Raytheon
BBN Technologies

Supporting Resiliency and Timeliness in Edge Applications with Dispersed Computing

Aaron Paulos

TREC4CPS Invited Talk

December 11, 2018

This work is being supported by DARPA under Contract No. HR0011-17-C-0049. The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Outline

Raytheon
BBN Technologies

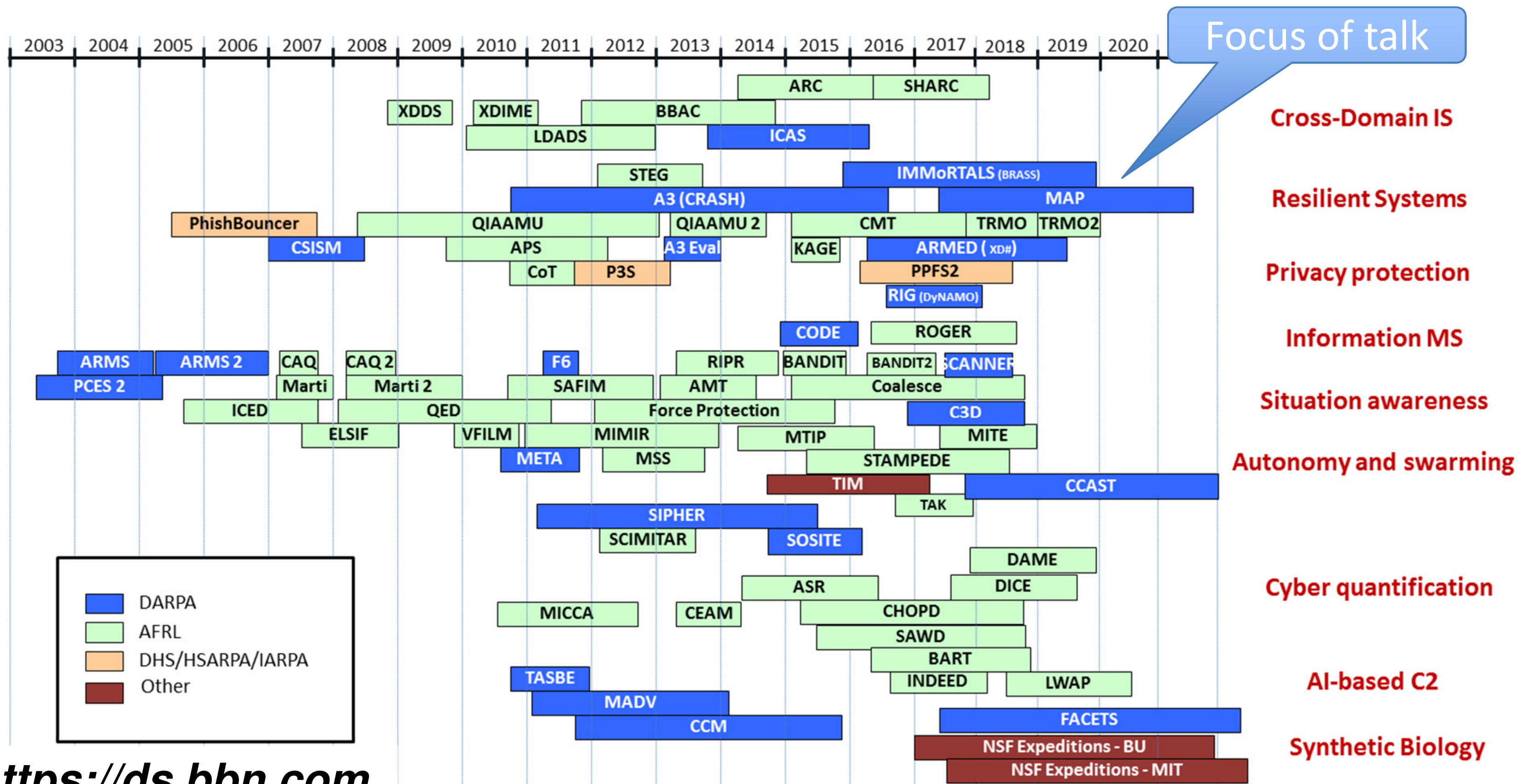
- **Introductions**
- Research Thrusts and CONOPS
- Initial Proof of Concept Results
- Takeaways

BBN Distributed Systems Group

Raytheon
BBN Technologies

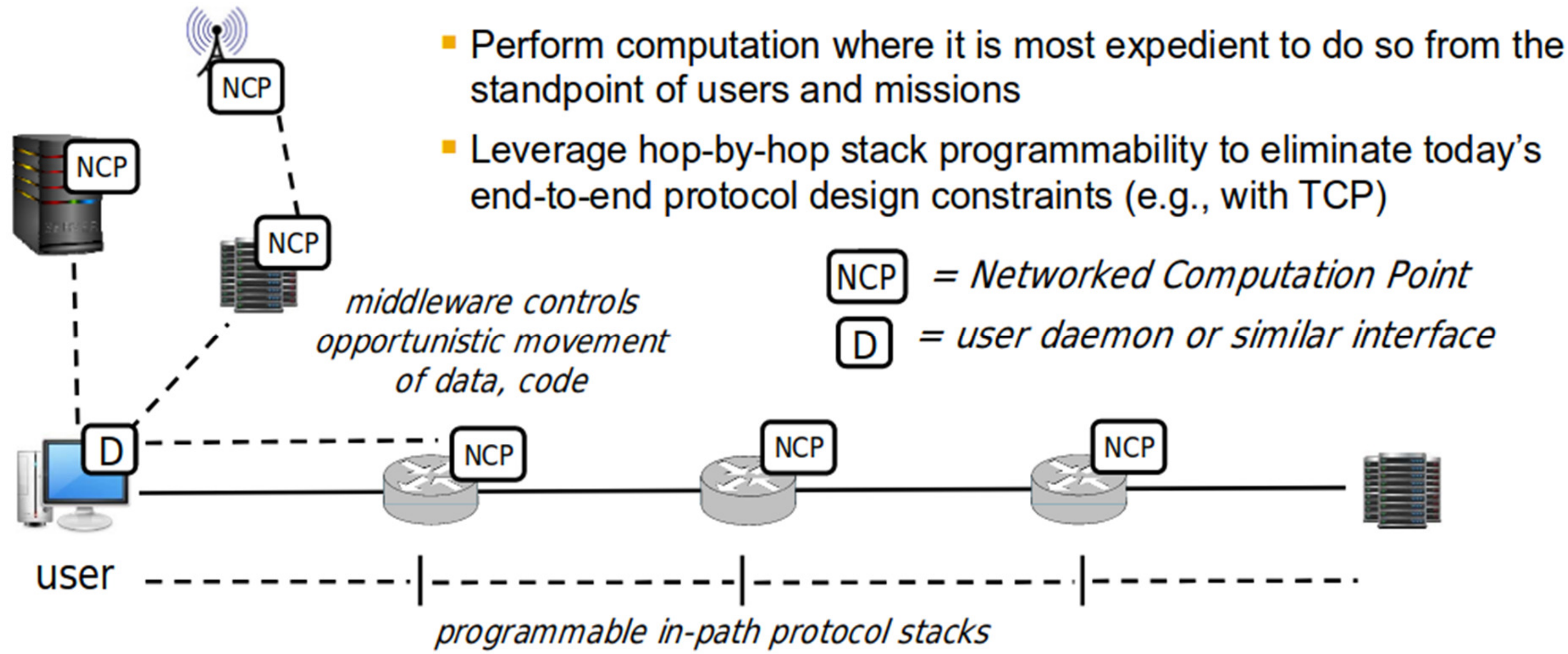
A group of ~20 researchers ***focused on dynamically changing, adaptive distributed systems problems***

