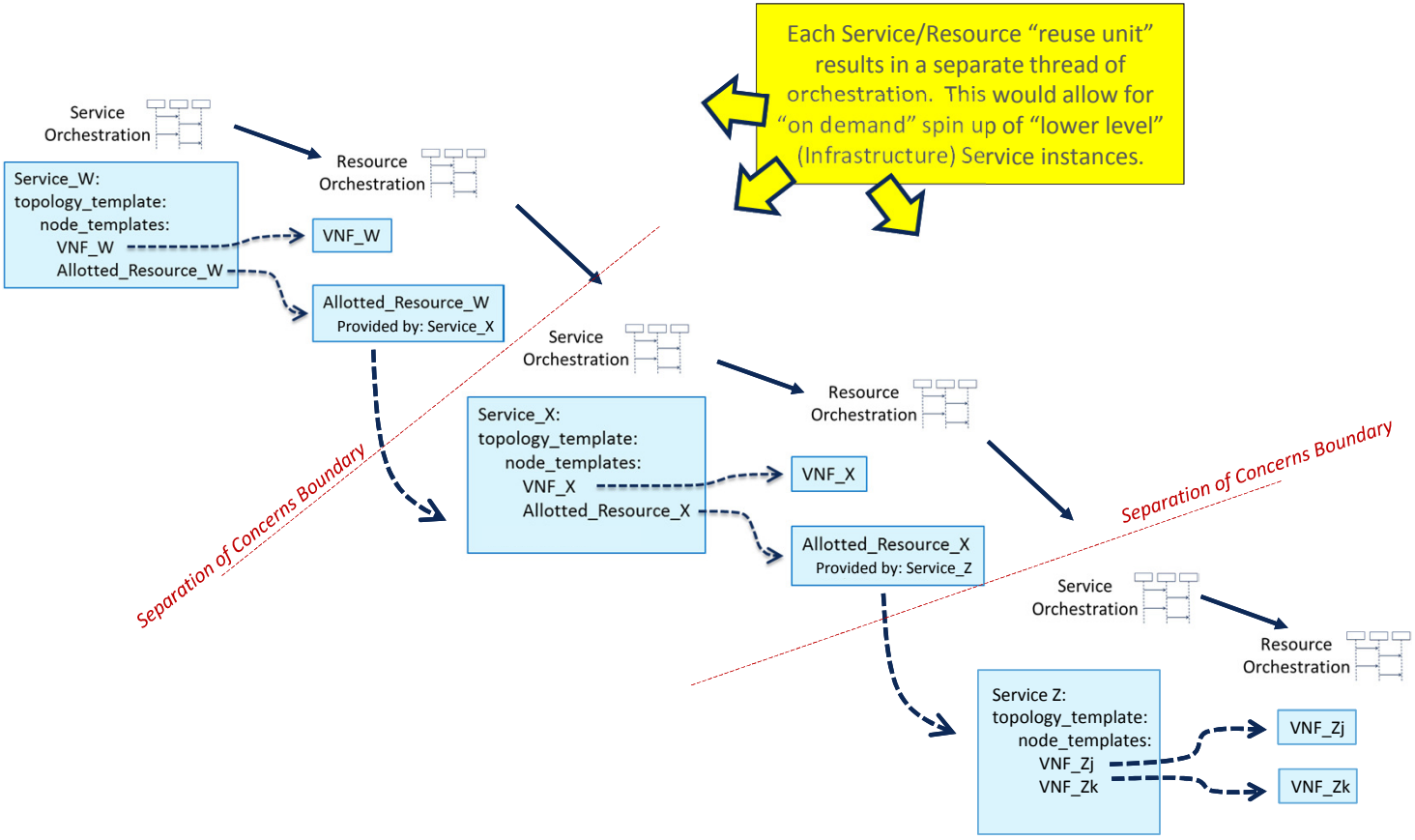
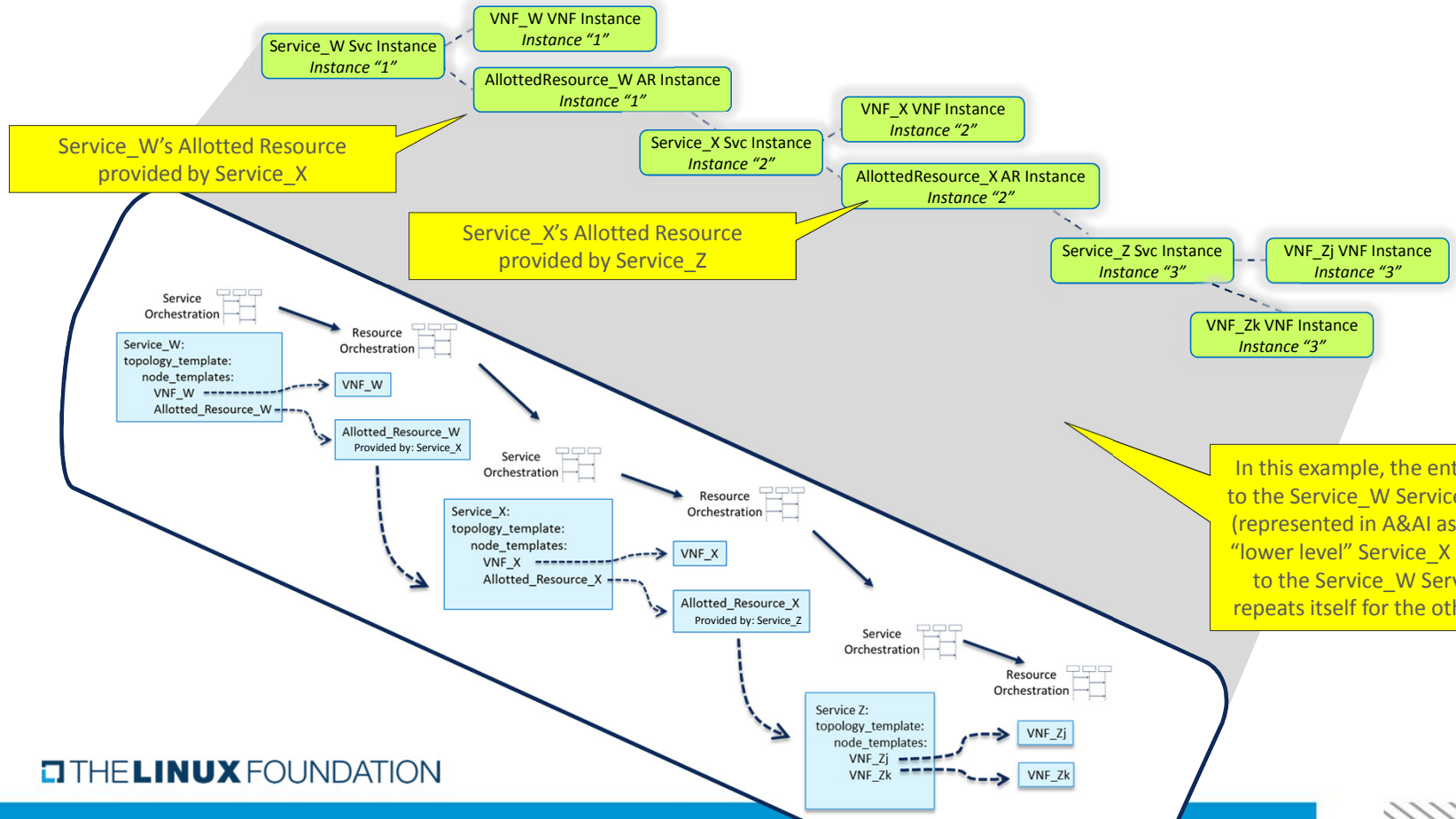
N-Level Run Time Nesting? Let The Service Providers Decide

