

Scalability, Accountability and Instant Information Access for Network Centric Warfare

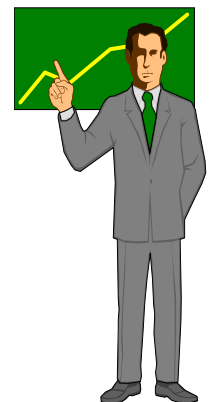
Yair Amir, Claudiu Danilov, Jon Kirsch, John Lane, Jonathan Shapiro

**Department of Computer Science
Johns Hopkins University**

*Chi-Bun Chan, Cristina Nita-Rotaru, Josh Olsen
David Zage*

**Department of Computer Science
Purdue University**

<http://www.cnds.jhu.edu>





Dealing with Insider Threats

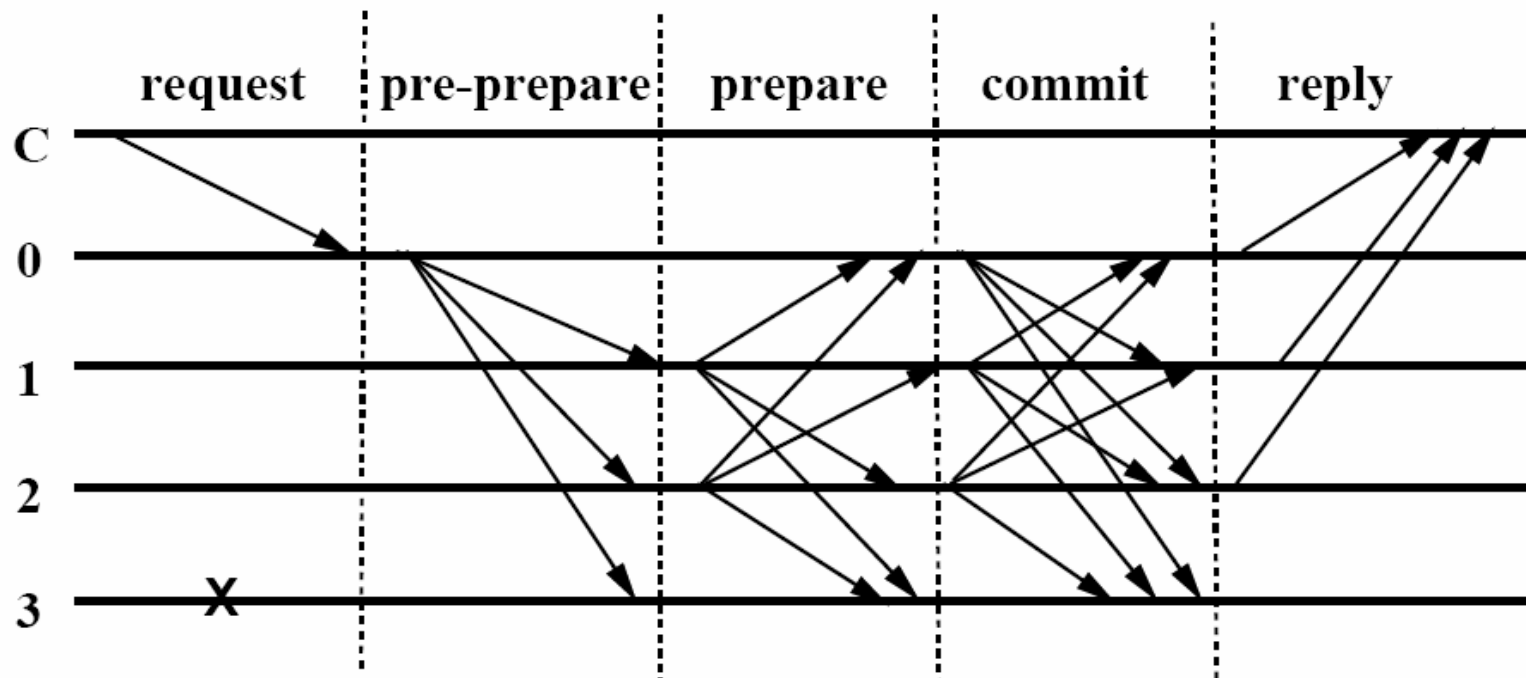
Project goals:

- Scaling survivable replication to wide area networks.
 - **Overcome 5 malicious replicas.**
 - **SRS goal: Improve latency by a factor of 3.**
 - **Self imposed goal: Improve throughput by a factor of 3.**
- Dealing with malicious clients.
 - **Compromised clients can inject authenticated but incorrect data - hard to detect on the fly.**
 - **Malicious or just an honest error?** Can be useful for both.
- Exploiting application update semantics for replication speedup in malicious environments.
 - **Weaker update semantics allows for immediate response.**

Here we focus on scaling survivable replication to wide area networks.
Introducing **Steward**: Survivable Technology for Wide Area Replication.