- E.g., information management, cyber-physical, cross-domain and federation
- With focus on QoS, Survivability and Resilience, Modeling and Analytics



https://ds.bbn.com

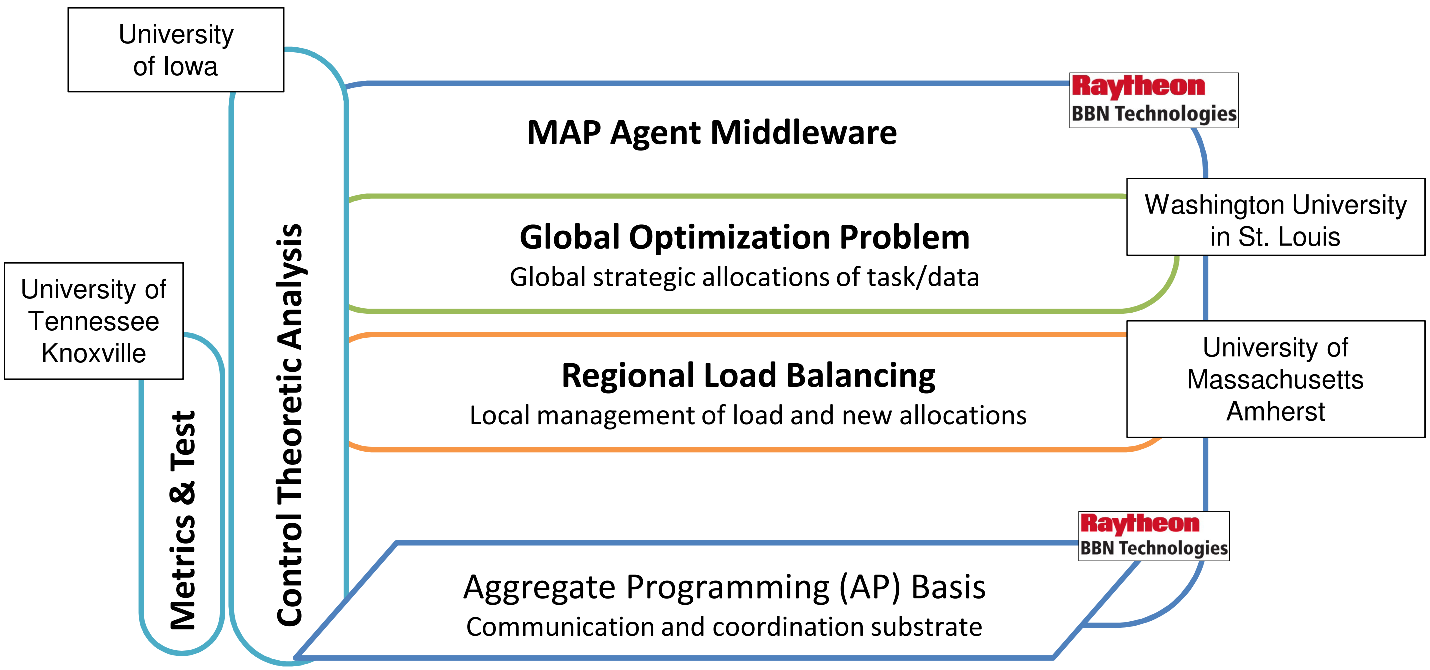
DARPA Dispersed Computing Proposer's Day Vision [1]



NCP Examples: Programmable Network Elements, Sensors w/ Embedded Programmable Processors, Micro/Nanoclusters, Smart Phones

[1] Dispersed Computing, https://www.darpa.mil/attachments/DispersedComputingProposersDay.pdf

Mission-Aware Adaptive Placement of Data and Tasks



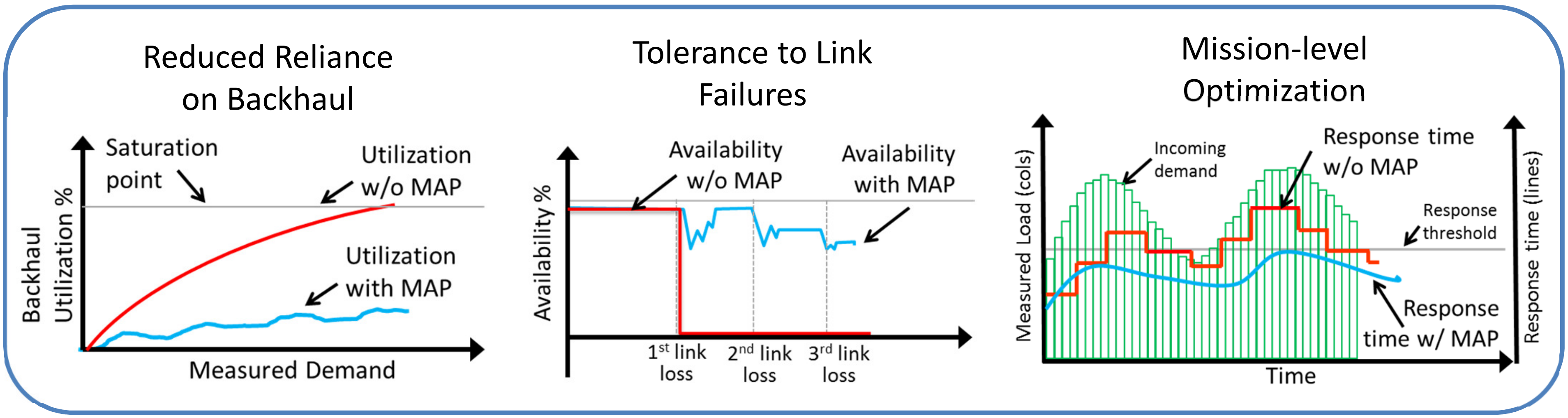
- Scalable, multi-layer, distributed resource mgmt. system
  - Calculate optimal mission-oriented task/data placement
  - Adaptively migrate application tasks and data
  - Monitor and manage compute/storage resources

Potential Benefits for Edge Computing and Cyber Physical Systems

Raytheon
BBN Technologies

- Algorithmic and middleware basis to ***reassign elasticity and load balancing into the network***
- Decentralize and disperse strategic and tactical decisions making to ***optimize bandwidth, CPU and storage use***
- Extensively embed mission-level requirements into decision making to ***maximize mission success/resilience***

Representative Benefits (Notional):

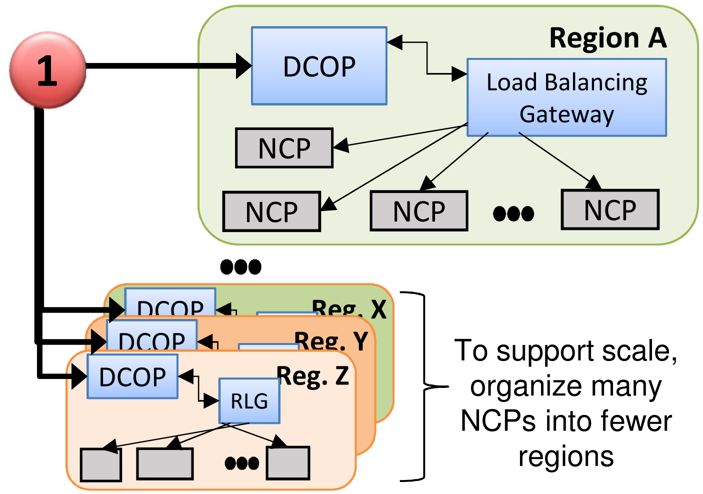


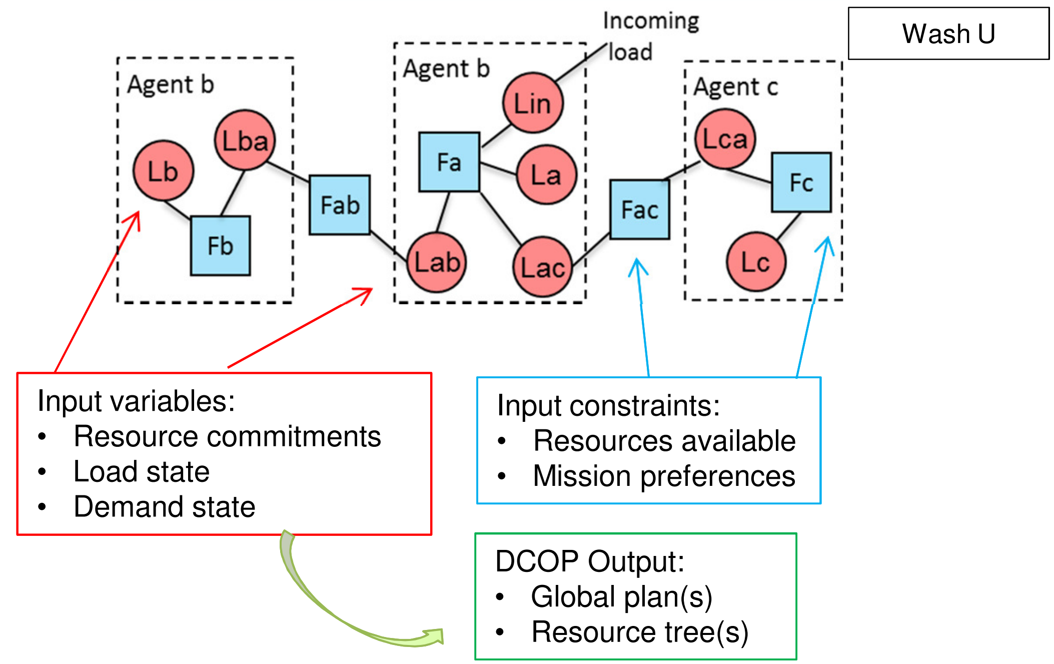
Outline

Raytheon
BBN Technologies

- Introductions
- **Research Thrusts and CONOPS**
- Initial Proof of Concept Results
- Takeaways