Service_W Modeling Example 1



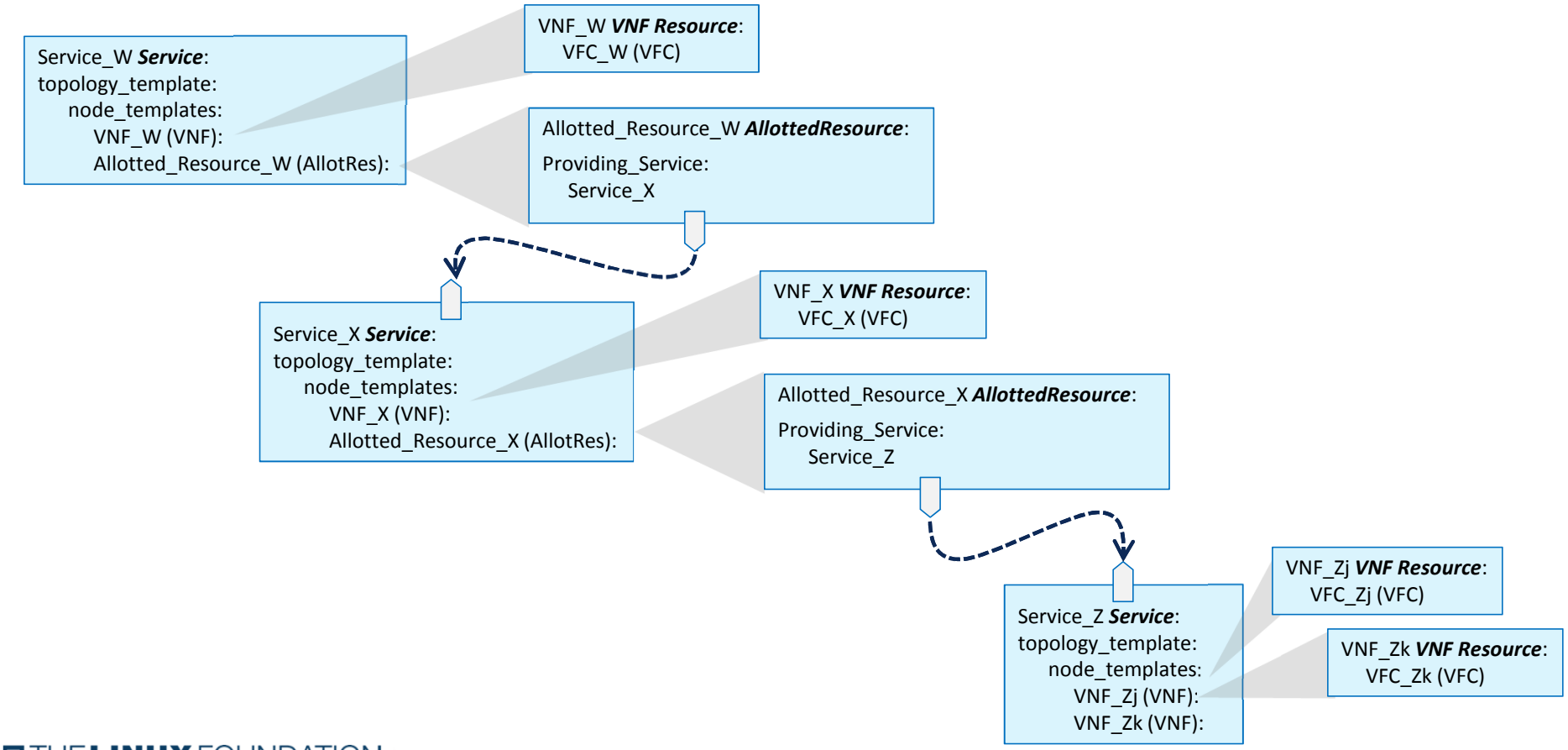
Could be extended to allow multiple "higher level Services" to each have a "share" of a "lower level Service's" instance

A&AI Instance Representation of Service_W Example 1

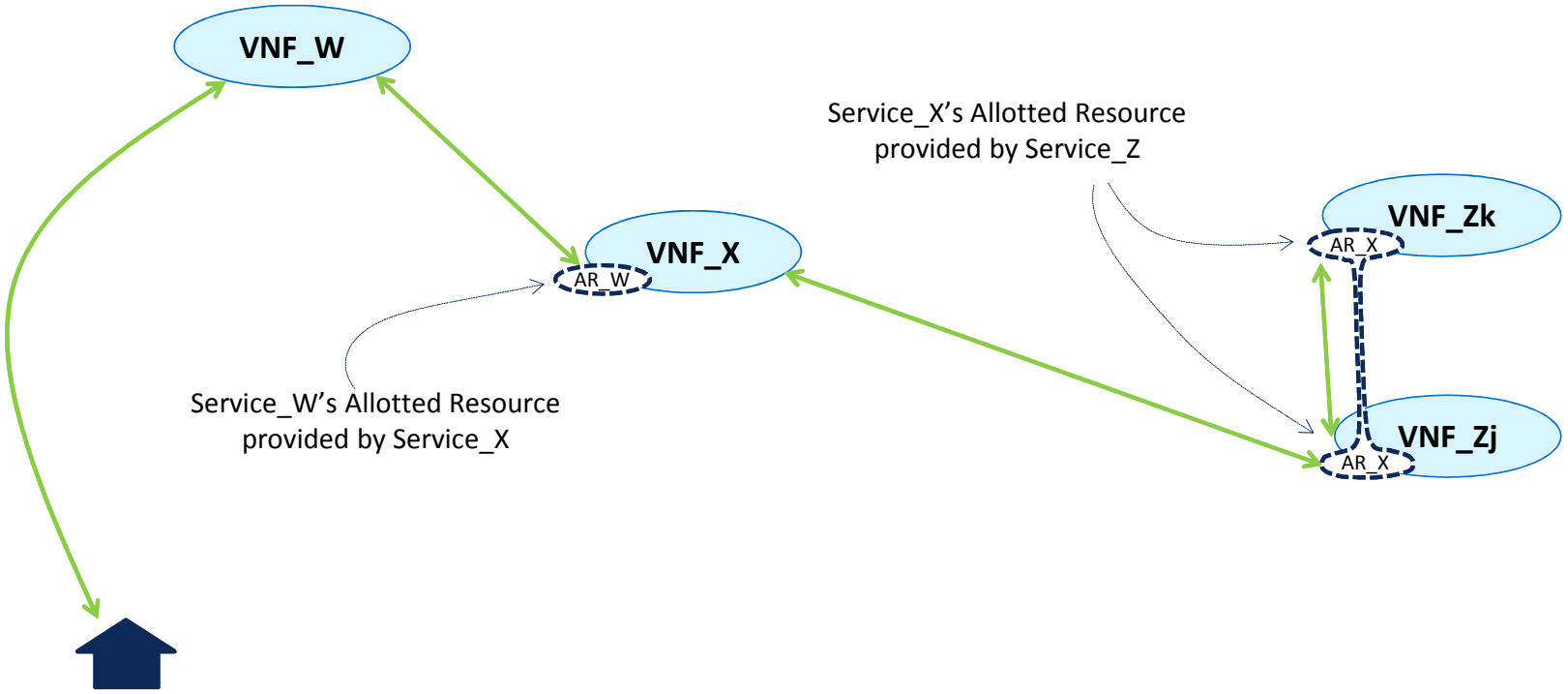


In this example, the entirety of VNF_W is dedicated to the Service_W Service Instance, but only a portion (represented in A&AI as AllottedResource_W) of the "lower level" Service_X Service Instance is dedicated to the Service_W Service Instance. This pattern repeats itself for the other Service Instances shown.

