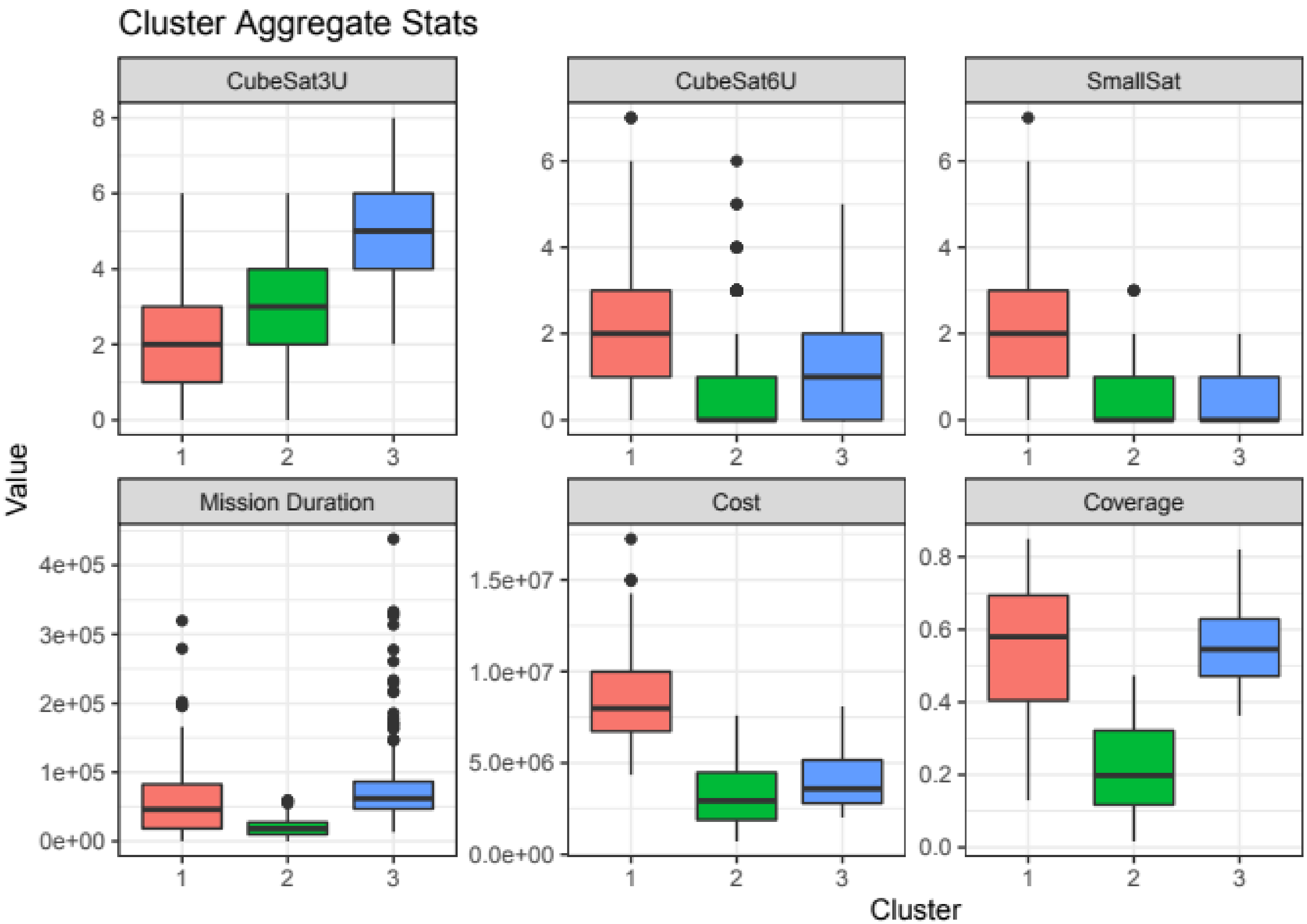


Compare ('00-demo-ant-xml/3-ant/build-74d714c.xmlant' - '00-demo-ant-xml/3-ant/build.xmlant')
Model differences (108 over 108 differences still to be merged — 53 differences filtered from view)
> Target test.dist
Target jar [depends add]
Target test.compile [depends add]
Java junit.textui.TestRunner [children move]
Call [children delete]
Call [children delete]
Model Compare (Containment Features)
00-demo-ant-xml/3-ant/build-74d714c.xmlant
Import common.xml
Path compile.classpath
Target jar
Target dist
Target test.dist
Target javadoc
00-demo-ant-xml/3-ant/build.xmlant
Path compile.classpath
Target jar
Target dist
Target test.dist
Target test.dist.run
Property old.api
Property new.api
Target jdiff
Target javadoc
Target no_aop

Graph-edit Distance

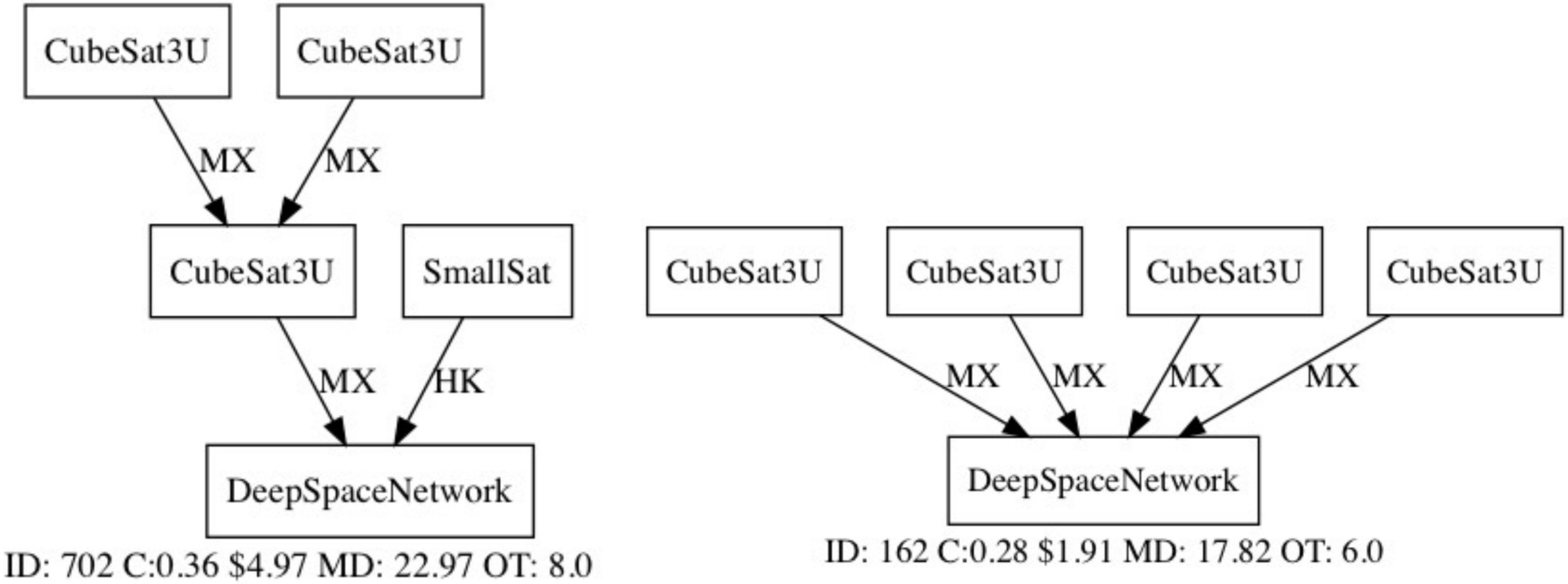


Feature Selection



Validation

- Manual clustering task
- Given pairs, assign a distance score
- Caveats
    - 31 pairs, two groups of 2-3