Research Thrusts and MAP Architecture

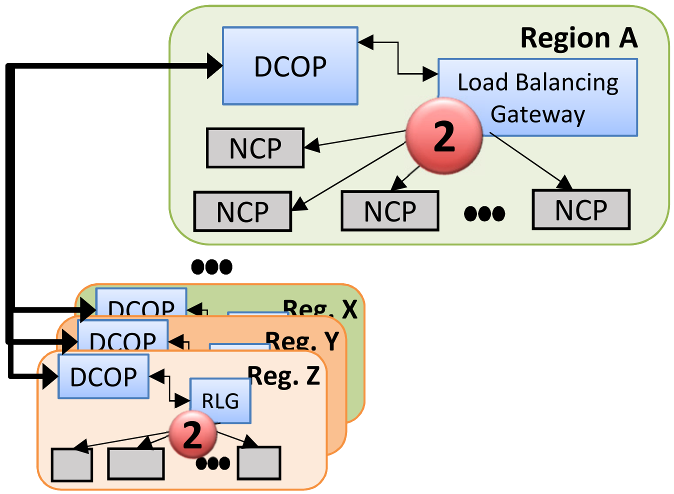


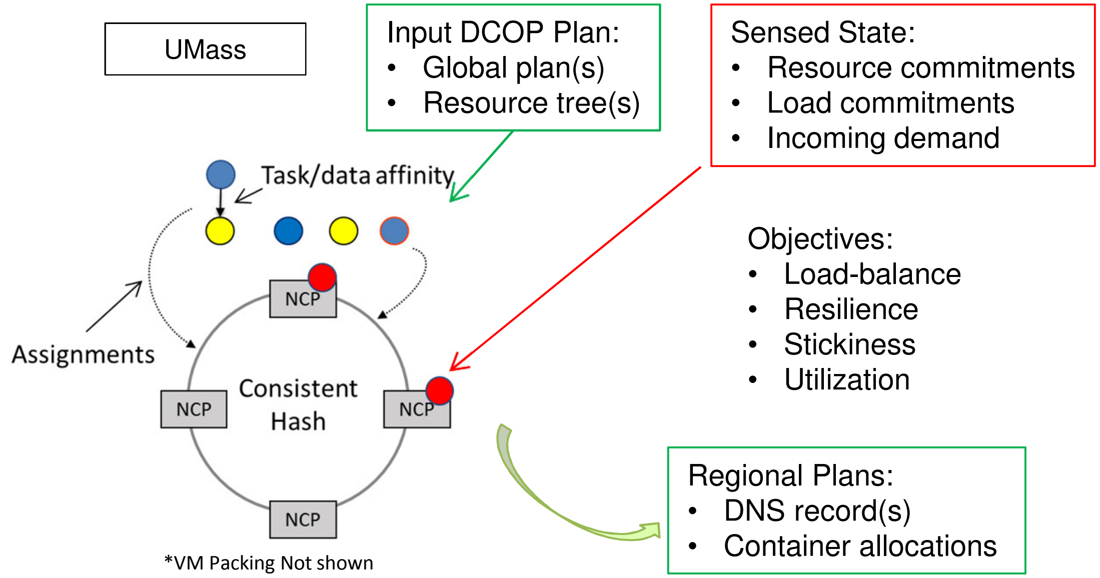


1 Extend ***DCOP to conduct multi-criteria optimizations that produce a pareto frontier of mission-focused solutions*** for many applications and objectives.

Research Thrusts and MAP Architecture

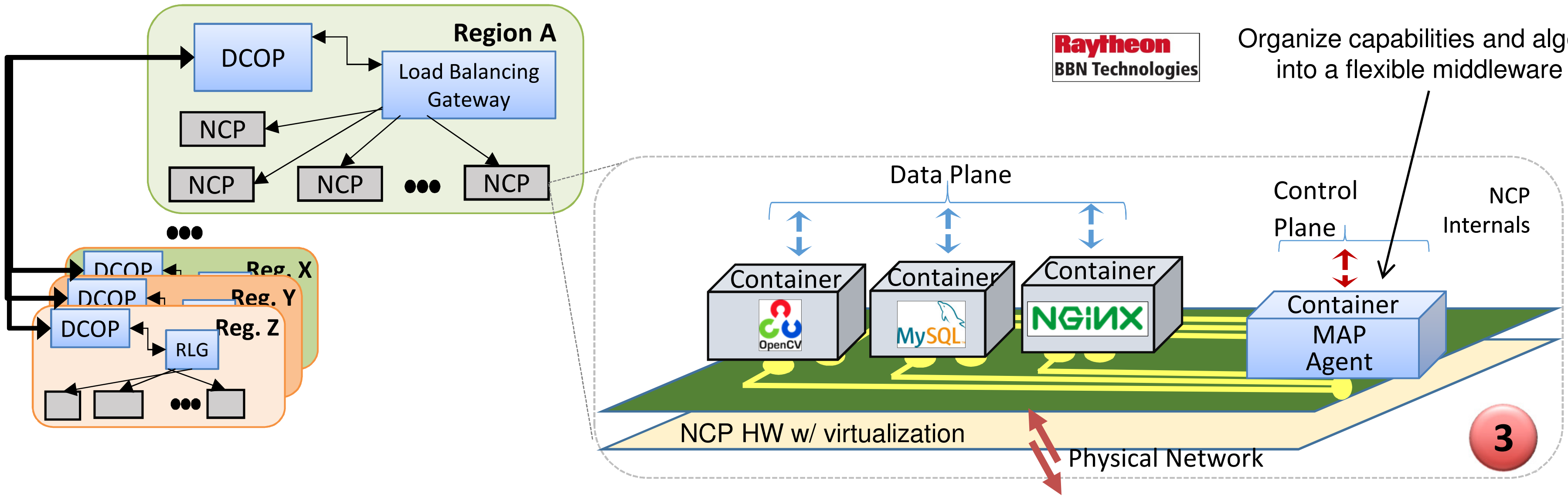
Raytheon BBN Technologies





2 Extend ***VM packing and consistent hashing techniques with mission affinities*** to enable ***fast in-region responses to failures and changes in demand***.