SDC Model View

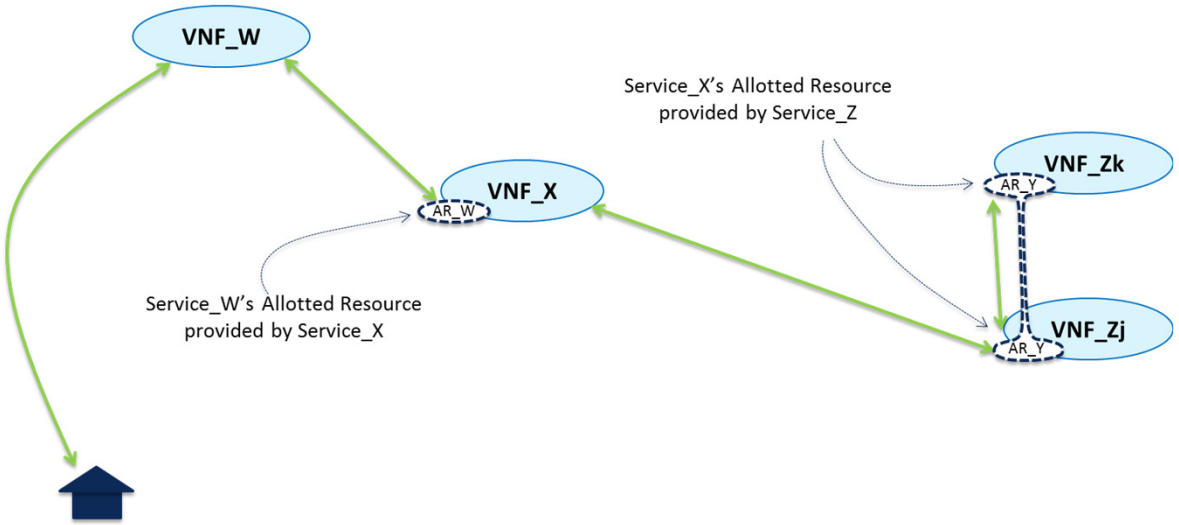


“VNF Chaining” Data Flow for Service_W Example 1



Modeling Network Latency Homing Constraints for Allotted Resources

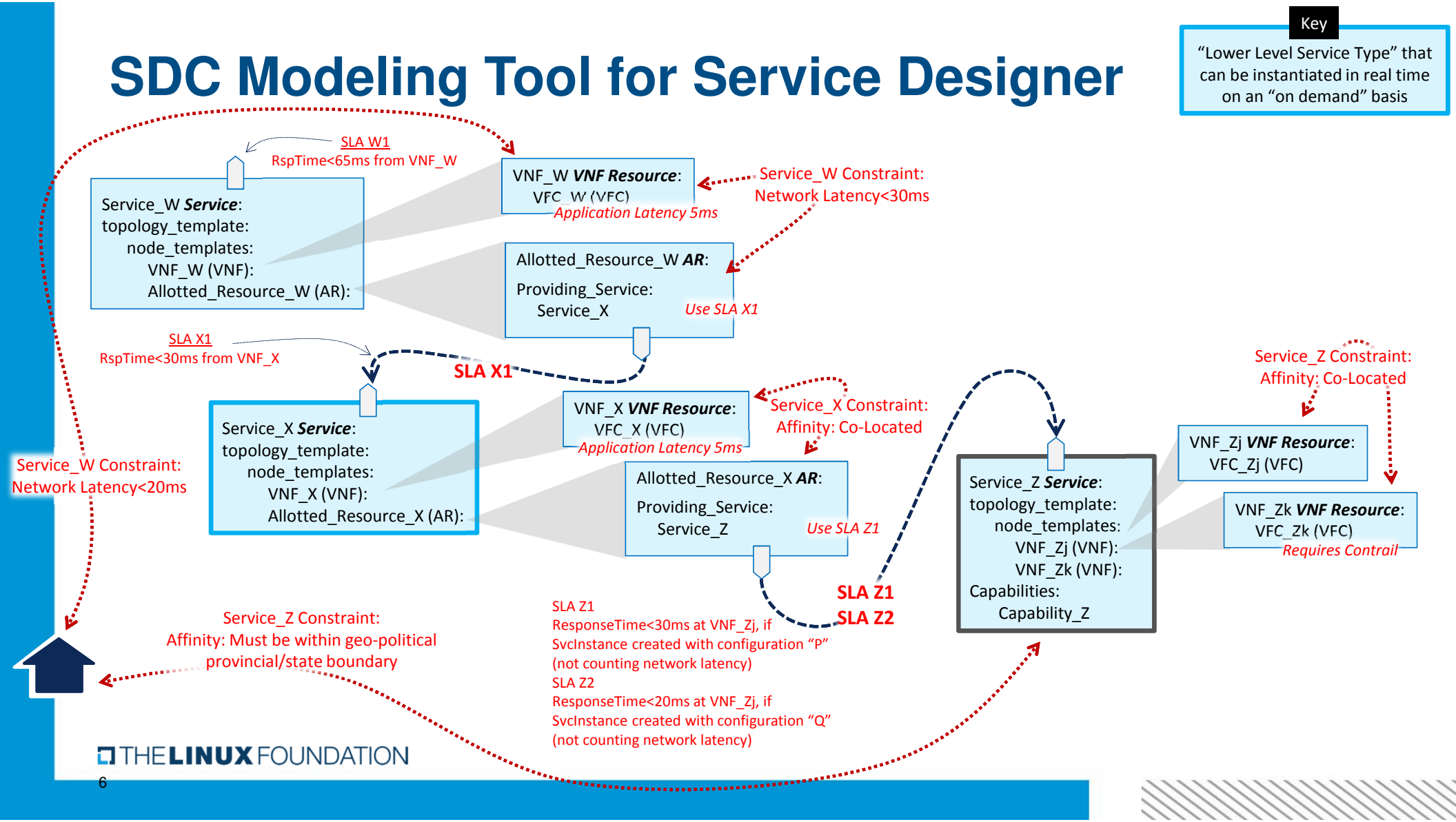
If Service_W is sensitive to network latency beween VNF_W and the VNF_X that hosts AR_W, then the homing algorithm will need to select only VNF_X instances that meet the Service_W constraint. However, we don't want to write any homing (or any other) policies for Service_W in terms of the internal structure of the underlying "lower order" Service type.



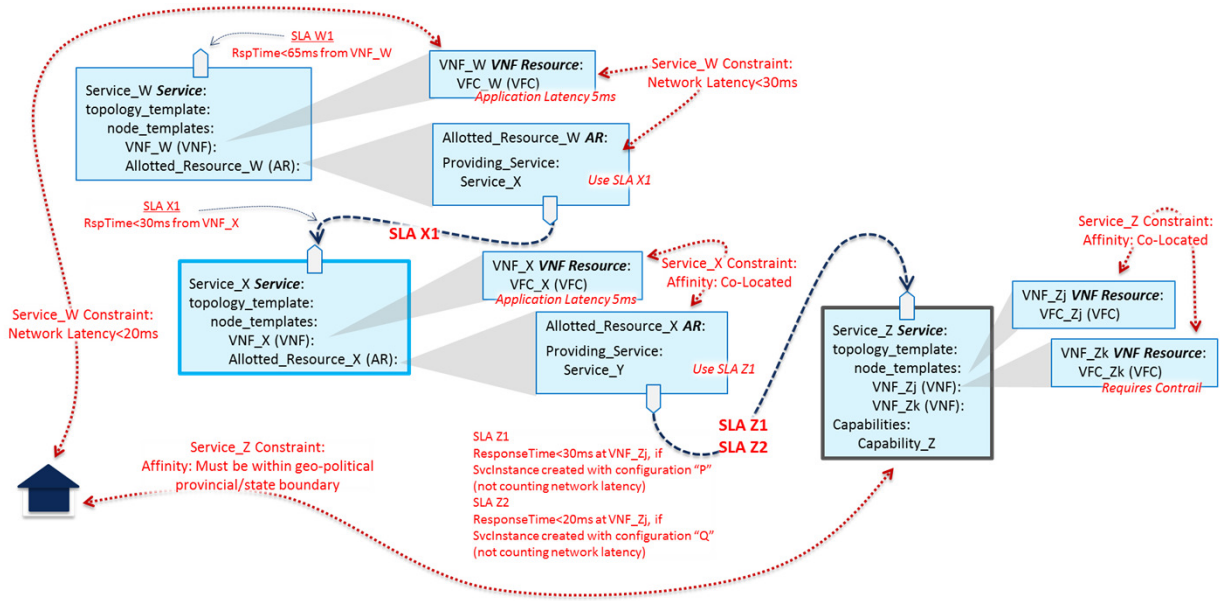
We can instead write the network latency constraint in terms of two policies, one a Service_W policy and one a Service_X policy.

Specifically, we will define the concept of an "SLA" that the lower order service will advertise. We will give the "higher order" Service a policy as to which SLA it requires from the "lower order" Service type. We will have the "lower order" Service type have a policy which indicates from which VNF the SLA is measured (mirroring the data path)