Results

	Group 1	Group 2	Features (All)	Features (Assets)	Features (Objectives)	Graph-edit Distance	EMF Compare
Group 1	1						
Group 2	**0.501**	1					
Features (All)	0.364	**0.386**	1				
Features (Assets)	0.263	**0.560**	**0.436**	1			
Features (Objectives)	0.304	0.223	**0.869**	0.341	1		
Graph-edit Distance	0.276	0.217	**0.464**	0.289	**0.429**	1	
EMF Compare	0.029	0.123	**0.536**	0.147	**0.424**	**0.789**	1

Insights from human designers

- Presence or absence of SmallSat
- Number of incoming / outgoing connections (relay)
- Number of bands of communication
- Difference influenced by:
  - Background
  - Goals

Keyword	Group 1	Group 2
relay	2	5
bands	2	3
layers / levels	2	6
SmallSats	2	2
threads	0	2

Conclusions

- Clustering has the potential to enable more through analysis of the architectural trade space
- Dissimilarity measures for space mission architectures are non-trivial, and have trade-offs in granularity, extensibility, and types of considered information
- Discussed insights from human clustering task, importance of a range of options

- Clustering is a promising approach for design space exploration

Cody Kinneer
ckinneer@cs.cmu.edu

Jet Propulsion Laboratory
California Institute of Technology

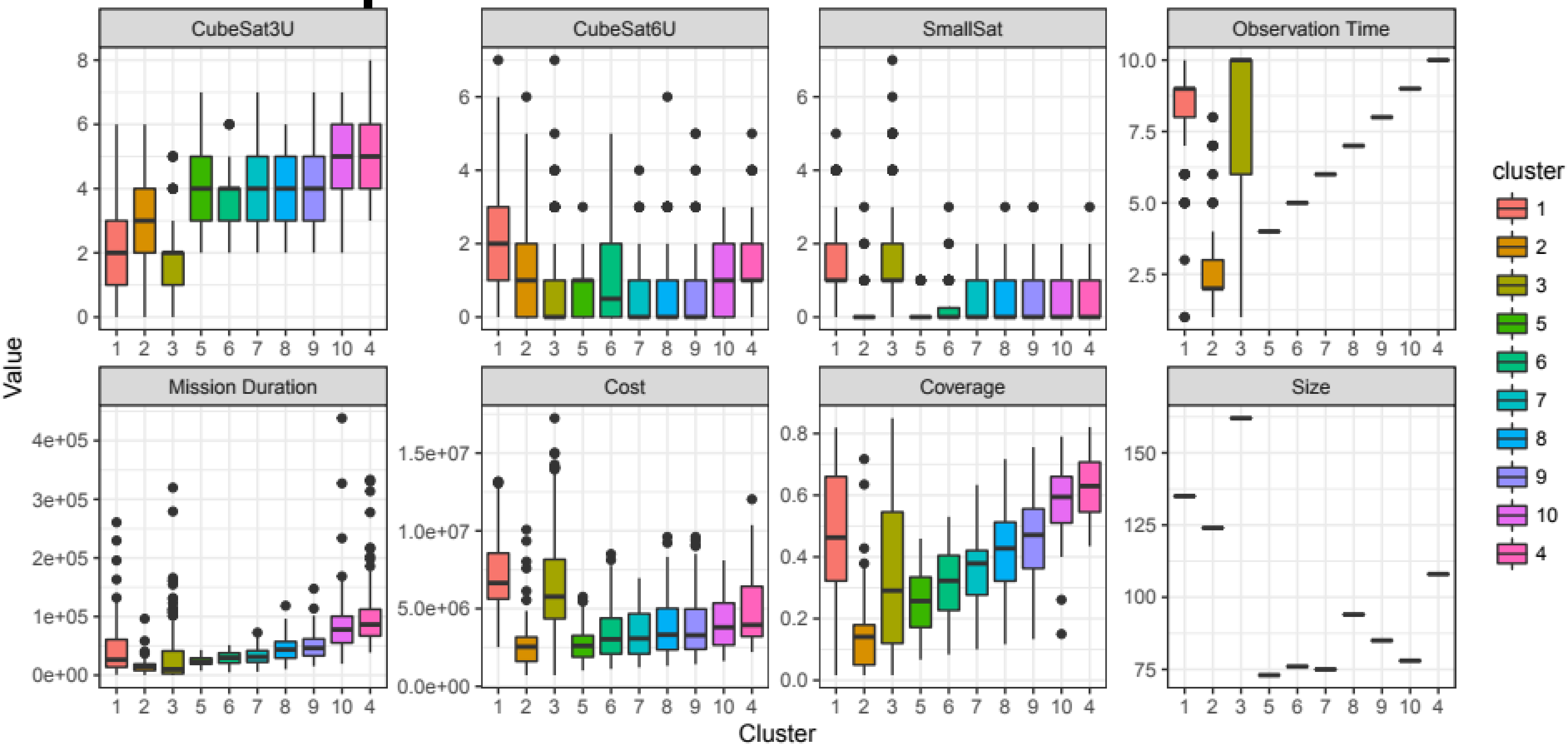
jpl.nasa.gov

Jet Propulsion Laboratory
California Institute of Technology

Backup Slides

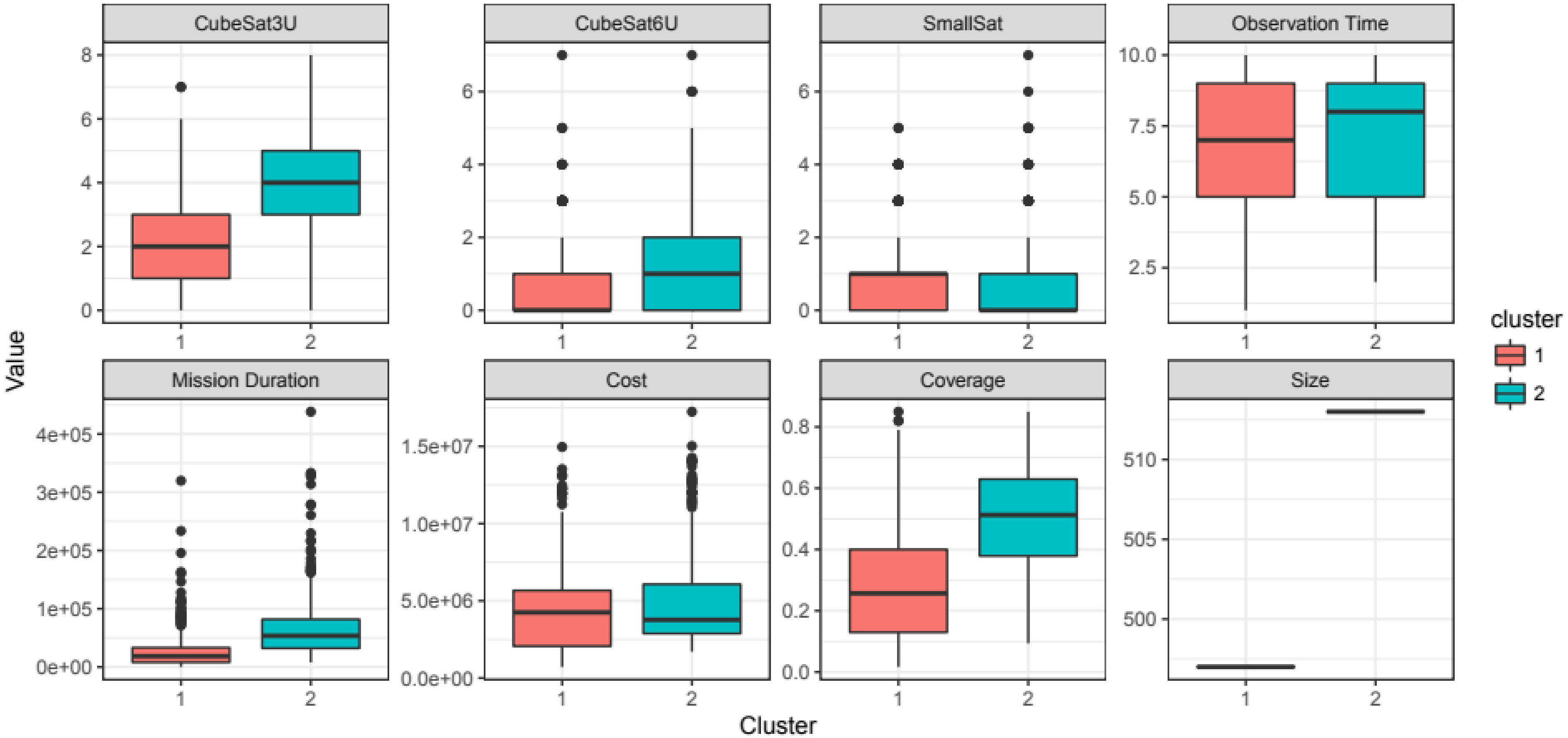
ACM/IEEE MODELS 2018 Presentation on *"Dissimilarity Measures for Clustering Space Mission Architectures"*