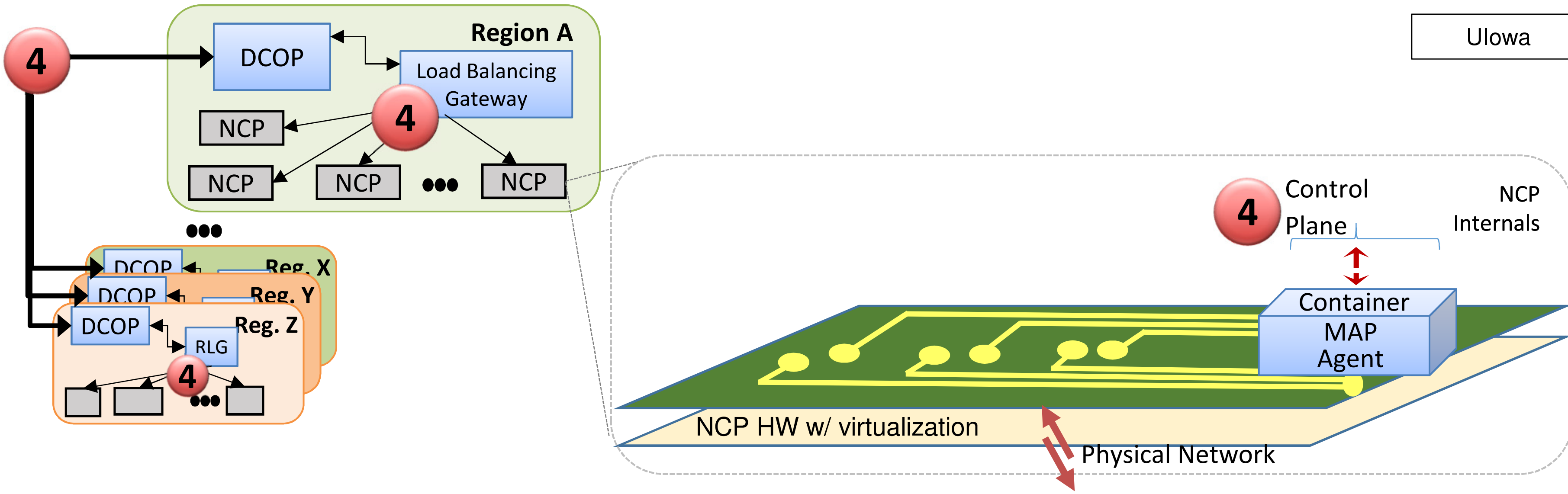
Research Thrusts and MAP Architecture



3 ***Develop middleware and an Aggregate Program (AP)*** to structure global and regional algorithms, sense and share state and failures, and manage tasks/data.

Research Thrusts and MAP Architecture

Raytheon BBN Technologies

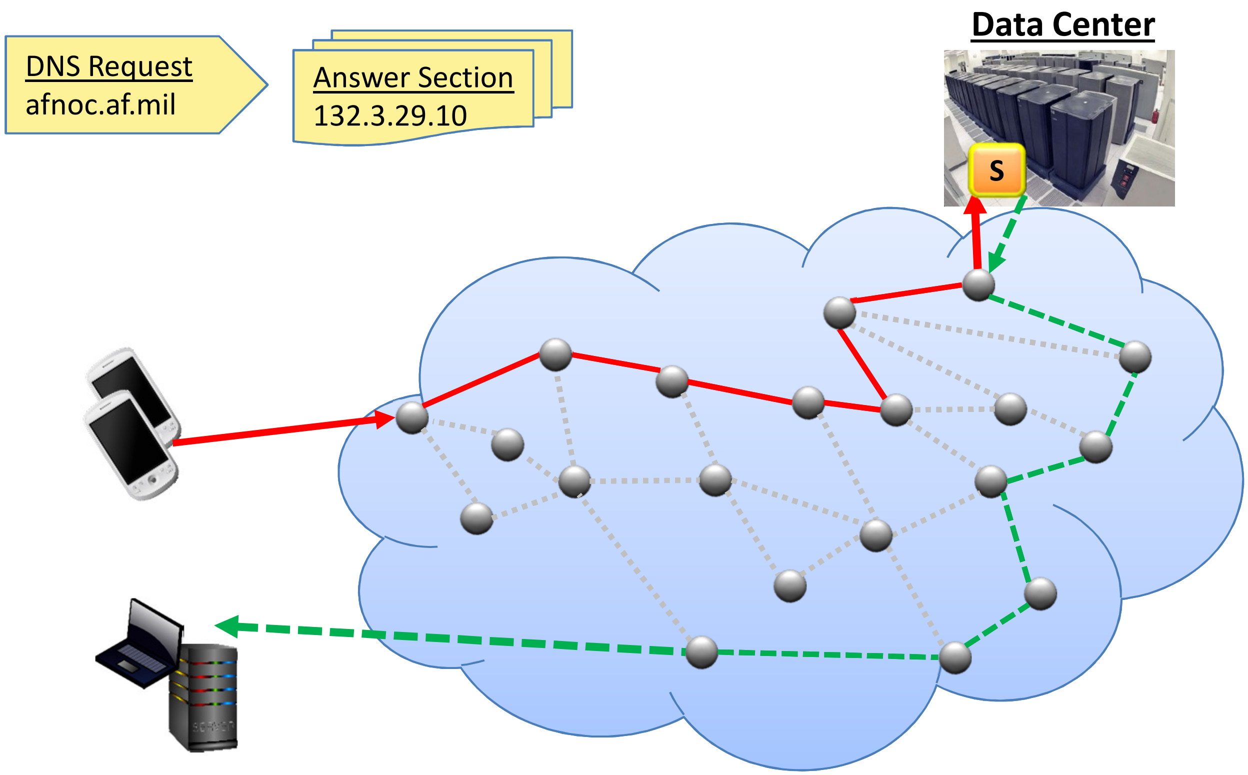


4 ***Apply control theoretic analysis*** to global/regional algorithms and the Agent at design time to ***identify potential sources of volatility that may destabilize MAP***.

Raytheon
BBN Technologies

Putting it all together!

Without DCOMP

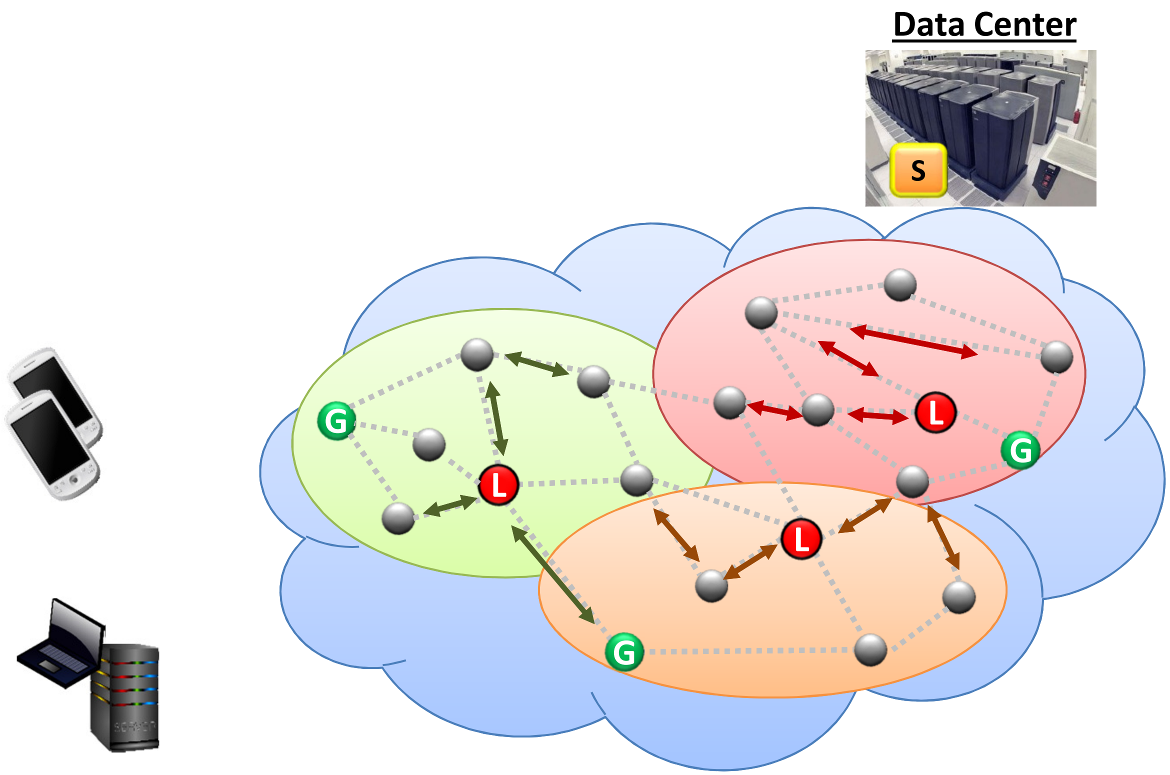


- Authoritative application resides in data center
- Long-haul links are used to access service