SDC Modeling Tool for Service Designer



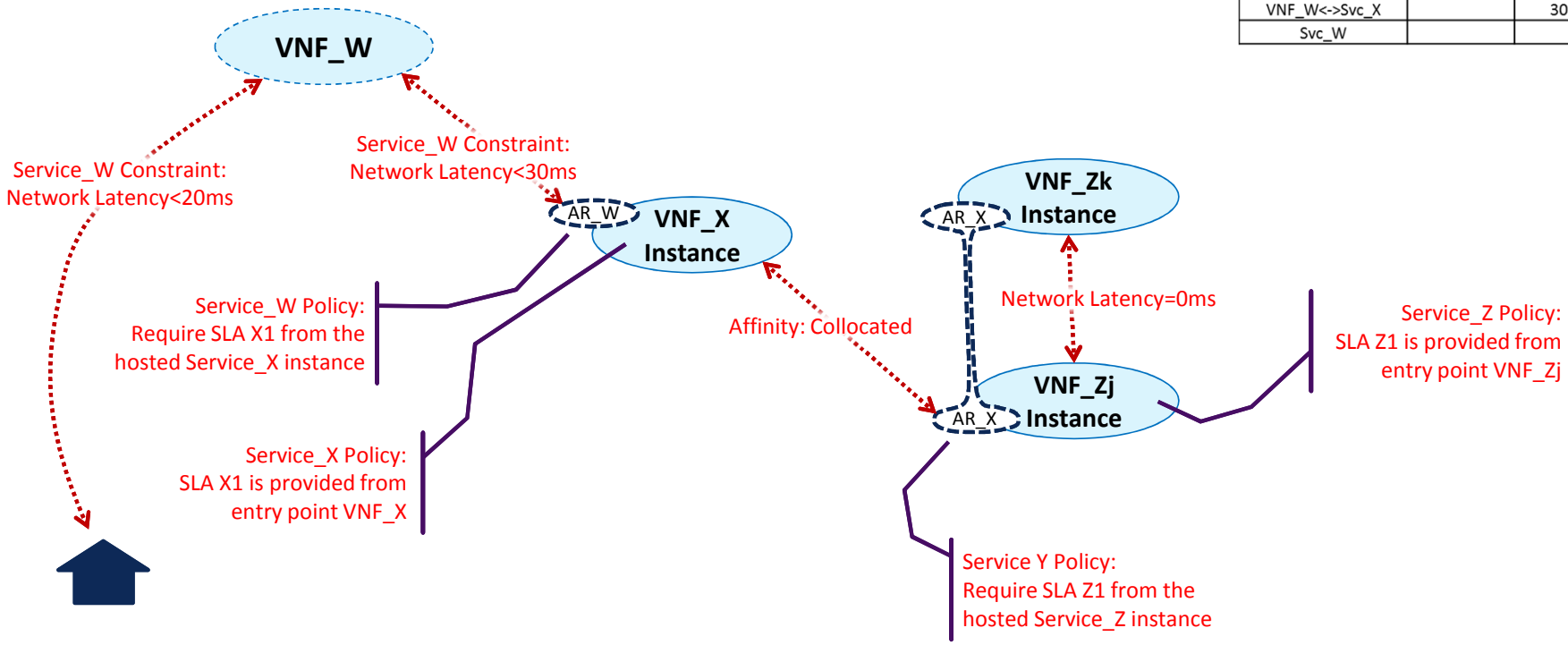
SDC Modeling Tool for Service Designer



	Application Latency	Network Latency	Cumulative Latency	Advertised SLA
VNF_Zk				
VNF_Zj				
VNF_Zk <-> VNF_Zj		0		
Svc_Z	Unknown			20
VNF_X	5			
VNF_X<->Svc_Z		0		
Svc_X			25	30
VNF_W	5			
VNF_W<->Svc_X		30		
Svc_W			65	65

SDC Homing Policy Calculator

	Application Latency	Network Latency	Cumulative Latency	Advertised SLA
VNF_Zk				
VNF_Zj				
VNF_Zk <-> VNF_Zj		0		
Svc_Z	Unknown			20
VNF_X	5			
VNF_X<->Svc_Z		0		
Svc_X			25	30
VNF_W	5			
VNF_W<->Svc_X		30		
Svc_W			65	65



Decomposition and Homing Approach

Note that, from a Service_W perspective, homing involves finding a cloud instance suitable for a new VNF_W instance such that the constraint:

Latency: [geographic point on map] <-> VNF_W < 20 ms

(where the geographic point is the location of the residence), and such that the "Network Latency" constraint of "VNF_W <-> AR_W < 30ms" is met. This involves knowing that the Providing Service for AR_W is Service_X. This processing would require decomposition to have created the Service_W rows in the decomposition example. If an appropriate cloud instance and Service_X service instance is found, then homing is complete.

However, if no such Service_X instance exists (i.e., OOF Service_X homing thread returns an exception), homing can determine that a new one should be created "on demand." In such a case, we want to take a separation of concerns approach whereby the Service_W homing thread can delegate down to a Service_X homing thread for further solutioning.

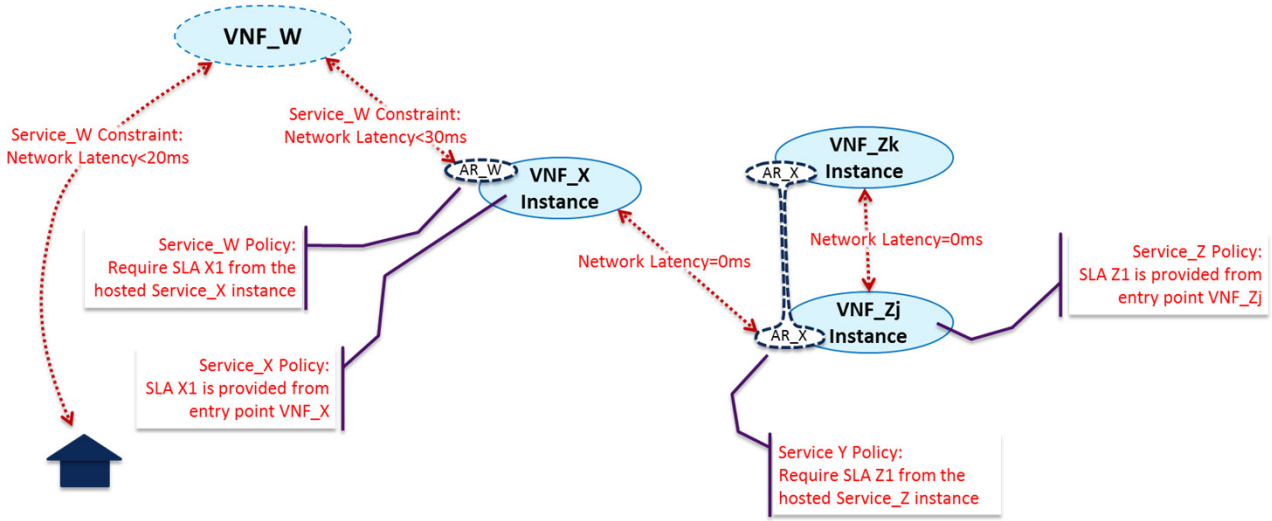
Homing of a new Service_X instance would similarly require that decomposition of Service_X had been performed (i.e., the Service_X rows oof the table). One goal of this homing is to find a cloud instance suitable for a new VNF_X instance such that the Service_W constraint that VNF_W <-> AR_W < 30ms is met. However, we don't want to violate separation of concerns between the Service_W and the Service_X processing, so we will have the Service_W homing thread pass to the Service_X homing thread a constraint that is written in terms that Service_X can understand:

Latency: [geographic point on map] <-> Service_X < 30 ms

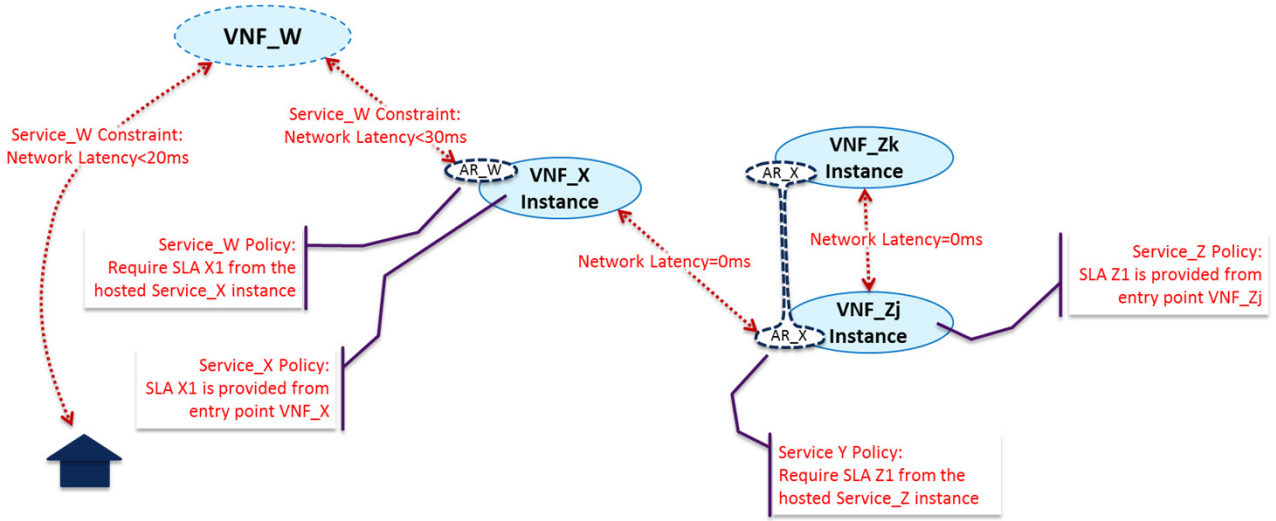
(where the geographic point is a "proposed" location of the VNF_W yet to be created). Because the optimal location of VNF_W has not yet been determined, this will likely require that the Service_W homing thread spawns multiple Service_X homing thread to solve the overall homing problem. Step by step processing can be seen on the following slides.