EMF Compare



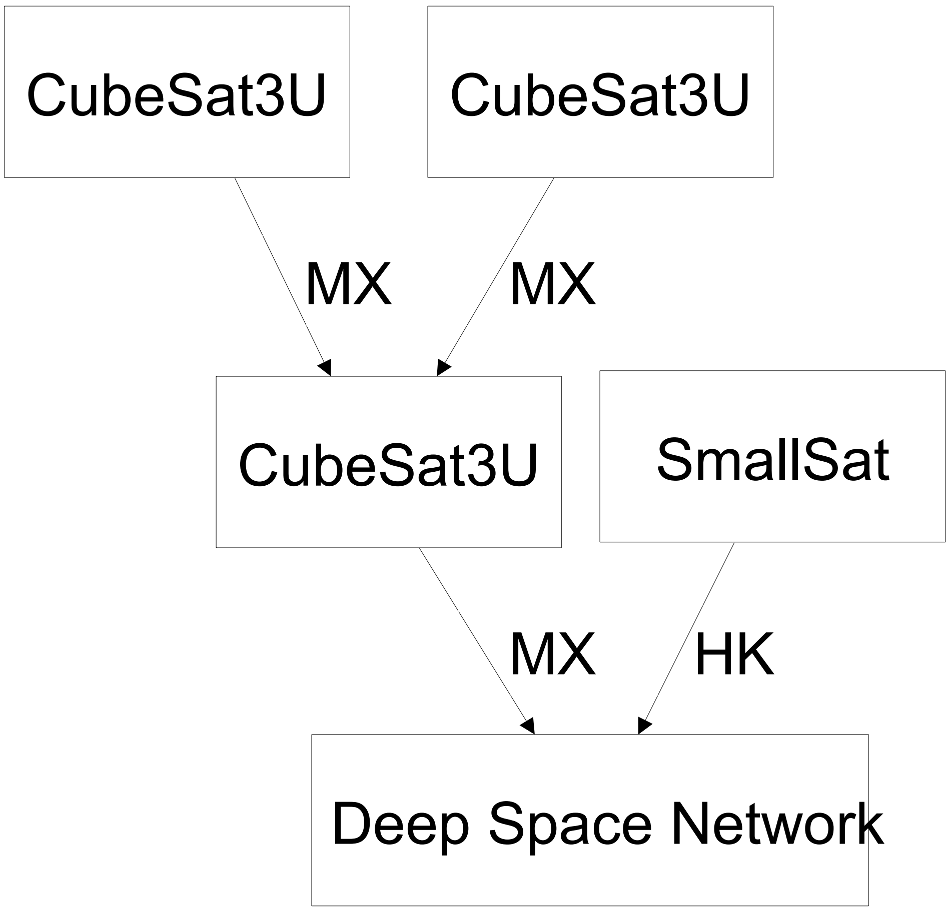
CubeSat3U
CubeSat6U
SmallSat
Observation Time
Mission Duration
Cost
Coverage
Size
Value
Cluster
cluster
1
2
3
5
6
7
8
9
10
4

Graph-edit Distance



Example Mission Architecture

- Number of spacecraft
- Type of spacecraft
- Directed communication links
- Communication equipment
    - Gain
    - Band
- Ground station
- Payload