State of the Art in Byzantine Replication

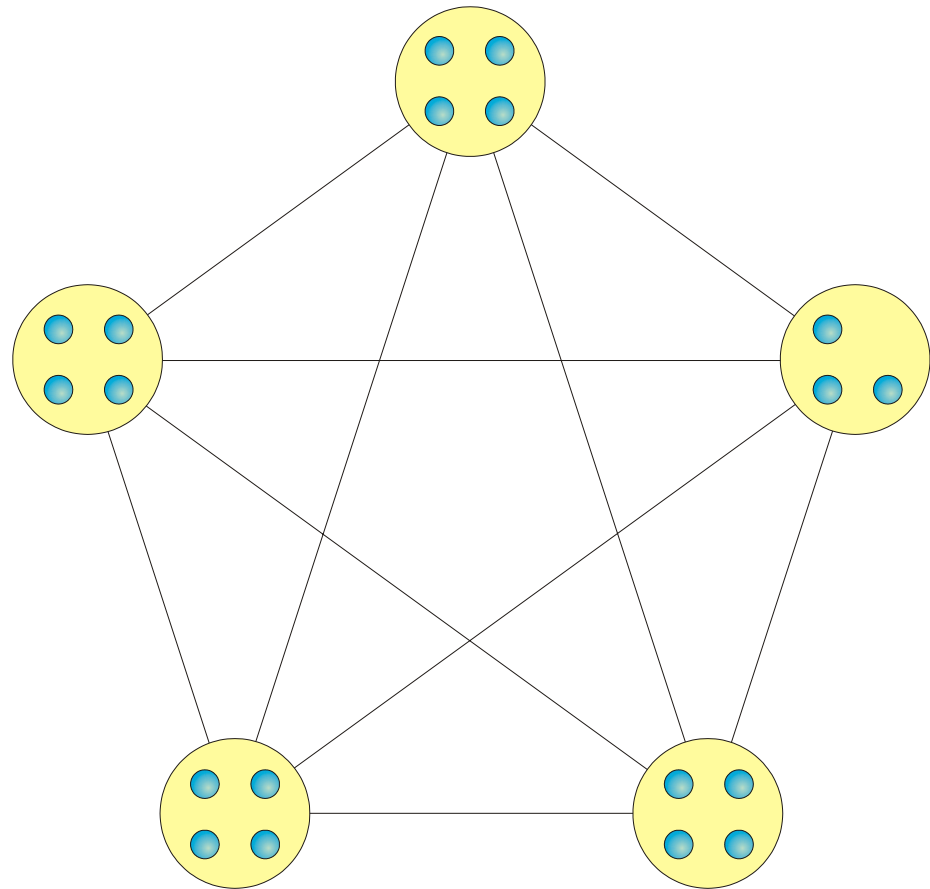
BFT [CL99]



Baseline technology

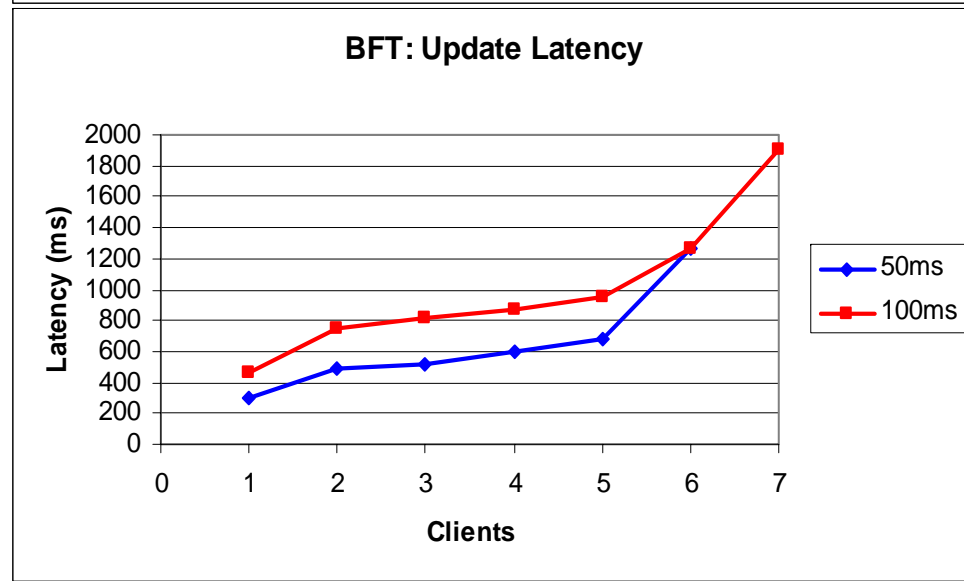
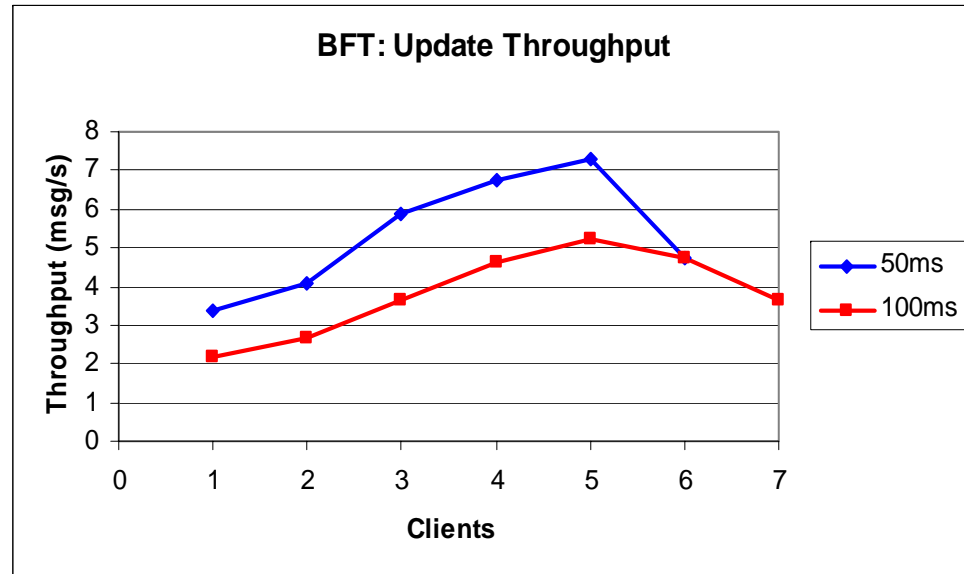
Evaluation Network 1: Symmetric Wide Area Network

- Synthetic network used for analysis and understanding.
- 5 sites, each of which connected to all other sites with equal latency links.
- Each site has 4 replicas (except one site with 3 replicas due to current BFT setup).
- Total – 19 replicas in the system.
- Each wide area link has a 10Mbits/sec capacity.
- Varied wide area latencies between 10ms - 400ms.