Homing Example Flow

SO sends OOF a Service_W homing request, providing as an input constraint the geographic location of the residence. OOF Service_W homing will comprise homing for VNF_W and AR_W. OOF homing for VNF_W will find eligible VNF_W cloud instances that meet the 20ms latency constraint with the residence. OOF homing for AR_W will, for each eligible VNF_W cloud instance, want to find the set of Service_X instances to provide that AR_W functionality that meet the 30ms latency constraint with that cloud instance.



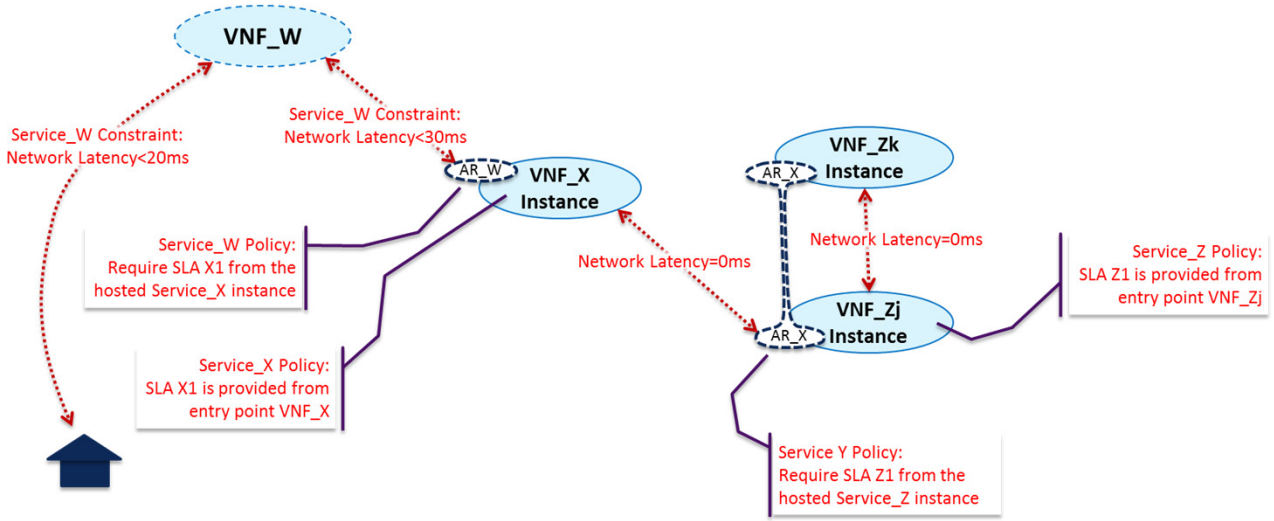
Homing Example Flow (Cont'd)



However, we want to maintain a separate of concerns approach, and the Service_W processing thread shouldn't know the implementation of AR_W such that it can measure latency to it. (This can be best seen in the Service_Z example to the right.) Thus, we will have the Service W OOF homing request thread delegate selection of the optimal Service_X instance to a subtending Service_X OOF thread. Thus, OOF can be seen as (logically) calling itself in parallel with multiple Service_X homing requests. Each such request can be seen as providing as input constraints the geographic location of the associated eligible VNF_W cloud instance and the SLA needed, in this case SLA X1.

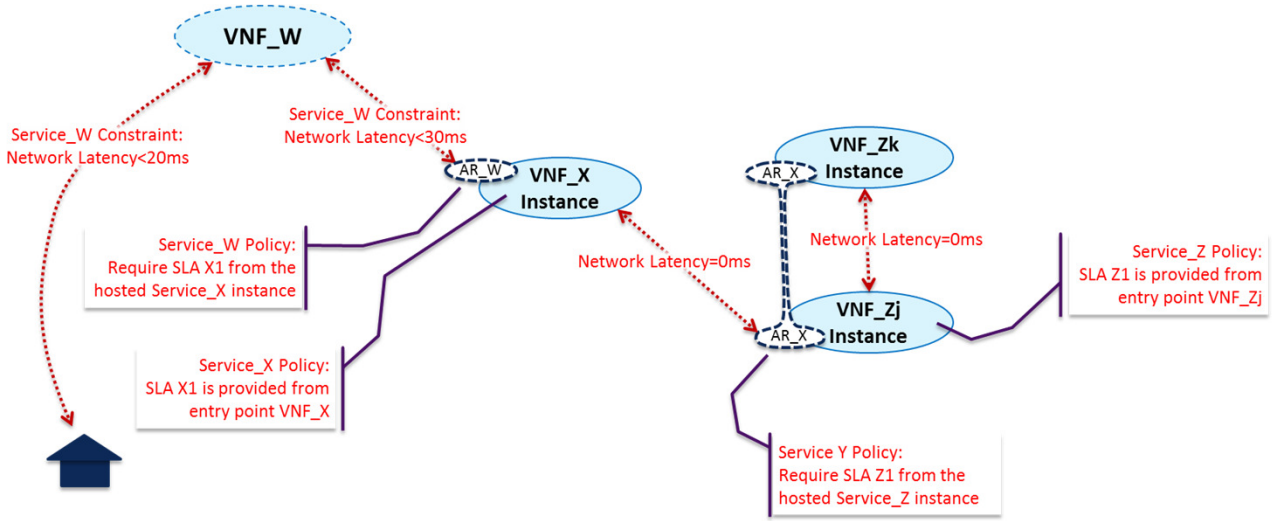
Homing Example Flow (Cont'd)

Service_X homing knows that SLA X1 is measured from an entry point on VNF_X. Thus Service_X homing is comprised of looking for the optimal Service_X instance whose VNF_X instance is within 30ms of the input geographic location. If at least one such Service X instance is found, homing is done (except for optimization).



If no such Service_X instance can be found, then homing will determine whether the Service_X service definition allows for dynamic instantiation of new Service_X instances. In this case we will assume "yes", so OOF would determine whether a new Service_X could be instantiated such that all constraints can be met.

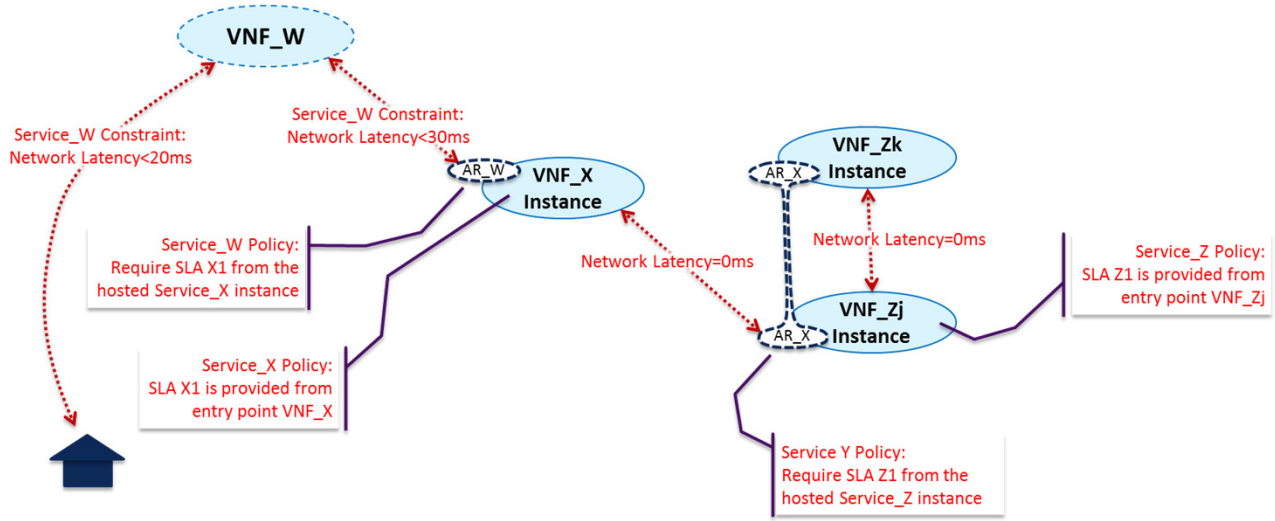
Homing Example Flow (Cont'd)



OOF Service_X homing will comprise homing for VNF_X and AR_X. OOF homing for VNF_X will find eligible VNF_X cloud instances that meet the 30ms latency constraint with the input geographic location. OOF homing for AR_X will, for each eligible VNF_X cloud instance, want to find the set of Service_Z instances to provide that AR_X functionality that meet the 0ms latency constraint with that cloud instance. The pattern recurs, however, that Service_X has no business understanding whether the 0ms latency constraint should be measured from VNF_Zj or VNF_Zk, or even in fact that there exists a VNF_Zj or VNF_Zk. In order to maintain separation of concerns, homing of AR_X will be delegated to a subtending request thread delegate selection of the optimal Service_Z instance to a subtending Service_Z OOF thread.