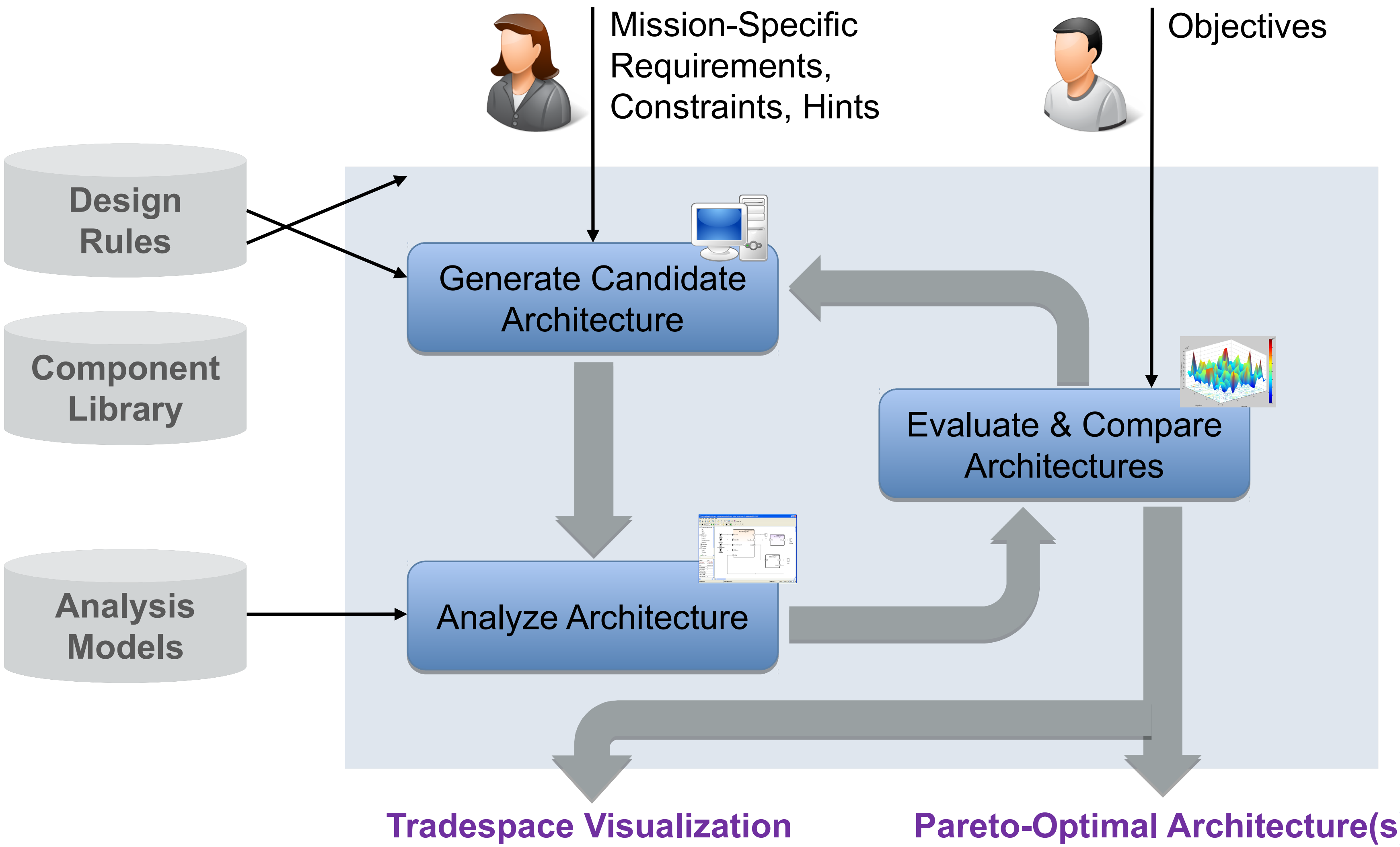
Implementation

Open Source Technologies Used in Implementation

- Representation of Domain
  - **Ecore / Eclipse EMF + OCL**
- Exploration Rules
  - **Henshin**
- Analyses / Fitness Functions
  - **Java**
- Optimization Using Genetic Algorithms
  - **MOMoT, MOEA**

Framework

CDS for Mission Architecture Design



Application to Case Study

Link Calculations

- Derived from standard link budget, assuming above average noise due to expected interference from Moon

Table 1. Computed communication rates. 385k km case assumes 72 dBi receive antenna gain for X-band, and 85 dBi for Ka-band (similar to DSN).

Transmitter Configuration	200 km	385k km
UHF, 3 W, 1 dBi	5 Mbps	-
X-Band, 5 W, 10 dBi	1.6 Mbps	0.7 Mbps
Ka-Band, 15 W, 25 dBi	220 Mbps	80 Mbps

Application to Case Study

Cost Calculations

- Cost per spacecraft calculation incorporates a learning curve
- Assuming $ 100,000 per hour of observation to estimate observation and data processing cost

$$c_i = c_{base,type(i)} \cdot n_{type(i)}^{-0.25} + c_{conf,i} \qquad (5)$$

$$c_{total} = \sum_{i=1}^{n_{sc}} c_i + 100,000 t_{obs} \qquad (6)$$

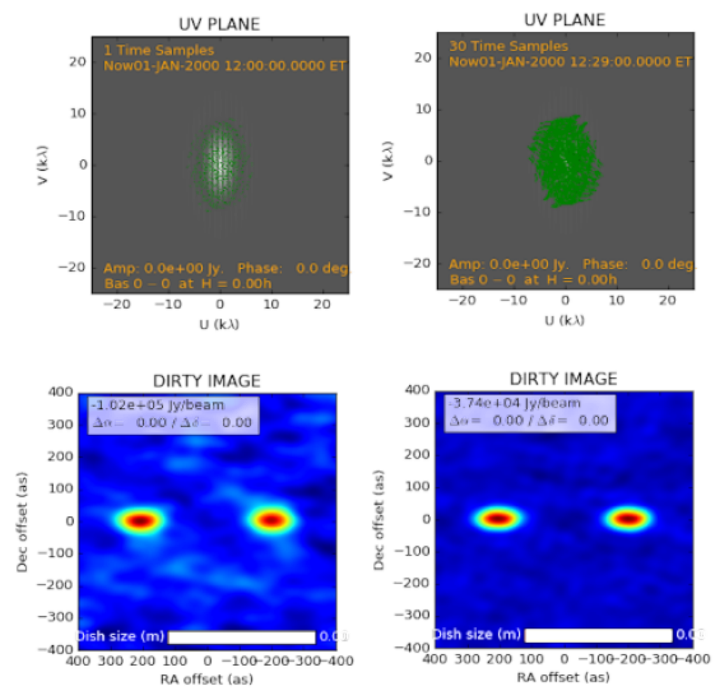
Application to Case Study

Coverage

- Simple coverage calculation

$$cov = (1 - \frac{2}{n_{obs}})^{1+9(1/t_{obs})} + 0.05\frac{t_{obs}}{3} \qquad (1)$$

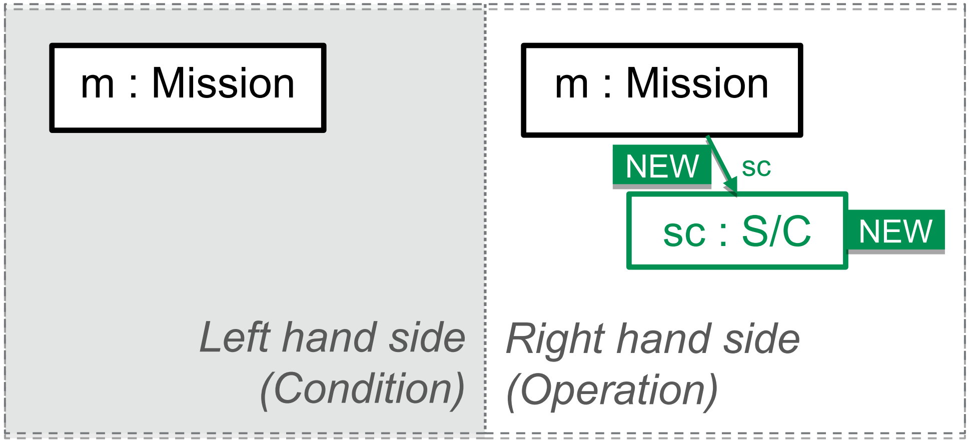
- Surrogate model that reflects trends observed from more sophisticated telescope array simulation performed by Alexander Hegedus (https://github.com/alexhege/Orbital-APSYNSIM/)



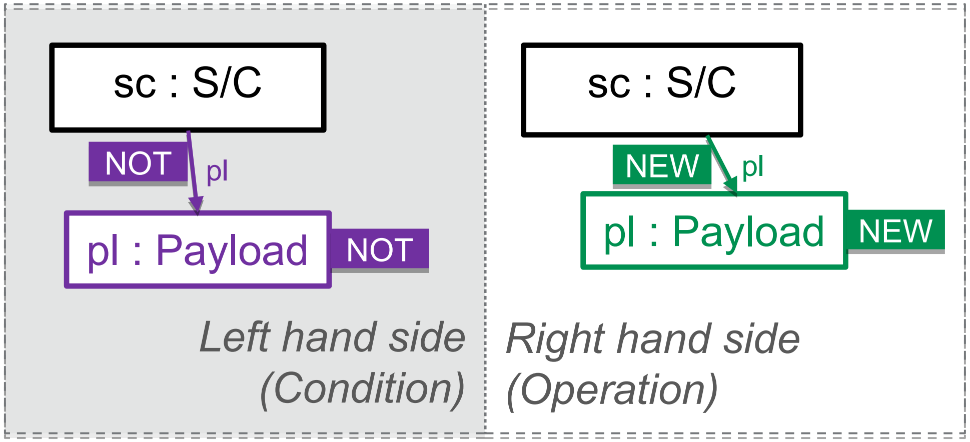
Model-Transformation-Based Exploration

Model Transformation Rules as Enablers for Evolving Solutions

- Transformation Rules
  - **LHS**: Condition for match in input model (e.g., *“find an element of type Mission”*)
  - **RHS**: Operation to be performed (e.g., *“create a new element of type S/C (Spacecraft) and attach it to the matched mission”*)
- Here: *endogenous* transformations
  - Source and target meta-models are the same
- Used for generating models in domain (~design rules)