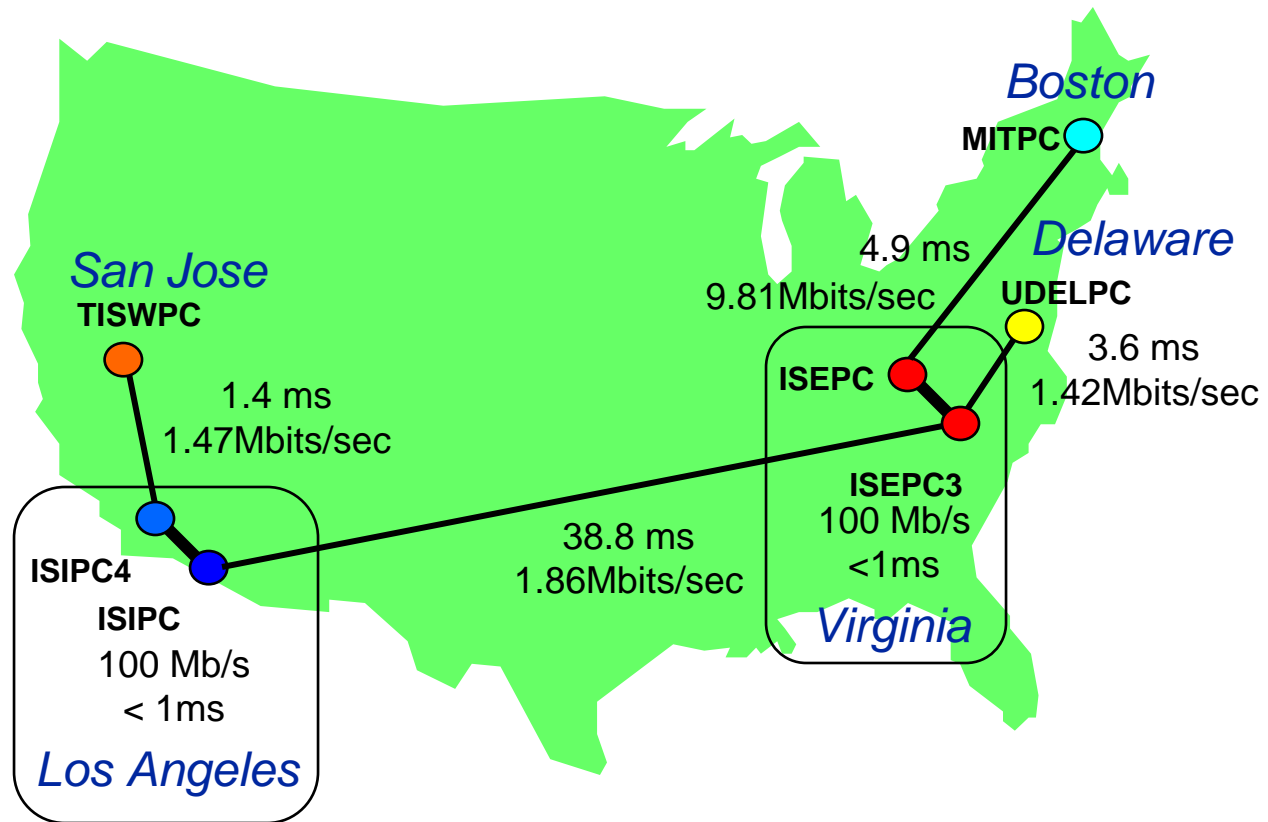


BFT Wide Area Performance

- Symmetric network.
- 5 sites.
- Total of 19 Replicas.
- Almost out of the box BFT, which is a **very good prototype software**.
- Update only performance (no disk writes).



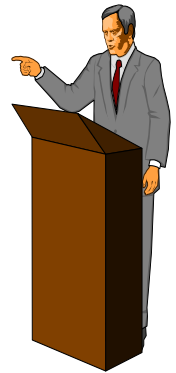
Evaluation Network 2: Practical Wide-Area Network