Prior to DCOMP/MAP, data and tasks are centralized in the data center

Concept of Operation (1/4)

Raytheon
BBN Technologies

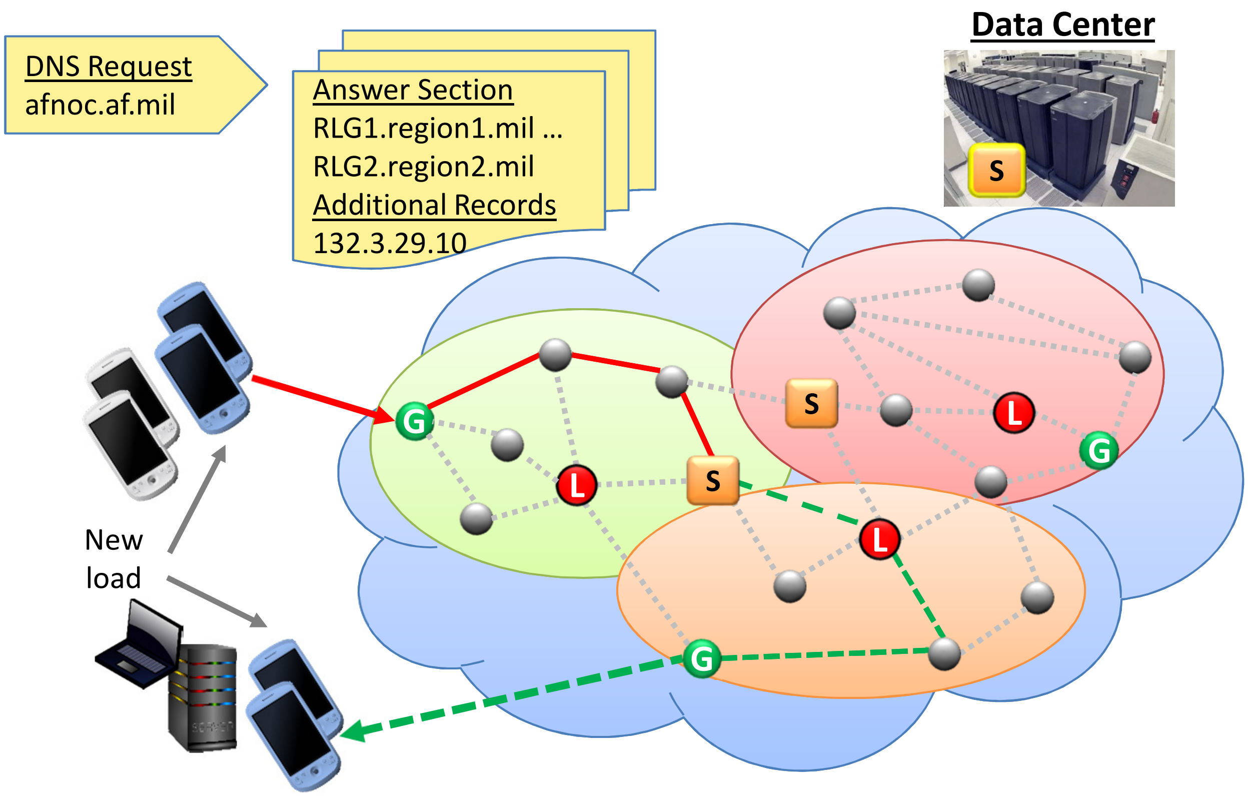


- Leaders and regions are designated
- Initial resource sharing begins

1 – Initially organize NCPs into regions, with ***region leaders*** and ***load-balancing gateways***

Concept of Operation (2/4)

Raytheon
BBN Technologies



First round of DCOP

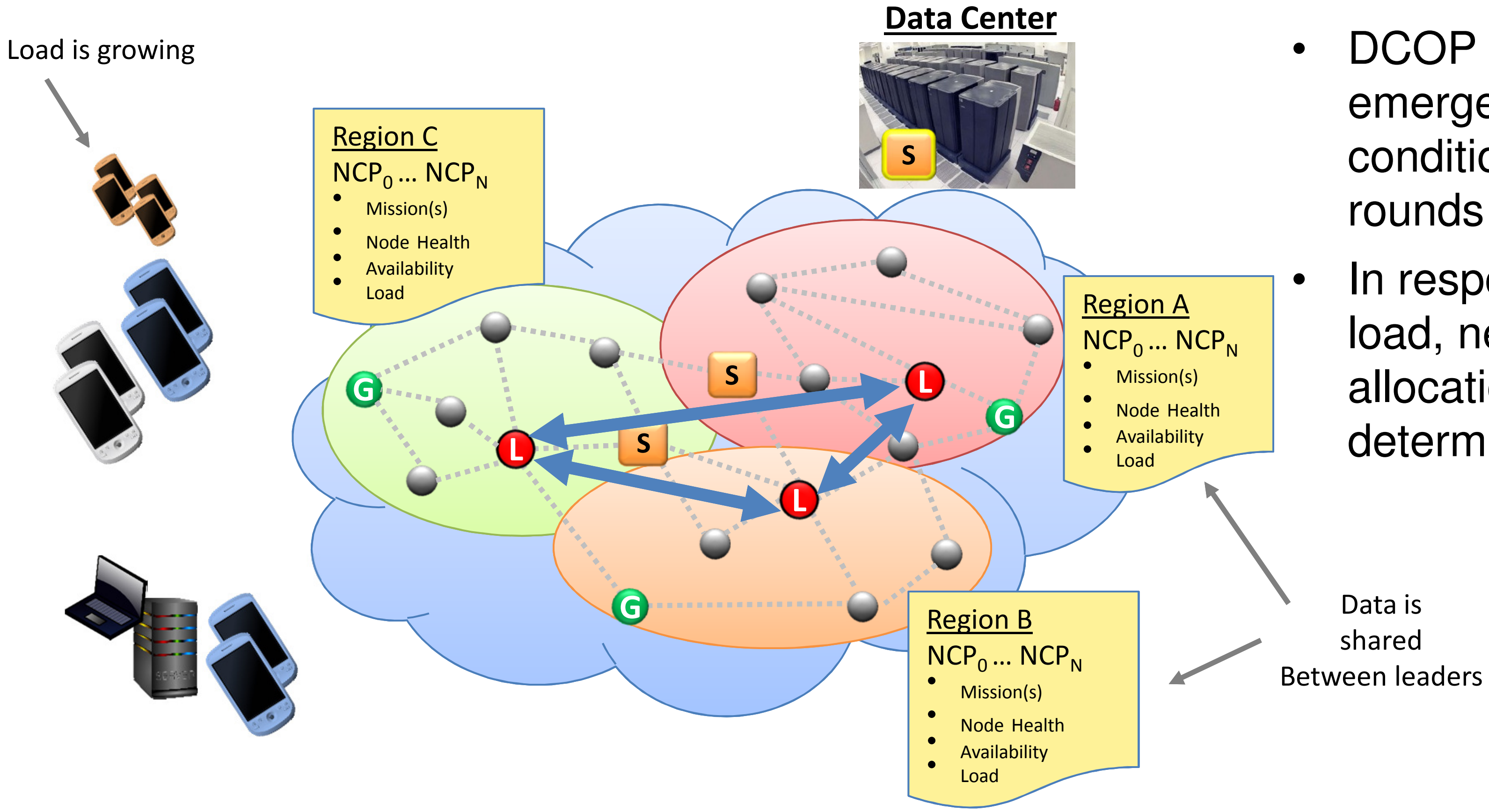
- Leaders deploy new services in network
- New load is balanced across new services

2 – Leaders will pre-position tasks/data at NCPs

- Incorporate mission requirements and application affinity
- Push out DNS zone updates to reference regional LBs

Concept of Operation (3/4)