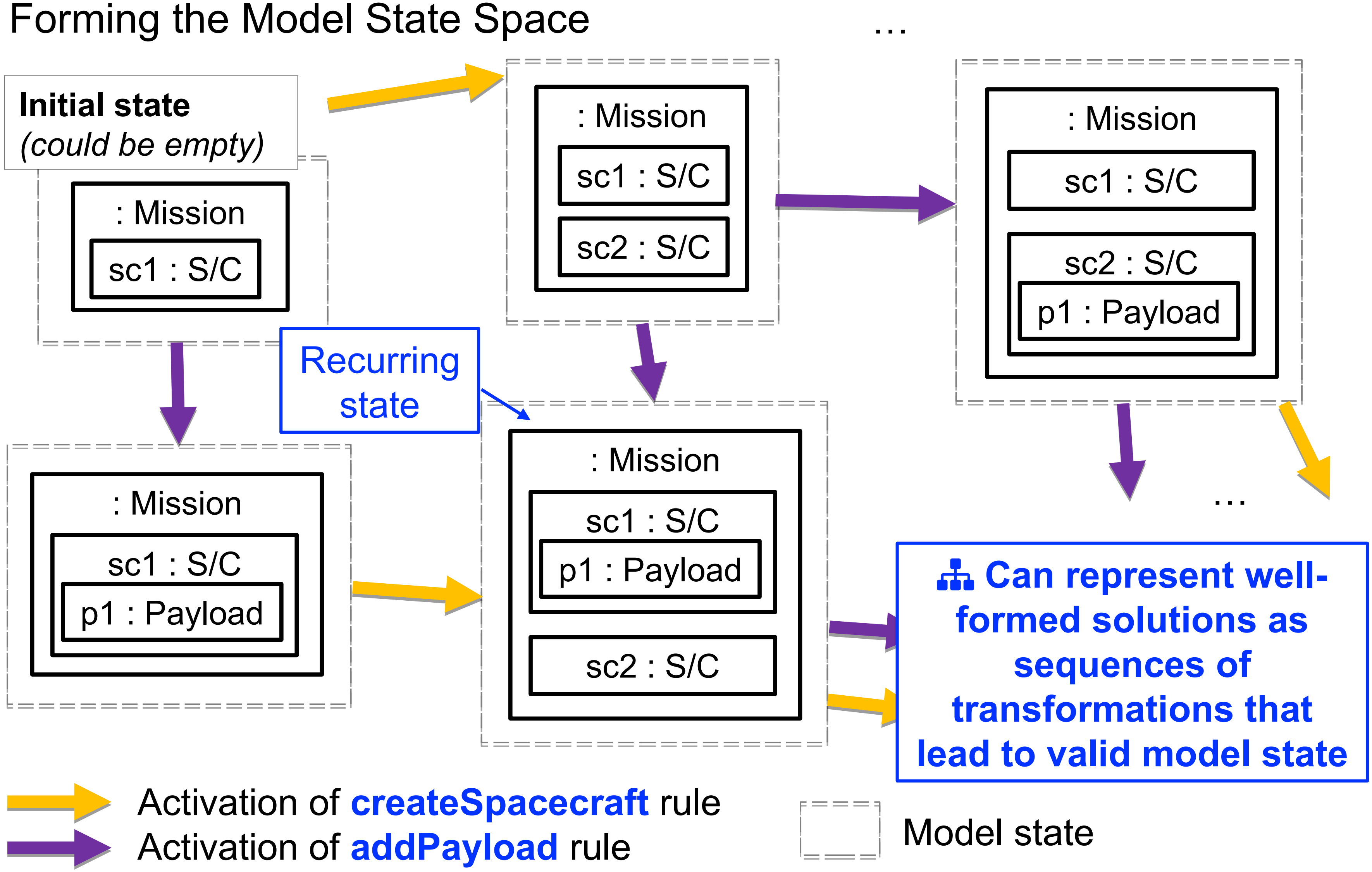
Rule “createSpacecraft”



Rule “addPayload”

Model-Transformation-Based Exploration

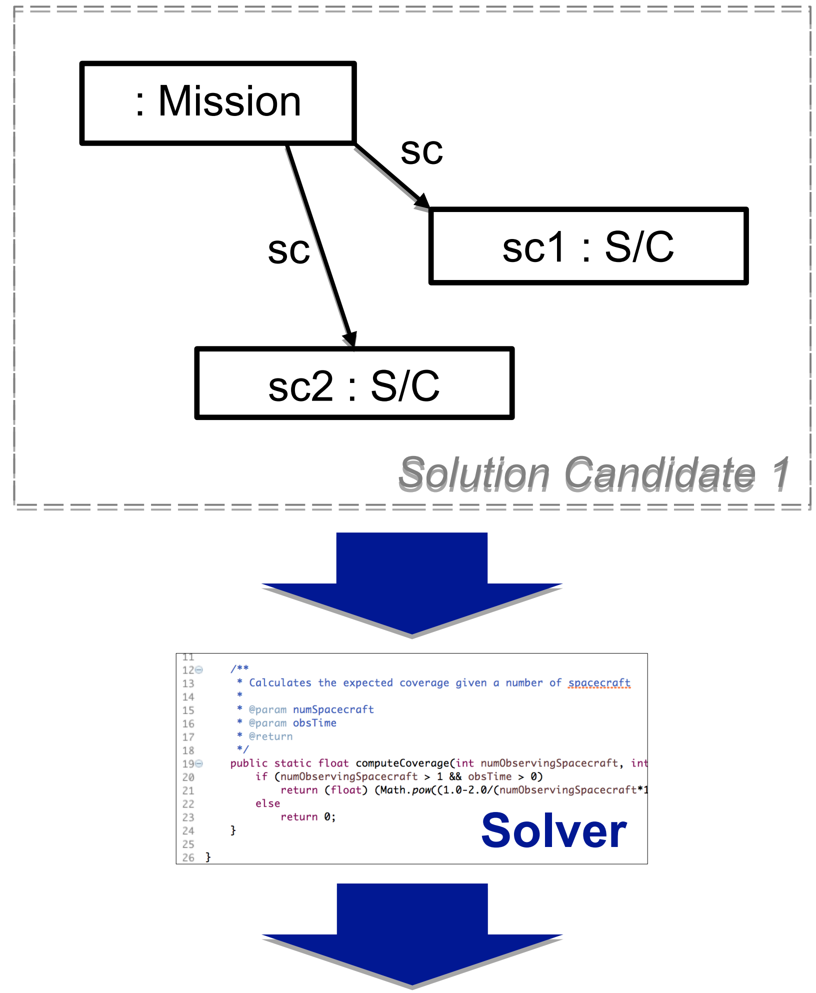
Forming the Model State Space



Evaluating the Objectives

- Evaluating objectives requires analysis of the candidate solution *(interpretation by a solver)*
  - Determine performance and determine values for measures of effectiveness
  - Determine objective function values

- Analyses defined at level of domain: part of formal interpretation of models within domain



"Scientific value of candidate 1 is 0.34"

Driving Exploration Towards Optima

Using Evolutionary Algorithms to find Pareto-Optimal Solutions

Crossover

Individual x:	Add Spacecraft	Add X-Band Comm	Add Spacecraft	Add Comm Link	*fitness=0.6*
(Selection from population)					*(Obj. Fct. Values)*
Individual y:	Add Spacecraft	Add Ka-Band Comm	Add Payload	Add Spacecraft	*fitness=0.5*