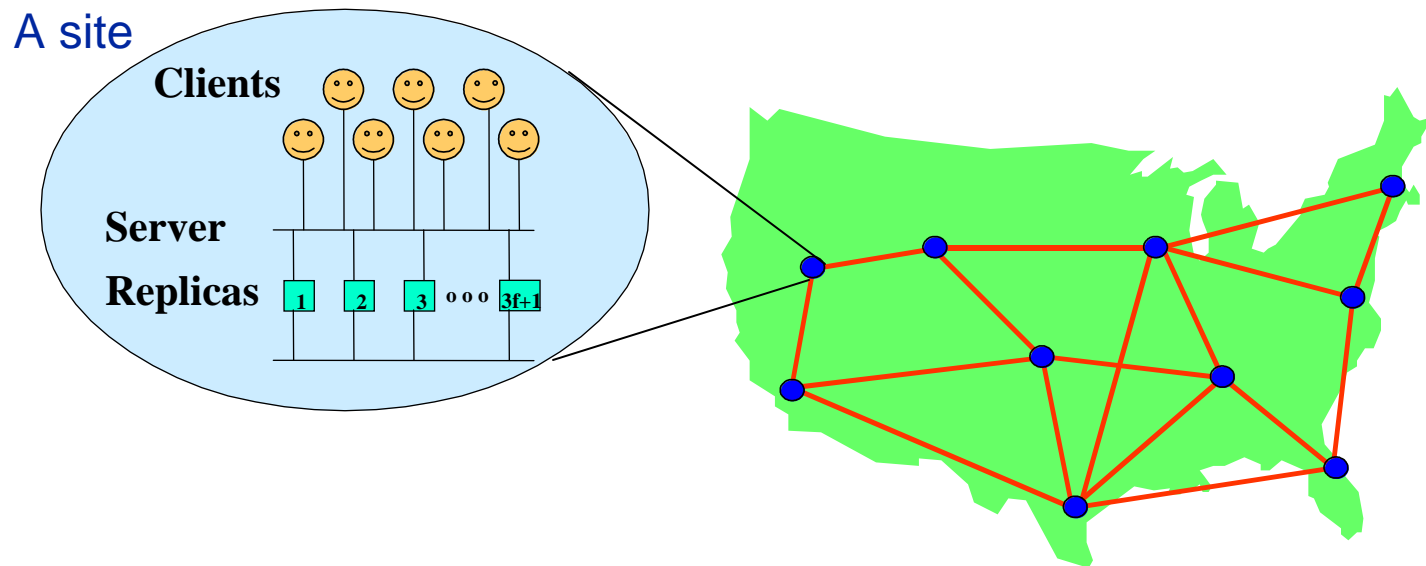
- Based on a real experimental network (CAIRN).
- Modeled in the Emulab facility.
- Capacity of wide area links was modified to be 10Mbits/sec to better reflect current realities.
- **Results are not shown here.**

Outline

- Project goals.
- Byzantine replication – current state of the art.
- **Steward** – a new hierarchical approach. ←
- Confining the malicious attack effects to the local site.
 - **BFT-inspired protocol for the local area site.**
 - **Threshold Cryptography for trusted sites.**
- Fault tolerant replication for the wide area.
 - **Initial thinking and snags.**
 - **A Paxos-based approach.**
- Putting it all together.
- Evaluation.
- Summary.