Homing Example Flow (Cont'd)

Service_Z homing will thus search for eligible Service_Z instances such that the 0ms constraint is measured from the input geographical location (in this case the potential cloud instance location for VNF_X) to an available VNF_Zj instance (the point from which the Service_Z SLA is measured.



If no such Service_Z instance can be found, then homing will determine whether the Service_Z service definition allows for dynamic instantiation of new Service_Z instances. In this case we will assume "no", so the OOF Service_Z homing thread would return an exception to the calling Service_X homing thread. Such an exception would likely not result in failure of the entire Service_W homing, but rather simply result in pruning a branch of the overall potential homing solution tree.

Decomposition Structure for Service_W Example 1

Svc Type	Rsc Type	AR Prov Svc	Advertised SLA	Homing Constraints	Capab Svc Struct
Service_W			W1: RspTime 65ms end to end	Ntw Latency: VNF_W <-> AR_W < 30ms	
Service_W	VNF_W			Ntw Latency: Residence <-> VNF_W< 20ms	
Service_W	AR_W	Service_X		Require SLA X1 from Service_X instance	

AR_W Provider's Svc Struct

Svc Type	Rsc Type	AR Prov Svc	Advertised SLA	Homing Constraints	Capab Svc Struct
Service_X			X1: RspTime 30ms end to end	Affinity: VNF_X, AR_X Co-Located	
Service_X	VNF_X				
Service_X	AR_X	Service_Z		Require SLA Y1 from Service_Z instance	

AR_X Provider's Svc Struct

Svc Type	Rsc Type	AR Prov Svc	SLA Policies	Homing Constraints	Capab Svc Struct
Service_Z			Z1: <30ms with config "Q" Z2: <20ms with config "P"	Affinity: VNF_Zj, VNF_Zk Co-Located	
Service_Z	VNF_Zj				
Service_Z	VNF_Zk				

Homing Solution Example for Service_W Example 1

Service Type	Resource Type	Allotted Resource Provider Service	Provider Service Struct	Homing Solution
Service_W	VNF_W			Cloud_Region_1
Service_W	Allotted_Resource_W	Service_X		Instantiation_Needed

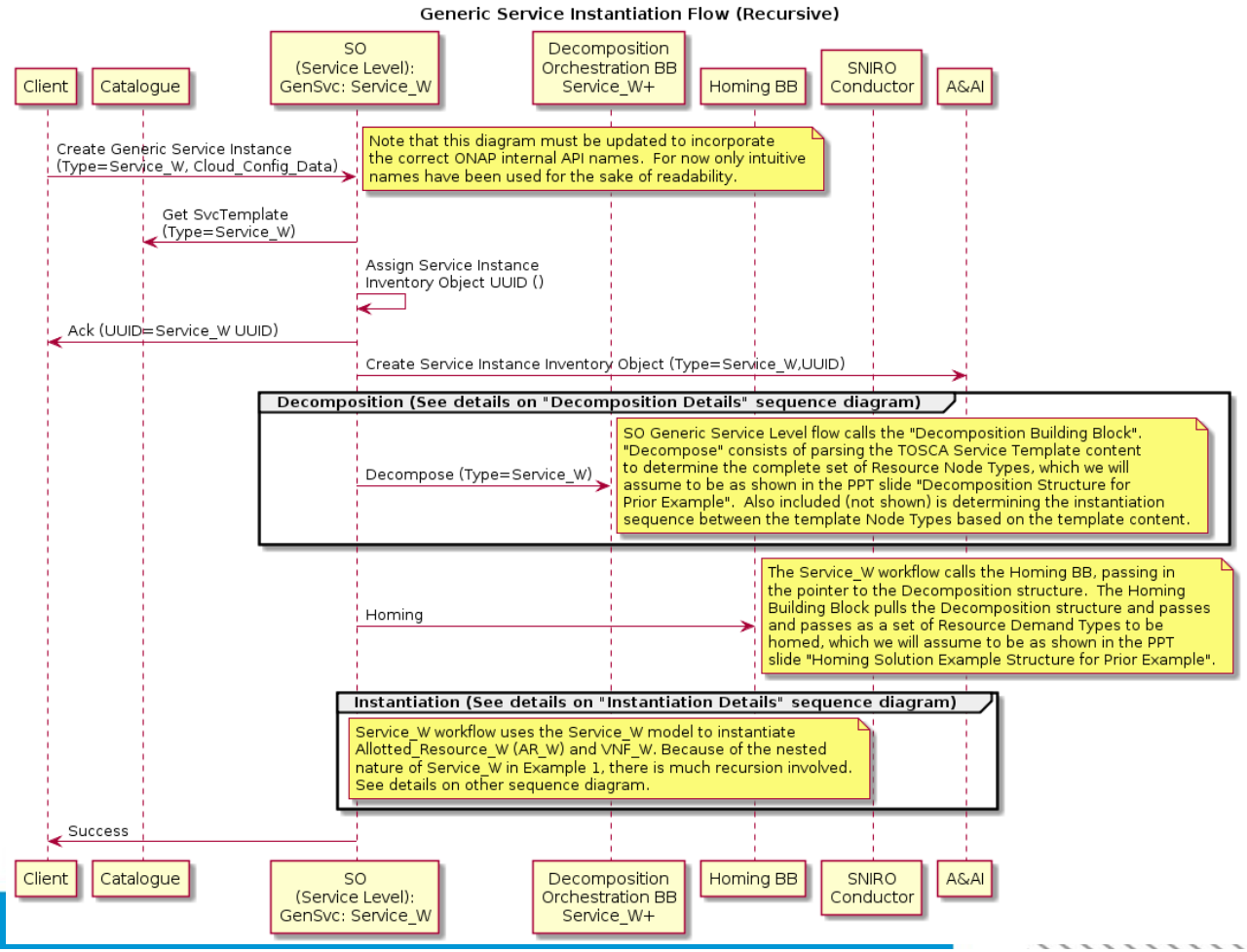
AR_W Homing Structure

Service Type	Resource Type	Allotted Resource Provider Service	Provider Service Struct	Homing Solution
Service_X	VNF_X			Cloud_Region_2
Service_X	Allotted_Resource_X	Service_Z		Service Z Instance Id

AR_X Homing Structure

Service Type	Resource Type	Allotted Resource Provider Service	Homing Solution
Service_Z	VNF_Zj		As Exists
Service_Z	VNF_Zk		As Exists

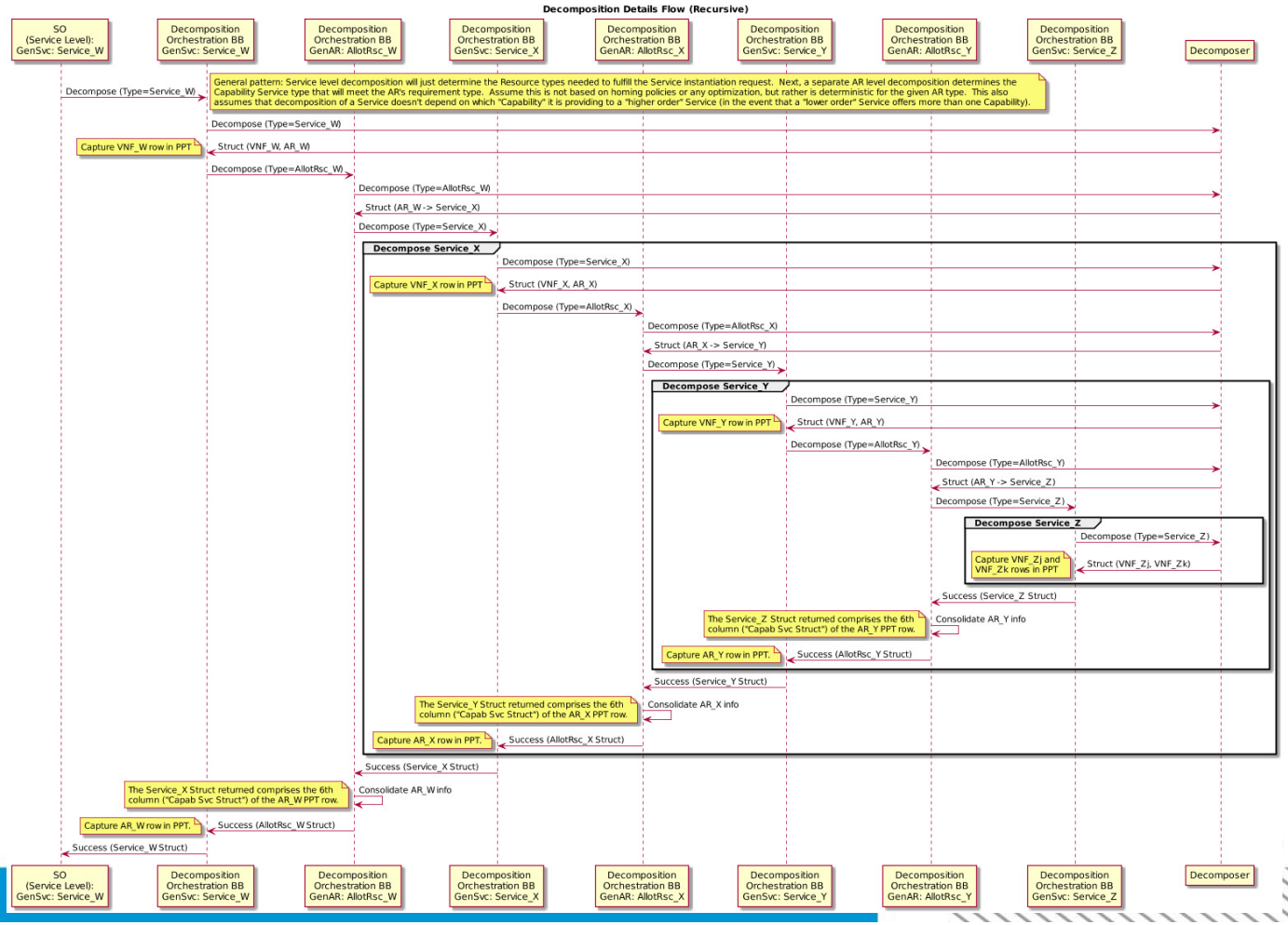
Generic Service Level Flow for Service_W Example 1