Here, individuals are **sequences of transformation rule activations**

→ Each genome in population is a variable with set of trafo rules as range

New: Add Ka-Band Comm *fitness=0.8*

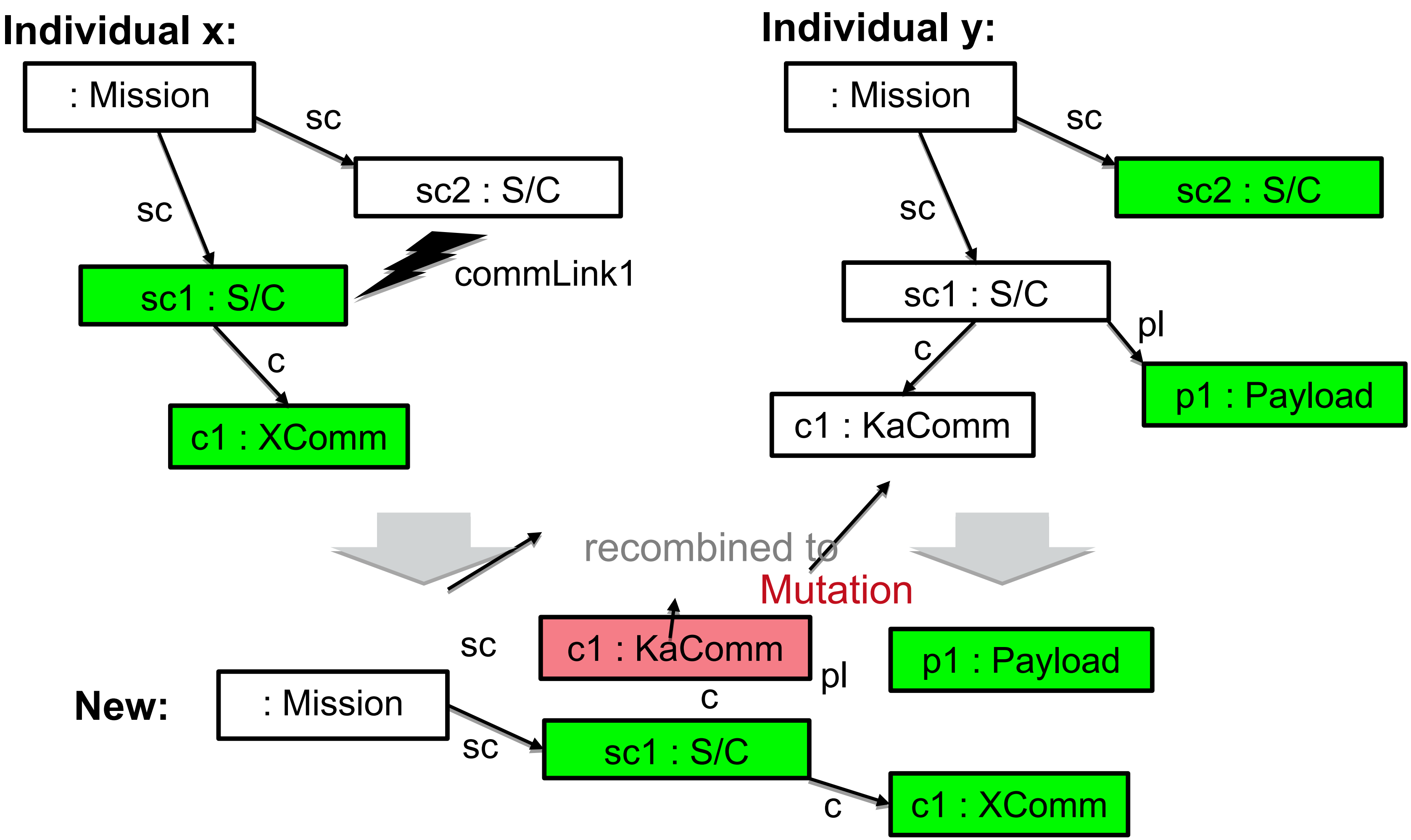
(Recombined individual in next generation)

Mutation

Could also be a "placeholder" transformation (= rule "do nothing")