Raytheon
BBN Technologies



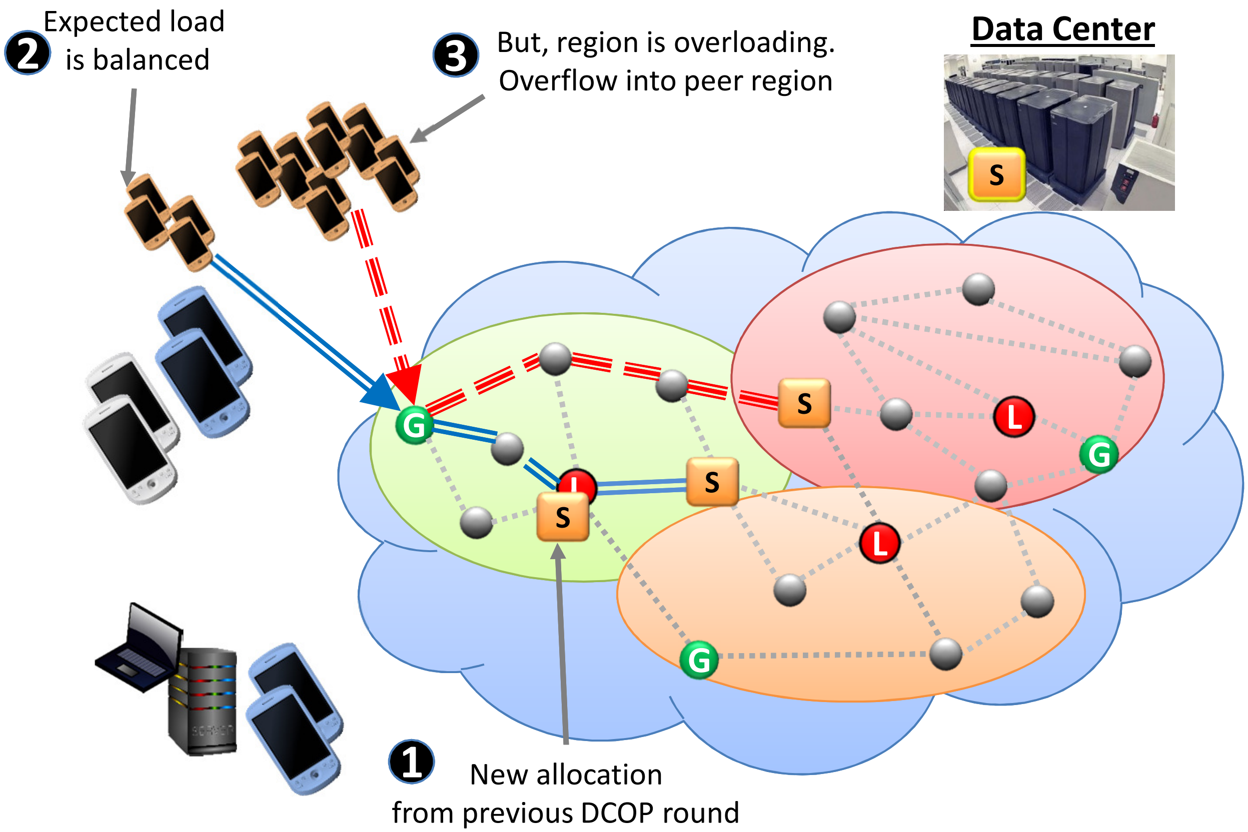
- DCOP considers emergent conditions in later rounds
- In response to load, new service allocations are determined

3a. – Periodically calculate task/data placement

- Monitor application usage and in-network resources
- Periodically determine solution for task and data placement

Concept of Operation (4/4)

Raytheon
BBN Technologies



1. Regional leaders execute new allocation plan
2. New load on system is load-balanced into the network
3. Overload or failure may be directed to ***overflow regions***

3b. – Quickly execute intra-regional algorithms

- Quickly provision global solution in region
- Tactically recover from failures and respond to increased load

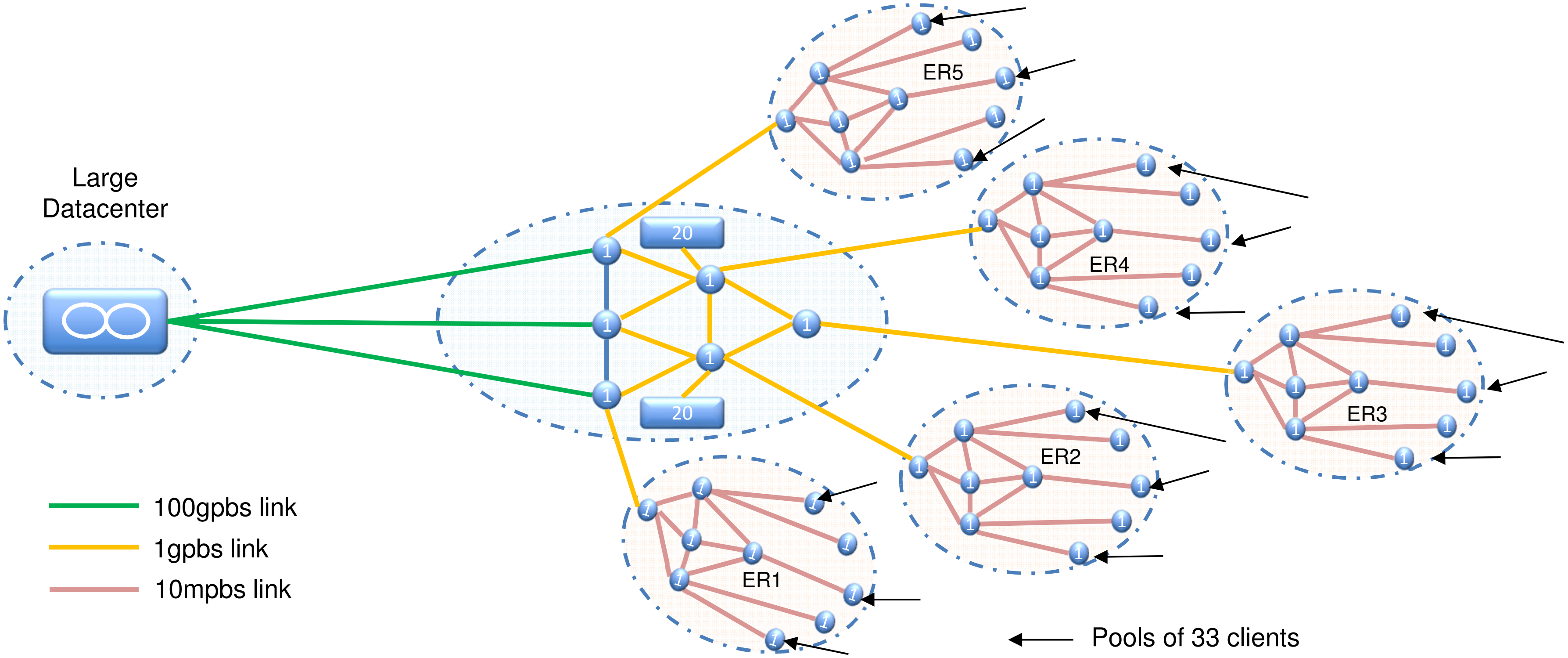
Outline

Raytheon
BBN Technologies

- Introductions
- Research Thrusts and CONOPS
- **Initial Proof of Concept Results**
- Takeaways