onap_uc_Generic_Service_Recursive_p1.html

Decomposition Detail Flow for Prior Example

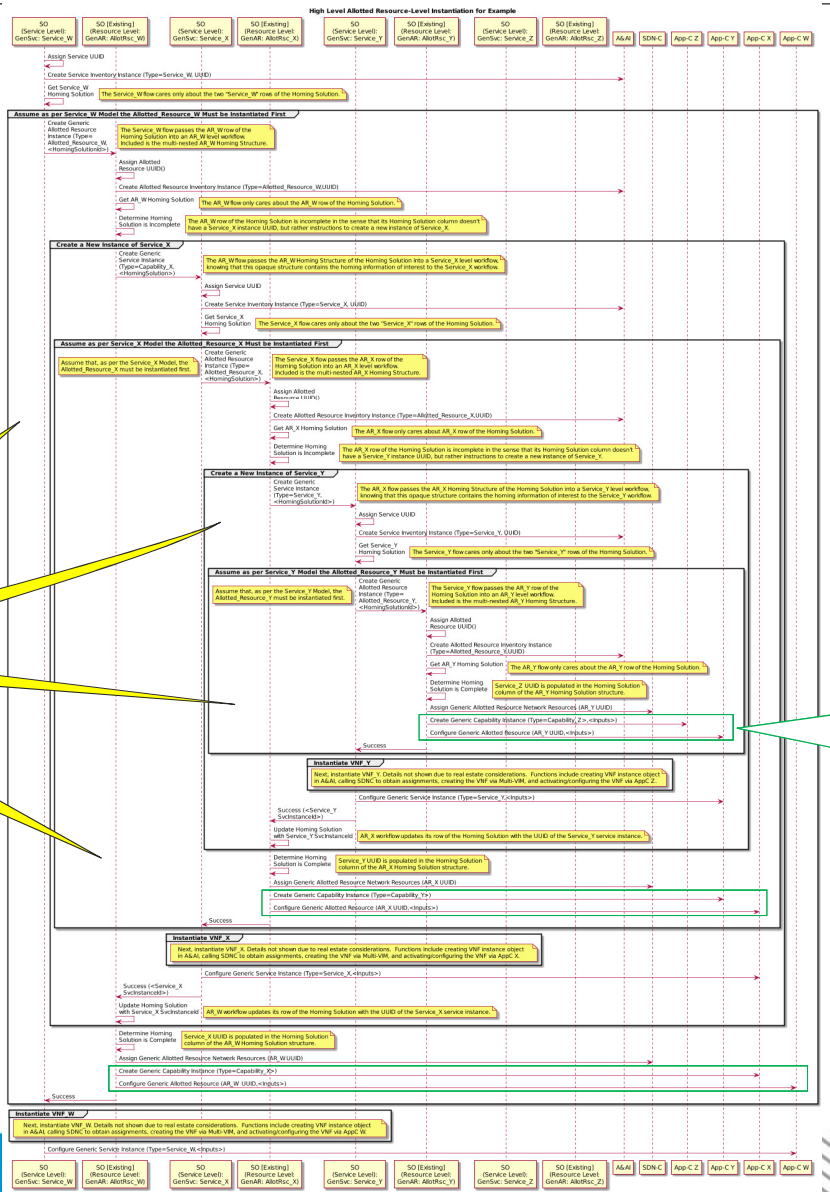
onap_uc_Generic_Service_Decomp_p1.html



Instantiation Detail Flow for Service_W Example 1

Scale PPT to 300% to view detail. ☺

Note recursion in the process



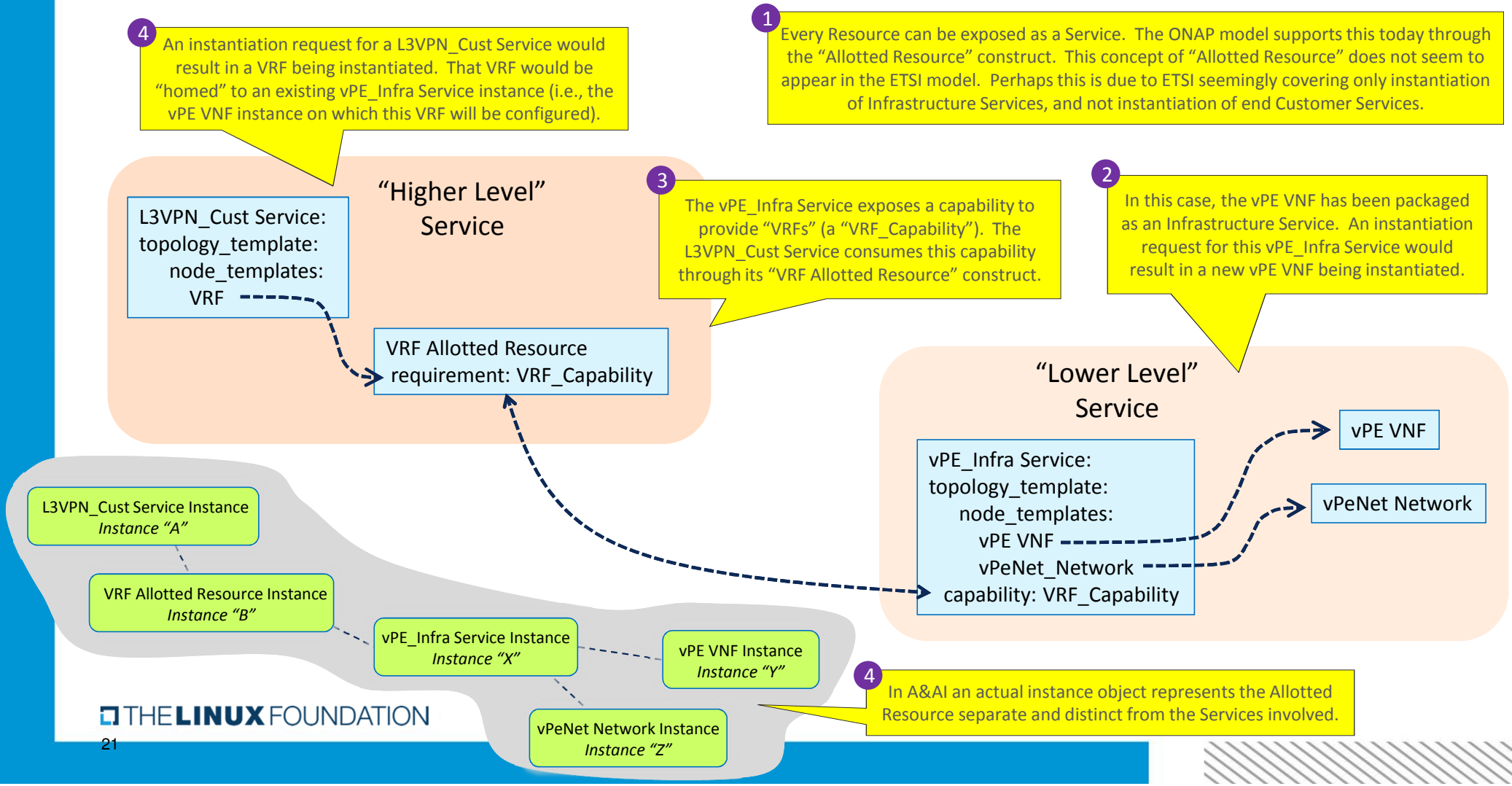
Potential insertion points for Cloudify? Do we want to do that as part of this POC? Or just stick with Camunda?