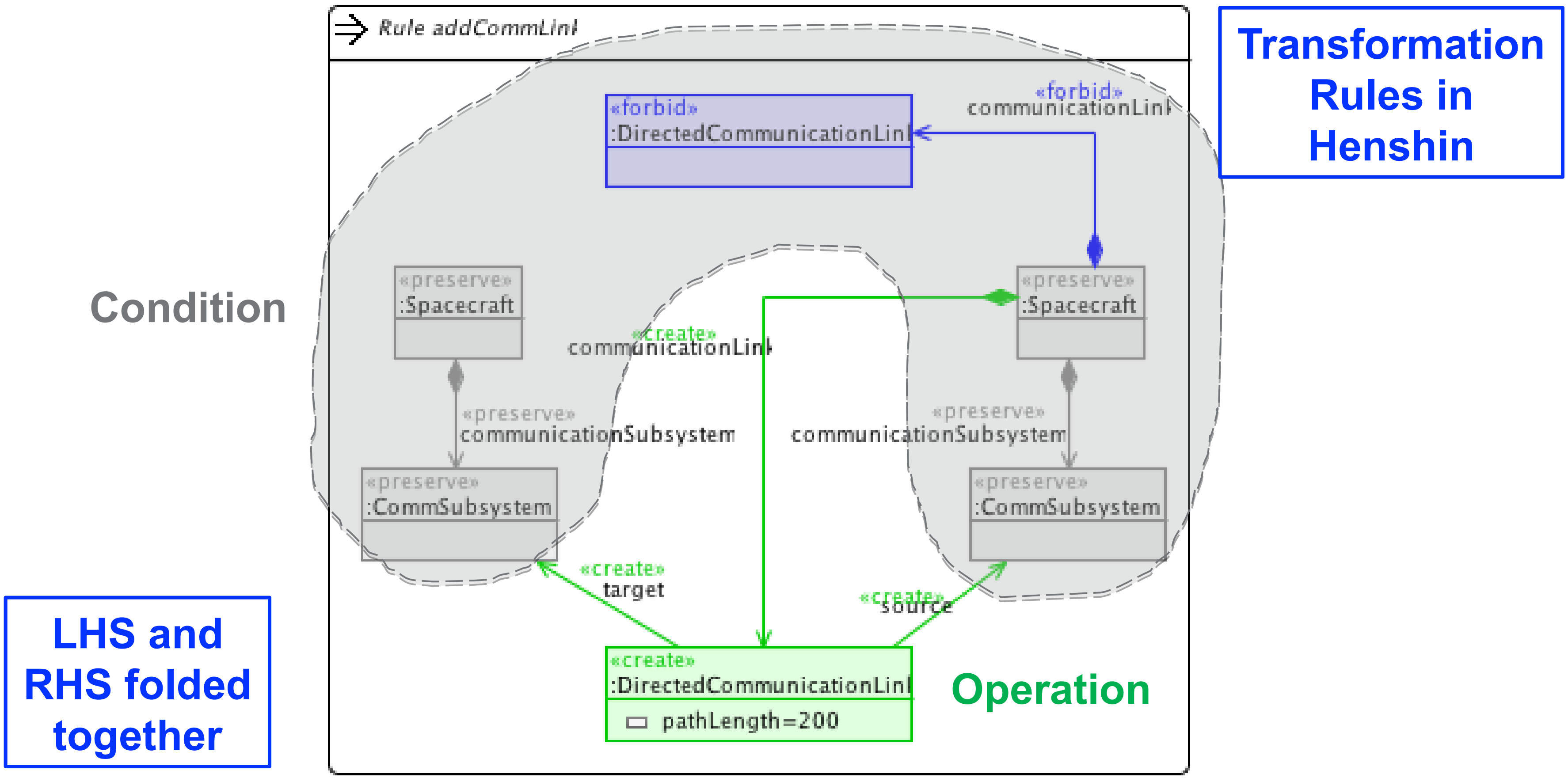
Driving Exploration Towards Optima

Models Resulting from Executing Transformations



Application to Case Study

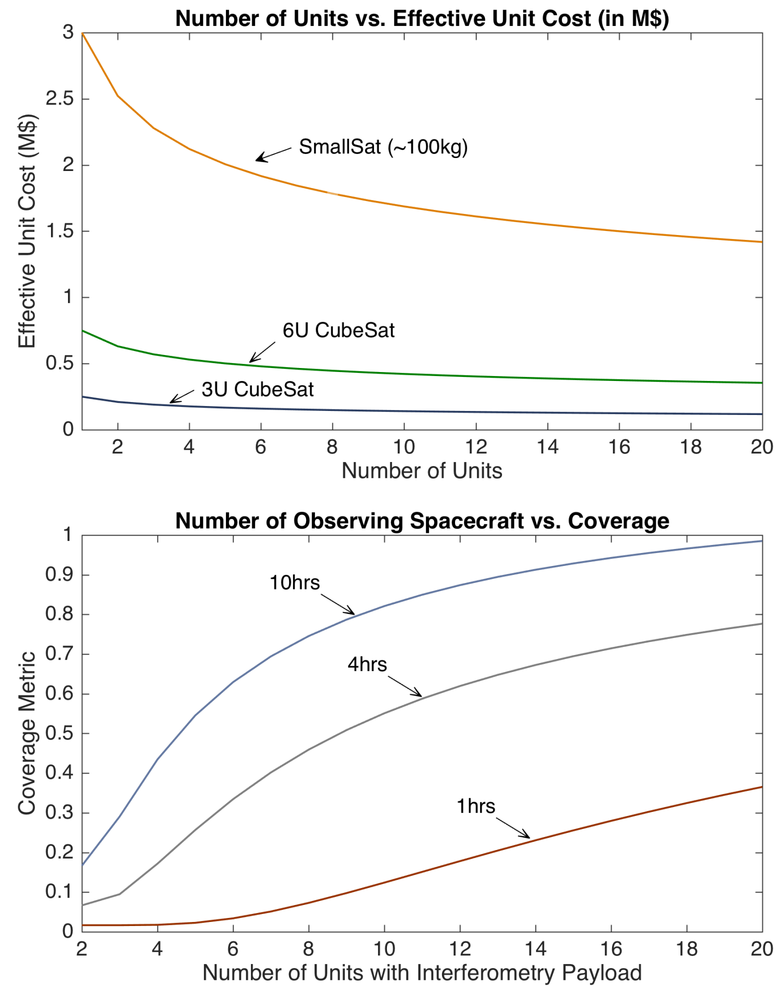
Transformation Rule Example (Henshin Syntax): Add Comm. Link



In Prose: *"Find 2 distinct spacecraft instances, and add a communication link between them"*

Application to Case Study

- Three objectives:
  - Minimize cost
  - Maximize coverage (measure of scientific benefit)
  - Minimize mission time
- Typical link budget for data rates
- Data collection & transfer model
- Abstracted away orbit design through coverage model
- Experiment setup:
  - 16 transformation rules
  - 180 variables per individual
  - NSGA-II with population size 1000, and 1000 generations
  - 30 runs, 7 minutes each*

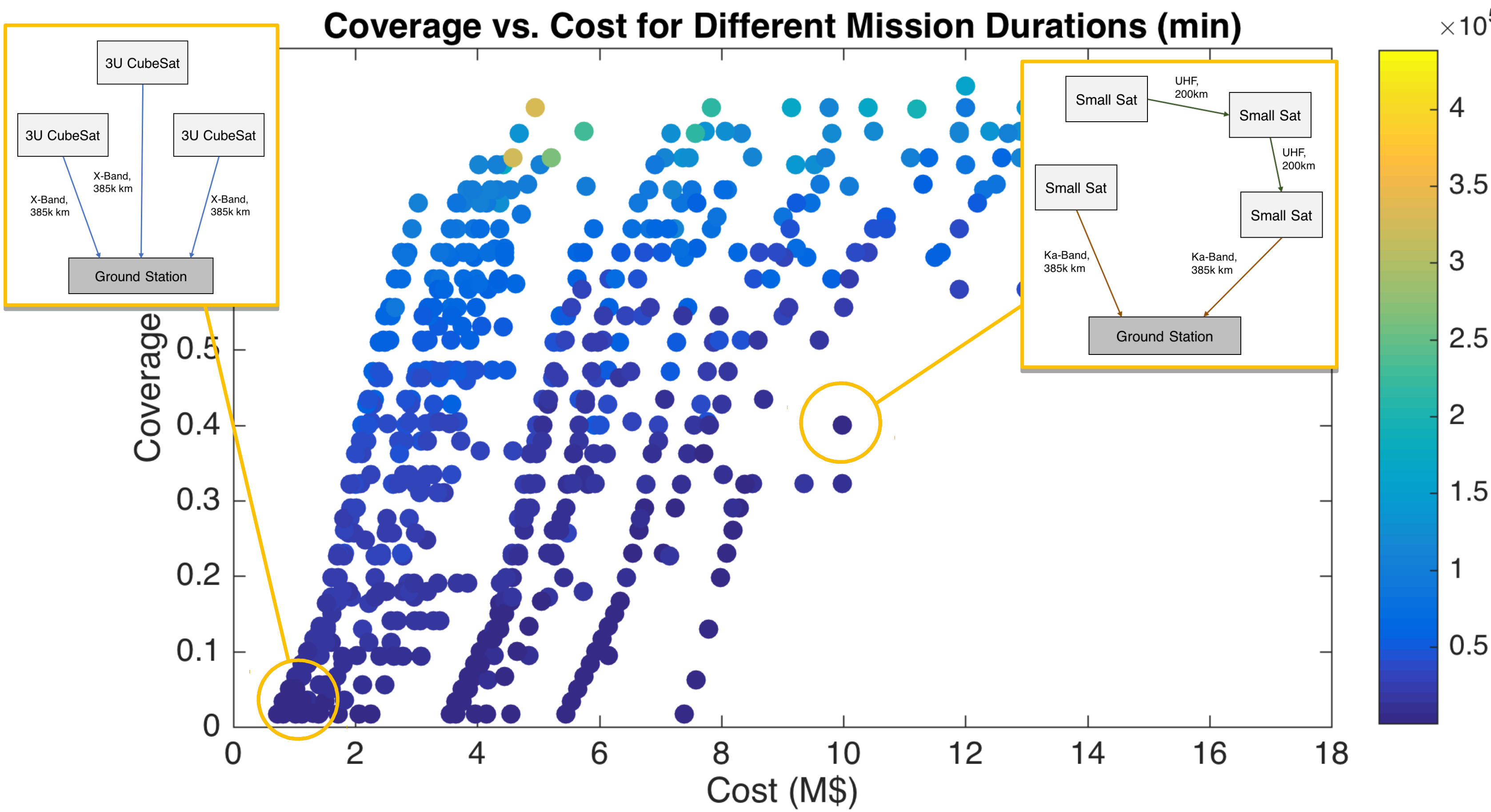


Fictitious cost model (top) and coverage model (bottom)