Lo-Fi Simulation Testing: 100 NCP and 500 Clients

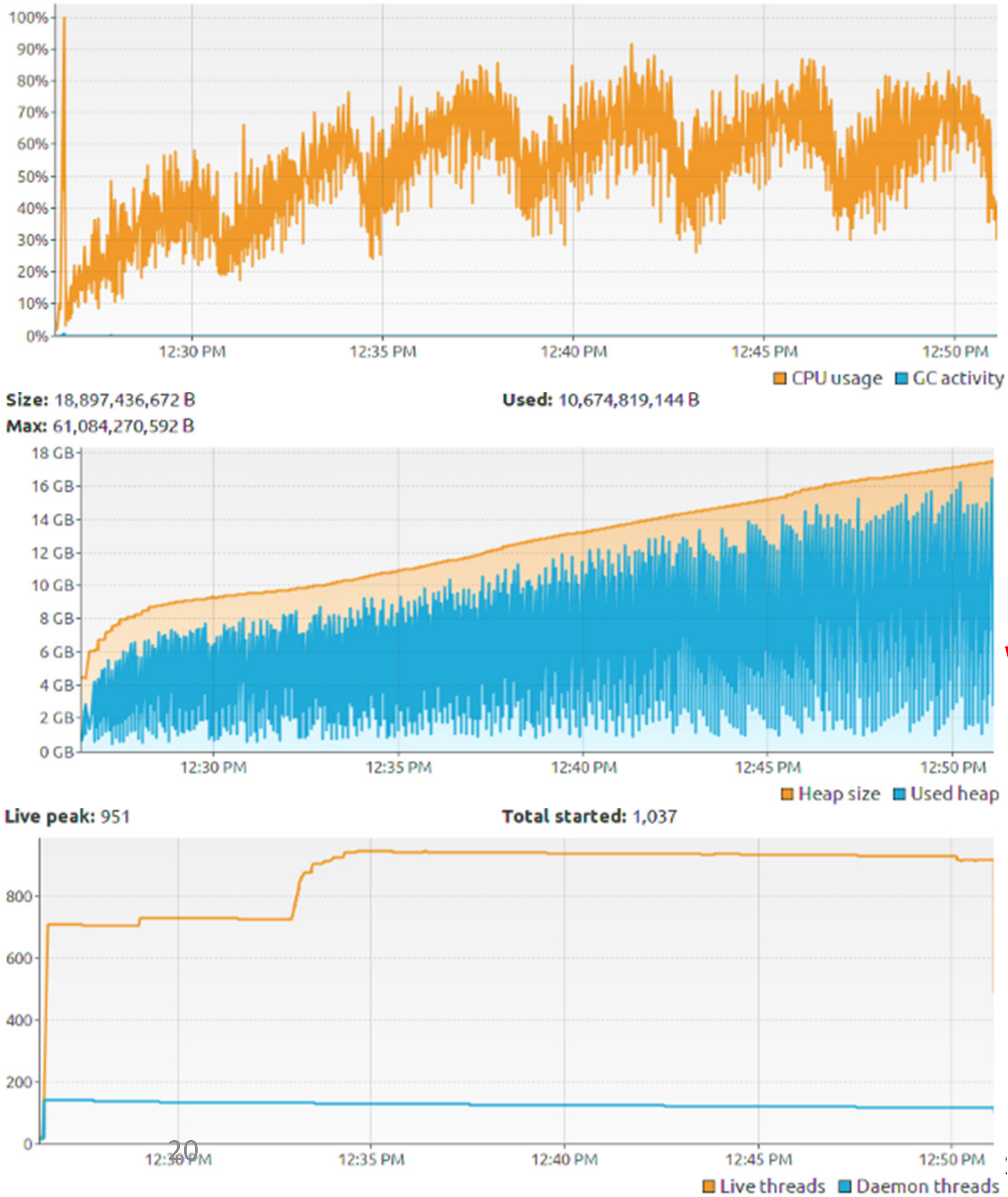
Raytheon
BBN Technologies



- Backhaul edge clients back to large data center
- MAP Agents on 99 NCPs and 1 in datacenter
- 3 pools of ~33 clients attached to the 5 edge regions
  - 24 minute demand for three simulated applications (50kbps each)

Lo-Fi Simulation Testing: Architectural Scaling (100NCPs,500clients)

Raytheon BBN Technologies



HW/SW Configuration

- c5.18xlarge
  - 72 vCPU, 144GiB
- JVM heap 64MB

Observations

- Simulation and Agent are stable at test scale ✓
- Memory growth is linear here, but well below heap bound
- Can fully simulate Program scale w/ lo-fi testbed ✓

Lo-Fi Simulation Testing: Estimating Backhaul Reduction (100NCP,500clients)

Raytheon BBN Technologies

	Baseline (Exp 1)	w/ MAP (Exp 2)			Estimating Backhaul Reduction				
Client Pool	Datacenter Requests	Regional Requests	Datacenter Requests	Ratio	Backhaul Path (shortest hop)	Region Path (ave hop)	Weighted Baseline	Weighted w/ MAP	Estimated Reduction
B0	1326	1155	166	0.87	6	3.8	15912	10770	0.32
B2	1298	1155	165	0.88	7	4.1	18172	11781	0.35
B4	1319	1155	146	0.89	6	3.7	15828	10299	0.35
C0	1301	1155	175	0.87	6	3.8	15612	10878	0.30
C2	1320	1155	173	0.87	7	4.1	18480	11893	0.36
C4	1320	1155	137	0.89	6	3.7	15840	10191	0.36
D0	1299	1122	185	0.86	7	3.8	18186	11117.2	0.39
D2	1314	1155	162	0.88	8	4.1	21024	12063	0.43
D4	1317	1155	228	0.84	7	3.7	18438	11739	0.36
E0	1345	1155	144	0.89	7	3.8	18830	10794	0.43
E2	1314	1155	153	0.88	8	4.1	21024	11919	0.43
E4	1288	1155	207	0.85	7	3.7	18032	11445	0.37
F0	1311	1155	183	0.86	7	3.8	18354	11340	0.38
F2	1349	1155	167	0.87	8	4.1	21584	12143	0.44
F4	1286	1155	135	0.90	7	3.7	18004	10437	0.42
								Ave Reduction	.38

- w/ MAP observed *87% reduction in data center requests*
  - Service migration and full DNS delegation completed ~5 minutes into test
- *Estimate 38% hop count reduction against test topology*
  - Note that backhaul reduction is a function of topology

Outline

- Introductions
- Research Thrusts and CONOPS
- Initial Proof of Concept Results
- **Takeaways**

Summary and Next Steps

Raytheon BBN Technologies

Investigating a *middleware* solution for dispersing computation and data into in-network NCPs

- ❑ Relocates data and tasks closer to use
  - Based on multi-layered distributed decision making
  - Has its algorithmic basis in DCOP, consistent hashing, control theory

Near term goal is release an open architecture for experimentation

- ❑ Interested in community involvement and new research thrusts

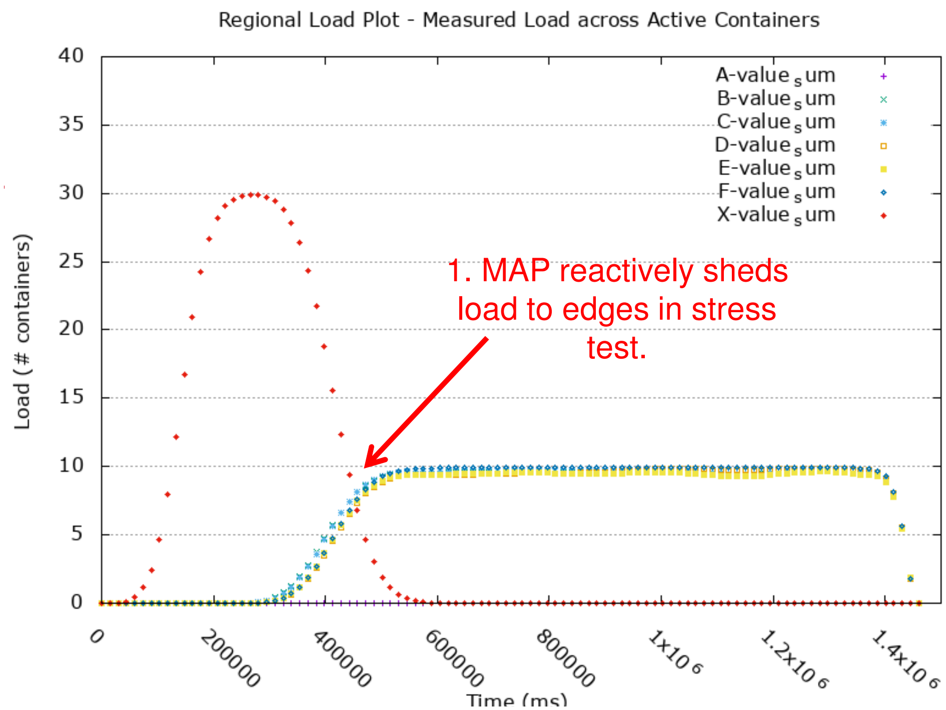
Discussion: Application to Applications with Edge Components