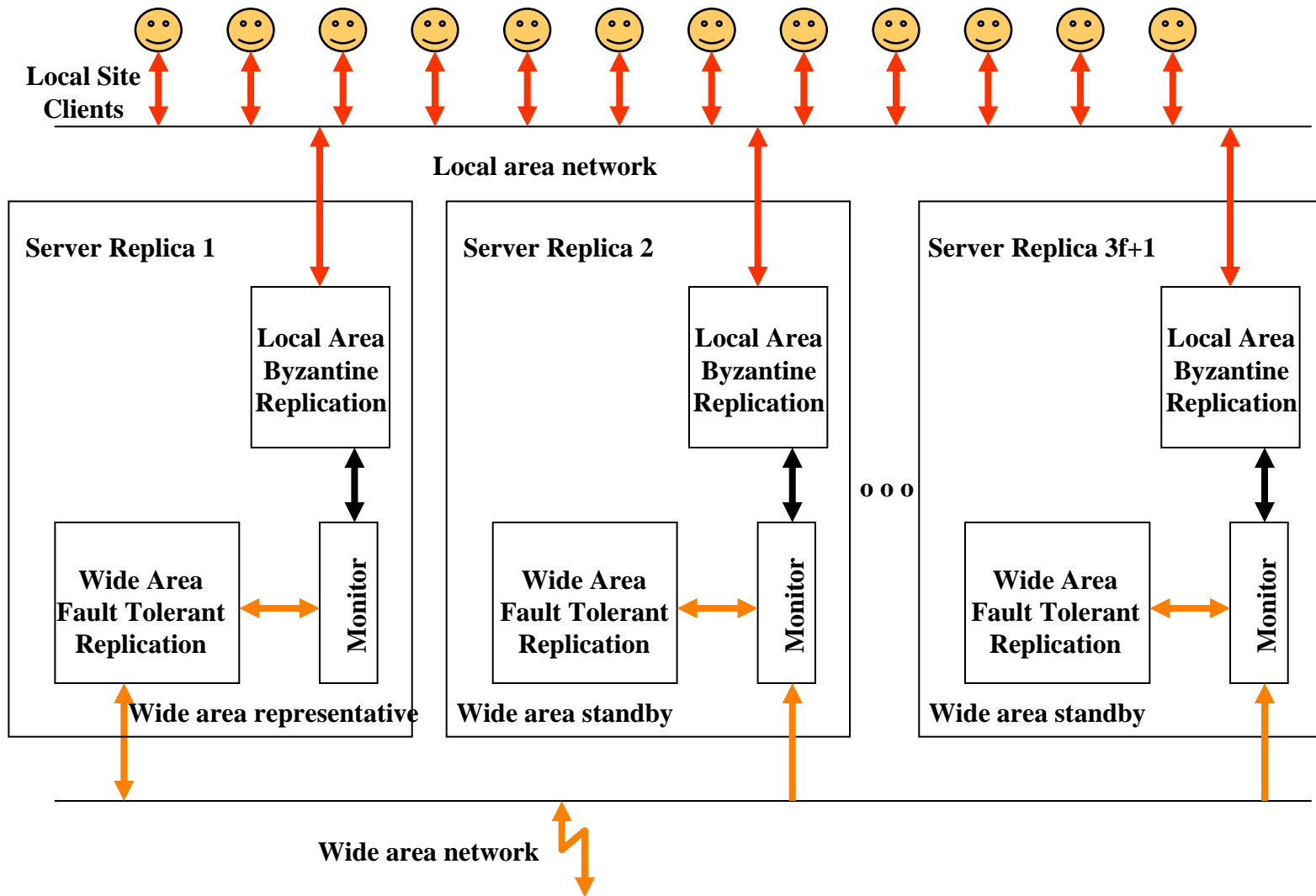


Steward: Survivable Technology for Wide Area Replication



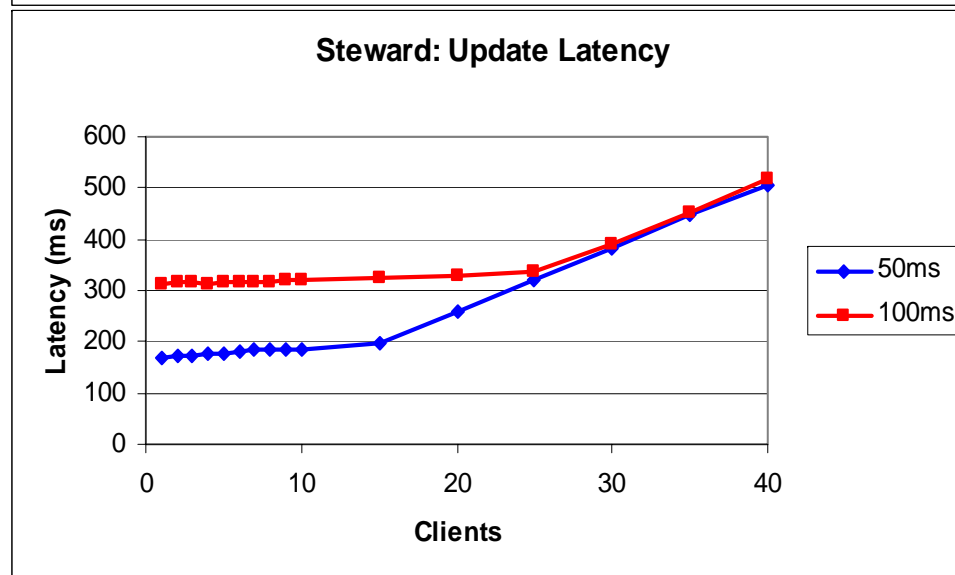
- Each site acts as a trusted logical unit that can crash or partition.
- Effects of malicious faults are confined to the local site.
- Between sites:
 - Fault-tolerant protocol between sites.
 - Alternatively – Byzantine protocols also between sites.
- There is no free lunch – we pay with more hardware...

Steward Architecture



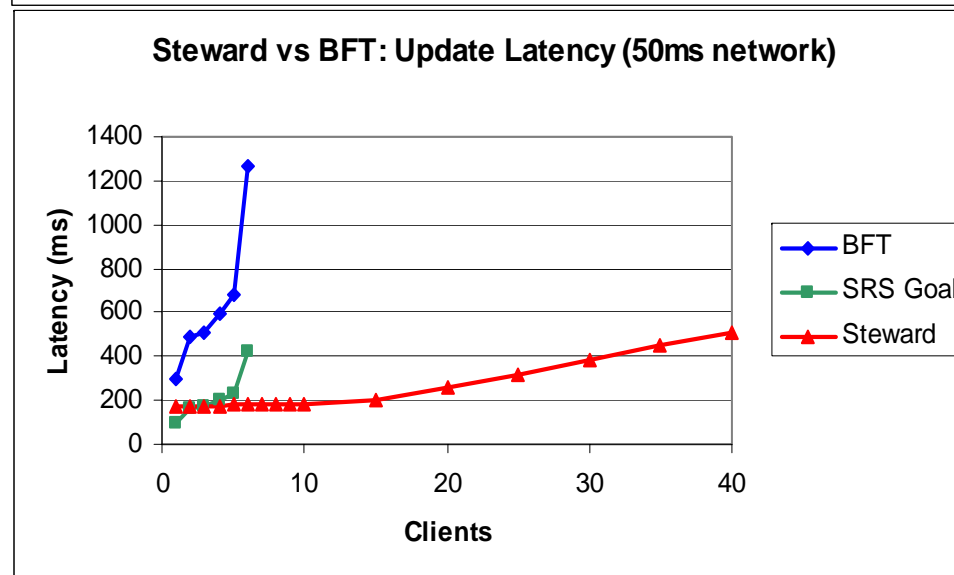
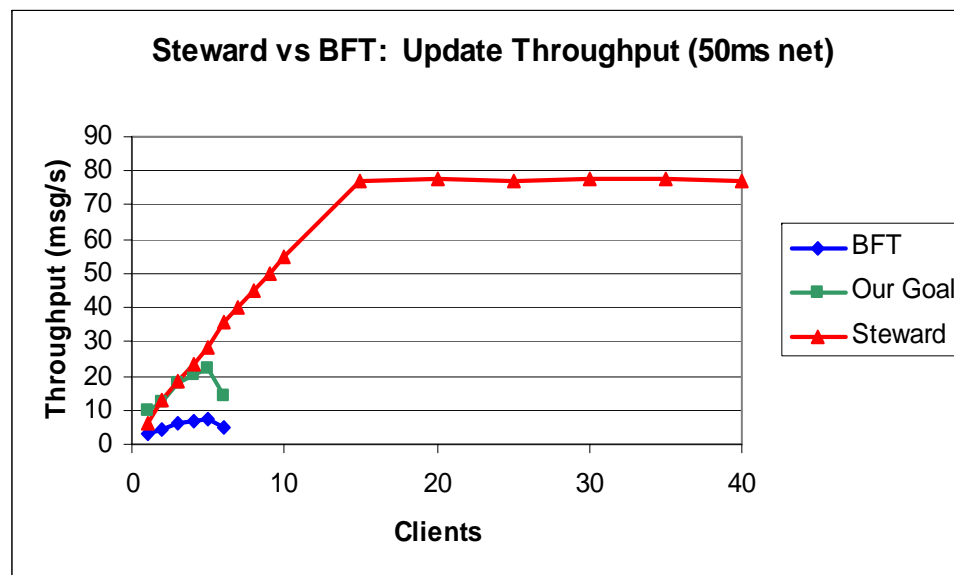
Steward Measured Performance