* 8 core Intel i7 @ 2.7Ghz, 16GB DDR3 RAM

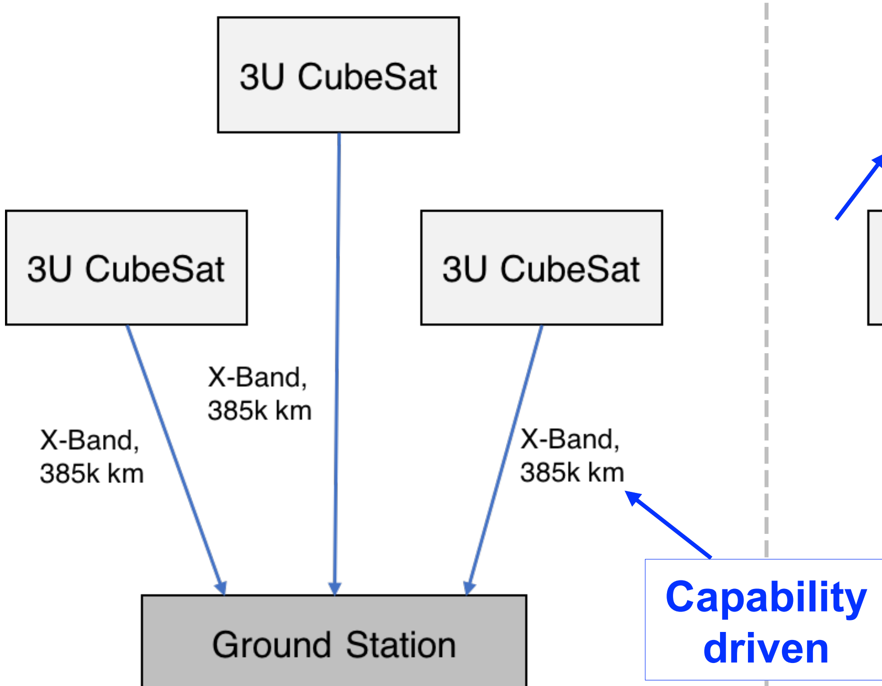
Results from Application to Case Study

Visualization of Trade Space

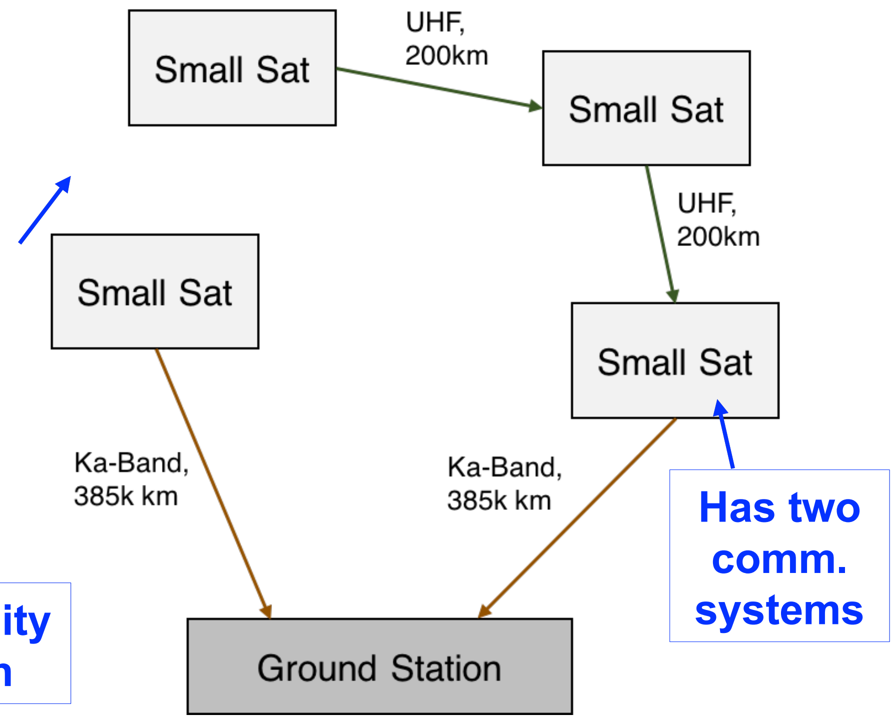


Results from Application to Case Study

Examples of Pareto-Optimal (Nondominated) Solutions



Candidate Solution #1
$1M, ~0.02 coverage

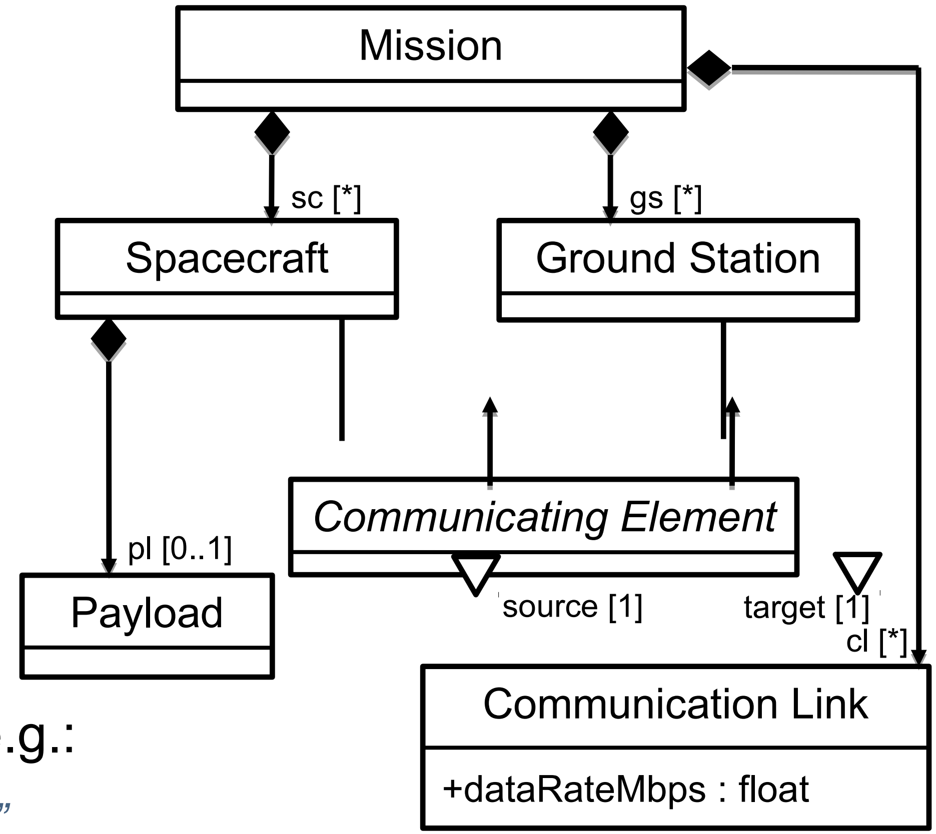


Candidate Solution #2
$10M, ~0.4 coverage

Similar mission duration, but #1 has much longer downlink time

Domain Model & Well-Formedness Constraints

- Domain model (meta-model)
  - Concepts
  - Associations / relations
  - Attributes
  - Describes a **universe of discourse**: many models in domain
  - Describes structural part of the problem



- Typically annotated with addl. well-formedness constraints, e.g.:

  "No communication loops may exist"

  "All spacecraft must (transitively) be connected to at least one ground station through a communication link"

Any model in the domain is a (structurally) valid solution