Raytheon BBN Technologies

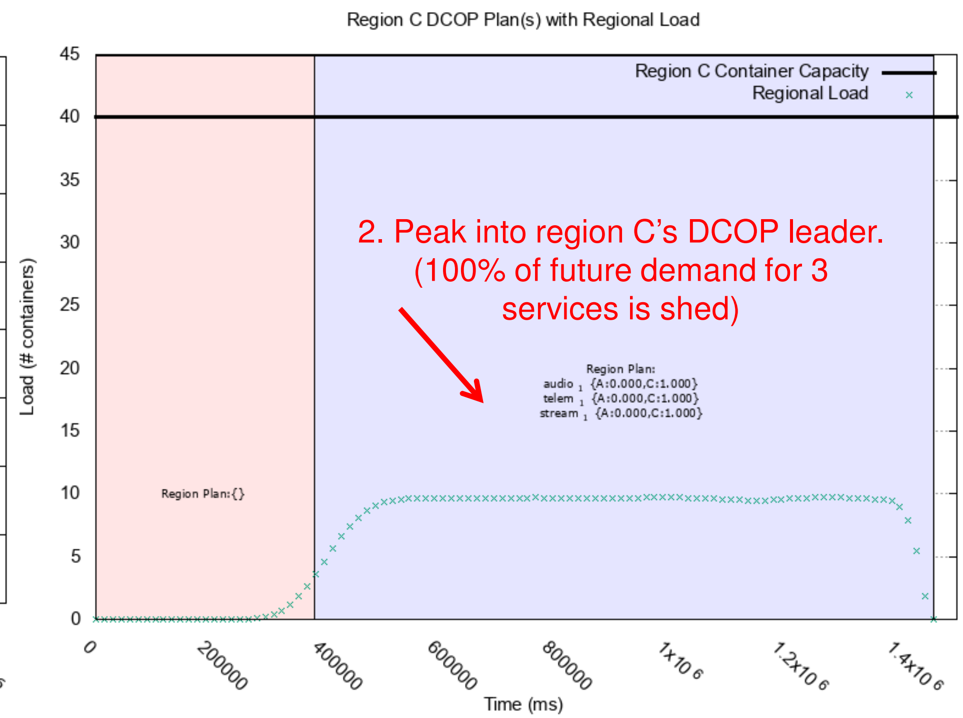
- New compute paradigm has the potential disrupt dependency on backhauled data-centers
  - Many potential benefits in QoS and resiliency for infrastructure and applications
- Motivates new thought in many areas
  - Algorithms for applications, e.g., partitioning tasks and data
  - Methods for managing applications, e.g., role of stakeholders
  - Security and trustworthiness, e.g., multi-tenancy

Presentation Q&A

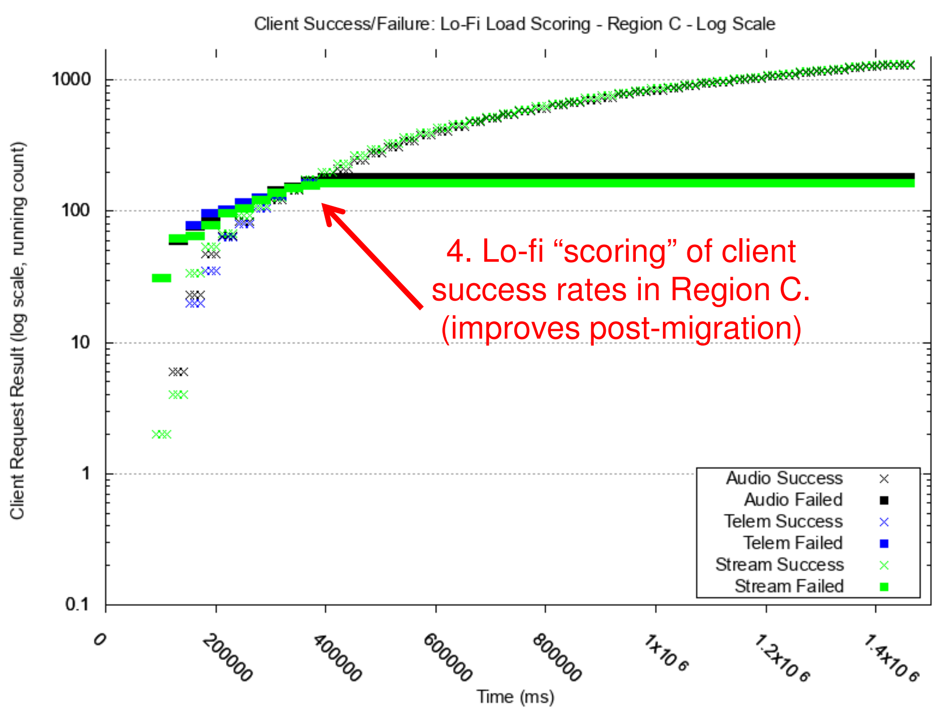
Distribution Statement “A” (Approved for Public Release, Distribution Unlimited)



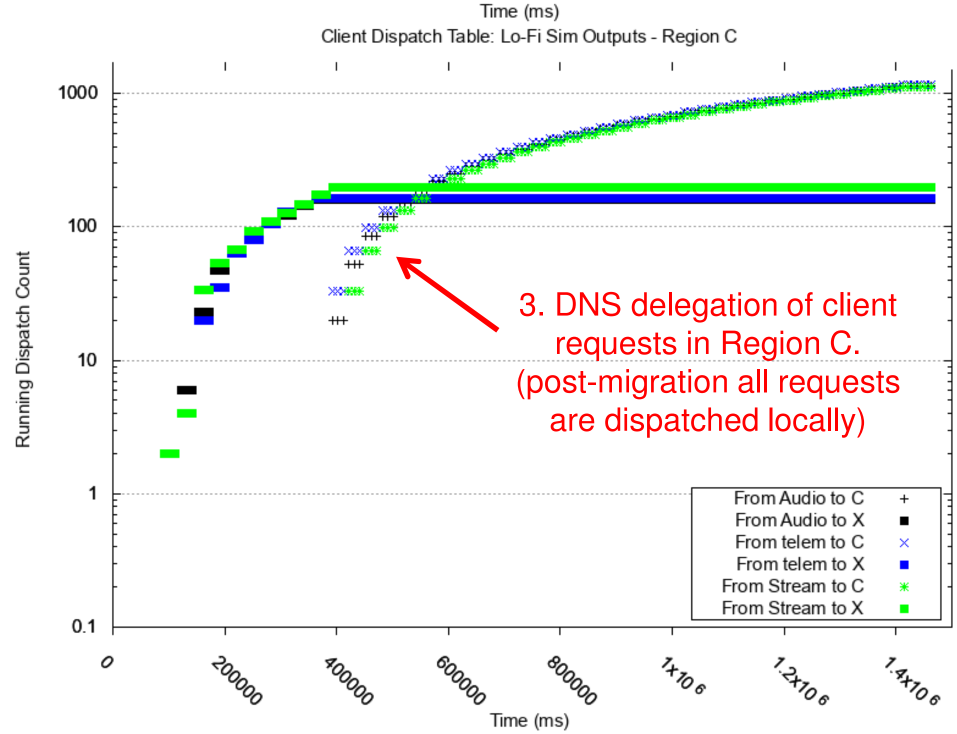
Regional Load Plot - Measured Load across Active Containers
Load (# containers)
Time (ms)
A-value_sum
B-value_sum
C-value_sum
D-value_sum
E-value_sum
F-value_sum
X-value_sum
1. MAP reactively sheds load to edges in stress test.



Region C DCOP Plan(s) with Regional Load
Load (# containers)
Time (ms)
Region C Container Capacity
Regional Load
2. Peak into region C's DCOP leader. (100% of future demand for 3 services is shed)
Region Plan:
audio 1 {A:0.000,C:1.000}
telem 1 {A:0.000,C:1.000}
stream 1 {A:0.000,C:1.000}
Region Plan:{}



Client Success/Failure: Lo-Fi Load Scoring - Region C - Log Scale
Client Request Result (log scale, running count)
Time (ms)
4. Lo-fi "scoring" of client success rates in Region C. (improves post-migration)
Audio Success
Audio Failed
Telem Success
Telem Failed
Stream Success
Stream Failed



Client Dispatch Table: Lo-Fi Sim Outputs - Region C
Running Dispatch Count
Time (ms)
3. DNS delegation of client requests in Region C. (post-migration all requests are dispatched locally)
From Audio to C
From Audio to X
From telem to C
From telem to X
From Stream to C
From Stream to X