- Symmetric network.
- 5 sites.
- 16 replicas per site.
- Total of 80 replicas.
- **Methodology:**
Leader site has 16 replicas. Each other site has 1 entity that performs busy-wait equal (**conservatively**) to the cost of a receiver site reply, including threshold cryptography.
- Actual computers: 20.
- Update only performance (no disk writes).



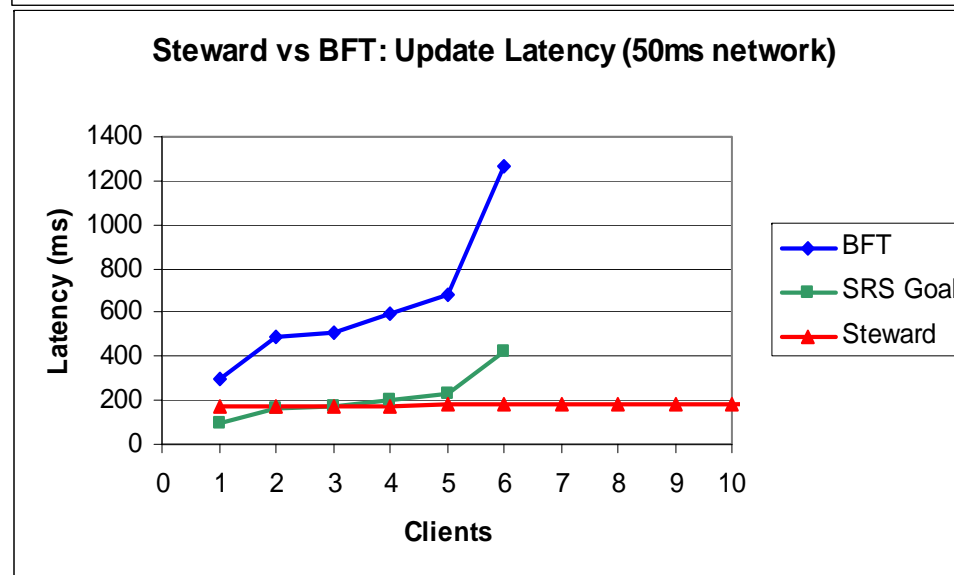
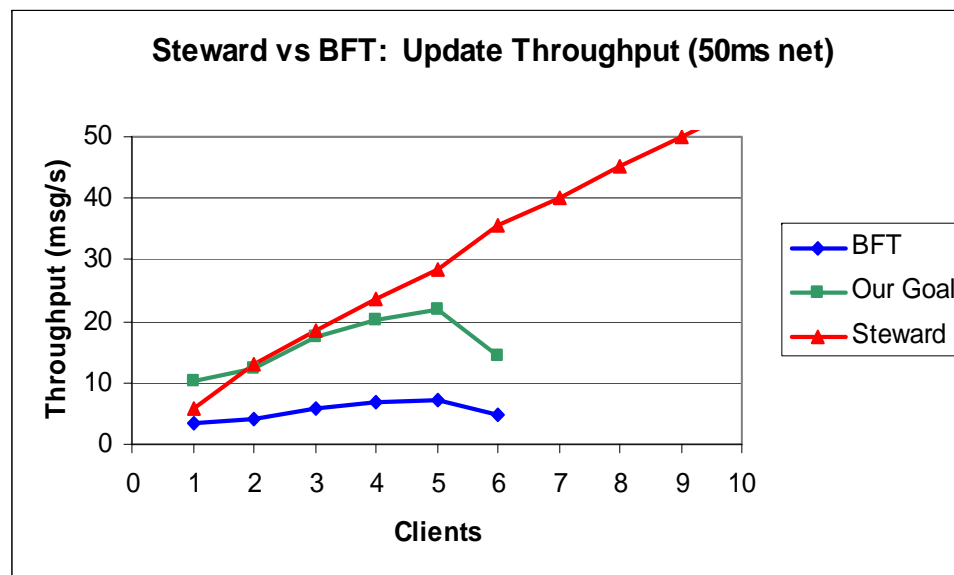
Head to Head Comparison (1)

- Symmetric network.
- 5 sites.
- 50ms distance between each site.
- 16 replicas per site.
- Total of 80 replicas.
- BFT broke after 6 clients.
- SRS goal: Factor of 3 improvement in latency.
- Self imposed goal: Factor of 3 improvement in throughput.
- **Bottom line:** Both goals are met once system has **more than one client**, and considerably exceeded thereafter.