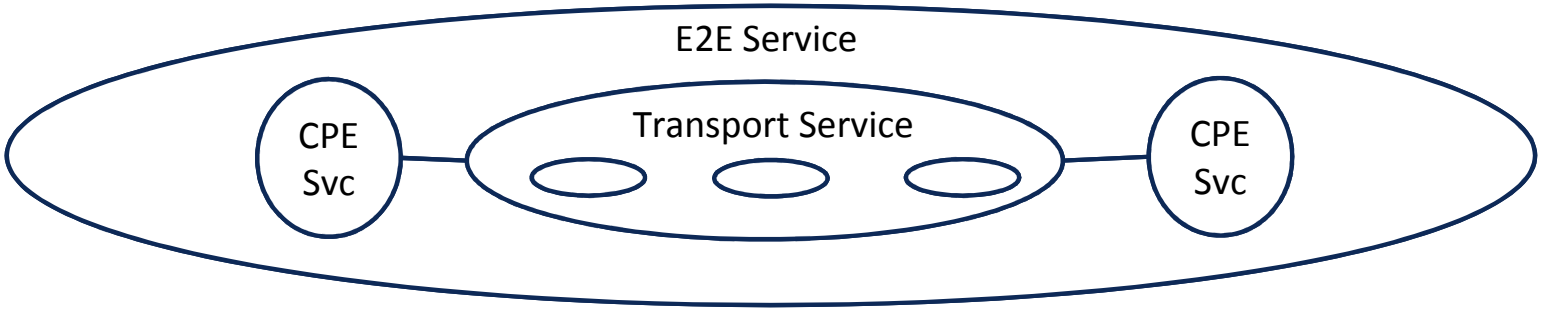
onap_uc_Generic_Resource_VNF_Recursive.html

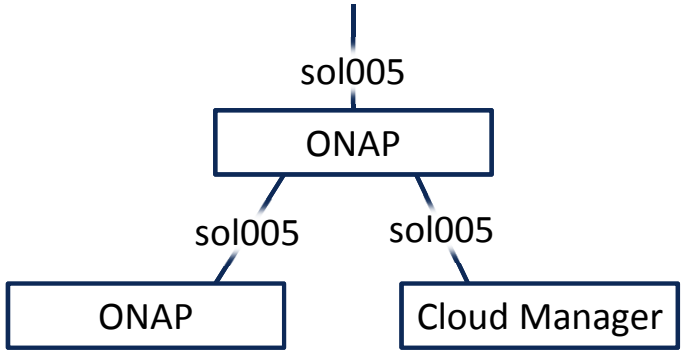
THE LINUX FOUNDATION

Backup Slides

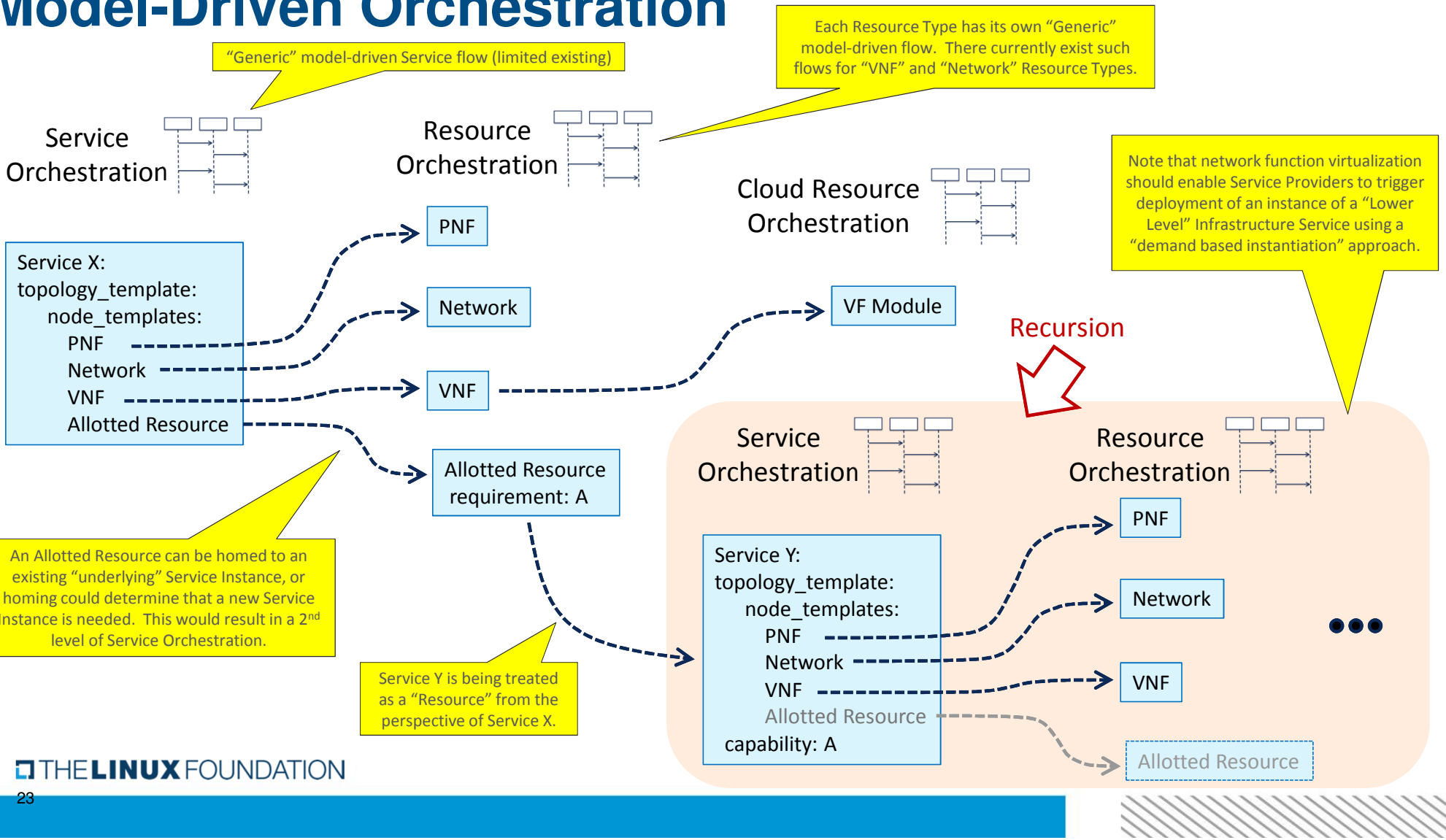
Allotted Resources – vPE/VRF Example




E2E Service
CPE
Svc
Transport Service
CPE
Svc


sol005
ONAP
sol005
sol005
ONAP
Cloud Manager

Model-Driven Orchestration



N-Level Run Time Nesting? Let The Service Providers Decide

Service_W Modeling Example 2

For the case whereby a "higher level Service" consumes the entirety of a "lower level Service's" instance, SDC should support the Design Time ability to construct an "upper" Service Definition from other Services definitions via substitution mapping (a.k.a., "Compile Time Nesting")

Service W:
topology_template:
node_templates:
VNF_W
Service_X

Service X:
topology_template:
node_templates:
VNF_X
Service_Y

Service Y:
topology_template:
node_templates:
VNF_Y
Service_Z

Service Z:
topology_template:
node_templates:
VNF_Zj
VNF_Zk

Substitution Mapping

Design Time

Distribution Time & Run Time

The "lower level Services" would not be visible at Distribution Time. Hence a "flattening" of the run-time orchestration would result.

Service_W:
topology_template:
node_template:
VNF_W
VNF_X
VNF_Y
VNF_Zj
VNF_Zk

VNF_W
VNF_X
VNF_Y
VNF_Zj
VNF_Zk

A&AI Instance Representation of Service_W Modeling Example 2

Service_W Svc Instance *Instance "1"*

VNF_W VNF Instance *Instance "1"*
VNF_X VNF Instance *Instance "2"*
VNF_Y VNF Instance *Instance "3"*
VNF_Zj VNF Instance *Instance "4"*
VNF_Zk VNF Instance *Instance "4"*

The entirety of these VNFs, not just a portion thereof, are dedicated to the Service_W Service Instance.