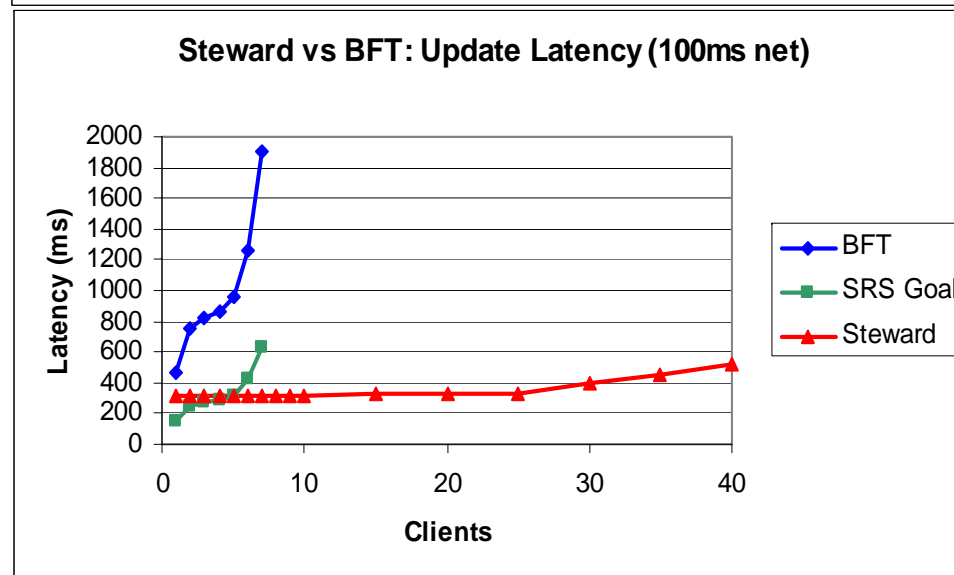
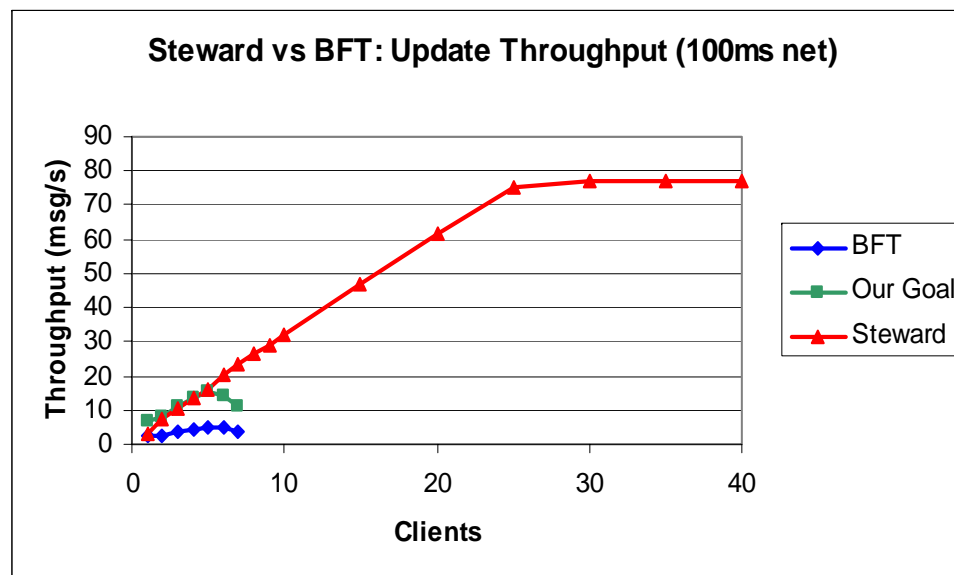
Head to Head Zoom (1)

- Symmetric network.
- 5 sites.
- 50ms distance between each site.
- 16 replicas per site.
- Total of 80 replicas.
- BFT broke after 6 clients.
- SRS goal: Factor of 3 improvement in latency.
- Self imposed goal: Factor of 3 improvement in throughput.
- **Bottom line:** Both goals are met once system has **more than one client**, and considerably exceeded thereafter.



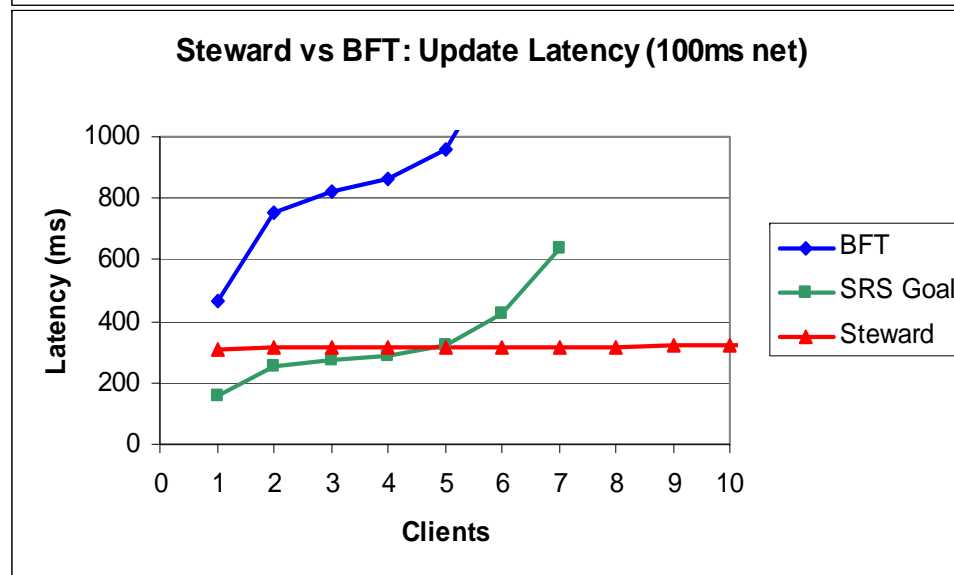
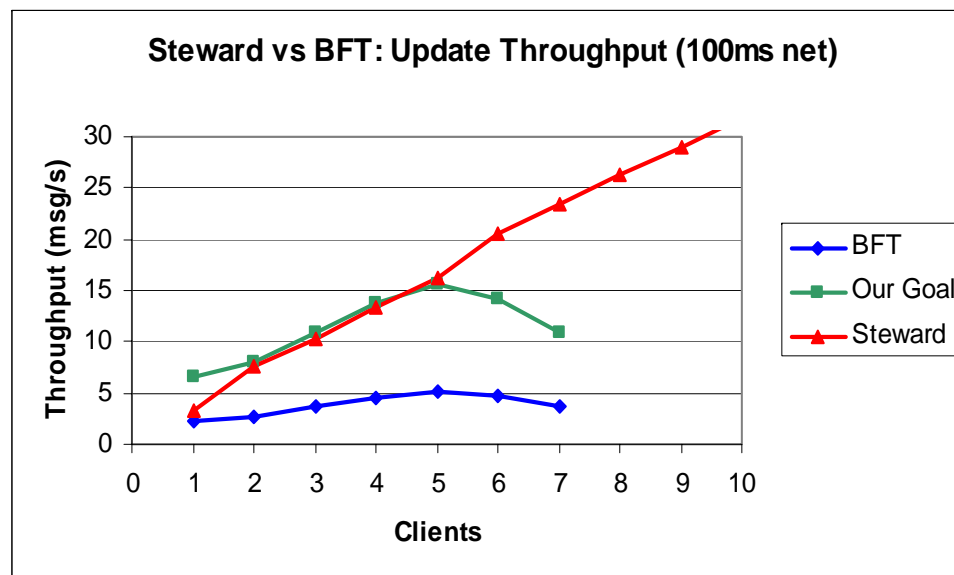
Head to Head Comparison (2)

- Symmetric network.
- 5 sites.
- 100ms distance between each site.
- 16 replicas per site.
- Total of 80 replicas.
- BFT broke after 7 clients.
- SRS goal: Factor of 3 improvement in latency.
- Self imposed goal: Factor of 3 improvement in throughput.
- **Bottom line:** Both goals are met once system has **one client per site**, and considerably exceeded thereafter.



Head to Head Zoom (2)

- Symmetric network.
- 5 sites.
- 100ms distance between each site.
- 16 replicas per site.
- Total of 80 replicas.
- BFT broke after 7 clients.
- SRS goal: Factor of 3 improvement in latency.
- Self imposed goal: Factor of 3 improvement in throughput.
- **Bottom line:** Both goals are met once system has **one client per site**, and considerably exceeded thereafter.



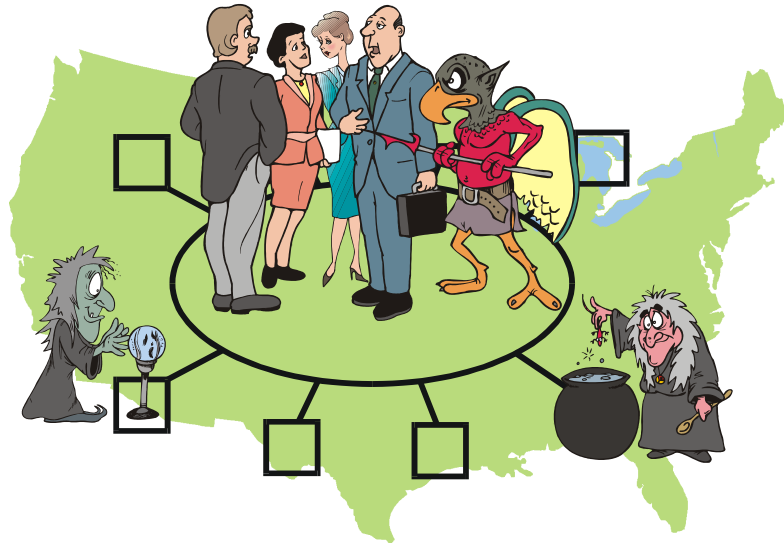


Factoring Queries In

- Steward:
 - **A query is answered locally after $f+1$ Threshold Cryptography operation.** Cost: 11ms.
- BFT:
 - **A query requires at least some remote answers in this setup.** Cost: at least 100ms (for 50ms network), 200ms (for 100ms network).
 - **We could change the setup to include 6 local members in each site (for a total of 30 replicas).** That will allow a local answer in BFT with a query cost similar to Steward, but then BFT performance will basically collapse on the updates.
- Bottom line prediction:
 - Both goals will be met once the system has **more than one client**, and will be considerably exceeded thereafter.

Scalability, Accountability and Instant Information Access for Network-Centric Warfare

New ideas



First scalable wide-area intrusion-tolerant replication architecture.

Providing accountability for authorized but malicious client updates.

Exploiting update semantics to provide instant and consistent information access.

Impact

Resulting systems with at least 3 times higher throughput, lower latency and high availability for updates over wide area networks.

Clear path for technology transitions into Military C3I systems such as the Army Future Combat System.

Schedule

